Equilibrium properties of DNA and other semiflexible polymers confined in nanochannels

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Abhiram Muralidhar

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: **Kevin D. Dorfman**

January, 2016

Acknowledgements

I'm grateful to my advisor, Kevin Dorfman, for providing me a stimulating environment to carry out scientific research and the freedom to express myself. I have learned a lot from him through some of the most formative years of my life, both from direct interactions with him, and from watching him work from a distance. He has also given me a number of opportunities for collaboration with bright researchers in various parts of the world, which have been productive and enjoyable. Kevin, thank you.

I would like to thank my other committee members Prof. Dave Morse, Prof. Xiang Cheng and Prof. Ilja Siepmann for reading and commenting on my work. I acknowledge help from Pranav Agrawal, Akash Arora, Damini Gupta and Xiaolan Li for proofreading parts of this thesis.

Various members of the Dorfman group over the years have made my tenure at Minnesota particularly memorable. Maggie Linak helped me get started with research in the first couple of years of my PhD. I'm deeply indebted to Doug Tree, who among other things wrote the original PERM and CFD code that I have used extensively in my research. I have also had many stimulating discussions with him about PERM, polymer physics and career, which were all quite enjoyable. Sarit Dutta has been a great friend and "philosopher". Our conversations over Starbucks coffee and otherwise kept me going through tough and happy times. I have enjoyed working with Yanwei Wang, Damini Gupta, Julian Sheats, Akash Arora, Aashish Jain, Xiaolan Li, Michael Quevillon, Pranav Agrawal and Seunghwan Shin at various points during my PhD.

I would like to thank my family, especially my mum and dad, for all their support from the other end of the world. Aside from the alphabet, my mum has taught me to always challenge myself and never give up, and these lessons have been extremely valuable to me. Special thanks to my aunt, who never made me feel I was far away from home here in Minnesota. Sonakka, thank you. I'm grateful to my friends and family, some of whom are here in the US, some in Europe and most of whom are back home

in India for their warmth and affection, which has had my stay here in America a lot easier.

Last but not least, I'd again like to thank Kevin Dorfman and the department of Chemical Engineering & Materials Science for the generous finanical support throughout my PhD. The work presented here was supported by grants from the National Institutes of Health (NIH) (R01-HG006851), the National Science Foundation (NSF) (CBET-1262286), and a David and Lucile Packard Foundation fellowship to KDD. A major portion of the computational work was done using resources provided by the University of Minnesota Supercomputing Institute.

In memory of D. R. Gopala Rao (*Thatha*) ...

Abstract

Recent developments in next-generation sequencing (NGS) techniques have opened the door for low-cost, high-throughput sequencing of genomes. However, these developments have also exposed the inability of NGS to track large scale genomic information, which are extremely important to understand the relationship between genotype and phenotype. Genome mapping offers a reliable way to obtain information about large-scale structural variations in a given genome. A promising variant of genome mapping involves confining single DNA molecules in nanochannels whose cross-sectional dimensions are approximately 50 nm. Despite the development and commercialization of nanochannel-based genome mapping technology, the polymer physics of DNA in confinement is only beginning to be understood.

Apart from its biological relevance, DNA is also used as a model polymer in experiments by polymer physicists. Indeed, the seminal experiments by Reisner et al. (2005) of DNA confined in nanochannels of different widths revealed discrepancies with the classical theories of Odijk and de Gennes for polymer confinement.

Picking up from the conclusions of the dissertation of Tree (2014), this dissertation addresses a number of key outstanding problems in the area of nanoconfined DNA. Adopting a Monte Carlo chain growth technique known as the pruned-enriched Rosenbluth method, we examine the equilibrium and near-equilibrium properties of DNA and other semiflexible polymers in nanochannel confinement. We begin by analyzing the dependence of molecular weight on various thermodynamic properties of confined semiflexible polymers. This allows us to point out the finite size effects that can occur when using low molecular weight DNA in experiments. We then analyze the statistics of backfolding and hairpin formation in the context of existing theories and discuss how our results can be used to engineer better conditions for genome mapping. Finally, we elucidate the diffusion behavior of confined semiflexible polymers by comparing and contrasting our results for asymptotically long chains with other similar studies in the literature.

We expect our findings to be not only beneficial to the design of better genome mapping devices, but also to the fundamental understanding of semiflexible polymers in confinement.

Contents

List of Tables

List of Figures

Statement of the Author's Contributions

Portions of this dissertation were written as a result of collaborative efforts between members of the Dorfman group. Some chapters have also appeared (or will appear) as articles in various journals. Here, I would like to point out the contributions of the different people involved.

I wrote most of Chapters 1, 2 and 3. Some figures in Chapter 1, Chapter 2 and 3 are reproduced from K. D. Dorfman, D Gupta, A Jain, A Muralidhar, and D. R. Tree, "Hydrodynamics of DNA confined in nanoslits and nanochannels", Eur. Phys. J. Spec. Top. **223**, 3179 (2014) and D. R. Tree, A. Muralidhar, P. S. Doyle, and K. D. Dorfman, "Is DNA a Good Model Polymer?", Macromolecules **46**, 8369 (2013). I co-authored both these articles. Some text in Chapter 2 is based on the former mini-review article from parts that were written by KDD, DG and me.

Chapter 4 was published as an article: A. Muralidhar, D. R. Tree, Y. Wang, and K. D. Dorfman, "Interplay between chain stiffness and excluded volume of semiflexible polymers confined in nanochannels", J. Chem. Phys. **140**, 084905 (2014). In this paper, we used pruned-enriched Rosenbluth method (PERM) simulations to investigate the effect of molecular weight on ideal and real chains in confinement. The original PERM code was written by DRT. The simulation results in this chapter came out of an updated version of this code, which I developed by optimizing some of the functionality of the code. I performed simulations and analyzed the data. Interpretation of the data was shared by all the authors. KDD and I wrote most of the paper.

Chapter 5 was published as A. Muralidhar, D. R. Tree, and K. D. Dorfman, "Backfolding of Wormlike Chains Confined in Nanochannels", Macromolecules **47**, 8446 (2014). This paper contained results from the aforementioned PERM code applied to study the behavior of backfolding and hairpin formation in confined semiflexible polymers. The ideas in the paper were motivated by initial discussions among DRT, KDD and me. I added functions to calculate the global persistence length to the PERM code. Thereafter, I performed the simulations and analyzed the data. While the interpretation of the data was shared by all the authors, the paper was mostly written by KDD and me, with comments from DRT.

Chapter 6 is based on A. Muralidhar and K. D. Dorfman, "Backfolding of DNA Confined in Nanotubes: Flory Theory versus the Two-State Cooperativity Model", Macro-

molecules (in press) (2016) 10.1021/acs.macromol.5b02556). I performed the PERM simulations and analyzed the data. KDD and I interpreted the data and wrote the article.

Chapter 7 is based on a manuscript that is under currently under review for publication in the MDPI journal, Polymers. The manuscript has been co-authored by Abhiram Muralidhar, Michael J. Quevillon and Kevin D Dorfman. This chapter contains simulations results for the backfolded Odijk regime in rectangular channels. MJQ performed preliminary simulations and analyzed the data. Based on these simulations, I performed more simulations and re-analyzed all the data. KDD and I interpreted the data and wrote the manuscript.

Chapter 8 appeared as A. Muralidhar and K. D. Dorfman, "Kirkwood Diffusivity of Long Semiflexible Chains in Nanochannel Confinement", Macromolecules **48**, 2829 (2015). This paper used numerical results from a combination of the PERM and a computational fluid dynamics (CFD) simulations to study the Kirkwood diffusivity of confined polymers. DRT wrote the CFD code for a previous publication (Tree et al., 2012). I incorporated the Kirkwood diffusivity methodology into the PERM code by combining it with the CFD code, performed all simulations and analyzed the data. The interpretation of the results and writing the paper was shared by KDD and me.

Chapter 1

Introduction and Motivation

1.1 Genome mapping

Ever since the discovery of deoxyribonucleic acid (DNA) as the carrier of genetic information (Watson and Crick, 1953; Crick, 1970), there has been continued interest in sequencing DNA to decode genetic information (França et al., 2002). Recent developments in next generation sequencing (NGS) techniques have revolutionized the fields of genomics and molecular biology because of their ability to sequence DNA much more cheaply and quickly compared to the previously used Sanger sequencing (Metzker, 2010). In fact, NGS methods have enabled massive strides in whole-genome sequencing (Ng and Kirkness, 2010), exome sequencing (Bamshad et al., 2011) and targeted sequencing (Levin et al., 2009), in addition to transcriptomics (Wang et al., 2009) and epigenomics (Hirst and Marra, 2010).

Figure 1.1 shows the workflow of the most widely used NGS platform, the Illumina Genome Analyzer. Other NGS techniques include single-molecule real time sequencing (Pacific Biosciences), pyrosequencing (Roche/454), sequencing by ligation (Applied Biosystems SOLiD^{TM} Sequencer) and Ion Torrent sequencing (Quail et al., 2012; Liu et al., 2012). Regardless of the different flavors of NGS, they share common unit operations — almost all NGS techniques rely on fragmentation and amplification of DNA followed by massively parallelized steps of sequencing. Such a workflow implies that NGS sacrifices long-range connectivity in the genome leading researchers to overlook many structural rearrangements such as deletions, inversions and copy-number polymorphism as shown in Figure 1.2.

Furthermore, *de novo* genome sequencing is important for samples with large struc-

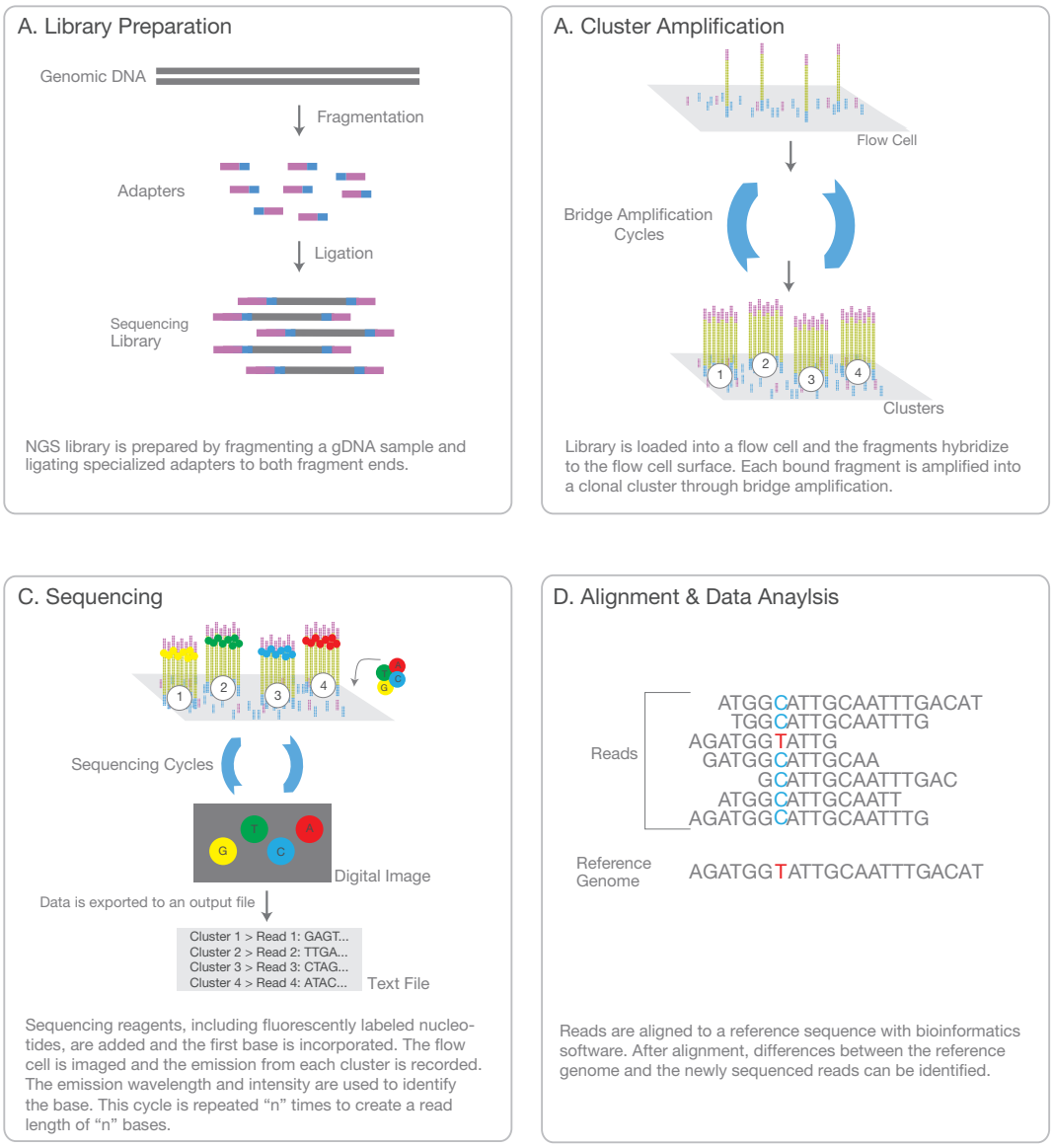


Figure 1.1. Schematic of the steps in Illumina sequencing, one of the widely used NGS techniques. Most other NGS methods share similar steps consisting of fragmentation, amplification, sequencing and sequence assembly. Reproduced from Illumina (2015).

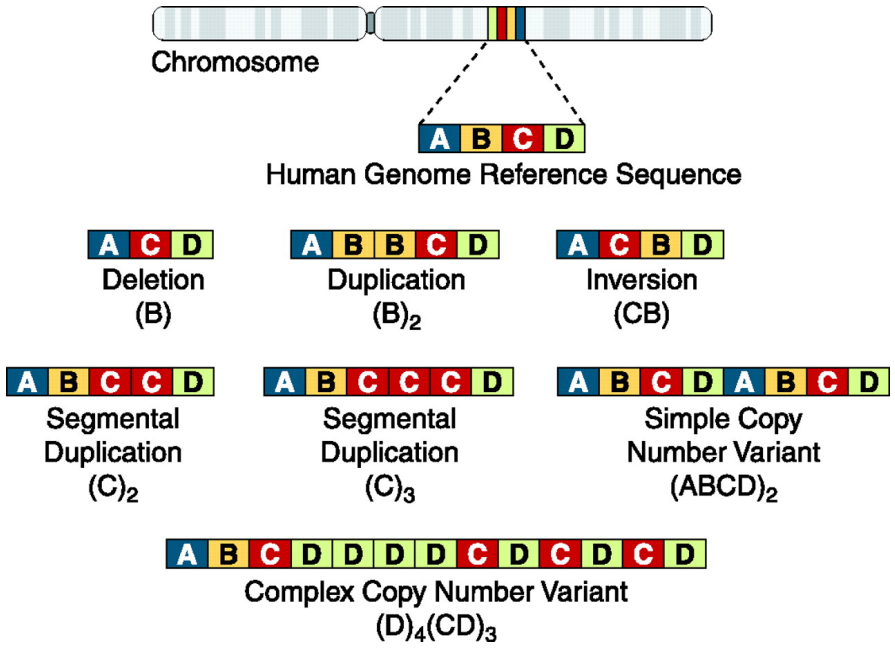


Figure 1.2. Various structural rearrangements that can occur with four segments (A, B, C and D) of a genome. NGS techniques can almost be blind to such rearrangements. Reproduced from Dierssen et al. (2009)

tural variation, e.g. cancer cells (Iafrate et al., 2004) and experimental strains containing genomes with high plasticity (Cazalet et al., 2004). In spite of the many indisputatble advantages that NGS possesses over conventional sequencing techniques, the short-length lengths and associated drawbacks such as inability to detect large inversions (Latreille et al., 2007) poses a challenge for obtaining *de novo* sequences with NGS. Indeed, the read-lengths achievable by the traditional Sanger sequencing ($\approx$ 1000 base pairs) (Karger and Guttman, 2009) are larger than the read-lengths characteristic of NGS ($\approx$ 500 base pairs) (Metzker, 2010; Niedringhaus et al., 2011). However, note that these read-lengths are orders of magnitude small compared to the size of a genome — the human genome, for instance, contains about 3 billion base pairs. Given that the information in N short reads does not equal the genomic information in a sequence N times as long, the shorter the read length, the more the loss of information. Amplification bias in some NGS methods further leads to loss of information. As a result, there are major challenges associated with using NGS techniques that have a resolution of the order of 1 kilo base pair (kbp) to assess genomic variations beyond the kbp range. Yet genomic rearrangments in the kilobase to megabase pair range are known to lead to important phenotypic traits and diseases (Feuk et al., 2006). Although recent single-molecule sequencing methods (Eid et al., 2009) attempt to address this problem by reaching read lengths in excess of 3 kbp (Koren et al., 2013), sequencing of large genomes suffers from the so-called "data deluge" problem (Dorfman et al., 2013) associated with dealing with large amounts of information from *de novo* sequences, which can make it hard to identify large-scale genomic arrangments.

A promising approach to address these drawbacks is to complement NGS with genome mapping techniques, where the basic idea is to determine the genomic distance between repeats of a given sequence as shown in Figure 1.3. Typically, this kind of mapping is first done a relatively large piece of DNA whose length ranges from several kilobase pairs up to a megabase pair. Once the mapping is done on a large number of such fragments of DNA, the information from the individual maps can be combined to come up with a genome map for the whole genome. The analogy used by Dorfman et al. (2013) (inspired by discusssion with Dr. Han Cao) perfectly sums up the relationship between genome mapping and NGS — "the connection between mapping and sequencing is analogous to exploring Google Maps with the zoom in/out functions." The short reads from NGS can be thought of as analogous to the localized 'street view' on Google maps, where one might still get lost if the houses look alike. However, the knowledge

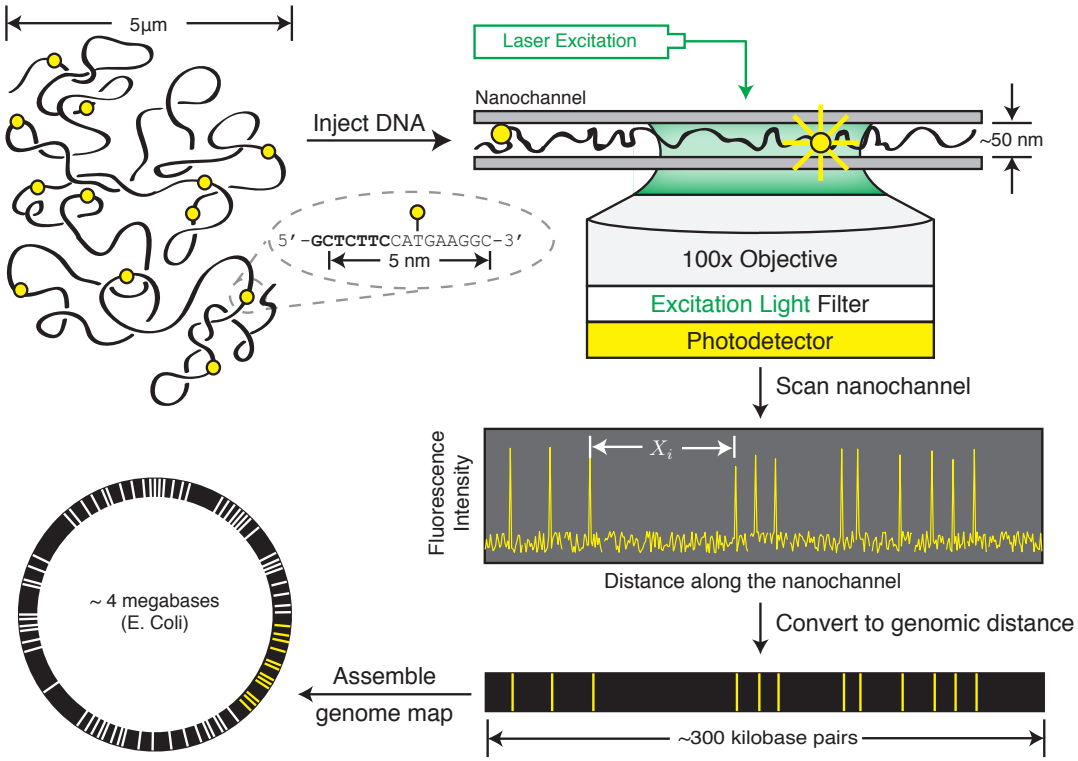


Figure 1.3. Schematic of genome mapping via DNA stretching in nanochannels. The example shown here corresponds to mapping of the *E. coli* genome. However genomes of many other organisms have been mapped (Dorfman et al., 2013). Reproduced from Dorfman et al. (2014).

gained from genome mapping, such as information about copy number amplifications and repeats, give us contextual information analogous to how the streets in a city are arranged.

Because the pattern obtained from some of the modern techniques of genome mapping resembles a barcode, as the output map shown in Figure 1.3, these methods are often referred to as "DNA barcoding" (Dorfman et al., 2013). DNA barcoding techniques essentially involve "linearizing" fragments of DNA, which are otherwise coiled up in free solution, by stretching them using an external stimulus. Once the fragments of DNA are linearized, the physical distance between fluorescent sequence-specific probes (repeats) can be measured using microscopy and can then be mapped to the genomic distance along the genome. DNA barcoding techniques include (i) molecular combing, where DNA is stretched by a receding contact line on a surface that preferentially absorbs DNA at one of its ends (Bensimon et al., 1994; Jing et al., 1998); (ii) extensional flow, wherein the DNA molecule is stretched by a velocity gradient in a hyperbolic channel (Chan et al., 2004) and (iii) nanochannel confinement, where the equilibrium conformation of the DNA molecule is shifted from a coiled-up state to a stretched conformation by confining the DNA molecule in a nanochannel whose cross-sectional size is approximately 50 nm (Tegenfeldt et al., 2004; Jo et al., 2007; Lam et al., 2012), as shown in Figure 1.3. Indeed the tremendous potential demonstrated by these techniques has led to commercialization of these technologies by OpGen and BioNano Genomics.

In nanochannel-based genome mapping, long fragments of DNA that are $\mathcal{O}(100$ kbp) long are labeled with sequence-specific fluorescent probes. As shown in Figure 1.4, the DNA molecule is then stretched and linearized by injecting it into a nanochannel. Injecting DNA into a nanochannel is a considerable challenge (radius of gyration of 50 kbp DNA is around 700nm) and it is achieved by electrophoretically passing the molecules through an entropic gradient as in Figure 1.4(a) and (b). Once the DNA molecules are confined in the channels, the fluorophores on the labeled nucleotides can be excited and observed with a laser of an appropriate wavelength. Typically, the backbone of the DNA molecule is labeled with an intercalating dye that fluoresces under light of a different color. As shown in Figure 1.4(c) and (d), by sequentially illuminating the backbone and fluorophores, the physical distance between the probes along the nanochannel can be measured, which can in turn be used to estimate the genomic distance between the probes along the molecule. By obtaining many overlapping DNA barcodes from different fragments of genomic DNA and aligning them onto the genome, a genome map can be

produced.

We are interested in the particular technique that uses nanochannel confinement for genome mapping as it offers several advantages over other barcoding techniques. First, DNA molecules in this method can remain in solution and at equilibrium. In contrast, other methods involve a stetched state far from equilibrium, which is typically produced by some kind of flow. Gathering data at equilibrium leads to high reproducibility making it a robust measurement. In fact, the fractional extension of DNA achieved in nanochannel-mapping has extremely small thermodynamic fluctuations, as opposed to substantial 'hydrodynamic' fluctuations (due to flow and associated hydrodynamic interactions) in other methods. Second, molecular alignment in nanochannel-barcoding is relatively much easier as the stretching is due to confinement by the walls. Third, there is great potential for throughput of this technique to be improved by using fluid flow and/or electrophoresis. Fourth, obtaining multiple measurements of the fluctuating molecule at equilibrium can reduce error in measurements (Tegenfeldt et al., 2004) and thereby increasing the accuracy of the final genome map. Indeed, recent work has shown that time-resolved data analysis improves genome map alignment, thus improving the accuracy of genome mapping, especially for complex samples such as cancer cells (Sheats et al., 2015). However, note that the current technology by BioNano Genomics uses 'static' information only from a single snapshot as opposed to 'dynamic' data of many different frames over time. Nonetheless, nanochannel-based genome mapping has been successfully applied for bacteriophage strain typing (Grunwald et al., 2015), to detect structural variation in a human genome (Cao et al., 2014; Mak et al., 2016) and for hybrid assembly combined with single-molecule sequencing (Pendleton et al., 2015; Steinberg et al., 2014).

1.2 DNA as a model polymer

Aside from the obvious relevance of DNA in genomics, DNA is also employed as a model polymer in numerous polymer physics experiments (Latinwo and Schroeder, 2011; Tree et al., 2013b). Because of its biological origin, relatively monodisperse samples of DNA (Gupta et al., 2015) can be obtained relatively easily over a wide range of molecular weights. In addition, the availability of fluorescent dyes such as YOYO-1 and TOTO-1 (Nyberg et al., 2013; Kundukad et al., 2014) means single molecules of DNA can be visualized under a microscope in a straightforward manner, as shown

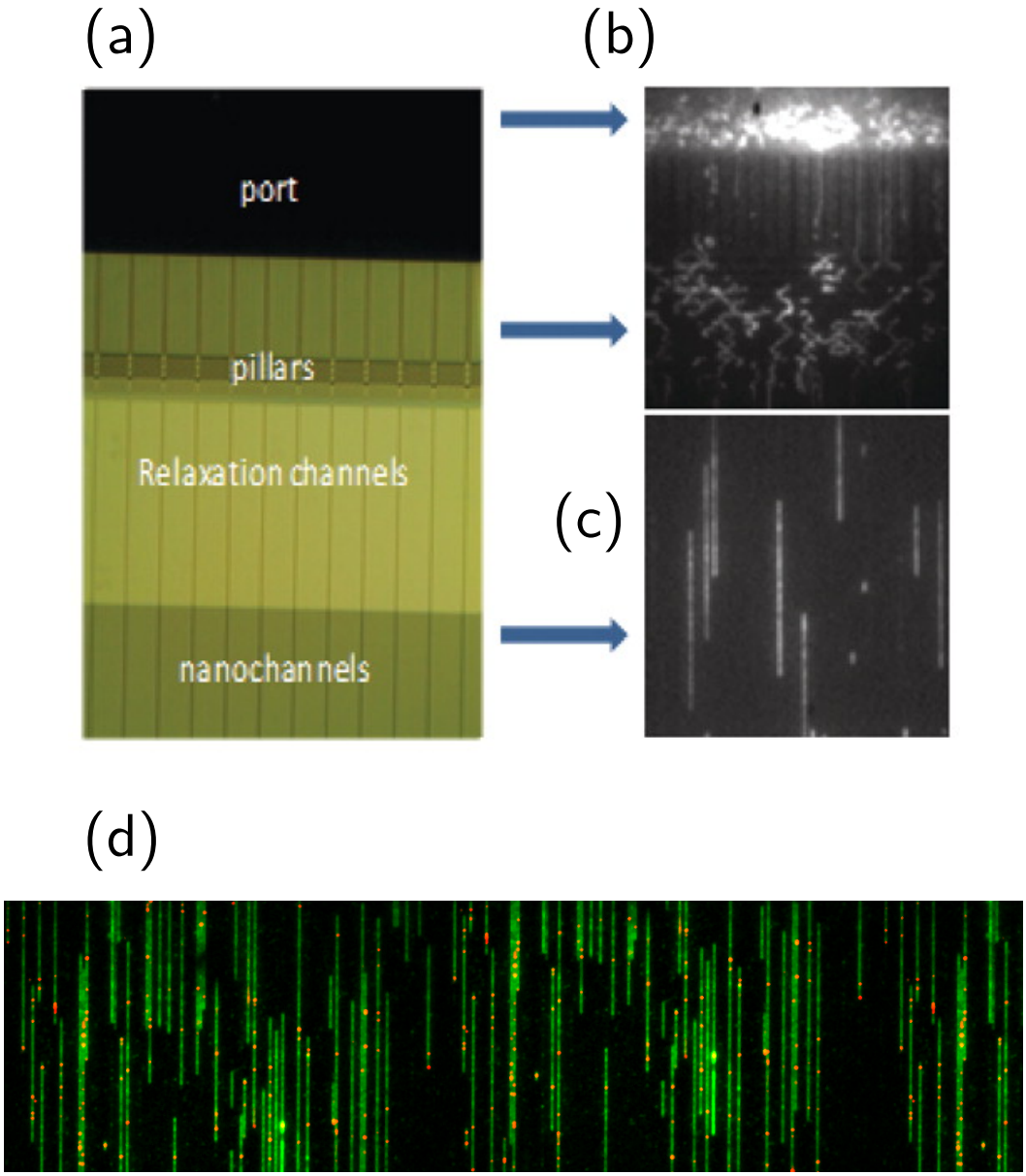


Figure 1.4. (a) The entropic gradient employed to inject DNA into nanochannels in the nanochannel device. (b) DNA translocating through these regions to the device (c) Linearized DNA molecules in a nanochannel array at equilibrium (d) Composite image of stretched DNA molecules in a nanochannel array. The backbone is in green and the red dots represent the probes. Parts (a), (b) and (c) were reproduced from Das et al. (2010). The image in (d) was taken from `http://www.bionanogenomics.com/`.

in Figure 1.4. As a result, DNA has long served as a model system for equilibrium (Smith et al., 1996; Yoshikawa et al., 1996) and non-equilibrium dynamics (Perkins et al., 1994; Perkins et al., 1995; Perkins et al., 1997; Schroeder et al., 2003; Robertson and Smith, 2007) of polymer solutions, particularly in single molecule studies (Shaqfeh, 2005).

With advances in nanofabrication techniques, DNA has naturally been adopted as a model polymer for studying the physics of confined polymers. Ever since the seminal studies from Bob Austin's group at Princeton University (Tegenfeldt et al., 2004), DNA has been used as a model system to understand the behavior of confined polymers in various environments. In fact, DNA has been used a model system to gain insight into statics and dynamics of polymers in channel (Reisner et al., 2005; Persson et al., 2009; Riehn et al., 2007; Khorshid et al., 2014; Alizadehheidari et al., 2015; Gupta et al., 2015; Iarko et al., 2015) and slit (Balducci et al., 2007; Hsieh and Doyle, 2008; Tang et al., 2010; Klotz et al., 2015) confinement in a number of studies.

1.3 Research Outline

Despite the success of nanochannel-based genome mapping technology, and the applicability of DNA as a model polymer to understand the physics of single polymers in confinement, some aspects of the behavior of DNA in confinement are still not well understood. In fact, the behavior of DNA in such environments is a complicated function of elecrostatic, hydrodynamic and entropic forces acting on the molecule. Furthermore, DNA is a semiflexible polymer characterized by a peristence length of about 50 nm. Often, the channel sizes used for DNA barcoding and polymer physics studies of nanoconfined DNA are in the range of 40-100 nm. This interplay of the length scales used to confine DNA and its persistence length further complicates the problem.

The objective of this dissertation is to use a coarse-grained model for DNA that we previously developed in our group (Tree et al., 2013b) to tackle some of the key unanswered questions surrounding DNA in nanochannel confinement via computer simulations. Our approach allows us to perform a wider sweep of the parameter space in a relatively short period of time as opposed to the expensive fabrication techniques that one has to adopt if one attempts to achieve the same goal through an experimental route. Moreover, our simulations can go beyond the realms of the phase space that are amenable to analytical solutions of statistical treatments of the confined polymer system. Since our model can be generically used to represent semiflexible polymers, our

work is not only relevant to confinement of DNA but to confinement of semiflexible polymers in general. We anticipate our approach and findings can also be extended to other biological systems, such as bacterial genome organization and chromosome aggregation (Marenduzzo et al., 2010), which contain biomolecules in confinement.

With aforementioned goals in mind, Chapter 2 summarizes recent developments in the field of nanoconfined DNA. We begin our discussion with the wormlike chain model, which has been the workhorse-model in relation to semiflexible polymers, such as DNA, for polymer physicists over many decades. We review experimental, theoretical and simulation studies surrounding channel confinement of semiflexible polymers, while pointing out open questions that will become the subject of this dissertation.

In Chapter 3, we describe the Monte Carlo simulation methods, which form the core of this dissertation. By illustrating elementary chain growth methods, such as simple sampling and Rosenbluth sampling, we motivate the need for a method that can simulate long chains that can reach molecular weights suitable for comparison with experiments and theories of polymer physics. We then explain the methodology of the pruned-enriched Rosenbluth method (PERM) in the context of a 2D self-avoiding walk. Recognizing that lattice methods are inadequate to study semiflexible polymers in confinement, we move on to discuss how we can perform off-lattice simulations using PERM. We first introduce off-lattice PERM for the simple case of a freely jointed chain. We end the chapter by elucidating our PERM algorithm for a discrete wormlike chain model by indicating a few tips and tricks that can be used to accelarate the simulation.

The properties of channel-confined semiflexible polymers are determined by a complicated interplay of chain stiffness and excluded volume effects. In Chapter 4, using PERM simulations, we study the equilibrium properties of channel-confined polymers by systematically controlling chain stiffness and excluded volume. Our calculations of chain extension and confinement free energy for freely jointed chains with and without excluded volume show excellent agreement with theoretical predictions. For ideal wormlike chains, the extension is seen to crossover from Odijk behavior (Odijk, 1983) in strong confinement to zero-stretching, bulk-like behavior in weak confinement. In contrast, for self-avoiding wormlike chains, we always observe that the linear scaling of the extension with the contour length is valid in the long-chain limit irrespective of the regime of confinement, owing to the coexistence of stiffness and excluded volume effects. We further propose that the long-chain limit for the extension corresponds to chain lengths wherein the projection of the end-to-end distance along the axis of the channel is nearly equal to

the mean span parallel to the axis. For DNA in nanochannels, this limit was identified using PERM simulations out to molecular weights of more than 1 megabase pairs; the molecular weight of λ-DNA is found to exhibit nearly asymptotic fractional extension for channels sizes used commonly in experiments.

Backfolding and hairpin formation is detrimental to genome mapping as this can scramble genomic information present in an otherwise linearized piece of DNA. As such, backfolding is the focus of the next three chapters. In Chapter 5, using PERM simulations of a discrete wormlike chain model, we provide compelling evidence in support of Odijk's prediction of two distinct Odijk regimes for a long wormlike chain confined in a nanochannel (Odijk, 2008). In both cases, the chain of persistence length l_p is renormalized into a series of deflection segments of characteristic length $D^{2/3}l_\mathrm{p}^{1/3}$, where D is the channel size. In the first (classic) Odijk regime, these deflection segments are linearly ordered. In the second Odijk regime, thin, long wormlike chains can backfold at a length scale quantified by the global persistence length. We have measured this quantity by simulations and modified Odijk's global persistence length theory to account for thermal fluctuations. The global persistence length, which is defined to be independent of the effect of excluded volume, provides the requisite closure to Odijk's scaling theory for the second regime and thus allows us to resolve much of the confusion surrounding the so-called "transition" regime for DNA confined in a nanochannel. We show that Odijk's theory for the backfolded regime correctly describes both the average chain extension and the variance about this extension for wormlike chains in channel sizes between the classic Odijk regime and the de Gennes blob regimes, with our data spanning several decades in terms of Odijk's scaling parameter ξ. Although the backfolded Odijk regime occupies a very narrow range of D/l_p, it is indeed a regime when viewed in terms of ξ and grows in size with increasing monomer anisotropy.

In Chapter 6, we compare two theories for backfolding of semiflexible polymers, such as DNA, confined in circular tubes of diameter of the order of the persistence length of the macromolecule. The first theory was proposed by Odijk on the basis of a one-dimensional analogue of Flory theory (Odijk, 2008), and the second theory is a cooperativity model of deflection segments and S-loops suggested by Dai et al. (2012). By performing Monte Carlo chain growth simulations of long chains, we find that Odijk's scaling theory not only captures the contour length dependence of extension of the confined chain, but also correctly predicts its asymptotic value. In contrast, the cooperativity model appears to quantify the extension only for the contour lengths that were used to parameterize

the model and systematically deviates from the simulation data as the contour length increases.

Moving beyond confinement in square channels and circular tubes, Chapter 7 discusses the effect of the asymmetry of the cross-section of the channel by considering confinement in rectangular channels. We confirm Odijk's scaling laws (Odijk, 2008) for (i) the average chain extension, (ii) the variance about the average extension, and (iii) the confinement free energy of a wormlike chain confined in a rectangular nanochannel smaller than its chain persistence length through PERM simulations of asymptotically long, discrete wormlike chains. In the course of this analysis, we also computed the global persistence length of ideal wormlike chains for the modestly rectangular channels that are used in many experimental systems. The results are relevant to genomic mapping systems that confine DNA in channel sizes around 50 nm, since fabrication constraints generally lead to rectangular cross-sections.

In Chapter 8, we switch gears to explain how the thermodynamic properties discussed in the previous chapters affect the diffusivity of channel-confined semiflexible polymers. Here, we compute the axial diffusivity of asymptotically long semiflexible polymers confined in square channels. Our calculations employ the Kirkwood approximation of the mobility tensor by combining computational fluid dynamics (CFD) calculations of the hydrodynamic tensor in channel confinement with PERM simulations of a discrete wormlike chain model. Three key results emerge from our study. First, for the classic de Gennes regime, we confirm that Brochard and de Gennes' blob theory (Brochard and de Gennes, 1977) correctly predicts the scaling of the axial diffusivity, contrary to the conclusions of previous analyses. Second, for the extended de Gennes regime (Dai et al., 2014), we show that a modified blob theory, which has been used to incorporate the effect of local stiffness on DNA diffusion in nanoslits, explains the deviation from the prediction of classic blob theory for diffusion in nanochannels. Third, we provide a calculation similar to the modified blob theory to explain the relative insensitivity of the diffusivity to channel size for channels between the extended de Gennes regime and the Odijk regime (Odijk, 1983), which is the most relevant regime for experiments and technological applications of DNA confinement in nanochannels. Our results are not only relevant to the dynamics of confined semiflexible polymers such as DNA, but also reveal interesting analogies between confinement in channels and slits.

Finally, we conclude by summarizing our findings and discuss possible directions for future work in Chapter 9.

Chapter 2

Background

This chapter provides a brief background and reviews recent literature concerning the physics of confined DNA and other semiflexible polymers, which are necessary to understand the material presented in the following chapters. This is by no means a comprehensive review and the interested reader is referred to magnificent review articles by Reisner et al. (2012) and Dorfman et al. (2013), which together provide an exhaustive discourse on the topic. Other articles of interest might be the mini-review of the dynamics of confined by DNA by Dorfman et al. (2014) and the critical review on the practice of manipulating DNA in nanofluidic systems by Levy and Craighead (2010). The doctoral theses of Tree (2014), Klotz (2015) and Werner (2015) also include excellent contemporary discussions on the subject.

2.1 DNA as a wormlike chain

DNA is often modeled as a wormlike chain with a local stiffness of the chain characterized in terms of a bending constant ϵ. It is convenient to view a wormlike chain as a differentiable space curve (Yamakawa, 1971). If $\mathbf{r}(s)$ is the position vector of a point of the space curve, where s denotes the contour distance from one end of the chain to the point under consideration, the tangent at any point with the parameter s is given by a unit vector

$$\mathbf{u} = \frac{\partial \mathbf{r}}{\partial s}, \tag{2.1}$$

where $|\mathbf{u}| = (\partial \mathbf{r}/\partial s)^2 = 1$. Because the bending energy per unit length is given by $\frac{1}{2}\epsilon\left(\frac{\partial \mathbf{u}}{\partial s}\right)^2$, where $\partial \mathbf{u}/\partial s$ is the local curvature vector, the potential energy of a chain with

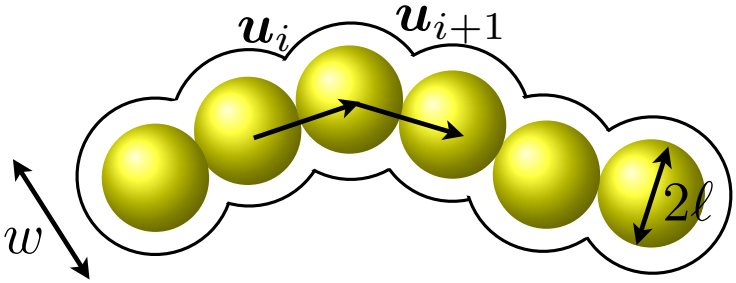


Figure 2.1. Illustration of the discrete wormlike chain (DWLC) model. It is a touching bead model with effective width w, bond length ℓ and persistence length l_p imposed by a bending potential between contiguous bonds.

contour length L can be obtained by summing up the bending energy over the whole chain,

$$U_{\text{bend}} = \frac{1}{2}\epsilon \int_0^L \left(\frac{\partial \mathbf{u}}{\partial s}\right)^2 ds. \tag{2.2}$$

This continuum approach is often employed in many statistical mechanical treatments of the wormlike chain (Yamakawa, 1971; Fredrickson, 2006). However, such a continuum formulation can not be used directly in mesoscopic simulations such as ours. Moreover, this model (Eq. 2.2), commonly known as the Kratky-Porod model, considers only phantom chains without excluded volume. In contrast, polymers such as DNA are made up of chemical entities that interact with each other and with the solvent through a variety of electrostatic and steric interactions, giving rise to an excluded volume (Rubinstein and Colby, 2003). As such, the discretization of the wormlike chain model and the incorporation of excluded volume effects is the subject of Section 2.1.1.

2.1.1 Discrete wormlike chain model

Consider a wormlike chain with a bending constant ϵ and a contour length L. Wormlike chains under good solvent conditions, which constitute an important part of this dissertation, exhibit repulsive interactions between monomer units. To account for these repulsive interactions, we associate a hard core effective width for the chain, w. Note the limit $w = 0$ represents an ideal chains with zero excluded volume. We then discretize the chain as a chain of $N + 1$ touching beads that are connected by rods of length ℓ, as shown in Figure 2.1.

Within this description, Eq. 2.2 can be discretized as

$$U_{\text{bend}} = \frac{\epsilon}{\ell} \sum_{i=1}^{N-1} \left(1 - \mathbf{u}_i \cdot \mathbf{u}_{i+1}\right), \tag{2.3}$$

$\mathbf{u}_i$ is the unit vector pointing in the direction of the bond vector connecting bead i and bead $i+1$. Note that we have expressed the square of the difference between the unit vectors in the discretized form of the derivative $|\mathbf{u}_{i+1} - \mathbf{u}_i|^2$ as $2\left(1 - \mathbf{u}_i \cdot \mathbf{u}_{i+1}\right)$. The hard core repulsive width for DNA is imposed between all pairs of non-contiguous beads by the potential U_{EV}, defined as

$$\beta U_{\text{EV}} = \begin{cases} \infty & \text{if } r_{ij} \leq w, \\ 0 & \text{otherwise.} \end{cases} \tag{2.4}$$

where r_{ij} is the distance between the centers of the beads with position vectors $\mathbf{r}_i$ and $\mathbf{r}_j$ respectively.

The configurational partition function of a wormlike chain with excluded volume is

$$Z = \int \exp\left[-\beta U\left(\{\mathbf{r}\}\right)\right] d\{\mathbf{r}\}, \tag{2.5}$$

where the integral is evaluated over all the configurations $\{\mathbf{r}\}$ of the chain. Here, $U(\{\mathbf{r}\})$ is the total potential energy of a configuration, given by $U(\{\mathbf{r}\}) = U_{\text{bend}}(\{\mathbf{r}\}) + U_{\text{EV}}(\{\mathbf{r}\})$. The ensemble average of a thermodynamic quantity, such as the mean square end to end distance, is calculated as

$$\left\langle R^2 \right\rangle = Z^{-1} \int \mathbf{R}\left(\{\mathbf{r}\}\right) \cdot \mathbf{R}\left(\{\mathbf{r}\}\right) \exp\left[-\beta U\left(\{\mathbf{r}\}\right)\right] d\{\mathbf{r}\}, \tag{2.6}$$

where $\mathbf{R}\left(\{\mathbf{r}\}\right)$ denotes the end-to-end vector of a given configuration.

We define a dimensionless quantity, $\kappa = \beta\epsilon/\ell$, where $\beta = 1/k_{\text{B}}T$ is the inverse of the thermal energy of the system. Substituting for ϵ/ℓ in Eq. 2.3, the bending potential becomes

$$\beta U_{\text{bend}} = \kappa \sum_{i=1}^{N-1} \left(1 - \mathbf{u}_i \cdot \mathbf{u}_{i+1}\right). \tag{2.7}$$

Wormlike chains can also be characterized by a length scale for bending commonly known as the persistence length, l_{p}. From the definition of the persistence length and the bending potential in Eq. 2.7 (see Appendix A), the persistence length of the chain

in terms of the parameter κ in the DWLC model can be expressed as

$$\frac{l_{\mathrm{p}}}{\ell} = \frac{\kappa}{\kappa - \kappa \coth(\kappa) + 1}. \tag{2.8}$$

The Kuhn length, b, which is often defined as twice the persistence length, is

$$\frac{b}{\ell} = \left(\frac{\kappa - 1 + \kappa \coth \kappa}{\kappa + 1 - \kappa \coth \kappa}\right). \tag{2.9}$$

Note that the Kuhn length is not exactly twice the persistence length, although this approximation is good enough for most practical purposes. For the definition of the Kuhn length and the derivation of Eq. 2.9, we refer the reader to Appendix A.

Note that the DWLC model does not consider many of the chemical details that are often important to determine the properties of short DNA, whose length is of the order of $\mathcal{O}(100)$ base pairs. We assume that such chemical details can be lumped into the two length scales in the model, the persistence, l_{p}, and the effective width, w. This kind of coarse-graining allows us to study length scales comparable to genomic length DNA, where the chemical details generally smooth out to reveal universal behavior (Freed, 1987; Tree et al., 2013b). The persistence length and the effective width of DNA are determined by a complex interplay of the charge on the DNA molecule (DNA is negatively charged) and the counter-ions present in the solvent (Baumann et al., 1997; Dorfman, 2010; Dorfman et al., 2013). Moreover, the persistence length and the effective width of DNA are known to be affected by the intercalating dyes that are often used to visualize DNA molecules in experiments. It is noteworthy that there is still some confusion as to how the solvent conditions and the presence of dyes affects these length scales (Hsieh et al., 2008; Nyberg et al., 2013; Kundukad et al., 2014; Iarko et al., 2015). In spite of the debate surrounding the dependence of the persistence length of DNA on solvent conditions, it is generally accepted that the persistence length of DNA in high ionic strength conditions is around 50 nm (Bustamante et al., 1994).

Figure 2.2 shows the dependence of the persistence length and the effective width on the ionic strength plotted using the empirical relation from Dobrynin (Dobrynin, 2006; Hsieh et al., 2008) and Stigter's theory (Stigter, 1977). We define a quantity, the monomer anistropy ratio, as l_{p}/w. It is a measure of how stiff the polymer is, with respect to its own width. The increase in the ionic strength leads to screening of electrostatic interactions within the DNA molecule, resulting in a decrease of both the persistence length and effective width with ionic strength. Quite strikingly, the l_{p}/w ratio decreases as

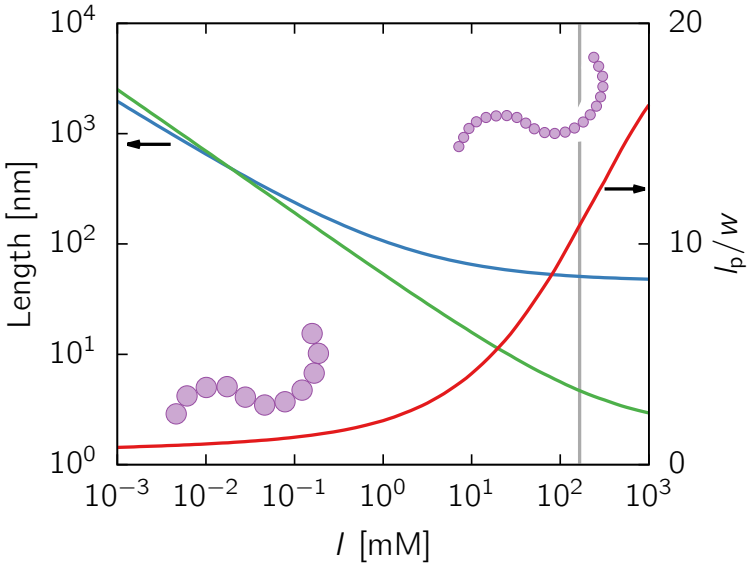


Figure 2.2. Ionic strength dependence of the persistence length (blue line) and effective width (green line) for DNA. The vertical gray line represents the ionic strength of 5×TBE (Hsieh et al., 2008). The red line is the monomer anistropy, l_p/w. The schematics show the increase in monomer anisotropy with ionic strength. Reproduced from Tree et al. (2013b).

the ionic strength is lowered because the effective width increases more rapidly than the persistence length at low ionic strengths. In fact, there is crossover a point at which the effective width seemingly becomes more than the persistence length in Fig 2.2. This absurd occurence is not observed in reality and is due to the fact that the linearization of the non-linear Poisson-Boltzmann equations in Stigter's theory (Stigter, 1975; Stigter, 1977; Schellman and Stigter, 1977) is no longer valid at such low ionic strengths.

2.2 Wormlike chains in free solution

2.2.1 Static properties

The properties of unconfined wormlike chains in free solution are well studied and are part of many polymer physics text books (Rubinstein and Colby, 2003; Hiemenz and Lodge, 2007; Grosberg and Khokhlov, 1994). One such property is the mean end-to-end distance, a common measure of the size of a polymer. The mean square end-to-end distance, $\langle R^2 \rangle$, of an ideal wormlike chain following the Kratky-Porod model can be

written as (Rubinstein and Colby, 2003)

$$\langle R^2 \rangle = 2l_p L - 2l_p^2 \left[1 - \exp\left(-\frac{L}{l_p} \right) \right]. \tag{2.10}$$

One can infer two simple limiting cases from Eq. 2.10. If $L \ll l_p$, the root mean square end to distance scales as $R \sim L$. In other words, the size of the polymer is proportional to contour length signifying the rod-like nature of the polymer at these length scales. However if $L \gg l_p$, $R \sim L^{1/2}$, which is characteristic of a random walk. Thus, a wormlike chain behaves like a rod for small contour lengths that are comparable to l_p, and behaves like a random walk (or a random coil) for much larger contour lengths. Therefore, wormlike chains are also often referred to as semiflexible polymers. Accordingly, we use the terms semiflexible and wormlike interchangeably in the rest of the thesis.

Flory proposed that the size of a polymer in a good solvent can be estimated by the balance of entropy and excluded volume (de Gennes, 1979; Rubinstein and Colby, 2003). The Flory approach, commonly known as Flory theory, consists of writing the free energy of the chain as sum of the contributions of the entropy and excluded volume to the free energy respectively as

$$\frac{F}{k_B T} \simeq \frac{R^2}{L l_p} + \frac{\mathcal{N}^2 v_{ex}}{R^3}. \tag{2.11}$$

The dimensionless parameter $\mathcal{N} = L/l_p$ represents the number of segments of the chain and $v_{ex} = l_p^2 w$ is the excluded volume of each segment. Minimizing the free energy yields the equilibrium end-to-end distance of the polymer

$$R \sim l_p \left(\frac{w}{l_p} \right)^{2\nu - 1} \left(\frac{L}{l_p} \right)^{\nu}, \tag{2.12}$$

where $\nu = 3/5$ is the Flory exponent. Despite the success of Flory theory in explaining experimental observations, it is now established that both the terms in Flory theory are wrong (Rubinstein and Colby, 2003; de Gennes, 1979; des Cloizeaux, 1976). In fact, Flory theory produces the (almost) correct result due to a fortuitous cancelation of errors. More sophisticated techniques, such as renormalization group theories and Monte Carlo simulations, have revealed that the actual value of ν is 0.587 597(7) (Clisby, 2010).

Note that the scaling in Eq. 2.12 is only valid in the long-chain limit when the chain is able to experience excluded volume interactions between distal segments. To quantify the length scale at which excluded volume interactions become important, a length scale

known as thermal blob is often used. The thermal blob contour length, l_{th} is obtained by equating the excluded volume free energy term in the Flory theory to $k_{\mathrm{B}}T$, which results in

$$l_{\mathrm{th}} \sim \frac{l_{\mathrm{p}}^3}{w^2}. \tag{2.13}$$

Therefore, a real wormlike chain exhibits three distinct scaling behaviors for the size of the chain as a function of contour length: (i) $R \sim L$ if $L < l_{\mathrm{p}}$; (ii) $R \sim L^{1/2}$ if $l_{\mathrm{p}} < L < l_{\mathrm{th}}$ and (iii) $R \sim L^{\nu}$ if $L > l_{\mathrm{th}}$ (Tree et al., 2013b). Another widely used metric for the size of the polymer is its radius of gyration, R_{g}, which we will frequently use in the rest of the thesis. The scaling of R_{g} with respect to L is identical to the scaling of R in the respective limits.

2.2.2 Dynamic properties

In addition to the static properties, such as the size of the polymer and its free energy, properties such as the diffusion coefficient are important to understand the motion of the polymer in the solvent.

We begin the discussion of the diffusion of a polymer chain by first considering a simple case of the motion of a particle. If a force $\mathbf{F}$ is applied to a small particle in a solvent, it soon achieves a constant velocity $\mathbf{v}$ that is related to the friction coefficient ζ by

$$\mathbf{F} = \zeta \mathbf{v}. \tag{2.14}$$

The diffusion coefficient $\mathcal{D}$ of the particle is related to the friction coefficient through the Einstein relation:

$$\mathcal{D} = \frac{k_{\mathrm{B}}T}{\zeta}. \tag{2.15}$$

A related quantity is the mobility of the particle defined as $\mu = \mathcal{D}/k_{\mathrm{B}}T$. Along the same lines, a force $\mathbf{F}_{\mathrm{CM}}$ on the center of mass of a polymer chain would result in a velocity $\mathbf{v}_{\mathrm{CM}}$ related by

$$\mathbf{F}_{\mathrm{CM}} = \zeta \mathbf{v}_{\mathrm{CM}}, \tag{2.16}$$

where ζ now is the center of mass friction coefficient of the chain. Likewise, one can define a center of mass diffusivity, $\mathcal{D}_{\mathrm{CM}}$, and a center of mass mobility μ_{CM}.

The diffusion of a polymer is different from that of a particle in that the motion of a segment of the polymer is affected by the collective motion of all other segments. For instance, force applied by a bead (assumed to be a point particle) on the solvent

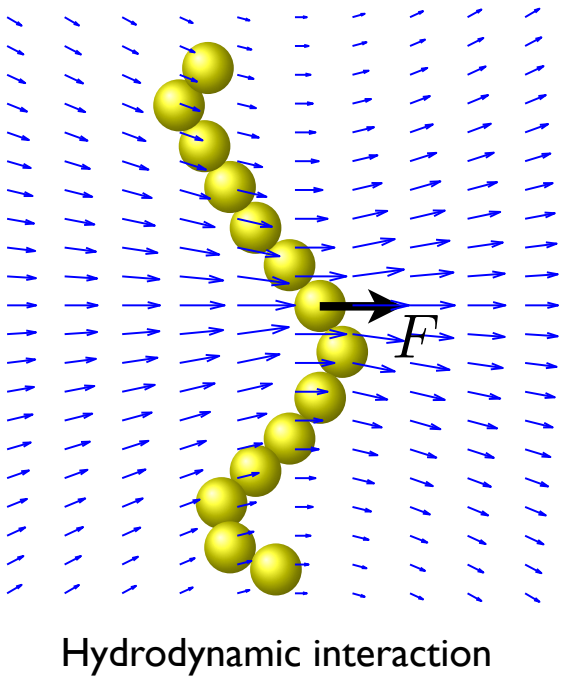


Figure 2.3. Velocity perturbation due to the point force applied by a single bead on the solvent affecting the motion of other beads.

sets up a velocity pertubation in the solvent, which affects the motion of other beads, as shown in Figure 2.3. These interactions, referred to as hydrodynamic interactions (HI), can occur even between distal segments of the chain. The velocity perturbation at any point with a position vector $\mathbf{r}$ with point source as the origin is quantified by the Oseen-Burgers tensor $\mathbf{\Omega}$ and is related to the force $\mathbf{F}$ applied by the particle on the solvent by

$$\mathbf{v}(\mathbf{r}) = \mathbf{\Omega}(\mathbf{r}) \cdot \mathbf{F}, \tag{2.17}$$

where

$$\mathbf{\Omega} = \frac{1}{8\pi\eta r}\left(\mathbf{I} + \frac{\mathbf{rr}}{r^2}\right) \tag{2.18}$$

For the motion of a single bead, Stokes law suggests that friction is related to the radius of the bead, a, by $\zeta_{\text{bead}} = 6\pi\eta a$. If we do not consider hydrodynamic interactions, the total friction factor of the chain is simply the sum of the friction factors of all the beads. In other words, the friction factor of the chain scales as the $\zeta_{\text{chain}} \sim N_{\text{b}}\eta a \sim \eta L$, where N_{b} is the number of beads in the chain with a hydrodynamic radius, a. This model of a chain with no HI is known as the Rouse model (Rouse, 1953), with the corresponding diffusivity $\mathcal{D}_{\text{CM}}^{\text{Rouse}}$ scaling as

$$\mathcal{D}_{\text{CM}}^{\text{Rouse}} \sim \frac{kT}{\eta L} \tag{2.19}$$

For long chains, the dynamics is typically dominated by HI. Zimm (1956) proposed that in this limit, the chain entrains the solvent inside it as it moves and thus behaves almost like a rigid sphere of radius, R_{h}, the hydrodynamic radius of the chain. This hydrodynamic radius also happens to be approximately proportional to the radius of gyration, R_{g}, of the polymer, although there are known subtleties associated with the relation between the two length scales (Tree et al., 2013b; Mansfield and Douglas, 2013). Therefore, the friction factor of the chain follows a similar form as the Stokes-Einstein relation for a sphere, $\zeta_{\text{chain}} \sim \eta R_{\text{h}}$ or $\zeta_{\text{chain}} \sim \eta R_{\text{g}}$. As as result, the diffusion coefficient is given by

$$\mathcal{D}_{\text{CM}}^{\text{Zimm}} \sim \frac{k_{\text{B}}T}{\eta R_{\text{g}}}. \tag{2.20}$$

A chain obeying the Zimm model is often called a non-draining chain as the nearly solid sphere-like behavior of the chain does not let the solvent drain through it. In contrast, the Rouse chain is called a freely-draining chain as the solvent drains all the segments of the chain.

Semiflexibility affects the hydrodynamic interactions between the segments of the

chain and thus the scaling of the diffusivity of semiflexible polymers is more subtle. For instance, for $L < l_p$, the diffusion is rod-like, where the diffusivity scales as (Yamakawa, 1971)

$$\mathcal{D}_{CM} \sim \frac{k_B T}{\eta L} \ln\left(\frac{L}{2a}\right), \tag{2.21}$$

in the limit $L \gg a$, where a is the local hydrodynamic radius of the rod. Longer chains that are not big enough to form a thermal blob behave as Gaussian coils with random walk statistics ($R \sim L^{1/2}$). In this case, the diffusivity is known to scale as (Kirkwood and Riseman, 1948; Yamakawa, 1971)

$$\mathcal{D}_{CM} \sim \frac{k_B T}{\eta L}\left[1 + c\left(\frac{L}{l_p}\right)^{1/2}\right], \tag{2.22}$$

where c is a constant. Eventually, wormlike chains with excluded volume reach the non-draining Zimm limit, where

$$\mathcal{D}_{CM} \sim \frac{k_B T}{\eta R_g} \sim \frac{k_B T}{\eta L^{\nu}}, \tag{2.23}$$

given that $R_g \sim L^{\nu}$ in swollen coil limit. Note that the simplified picture of dynamics outlined here neglects the subtleties associated with the interplay of excluded volume and HI (Wang et al., 1986; Mansfield and Douglas, 2013; Mansfield et al., 2007), and some of the preaveraging approximations in the theory of Kirkwood and Riseman (1948) used to arrive at the results presented here (Dorfman et al., 2014).

2.3 DNA and other semiflexible polymers in nanochannel confinement

We now turn our attention to the physical properties of DNA and such semiflexible polymers in confinement. The challenge in describing the properties of a confined DNA lies in the length scales characterizing the problem, as illustrated in Figure 2.4. While the backbone of the DNA has a bare width of around 2 nm, DNA is highly charged in solution and its electrostatic interactions give rise to an effective width, w, of about 5 nm. The persistence length of double stranded DNA in a high ionic strength buffer, $l_p = 50$ nm, is an order of magnitude higher than the effective width.

We are interested here only in rather long molecules of DNA containing many per-

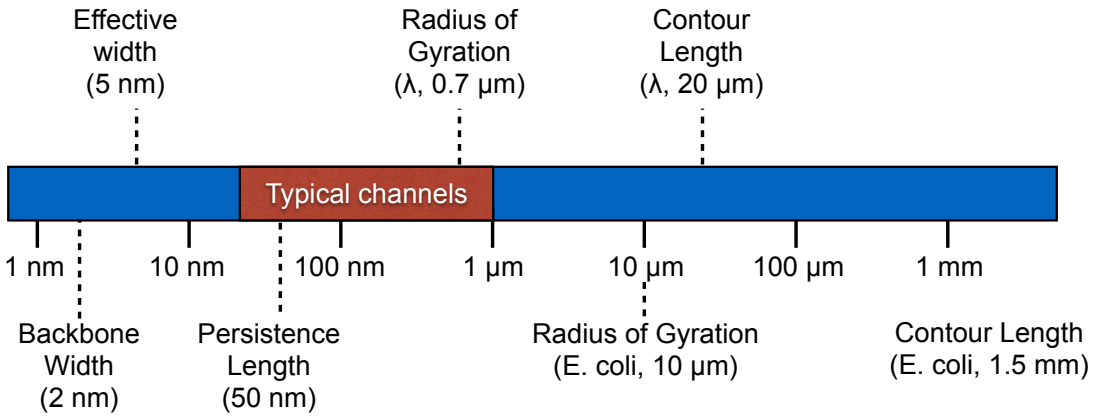


Figure 2.4. Length scales of the bare backbone width, the effective width and the persistence length associated with a DNA molecule under typical solvent conditions. The contour length and radius of gyration of λ-DNA and genomic length E. coli DNA are also shown. Typical channel sizes used in experiments span sizes starting from about 30 nm (strong confinement) to sizes as big as 1000 nm (weak confinement). Reproduced from Dorfman et al. (2014).

sistence lengths, which introduces two more length scales much larger than l_p: (i) the radius of gyration, R_g, of the coiled polymer in free solution and (ii) its contour length, L. The DNA obtained from the λ-phage virus is one of the most widely used DNA molecules in biophysics experiments, including many experiments on confined DNA. Commonly known as λ-DNA, this molecule has 48,502 base pairs and it is readily available at high concentrations. The radius of gyration, R_g, of λ-DNA is around 700 nm (Smith et al., 1996), although the assumptions underlying this analysis have been called into question recently (Tree et al., 2013b; Mansfield and Douglas, 2013). When λ-DNA is stained with a fluorescent dye for visualization, its contour length L is approximately $20\,\mu$m.

We only consider channel sizes, D, such that the confined chain feels the effect of the walls from all sides. This condition can be expressed as $D < R_g$. Accordingly, on the upper end of the range of channel sizes in Figure 2.4, we have a weak confinement regime, $D \approx R_g$. As the confinement increases, the chain becomes increasingly stretched in the unconfined directions. The stretching of the chain depends on the confinement size and other physical properties of the chain. The behavior of the confined chain can be classified into various regimes depending on the relative sizes of the length scales involved, as described below.

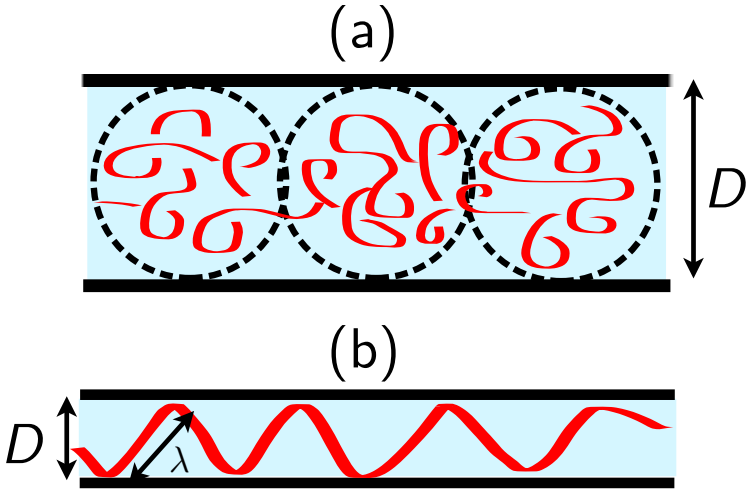


Figure 2.5. (a) Illustration of blobs in de Gennes' theory (b) Illustration of deflection segments in the Odijk regime.

2.3.1 de Gennes' blob theory

Almost 40 years have passed since Brochard and de Gennes (1977) considered the problem of a flexible polymer confined in a channel (or tube). Later, de Gennes (1979) summarized the results of their theory in his book based on the seemingly simple blob theory. To recapitulate the results from the work of Brochard and de Gennes (1977) via the blob theory approach, we consider confinement of a chain in a channel with square cross section of size length D. However, note that their results can be generalized to confinement in circular tubes (de Gennes, 1979) and rectangular channels (Werner and Mehlig, 2015) as well.

For channel confinement of a chain with persistence length l_p and effective width w, such that $D \gg l_\mathrm{p}$, the Brochard and de Gennes blob theory consists of the following assumptions:

1. The blob size is equal to the confinement size, D, where the conformation of the chain inside a blob is given by that of a three-dimensional, self-avoiding coil in good solvent conditions such that the blob size is $D \sim \mathcal{N}_\mathrm{blob}^{\nu}$, where $\mathcal{N}_\mathrm{blob}$ is the number of persistence lengths inside the blob.

2. The blobs themselves form a 1D self-avoiding walk within the channel, as shown in Figure 2.5(a).

3. Each blob is a hydrodynamically isolated object such that total friction on each blob is $\zeta_{\text{blob}} \sim \eta D$, where η is the viscosity of the fluid.

The theory applies for relatively weak confinement for channel sizes larger than $D \gtrsim l_{\text{p}}^2/w$ (Odijk, 2008).

The theoretical analysis proceeds by first computing the number of persistence lengths in a blob,

$$\mathcal{N}_{\text{blob}} \sim D^{5/3} w^{-1/3} l_{\text{p}}^{-4/3}, \tag{2.24}$$

where we use the classical Flory exponent $\nu = 3/5$ for convenience and add the standard corrections for a wormlike chain (Schaefer et al., 1980). The total number of blobs $\mathcal{N}_{\text{blob}} = \mathcal{N}/\mathcal{N}_{\text{blob}}$, where $\mathcal{N} = L/l_p$ is the number of persistence lengths in a chain of contour length L. The average extension in channels is given by

$$\langle X \rangle \simeq (n_{\text{blob}}) D \sim L \left(\frac{w l_{\text{p}}}{D^2} \right)^{1/3}. \tag{2.25}$$

The total friction on the chain is given by the sum of the friction on all the blobs,

$$\zeta \simeq n_{\text{blob}} \zeta_{\text{blob}} \sim \eta L \left(\frac{w l_p}{D^2} \right)^{1/3}. \tag{2.26}$$

The axial (channel) mobility of a long chain is thus given by

$$\mu \equiv \frac{1}{\zeta} \sim \frac{1}{\eta L} \left(\frac{D^2}{w l_{\text{p}}} \right)^{1/3}, \tag{2.27}$$

2.3.2 Odijk's theory of deflection segments

Odijk proposed that chain forms a linearly ordered array of deflection segments when it is confined in a channel with size $D \ll l_{\text{p}}$ (Odijk, 1983), as shown in Figure 2.5(b). We refer to this range of confinement as the 'Odijk regime' or the 'classic Odijk regime' owing to recent modifications (Muralidhar et al., 2014a) of the classical theory of Odijk. By using the results of Yamakawa and Fujii (1976) for the properties of a wormlike chain near the rod limit, Odijk estimated that the deflection segment scales as $\lambda \simeq D^{2/3} l_{\text{p}}^{1/3}$. One can then write an expression for the average extension of the chain as confined in a square channel

$$\langle X \rangle = L \left[1 - \alpha_{\square} (D/l_{\text{p}})^{2/3} \right], \tag{2.28}$$

where the geometry-dependent prefactor α has been computed by recent studies as $\alpha_{\square} = 0.18274$ (Burkhardt, 1997; Burkhardt et al., 2010). Because of the extremely tight confinement, the chain is almost fully elongated in this regime. For instance, for a channel size that is half the persistence length, the fractional extension, $\langle X \rangle / L \approx 0.9$. The variance of the extension scales as

$$\delta X^2 = \beta_{\square} \frac{D^2}{l_{\mathrm{p}}} L. \tag{2.29}$$

Burkhardt and co-workers (Burkhardt et al., 2010) have also computed the prefactor for the variance of the extension as $\beta_{\square} = 0.00956$. Naturally, the Odijk regime is the targeted regime of confinement for DNA barcoding technology, owing to the complete linearization of the chain and the low variance of extension exhibited by DNA in this range of channel sizes.

For channels corresponding to the Odijk regime ($D \ll l_{\mathrm{p}}$), the diffusion coefficient or the mobility can be estimated using the expression for motion of a single semiflexible polymer in a tube given by (Reisner et al., 2012; Morse, 1998)

$$\mu \simeq \frac{1}{2\pi\eta\langle X \rangle} \ln\left(\frac{H}{2a}\right), \tag{2.30}$$

where a is the local hydrodynamic radius of the chain (Dorfman et al., 2014). Since the chain is almost at full extension, it is common to replace the extension with the contour length in Eq. 2.30.

2.3.3 Recent studies on semiflexible polymers in nanochannels

The theories of de Gennes (Daoud and de Gennes, 1977; Brochard and de Gennes, 1977) and Odijk (Odijk, 1983) were accepted dogma until experiments from Reisner et al. (2005) revealed discrepancies with predictions from both approaches. In particular, their experiments showed an apparent scaling of $\langle X \rangle / L \sim D^{-0.85}$ as opposed to the predicted slope of -0.70 in the de Gennes regime, as shown in Figure 2.6 (a). These findings and related discrepancies have been reproduced with better statistics in tapered channels (Persson et al., 2009; Gupta et al., 2014). This led to a number of theoretical, experimental and simulation studies with the aim of resolving these discrepancies.

Odijk (Odijk, 2006) again led this resurgence of interest in the problem by suggesting that DNA and other semiflexible polymers can form hairpins and the formation of such

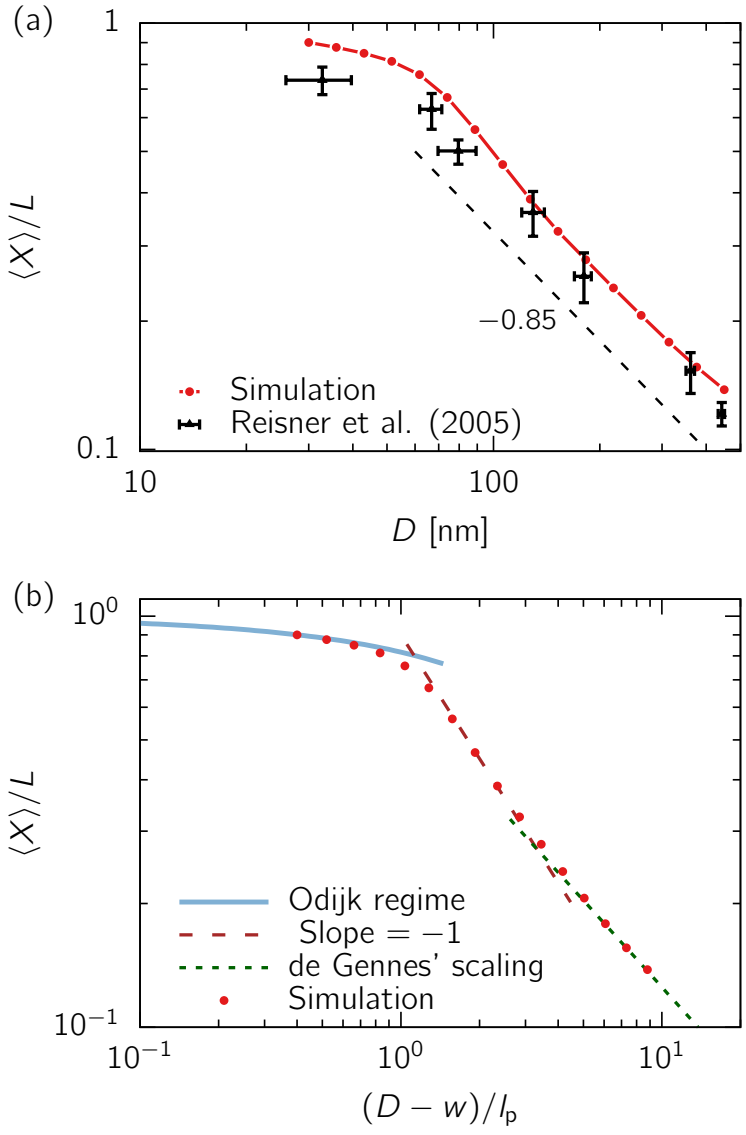


Figure 2.6. (a) Fractional extension from experiments of Reisner et al. (2005) and simulation data (Tree et al., 2013c). The apparent slope of -0.85 in the data of Reisner et al. (2005) is shown as a dashed line. (b) Simulation data showing the apparent slope, -1, in the so-called transition regime. The prediction for extension in the Odijk regime (Eq. 2.28) and the slope of the data in the de Gennes regime, $1 - 1/\nu = -0.701$, obtained from the actual value of $\nu = 0.5876$ are also shown (Eq. 2.25). Simulation parameters: $l_p = 50\,\text{nm}$; $w = 10\,\text{nm}$; $L = 18.63\,\mu\text{m}$. Image reproduced from Tree et al. (2013c).

hairpins can be used to explain the aforementioned discrepancies of the experimental data with the classical theories of Odijk and de Gennes. He then followed it up by a comprehensive theory for the thermodynamics of confined semiflexible polymers in nanochannels and nanoslits (Odijk, 2008). Odijk's theoretical work was followed by a number of simulation studies attempting to explain the scaling of the properties of DNA and other semiflexible polymers in nanoconfinement (Cifra et al., 2009; Cifra, 2009; Cifra et al., 2010; Wang et al., 2011; Tree et al., 2013a; Tree et al., 2013c; Dai et al., 2014). These simulation studies quite unanimously verified the predictions for the extension in the de Gennes and Odijk regimes embodied in Eq. 2.25 and Eq. 2.28 (Wang et al., 2011; Cifra, 2009; Tree et al., 2013a; Dai et al., 2014).

Of particular interest to us is the work from Wang et al. (2011) and Tree et al. (2013c) that explained the apparent discrepancies in the experimental data of Reisner et al. (2005). They suggested that the observed slope -0.85 is a conflation of the supposed slope of -0.70 in the blob regime for larger channel sizes, and an apparent slope of -1 in intermediate channel sizes between the Odijk regime and the blob regime. Their argument was supported by simulation data, as shown in Figure 2.6(b).

Wang et al. (2011) also provided evidence for a different regime of behavior of the confined chain for channel sizes with $D \gg l_\mathrm{p}$ and $D \ll l_\mathrm{p}^2/w$, confirming Odijk's (Odijk, 2008) earlier predictions. Odijk proposed that properties of the chain in this regime can be estimated by depicting the confined chain as a series of anisometric blobs of size $H \simeq (Dl_\mathrm{p})^{2/3}/w^{1/3}$. Wang et al. (2011) subsequently coined the term 'extended de Gennes regime' for this range of channel sizes since the scaling of the extension of the chain is exactly the same as the de Gennes regime (Eq. 2.25). However, the variance of extension in the extended de Gennes regime is independent of the channel size and scales as (Wang et al., 2011; Reisner et al., 2012)

$$\delta X^2 \sim L l_\mathrm{p}. \tag{2.31}$$

There was some confusion and controversy surrounding whether the extended de Gennes behavior is observed in simulations. Tree et al. (2013a) argued that the extension of confined semiflexible chains is characterized by a scaling of $\langle X \rangle \sim D^{-1}$, owing to a linear arrangement of ideal blobs in the channel. They suggested that a Gauss de Gennes regime briges the behavior between the Odijk and de Gennes regimes, as they did not seem to see any evidence for the extended de Gennes regime in the scaling of the confinement free energy. The debate appears to have settled now, as Dai et al. (2014)

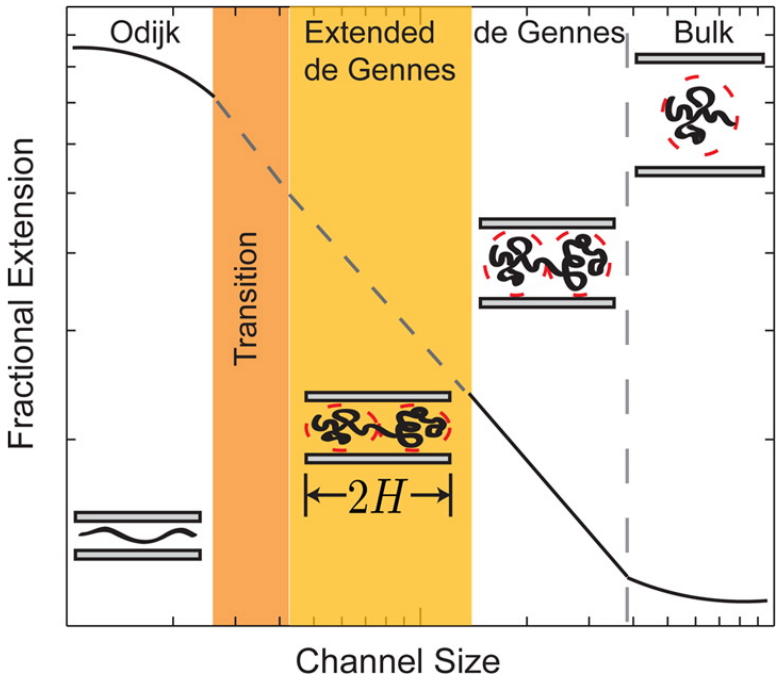


Figure 2.7. Schematic of a plot of the fractional extension of the chain against channel size for semiflexible polymer confined in a channel showing the different regimes of behavior. Image reproduced from Dorfman et al. (2013).

showed that there is indeed simulation evidence in support of the extended de Gennes regime. Dai et al. (2014) further found that the confusion about the Gauss de Gennes regime was caused due to the misinterpretation of the confinement free energy in the work of Tree et al. (2014). More recent work by Mehlig and co-workers demonstrated that the scaling of thermodynamic properties in the extended de Gennes regime can be obtained via an exactly solvable one-dimensional model (Werner and Mehlig, 2014; Smithe et al., 2015). Likewise, the scaling of the mean and variance of extension in this regime have also been corroborated in experiments (Gupta et al., 2015; Iarko et al., 2015).

Figure 2.7 summarizes our discussion about the scaling of the mean extension with respect to channel size. Note that there is still an ongoing debate regarding the so-called transition regime in between the Odijk and extended de Gennes regimes, which forms the subject of much of this thesis. The aforementioned Gauss-de Gennes model (Tree et al., 2013a), a cooperativity model based on the Zimm-Bragg model for helix-coil transitions (Dai et al., 2012), and a Flory-theory based scaling approach (Odijk, 2008) have all been employed to describe the behavior of the chain in this regime. In consideration

of the confusion surrounding the behavior, the thermodynamic properties of the chain in this regime is the subject of Chapters 5, 6 and 8.

With respect to the dynamic properties, Tree et al. (2012) computed the center of mass mobility of confined semiflexible polymers as a function of channel size. Given that the fractional extension of a chain has a one-to-one relationship with the channel size, the dependence of μ on D can also be viewed as the dependence of μ on $\langle X \rangle / L$. Their results are summarized in Figure 2.8. Their work confirmed the scaling of the mobility in the de Gennes regime predicted by blob theory (Eq. 2.27). In addition they were also able to corroborate the mobility in the Odijk regime (Eq. 2.30). Intriguingly, their work also revealed an anomalous dependence of the mobility in the transition regime. They observed that the mobility of the molecules becomes independent of the size (extension) of the chain as the stiffness (persistence length) of the chain increases. They attributed this plateau of the mobility to the screening of hydrodynamic interactions between segments of the chain due to the presence of channel walls, leading to Rouse-like behavior. Although their work attempts to explain the axial diffusivity of the chain for the entire range of confinement, it opened up a lot more questions about other details surrounding the dynamics. For instance, is the scaling of the mobility in the extended de Gennes regime similar to the scaling in the de Gennes regime? We answer questions such as this in Chapter 8.

Interestingly, no experimental data exist for the axial diffusion coefficient of a DNA molecule in nanochannels. However, measuring the relaxation time provides an indirect tool to analyze the hydrodynamics of a channel-confined chain. In fact, the experimental data of the relaxation time of DNA confined in nanochannels by Reisner et al. (2005) were explained by the mobility calculations of Tree et al. (2013c). The experiments showed an apparent slope of $\tau \sim D^{-0.9}$ for larger channels (Reisner et al., 2005), way off from the expected scaling of $D^{-1/3}$ in the de Gennes regime (Reisner et al., 2012). By using a dumbbell model, Tree et al. (2013c) showed that the longest relaxation time of the confined DNA molecule scales as the product of the normalized variance of extension and the normalized friction of the chain,

$$\tau \sim \frac{\delta X^2}{L l_\mathrm{p}} \frac{2\zeta}{\eta L}. \tag{2.32}$$

By using an $\mathcal{O}(1)$ prefactor, they were able to obtain excellent agreement with the experimental data, as shown in Figure 2.9(a). Thus, the non-monotonicity in the relaxation

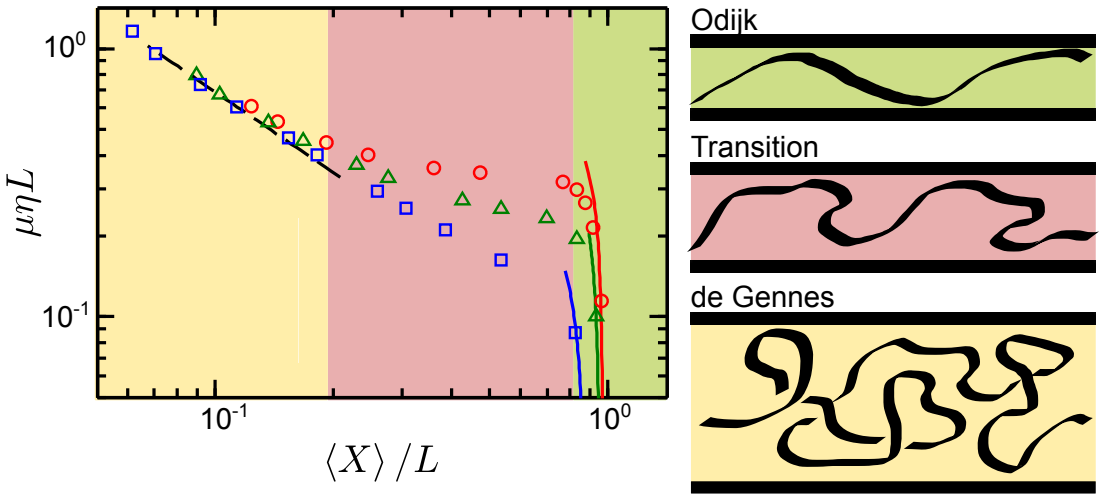


Figure 2.8. Hydrodynamic mobility, made dimensionless with the Rouse mobility, as a function of fractional extension of the chain. All simulations correspond to an effective width w = 4.6 nm, hydrodynamic radius a = 1.38 nm and L = 9.42 μm ($N + 1$ = 2048 beads). The data correspond to three different persistence lengths: 53 nm (red circles), 26 nm (green triangles), and 5 nm (blue squares). The dotted line is the scaling of Eq. 2.27. The solid lines are the approximation in Eq. 2.30 with $X = L$. The shaded regions of the plot, along with the corresponding schematics, represent different regimes of confinement. Reproduced from Dorfman et al. (2013).

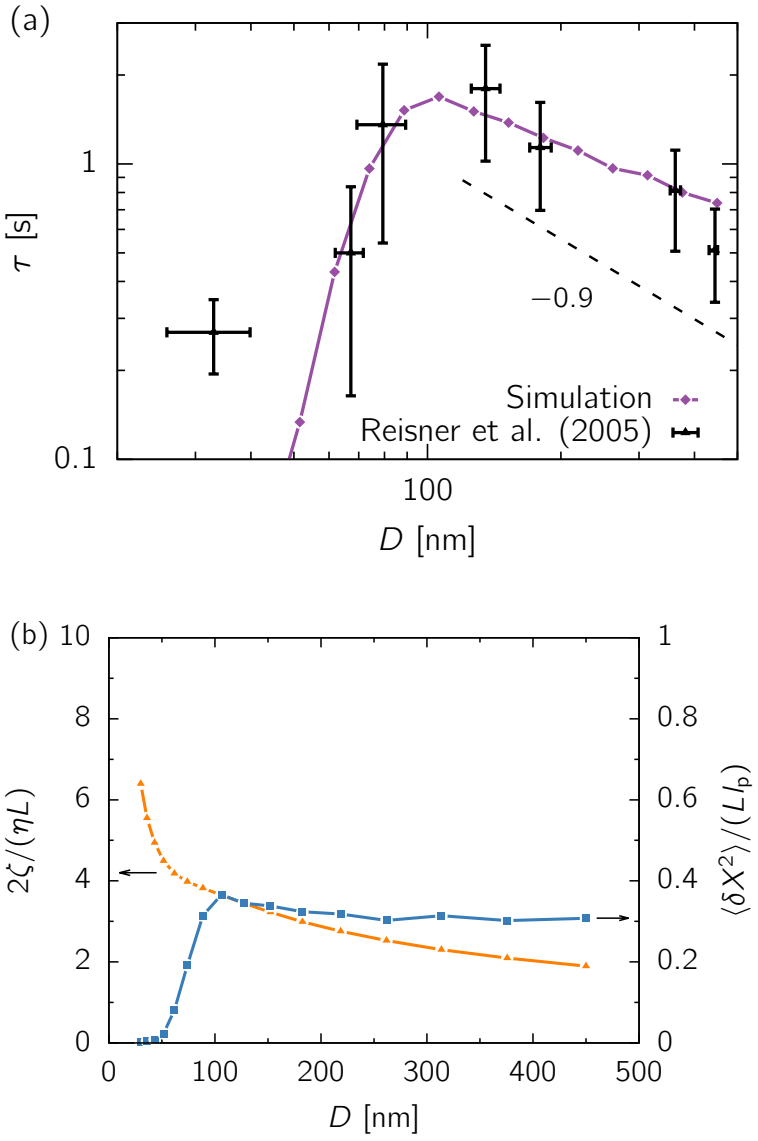


Figure 2.9. (a) Log-log plot of the relaxation time against channel size. The experimental data are from Reisner et al. (2005). The slope of -0.9 is the best-fit slope to the experimental data. The purple curve represents simulation data of Tree et al. (2013c). (b) Normalized friction of the chain (orange traingles, left panel) and normalized variance of extension (blue squares, right panel) from simulations. Reproduced from Tree et al. (2013c).

time can be attributed to non-monotonicity of the variance of extension (Figure 2.9(b), as the friction monotonically decreases with channel size. Considering that the behavior of the friction (inverse of the mobility) was reasonably well explained earlier (Tree et al., 2012), the outstanding question relates to the non-monotonicity of the variance. Accordingly, a part of Chapter 5 will be directed towards elucidating this sharp transition of the variance, which gives rise to the peak in relaxation time for channel sizes close to twice the persistence length ($D \approx 2l_{\mathrm{p}}$ assuming $l_{\mathrm{p}} = 50\,\mathrm{nm}$ for DNA) of the confined molecule.

Chapter 3

Simulation methods

Although much insight has been gained by the numerical solution of the modified diffusion equation for the wormlike chain model *without* excluded volume to calculate the properties of stiff polymers (Chen and Sullivan, 2006; Cui et al., 2008; Chen, 2013a; Gao et al., 2014) in confinement, this method is intractable in the presence of excluded volume (Werner et al., 2013). Furthermore, an analogy between a confined stiff polymer in the classic Odijk regime and a randomly accelarated particle has also been successfully used to compute the confinement free energy, the average extension and the variance of the extension for confinement in circular tubes (Yang et al., 2007) and rectangular channels (Burkhardt et al., 2010). However, this approach has limited applicability — it can only be applied to extreme confinement, where $D \ll l_\mathrm{p}$. These challenges in obtaining analytical or numerical solutions of the wormlike chain model with excluded volume in confinement means that molecular and mesoscopic simulations remain the only viable theoretical approach. The remainder of this thesis will address several open problems surrounding semiflexbile polymer confinement via a mesoscopic simulation approach.

In this thesis, we are interested mainly in the properties of semiflexible polymers in quasi-one dimensional confinement in the entire range of confinement sizes, from $D \ll l_\mathrm{p}$ to $D \gg l_\mathrm{p}$. In particular, our starting point is a polymer chain already at equilibrium in channel or tube confinement. We do not consider highly non-equilbrium phonomena such as the entrance of a DNA molecule inside a nanochannel (Mannion et al., 2006; Levy et al., 2008), even though such a non-equilibrium injection process is necessary to place the DNA molecule at equilibrium inside a nanochannel. For the calculation of equilibrium properties with the discrete wormlike chain model, there exist several methods as potential candidates. On the one hand, Metropolis Monte Carlo

methods have been used extensively to calculate thermodynamic averages in polymer solutions and melts (Binder, 1986). By permitting unphysical moves in phase space, these simulations can accelerate the process of 'importance sampling', where one attempts to efficiently sample the more favorable parts of the phase space that contribute most to the thermodynamic average (Frenkel and Smit, 2001). On the other hand, Brownian dynamics simulations can also be used to estimate equilibrium properties, with the added advantage that one can even study systems out of equilibrium accounting for the hydrodynamics of the solvent (Ermak and McCammon, 1978). However, in order to simulate experimentally relevant molecular weights of DNA and to be able to reach the long-chain limit in which the theories of polymer physics are valid (Tree et al., 2013b), the $\mathcal{O}(10^3)$ beads afforded by these methods are not sufficient. Accordingly, we employ a Monte Carlo chain growth technique, the pruned-enriched Rosenbluth Method (PERM) (Grassberger, 1997), which is known to simulate $\mathcal{O}(10^5)$ beads (Grassberger, 1997; Tree et al., 2013b).

In this chapter, our focus will be on describing the PERM methodology and in particular, explaining our implementation of off-lattice PERM, which is rather specific for our simulations of wormlike chains. We begin with motivating the need for the pruned-enriched Rosenbluth method by introducing various Monte Carlo chain growth methods in Section 3.1 such as simple sampling (Section 3.1.1), Rosenbluth sampling (Section 3.1.2 and PERM (Section 3.1.3) in the context of a 2D self avoiding walk on a lattice. In Section 3.2.1, we explain the modifications that we did to move beyond simple lattice simulations to off-lattice simulations of a self-avoiding walk by taking the example of a freely-jointed chain. Finally, in Section 3.2.2, we discuss how the discrete wormlike chain model was used to perform highly optimized PERM simulations via a parallelized Fortran code.

3.1 Monte Carlo chain growth methods

We will begin by illustrating Monte Carlo chain growth methods in the context of simulation of a 2D self-avoiding walk (SAW) on a lattice. In other words, our model polymer is a chain on a 2D lattice such that no two monomers (beads) can occupy the same point on the lattice. Our goal is to compute the average thermodynamic properties, such as the mean end to end distance, $\langle R \rangle$, of a 2D SAW of a chain of contour length 200 beads. Our discussion in this chapter is mostly inspired by the notes on PERM by Prellberg

(2013).

3.1.1 Simple sampling

The first approach we consider is 'simple sampling'. Here, we place the first bead at a random point on the 2D lattice, which we fix as the origin of our coordinate system. To 'grow' the chain, we add a bead at random on one of the four neighboring sites to the first bead. As shown in Figure 3.1(a), we then keeping adding the $(i+1)$th bead at one of the four neighboring lattice sites to the ith bead until we reach the preset maximum number of beads, which in this case is 200, or we end up adding a bead to a site that is already occupied. In case the $(i+1)$th bead overlaps with any of the first i beads or we reach the maximum molecular weight, we stop our chain growth. One such realization of chain growth starting from the first bead until we stop the chain growth is called a 'tour'. We then continue growing a fresh tour from the very beginning until we have enough tours to obtain good estimates of thermodynamic averages. Because this is a chain growth method, we can estimate theromodynamic averages as a function of the molecular weight of the chain. Note that this ability to compute averages as a function of molecular weight in a single simulation is characteristic of chain growth methods. In other traditional methods such as molecular dynamics, Metropolis Monte Carlo and Brownian dynamics techniques, one has to perform independent simulations for different molecular weights.

If we perform chain growth for N_t tours, the average end to end distance of a chain with n steps $(n+1$ beads$)$ R_n is given by

$$\langle R_n \rangle = \frac{\sum_{t=1}^{t=N_\mathrm{t}} R_n^t}{N_\mathrm{t}}, \tag{3.1}$$

where R_n^t is the end to end distance of a chain with n beads in the tth tour. We call this 'simple sampling' because all the walks generated by this method have equal probability. In other words, each n-step walk generated by this method has a probability of $1/4^n$. However, because there are 2.638^n 2D self-avoiding walks for 4^n n-step random walks (Prellberg, 2013), the probability of sampling a SAW exponentially decreases with n. In fact, the ratio of the number of self avoiding random walks to the number of random walks for a 50-step chain is as low as $(2.638/4)^n = 9.14 \times 10^{-10}$. This extremely low probability is reflected in the difficulty to sample chains with high n. Figure 3.2 shows

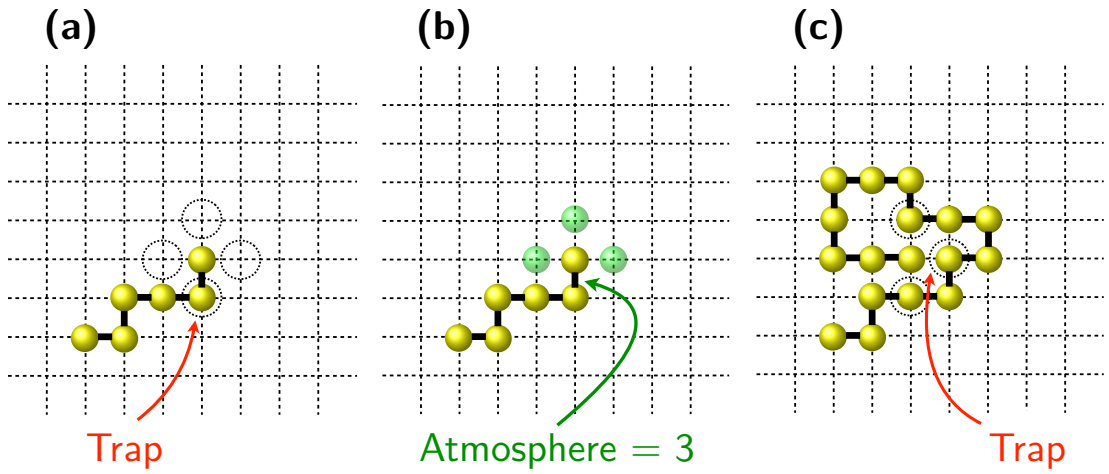


Figure 3.1. (a) Simple sampling of a 2D SAW. Because simple sampling chooses all 4 neighboring sites with equal probability, it often gets into traps. (b) Rosenbluth sampling of the same SAW biases chain growth towards self-avoiding configurations. (c) As the chain gets longer, even Rosenbluth sampling gets stuck in traps.

the number of samples as a function of the length of the chain. We observe that simple sampling is extremely inefficient to sample any chain size beyond a meagre chain length of about 10 beads.

3.1.2 Rosenbluth sampling

In order to address this 'attrition' problem, Rosenbluth and Rosenbluth (1955) proposed to bias the chain growth towards self-avoiding configurations. In their method, commonly referred to as 'Rosenbluth sampling', the chain growth at each step is chosen from one of the available (unoccupied) neighboring sites randomly, as shown in Figure 3.1(b). For a chain of $n+1$ beads or n steps, the number of unoccupied neighboring sites known as the 'atmosphere', a_n, of the configuration is calculated. The $(n+2)$th bead is then added to one of these a_n sites by picking an available site randomly. Because the growth is biased towards self-avoiding configurations, the probability of generating each configuration is not uniform as in the case of simple sampling. As a result, each configuration is associated with an appropriate weight to account for the bias.

It is convenient to view the Rosenbluth algorithm as approximate counting, wherein the weights of configurations represent the number of possible self-avoiding configurations. For instance, if we want to exactly enumerate the total number of configurations,

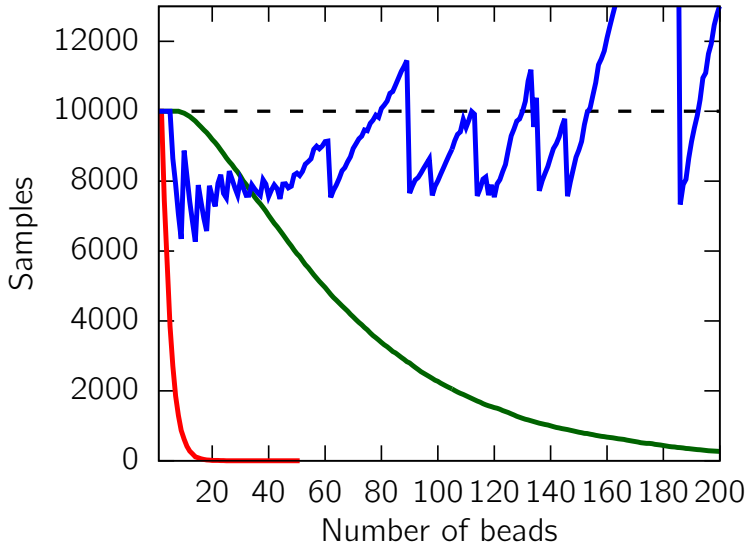


Figure 3.2. The number of samples in different chain growth methods against the number of beads. The red curve is from simple sampling, the green from Rosenbluth sampling, and the blue curve is from PERM. The black dashed line is the number of tours in each of the simulations.

a naïve way would be to choose all possible continuations and count all configurations with equal weight. However, if we choose fewer configurations, we need to change the weight of these configurations appropriately, accounting for the configurations we did not consider. Therefore, if we pick one step out of a possible growth steps, we need to replace the a possible configurations with equal weight by one configuration with a-fold weight. Thus, the weight of a configuration is an approximate count of the total number of available configurations. Accordingly, the average of the sum of the weights over many possible configurations should give us an estimate of the partition function, with the reference state of an ideal chain that permits overlap of beads.

The approximate count or the weight of a configuration after j growth steps ($j+1$ beads) is the product of the atmospheres at each step of chain growth,

$$w_n = \prod_{j=1}^{j=n} a_j, \tag{3.2}$$

where a_j is the atmosphere of the configuration after j growth steps. Since the probability of picking a growth step out of a_j is $1/a_j$, the probability with which a n-step

walk is generated is

$$p_n = \prod_{j=1}^{j=n} \frac{1}{a_j}, \tag{3.3}$$

ensuring that the condition $p_n w_n = 1$ is satisfied. The average of a thermodynamic quantity such as the end-to-end distance over N_t tours is calculated as

$$\langle R_n \rangle = \frac{\sum_{t=1}^{N_\mathrm{t}} w_n^t R_n^t}{\sum_{t=1}^{N_\mathrm{t}} w_n^t}, \tag{3.4}$$

where w_n^t is the weight of a n-step walk in the tth tour. Likewise, the partition function, the total number of possible configurations of a n-step walk, Z_n, can be estimated as

$$Z_n = \frac{\sum_{t=1}^{N_\mathrm{t}} w_n^t}{N_\mathrm{t}}. \tag{3.5}$$

Figure 3.2 shows that Rosenbluth sampling achieves significant improvement over simple sampling in the molecular weights that it is able to sample. However, the introduction of the Rosenbluth bias only advances the realizable molecular weights by about 100 beads as the molecules can still get into traps, as shown in Figure 3.1(c), leading to exponential attrition with increase in chain length.

3.1.3 Pruned-enriched Rosenbluth method

As Prellberg and Krawczyk (2004) put it, there are two key problems associated with Rosenbluth sampling. First, the fluctuations in the weights can be very large and therefore the averages can be dominated by a handful of samples with large weight. Second, the walks generated often get trapped in configurations with zero atmosphere, which can lead to exponential attrition with increase in chain length.

Grassberger (1997) introduced pruning and enrichment steps to counter these obstacles in Rosenbluth sampling. The basic idea of his strategy is to (i) prune or stop the chain growth when the weight of the chain is too low and (ii) enrich or introduce branches to the chain growth when the weight of the chain is too high. Grassberger (1997) advocated the use of certain parameters to decide whether to prune or enrich the chain during chain growth. Several years following the publication of Grassberger (1997),

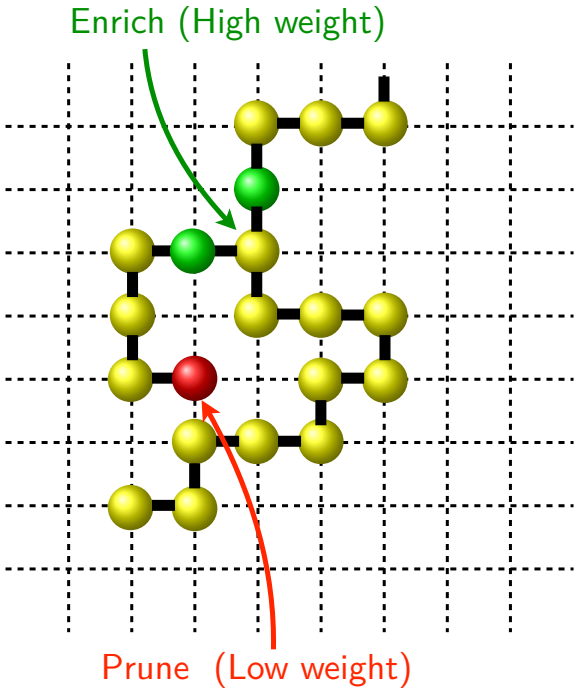


Figure 3.3. Illustration of PERM of a 2D SAW. Unlike Rosenbluth sampling, PERM enriches/branches configurations with high weight and prunes configurations with low weight.

Prellberg and Krawczyk (2004) modified the method to come up with a parameter-free version of PERM, which is the variant of PERM that we use in our work.

Consider a n-step walk that has been generated by Rosenbluth sampling with a weight w_n. Let W_n be the average weight of all such SAWs. W_n is essentially the partition function of n-step walks. In the ideal scenario, we want w_n to be close to W_n with minimal fluctuations in the weights. However, if the ratio $R_n = w_n/W_n < 1$, the current weight, w_n, is too small, or if $R_n = w_n/W_n > 1$, w_n is too large. Accordingly, Prellberg (2013) suggests the following modification to the Rosenbluth algorithm:

1. If $R_n = w_n/W_n < 1$, stop growing (prune) with probability $1 - R_n$ or continue growing with probability R_n.

2. If $R_n = w_n/W_n > 1$, make $\lfloor R_n \rfloor + 1$ copies with probability $p_n = R_n - \lfloor R_n \rfloor$ and $\lfloor R_n \rfloor$ copies with probability $1 - p_n$. Thereafter, continue growing each copy with the weight of each copy set to W_n.

With the introduction of pruning and enrichment steps, we expect to obtain a uniform distribution in the number of samples with varying chain size. However, as shown in Figure 3.2, it's not quite the case and the number of samples fluctuates around the targeted value of 10000. This is due to the fact we do not typically know the target weight, W_n, before we start the simulation. Indeed, the simulation corresponding to Figure 3.2 was a simulation in which the target weight was acquired on the go. In Section 3.2.2, we will discuss how this 'chicken and egg' problem of not knowing the target weights beforehand and then acquiring the target weight in a PERM simulation can be solved to run PERM simulations that yield reliable estimates of weights and other observables. In spite of these difficulties, it is evident from Figure 3.2 that PERM enables us to obtain a sizeable number of samples for the full range of molecular weights by overcoming the attrition problem that plagued Rosenbluth sampling.

In PERM, we can have more than one sample for a given chain size from the same tour, because enrichment steps introduce branches in chain growth resulting in a tree-like structure. Accordingly, the average end-to-end distance is estimated as the weighted sum over all samples, N_s, as

$$\langle R_n \rangle = \frac{\sum_{s=1}^{N_\mathrm{s}} w_n^s R_n^s}{\sum_{s=1}^{N_\mathrm{s}} w_n^s}, \tag{3.6}$$

where w_n^s and R_n^s represent the weight and end-to-end distance of the sth sample of a n-step walk respectively. The partition function estimate can be obtained by

$$Z_n = \frac{\sum_{s=1}^{N_s} w_n^s}{N_t}. \tag{3.7}$$

3.2 Off-lattice PERM

Thus far, the focus of our discussion was the simulation of a 2D SAW on a lattice using PERM. Although PERM simulations with 3D and 2D lattice models have been successfully applied to study properties of polymers in confinement (Hsu and Grassberger, 2003; Hsu and Grassberger, 2004; Hsu and Binder, 2013), lattice simulations of semiflexible polymers introduce artifacts because the bending angles are restricted to multiples of 90° (Hsu and Binder, 2013). Given the absence of continuous bonding angles, lattice models do not possess the ability to reproduce sub-persistence length features such as deflection segments (Hsu and Binder, 2013), which are critical to our work. In what follows, we explain our methodology of off-lattice PERM that we employ in simulations of semiflexible polymers in confinement.

3.2.1 PERM simulation of freely jointed chains

In the interest of simplicity, we first begin our elucidation of off-lattice PERM with an application of PERM to an off-lattice 3D SAW . The model we first use here is that of a freely jointed chain, i.e. there is no stiffness associated with the bonds of the chain. In this case, the potential energy of interaction between two beads with indices i and j with position vectors $\mathbf{r}_i$ and $\mathbf{r}_j$ is given by

$$\beta U_{\mathrm{EV}}(\mathbf{r}_i, \mathbf{r}_j) = \begin{cases} \infty & \text{if } |\mathbf{r}_i - \mathbf{r}_j| \leq w, \\ 0 & \text{otherwise,} \end{cases} \tag{3.8}$$

where $\beta = 1/k_B T$ and w is the width of the chain. The self-avoidance part can be accounted for in a similar way as in Section 3.1.3. However, because of the off-lattice nature of the method, the possible growth steps are not limited to a fixed number of lattice sites as in the case of PERM on a 2D lattice. The possible growth steps for a

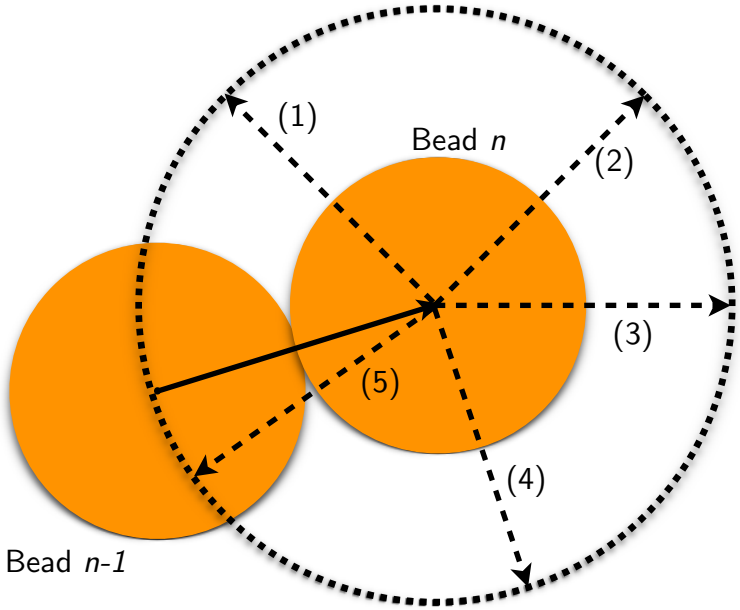


Figure 3.4. Off-lattice PERM of a freely jointed chain. The center of the $(n+1)$th can be any point on the surface of a sphere whose center coincides with the center of the nth bead and has a radius w. In PERM, we generate m trial sites on the surface of this sphere. In this case $m = 5$. It is likely that sites such as (1) and (5) are not chosen as such sites would lead to overlap of the $(n+1)$th with the $(n-1)$th bead.

n-bead chain can be chosen from any of the (infinite) points on the surface a sphere with the center being the center of the nth bead, as shown in Figure 3.4.

To account for the off-lattice nature, we discretize the surface of this sphere by picking a maximum of m possible trial positions on the sphere for the next growth step. For PERM on a 2D lattice, the value of m is fixed ($m = 4$) as there are 4 possible neighboring lattice positions for a growth step. Assuming all m trial positions are free, we need to pick one of these $k_n = 1 \ldots m$ trial positions as the next growth step. However, not all m positions are free in general. If we choose each free neighbor, $k_n = 1, \ldots, m_n$, at the nth step with a probability, p_k, then we need to choose the corresponding weight for the nth step, $\omega_n^{(k_n)}$, to counter this bias. The weight of the n-step long chain is the product of the weights associated with each step of growth given by

$$w_n = \prod_{j=1}^{n} \omega_j^{(k_j)}. \tag{3.9}$$

In Section 3.1.3, we mentioned that the condition $p_n w_n = 1$ must be satisfied in order

to assign the correct weights depending on how the growth steps are picked. A more general condition is expressed as (Grassberger, 1997)

$$p_k \omega_n^{(k_n)} = c_n \exp(-U_{k_n}/k_\mathrm{B}T), \tag{3.10}$$

with a prefactor c_n independent of k. Here, $\omega_n^{(k_n)}$ is the weight of the kth trial step at the nth step of chain growth and U_{k_n} is the increment in the potential energy associated with it. In our work, we also fix $c_n = 1$ to be independent of n, which essentially expresses the condition for the grand canonical ensemble (Grassberger, 1997).

For a 3D SAW, the hard-core potential means that the Boltzmann factor for all free sites is $\exp(-U_k/k_\mathrm{B}T) = 1$. Thus, if we choose each free neighbor, $k = 1, \ldots, m_n$, with the same probability $p_k = 1/m_n$, Eq. 3.10 reduces to the condition that the weight is equal to the total number of free neighboring sites or the atmosphere. Accordingly, $\omega_n^{(k_n)} = m_n = a_n$. With the aforementioned steps, the PERM chain growth is performed with exactly the same pruning and enrichment strategies as outlined in Section 3.1.3.

In our implementation of our algorithm, we set the number of trial neighboring sites to $m = 5$. Note that the complexity of the algorithm is linear in m. Choosing a very low value of m such as 1 or 2 can often lead to situations where the atmosphere is 0, resulting in forced pruning of the chain. Such pruning further leads to an equivalent amount of enrichment to compensate for the pruning as the PERM algorithm tries to maintain a constant number of samples as a function of molecular weight. Choosing a high value of m enables better sampling, but at the cost of computational time. Choosing an optimal value of m is somewhat of an art as it depends on the relative strengths of excluded volume and stiffness of the chain. Nonetheless, the optimal value of m can be found empirically. Because of the narrowness of the WLC distribution for moderate values of κ, a value of m between 5 and 10 achieves satisfactory sampling.

Furthermore, to calculate the atmosphere of the chain at the nth step of chain growth, we need to check if each of the m trial steps result in an overlap of the potential $(n+1)$th bead with the rest of the $1, \ldots, n$ beads. Because the self-avoidance problem is non-Markovian, this step of checking for the collisions of the chain itself needs to be repeated at each step of the chain growth. Each of these steps involves a loop over n beads to check for collisions. As such, the complexity of the PERM algorithm is $\mathcal{O}(n^2)$, which makes it infeasible to perform simulations beyond 10000 beads. To address this challenge, we employ a neighbor list that is loosely based on the Verlet neighbor list (Allen and Tildesley, 1987), as described below.

Since PERM is a chain growth algorithm, the neighbor list needs to be constructed such that it can be updated during chain growth. To construct a neighbor list, we use two length scales — the cutoff radius, r_c, and the skin depth, d_s, as shown in Figure 3.5. In all our simulations, we fix the values of these parameters to be $r_c = 22w$ and $d_s = 2w$. With the start of chain growth, we keep track of all beads within a radius of r_c from the center of the first bead. Note that this list also includes the first bead. As chain growth continues and we add more beads to the chain, we update the neighbor list with the indices of the newly added beads. We continue updating the neighbor list with newly added beads until for some value $n_1 > 1$, the position of the n_1th bead goes outside the sphere of radius $r_c - d_s$ that was constructed around the first bead. At this point, we construct a new neighbor list that keeps track of all the beads within a sphere of radius r_c centred around the n_1th bead. This neighbor list is further updated as beads are added to the chain. Again, when a certain n_2th bead goes outside the sphere of radius $r_c - d_s$ centered around n_1th bead, a new neighbor list is constructed by considering another sphere of radius r_c centered around the n_2th bead. This process of updating or freshly constructing new neighbor lists is repeated until the tour ends. By constructing such a neighbor list, we only check for bead overlap with the members in the neighbor list during chain growth. This makes of our algorithm much more efficient compared to the PERM code without the neighbor list. Indeed, the computational time of our code scales roughly linearly with chain size with the incorporation of the neighbor list. Finally, note that a skin depth of $d_s \geq w$ is necessary to ensure that there are no artifacts in the simulation due to overlaps with beads that are not in the neighbor list at the boundary of the sphere.

3.2.2 PERM simulation using the discrete wormlike chain model

The DWLC model differs from the freely jointed model in that there is an additional bending potential between contiguous triplets of beads with indices i, $i+1$ and $i+2$ in the chain given by

$$\beta U_{\text{bend}} = \kappa(1 - \cos\theta_i), \tag{3.11}$$

where κ is the bending constant, $\cos\theta_i = (\mathbf{r}_{i+2} - \mathbf{r}_{i+1}) \cdot (\mathbf{r}_{i+1} - \mathbf{r}_i) / \ell^2$. Here ℓ is the bond length of the chain.

To incorporate the effect of the bending potential at the nth growth step in PERM for a configuration such as the one shown in Figure 3.6, the simplest approach would

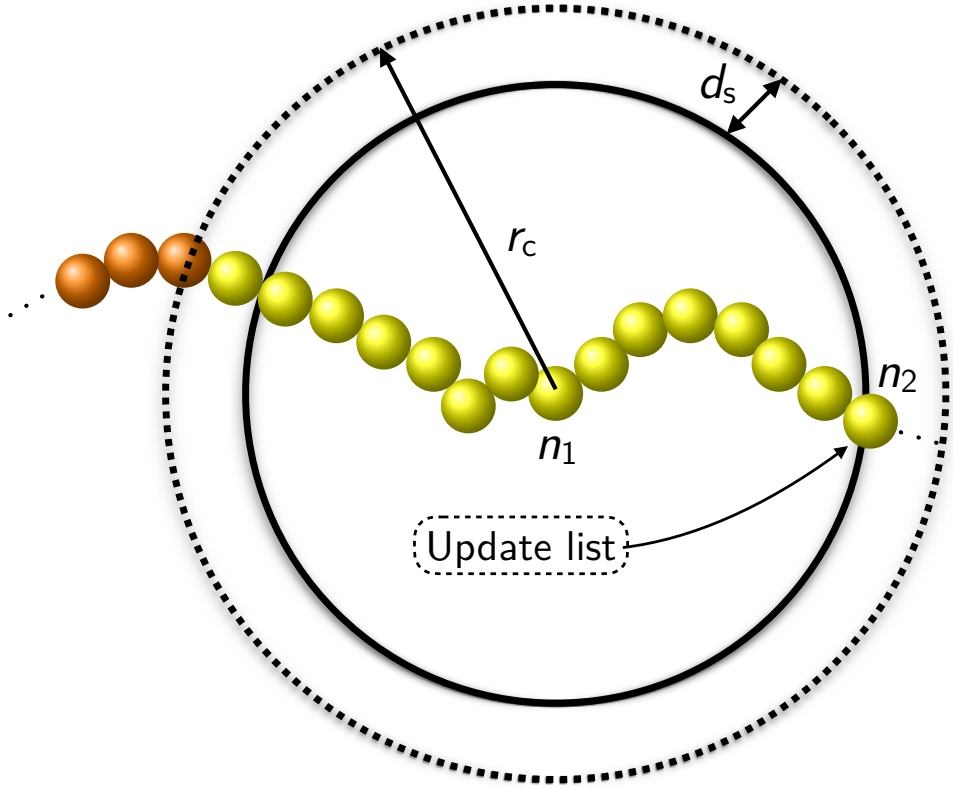
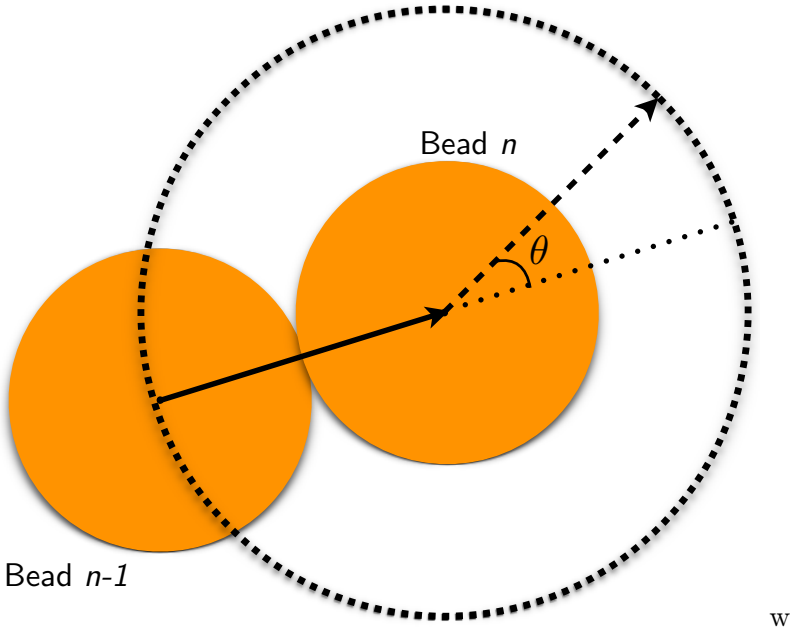


Figure 3.5. Illustration of the neighbor list. The neighbor list here consists of all beads within a sphere of radius r_c with its center being the center of the n_1th bead, as indicated by the dashed circle. The yellow beads are part of the neighbor list and the orange beads are not part of the neighbor list. As the n_2th bead crosses the sphere with radius $r_c - d_s$ (circle with solid black line), a new neighbor list is constructed with a sphere centred around the n_2th bead.



w

Figure 3.6. Off-lattice PERM of a wormlike chain. The probability of a trail step for the $(n+1)$th growth step depends on its orientation. The orientation of the nth bond vector is likely to be highly correlated to the $(n-1)$th bond vector.

be to (i) pick m trial sites on the surface of a sphere randomly as is done for freely jointed chains (Figure 3.4); (ii) assuming there are $k = 1, \ldots, m_n$ free sites out of the m sites, we can pick one of these sites with the same probability $p_k = 1/m_n$ and (iii) assign an appropriate weight, $\omega_n^{(k_n)} = \exp(-U_{k,\text{bend}}/k_\text{B}T)$, to ensure consistency with Eq. 3.10 before continuing the chain growth to $n+1$ beads. This approach sounds perfectly reasonable until we look at the distribution of the weights in this approach. For instance, for a typical value of $\kappa = 10$, the weight arising from the bending potential can be as high as $\exp(0) = 1$ for $\theta = 0°$ and as low as $\exp(-10) \approx 4.5 \times 10^{-5}$ for $\theta = 90°$, a remarkable 5 orders of magnitude difference. Thus in this approach, the distribution of weights can range over several orders of magnitude. In other words, the distribution of favorable weights is extremely narrow and is centred around the direction of the previous bond vector, as shown in Figure 3.7. Such unfavorable distribution of weights is also the reason for the failure of Rosenbluth sampling, which we know leads to exponential attrition.

Another approach would be to replace step (ii) in the latter approach by picking the m_n free sites with a probability that is proportional to the Boltzmann weight, i.e.

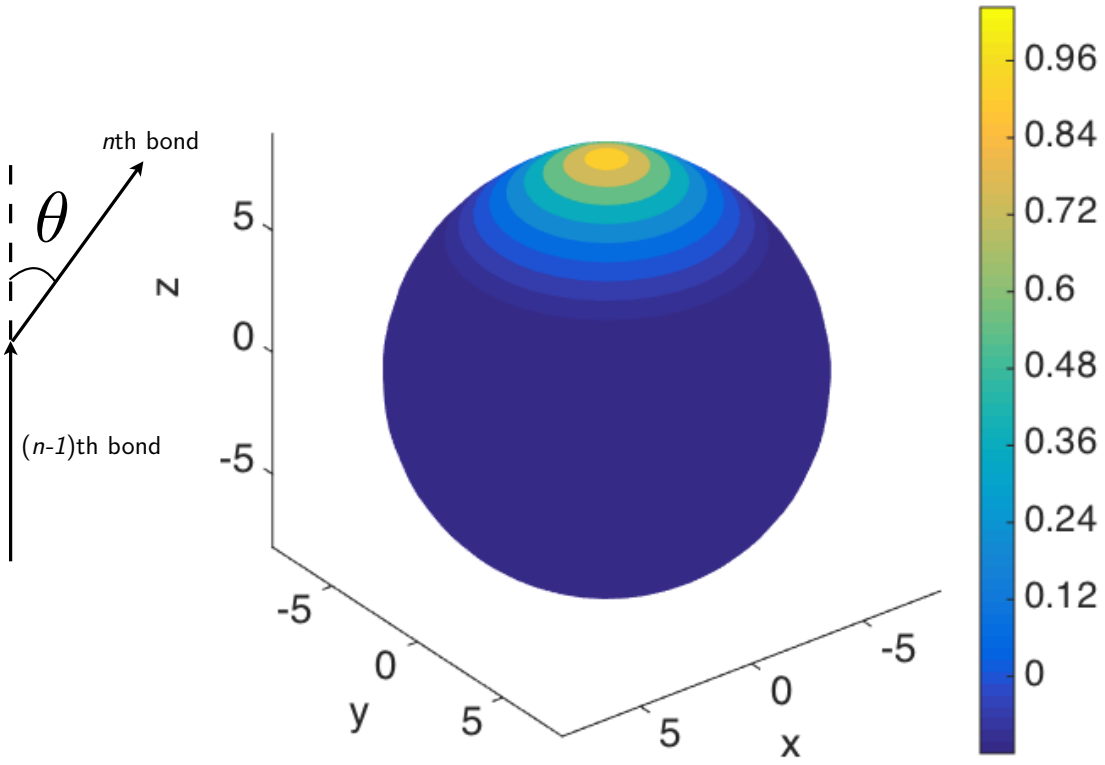


Figure 3.7. The distribution of $\exp(-\beta U_{\text{bend}})$ for $\kappa = 10$. Here the $(n-1)$th bond is oriented along the z axis on the $x = 0$ plane from $z = -8$ to $z = 0$. The bond length in units of length shown in the figure is 8. The end point of the nth bond vector ('important' part of the sphere) is likely to be in a very narrow part of the sphere as shown by the contour of the Boltzmann weight corresponding to the bending potential.

$p_k \propto \exp(-U_{k,\text{bend}}/k_\text{B}T)$ or $p_k = \exp(-U_{k,\text{bend}}/k_\text{B}T)/\sum_{k'=1}^{m_n}\exp(-U_{k',\text{bend}}/k_\text{B}T)$. Because we choose the trial steps with a probability proportional to its Boltzmann weight, this is nothing but importance sampling, which is characteristic of Metropolis Monte Carlo simulations. As a result, the weights $\omega_n^{(k_n)} = 1$ would be independent of k in accordance with Eq. 3.10. Unlike the approach mentioned in the previous paragraph, this method would not result in a large distribution of weights, but it does make it challenging to sample the 'important' parts of the phase space if the distribution of the Boltzmann weight is narrow, as is typically the case for chains with $\kappa \gtrsim 5$. To perform importance sampling efficiently at each growth step, one can still choose a high enough value of m so that one has enough trial steps to be able to numerically sample the distribution of Boltzmann weights. However, the higher the value of κ, the narrower this distribution gets, necessitating a value of m that is unreasonable for simulation of long chains.

An alternative approach would be to generate m random orientations (trial steps) that inherently satisfy the condition $p_k = \exp(-U_{k,\text{bend}}/k_\text{B}T)/\sum_{k'=1}^{m}\exp(-U_{k',\text{bend}}/k_\text{B}T)$. Fortunately, this can be achieved because the Boltzmann factor for the bending potential, $\exp\left[-\kappa\left(1-\cos\theta\right)\right]$, is analytically invertible. Therefore, the inverted form can be used to generate random numbers that lead to a distribution of orientations whose probability is proportional to their respective Boltzmann weights. Thereafter, identifying m_n out of the m free sites yields us the atmosphere for the growth step. One of these m_n trial steps can be chosen randomly with equal probability and assigned a weight of $\omega_n^{(k_n)} = m_n$ maintaining consistency with Eq. 3.10. This method circumvents the problems associated with the previous two approaches discussed here, thereby allowing us to grow very long chains with contour length of $\mathcal{O}(10^5)$ beads with a moderate value of $m = 5$.

Finally, using the pruning and enrichment strategies outlined in Section 3.1.3, we perform PERM simulations and calculate thermodynamic averages based on Eqs. 3.6 and 3.7. Our PERM algorithm is summarized in the form of a pseudo code along the lines of the method of Prellberg (2013) in Algorithm 1. Note that the algorithm depends on a target weight (or partition function), W_n, as a function of chain length in order to decide whether to prune or enrich or to continue the chain growth. Because we do not know the target partition function beforehand, we perform three kinds of runs.

1. To begin with, we perform "blind-PERM" simulations, where we acquire an estimate of the partition function (or the target weight), for a short chain of 50 or 100 beads. Here, the estimate of the partition function is acquired during run-time

Algorithm 1 Implementation of off-lattice PERM

$s. \leftarrow 0$; $w. \leftarrow 0$; $Tours \leftarrow 0$; $n \leftarrow 1$; $w_1 \leftarrow 1$; $a \leftarrow 0$; $Copy_1 \leftarrow 1$; $s_1 \leftarrow s_1 + 1$
$r_1 \leftarrow$ Set first bead position; Start neighbor list; Sum.$Z_1 \leftarrow$ Sum.$Z_1 + w_1$
while $Tours < MaxTours$ **do**
  if $n = MaxBeads$ or $a = 0$ **then** $Copy_n \leftarrow 0$
  else $Ratio \leftarrow w_n/W_n$; $p \leftarrow Ratio$ mod 1; Generate random number $\rho \in [0;1]$
    if $\rho < p$ **then** $Copy_n \leftarrow \lfloor Ratio \rfloor + 1$
    else $Copy_n \leftarrow \lfloor Ratio \rfloor$
    end if
    if (Blind) **then** $W_n =$ Sum.$Z_n/(Tours + 1)$; $w_n = W_n$
    else $w_n = W_n$
    end if
  end if
  if $Copy_n = 0$ **then**
    while $n > 1$ and $Copy_n = 0$ **do** $n \leftarrow n - 1$
    end while
    if $n > 1$ **then** Construct new neighbor list
    end if
  end if
  if $n = 1$ and $Copy_n = 0$ **then**
    $Tours \leftarrow Tours + 1$; $a \leftarrow 0$; $Copy_1 \leftarrow 1$; $s_1 \leftarrow s_1 + 1$; Sum.$Z_1 \leftarrow$ Sum.$Z_1 + w_1$
  else
    Update/Construct new neighbor list
    if a > 0 **then**
      $Copy_n \leftarrow Copy_n - 1$
      Draw 5 sites on the surface of a sphere from the DWLC distribution
      Pick one of the non-overlapping sites randomly with uniform probability
      $n \leftarrow n+1$; $w_n \leftarrow w_n \times a$; $r_n \leftarrow$ Picked site; $s_n \leftarrow s_n+1$; Sum.$Z_n \leftarrow$ Sum.$Z_n + w_n$
      if (Production run) **then** # Collect statistics for thermodynamic averaging
        Sum.$R_n \leftarrow$ Sum.$R_n + w_n \times R_n$
      end if
    end if
  end if
end while
Sum.$Z_n \leftarrow$ Sum.$Z_n/Tours$
if (Production run) **then** Sum.$R_n \leftarrow$ Sum.R_n/Sum.Z_n
end if

as the average of the weights over all the completed tours. Although initially the estimate of the partition function, W_n, is dominated by only a small number of tours and can be far away from the actual value, it does converge if the PERM simulation is run for thousands of tours. In other words, the PERM algorithm is "self-tuning" (Prellberg and Krawczyk, 2004).

2. We then run "non-blind" simulations by increasing the molecular weight of the simulated chain step by step in successive runs. We term these simulations "non-blind" as we use the estimates of the partition function acquired in a previous run, $W_{1,\ldots,n_1}$, and extrapolate them to a higher molecular weight n_2 to obtain an estimate of the partition function for molecular weights between n_1 to n_2, $W_{n_1+1,\ldots,n_2}$. Typically, we perform our blind simulations for a contour length of 50 beads and get an estimate of the partition function for chains up to a length of 50 beads. Thereafter, we perform non-blind simulations for 100 beads, 200 beads, 400 beads and so forth in succession by linearly extrapolating the logarithm of the partition function, $\ln Z_n$, linearly with molecular weight. Here, we use the fact the logarithm of the partition function scales linearly in the molecular weight in the long-chain limit, i.e. $\ln Z_n \sim n$ as $n \to \infty$ (Li et al., 1995; Madras and Slade, 2013). Even if the extrapolated values of the weights are not accurate possibly due to the non-linear nature of $\ln Z_n$ for simulations at low molecular weights, the successive non-blind run tunes the weights towards the correct estimate of the partition function. For instance, after the completion of a non-blind run for a 100-bead simulation that uses extrapolation of the partition function from a 50-bead simulation, the resulting estimates of the partition function, $\ln Z_{51,\ldots,100}$, are accurate even if the originally extrapolated values deviate from the actual values.

3. Finally when we reach the desired molecular weight (typically $\mathcal{O}\left(10^4\right)$) after gradually increasing the molecular weight through a series of non-blind simulations, we perform "production" runs, where we acquire averages of observables such as the end-to-end distance using Eq. 3.6.

With the aforementioned steps and the algorithm outlined in Algorithm 1, we wrote a parallelized Fortran code based on the message passing paradigm using message passing inteface (MPI) directives. We divide the total number of tours in our simulations into several independent sets, so that we can estimate the statistical error in the simulations from independent sets of data. Typically a simulation of 1 million tours is divided into

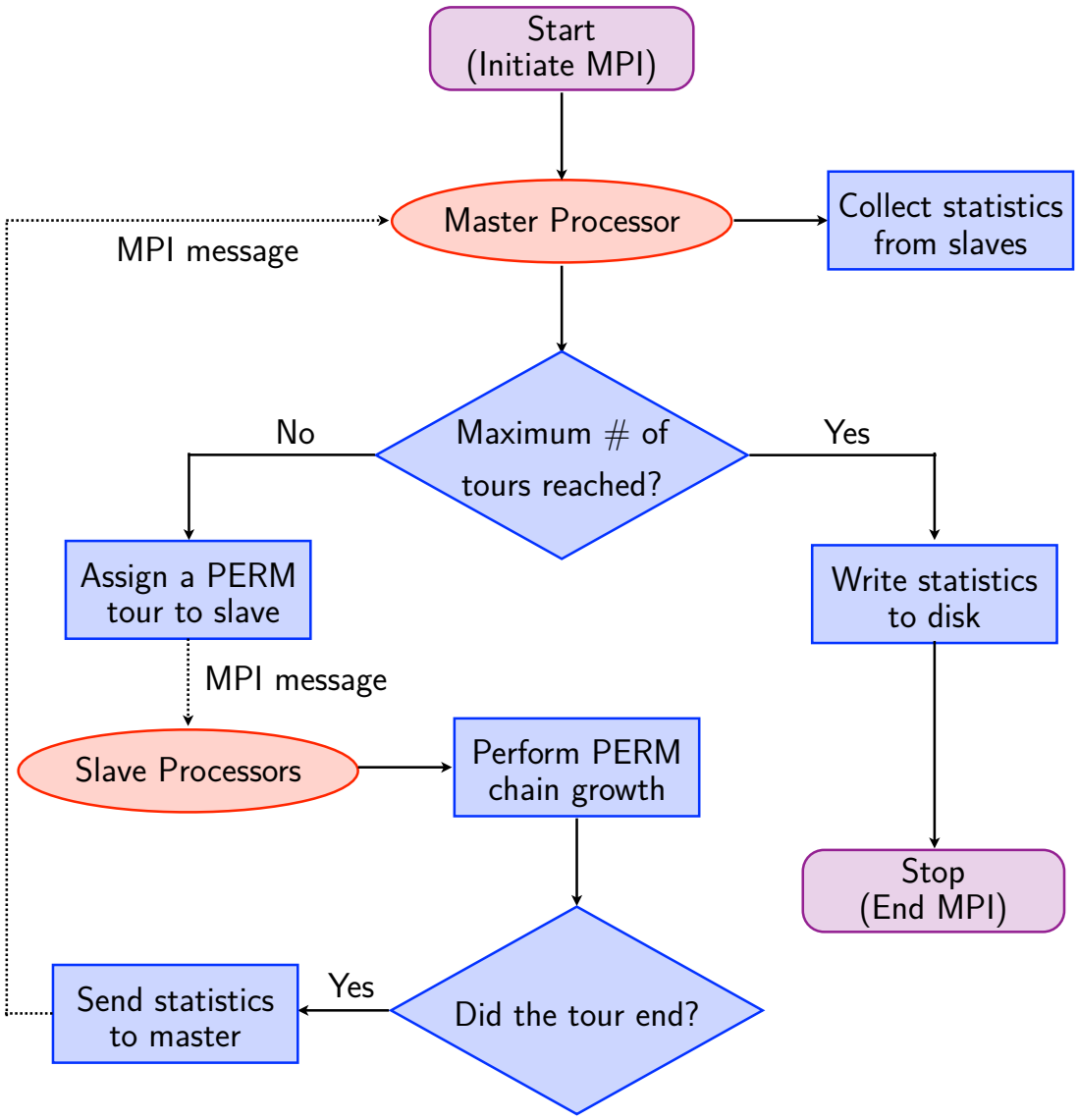


Figure 3.8. Flowchart of our PERM algorithm for each set of our data. The dotted lines denote MPI messages between processors. Each set of our data consists of a master processor that assigns jobs to slave processors, in addition to keeping track of all the statistics. Each of the slave processors perform the PERM simulation until the maximum number of tours is reached.

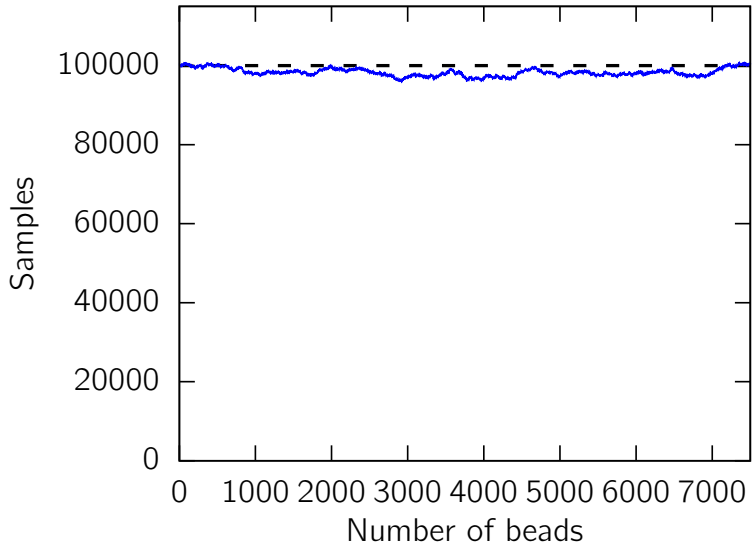


Figure 3.9. Number of samples (blue line) from a PERM simulation with 5 sets of 20,000 tours each for a maximum contour length of 7500 beads. The black dashed line shows the total number of tours in the simulation (100,000).

5 sets of 200,000 tours each. We run the stipulated number of tours in each set with a master-slave algorithm as illustrated in the form of a flowchart in Figure 3.8. The master processor assigns jobs to the slave processors and collects statistics such as the weights and the values of observables from each tour. Each job is a single tour of the PERM simulation. The master processor assigns a tour to all free processors and waits until the slave processors get back to the master processor with a finished tour and all associated statistics. This back and forth message passing continues until the maximum number of tours is reached, upon which the master stops assigning jobs to the slaves. Once all the slaves complete their respective tours, the master writes all the statistics to disk. With these modifications to the basic PERM algorithm described in Section 3.1.3, we obtain an approximately uniform distribution in the number of samples, as shown in Figure 3.9.

3.2.3 Comparison with experiments

As PERM can natively generate data as a function of molecular weight, it enables us to compare our simulation data from a single simulation with experimental data from different sources, corresponding to different molecular weights. In Figure 3.10, we plot

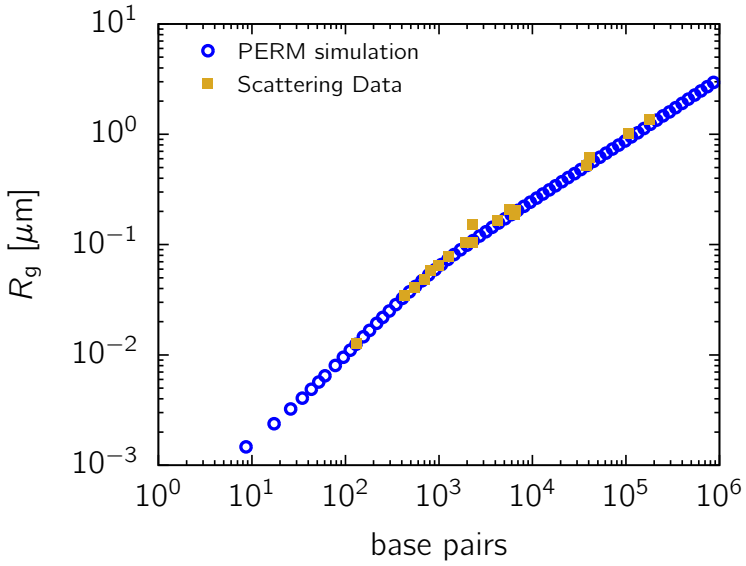


Figure 3.10. PERM simulation data against experimental data for DNA from various light and neutron scattering studies (Harpst, 1980; Godfrey, 1976; Godfrey and Eisenberg, 1976; Jolly and Eisenberg, 1976; Kam et al., 1981; Schmid et al., 1971; Gray Jr. and Hearst, 1968; Allison et al., 1990; Sorlie and Pecora, 1990; Seils and Dorfmüller, 1991; Voordouw et al., 1978; Lederer et al., 1986). Reproduced from Tree et al. (2013b).

the radius of gyration against the molecular weight of the chain from our simulations for parameters of DNA in a high ionic strength buffer — an effective width, $w = 4.6\,\mathrm{nm}$, and a Kuhn length, $b = 106\,\mathrm{nm}$. We observe very good agreement with experimental data at various ionic strengths ($\geq 100\,\mathrm{mM}$) lending us confidence in the ability of the DWLC model and the accompanying PERM simulation technique to study the physics of DNA in diverse environments.

Chapter 4

Effect of contour length on the static properties of semiflexible polymers confined in nanochannels

4.1 Introduction

* A long polymer confined inside a narrow channel extends along the channel axis due to constraints imposed by the walls (Reisner et al., 2012). In weak confinement, where the channel size is somewhat smaller than the polymer's radius of gyration, de Gennes suggested that the resulting chain configuration could be represented by a linearly ordered series of blobs, where the subchain inside the blob obeys self-avoiding random walk statistics (de Gennes, 1979; Daoud and de Gennes, 1977). Thus, the de Gennes regime arises solely from excluded volume interactions. In strong confinement, where the channel size is small compared to the persistence length of the chain, Odijk proposed that the chain configuration is instead described by a series of deflection segments that obey ideal wormlike chain statistics (Odijk, 1983). In contrast to the de Gennes regime, the Odijk regime produces chain extension even in the absence of excluded volume interactions due to chain stiffness. When viewed in this light, it becomes clear that the Odijk and de Gennes regimes are limiting cases for polymer extension corresponding to the dominance of either chain stiffness or excluded volume.

*This chapter is based on A. Muralidhar, D. R. Tree, Y. Wang, and K. D. Dorfman, "Interplay between chain stiffness and excluded volume of semiflexible polymers confined in nanochannels", J. Chem. Phys. **140**, 084905 (2014)

Unfortunately, neither the Odijk theory nor the de Gennes theory appears to describe the existing experimental data for DNA extension in a nanochannel (Reisner et al., 2012; Levy and Craighead, 2010; Persson and Tegenfeldt, 2010; Dorfman et al., 2013; Reisner et al., 2005; Reisner et al., 2007; Odijk, 2006; Odijk, 2008; Zhang et al., 2008; Cifra et al., 2009; Persson et al., 2009; Cifra et al., 2010; Cifra and Bleha, 2012; Wang et al., 2011; Su et al., 2011; Tree et al., 2013a; Dai et al., 2012; Werner et al., 2012; Benková and Cifra, 2012). The apparent failure of the classic de Gennes and Odijk theories, when applied to channel-confined DNA, lies in the relative length scales characterizing the DNA experiments; the persistence length of double-stranded DNA ($l_\mathrm{p} \approx 50\,\mathrm{nm}$) (Bustamante et al., 1994) is of the same order of magnitude as the channel sizes used in experiments. As a result, the extension of DNA in nanochannel experiments results from a competition between excluded volume effects and chain stiffness for all experimentally relevant channel sizes and molecular weights. For this reason, we recently suggested that it is useful to think of the transition from the Odijk regime to the de Gennes regime as a rod-to-coil transition for the subchain used to renormalize the problem into a series of deflection segments (the rods) or excluded volume de Gennes blobs (the coils) (Tree et al., 2013a). While the details of the physics in the transition region remain a topic of debate, the existence of at least one additional regime between the Odijk and de Gennes regimes now seems very likely (Odijk, 2008; Cifra et al., 2009; Wang et al., 2011; Tree et al., 2013a; Brochard-Wyart et al., 2005; Dai and Doyle, 2013).

In the present contribution, we explore this competition between excluded volume and chain stiffness in considerably more detail than our recent letter (Tree et al., 2013a). Our exposition is two-fold. In the first part, we use Pruned-Enriched Rosenbluth method (PERM) simulations of a touching bead model to directly test the Odijk and de Gennes theories within the limits of their underlying assumptions: for the Odijk regime, we consider an ideal wormlike chain; for the de Gennes regime, we consider a self-avoiding, freely jointed chain. So long as the assumptions of the theories are satisfied, we show that their predictions for the chain extension and confinement free energy agree with the simulation data for sufficiently long chains. We also show that our PERM calculations produce the undisputed results for the confinement free energy and extension of an ideal, freely jointed chain. In the second part, we compare the predictions of the chain extension for these idealized cases to the extension of a self-avoiding wormlike chain in a nanochannel as a function of molecular weight and channel size. Our simulations of real wormlike chains clearly produce the Odijk and de Gennes regimes as

the limiting cases. Analogous to DNA nanochannel experiments, much of the phase space of our simulations is occupied by chains that satisfy neither theory. We also show that the projection of the end-to-end distance of the confined chain along the channel axis approaches its mean span in the long-chain limit, owing to the difficulty to form hairpins (when stiffness dominates) or the weak penetration of the end of the chain into the linearly ordered, self-avoiding blobs (when excluded volume dominates). We thus propose that the "long-chain" limit corresponds to molecular weights where these two size metrics are approximxately equal; data obtained outside of this limit is subject to artifacts due to finite chain lengths. Remarkably, while recent simulations indicate that λ-DNA (48,500 base pairs, bp) is well below the long-chain limit in free solution (~ 1 Mbp) (Tree et al., 2013b), the amplification of excluded volume effects in confinement (Reisner et al., 2012) lead to λ-DNA being a sufficiently long chain for most nanochannel experiments measuring the extension.

4.2 Model and simulation method

4.2.1 Discrete wormlike chain model

Consider a wormlike chain of contour length L, persistence length l_p and width w confined in a square channel of size D as shown in Figure 4.1(a). The length of the channel is much greater than the contour length of the chain and hence there are no end-effects due to finite length of the channel. For simplicity, we assume that the depletion length between the wall and polymer due to excluded volume is $w/2$, which is not necessarily true in general. As a result, the effective width of the channel accessible to the center of the chain is $D_\mathrm{eff} = D - w$.

To carry out the PERM simulations, we use an off-lattice, discrete wormlike chain (DWLC) model (Tree et al., 2013b; Wang and Gao, 2005). PERM simulations with lattice models have also been successfully used to answer a wide variety of questions surrounding polymer confinement (Hsu and Grassberger, 2003; Hsu and Grassberger, 2004; Hsu and Grassberger, 2011; Hsu and Binder, 2013). However, these models are known to introduce lattice artifacts, particularly where bending stiffness plays a critical role in determining the properties of the polymer chain (Hsu and Binder, 2013). Because our work demands an accurate depiction of chain stiffness in the model, we sacrifice the computational efficiency of a lattice model for the precision of an off-lattice model.

Our model is a touching-bead/necklace model of $N + 1$ beads connected by rigid

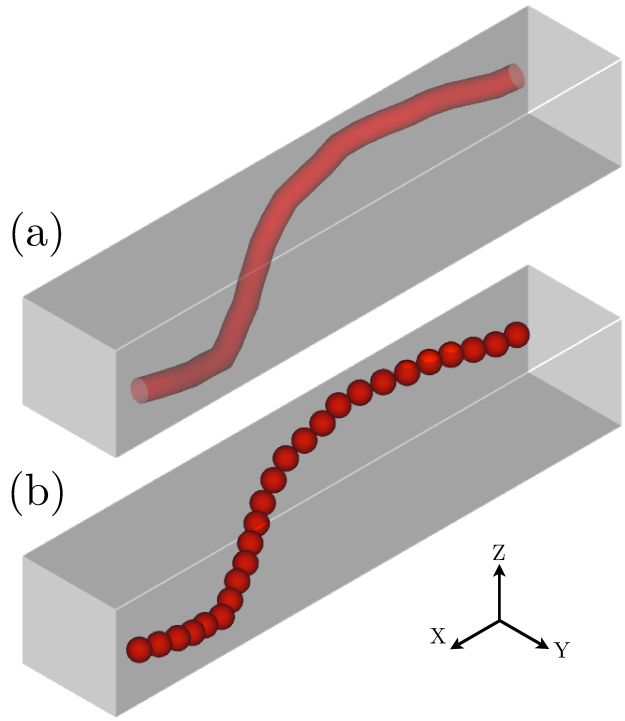


Figure 4.1. (a) A WLC with contour length L, persistence length l_{p}, and width w confined in a square channel of size D. (b) Representation of the same chain in our DWLC model. After discretization, the bond length a is set to the value of the width w.

bonds of length a as shown in Figure 4.1(b). Stiffness is imposed by a bending potential given by

$$\beta U_{\text{bend}} = \kappa \sum_{i=1}^{N-1} (1 - \cos\theta_i) \tag{4.1}$$

between consecutive bonds. The parameter θ_i is the angle between bonds i and $i+1$, determined by $\cos\theta_i = |(\mathbf{r}_{i+2} - \mathbf{r}_{i+1}) \cdot (\mathbf{r}_{i+1} - \mathbf{r}_i)|/a^2$, where $\mathbf{r}_k$ is the position vector of the k^{th} bead. The quantity $\beta = (k_{\text{B}}T)^{-1}$, where k_{B} is the Boltzmann constant and T is the absolute temperature. The bending constant κ is obtained by solving the equation

$$\frac{b}{a} = \frac{\kappa - 1 + \kappa\coth\kappa}{\kappa + 1 - \kappa\coth\kappa}, \tag{4.2}$$

where $b = 2l_{\text{p}}$ is the Kuhn length of the chain (Schellman, 1974; Tree et al., 2013b). Note that in the limit $\kappa \to 0$, Eq. 4.2 reduces to the familiar result for a freely joined chain (FJC), $b = a$. The effect of self-avoidance is incorporated by a hard sphere potential of the form

$$\beta U_{\text{EV}} = \begin{cases} \infty & \text{if } r_{ij} \le w, \\ 0 & \text{otherwise,} \end{cases} \tag{4.3}$$

where r_{ij} is the distance between the centers of the beads i and j, and w is the width of the chain.

4.2.2 Variants of the DWLC model

In order to investigate the effects of stiffness and excluded volume on properties of confined chains, we adopted a four-step approach in which we independently turned on/off stiffness and excluded volume. Specifically we performed simulations for (i) ideal freely jointed chains (IFJCs); (ii) real freely jointed chains (RFJCs); (iii) ideal wormlike chains (IWLCs) and (iv) real wormlike chains (RWLCs), all confined in square channels of various sizes.

Since the words "ideal" and "real" appear in different contexts in various areas of polymer physics (Rubinstein and Colby, 2003; Grosberg and Khokhlov, 1994; Doi and Edwards, 1986), let us clarify what we mean by these terms. We use the word "ideal" to refer to chains in which there are no excluded volume interactions between the beads. In contrast, "real" refers to chains in which there exist repulsive interactions between non-neighboring beads along the contour, as in a polymer dissolved in a good solvent

Table 4.1: Presence or absence of stiffness and excluded volume in the four models considered here.

	Bending stiffness (κ)	Excluded volume (w)
Ideal Freely Jointed Chain (IFJC)	$= 0$	$= 0$
Real Freely Jointed Chain (RFJC)	$= 0$	> 0
Ideal Wormlike Chain (IWLC)	> 0	$= 0$
Real Wormlike Chain (RWLC)	> 0	> 0

(Rubinstein and Colby, 2003). Thus, a given chain is ideal in the limit $w = 0$, and is real for $w > 0$. Similarly, in the limit $\kappa = 0$, we recover a freely jointed chain (FJC) from the DWLC model, and a wormlike chain (WLC) for $\kappa > 0$.

Note that the term "RFJC model" is somewhat of a misnomer, since one normally assumes that there are no correlations between bond vectors along the contour in a freely jointed chain model (Rubinstein and Colby, 2003; Grosberg and Khokhlov, 1994; Doi and Edwards, 1986). The excluded volume interactions do introduce correlations between distal segments of the chain. However, we use the term RFJC to mean that the chain behaves like a FJC locally, as there is no bending stiffness between different beads of the chain. Therefore, the four kinds of models considered in our work are all special cases of the DWLC model for different values of κ and w. Table 4.1 summarizes the properties of the four variants of the DWLC model used here.

4.2.3 Pruned-Enriched Rosenbluth Method (PERM) Simulation

As in our previous work (Tree et al., 2013a; Tree et al., 2013b), we use an off-lattice PERM algorithm to compute equilibrium properties of single polymer chains in confinement. PERM is a Monte Carlo chain-growth algorithm that enables growth of a collection of single polymer molecules which approximately follow a given probability distribution (Grassberger, 1997; Prellberg and Krawczyk, 2004). In contrast, the more common Metropolis Monte Carlo (MMC) algorithm employs "moves" in the conformations of a polymer chain of a certain length starting from an initial condition (Binder, 1995; Frenkel and Smit, 2001). Algorithms based on the Metropolis method are referred to as dynamic Monte Carlo algorithms as they sample states from a Markov chain of states, i.e. a move is chosen probabilistically depending on the previous state (van Rensburg, 2009). On the other hand, PERM belongs to the class of so-called approximate counting algorithms

(van Rensburg, 2009). Every sample in a PERM simulation consists of a "tour" of configurations grown from a single bead to a maximum of $N+1$ beads. Each step of growth ranging from $i = 0$ to $i = N$ in a tour is associated with a weight W_i, the average $\langle W_i \rangle$ of which is an estimate of the partition function, Z_i, relative to the reference state.

PERM is a powerful technique for single molecule simulation of polymers and has several advantages over conventional MMC methods. First, the free energy of a system (with respect to a reference state) can be estimated directly from the simulation using

$$\beta \Delta F_i = -\ln \frac{Z_i}{Z_{\mathrm{ref},i}} \approx -\ln \left\langle \frac{W_i}{W_{\mathrm{ref},i}} \right\rangle. \tag{4.4}$$

Here ΔF_i is the free energy of the polymer of length i, and $W_{\mathrm{ref},i}$ and $Z_{\mathrm{ref},i}$ are the weight and the partition function corresponding to the reference state. The average $W_i/W_{\mathrm{ref},i}$ is computed over a total of N_t tours. The reference state we choose in our simulations depends on the variant of the DWLC model under consideration as we explain later in Section 4.2.4. Second, all equilibrium properties can be estimated as a function of contour length in a single simulation. For example, the mean span X_i at any step i can be estimated as

$$\langle X_i \rangle = \frac{\sum_{t=1}^{N_t} W_i^t X_i^t}{\sum_{t=1}^{N_t} W_i^t}, \tag{4.5}$$

where W_i^t is the weight associated with the i^{th} step after the t^{th} tour among a maximum of N_t tours. This attribute of PERM of calculating properties as a function of contour length forms the methodological cornerstone of this paper.

Our PERM simulations are based on a parallel master-slave algorithm consisting of three stages, similar to our previous paper (Tree et al., 2013b): (i) an initial "blind" run to obtain estimates of the partition function for a few hundred beads; (ii) several "non-blind" runs where we incrementally increase the number of beads until the maximum number of beads are reached; (iii) finally, a "full" run wherein properties of interest such as span or end-to-end distance are calculated (Prellberg and Krawczyk, 2004; Prellberg, 2013). Our algorithm employs a neighbor list resembling the Verlet list (Allen and Tildesley, 1987), unlike the hashing-based neighbor lists used in lattice simulations (Grassberger and Hegger, 1995; Grassberger, 1997; Madras and Sokal, 1988). Nonetheless, a typical simulation in a given channel with 10^5 beads for 4×10^5 tours running on

96 6-core Intel Xeon X7542 "Westmere" processors takes between 6-8 hours of walltime, which is equivalent to about 700 CPU hours. This large number of tours produce thousands of independent realizations of chains spanning the configuration space, resulting in small error bars. Unless otherwise specified, error bars are smaller than symbol sizes in all our figures. For further details on our implementation of the simulation, we direct the reader to previous work from our group (Tree et al., 2013b; Tree et al., 2013a).

4.2.4 Free energy calculations

In all our simulations, the reference state for the estimation of the partition function is an unconfined chain with no excluded volume. Other attributes of the reference chain, such as a and l_p, are set equal to the chain under consideration (Tree et al., 2013b). In other words, the reference state for simulations of both RWLCs and IWLCs is an unconfined IWLC with the same persistence length. Likewise, the reference state for both IFJCs and RFJCs is an unconfined IFJC.

The confinement free energy of a chain without excluded volume, $\Delta F_{\mathrm{conf,\ id}}$, is calculated by

$$\beta\Delta F_{\mathrm{conf,\ id}} = -\ln\frac{Z_\mathrm{conf}}{Z_\mathrm{ref}} \approx -\ln\left\langle\frac{W_\mathrm{conf}}{W_\mathrm{ref}}\right\rangle, \tag{4.6}$$

where the average of ratio of the weights of the confined chain and the unconfined chain, $W_\mathrm{conf}/W_\mathrm{ref}$, is evaluated in the course of the simulation. Note that $\langle W_\mathrm{conf}/W_\mathrm{ref}\rangle$ is an estimate of the ratio of the respective partition functions, $Z_\mathrm{conf}/Z_\mathrm{ref}$.

Calculation of the confinement free energy of chains with excluded volume involves an additional step as the reference state in our simulations is always an unconfined, ideal chain with the same persistence length. First, the free energy of an unconfined chain with excluded volume, ΔF_EV, is obtained from a free-solution PERM simulation by

$$\beta\Delta F_\mathrm{EV} = -\ln\frac{Z_\mathrm{EV}}{Z_\mathrm{ref}} \approx -\ln\left\langle\frac{W_\mathrm{EV}}{W_\mathrm{ref}}\right\rangle, \tag{4.7}$$

where Z_EV and W_EV are the partition function and the weight of the unconfined real chain respectively. Second, the free energy of a confined real chain, $\Delta F_\mathrm{conf+EV}$, is calculated through the estimation of the partition function, $Z_\mathrm{conf+EV}$, from a PERM simulation of the confined chain as

$$\beta\Delta F_\mathrm{conf+EV} = -\ln\frac{Z_\mathrm{conf+EV}}{Z_\mathrm{ref}} \approx -\ln\left\langle\frac{W_\mathrm{conf+EV}}{W_\mathrm{ref}}\right\rangle, \tag{4.8}$$

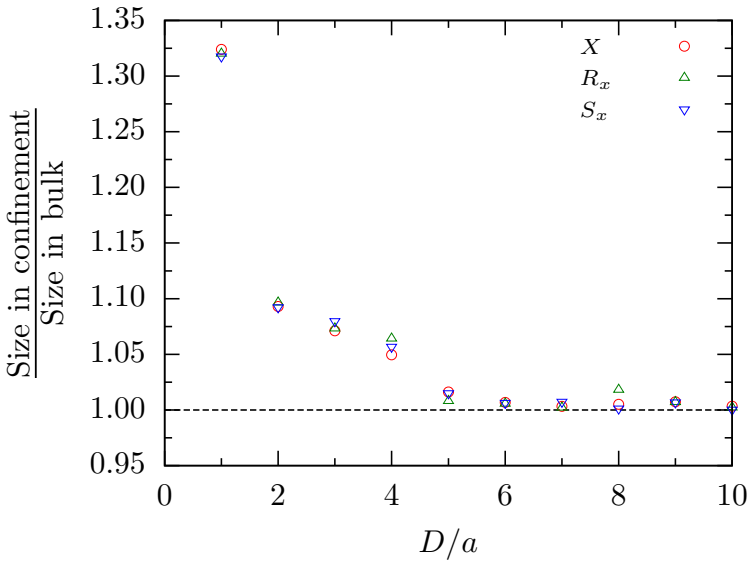


Figure 4.2. Mean span (X), x-projection of the end-to-end distance (R_x) and $x-x$ component of the gyration tensor (S_x), normalized by the corresponding quantities in free solution for an ideal FJC, plotted against D/a. The dashed line at unity is the expected result in the absence of discretization error(de Gennes, 1979). All data points show simulation results for a contour length $L/a = 5000$.

where $W_{\text{conf+EV}}$ is the corresponding Rosenbluth weight. Finally, the confinement free energy of the real chain, $F_{\text{conf, EV}}$, is calculated as

$$\Delta F_{\text{conf, EV}} = \Delta F_{\text{conf+EV}} - \Delta F_{\text{EV}}. \tag{4.9}$$

4.2.5 Discretization

Three independent parameters characterize our DWLC model: l_{p}, w and a. Given l_{p} and w, the ratios l_{p}/a and w/a determine the number of discrete units per persistence length and width respectively. By definition, the Kuhn length b must satisfy $b \geq a$, because in the limiting case of a freely joined chain (FJC), we have $b = a$. Moreover, if $w < a$ there is a possibility of the chain crossing itself, and this may result in an erroneous representation of self-repulsion effects. This implies that $w \geq a$. Here, we choose $a = w$, (Tree et al., 2013a) as this yields the best compromise between correct physical representation and computational tractability. In all our chains, $b \geq w$, and hence $a = w$ satisfies the constraints imposed by both stiffness and excluded volume on the bond length.

Channel confinement introduces an additional parameter D, the channel size. A very small D/a ratio can lead to discretization artifacts in the simulation results. In addition, obtaining a reliable number of samples from PERM becomes increasingly difficult as D/a decreases because, for small channels, the attrition rate is high due to collision with walls. Therefore, to determine the role of D/a ratio in our simulations, we calculated the size of ideal FJCs in confinement through PERM simulations.

Here, we make use of the fact that the size of an ideal flexible chain in the direction parallel to the channel is unaffected by confinement (de Gennes, 1979). Note that the FJC does not model a truly "flexible chain" because of the rod-like nature of the model for length scales smaller than the step size, b (Fredrickson, 2006). However, an unconfined IFJC reproduces the properties of a flexible polymer model such as a continuous Gaussian chain, given that $L \gg b$ (Doi and Edwards, 1986; Fredrickson, 2006). It is in this context that we use FJC as a model for a flexible polymer. Consequently for an IFJC, we expect that the mean span X,

$$X \equiv \langle \max(x_i) - \min(x_i) \rangle \,, \text{ where } (i \in [1, N+1]), \tag{4.10}$$

the x-projection of the end-to-end distance R_x,

$$R_x \equiv \sqrt{\left\langle (x_{N+1} - x_1)^2 \right\rangle}, \tag{4.11}$$

and the $x - x$ component of the gyration tensor S_x,

$$S_x \equiv \sqrt{\left\langle \frac{1}{N+1} \sum_{i=1}^{N+1} (x_i - x_{\mathrm{cm}})^2 \right\rangle} \tag{4.12}$$

in confinement must be identical to their respective values in free solution. Here, x_i is the x-coordinate of the i^{th} bead, among a total of $N+1$ beads. However, when the bond length is of the order of the confinement size, the assumption that the chain is flexible at the length scale of the channel is no longer valid. Hence the size in confinement is not equal to the size in the bulk.

Figure 4.2 shows the ratio of these quantities with the corresponding quantities in free solution. As expected, the discretization error increases as D/a decreases. Moreover, for $D/a \geq 5$ there is no significant change in the ratio as D/a increases, and in this region, the ratio of the sizes of the confined and unconfined chains is approximately 1. We use this condition as a rule of thumb to choose the minimum D/a ratio for our simulations

of FJCs. Accordingly, all our simulations for FJCs in this paper are done with a ratio $D/a \geq 5$.

4.3 Results and discussion

In Section 4.2.2, we classified the DWLC model into 4 broad variants, depending on the values of w and κ. Here in Section 4.3.1, we examine the properties of confined chains when either of w or κ (or both) are 0. Thus the limiting cases of no stiffness and no excluded volume encompass (i) IFJCs (w = 0 and κ = 0), (ii) RFJCs (κ = 0), and (iii) IWLCs (w = 0). In Section 4.3.2, we show simulation results for RWLCs with non-zero values of both w and κ. We then end our discussion in Section 4.3.2.2 by outlining the implications of our findings to the case of DNA confined in nanochannels.

4.3.1 Limiting cases

4.3.1.1 Ideal freely jointed chains (IFJC)

With a carefully discretized DWLC model, we begin with a review of the extension and free energy of ideal flexible chains in confinement. The purpose of laying out the properties of IFJCs is twofold. First, we wish to recapitulate well known results for IFJCs in confinement, which will prove beneficial in elucidating the behavior in the presence of stiffness and excluded volume. Second, because exact analytical results are available for properties of ideal flexible chains in confinement, IFJCs are an ideal candidate for use in validating our simulation code to high precision.

For a flexible ideal chain, confinement does not affect extension parallel to the axis of the channel, i.e. the extension is identical to that in free solution (Daoud and de Gennes, 1977; de Gennes, 1979). This can be readily recognized by examining the probability distribution of the end-to-end vectors $\mathbf{R}$ in 3 dimensions, $P_{3D}(N, \mathbf{R})$, for a freely jointed chain with N segments. For an IFJC, the distribution functions in the three orthonormal directions $P_{1D}(N, \mathbf{R}_x)$, $P_{1D}(N, \mathbf{R}_y)$ and $P_{1D}(N, \mathbf{R}_z)$ are independent of each other, i.e., $P_{3D}(N, \mathbf{R}) = P_{1D}(N, \mathbf{R}_x)P_{1D}(N, \mathbf{R}_y)P_{1D}(N, \mathbf{R}_z)$ (Doi and Edwards, 1986; Werner et al., 2013). Therefore, for an IFJC confined in a square conduit as in Figure 4.1, confinement alters the distribution functions only in the y and z directions. In other words, $P_{1D}^{\text{conf}}(N, \mathbf{R}_x) = P_{1D}^{\text{free}}(N, \mathbf{R}_x)$. Consequently, metrics of size such as the averages of span (X), radius of gyration (S_x) and end-to-end distance (R_x) are unaltered

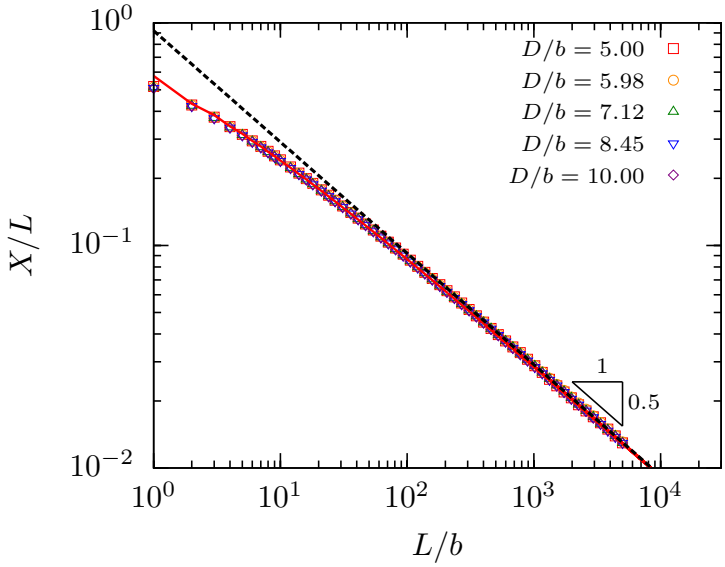


Figure 4.3. Average fractional extension as a function of contour length for IFJCs confined in square channels of various sizes. Most data points are not visible as they overlap each other exactly. The red line denotes the theoretical prediction in Eq. 4.13 for IFJCs in free solution (Daniels and Smithies, 1941). The dashed line is the result for the mean span in the long-chain limit given by $X = \sqrt{8Lb/3\pi}$ (Daniels and Smithies, 1941). All results show a scaling of $X \sim L^{0.5}$ in the long-chain limit, as expected.

due to confinement (Daoud and de Gennes, 1977; de Gennes, 1979).

We confirm this prediction by comparing extension of IFJCs confined in channels against the prediction for the mean span of IFJCs in free solution (Daniels and Smithies, 1941). To do so, we make use of an expression derived by Daniels and Smithies (1941) (Daniels and Smithies, 1941) for the mean span of a 1D random walk. We convert these equations to evaluate the mean span in 3D by dividing by $\sqrt{3}$ as

$$X = \begin{cases} \dfrac{1}{\sqrt{3}} \dfrac{2\mathcal{N}+1}{2^{\mathcal{N}}} \dfrac{\mathcal{N}!}{\left(\frac{\mathcal{N}}{2}\right)!\left(\frac{\mathcal{N}}{2}\right)!} - 1 \ (\mathcal{N} \text{ even}), \\ \dfrac{1}{\sqrt{3}} \dfrac{2\mathcal{N}+2}{2^{\mathcal{N}}} \dfrac{\mathcal{N}!}{\left(\frac{\mathcal{N}}{2}-\frac{1}{2}\right)!\left(\frac{\mathcal{N}}{2}+\frac{1}{2}\right)!} - 1 \ (\mathcal{N} \text{ odd}), \end{cases} \tag{4.13}$$

where $\mathcal{N}$ is the normalized contour length L/b. In Figure 4.3, we observe that the fractional extension of IFJCs parallel to the channel axis is independent of the channel size and is unperturbed by confinement in the channel. However, the prediction of Eq. 4.13 for the mean span in 3D for $L/b \approx 1$ is not accurate as the prefactor $\sqrt{3}$ is valid only when $L \gg b$. Nonetheless, the collapse of the extension data for all channels gives further credence to the absence of discretization artifacts in these simulations.

Although the extension parallel to the channel is unchanged, there is an entropic penalty to squeeze the chain into the channel, and hence the confined chain has a higher free energy than the chain in the bulk. The change in free energy due to confinement in a channel for a FJC with $D \gg b$ and $L > b$ can be approximated as (Gorbunov and Skvortsov, 1995)

$$\begin{aligned} \beta\Delta F &= -\ln\left[\frac{P_{1D}^{\text{conf}}(N,\mathbf{R}_x)P_{1D}^{\text{conf}}(N,\mathbf{R}_y)P_{1D}^{\text{conf}}(N,\mathbf{R}_z)}{P_{1D}^{\text{free}}(N,\mathbf{R}_x)P_{1D}^{\text{free}}(N,\mathbf{R}_y)P_{1D}^{\text{free}}(N,\mathbf{R}_z)}\right] \\ &= -\ln\left(\frac{P_{1D}^{\text{conf}}(N,\mathbf{R}_y)}{P_{1D}^{\text{free}}(N,\mathbf{R}_y)}\right) - \ln\left(\frac{P_{1D}^{\text{conf}}(N,\mathbf{R}_z)}{P_{1D}^{\text{free}}(N,\mathbf{R}_z)}\right). \end{aligned} \tag{4.14}$$

It is evident from Eq. 4.14 that the free energy of confinement arises due to constraints imposed in the y and z directions by the walls of the channel. Figure 4.4(a) shows the free energy of a confined IFJC for various channel sizes as a function of contour length L. Unlike extension, the free energy of confinement in different channels are never equal. Even at extremely small contour lengths (order of a few beads), the free energy values are unequal in different channels because of a different degree of entropy loss. However, in the long-chain limit, $\Delta F/L$ asymptotes to a constant value (Doi and Edwards, 1986)

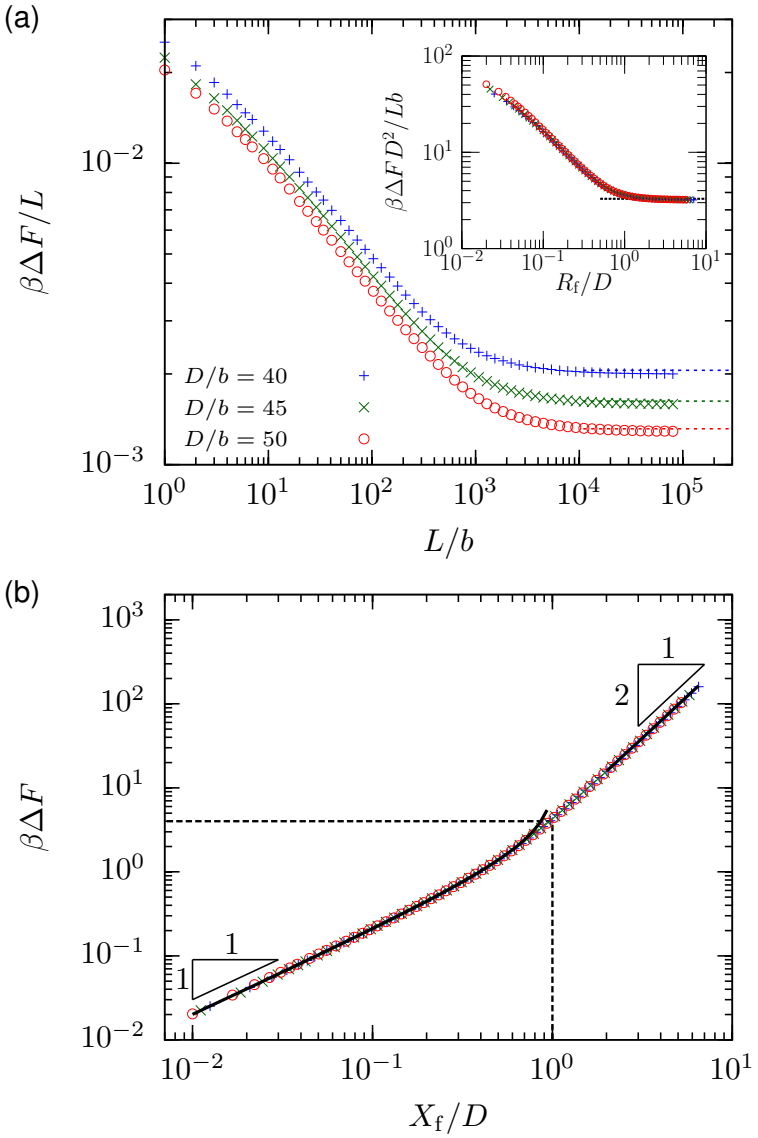


Figure 4.4. (a) Free energy of ideal FJCs in 3 square channels of different sizes. The free energy values asymptote to the value given by Eq. 4.15 (dashed lines) (Chen and Sullivan, 2006; Tree et al., 2013a). In the inset, the free energy values are shown to reach asymptotic scaling for all 3 channels when $R_f \approx D$. The dashed line signifies the theoretically expected limit, $\pi^2/3$. (b) Free energy data for the 3 channels fall on a universal curve as per scaling theory, with an asymptotic scaling $\Delta F \sim (X_f/D)^2$ for $X_f \gg D$. The dashed lines show that the free energy of confinement is approximately $4k_BT$ for $X_f/D = 1$. The solid black lines show theoretical predictions: Eq. 4.17 for $X_f/D \lesssim 1$ (Gorbunov and Skvortsov, 1995; Wang et al., 2010) and Eq. 4.15 for $X_f/D \gtrsim 1$. A scaling of $\Delta F \sim (X_f/D)^1$ is also seen for $X_f \ll D$.

and this asymptotic value when $L \gg D \gg b$ is given by (Chen and Sullivan, 2006; Tree et al., 2013a)

$$\beta\Delta F = \frac{\pi^2 Lb}{3D^2}. \tag{4.15}$$

Our simulation results in Figure 4.4 show excellent agreement with this prediction.

Furthermore, akin to de Gennes' scaling theory for flexible real chains (de Gennes, 1979; Daoud and de Gennes, 1977; Freed, 1987), a scaling equation of the form

$$\beta\Delta F = \phi_{\Delta F}\left(\frac{R_\mathrm{f}}{D}\right) \tag{4.16}$$

can be written, where R_f is the root mean square end-to-distance in free solution and $\phi_{\Delta F}$ is a function such that $\phi_{\Delta F}(x) \sim x^2$ as $x \to \infty$ in the limit $R_\mathrm{f} \gg D \gg b$ (Gorbunov and Skvortsov, 1995). However, Eq. 4.16 does not hold if $R_\mathrm{f} \ll D$. Fortunately, in the limit $R_\mathrm{f} \ll D$ or more precisely $X_\mathrm{f} \ll D$, the confinement free energy in a channel is given by

$$\beta\Delta F = -\ln\left(1 - \frac{X_\mathrm{f}}{D}\right)^2, \tag{4.17}$$

where X_f is the mean span in free solution (Gorbunov and Skvortsov, 1995; Wang et al., 2010). Now, the confinement free energy can be written as a similar function of the mean span in free solution as

$$\beta\Delta F = \Phi_{\Delta F}\left(\frac{X_\mathrm{f}}{D}\right). \tag{4.18}$$

Since the ratio $X_\mathrm{f}/R_\mathrm{f}$ is a constant for $R_\mathrm{f}, X_\mathrm{f} \gg b$, i.e $X_\mathrm{f}/R_\mathrm{f} \to \sqrt{8/3\pi}$ (Daniels and Smithies, 1941), $\Phi_{\Delta F}(x)$ has the following properties:

$$\Phi(x): \begin{cases} = -\ln\left(1 - x^2\right) & \text{if } x \ll 1, \\ \sim x^2 & \text{if } x \gg 1. \end{cases} \tag{4.19}$$

To verify Eq. 4.19, we plot the confinement free energy against X_f/D for 3 channel sizes as shown in Figure 4.4(b). We observe that the free energy values for the three channels considered all fall on the same curve irrespective of the contour length.

Plotting free energy as a function of X_f/D also provides us with a way to identify cut-offs for the onset of asymptotic scaling. We find two distinct regimes in Figure 4.4(b): (i) a weak confinement regime, where the bulk size of the chain is less than that of

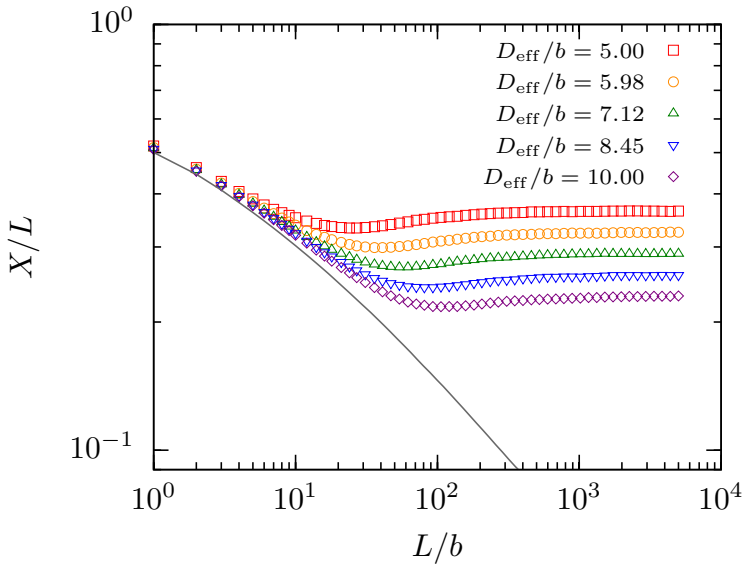


Figure 4.5. Fractional extension of RFJCs (a = w = b) confined in channels with same D_{eff} values as in Figure 4.3. The gray line corresponds to simulation results of a RFJC in free solution.

the channel width, and (ii) a strong confinement regime, where the average size of the chain in the bulk is greater than D. In the former, our free energy calculations agree with Eq. 4.17. Moreover, our data in Figure 4.4(b) clearly show a distinct scaling of $\beta\Delta F \sim X_{\mathrm{f}}/D$ in the weak confinement regime, because $\beta\Delta F = -\ln(1 - X_{\mathrm{f}}/D)^2 \approx 2X_{\mathrm{f}}/D$ if $X_{\mathrm{f}} \ll D$.

In the strong confinement regime ($X_{\mathrm{f}} > D$), the free energy values exhibit asymptotic scaling for $X_{\mathrm{f}}/D \gtrsim 1$, endorsing the fact the free energy attains limiting scaling when the chain as a whole feels the effect of confinement. Figure 4.4(b) also shows that $\Delta F \approx 4k_{\mathrm{B}}T$ when $X_{\mathrm{f}} = D$ thus supporting the widely prevalent notion that the confinement free energy is $\mathcal{O}(k_{\mathrm{B}}T)$ for a chain of unconfined size D.

4.3.1.2 Real freely jointed chains (RFJC)

In the presence of repulsive excluded volume interactions, confined flexible chains follow statistics outlined by Daoud and de Gennes (1977) in their seminal paper. de Gennes (de Gennes, 1979) proposed that for a chain confined in a cylinder of diameter D_{cyl}, in the long chain limit ($L \gg D_{\mathrm{cyl}}$), "the chain behaves as a sequence of blobs of diameter D_{cyl} which act as hard spheres and pack into a regular one-dimensional array (de Gennes, 1979)" A similar depiction of the chain as a 1D self-avoiding walk of blobs in

square confinement leads to a predicted scaling, $X \sim L$. Figure 4.5 shows fractional extension of RFJCs confined in square channels with the same D_{eff} values as in Figure 4.3. We observe that at very small chain lengths, the extension of the chain is the same as in free solution. However, as the chain length increases, the chain begins to backfold because of collisions with the walls. In contrast to the IFJC, the increase in backfolding induces self-repulsion, resulting in a rapid increase in the extension. Once the average monomer concentration attains a threshold value, backfolding at the length scale of the channel size D is no longer possible due to self-avoidance (Hsu and Grassberger, 2003; Zhang et al., 2008; Hsu and Binder, 2013). This leads to a linear ordering of segments of the chain resulting in a linear scaling, $X \sim L$.

Analogous to the case of ideal flexible chains, the properties of real flexible chains in confinement have been studied with the use of diffusion-annihilation equations (Edmund A. DiMarzio, 1965; Werner et al., 2013). However, because of the difficulty in determining the annihilation term, a mean-field approximation is employed to account for the excluded volume interactions by means of self-consistent field theory (Werner et al., 2013). Werner et al. (2013) recently showed that mean-field approximation does not yield satisfactory results as it neglects intrachain correlations, which are especially critical in determining properties of self-avoiding chains in confinement.

Nonetheless, scaling theory can still be applied to interpret the behavior of RFJCs in confinement. For $X_{\text{f}} \gg D_{\text{eff}} \gg b$, the extension parallel to the axis of a channel was written by de Gennes in a scaling form as (Daoud and de Gennes, 1977; de Gennes, 1979; de Gennes, 1999)

$$\frac{X}{X_{\text{f}}} = \Phi_X\left(\frac{X_{\text{f}}}{D_{\text{eff}}}\right), \tag{4.20}$$

where X is the extension (mean span) in confinement, and X_{f} is the mean span in free solution. Similarly, a relation analogous to Eq. 4.18 for confinement free energy ΔF is (Daoud and de Gennes, 1977; de Gennes, 1979)

$$\beta\Delta F = \Phi_{\Delta F}\left(\frac{X_{\text{f}}}{D_{\text{eff}}}\right), \tag{4.21}$$

where the functions Φ are defined such that $\Phi_X(x) \sim x^{1/\nu_{\text{F}}-1}$ and $\Phi_{\Delta F}(x) \sim x^{1/\nu_{\text{F}}}$ for $x \gg 1$ in accordance with the scaling laws in the de Gennes regime. Here, $\nu_{\text{F}} \simeq 0.5877$ is the Flory exponent (Li et al., 1995; Clisby, 2010). Also, recall from Section 4.2.1 that $D_{\text{eff}} = D - w$ is the effective width of the channel accessible to the centerline of the

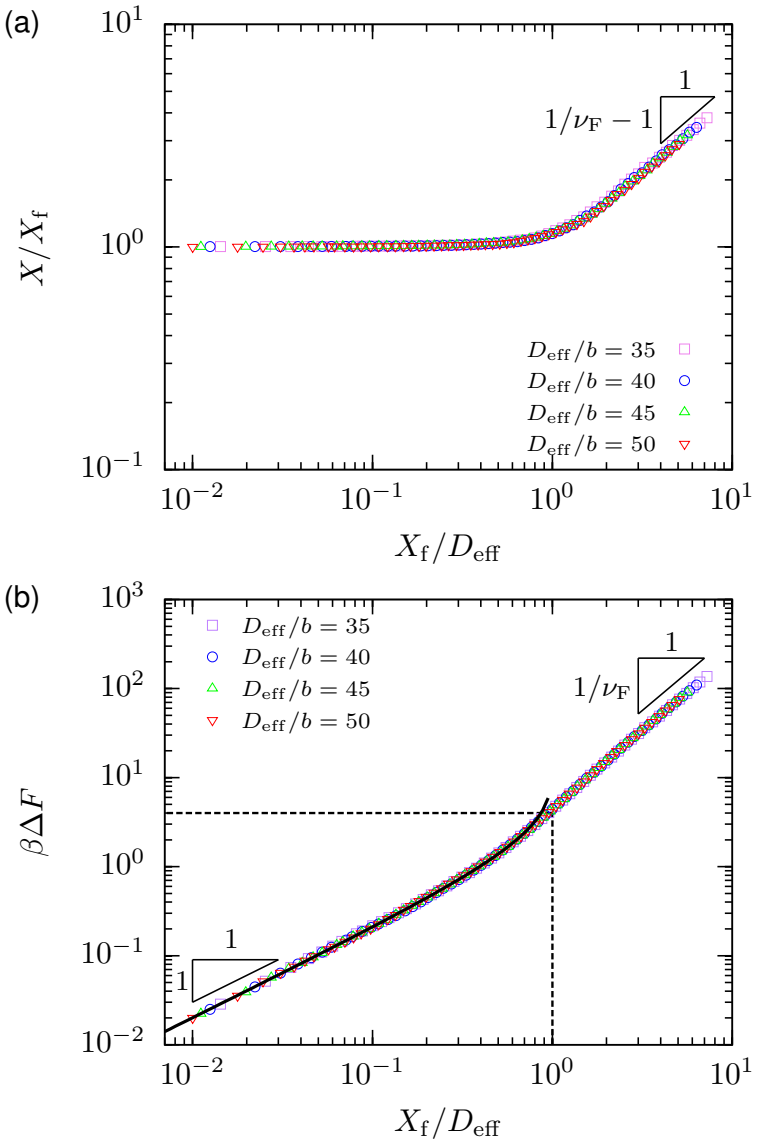


Figure 4.6. (a) Ratio of extension in confinement with extension in the bulk versus ratio of the free-solution mean span with the channel size for 4 channels. (b) Confinement free energy for the chains under the same conditions. The dashed lines show that the confinement free energy is $\mathcal{O}(k_\mathrm{B}T)$ when $X_\mathrm{f}/D_\mathrm{eff} = 1$. The prediction in Eq. 4.22 is shown as a solid black line. Also shown is the scaling, $\beta\Delta F \sim X_\mathrm{f}/D_\mathrm{eff}$, in the weak confinement regime. Notice the scaling of both properties in the long-chain limit: $X/X_\mathrm{f} \sim (X_\mathrm{f}/D_\mathrm{eff})^{1/\nu_\mathrm{F}-1}$, and $\Delta F \sim (X_\mathrm{f}/D_\mathrm{eff})^{1/\nu_\mathrm{F}}$. The data shown here corresponds to RFJCs of contour length L ranging from $L = b$ to $L = 5000b$.

backbone of the polymer molecule. Note that for $x \ll 1$, $\Phi_{\Delta F}(x) = (1-x)^2$, wherein the quantity D in Eq. 4.17 has been replaced by D_{eff}. Agreement with Eqs. 4.20 and 4.21 is indicated by the collapse of the X/X_{f} curves in Figure 4.6(a) and the $\beta\Delta F$ curves in Figure 4.6(b) respectively.

Furthermore, we observe in Figure 4.6(a) that confined RFJCs behave virtually like unconfined chains until the unconfined chain size X_{f} is approximately equal to the channel size D_{eff}. In other words, $X \approx X_{\text{f}}$ for $X_{\text{f}} \lesssim D_{\text{eff}}$. Beyond this threshold value, i.e. when $X_{\text{f}} \gtrsim D_{\text{eff}}$, there is a sharp increase in the extension of the confined chain culminating in a scaling of $X/X_{\text{f}} \sim (X_{\text{f}}/D_{\text{eff}})^{1/\nu_{\text{F}}-1}$.

The confinement free energy also exhibits two distinct regimes as shown in Figure 4.6(b). We find that free energy in the weak confinement regime ($X_{\text{f}} \lesssim D_{\text{eff}}$) agrees well with the following relation, which is essentially a restatement of Eq. 4.17:

$$\beta\Delta F = -\ln\left(1 - X_{\text{f}}/D_{\text{eff}}\right)^2 . \tag{4.22}$$

We also note that the slope of the ΔF curve abruptly changes when $X_{\text{f}}/D_{\text{eff}} \approx 1$ to a slope $1/\nu_{\text{F}}$ as predicted by Daoud and de Gennes (1977) (Daoud and de Gennes, 1977), indicating that the statistics of confined RFJCs transition quickly to the long-chain limit once the chain size is large enough to form a blob of size D_{eff}. A careful examination of Figure 4.6(b) shows that $\Delta F \approx 4k_{\text{B}}T$ for $X_{\text{f}} \approx D_{\text{eff}}$, which is in agreement with previous predictions that the confinement free energy of a blob is $\mathcal{O}(k_{\text{B}}T)$ (de Gennes, 1979; Daoud and de Gennes, 1977; Raphael and Pincus, 1992; Zhang et al., 2008).

4.3.1.3 Ideal wormlike chains (IWLC)

In the seminal paper on properties of confined stiff polymers in extreme confinement, Odijk (1983) introduced a new length scale, $\lambda \simeq D^{2/3} l_{\text{p}}^{1/3}$, that corresponds to the length of a deflection segment (Odijk, 1983). According to Odijk's theory, the extension of IWLCs with $L \gg \lambda$ in a square channel of size D can be written as

$$X = L\left[1 - \alpha_{\square}\left(\frac{D}{l_{\text{p}}}\right)^{2/3}\right], \tag{4.23}$$

when $D \ll l_{\text{p}}$. The prefactor has been calculated by Yang et al. (2007) as $\alpha_{\square} = 0.18274$. Equation 4.23 shows that the scaling of the extension, $X \sim L$, is markedly different for IWLCs in comparison to IFJCs, where $X \sim L^{0.5}$ in the long-chain limit. Moreover, the

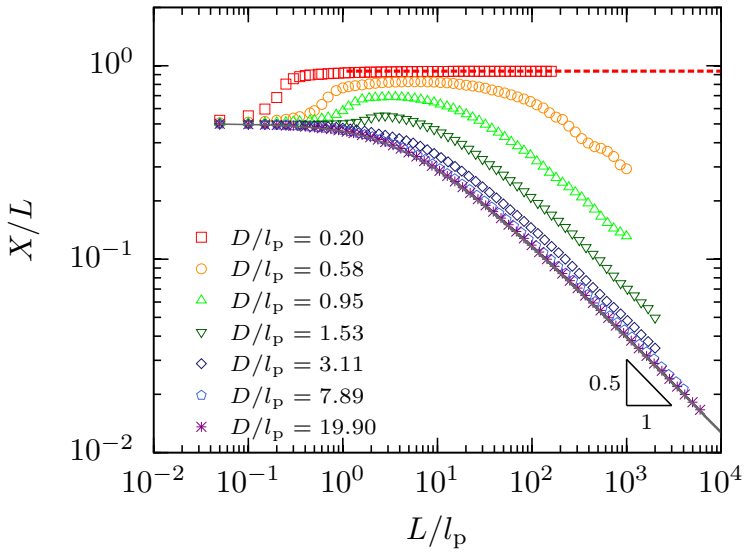


Figure 4.7. Fractional extension of IWLCs with $l_p/a = 20$ versus contour length for various channel sizes. The gray line shows simulation results for fractional extension of the chain in free solution with an expected slope of -0.5 in the long-chain limit. The red dashed line corresponds to the prediction in the Odijk regime given by Eq. 4.23 (Yang et al., 2007).

extension of IFJCs is equal to the bulk value and therefore does not depend on D, unlike the extension of IWLCs in the Odijk regime. Furthermore, little is known about the behavior of IWLCs when $l_p \gtrsim D$, as most studies with WLCs involve chains with excluded volume (Wang et al., 2011; Cifra et al., 2009; Cifra, 2009; Cifra and Bleha, 2012). Here, we study the extension behavior of IWLCs in confinement as a function of channel size and contour length, in the interest of demarcating the effect of stiffness on the properties of confined semiflexible polymers.

The fractional extension of stiff ideal chains confined in channels of various sizes is shown in Figure 4.7. We note here that unlike the figures in Sections 4.3.1.1 and 4.3.1.2, we plot our data with the contour length normalized with the persistence length, l_p, as is the norm for wormlike chains. For all the channels, we observe that X/L starts at 0.5, which is the value of the fractional extension in the bulk for a rod given by

$$\frac{X}{L} = \frac{\int_0^\pi |\cos\theta| \sin\theta \mathrm{d}\theta}{\int_0^\pi \sin\theta \mathrm{d}\theta} = 0.5, \tag{4.24}$$

where θ is the polar angle with the x axis (Wang et al., 2008). At these values of the

contour length L, the chain is almost unaffected by confinement and has virtually the same rotational freedom as in free solution. For $D \lesssim l_p$, as the chain length increases, the chain begins to feel the walls and aligns with the axis of the channel. This is reflected in Figure 4.7 as a rise in the fractional extension from the bulk value, at short chain lengths. For $D/l_p = 0.20$, the fractional extension is practically at the asymptotic value when L/l_p is approximately 0.5. At this value of L, the number of deflection segments ($\lambda \approx l_p^{1/3} D^{2/3}$) is approximately $L/(l_p^{1/3} D^{2/3}) \approx 1.5$, even though the chain is smaller than a single persistence length. This implies that the transition to asymptotic Odijk extension is rather quick, i.e. a chain that can form a few deflection segments is sufficiently long to exhibit behavior as in Eq. 4.23.

For $D/l_p = 0.58$ and 0.95, we observe significant stretching in the unconfined direction, i.e. $X/X_f > 1$, as shown in Figure 4.7. Particularly in the case of $D/l_p = 0.58$, the chain seems to follow Odijk-like $X \sim L$ behavior till about $L/l_p = 10$. However, after the chain reaches a certain length, which Odijk (2006) terms the global persistence length (Odijk, 2006), the entropic constraints of the walls are not strong enough to prevent formation of hairpins. Formation of hairpins by ideal chains leads to an extension X that is not linear in the contour length (Odijk, 2008; Cifra and Bleha, 2012). As the channel size increases, the amount of backfolding of these IWLCs increases, leading to decreasing fractional extension X/L with L. We note here that the noise in the data shown in Figure 4.7 at high L for intermediate channel sizes is due to high fluctuations of extension in this region of confinement. Similar behavior has been observed before, both in simulations (Tree et al., 2013c) and experiments (Reisner et al., 2012; Reisner et al., 2005).

The fluctuations in the extension decrease as the channel size gets much bigger than the persistence length. When $D/l_p = 19.90$, the fractional extension of the chain is indistinguishable from the free solution curve, indicating stiffness has very little effect on stretching of the chain in this regime. In other words, the chain behaves exactly like an IFJC confined in a wide channel. For channel sizes such as $D/l_p = 3.11$ and $D/l_p = 7.89$, however, we observe that the fractional extension is almost identical to that of a chain in bulk, though there is moderate stretching.

Confinement free energy of channel-confined IWLCs as a function of channel size has been studied by our group previously (Tree et al., 2013a). It was found that the confinement free energy transitions from $\Delta F_{conf} \sim D^{-2/3}$ in the Odijk regime to $\Delta F_{conf} \sim D^{-2}$ in weak confinement (Tree et al., 2013a), in agreement with free energy calculations

of IWLCs confined in slits (Chen and Sullivan, 2006). In addition, these results are also in qualitative agreement with recent work on confinement free energy of IWLCs in channels (Chen, 2013a). This D^{-2} scaling of the free energy is analogous to the case of confined IFJCs shown in Section 4.3.1.1, again reinforcing the fact that confined IWLCs show the same properties of equivalent IFJCs in weak confinement.

In summary, extension of IWLCs transitions from the Odijk regime in extreme confinement ($D \ll l_p$) to the behavior of unconfined IWLCs in weak confinement ($D \gg l_p$). We term this regime of weak confinement the "bulk-like" regime for IWLCs, as the extension is given by the bulk extension, just as in the case of confined IFJCs. It is interesting to note that our results also indicate the presence of a broad transition regime ($0.5 \lesssim D/l_p \lesssim 20$) between the Odijk regime and the bulk-like regime.

4.3.2 Real wormlike chains (RWLC)

4.3.2.1 Theoretical aspects

We now turn our attention to the case of RWLCs, which exhibit a nontrivial combination of the properties of RFJCs and IWLCs shown in Sections 4.3.1.2 and 4.3.1.3 respectively. When $D \ll l_p$, the chain is in the Odijk regime where the properties are dominated by chain stiffness. In contrast, when $D \gg l_p$ and $L \gg D$, we recover the de Gennes regime, which is governed by excluded volume. However, in the region $D \approx l_p$, there are two contributing effects leading to stretching in the unconfined direction, namely stiffness and excluded volume. This complicated interplay between stiffness and excluded volume is not well understood, and hence there is little consensus about the mechanism of stretching in this transition region.

Here, we seek to gain further understanding of the mechanism of stretching in confinement by examining the effect of contour length on the extension. Figure 4.8 shows the fractional extension of real wormlike chains (RWLCs) for the same D_{eff}/l_p values as in Figure 4.7. For $D_{eff}/l_p = 0.20$, the curve is identical to the case of an IWLC hinting that excluded volume does not have any effect on stretching of the chain in the Odijk regime. For bigger channels considered in Section 4.3.1.3, recall that we did not observe the scaling $X \sim L$, presumably due to long-range backfolding of the chain. In contrast, RWLCs reach an asymptotic value of the fractional extension in the long-chain limit, regardless of the channel size. For instance, in the case of $D_{eff}/l_p = 0.58$, the IWLC shows a drop in the fractional extension around $L/l_p = 20$ as shown in Figure 4.7. This descent

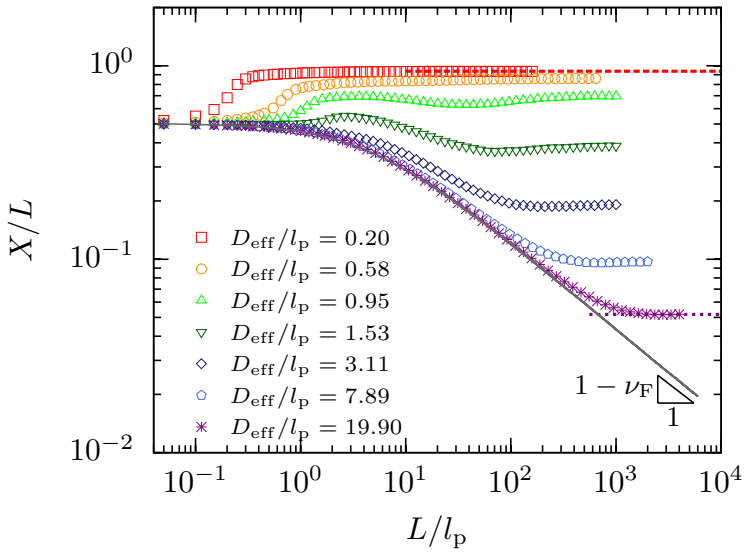


Figure 4.8. Fractional extension of RWLCs with $l_\mathrm{p}/a = 20$ versus contour length for the same D_eff values as in Figure 4.7. The gray line depicts the fractional extension for the chain in free solution with an expected slope of $-(1-\nu_\mathrm{F}) \simeq -0.42$ in the long-chain limit. Our results agree with predicted extension in the Odijk regime (red dashed line) (Yang et al., 2007) and prior results for the long-chain limit in the de Gennes regime (purple dotted line) (Tree et al., 2013a). However, most other curves fall in the transition regime.

of X/L is absent in Figure 4.8, showing that backfolding is minimal for RWLCs at this channel size. This stark contrast between extension behavior of IWLCs and RWLCs demonstrates that excluded volume between distal segments suppresses backfolding of confined semiflexible chains.

For larger channels such as the one with size $D_{\text{eff}}/l_{\text{p}} = 1.53$, the drop in X/L around $L/l_{\text{p}} = 10$ in Figure 4.8 signifies backfolding. However, as the chain length increases, this backfolding in turn leads to an amplification of excluded volume interactions. Accordingly, the curve plateaus to an asymptotic fractional extension with increasing chain length, showing a linear scaling in the contour length, $X \sim L$.

For the biggest channel, i.e. when $D_{\text{eff}}/l_{\text{p}} = 19.90$, we observe that the chain follows the free solution curve up to about $L/l_{\text{p}} = 100$, which is reminiscent of the behavior of IWLCs in weak confinement. However, self-avoidance induced by backfolding eventually causes the chain to swell in the direction parallel to the axis of the channel. Thereafter, the X/L curve in Figure 4.8 branches off from the free-solution curve, plateauing to the value previously observed in the de Gennes regime (Tree et al., 2013a). The resulting constant fractional extension is a consequence of the cigar-like conformation formed by the chain in this regime (de Gennes, 1979).

4.3.2.2 DNA in nanochannels and other practical implications

Following the seminal experiments by Reisner et al. (2005) (Reisner et al., 2005), there has been tremendous interest in understanding the static and dynamic properties of nanochannel-confined wormlike chains, particularly DNA (Köster et al., 2005; Odijk, 2006; Odijk, 2008; Cifra, 2009; Cifra and Bleha, 2012; Wang et al., 2011; Su et al., 2011; Tree et al., 2012; Tree et al., 2013a; Reinhart et al., 2013). In addition, nascent genome mapping technologies such as DNA barcoding in nanochannels are based on the assumption that there is a one-to-one linear relationship between the separation of markers in the channel and the genomic length between two points on a DNA molecule (Reisner et al., 2012; Dorfman et al., 2013; Wang et al., 2012). We know from results reported in the preceding sections that the linear scaling, $X \sim L$, for RWLCs in confinement is valid when the chain is long enough to contain several statistical segments such as deflection segments or blobs, depending on the regime of confinement. One of the aspects that is often overlooked is whether the data obtained from the aforementioned experiments and simulations satisfies conditions for the asymptotic linear scaling, $X \sim L$, in channel confinement.

We recently showed that the onset of asymptotic scaling, $X \sim L^{\nu_F}$, for unconfined DNA in a high ionic strength buffer occurs at a molecular weight of approximately 1 Mbp (Tree et al., 2013b). One of the consequences of this outcome is that λ-DNA, the experimental workhorse, lies well within the transition regime connecting $\nu_{id} = 0.5$, for an ideal coil and $\nu_F = 0.5877$, the universal exponent for swollen coils in free solution. It is unclear how these molecular-weight dependent results for chains in the bulk reconcile with those in quasi-1D square confinement as in a nanochannel. To resolve such uncertainties, we examine the finite length effects of DNA in nanochannels to establish limits for the validity of the linear scaling of the extension, $X \sim L$.

In order to simulate DNA in nanochannels, it is important to choose the appropriate set of parameters that account for the interactions present in an experimental system. As in our previous work (Wang et al., 2011; Tree et al., 2013a; Tree et al., 2013b), we use parameters for DNA in a high ionic strength (165 mM) buffer, 5×TBE, in which electrostatic interactions are highly screened. Such a value of the ionic strength allows us to use a hard-core repulsion for the excluded volume potential as a good approximation. Further, the value of the persistence length at high ionic strength is less ambiguous, as there exist discrepancies for the value of l_p at low ionic strengths (Tree et al., 2013b; Dobrynin, 2006; Skolnick and Fixman, 1977). Consequently, we use an effective width of $w = 4.6$ nm from Stigter's theory (Stigter, 1977) and a persistence length of $l_p = 53$ nm (Dobrynin, 2006; Hsieh et al., 2008) corresponding to the aforementioned ionic strength. Accordingly, the contour length of DNA in units of base pairs (bp) is obtained from Nw/l_{bp}, where $l_{bp} = 0.34$ nm is the rise per base pair in DNA.

In our recent work, simulations employing these parameters yielded good agreement with experimental scattering data for the radius of gyration of unconfined DNA (Tree et al., 2013b). Notwithstanding this agreement, our simulation results did not conform with single molecule fluorescence microscopy data (Tree et al., 2013b). We believe that this discrepancy may be due to the presence of intercalating dyes in these experiments. These dyes alter the size of DNA by changing the contour length (L) (Bakajin et al., 1998; Nyberg et al., 2013; Bennink et al., 1999; Günther et al., 2010), and conceivably even properties such as the width (w) and the persistence length (l_p) (Nyberg et al., 2013). There is also contrary evidence suggesting that the persistence length more or less remains constant and is independent of the amount of bound YOYO-1 (Günther et al., 2010). Additionally, the dependence of w on the dye concentration is still unclear. This makes direct comparison of simulations with experiments difficult,

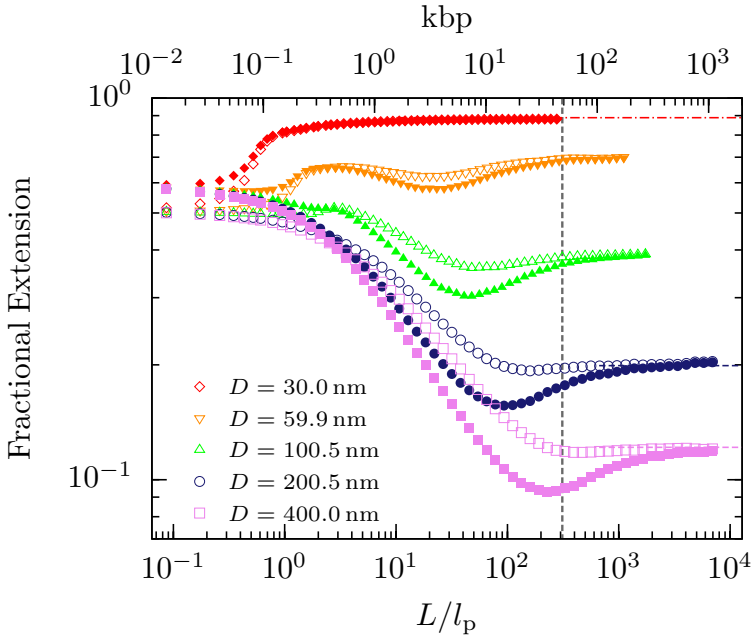


Figure 4.9. Mean span X (unfilled symbols) and projection of the end-to-end distance on the axis of the channel R_x (filled symbols) normalized by the contour length for confined, undyed DNA in a high ionic strength buffer. Note that the x axis at the bottom indicates the number of persistence lengths of the chain and the x axis at the top denotes the length of DNA in units of kilobase pairs. The red dashed-dotted line corresponds to the asymptotic fractional extension in the Odijk regime (Yang et al., 2007). The dashed lines corresponding to the two biggest channels show the empirical fit obtained by Tree et al. (2013a)] for the long-chain limit in the de Gennes regime. The molecular weight of λ-DNA is shown as a vertical gray dashed line.

as most experiments with DNA in nanochannels employ dyed DNA. Moreover, the effect of electrostatics on the wall-DNA interaction is still unclear (Reisner et al., 2012; Wang et al., 2011), and this further complicates modeling of DNA in a nanochannel. Our simulations thus make the simplifying assumption that the depletion width due to electrostatic interactions between the wall and DNA is $w/2$. In view of the confusion surrounding the right parameters for dyed DNA, we stick to the parameters for undyed, "naked" DNA listed in the previous paragraph.

Figure 4.9 displays results for extension of undyed DNA calculated using two metrics: (i) the mean span (X); and (ii) x-projection of the end-to-end distance (R_x). As expected, we see a trend similar to Figure 4.8. Among the two metrics shown, the span (X), which is the common experimental measure (Tegenfeldt et al., 2004; Reisner et al., 2005; Reisner et al., 2007; Wang et al., 2011), is less sensitive to backfolding com-

pared to the end-to-end distance projection (R_x) on the channel axis. Unsurprisingly, in the Odijk regime, X and R_x are almost identical except at very short lengths of the chain ($L < O(\lambda)$) due to the absence of backfolding. In contrast, for the large channels, we observe that X and R_x asymptotically approach a limiting value which in turn agrees with the predicted extension in the de Gennes regime, independent of the metric used to describe the extension (Tree et al., 2013a).

An interesting feature of Figure 4.9 is that X and R_x are equal in the long chain limit irrespective of the confinement regime. This result is a testament to the notion of linear ordering of blobs for self-avoiding chains in quasi-1D confinement, outside the Odijk regime. Correspondingly, the apparent confusion surrounding the earliest reports of simulations on DNA extension suggesting that $X \not\sim L$ seems be an effect of finite chain length as the longest chain used in that work was 14.8 kbp long (Cifra et al., 2009). Figure 4.9 also furnishes us with a convenient rubric to check if a chain is sufficiently long to be in the asymptotic region of the blob regime in which $X \sim L$ applies. This is especially important for conventional Metropolis Monte Carlo simulations, where each simulation gathers data for a single chain length. In order to ascertain whether the chain is long enough to be in the asymptotic region, one might have to perform multiple simulations at various lengths and verify that the properties fall on a universal curve. A relatively simple and convenient way to ensure the validity of asymptotic behavior ($X \sim L$) is to calculate X and R_x from the same simulation, and check if the two quantities are the same within the margin of error.

It is customary to presume that the span measured from fluorescence microscopy experiments is the same as the end-to-distance, which itself is based on the assumption that the chain has reached the asymptotic limit (Reisner et al., 2012; Tegenfeldt et al., 2004; Reisner et al., 2005). However, our findings suggest that X and R_x are not necessarily equal for λ-DNA, as there is a gradual deviation of R_x from X with increase in channel size. We find from Figure 4.9 that the relative error in fractional extension for λ-DNA from the asymptotic value is less than 3% for the channel sizes considered here, given that it is calculated from the mean span. In contrast, the error is as high as 20% for the largest channel, if the fractional extension is calculated from the x-projection of the end-to-end distance. This difference in the relative error of the fractional extension obtained from the two metrics is apparent when the data is plotted with a logarithmic axis as shown in Figure 4.9. These results demonstrate that although λ-DNA is not truly in the asymptotic limit, the fractional extension measured from the span is within

3% of the asymptotic value for channel sizes typically used in experiments.

Furthermore, the w/l_{p} ratio increases with decreasing ionic strength for DNA, indicating that the excluded volume becomes increasingly more important at low ionic strengths (Tree et al., 2013b). As most experiments are performed at ionic strengths lower than the one used in our simulations (Tegenfeldt et al., 2004; Reisner et al., 2005; Reisner et al., 2007), we expect excluded volume to have a stronger effect on extension data obtained from experiments. This implies that the onset of asymptotic scaling, $X \sim L$, occurs for chains with lower L compared to the ones dealt in our work. Therefore, we conclude that if the fractional extension is inferred from the measurement of the span, chains of the contour length of λ-DNA or above yield nearly asymptotic results for X/L.

4.4 Concluding remarks

By independently controlling the effects of excluded volume and stiffness, we studied the equilibrium properties of single polymer chains in channel confinement. Through our large scale PERM simulations, we verified that confinement does not have any effect on the extension of ideal freely jointed chains, provided that the confinement size is not comparable to the Kuhn length. Further, our extension and free energy calculations for self-avoiding flexible chains agree well with de Gennes' theory. We also demonstrated that as the confinement width increases, the extension of ideal stiff chains exhibits a broad transition regime between the Odijk regime and a stretch-free bulk-like regime.

For self-avoiding stiff chains, we confirmed that the extension is linear in the contour length in the long-chain limit regardless of the regime of confinement, albeit due to different underlying mechanisms. In the Odijk regime, this is due to the linear ordering of deflection segments whereas outside the Odijk regime the arrangement of nearly non-interpenetrating blobs in a channel leads to this linear scaling. The onset of this asymptotic scaling in extension in the Odijk and the de Gennes regimes correspond to the chain length at which the chain contains a few deflection segments or de Gennes blobs, as the case may be. However, the behavior within the transition regime is a function of the interplay between stiffness and excluded volume (Odijk, 2008).

Irrespective of the degree of confinement the projection of the end-to-distance on the axis and the mean span parallel to the axis of the channel were found to be equal in the long-chain limit. Although the chain length of λ-DNA is not sufficient to exactly

satisfy this condition for all the channel sizes considered here, we inferred that λ-DNA is long enough to exhibit nearly asymptotic value of the fractional extension. This is because the fractional extension measured from the mean span for λ-DNA in a high ionic strength buffer is within 3% of the asymptotic value for square channels of sizes as high as 400 nm. The systematic error in the fractional extension due to finite chain length of λ-DNA measured in experiments is therefore negligible especially considering various other kinds of errors introduced in quantities such as the exact geometry of the nanofabricated device, and the image processing required to extract the mean extension from fluorescence intensity data.

Chapter 5

Backfolding and hairpin formation of wormlike chains confined in square nanochannels

5.1 Introduction

* The problem of a wormlike chain confined in a channel has attracted significant attention since seminal experiments on DNA in nanochannels revealed a discrepancy between theory and experiment (Reisner et al., 2005). Subsequent research, driven primarily by simulation, has explained much of the discrepancy. For a wormlike chain of persistence length l_p and width w confined in a square channel of size D, there are now simulation data that support the existence of at least three regimes. The chain statistics in very tight channels, where $D \ll l_\mathrm{p}$, correspond to the classic Odijk regime (Odijk, 1983) of linearly ordered deflection segments. The prefactors for the chain statistics in the classic Odijk regime are known (Burkhardt et al., 2010) and in quantitative agreement with simulations (Tree et al., 2013a). When the confinement is relatively weak, confined chains are in the de Gennes regime and arrange themselves into a one-dimensional array of compression blobs (Daoud and de Gennes, 1977). While the original de Gennes model of isometric compression blobs only extends down to channels of size $D \approx l_\mathrm{p}^2/w$, there is now convincing evidence (Dai and Doyle, 2013; Dai et al., 2014; Werner and

*This chapter is based on A. Muralidhar, D. R. Tree, and K. D. Dorfman, "Backfolding of Wormlike Chains Confined in Nanochannels", Macromolecules **47**, 8446 (2014)

Mehlig, 2014) for a so-called "extended de Gennes" regime (Brochard-Wyart et al., 2005; Odijk, 2008; Wang et al., 2011) consisting of anisometric blobs that describes the chain statistics over the range $l_\mathrm{p} \lesssim D \lesssim l_\mathrm{p}^2/w$.

The key outstanding question for channel-confined wormlike chains, which we address here, concerns the chain statistics for channel sizes between $D/l_\mathrm{p} \ll 1$ and $D/l_\mathrm{p} \approx 1$. In the context of DNA in a nanochannel (Wang et al., 2011; Reisner et al., 2012), this range of channel sizes is often referred to as the "transition regime" despite the fact that simulations of DNA models in channel confinement (Cifra et al., 2009; Wang et al., 2011; Tree et al., 2013a; Dai et al., 2014) indicate that the transition from the extended de Gennes regime to the classic Odijk regime spans less than a decade in D/l_p. Indeed, since DNA is generally taken as the model polymer for studying channel-confined chains at the single molecule level, it is not even clear whether there exists a universal regime spanning channel sizes from $D/l_\mathrm{p} \ll 1$ to $D/l_\mathrm{p} \approx 1$ or just some gradual transition from the classic Odijk regime to the extended de Gennes regime. In this sense, the moniker "transition regime" splits the difference; these channel sizes either correspond to a transition or a regime, but probably not both. While the range of channel sizes spanning the gap between the classic Odijk regime and the extended de Gennes regime may be narrow, these channel sizes are highly relevant for practical purposes, in particular in genomics (Dorfman et al., 2013; Lam et al., 2012; Reisner et al., 2012), since they encompass fractional extensions from about 20% to 85% for DNA (Tree et al., 2013a).

There exist three different explanations for the transition regime: (i) an ideal blob (Gauss-de Gennes) model (Zhang et al., 2008; Tree et al., 2013a; Raphael and Pincus, 1992), which is inspired by simulation data (Cifra et al., 2009; Wang et al., 2011; Tree et al., 2013a) for DNA indicating that the mean extension in the transition regime scales like $\langle X \rangle \sim D^{-1}$, albeit over a narrow range of channel sizes; (ii) a cooperative backfolding model (Dai et al., 2012), which was inferred from parameterized simulation data for chains with the relatively small ratios of l_p/w that characterize DNA; and (iii) a theory by Odijk (Odijk, 2008) where the deflection segments can form hairpins with a characteristic length scale g, which Odijk calls the global persistence length (Odijk, 2006). Explanations (i-iii) are conveniently divided into two categories, based upon their fundamental explanation of the transition regime. Both the ideal blob and Odijk explanations posit a universal regime, in the sense that the behavior should persist over many decades in some relevant scaling parameter. By contrast, the backfolding model supposes that no such universal regime exists, and aims to describe a transition between the classic

Odijk and extended de Gennes regimes.

Of the arguments supporting a universal regime, Odijk's theory (Odijk, 2008) seems the most persuasive, since it merges Flory theory arguments that have been successful in explaining the blob regimes (Reisner et al., 2012) with the undisputed presence of deflection segments in very tight confinement. Moreover, Odijk's theory suggests that the relevant scaling parameter for backfolded chains is not $D/l_{\rm p}$, and thus permits the possibility of a universal regime spanning many decades in another scaling parameter, ξ, for sufficiently large ratios of $l_{\rm p}/w$, even though it only may span a narrow range in $D/l_{\rm p}$. Unfortunately, the closure of the Odijk scaling theory (Odijk, 2008) requires a model for the global persistence length that correctly accounts for the free energy costs for backfolding in the absence of excluded volume. Odijk (Odijk, 2006) provided a mechanical approximation for the global persistence length, which indicates that an enormous contour length is required to observe backfolding of deflection segments (Odijk, 2008; Odijk, 2006). However, Odijk was rather critical of his mechanical model of the global persistence length (Odijk, 2006), since it failed to account for fluctuation effects.

In addition to its aforementioned positive qualities, Odijk's global persistence length theory is particularly appealing in the light of recent work on strongly confined wormlike chains in slits (Tree et al., 2014). In agreement with a Flory theory by Odijk (Odijk, 2008), there is now numerical evidence that strongly confined chains perform a two-dimensional walk of deflection segments (Tree et al., 2014). This walk gives rise to a number of different "sub-regimes," depending on the chain contour length, persistence length and excluded volume strength. Similar to the way in which blob regimes in channels parallel those in slits, we expect to find multiple regimes of strongly confined wormlike chains in channels. Indeed, Odijk's theory for channel confinement (Odijk, 2006; Odijk, 2008) posits that a wormlike chain performs a one-dimensional walk of deflection segments with regimes that depend the chain contour length and excluded volume strength.

In the present contribution, we compute the global persistence length via off-lattice, Pruned-Enriched Rosenbluth Method (PERM) simulations (Grassberger, 1997; Prellberg and Krawczyk, 2004) of a discrete wormlike chain model (Wang and Gao, 2005), which we have used recently (Tree et al., 2013a; Muralidhar et al., 2014b; Gupta et al., 2014; Tree et al., 2014) to study very long wormlike chains in confinement. We find that Odijk's mechanical theory (Odijk, 2006) overestimates the global persistence length by several orders of magnitude over the range of channel sizes where chains of contour length $L \gg g$

are still accessible by simulation. By correcting Odijk's theory for thermal fluctuations, (Odijk, 2006) we provide the requisite closure to test Odijk's proposition of a regime of backfolded deflection segments (Odijk, 2008). We show that this regime indeed exists, as the scaling theory not only collapses the data for both the average chain extension and the variance about the mean extension for channel sizes $D \lesssim l_\mathrm{p}$, but provides a collapse of the data over more than a decade in terms of Odijk's scaling parameter ξ (see Eq. 5.15). These results suggest that Odijk's theory (Odijk, 2006; Odijk, 2008), as modified herein, is the correct description of the so-called "transition" regime for DNA in a nanochannel. Moreover, moving beyond the specific case of DNA, we confirm that the monomer anisotropy l_p/w plays a key role in demarcating the transition to the classic Odijk regime of linearly ordered deflection segments, with the width of the backfolded regime increasing with the monomer anisotropy. As a result, the backfolded Odijk regime is very narrow for almost all practical situations, most notably for DNA.

5.2 Theoretical Background

5.2.1 Scaling theory

Odijk (Odijk, 2008) classified channel confinement for $D \lesssim l_\mathrm{p}$ into two regimes, based on whether or not the chains should be able to backfold. The regime with no backfolding of deflection segments is the classic Odijk regime (Odijk, 1983), which has been studied and confirmed in various theoretical (Wang et al., 2011; Cifra, 2012; Dai et al., 2012; Wagner et al., 2007) and experimental studies (Kim et al., 2011; Nöding and Köster, 2012; Köster et al., 2007; Köster et al., 2005). The second regime, characterized by backfolded deflection segments, was cast by Odijk into a modified Flory theory where the polymer is treated as a renormalized "chain of deflection segments" confined in a channel.

To see clearly the connection between Odijk's theory for a confined chain and standard Flory theory, let us first recall the well established Flory theory results for a real chain in free solution (Rubinstein and Colby, 2003). The chain is described as a series of segments of width w with a typical bending length scale l_p. The persistence length l_p is defined independent of excluded volume effects, arising from the intrinsic stiffness of the polymer (Hsu et al., 2010b; Hsu et al., 2010a). The excluded volume strength can

be quantified by the z parameter,

$$z = \frac{N_{\mathrm{lp}}^2 v'_{\mathrm{ex}}}{R^3} = \frac{w}{l_{\mathrm{p}}} \left(\frac{L}{l_{\mathrm{p}}} \right)^{1/2} \tag{5.1}$$

which measures the total excluded volume of a "gas" of $N_{\mathrm{lp}} = L/l_{\mathrm{p}}$ monomers uniformly distributed in a volume R^3 when the coil obeys ideal chain statistics, $R^2 = Ll_{\mathrm{p}}$, with an excluded volume per contact of

$$v'_{\mathrm{ex}} = l_{\mathrm{p}}^2 w \tag{5.2}$$

The chain exhibits swollen coil behavior if $z > 1$ and Gaussian coil statistics if $z < 1$.

In free solution, the Flory free energy is the sum the entropic and excluded volume energetic contributions,

$$\frac{F}{k_{\mathrm{B}}T} = \frac{R^2}{Ll_{\mathrm{p}}} + \frac{N_{\mathrm{lp}}^2 v'_{\mathrm{ex}}}{R^3} \tag{5.3}$$

where R^2 is the square of the end-to-end distance of the chain, k_{B} is the Boltzmann constant and T is the absolute temperature. Recall that the first term corresponds to the entropic elasticity of an ideal chain and the second term is the z parameter evaluated at some value of R. As noted by de Gennes, the first term is inaccurate because of the assumption of ideality, while the second term is erroneous as it does not consider correlations between monomer positions (de Gennes, 1979). However, cancellation of errors in the Flory free energy gives rise to remarkably good scaling estimates of properties such as the average end-to-end distance (de Gennes, 1979; Rubinstein and Colby, 2003).

Odijk's theory (Odijk, 2008) for a backfolded wormlike chain confined in a channel of size D between $D \ll l_p$ and $D \approx l_{\mathrm{p}}$ is built on an analogy with the Flory theory in free solution (Eq. 5.3). However, instead of considering a chain of monomers of size w and persistence length l_{p}, Odijk considers a chain of deflection segments, where the renormalized monomers have a length $\lambda = D^{2/3} l_{\mathrm{p}}^{1/3}$ and a width w. The excluded volume per contact is

$$v_{\mathrm{ex}} \approx \lambda^2 w \, (D/l_{\mathrm{p}})^{1/3} \tag{5.4}$$

where the additional factor $\sin\delta = (D/l_{\mathrm{p}})^{1/3}$ accounts for the orientation of deflection segments, which are aligned at an average angle, δ, between each other (Odijk, 2008). The characteristic length scale for bending in this renormalized chain of deflection segments, which corresponds to the formation of a hairpin, is the global persistence length g. Like the persistence length in free solution, the global persistence length in confinement

is defined independent of excluded volume interactions.

For a confined chain of deflection segments, the analogy to the z parameter in Eq. 5.1 is (Odijk, 2008)

$$Z = \frac{N_\lambda^2 v_{\rm ex}}{(Lg)^{1/2} D^2} \tag{5.5}$$

where $N_\lambda \equiv L/\lambda$ is the number of deflection segments in a chain of length L and $R^2 = Lg$ is the size of an ideal chain of such deflection segments with global persistence length g. Accordingly, the denominator of Eq. 5.5 is the total volume available to such an ideal chain in a channel of size D. In the development of his theory, Odijk (Odijk, 2008) also defined a second parameter related to the excluded volume interactions for a single hairpin of length g,

$$\xi = \frac{n_\lambda^2 v_{\rm ex}}{gD^2} = \frac{gw}{D^{5/3} l_{\rm p}^{1/3}} \tag{5.6}$$

where $n_\lambda \simeq g/\lambda$ is the number of deflection segments in a single hairpin and gD^2 is the volume occupied by that hairpin.

Excluded volume thus plays a role at two different length scales in the Odijk theory. On one hand, the ability to backfold is controlled by ξ; if $\xi > 1$ then the volume gD^2 available to a hairpin is too small to contain the excluded volume $n_\lambda^2 v_{\rm ex}$ caused by the turn. On the other hand, if the chain can backfold, the swelling of the chain of deflection segments by excluded volume is controlled by Z; if $Z > 1$, then excluded volume amongst all of the deflection segments plays an important role. Odijk's theory (Odijk, 2008) thus consists of three possible cases: (i) $\xi > 1$, where the chains cannot backfold, (ii) $\xi < 1$ and $Z > 1$, where the chains backfold and experience excluded volume interactions, and (iii) $\xi < 1$ and $Z < 1$, where the chains backfold but do not experience significant excluded volume interactions. Let us examine these three cases in turn.

If $\xi > 1$, then the excluded volume caused by the backfolded deflection segments in a hairpin exceeds the volume available in the channel containing those deflection segments. Backfolding is thus suppressed and the deflection segments are linearly ordered (Odijk, 1983; Odijk, 2008). When excluded volume interactions are strong enough to prevent backfolding, the mean extension of the confined chain is close to its contour length L (Odijk, 1983; Burkhardt et al., 2010),

$$\langle X \rangle = L \left[1 - \alpha \left(\frac{D}{l_{\rm p}} \right)^{2/3} \right] \tag{5.7}$$

and the variance $\delta X^2 \equiv \langle (X - \langle X \rangle)^2 \rangle$ about this average extension is

$$\delta X^2 = \beta \frac{D^2}{l_\mathrm{p}} L. \tag{5.8}$$

Burkhardt et al. (2010) computed the prefactors for square channels to be $\alpha = 0.18274$ and $\beta = 0.00956$. In the above expressions, we have used the mean span

$$\langle X \rangle \equiv \langle x_\mathrm{max} - x_\mathrm{min} \rangle \tag{5.9}$$

as the measure of the extension of the chain, where x_max and x_min refer to the maximum and minimal axial positions respectively (Wang et al., 2011). For strongly stretched chains, or for asymptotically long chains in channel-confinement, the mean span and mean end-to-end distance are indistinguishable (Muralidhar et al., 2014b).

If $\xi < 1$, Odijk's modification of the Flory theory in Eq. 5.3 leads to (Odijk, 2008)

$$\frac{F}{k_\mathrm{B}T} \cong \frac{L}{\lambda} + \frac{X^2}{Lg} + \frac{N_\lambda^2 v_\mathrm{ex}}{XD^2}. \tag{5.10}$$

The first term in Eq. 5.10, representing the confinement of the deflection segments (Odijk, 1983), does not appear in the original theory (Odijk, 2008) because it does not contribute to the calculation of the average chain extension $\langle X \rangle$, even though it is the dominant term in the calculation of F. We include it here since (i) we know from previous simulation work (Tree et al., 2013a) that the confinement free energy is extensive in chain length and exhibits Odijk scaling $F \sim D^{-2/3}$ well past $D/l_\mathrm{p} \ll 1$ and (ii) the neglect of such terms in the Flory theories for the confinement free energy in blob regimes (Brochard-Wyart et al., 2005) led to some confusion about the existence of the extended de Gennes regime (Tree et al., 2013a; Dai et al., 2014).

Minimizing the free energy in Eq. 5.10 with respect to X leads to the average extension (Odijk, 2008)

$$\langle X \rangle \sim L\xi^{1/3}. \tag{5.11}$$

and the second derivative of the free energy produces the variance

$$\delta X^2 \sim gL, \tag{5.12}$$

The latter results are valid for $\xi < 1$ and $Z > 1$, i..e when the chain can backfold but

excluded volume interactions amongst all the backfolds swells the confined chain. For the case where $\xi < 1$ and $Z < 1$, the extension of the chain of deflection segments will exhibit ideal scaling (Odijk, 2006)

$$\langle X \rangle \sim \sqrt{gL}, \tag{5.13}$$

with a persistence length g. Accordingly, the variance for $\xi < 1$ and $Z < 1$ is the same as Eq. 5.12. The crossover contour length, L_*, between Eq. 5.11 and Eq. 5.13, is given by $Z = 1$, namely (Odijk, 2008)

$$L_* \cong g\xi^{-2/3}. \tag{5.14}$$

For the backfolded Odijk regime ($\xi < 1$), in addition to L_*, there is also a second crossover length $L_{**} = g$ where the chains are so short that they cannot even form a global persistence length. Interestingly, for $L < L_{**}$, so long as these chains have a contour length of at least $L > \lambda$, they exhibit the same extension and variance as the classic Odijk regime given by Eq. 5.7 and Eq. 5.8 respectively, independent of the value of ξ. The linear ordering of such short chains is driven by an insufficient number of segments to form a global hairpin, rather than excluded volume interactions suppressing backfolding. In addition, there should exist other sub-regimes for even shorter chains $L < \lambda$ corresponding to the wall-induced orientation of a rod, analogous to the case of a slit (Tree et al., 2014).

Figure 5.1 summarizes our discussion of Odijk's scaling theory (Odijk, 2008) so far. This figure emphasizes that the chain statistics depend on two factors: (i) the channel size relative to the persistence length, D/l_p, and (ii) the relative strength of excluded volume to the channel volume in a hairpin, ξ. Note that while both parameters are required to classify the various regimes, ξ and D/l_p are not independent variables. A rearrangement of Eq. 5.6 yields

$$\xi \simeq \left(\frac{g}{l_\mathrm{p}}\right)\left(\frac{l_\mathrm{p}}{D}\right)^{5/3}\left(\frac{w}{l_\mathrm{p}}\right), \tag{5.15}$$

which shows how ξ is related to other relevant dimensionless length scales. In particular, ξ has both an explicit dependence of $(D/l_\mathrm{p})^{-5/3}$ and an implicit dependence on D/l_p through the global persistence length. As the channel size increases and the chain exits the Odijk regimes, we expect that $g \to l_\mathrm{p}$ (Odijk, 2006). In other words, the global persistence length limits to the persistence length of the chain in the absence of

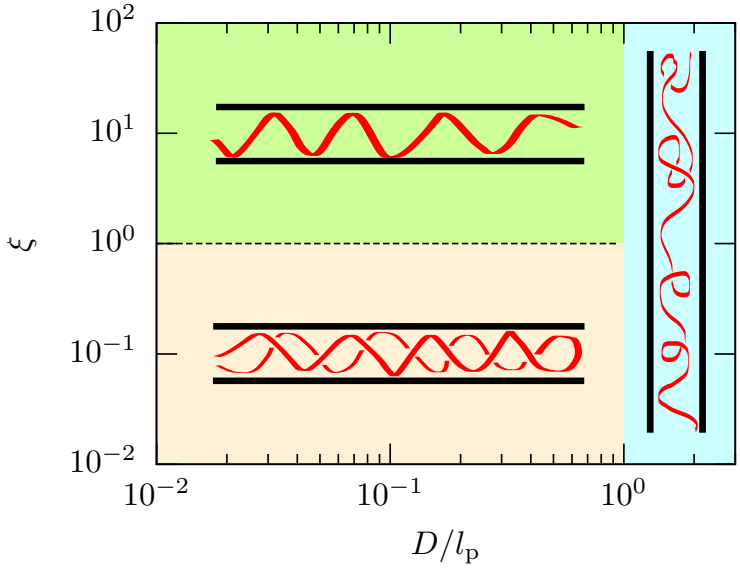


Figure 5.1. Phase diagram for confined wormlike chains based on the scaling theory of Odijk (Odijk, 2008). Channel sizes $D/l_\mathrm{p} \lesssim 1$ correspond to Odijk regimes, where $\xi \gtrsim 1$ is the classic Odijk regime (Odijk, 1983) and $\xi \lesssim 1$ is the backfolded Odijk regime. For wider channels with $D \gtrsim l_\mathrm{p}$, there is a gradual transition to blob regimes. This schematic is simplified to emphasize the importance of both D/l_p and ξ; note that ξ depends on D/l_p via Eq. 5.15.

confinement, which further emphasizes that both types of persistence lengths are defined independent of excluded volume interactions.

5.2.2 Odijk theory for the Global Persistence Length

The challenge in applying Odijk's theory (Odijk, 2008) is that it requires a model for the global persistence length g to close the system of equations. Odijk developed a theory for the statistical mechanics of hairpins for ideal wormlike chains confined in square channels, an outcome of which is an expression for g (Odijk, 2006). By determining the optimum chord length of the hairpin that minimizes the bending penalty for hairpin formation, and estimating the free energy of confinement of these hairpins, Odijk arrived at an expression for the global persistence length (Odijk, 2006),

$$g = \alpha \bar{r} \exp\left(\frac{\bar{F}}{k_\mathrm{B}T}\right). \tag{5.16}$$

The first term in Eq. 5.16 is the constant $\alpha = 3.3082$, which Odijk obtained by numerical integration from his model. The quantity $\bar{r}(D/l_\mathrm{p})$ is the average length of a hairpin chord,

$$\frac{\bar{r}}{l_\mathrm{p}} = \frac{1}{6}\left\{\left[E_m^2 + 6\sqrt{2}E_m\left(\frac{D}{l_\mathrm{p}}\right)\right]^{1/2} - E_m\right\}, \tag{5.17}$$

where a separate numerical integration produced $E_m = 1.5071$. Odijk's derivation corresponds to the mechanical limit, $\bar{F} = \bar{F}_\mathrm{mc}$, where $\bar{F}_\mathrm{mc}(D/l_p)$ has the form

$$\frac{\bar{F}_\mathrm{mc}}{k_\mathrm{B}T} = E_m\left(\frac{l_\mathrm{p}}{\bar{r}}\right) - 3\ln\left(\frac{D-\bar{r}\sqrt{2}}{D}\right) - \ln\left(\frac{8}{3\pi}\right) \tag{5.18}$$

and accounts for bending energy of the hairpin in addition to its translational and orientational entropy.

Odijk's mechanical approximation for the global persistence length has no adjustable parameters, which makes it a remarkable result. However, at the conclusion of his paper (Odijk, 2006), Odijk provides an enumerated list of shortcomings in this model. Thus, in order to test Odijk's scaling theory, we need to first validate (and, if necessary, modify) Odijk's mechanical model for the global persistence length.

5.3 Simulation Methodology

From a simulation standpoint, testing Odijk's theory is reasonably difficult. The challenge becomes clear when we consider Eq. 5.16 in the limit of a very small channel,

$$\frac{g}{l_\mathrm{p}} \simeq c_1\left(\frac{l_\mathrm{p}}{D}\right)^2 \exp\left[c_2\left(\frac{l_\mathrm{p}}{D}\right)\right] \quad \text{for } D/l_\mathrm{p} \to 0 \tag{5.19}$$

where c_1 and c_2 are $O(1)$ constants. Thus, in the limit of narrow channels, the contour length between hairpins diverges exponentially and hairpin formation becomes a rare event. From a practical standpoint, this exponential divergence makes the observation of hairpins in simulations increasingly difficult with decreasing channel size, necessitating simulation of extraordinarily long chains. Coarse-graining of the model to access these length scales, however, is limited by the ability of a coarse-grained model to resolve the sub-persistence length scale information characterizing deflection segments, which is critical in this extreme-confinement regime (Dai et al., 2014).

To address this challenge, we exploit the capability of the Pruned-Enriched Rosen-

bluth Method (PERM) (Grassberger, 1997; Prellberg and Krawczyk, 2004; Tree et al., 2013b) to simulate long wormlike chains in narrow square channels with a sufficiently fine-grained model, namely the discrete wormlike chain model (Wang and Gao, 2005). Our model consists of a discretized chain of beads connected by rods of length a. For real chains, we use $a = w$, and in the case of ideal chains, a is chosen such that $2l_p \geq a$. Particularly for simulations in the Odijk regime, we opt for $D/a \geq 5$ in order to resolve deflection segments (Muralidhar et al., 2014b). We have described our simulation approach in some detail elsewhere (Tree et al., 2013a; Tree et al., 2013b; Muralidhar et al., 2014b) and provide a brief description of the methodology in the Supporting Information.

5.4 Results

5.4.1 Measuring the Global Persistence Length

The global persistence length g is defined independent of any excluded volume, as it plays the one-dimensional analog to the three-dimensional persistence length l_p in free solution, and the effect of excluded volume only enters the scaling theory through the third term in Eq. 5.10. We thus only need to work for the moment with ideal wormlike chains ($w = 0$). In the context of the phase diagram in Figure 5.1, ideal wormlike chains are either linearly ordered (if $L < g$) or backfolded ($L > g$) since $\xi = 0$ for $w = 0$ by Eq. 5.6.

To measure g as a function of the confinement, D/l_p, we take advantage of qualitative similarities between the properties of a wormlike chain in channel confinement and the properties of single wormlike chains in the nematic phase of lyotropic liquid crystalline solutions, which has been studied extensively (Odijk, 2006; Odijk, 2008; Vroege and Odijk, 1988; Odijk, 1996; Spakowitz and Wang, 2003). Just as wormlike chains align parallel to the director axis in the nematic phase (de Gennes, 1982), wormlike chains in confinement align parallel to the axis of the channel. There is no polarity associated with the channel axis, i.e. the unit vectors $\hat{\mathbf{e}}_x$ and $-\hat{\mathbf{e}}_x$ are equivalent for a channel whose axis is parallel to the x axis, much the same as the director vector. Furthermore, wormlike chains in the nematic phase form hairpins that are separated by a length scale resembling the global persistence length (Vroege and Odijk, 1988; Spakowitz and Wang, 2003). Accordingly, it is convenient to define an orientational order parameter for the channel confined chains as the second Legendre polynomial of $\cos\theta$ (Spakowitz and Wang, 2003;

Wagner et al., 2007; Chen, 2013b),

$$m \equiv \langle P_2(\cos\theta)\rangle = \frac{1}{2}\left(3\left\langle\cos^2\theta\right\rangle - 1\right), \tag{5.20}$$

where θ is the angle between the channel axis and a tangent to the polymer molecule. The angle brackets in Eq. B.1 denote both an ensemble average over multiple chains and an average over the whole contour for long chains.

In the long-chain limit, the confined chain behaves like a one-dimensional ideal wormlike chain aligned along the x-axis with an effective persistence length of g, albeit with small-scale excursions perpendicular to the channel (Odijk, 2006; Odijk, 2008; Vroege and Odijk, 1988; Odijk, 1996). From the analogy between a wormlike chain in the nematic phase and a wormlike chain in channel confinement, the projection of the mean square end-to-end distance on the x-axis in the limit $L \gg g$ can be written as (Spakowitz and Wang, 2003; Tkachenko and Rabin, 1995)

$$R_x^2 = \frac{1}{3}(1+2m)\left[2gL - 2g^2\left(1-\exp\left(-L/g\right)\right)\right]. \tag{5.21}$$

The prefactor $(1+2m)/3 = \langle\cos^2\theta\rangle$ in Eq. C.1, though not considered by Odijk (2008) and Odijk (2006), accounts for the alignment of the chain along the channel axis. For a completely aligned chain ($m = 1$), Eq. C.1 reduces to the mean square end-to-end distance of a 1D wormlike chain with persistence length g. On the contrary, for an isotropic chain in free solution ($m = 0$), Eq. C.1 gives the expression for the projection of the end-to-end distance for a 3D wormlike chain with persistence length g. It should be noted that the scaling of R_x and X must be the same in the long-chain limit; we use R_x^2 here as it has the relatively simple relation to g embodied in Eq. C.1.

Having first established that our simulations give the expected result for the orientational order parameter (see Supporting Information), we calculated g by fitting our PERM simulation data for various values of l_p and D to Eq. C.1. Figure 5.2 shows the rescaled value of R_x^2 plotted as a function of L/g, for ideal wormlike chains confined in channels of sizes ranging from $D/l_\mathrm{p} = 0.5$ to $D/l_\mathrm{p} = 20$. In accordance with Eq. C.1, the data for all chains collapse onto a single curve when $L/g \gtrsim 10$, since this collapse is enforced in our fits (see Supporting Information). We also observe that the end-to-end distance in the unconfined direction approaches the curve for free solution ($m = 0$) as one increases the channel size, confirming that the size of ideal wormlike chains parallel to the axis of the channel is unaffected by confinement for $D \gg l_\mathrm{p}$ (Muralidhar

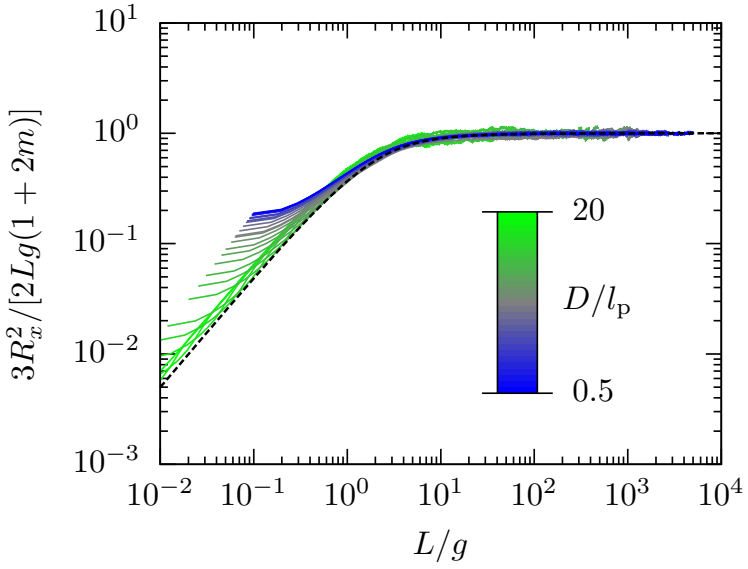


Figure 5.2. RMS end-to-end distance versus the contour length rescaled using the global persistence length for ideal wormlike chains with 10 beads per persistence length (l_p/a = 10). The colored curves correspond to simulation data for chains confined in 29 channels logarithmically spaced between D/l_p = 0.5 to D/l_p = 20. The black dashed line shows the result from Eq. C.1 for m = 0, which is the expected behavior in free solution.

et al., 2014b).

As we obtained the values for g for each value of D/l_p to produce Figure 5.2, we are now in a position to test Odijk's theory (Odijk, 2006) for the global persistence length. Figure 5.3 demonstrates that the global persistence length is indeed a universal quantity, as all our data fall on a single curve as a function of D/l_p, irrespective of the value of l_p. It is worth pointing out that g is almost 10 times l_p even for channels as big as $D = l_\mathrm{p}$. These results are consistent with previous observations of the increase in the apparent persistence length of wormlike chains in channel and tube confinement (Cifra et al., 2008; Cifra et al., 2010). Moreover, for $D \gg l_\mathrm{p}$, the global persistence length approaches the native persistence length of the molecule, suggesting that the length scale for backfolding approaches the persistence length in the de Gennes regime (Odijk, 2006).

Our simulation method for computing g has some advantages compared to Odijk's theoretical approach (Odijk, 2006) since we can reliably determine g in the wide-channel limit and thus obtain the right limiting value of $g = l_\mathrm{p}$. However, our method is much more limited in the small-channel limit. We cannot directly verify the behavior of g/l_p as $D/l_\mathrm{p} \to 0$, since our calculation of g relies on simulations of chains of contour length

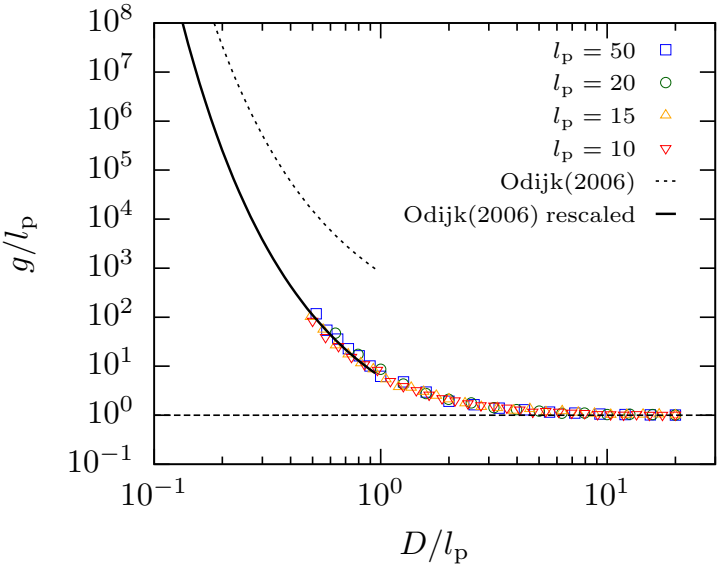


Figure 5.3. Collapse of global persistence length data for different values of l_p and D. The dotted line shows the result from Eq. 5.16. The solid line corresponds to Eq. 5.23. An extrapolation of Eq. 5.23 is also shown for $D/l_\mathrm{p} < 0.5$, where simulation data are absent. All values of l_p are in units of a.

of the order of several global persistence lengths. Because the value of g exponentially increases with decreasing D/l_p (Eq. 5.19), obtaining an estimate of g for $D/l_\mathrm{p} < 0.5$ becomes prohibitively expensive. This exponential growth poses an intrinsic problem for simulations of the global persistence length. For example, while other coarse-grained models such as the freely-jointed rod model (Dai and Doyle, 2013; Dai et al., 2014) have been very successful at modeling the extended de Gennes regime, they cannot be used to overcome this issue either, since these models lack the sub-persistence length information critical in resolving the deflection segments.

Figure 5.3 also compares our simulation data with Eq. 5.16 for $D < l_\mathrm{p}$, as this is roughly the range in which Odijk's theory (Odijk, 2006) should be valid. Although our simulation data do not agree quantitatively with his theory, we find a similar qualitative trend — in both Odijk's theory and our simulations, g/l_p rapidly increases as D/l_p decreases to values much lesser than unity. However, it appears from Figure 5.3 that Odijk's theory overestimates g by about two orders of magnitude in the range of our simulation data.

Odijk derived the expression in Eq. 5.16 in the mechanical limit, neglecting any

fluctuations of the hairpin. In addition, the approximations used by Odijk en route to Eq. 5.16 consider only the leading order terms in the perturbations to the shape of the hairpin about the lowest energy state, resulting in systematic errors in the estimation of g. Odijk recognized these potential shortcomings in his derivation (Odijk, 2008) and proposed that the free energy in the global persistence length of Eq. 5.16 should be of the form

$$\bar{F} = \bar{F}_{\mathrm{mc}} + H(D/l_{\mathrm{p}}) \tag{5.22}$$

where the second term corrects for the approximations in the mechanical limit (Odijk, 2006).

Although there is no obvious way to compute $H(D/l_{\mathrm{p}})$ theoretically (Odijk, 2006), we can easily estimate this quantity from the simulation results in Figure 5.3. For each channel size D/l_{p}, we use Eq. 5.22 as the free energy in Eq. 5.16 and solve for the correction term H. We find that H is a weak function of D/l_{p} for the channel sizes $D \leq l_{\mathrm{p}}$ where we could measure the global persistence length, and it is almost constant in the range of our simulations (see Supporting Information). The best fit constant to the data is $H/k_{\mathrm{B}}T = -4.91$, leading to the global persistence length

$$g = \alpha \bar{r} \exp\left(\frac{\bar{F}_{\mathrm{mc}}}{k_{\mathrm{B}}T} - 4.91\right). \tag{5.23}$$

As we see in Figure 5.3, this corrected form of the global persistence length for $0.5 < D/l_{\mathrm{p}} < 1$ is in good quantitative agreement with our simulation data, as would be expected from the relative insensitivity of H to the channel size.

From a practical standpoint, we can treat Eq. 5.23 as a useful functional form, valid over the range of D/l_{p} where we have obtained our data. In other words, Eq. 5.23 provides the closure to Odijk's Flory theory and thus allows us to evaluate the predictions of that theory independent of the theory for the global persistence length. However, we need to be cautious in extrapolating Eq. 5.23 for channels where $D/l_{\mathrm{p}} < 0.5$. Both Odijk's model in Eq. 5.16 and our modified version in Eq. 5.23 agree with the asymptotic result in Eq. 5.19, albeit with different prefactors. However, it seems reasonable that $H \to 0$ as $D/l_{\mathrm{p}} \to 0$, since extreme confinement would suppress the thermal fluctuations. This logic implies that H is indeed a function of D/l_{p} over some range of channel sizes. Thus, we would expect that the actual global persistence length for $D < l_{\mathrm{p}}/2$ would lie somewhere between the extrapolation of Eq. 5.23 in Figure 5.3 and the mechanical model.

Nonetheless, the quantitative difference between the value of g from our simulations and Odijk's mechanical theory seems to clarify some apparent discrepancies in the lit-

erature. Odijk used Eqs. 5.16-5.18 and an equation similar to Eq. C.1 to calculate the mean end-to-end distance of "ideal" wormlike chains with the same persistence length of DNA and showed that resultant values are in agreement with experimental data from Reisner et al. (2005). This agreement now appears like a fortuitous cancellation of errors, as Odijk's approach for estimating the extension of DNA did not account for excluded volume, while making use of extraordinarily high values for g. Furthermore, Wang et al. (2011) observed that the values for the extension of DNA in their simulations were much lower than what one would predict from the mechanical theory. Our results for g, which are two orders of magnitude less than Odijk's prediction, are consistent with this observation.

5.4.2 Test of Scaling Theory

Now that we have computed the global persistence length, we are in a position to evaluate Odijk's scaling theory for the backfolding regime. Because we are addressing the problem numerically, we can only evaluate his theory for cases where we can obtain g by the procedure outlined above. Fortunately, this does not limit our analysis of real chains ($w \neq 0$), since the computationally accessible contour lengths of real chains are a subset of those we can achieve with ideal chains. This is due to the fact that ideal chain simulations are cheaper than real chain simulations, since we do not need to make expensive excluded volume calculations. As a result, if we cannot simulate an ideal chain to sufficient length to measure g for a given channel size, then we also cannot simulate a real chain to sufficient length $L > L_*$ to reach the long chain limit, since $L_* \gg g$.

In what follows, we only present real chain data where we reach the long chain limit $\langle X \rangle \sim L$. Furthermore for the variance, we restrict ourselves to the subset of the latter data where we also reach the long-chain limit $\delta X^2 \sim L$ (see Supporting Information). The value of g for all real chains confined in channels of size D has been computed for equivalent ideal chains confined in channels with size D_{eff}, except for cases where Eq. 5.23 has been used as an approximation. Note that we do not need to test the confinement free energy, since we showed that $F \sim D^{-2/3}$ for these channel sizes in a previous publication (Tree et al., 2013a). This scaling of the confinement free energy corresponds to the confinement of ideal wormlike chains, which is accounted for by the L/λ term in Eq. 5.10. The magnitude of the excluded volume term is small compared to this term and thus does not contribute at the leading order to the scaling of the free energy (Tree et al., 2014).

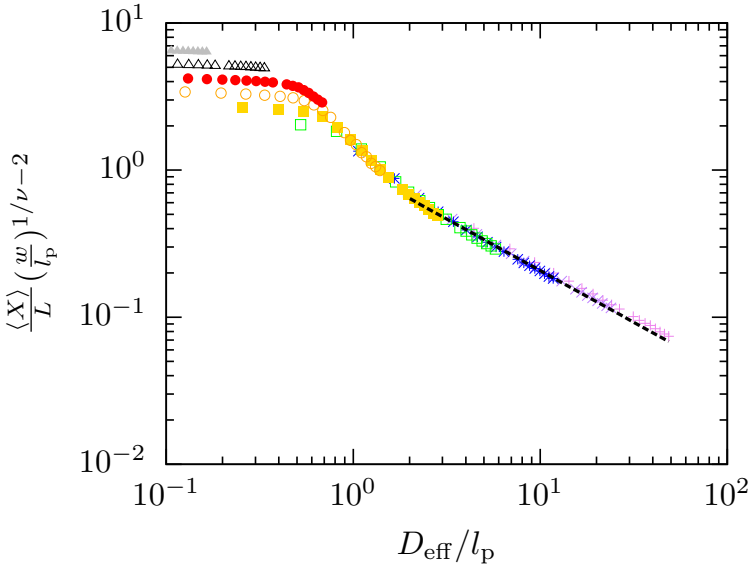


Figure 5.4. Collapse of extension data using the scaling for the blob regimes. The dashed line indicates the scaling $\langle X\rangle/L = 1.04278\,(D_\mathrm{eff}/l_\mathrm{p})^{-0.701}\,(w/l_\mathrm{p})^{0.298}$. The l_p/w values used here are 2.04 (violet +), 4.15 (purple ×), 8.45 (blue ∗), 17.20 (green □), 35.04 (gold ■), 71.35 (orange ○), 145.3 (red •), 295.5 (black △) and 602.7 (gray ▲).

We begin our test of the scaling theory with the extension. We first need to identify the blob-to-deflection segment transition in the phase diagram of Figure 5.1. To do this, we repeat the analysis by Dai et al. (2014) for our results with touching beads ($a = w$) and plot our extension data rescaled in the same way as Dai et al. (2014) in Figure 5.4. For real chains, the channel size $D_\mathrm{eff} = D - w$ represents the cross-sectional width available to the chain. For the extended de Gennes regime (Brochard-Wyart et al., 2005; Wang et al., 2011; Dai et al., 2014), we would expect extension to have the scaling $\langle X\rangle \sim D_\mathrm{eff}^{(\nu-1)/\nu}$, where $\nu = 0.587597(7)$ is the Flory exponent (Clisby, 2010). We observe that although the data seem to collapse all the way down to channel sizes $D \approx 0.8 l_\mathrm{p}$, there is a clear change in slope around $D_\mathrm{eff} \approx 2 l_\mathrm{p}$, indicating that the chains are losing their blob character and are are no longer in the extended de Gennes regime for smaller channels. This value, $D_\mathrm{eff} \approx 2 l_\mathrm{p}$, for the transition out of the extended de Gennes regime is in agreement with the freely-jointed rod model of Dai et al. (2014). These results show that the data are insensitive to the detailed model of the chain , confirming the universal nature of the de Gennes blob regime.

Having confirmed that there is a transition to the blob regime at $D/l_\mathrm{p} \approx 2$, we now

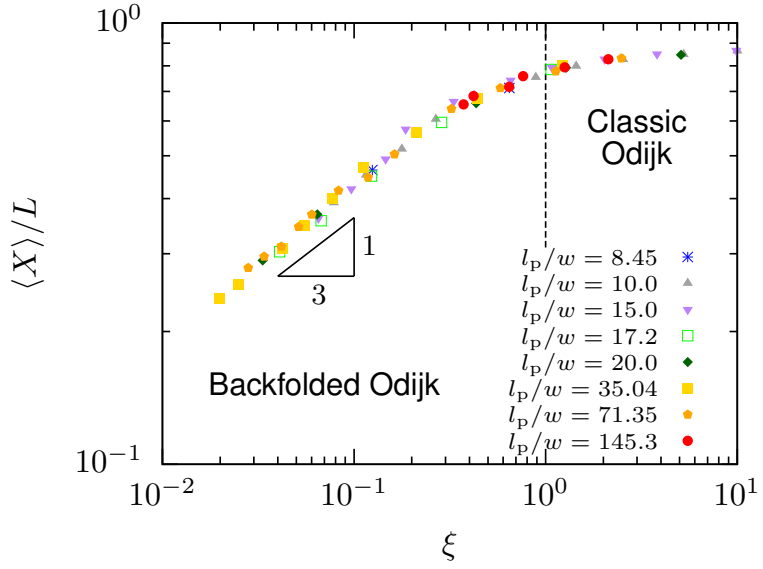


Figure 5.5. Plot of the fractional extension versus the parameter ξ for real wormlike chains with different monomer anisotropies. The vertical dashed line denotes the boundary between the two regimes according to scaling theory, $\xi = 1$. We calculated ξ from Eq. 5.6 by using g data obtained from simulations of ideal wormlike chains for the same $D_{\text{eff}}/l_{\text{p}}$ values as those of the corresponding real chains. Only data from channel sizes with $D_{\text{eff}} < 2l_{\text{p}}$ are shown. The backfolded Odijk regime encompasses fractional extensions from 20% to 80%, which is the typical range of the "transition" regime. A fit to the data for $\xi < 0.2$ reveals a relation given by $\langle X \rangle / L = 0.92\xi^{1/3}$.

turn our attention to examining Odijk's scaling theory for backfolded chains. Figure 5.5 shows that the Odijk backfolding model in Eq. 5.11 with computed global persistence length values from ideal wormlike chain simulations collapses all of the data for $\xi \ll 1$ and $D_{\text{eff}} < 2l_{\text{p}}$, independent of the particular values of l_{p}/w and D/l_{p}. As we would expect, the data do not exhibit the scaling $\langle X \rangle / L \sim \xi^{1/3}$ for $D_{\text{eff}} > 2l_{\text{p}}$, even when $\xi \ll 1$, since these channels correspond to the blob regimes (see Supporting Information). While the range of $D_{\text{eff}}/l_{\text{p}}$ for the backfolded Odijk regime in Figure 5.4 may be small, this regime spans many decades in ξ, thereby identifying ξ as the proper scaling variable. The lower bound on the value of ξ for the backfolded Odijk regime for a given l_{p}/w ratio can be estimated by using a channel size $D \approx 2l_{\text{p}}$ as the boundary. At this channel size, $g \approx 2l_{\text{p}}$ (from Figure 5.3), implying that the lower bound is $\xi \approx 0.63(w/l_{\text{p}})$ from Eq. 5.15. Note that the range of fractional extensions corresponding to the backfolded Odijk regime in Figure 5.5 is between 0.2 and 0.8, which is reminiscent of DNA extensions in the transition regime (Wang et al., 2011; Tree et al., 2012). This rapid increase in the

extension over a small range of channel sizes is due to the brisk increase in the scaling parameter ξ with decrease in D, as the extension is related to ξ by Eq. 5.11. The rise in ξ is in turn related to the exponential upsurge of g, since ξ is proportional to g (Eq. 5.6).

The results for the Odijk backfolding theory are even more impressive for the chain variance in Figure 5.6. For relatively large channels ($D_{\text{eff}}/l_{\text{p}} \approx 3$ to $D_{\text{eff}}/l_{\text{p}} \approx 10$), the variance is independent of channel size, as observed previously (Dai and Doyle, 2013) in studies of the extended de Gennes regime. Our results also confirm the expected scaling in the de Gennes regime, $\delta X^2 \sim D_{\text{eff}}^{1/3}$ (Reisner et al., 2012), for the largest channels. Once the chain passes into the backfolded Odijk regime, the fluctuations in the extension increase rapidly; note that the ordinate in Figure 5.6(a) spans nine orders of magnitude. As real chains cross over from the extended de Gennes regime to the backfolded Odijk regime, their rescaled variance $\delta X^2/Ll_{\text{p}}$ follows the $g/4l_{\text{p}}$ curve, as shown in Figure 5.6(b). This implies that

$$\delta X^2 \approx 0.25Lg \tag{5.24}$$

in accordance with Eq. 5.12. The collapse of the data onto Odijk's theory is even more apparent in Figure 5.7, where we use the scaling variable ξ rather than the channel size $D_{\text{eff}}/l_{\text{p}}$. Similar to what we saw for the extension in Figure 5.5, the data collapse onto the Odijk theory for $\xi < 0.1$, given that $D_{\text{eff}} < 2l_{\text{p}}$. The collapse of the data for the variance in Figure 5.7 looks sparser than that for the extension in Figure 5.5 for two reasons. First, it is difficult to reach the asymptotic limit for the chain variance (see Supporting Information). Second, the scaling of the variance in the backfolded Odijk regime given by Eq. 5.12 holds only for $\xi < 0.1$, while the scaling of the extension in Eq. 5.11 is valid up to $\xi \approx 0.4$. While this result is indeed a scaling law in terms of g, the global persistence length itself is not a power law in channel size D, so there is no simple power law of the form $\delta X^2 \sim D^x$. Note that the variance data in Figure 5.6(b) fall off the $g/4l_{\text{p}}$ curve at different values of $D_{\text{eff}}/l_{\text{p}}$ depending on the monomer anisotropy, l_{p}/w, suggesting that the width of the backfolded regime is a function of l_{p}/w. This dependence of the width of the backfolded Odijk regime on the monomer anistropy will be addressed in detail in Section 5.5.1.

Eventually, all real chains need to reach a channel size where $\xi > 1$. Since excluded volume terminates the backfolded Odijk regime, chains with higher monomer anisotropy remain in the backfolded Odijk regime down to smaller channel widths. The global persistence length increases almost exponentially as the channel size decreases,

so the maximum in the variance of the extension also increases rapidly with monomer anisotropy. This peak in the variance in chain extension, taken together with the freely draining hydrodynamics for these channel sizes (Tree et al., 2012), also explain the peak in the relaxation time for DNA in nanochannels around 100 nm (Reisner et al., 2005; Tree et al., 2013c). Similar to the chain extension, there is a broad transition for the δX^2 out of the backfolded Odijk regime and into the classic Odijk regime. As we see in Figure 5.6, all real chains eventually collapse onto Eq. 5.8 for sufficiently small channels. As was the case with the chain extension, the prefactor computed by Burkhardt et al. (2010) is in remarkable agreement with our simulation data.

The peak in the variance of the chain extension around $D \approx l_\mathrm{p}$ might lead one to think that it represents a discontinuity in the second derivative of the confinement free energy. This effect, coupled with the smoothness of the confinement free energy (Tree et al., 2013a) and the mean extension over the same range of channel sizes, may lead one to speculate further that the blob-to-deflection segment transition is a second-order phase transition. This supposition is supported by the difficulty in obtaining variance data in the long-chain limit (see Supporting Information). Indeed, the confinement free energy reaches the long chain limit very quickly (Tree et al., 2013a) compared to the variance in extension. However, the Odijk scaling theory makes it clear that no such phase transition exists at $D \approx l_\mathrm{p}$. The variance in the extension increases exponentially for small channel sizes through the scaling of Eqs. 5.12 and 5.19. However, the variance in the extension only diverges for ideal chains ($w = 0$), and this divergence only takes place in the limit of vanishing channel size. As a result, our current understanding of the Odijk theory and our simulation data suggest that there is no phase transition between blobs and deflection segments.

5.5 Discussion

5.5.1 Possible Experimental Tests of the Odijk Theory

From a practical standpoint, is it possible to verify the backfolded Odijk regime experimentally? To help answer this question, we computed the channel size corresponding to $\xi = 1$ for a number of experimentally relevant wormlike chains, from the fairly flexible example of double-stranded DNA to carbon nanotubes. In most cases, we expect the backfolded Odijk regime to extend below the channel sizes where we obtained data for g. In these instances, we have extrapolated our result in Eq. 5.23 to smaller channel

sizes, with the caveat that this extrapolation must fail at some point.

As we can see in Figure 5.8, the backfolded Odijk regime occupies a very narrow range in channel sizes due to the exponential increase in the global persistence length as the channel size decreases, even for extremely stiff molecules such as carbon nanotubes. The inset of Figure 5.8 highlights the narrowness of the backfolded Odijk regime; in order for the backfolded Odijk regime to span two decades in channel size, we would need to have a monomer anisotropy $l_p/w = O(10^{100})$. This is clearly well outside the range of any of the typical materials used to study channel confinement. Likewise, while DNA ($l_p/w = 10$) is considered a stiff molecule in many polymer physics situations, the l_p/w ratio of DNA in this context is so low that the backfolded Odijk regime is almost non-existent. The compression of the backfolded Odijk regime thus poses a significant problem for experiments if we use the channel size as the control parameter; we can explore a wide range of ξ values over a narrow range of channel sizes, but doing so in experiments requires an exquisite control over the channel size that is unlikely to be achieved (Reisner et al., 2005) for the persistence length of DNA. We can open up a wider range of channel sizes by increasing the persistence length, provided that wormlike chains with these larger persistence lengths are available with contour lengths such that $L \gg g$.

While it seems unlikely that the backfolded Odijk regime will be verified by experiments that involve changing the channel size, it may be possible to test Odijk's theory experimentally instead by fixing the channel size and changing the molecular weight of the confined chain. For example, consider the phase diagram for RecA-DNA in Figure 5.9. Relatively short RecA-coated DNA has recently been used to study confined chains in the Odijk regime (Frykholm et al., 2014). To construct this phase diagram, we assumed that the effective width of RecA-coated DNA is the same as naked DNA in the absence of a better estimate. For chains with $L < g$, we would expect to observe the classic Odijk regime ($\langle X \rangle \sim L$) because there are not a sufficient number of deflection segments to produce a global hairpin. This appears to be the case in experiments so far using RecA-DNA (Frykholm et al., 2014) (and actin filaments (Nöding and Köster, 2012)) based on the microscopy images. However, if we increase the contour length, such stiff molecules would eventually exhibit a regime with backfolding and weak excluded volume interactions (i.e., $g < L < L_*$), where $\langle X \rangle \sim L^{1/2}$. A further increase in contour length would lead to an onset of excluded volume interactions and a return to the scaling $\langle X \rangle \sim L$ for the backfolded Odijk regime. Counterintuitively, this test

of the Odijk regime is best done in relatively wide channels ($D \approx l_p$) rather than very small channels; as we see in Figure 5.9, the range of molecular weights with $\langle X \rangle \sim L^{1/2}$ widens substantially as the channel size increases. Indeed, with the estimates we used to construct Figure 5.9, it may be possible to observe the first transition using RecA-coated DNA at reasonable molecular weights. It seems unlikely that we would reach the transition back to $\langle X \rangle \sim L$ at L_*, since this would require RecA-coated DNA of several hundred thousand base pairs. The upside of such an experiment is that, in order to observe a reasonable range of molecular weights where $g < L < L_*$, the experiment needs to be performed in a channel close to the persistence length. Thus, the fabrication of the experimental system should be relatively easy, with the biochemistry being the limiting step.

5.5.2 Implications for Genomic Mapping

While the exponential increase in the global persistence length as the channel size decreases makes fundamental studies of the different Odijk regimes difficult, it is critical to recent successes using 45 nm channels for genomic mapping of DNA (Lam et al., 2012; Hastie et al., 2013). In nanochannel mapping (Jo et al., 2007; Das et al., 2010), DNA are labeled with sequence specific probes and stretched by injecting them into a channel. The physical distances between neighboring probes are measured using fluorescence microscopy, and these data can be converted to genomic distances from the fractional extension of the chain. When implemented in a massively parallel way (Lam et al., 2012), this method can be used to obtain information on structural variations in genomes that are difficult to assess by other techniques.

Clearly, the key to successful genomic mapping is to obtain linearly ordered probes; if the probes are disordered due to backfolding of the chain, this may be interpreted as a genomic reorganization rather than just a physical reorganization inside the channel. Based on Odijk's theory, there are two strategies to obtain linearly ordered deflection segments. In the first strategy, we can work with a chain in the Odijk regime $D \lesssim l_\mathrm{p}$ so long as the chain length satisfies $L \ll g$, even if $\xi < 1$. This approach will lead to ordered deflection segments and, in principle, can be implemented in circa 100 nm channels. Unfortunately, working with such small fragments of DNA makes genome assembly more difficult and leads to a large variance in chain extension (Su et al., 2011). The second strategy involves working with very long molecules, hundreds of kilobases in size, with the largest value of ξ achievable in experiments. It is obvious that forming

a global hairpin is a problem, since it leads to scrambling of the genomic information. However, since global hairpin formation is a stochastic process, it is possible to get linearly ordered genomic data from a single molecule that did not happen to form a global hairpin. While this appears to be a possible strategy, we know from Odjik's theory that there will still be large fluctuations in the distance between barcodes *in the absence* of a global hairpin. It is the second derivative of the Flory free energy in Eq. 5.10 that drives the fluctuations, which is affected by the statistical probability of forming a hairpin but does not require that such a hairpin actually form. This theoretical insight is supported by the evolution of the genome mapping technology, where large fluctuations (Su et al., 2011) motivated the switch from 100 nm channels (Das et al., 2010) to 45 nm channels (Lam et al., 2012) in the current commercial technology. Once the system is firmly entrenched in the classic Odijk regime, not only are the deflection segments linearly ordered but the variance in the extension also drops precipitiously. As can be seen from Figure 5.6(a) a small decrease in channel size going from the backfolded Odijk regime to the classic Odijk regime can result in a decrease in variance by 3 to 4 orders of magnitude.

Since DNA is a polyelectrolyte, its persistence length and effective width can be manipulated by changing the ionic strength of the buffer. Thus, in an experimental system, the extent of confinement can be controlled by either changing the channel size or changing the ionic strength of the system (Jo et al., 2007; Reisner et al., 2007; Kim et al., 2011). Figure 5.10 provides a phase diagram for DNA in channel confinement as a function of ionic strength and channel size. To construct this diagram, we used Dobrynin's theory (Dobrynin, 2006) for the persistence length of DNA (Hsieh et al., 2008) and Stigter's theory (Stigter, 1977) for the effective width. Our results indicate that double-stranded DNA rapidly reaches the classic Odijk regime and thus should not exhibit many hairpins in 45 nm channels at 100 mM salt, which is similar to the point where the genomic mapping technology operates. Hairpin formation could still occur due to kinetically trapped, frozen hairpin states as the DNA enters the nanochannel (Levy et al., 2008). Our work certainly did not account for such cases, as our results are restricted to equilibrium thermodynamics.

5.5.3 Implications for Theories of the Transition Regime

Our results strongly support Odijk's theory of two deflection segment regimes, one with linearly ordered deflection segments and a second with backfolding. In this sense, they

clarify the confusion in the literature surrounding the so-called "transition regime" for DNA in a nanochannel. Only by considering a wide range of monomer anisotropies, up to approximately fifteen times stiffer than DNA, using many beads per persistence length, were we able to observe the generality of Odijk's scaling theory in Figures 5.5 and 5.6. We thus recognize many of the prior generalizations related to the transition regime are extrapolations of artifacts arising from the modest monomer anisotropy of DNA and the limited chain lengths used in previous work.

With respect to the ideal blob (Gauss-de Gennes) model (Zhang et al., 2008; Tree et al., 2013a; Cifra and Bleha, 2012), there is indeed an apparent exponent $\langle X \rangle \sim D^{-1}$ at the start of the backfolded Odijk regime in Figure 5.4 for monomer anisotropies similar to DNA (Cifra et al., 2009; Wang et al., 2011; Tree et al., 2013a; Dai et al., 2014) using many different polymer models. Moreover, one of the striking results from Dai et al. (2014) is the collapse of the data in the transition regime using the ordinate and abscissa of Figure 5.4, which we initially thought might support a regime with the scaling

$$\frac{\langle X \rangle}{L} \sim \left(\frac{D}{l_\mathrm{p}}\right)^{x} \left(\frac{l_\mathrm{p}}{w}\right)^{1/\nu - 2}. \tag{5.25}$$

However, in order for Eq. 5.25 to be congruent with Odijk's theory with a scaling exponent $x = -1$, the global persistence length would need to follow the power law $g/l_\mathrm{p} \sim (l_\mathrm{p}/D)^{4/3}$. This is clearly not the case from the simulation data in Figure 5.3, even for channel sizes $D \approx l_\mathrm{p}$.

It is thus natural to wonder why the data collapse in both the blob regimes and the backfolded Odijk regime using the approach in Figure 5.4. First, note that the exponent in the ordinate of Figure 5.4 was chosen to ensure collapse in the blob regimes (Dai et al., 2014). If we use the the classical value of the Flory exponent, $\nu = 3/5$, then $1/\nu - 2 = -1/3$. Second, the mean extension in the backfolded Odijk regime (Eq. 5.11) can be rewritten as

$$\frac{\langle X \rangle}{L} \left(\frac{w}{l_\mathrm{p}}\right)^{-1/3} \simeq \left(\frac{g}{l_\mathrm{p}}\right)^{1/3} \left(\frac{l_\mathrm{p}}{D}\right)^{5/9} = f\left(\frac{D}{l_\mathrm{p}}\right), \tag{5.26}$$

where the left hand hand side is the ordinate of Figure 5.4 and f is some function of D/l_p. Thus, the collapse of the data for different l_p/w values in Figure 5.4 in the range $D_\mathrm{eff} \approx l_\mathrm{p}$ to $D_\mathrm{eff} \approx 2l_\mathrm{p}$ is a serendipitous consequence of the scaling, $\langle X \rangle \sim \xi^{1/3}$, in the backfolded Odijk regime rather than evidence of a scaling law for $\langle X \rangle$ in terms of D/l_p.

Unlike the ideal-blob model, the cooperative backfolding model of Dai et al. (2012) is not directly contradicted by the results presented herein. Furthermore, Odijk's theory clarifies the region in phase space where the cooperativity model is potentially applicable. In the cooperativity model, the extension of the chain proximate to the classic Odijk regime is described by conformations consisting of regions of S-loops, and regions of linearly ordered deflection segments (Dai et al., 2012). This premise is consistent with the physics of the transition at $\xi \approx 1$ in Odijk's theory (Odijk, 2008). For $\xi \ll 1$, the chain can have many hairpins present simultaneously at some channel location x. Indeed, this is the basis for the Z parameter in Eq. 5.5. As ξ increases, the ability to backfold decreases. As a result, the number of hairpins present at a cross-sectional slice, x, should decrease, with an S-loop being the limiting case when $\xi \approx 1$.

In the context of elucidating the universal behavior of a confined wormlike chain, modeling the transitions between regimes is generally not a priority. However, in the context of DNA in a nanochannel, Figure 5.8 shows that the backfolded Odijk regime is almost entirely suppressed and Figure 5.10 shows that the genomic mapping technologies lie near $\xi = 1$. Thus, there is a practical impetus to develop a model for the transition to the classic Odijk regime for DNA. Intriguingly, Figure 5.5 suggests that ξ may still be the relevant parameter for describing the extension even as ξ approaches unity. While the extension data in Figure 5.5 no longer follow the scaling law $\langle X/L \rangle \sim \xi^{1/3}$ as $\xi \to 1$, the data still collapse when plotted as a function of ξ well past $\xi = 1$. While the Flory theory in Eq. 5.10 is not the correct approach to describing the transition at $\xi = 1$, the physics embodied in ξ (and, more importantly, the global persistence length g) suggests the possibility of a more fundamental, predictive model of the transition to the classic Odijk regime at $\xi \approx 1$.

5.5.4 Comparison with Slit Confinement

It is worthwhile to briefly compare Odijk's theory (Odijk, 2008) for the extreme confinement limits, $D \lesssim l_\mathrm{p}$ to recent results for confinement in slits (Tree et al., 2014) of size $H \lesssim l_\mathrm{p}$. Similar to the case in channels, slit confinement also leads to different types of behavior in strong confinement depending on the chain length, channel size and excluded volume interactions. One of the advantages in describing slit confinement is that we know that the global persistence length in a slit is bounded between $g = 2l_\mathrm{p}$, i.e. the persistence length of a wormlike chain in two dimensions, and $g = l_\mathrm{p}$.

The situation for short chains ($\lambda \ll L \ll g$) is identical for slits and channels; the

walls orient the stiff chain, leading to an extension of the form of Eq. 5.7. The key difference is the upper bound in molecular weight; since the global persistence length grows rapidly in a channel as its cross-section decreases, the short-chain behavior persists up to much higher molecular weights in a channel than it does in a slit.

As the chain length increases beyond the global persistence length, both slit-confined and channel-confined chains undergo ideal random walks in the unconfined dimension(s) provided that the chain is thin enough so that excluded volume interactions are weak ($\xi < 1$). The random walk in the channel is one-dimensional, while the random walk in the slit is two-dimensional, but the scaling of the size of the chain with molecular weight in both cases is $L^{1/2}$. In both slits and channels, the ideality of the random walk arises because the chains can bend along the unconfined dimension(s) but they are too short to exhibit substantial excluded volume interactions. Naturally, in both slits and channels, this scaling persists to high molecular weights as $w \to 0$, since the chain can bend many times without experiencing substantial excluded volume interactions.

The key distinction between slits and channels takes place for long chains. In the slit case, there exists a single long chain regime (Tree et al., 2014). Here, chains with contour length $L \gg Hl_{\mathrm{p}}/w$ cross-over to a self-avoiding random walk in two dimensions, where the RMS end-to-end distance scales like $R \sim L^{3/4}$. In channels, Odijk's theory (Odijk, 2008) also predicts that chains with $L > L_*$ and $\xi < 1$ have the scaling for a self-avoiding random walk in one-dimension, namely $\langle X \rangle \sim L$. Likewise, chains with $\xi > 1$ always have this extensive scaling, independent of L, as long as $L > \lambda$. However, the dependence of $\langle X \rangle$ on the parameters w, l_{p}, and D depends on whether excluded volume interactions are strong ($\xi > 1$) or weak ($\xi < 1$).

5.6 Concluding Remarks

In the present contribution, we have used large scale PERM simulations of an ideal wormlike chain model to correct Odijk's theory for the global persistence length (Odijk, 2006) to account for fluctuation effects and other approximations down to channel sizes $D/l_{\mathrm{p}} = 0.5$. This result provided the closure for Odijk's scaling theory for confined wormlike chains (Odijk, 2008). The theory correctly predicts the confinement free energy (Tree et al., 2013a), average chain extension, and the variance in the extension, as well as the point at which excluded volume interactions suppress backfolding. We have tested his theory from $l_{\mathrm{p}}/w = 8.45$ to $l_{\mathrm{p}}/w = 145$ using a simulation model that resolves highly

anisotropic chains with many beads per persistence length.

The concordance between the results presented here and Odijk's theory implies that prior models of the so-called transition regime, (Zhang et al., 2008; Tree et al., 2013a; Dai et al., 2012) while in agreement with simulation data for a limited range of monomer anisotropies, are not the correct description for the full range of parameters. In fact, the sharp change in extension over a narrow range of channel sizes in the transition regime is a result of the rapid increase of the global persistence length, as one enters the backfolded Odijk regime from the extended de Gennes regime. Consequently, given the confusion disseminated by the term "transition regime", we propose abandoning it in favor of the label "backfolded Odijk" regime for the universal regime when $\xi \ll 1$ and reserve the term "transition" for crossovers between universal regimes.

Our results, when combined with recent work on the extended de Gennes regime (Dai et al., 2014; Werner and Mehlig, 2014), validate Odijk's description of channel-confined wormlike chains (Odijk, 2008) in Figure 5.11. While the classic Odijk regime (Odijk, 1983) and the de Gennes regime (Daoud and de Gennes, 1977) have been accepted dogma for several decades, they represent limiting cases that are difficult to access experimentally and thus not a good description of typical experimental data (Reisner et al., 2005; Persson et al., 2009; Werner et al., 2013; Gupta et al., 2014). Although the effect of excluded volume was not considered in the original theory of the classic Odijk regime (Odijk, 1983), our results indicate that excluded volume is indeed necessary to suppress backfolding of long chains in this regime. Most experimental systems either lie in the extended de Gennes regime or, for a small range of channel sizes (but a large range of $\langle X \rangle$ and ξ), the backfolded Odijk regime. Moreover, since the transitions between confinement regimes are not sharp and the molecular weights required to reach the long-chain limit are large, a substantial portion of the experimental data do not lie firmly in any well defined regime.

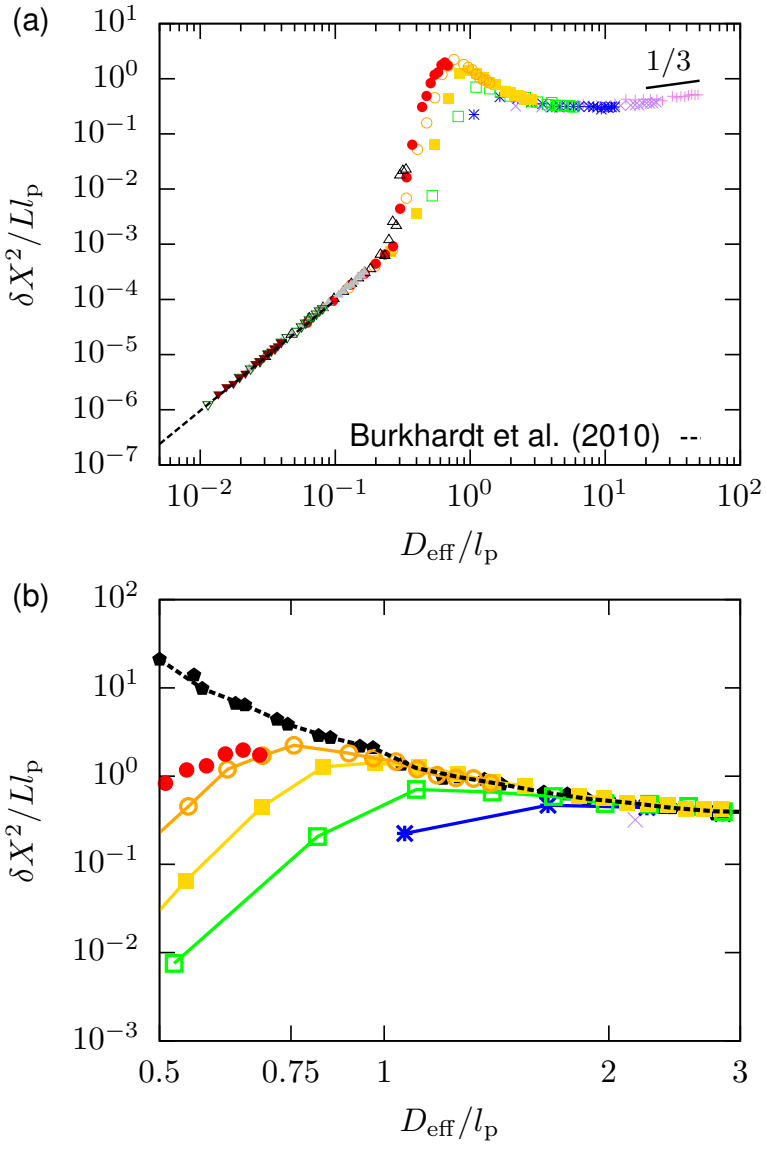


Figure 5.6. (a) Normalized variance in the chain extension as a function of channel size. The l_p/w values 2.04 (violet +), 4.15 (purple ×), 8.45 (blue ∗), 17.20 (green □), 35.04 (gold ■), 71.35 (orange ○), 145.3 (red •), 295.5 (black △) and 602.7 (gray ▲), 1227.54 (dark-green ▽), 2500 (brown ▼). The dashed line is Eq. 5.8 for the classic Odijk regime. The solid line shows the slope $\delta X^2 \sim D_{eff}^{1/3}$ expected in the de Gennes regime. (b) The region in part (a) where the normalized variance appears to rise up with decreasing channel size. The solid black pentagons are simulation data for $g/4l_p$ for ideal chains of $l_p = 15$ and $l_p = 10$. The black dashed line, which joins the simulation data for $g/4l_p$, and the colored solid lines for the variance data are meant as a guide to the eye.

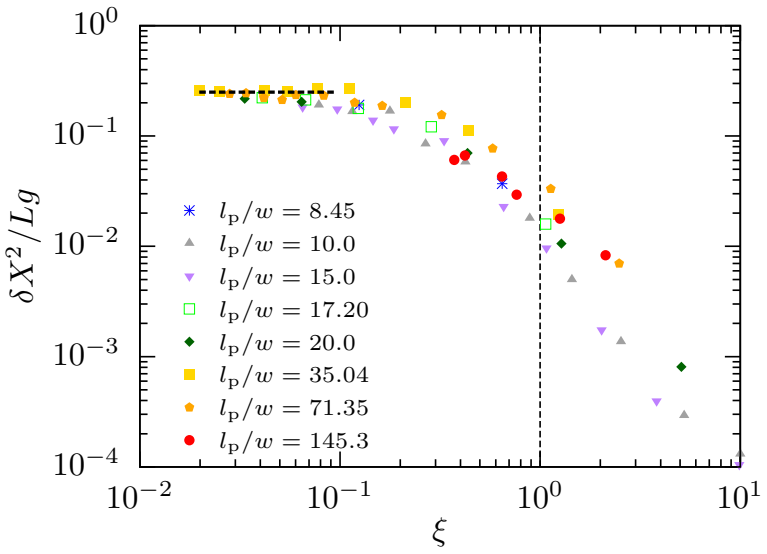


Figure 5.7. Collapse of the variance in the chain extension to Eq. 5.12 for $\xi < 0.1$ and $D_{\text{eff}} < 2l_p$. The horizontal black dashed line denotes $\delta X^2/Lg = 0.25$ from Eq. 5.24 and the vertical dashed line shows the boundary between the backfolded and classic Odijk regimes according to scaling theory ($\xi = 1$).

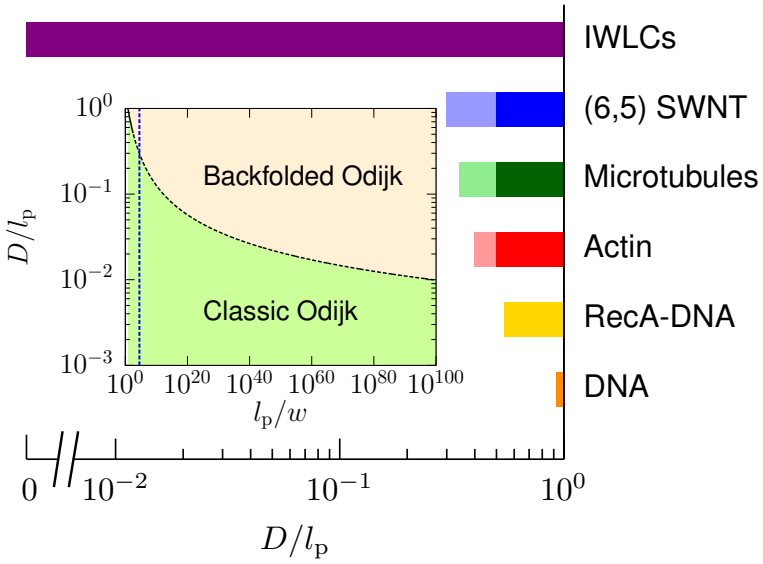


Figure 5.8. The range of the backfolded Odijk regime for DNA in a high ionic strength buffer (l_p/w = 10) (Tree et al., 2013a), RecA-DNA complex (l_p/w = 230 assuming w = 5 nm) (Frykholm et al., 2014), actin (l_p/w = 2375) (Köster et al., 2005; Köster et al., 2007), microtubules (l_p/w =9600) (van den Heuvel et al., 2007; Köster et al., 2005) and (6,5) single-walled carbon nanotubes (SWNT, l_p/w = 34210) (Fakhri et al., 2010). For ideal wormlike chains, l_p/w = ∞ and the backfolded Odijk regime extends all the way down to D/l_p = 0. Darker shade indicates the region in which we have simulation data for g, while lighter shade signifies the region wherein Eq. 5.23 has been extrapolated to obtain the value of g. The inset shows the range of the Odijk regimes as a function of the value of monomer aspect ratio, l_p/w. The vertical blue line corresponds to the (6,5) SWNT, the wormlike macromolecule with the highest l_p/w value considered here. We set the upper limit of the backfolded Odijk regime to the scaling theory value of D/l_p = 1 for simplicity, although this limit is closer to $D/l_\mathrm{p} \approx 2$.

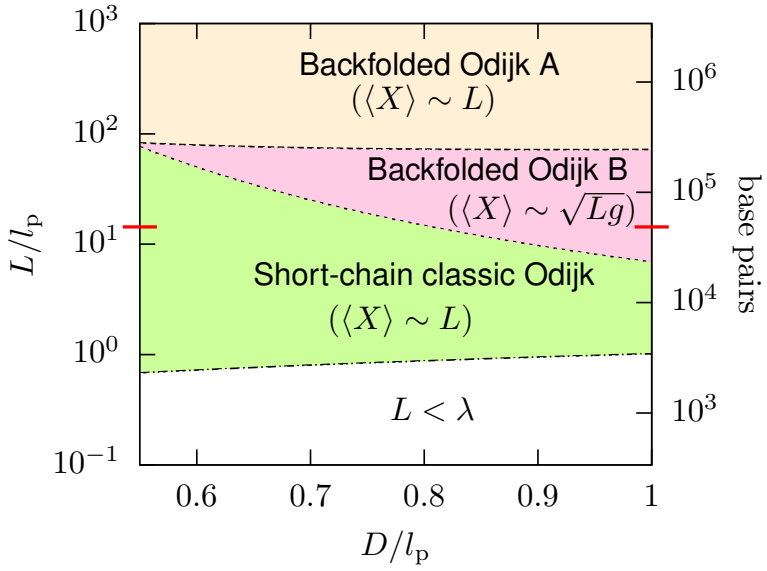


Figure 5.9. Scaling of the end-to-end distance as a function of contour length of RecA-DNA (l_p/w = 210) (Frykholm et al., 2014). The red lines on the y-axes correspond to λ-DNA, which is most commonly used in experiments.

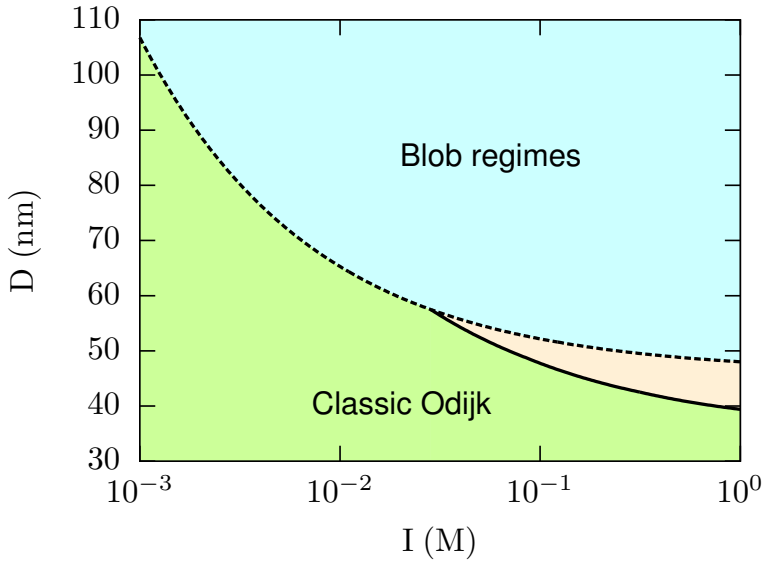


Figure 5.10. Phase diagram for DNA confinement in a nanochannel as a function of channel size and the ionic strength of the buffer. The onset of the blob regimes is denoted as $D = l_\mathrm{p}$ and the transition between Odijk regimes is at $\xi = 1$ using Eq. 5.23. The beige-shaded region is the backfolded Odijk regime.

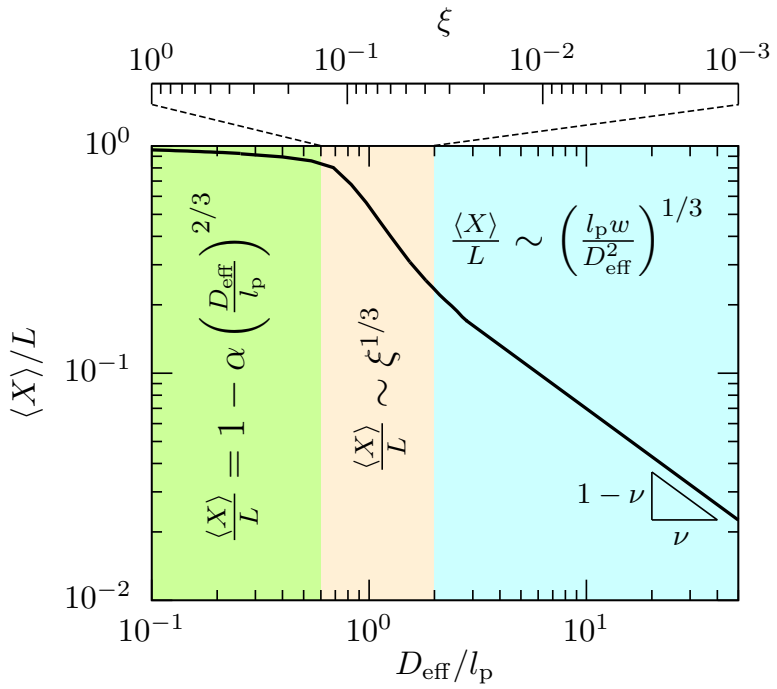


Figure 5.11. Summary of the regimes of confinement for a wormlike chain in a nanochannel and the corresponding fractional extension. The blue-shaded region corresponds to the de Gennes regime and the extended de Gennes regime, as they exhibit identical scaling of the extension. Note that the lower bound in ξ for the backfolded Odijk regime is a function of $w/l_{\rm p}$ and is approximately given by $0.63(w/l_{\rm p})$. However, we used the lower bound value of 10^{-3} only to illustrate that this regime can span many decades in ξ.

Chapter 6

Backfolding of DNA confined in nanotubes: Flory theory versus the two-state cooperativity model

6.1 Introduction

* The thermodynamics of DNA confined in a circular tube has received significant attention over the past decade due to its importance in genome mapping and semiflexible polymer physics (Reisner et al., 2012; Dorfman et al., 2013). The problem involves two important length scales, the tube diameter, D, and the persistence length of the confined chain, l_{p}. The description of the behavior of the chain is reasonably clear when there is a broad separation between these length scales. On the one hand, if $D \ll l_{\mathrm{p}}$, Odijk (1983) showed that chain is a linear arrangement of deflection segments. On the other hand, if $D \gg l_{\mathrm{p}}$, the conformation of the confined chain can be explained by de Gennes' blob theory (Daoud and de Gennes, 1977) and modifications thereof (Brochard-Wyart et al., 2005; Wang et al., 2011; Reisner et al., 2012; Dai et al., 2014; Werner and Mehlig, 2014). However, the situation is considerably more complicated and somewhat controversial when the two length scales are comparable, i.e. $D \approx l_{\mathrm{p}}$. In order to elucidate the extension of DNA and other semiflexible polymers in this region of the phase space, analogies have been drawn with other problems in polymer physics such as ne-

*This chapter is based on A. Muralidhar and K. D. Dorfman, "Backfolding of DNA Confined in Nanotubes: Flory Theory versus the Two-State Cooperativity Model", Macromolecules (in press) (2016) 10.1021/acs.macromol.5b02556)

matically confined stiff chains, most notably by Odijk (Odijk, 2006; Odijk, 2008), and the Zimm-Bragg model (Zimm and Bragg, 1959; Grosberg and Khokhlov, 1994) for helix-coil transitions in proteins by Dai et al. (2012). In the present contribution, we use simulations of confined wormlike chains to assess the ability of these theories to capture the extension of the chain as a function of contour length.

Odijk proposed that a chain confined in a tube with $D \approx l_p$ can be thought of as a one-dimensional chain of linearly ordered deflection segments with occasional hairpin turns (Odijk, 2006). The hairpin turns can be viewed as defects in an otherwise ordered arrangement of deflection segments, similar to the hairpins in a nematically ordered solution of stiff polymers (Spakowitz and Wang, 2003). He then calculated the frequency of backfolding by estimating the bending energy penalty for hairpin formation and the associated entropy gain for confined ideal chains. Accordingly, Odijk (2006) defined a global persistence length, g, that measures the average contour length for backfolding, as illustrated in Figure 6.1(a). Odijk then used the latter length scale to predict the average extension of a confined chain with excluded volume through a one-dimensional Flory theory (Odijk, 2008). We recently showed strong evidence supporting Odijk's theory for long chains in square channels and consequently termed this regime the backfolded Odijk regime (Muralidhar et al., 2014a).

Alternately, Dai et al. (2012) suggested that the extension of DNA confined in tubes of diameter close to the persistence length is governed by the nucleation and propagation of S-shaped loops. They posited that the chain consists of two states: deflection segments and S-loops, as shown in Figure 6.1(b). By dividing the chain into units of length πD, they put forward a cooperativity model based on the Zimm-Bragg model for helix-coil transition (Zimm and Bragg, 1959). Inherent in their model is the assumption that the S-loop and deflection units promote units of the same kind in their neighborhood. They then derived an expression for the extension of the chain by solving a model along the lines of the Zimm-Bragg model (Zimm and Bragg, 1959). After correcting their prediction of extension for the existence of C-loops at the two ends of the chain (see Figure 6.2), they showed consistency between their simulations and the model for chains with a finite contour length (Dai et al., 2012).

Given that there is no backfolding in the classic Odijk regime and the chain is coiled up in the blob regimes, it is natural to expect backfolding of deflection segments as the channel size increases and the confined chain transitions out of the classic Odijk regime (Cifra and Bleha, 2012). Both the approaches of Odijk (Odijk, 2008; Muralidhar

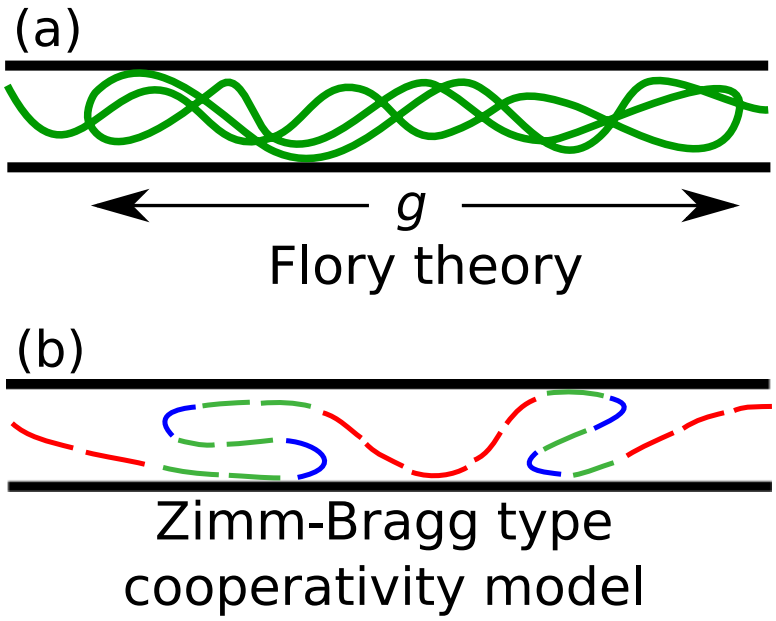


Figure 6.1. (a) Illustration of backfolding in Odijk's theory. The confined chain that forms deflection segments by collisions with the walls can also form hairpins. The average separation between hairpin turns is g. A one-dimensional Flory theory was applied to estimate the average extension (Odijk, 2008). (b) The two-state model divides the confined chain into deflection segment (red) and S-loop (green) states. Hairpin turns are in blue. Each dashed line along the chain contour contains a contour length πD. Note that the units containing contour length πD are not drawn to scale. A Zimm-Bragg type cooperativity model (Zimm and Bragg, 1959) was applied to predict the extension (Dai et al., 2012).

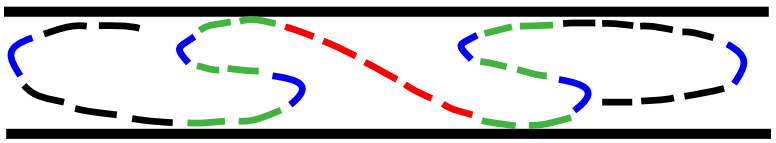


Figure 6.2. Illustration of C-loop units (in black) at the two ends of the chain (the units of contour length πD are not drawn to scale). Note the S-loop and deflection segment units in the interior of the chain.

et al., 2014a) and Dai et al. (2012) are consistent with this notion of backfolding of deflection segments. However, the two theories are fundamentally different. Odijk's Flory theory is based on the balance of entropy of the chain and excluded volume interactions between deflection segments, and it gives rise to a scaling law for the extension of long chains in terms of a new scaling variable, ξ, which is a function of D, l_{p}, g and the chain width, w (Odijk, 2008). In contrast, the model of Dai et al. (2012) hinges on the assumption of nucleation and propagation of S-loops and depends on fitting parameters that are obtained by simulations. Their model is not a scaling law but rather proposes that backfolding introduces a perturbation to the extension in the classic Odijk regime. Moreover, the differences in the two approaches are not restricted to predictions for long chains. While Odijk's theory predicts a gradual crossover from a rod to a random walk to a self-avoiding walk of deflection segments as a function of contour length, the cooperativity model considers C-loops to describe the finite-length behavior of the confined chain.

Here, we reconcile the two different threads in the literature by analyzing simulation data in the context of the two theories. We achieve this by performing Pruned-Enriched Rosenbluth Method (PERM) simulations of a discrete wormlike chain model using exactly the same parameters as the Metropolis Monte Carlo simulations of Dai et al. (2012), but now as a function of molecular weight rather than for a fixed chain size (Dai et al., 2012). We chose $l_{\mathrm{p}}/w = 10$ for much of our analysis, which also conveniently corresponds to DNA in a high ionic strength buffer (Wang et al., 2011). We observe that Odijk's scaling theory not only agrees with our data in the long-chain limit but also predicts the contour length dependence of the extension in the backfolded Odijk regime. In contrast, we find that the cooperativity model furnishes the correct extension only for the contour lengths around the region where the model was parameterized.

6.2 Methods

In order to compare the aforementioned theories, we carried out PERM simulations of long semiflexible chains (> 10000 beads) confined in circular tubes for a subset of the parameter space explored by Dai et al. (2012), who used Metropolis Monte Carlo simulations for chains up to a length of 1601 beads. To this end, we employed the discrete wormlike chain model that we have used successfully before (Muralidhar and Dorfman, 2015). We have explained our PERM methodology at length elsewhere and we

refer the reader to our previous publications for more details (Tree et al., 2014; Muralidhar et al., 2014b; Muralidhar et al., 2014a). In the case of chains with finite width, we use an effective tube size, $D_{\text{eff}} = D - w$, for the actual cross section of the tube available to the chain (Muralidhar et al., 2014a). Here, we performed simulations for 11 tube sizes equally spaced between $D_{\text{eff}}/l_{\text{p}} = 1$ and $D_{\text{eff}}/l_{\text{p}} = 2$, for a fixed monomer anisotropy ratio $l_{\text{p}}/w = 10$, which is exactly the range of tube sizes considered by Dai et al. (2012). Note that the range of validity of the cooperativity model is also $1 \leq D_{\text{eff}}/l_{\text{p}} \leq 2$ (Dai et al., 2012). Each simulation for a given tube size, D, consists of a total of 2.4 million tours divided into 6 independent sets of 400,000 tours each to estimate sampling errors.

We use two metrics for the average extension of the chain— (i) the projection of the end-to-end distance on the axis of the tube (x-axis), R_x, and (ii) the mean span of the chain along the axis, X. In our simulations of chains with N beads, the mean square end to end distance R_x^2 is calculated as

$$R_x^2 \equiv \left\langle (x_N - x_1)^2 \right\rangle, \tag{6.1}$$

where x_i refers to the x coordinate of the ith bead. The mean span, X, is obtained as

$$X \equiv \left\langle \max_i (x_i) - \min_i (x_i) \right\rangle, \tag{6.2}$$

where the bead index i is an integer such that $i \in [1, N]$.

Comparison of our simulation data with Odijk's Flory theory requires calculating the global persistence length, g, as a function of channel size (Odijk, 2008; Muralidhar et al., 2014a). Recently, we showed that the global persistence length can be calculated by measuring the x-projection of the end-to-end distance of an ideal confined chain as a function of contour length by simulation and then extracting g as a fitting parameter from the relation (Muralidhar et al., 2014a; Spakowitz and Wang, 2003)

$$R_x^2 = \frac{1}{3}(1 + 2m)[2gL - 2g^2\{1 - \exp(-L/g)\}]. \tag{6.3}$$

The orientational order parameter, $m = \left(3\langle \cos^2\theta \rangle - 1\right)/2$, where θ is the angle between the tangent to the contour of the chain and the tube axis, can also be measured by the same simulation by computing the angle between bond vectors. Note that the global persistence length is defined for ideal chains, analogous to how the persistence length is defined as the length scale of bending for ideal unconfined chains (Hiemenz and

Lodge, 2007). Accordingly, the global persistence length corresponding to a real chain of persistence length l_p confined in a tube of diameter D is computed from simulations on an equivalent ideal chain confined in a tube with diameter $D_{eff} = D - w$. To calculate the global persistence length, we performed PERM simulations of ideal chains confined in tubes with $1 \leq D_{eff}/l_p \leq 2$ for a total of 1.2 million tours in 6 sets of 200,000 tours each. We then computed g by fitting our simulation data for the x-projection of the end-to-end distance to Eq. 6.3 (see Supporting Information). Note that our previous work (Muralidhar et al., 2014a) was restricted to simulations in square channels where the hairpins are expected to form predominantly along the diagonals of the channel (Odijk, 2006). Although we expect the qualitative nature of the dependence of g on D_{eff}/l_p to remain the same in a circular tube, the different entropic effects in circular tubes (Odijk, 2006) lead to the quantitatively different behavior shown in Supporting Information.

6.3 Results and Discussion

6.3.1 Comparison with Odijk's scaling theory

We begin our comparison with Odijk's theory by examining the dependence of the extension of confined real chains as a function of molecular weight. For short-chain properties, it is convenient to use R_x as a metric to ascertain the molecular weight dependence of the extension. This is because an analytical expression is available for the end-to-end distance of ideal chains in the backfolded Odijk regime (Eq. 6.3), which proves to be a useful tool for comparison. In addition, the end-to-end distance exhibits a more pronounced ideal chain regime for relatively short chains as opposed to the span (Tree et al., 2013b). When we consider the asymptotic properties later, we will use the span as this is the most experimentally relevant quantity (Wang et al., 2011).

Figure 6.3(a) shows the dependence of the square of the x-projection of the end-to-end distance, R_x^2, of a real chain with $l_p/w = 10$ confined in tubes with diameter, $l_p \leq D_{eff} \leq 2l_p$. The x-axis is the contour length normalized by the global persistence length, g. Odijk's theory predicts three sub-regimes as a function of contour length. If $L < g$, hairpin formation is extremely rare and we expect the essentially linear chain to exhibit a scaling of $R_x \sim L$ resembling a rod. As expected, all curves in Figure 6.3(a) exhibit this rod-like regime for $L < g$. If $L > g$, it is entropically more favorable for the chain to form hairpins. Thus the chain can form a one-dimensional random walk

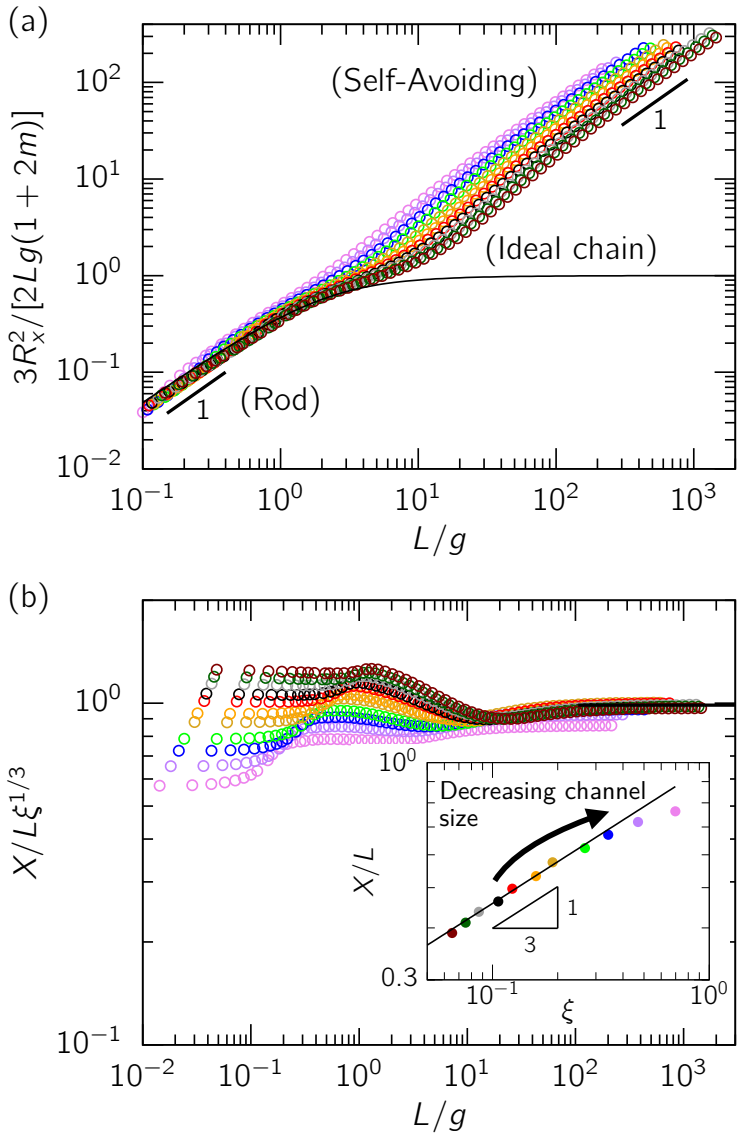


Figure 6.3. (a) Square of the x-projection of the end-to-end distance against the contour length normalized by the global persistence length. On the y-axis, R_x^2 is normalized using the prefactor on the right hand side of Eq. 6.3. The black curve shows the result for an ideal chain corresponding to Eq. 6.3 with $m = 0$. The slope of unity corresponds to the rod and self-avoiding limit respectively. (b) The mean span rescaled using Eq. 6.6 versus the contour length. The results are for a chain with $l_p/w = 10$ confined in a tube with following D_{eff}/l_p values: 1 (violet), 1.1 (purple), 1.2 (blue), 1.3 (green), 1.4 (gold), 1.5 (orange), 1.6 (red), 1.7 (black), 1.8 (gray), 1.9 (dark green), 2.0 (brown). The inset shows the asymptotic fractional extension against ξ. The colors denoting the different tube sizes are consistent in both (a) and (b) including the inset. The black line in the inset is $X/L = 0.99\xi^{1/3}$.

with g being the characteristic length scale for hairpin turns. In this sub-regime, we expect $R_x \sim L^{1/2}$. However, the width of this ideal chain sub-regime depends on the strength of excluded volume interactions between deflection segments. As can be seen from Figure 6.3(a), for ideal chains this sub-regime extends to $L \to \infty$.

In the case of real chains, the extent of the ideal regime is determined by the Z parameter measuring the strength of the excluded volume interactions between deflection segments for the whole chain (Odijk, 2008; Muralidhar et al., 2014a). It is defined as

$$Z = \frac{N_\lambda^2 v_{\text{ex}}}{(Lg)^{1/2} D^2}, \tag{6.4}$$

where $\lambda \simeq D^{2/3} l_{\text{p}}^{1/3}$ is the deflection segment length, $v_{\text{ex}} \simeq \lambda^2 w (D/l_{\text{p}})^{1/3}$ is the excluded volume between nearby deflection segments accounting for the orientation between them, and $N_\lambda \simeq L/\lambda$ is the total number of deflection segments in the chain (Odijk, 2008). At a critical contour length L^*, the behavior of the chain crosses over from that of an ideal random walk to a one-dimensional self-avoiding walk. This crossover is given by the condition $Z \simeq 1$, which translates to $L^* \simeq g\xi^{-2/3}$, where ξ is the dimensionless parameter defined below in eq 6.7. For the range of D used here, L^* increases with tube size leading to a more gradual crossover to self-avoiding behavior with increasing tube size. This naturally follows from the fact that ξ monotonically decreases with D for a fixed l_{p}/w as portrayed in the inset of Figure 6.3(b). Note that none of the curves in Figure 6.3(a) exhibit truly ideal chain scaling, a result that is reminiscent of dilute solution behavior of DNA (Tree et al., 2013b). However, they do show signatures of ideal chain scaling to varying degrees depending on the value of ξ.

In order to estimate the extension of the confined chain in the self-avoiding limit ($L > L^*$), Odijk suggested that the equilibrium conformations are governed by a competition between the tendency of the chain to maximize entropy by backfolding, and the excluded volume interactions between proximate segments of the chain (Odijk, 2008). Accordingly, the Flory free energy for a chain of contour length L is (Odijk, 2008)

$$\frac{F}{kT} \sim \frac{L}{\lambda} + \frac{X^2}{Lg} + \frac{N_\lambda^2 v_{\text{ex}}}{X D^2}, \tag{6.5}$$

where kT is the Boltzmann factor. Minimizing the free energy with respect to the

extension, X, yields (Odijk, 2008; Muralidhar et al., 2014a)

$$\frac{X}{L} \sim \xi^{1/3}, \tag{6.6}$$

where ξ is a dimensionless parameter defined as

$$\xi \cong \frac{(g/\lambda)^2 v_{\text{ex}}}{gD^2} = \frac{gw}{D^{5/3} l_{\text{p}}^{1/3}} \tag{6.7}$$

Physically, ξ is the Z parameter for a segment of the chain with g/λ deflection segments and extension g. If $\xi > 1$, the excluded volume inside a potential hairpin is so strong that hairpin formation is highly improbable. Therefore $\xi > 1$ represents the condition for the classic Odijk regime in Odijk's scaling theory. Conversely, when $\xi < 1$, the chain is able to backfold. Therefore, the Flory theory in Eq. 6.5 and the resulting average extension in Eq. 6.6 are valid only if $\xi < 1$. In the long-chain limit, all the chains form a self-avoiding walk of hairpins exhibiting a scaling of $R_x \sim L$, as depicted by a slope of unity in Figure 6.3(a).

To test whether these data conform with the predicted scaling in the backfolded Odijk regime (Eq. 6.6), we plot the normalized mean span $X/(L\xi^{1/3})$ against contour length in Figure 6.3(b). Henceforth, we use the mean span instead of the end-to-end distance as the metric of extension for we know that the span converges to the long-chain limit quicker than the end-to-end distance (Muralidhar et al., 2014b; Smithe et al., 2015). We observe that the fractional extension collapses to a constant value ($X/L = 0.99\xi^{1/3}$) for chains with $L \gg g$ if $\xi \ll 1$. The inset of Figure 6.3(b) confirms that the data points corresponding to $\xi < 0.3$ exhibit the scaling predicted by Eq. 6.6. Thus the scaling theory by Odijk not only correctly predicts the asymptotic extension, but also explains the molecular weight dependence of the extension of confined chains in the backfolded Odijk regime. Our calculations represent the first test of the Odijk theory for circular channels.

6.3.2 Comparison with the cooperativity model

In the model of Dai et al. (2012), the confined chain is assumed to consist of units of S-loops and deflection states, each having a unit contour length of πD. In analogy with the Zimm-Bragg model for helix-coil transitions (Zimm and Bragg, 1959), the S-loop state in their model is similar to the helix state while the deflection state is analogous

to the coil state. Accordingly, there are several assumptions inherent in the two-state model. First, just as the presence of a helical turn facilitates the formation of another helical turn in its neighborhood in the Zimm-Bragg model, the presence of an S-loop unit is assumed to promote the presence of other S-loop units in its surroundings. Second, the free energy difference between a hairpin turn and the deflection state is assumed to arise from the bending energy for hairpin formation and is written as

$$\frac{F_{\mathrm{u}}}{kT} = \frac{c_1 \pi l_{\mathrm{p}}}{D} - c_2, \tag{6.8}$$

where c_1 and c_2 are $O(1)$ constants. The constant c_2 is the entropic contribution. Note that F_{u} corresponds to the free energy of the boundary between the S-loop state and the deflection state. Third, Dai et al. (2012) argue that the growth of the S-loop is determined by the excluded-volume interactions between the parallel segments within the S-loop. Accounting for the orientation of deflection segments (Odijk, 2008), the excluded volume between two proximate segments in the S-loop can be written as $E_x \approx 1/2(\pi D)^2 w (D/l_{\mathrm{p}})^{1/3}$. The volume available for a unit of the S-loop is given by $(\pi D^2/4)(\pi D)$. They then proposed a Flory-type free energy for the free energy of a unit in the S-loop state as the ratio of the excluded volume between 3 segments in a S-loop and the total volume available,

$$\frac{F_{\mathrm{s}}}{kT} \approx 3\frac{E_x}{\pi^2 D^3/4} = c_3 \frac{6w}{D^{2/3} l_{\mathrm{p}}^{1/3}}, \tag{6.9}$$

where c_3 is a third $O(1)$ prefactor.

Dai et al. (2012) estimated the extension of the confined chain with S-loops as (Dai et al., 2012)

$$\left(\frac{X}{L}\right)_{\mathrm{S}} = \left(\frac{X}{L}\right)_{\mathrm{CO}} \left(1 - \frac{2}{3} f_{\mathrm{s}} L_{\mathrm{s}} - \frac{1}{3}\pi D f_{\mathrm{s}}\right), \tag{6.10}$$

where f_{s} is the number of S-loops per unit contour length and L_{s} is the average contour length stored in a S-loop. The fractional extension in the classic Odijk regime is given by

$$\left(\frac{X}{L}\right)_{\mathrm{CO}} = 1 - \alpha_{\circ} \left(\frac{D}{l_{\mathrm{p}}}\right)^{2/3}, \tag{6.11}$$

where $\alpha_{\circ} = 0.1701$ is the prefactor for confinement in circular tubes (Yang et al., 2007). The unknowns f_{s} and L_{s} in Eq. 6.10 were measured by Dai et al. (2012) from simulations in the range $1 \leq D/l_{\mathrm{p}} \leq 2$ (Dai et al., 2012). By solving their model along the lines of the

Zimm-Bragg model, f_s and L_s can be written as

$$f_s = \frac{\langle N_d \rangle}{L} = \frac{u}{2\pi D(1-u)} \left[\frac{1+s}{\{(1-s)^2 + 4su\}^{1/2}} - 1 \right] \tag{6.12}$$

and

$$L_s = \pi D \frac{\langle N_s \rangle}{\langle N_d \rangle} = \pi D \left[1 + \frac{1}{2u} \left\{ s - 1 + \left((1-s)^2 + 4su \right)^{1/2} \right\} \right], \tag{6.13}$$

where

$$u = \exp(-2F_u/kT) \text{ and } s = \exp(-F_s/kT) \tag{6.14}$$

are the Zimm-Bragg parameters for nucleation and propagation respectively. Equations 6.12 and 6.13 allowed them to compute c_1, c_2 and c_3 as fitting parameters by relating the parameters of the cooperativity model to values of f_s and L_s obtained from the simulation data.

Next, Dai et al. (2012) substituted for f_s and L_s in terms of the numerical values of c_1, c_2 and c_3 in Eq. 6.10 (Dai et al., 2012). However, they found that Eq. 6.10 does not completely explain the reduction in extension due to backfolding for the finite chain lengths used in their simulations. They then argued that the extension of such short chains is further affected by the presence of C-loops at the two ends, as shown in Figure 6.2. By assuming that the fractional extension of the chain is Boltzmann distributed, the fractional extension of a C-loop was written as

$$\frac{X_c}{L_c} = \left(\frac{X}{L} \right)_{CO} \left(\frac{s'}{1+s'} \right), \tag{6.15}$$

where the average contour length in the C-loop state is

$$L_c = \frac{\pi D u'(1+s')}{2(1-s')(1+u'-s')}. \tag{6.16}$$

Here $u' = \exp(-F_u/kT)$ and $s' = \exp(-F_s/3kT)$ are respectively the nucleation and propagation parameters for C-loop formation. The fractional extension in the presence of both C-loops and S-loops is the weighted average of the middle and end sections of the chain given by

$$\left(\frac{X}{L} \right)_{S\text{-}C} = \left(1 - \frac{2L_c}{L} \right) \left(\frac{X}{L} \right)_S + \left(\frac{2L_c}{L} \right) \left(\frac{X_c}{L_c} \right). \tag{6.17}$$

The C-loop correction to the S-loop model in Eq. 6.17 yielded good agreement with their simulation data (Dai et al., 2012). Henceforth, we refer to the model with S-loops and deflection segments as the 'S-loop model' (Eq. 6.10) and its variation with the C-loop correction as the 'C-loop model' (Eq. 6.17).

Dai et al. (2012) performed simulations for model polymers with 3 different l_p/w values. Here, we choose $l_p/w = 10$ for the sake of comparison, and for the moment fix the contour length to be the same as Dai et al. (2012) (1601 beads). In addition, we use the values furnished by Dai et al. (2012), $c_1 = 1.2$, $c_2 = 1.7$ and $c_3 = 1.08$ (Dai et al., 2012), for the constants in eqs 6.8 and 6.9. Figure 6.4(a) suggests that there is reasonable agreement between the simulation data obtained using the two different methods. The slight disagreement can be attributed to the difference in the methodology between our PERM simulations and their Metropolis Monte Carlo method. Indeed, since PERM is superior at sampling rare events (Prellberg, 2013), the systematic marginal difference in the fractional extension from the two methods points to PERM's greater efficiency in sampling the backfolded states that are difficult to sample using the Metropolis Monte Carlo algorithm. Figure 6.4(a) also confirms the consistency between the C-loop model and simulations for a contour length of 1601 beads (Dai et al., 2012).

The key question is whether the cooperativity model accurately models the chain extension for all contour lengths. Our PERM simulations furnish data for a range of molecular weights well beyond 1601 beads as we earlier illustrated in Figure 6.3. Indeed, our simulation data reach the long-chain limit where the scaling $X \sim L$ is valid. This capability allows us to compare the extension in the long-chain limit, where the S-loop model should be valid. Figure 6.4(b) plots the prediction of the S-loop model (Eq. 6.10) against our simulation data for long chains (15001 beads or more) and the simulation data of Dai et al. (2012) for 1601 beads. First, we verify the observation made by Dai et al. (2012) that the S-loop model fails to predict the extension for chains of finite contour length, as shown by the deviation of the simulation data of Dai et al. (2012) from the $y = x$ line in Figure 6.4(b). Interestingly, the S-loop model also does not predict the extension of long chains, which are at least 15,000 beads in length. Indeed, there is very little difference in fractional extension even when the contour length is increased by an order of magnitude, indicating that chains of 1601 beads are already quite close to the asymptotic limit.

In order to investigate how our simulation data relate to the cooperativity model in more detail, Figure 6.5 plots our simulation data and the prediction of the S-loop and

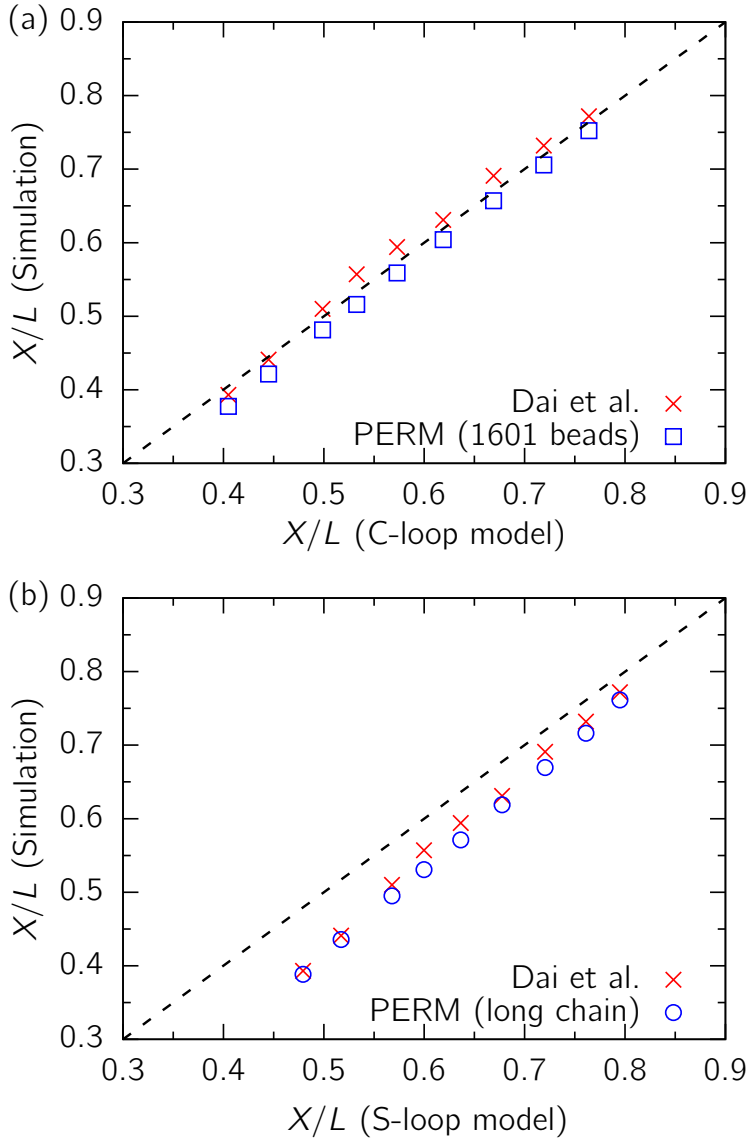


Figure 6.4. (a) Simulation data from Dai et al. (2012) (1601 beads) and PERM (1601 beads) versus the predictions of the C-loop model for a contour length of 1601 beads (Eq. 6.17). (b) PERM data (15001 beads or more) against the predictions of the S-loop model (Eq. 6.10). The dashed line ($y = x$) is the expected result if the model agrees with the simulation. The data shown here correspond exactly to the channel sizes considered by Dai et al. (2012) ($D_{\mathrm{eff}}/l_{\mathrm{p}} = 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.8, 2.0$).

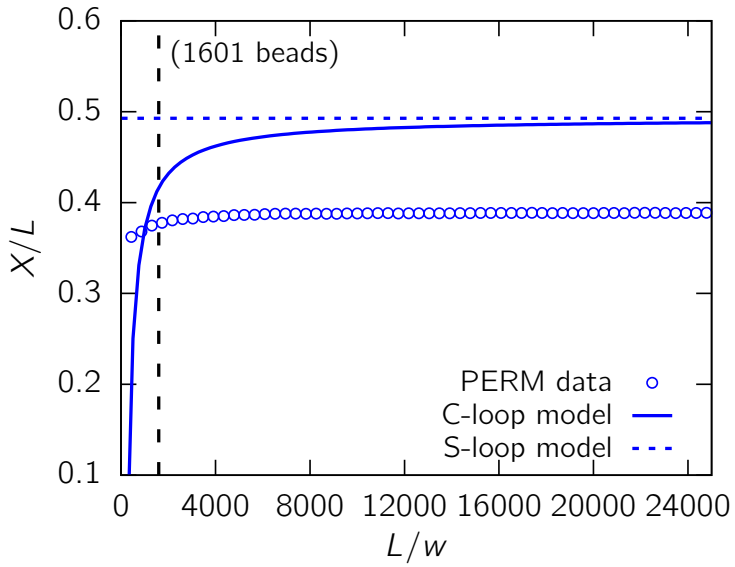


Figure 6.5. Fractional extension as a function of the contour length for a chain with l_p/w = 10 confined in a tube of size, D_{eff}/l_p = 2. The circles are PERM simulation data, the solid line denotes the C-loop model (Eq. 6.17) and the blue dashed line indicates the prediction of the S-loop model (Eq. 6.10). The vertical black dashed line shows the contour length (1601 beads) for which the C-loop model has been parameterized (Dai et al., 2012).

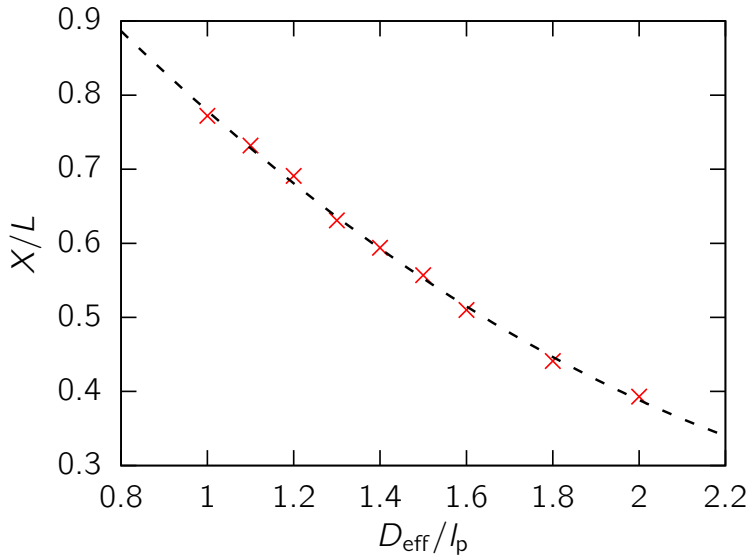


Figure 6.6. Fractional extension against dimensionless tube size taken from Dai et al. (2012) for $l_p/w = 10$ (Dai et al., 2012). The dashed line is a quadractic polynomial fit, $X/L = 0.124(D/l_p)^2 - 0.762(D/l_p) + 1.417$.

C-loop models against contour length for a fixed channel size. It is evident that the fractional extension in our simulations plateaus to a constant value over thousands of beads. Furthermore, the predictions for the S-loop model and the C-loop model converge in the long-chain limit as expected. However, the S-loop model severely overestimates the long-chain fractional extension. Likewise, the C-loop model does not seem to correctly quantify the contour length dependence of the extension. Therefore, it appears that the agreement between simulations and the C-loop model in Figure 6.4(a) is specious, because the curve predicted by Eq. 6.17 (C-loop model) intersects the simulation data only in the region where the C-loop model was parameterized, i.e. for 1601 beads (Dai et al., 2012) as indicated by the vertical dashed line in Figure 6.5. This agreement between the C-loop model and the simulation data only around 1601 beads is not limited to the tube size ($D_{\text{eff}}/l_p = 2$) considered here; we observe a similar trend for other tube sizes as well (see Supporting Information).

6.3.3 Modeling experimental data

While our results thus far indicate that Odijk's scaling theory is the better description for the extension of the chain both as a function of molecular weight and in the long-

chain limit, it still leaves open a more subtle question about which theory does a better job describing experiments for confined DNA. Explicitly, Dai et al. (2012) developed the C-loop model by using a persistence length value of $l_p = 50\,\text{nm}$ and widths of 5 nm, 7.5 nm and 10 nm for chains of a fixed contour length $L = 8\,\mu\text{m}$. These values clearly resemble the parameters one would associate with DNA in experimentally relevant ionic strengths, albeit for a somewhat shorter contour length than typical experiments (Reisner et al., 2005; Gupta et al., 2014; Gupta et al., 2015; Iarko et al., 2015). Even though their model does not capture the contour length dependence of the extension, their work furnishes a model and the associated constants (c_1, c_2 and c_3) corresponding to parameters of DNA. In contrast, Odijk's theory is a scaling theory, and suffers from the typical shortcomings of such theories. In order to get the scaling theory to span a wide range of channel sizes, the monomer anisotropy, l_p/w, must be much larger than the actual values for DNA (Tree et al., 2013b). In practice, this makes Odijk's theory challenging to apply to DNA, since it is difficult to determine what part of the rather narrow range of channel sizes over which Odijk's theory applies corresponds to the theory, and which parts represent transitions into and out of the backfolded Odijk regime (Muralidhar et al., 2014a). Indeed, testing Odijk's theory using experiments with DNA is particularly challenging, since the range of channel sizes is small and puts stringent demands on the nanofabrication. Therefore, it might be attractive to use the C-loop model in practice as opposed to the prediction in the backfolded Odijk regime, simply because it provides a reasonably accurate prediction over the relevant range of channel sizes and contour lengths for DNA-like values of l_p and w.

Recognizing that the C-loop model consists of 3 free parameters c_1, c_2 and c_3, we performed a least-squares fit to the simulation data from Dai et al. (2012) using a quadratic with the same number of fitting parameters. Our simulation data for fractional extension and the resulting quadratic fit are shown in Figure 6.6. We then computed the root mean squared error (RMSE) between the data and the C-loop model (Eq. 6.17). Intriguingly, the RMSE between the data and the best fit quadratic (6×10^{-3}) is an order of magnitude lower than the RMSE between the simulation data and the C-loop model (3×10^{-2}). Given that the C-loop model is only valid in the range $1 \leq D/l_p \leq 2$ for a fixed contour length and includes 3 fitting parameters that are function of l_p/w, using a quadratic that is a better fit of the simulation data might be a more suitable alternative in practice compared to the C-loop model. As was the case with the C-loop model, the quadratic fit would need to be re-parameterized if the l_p/w ratio changes. Thus, we

only advocate a fitting approach for regions of the phase space that are experimentally relevant. For generic analysis, Odijk's theory remains superior since it captures the overall trends as a function of molecular weight, channel size, and monomer anisotropy.

6.4 Concluding remarks

In summary, our simulations of model DNA confined in circular tubes with diameter in the range $l_p \leq D \leq 2l_p$ show good agreement with Odijk's Flory theory for backfolding of semiflexible polymers. Our work also exposes the challenges inherent in using the two-state cooperativity model to predict the contour length dependence of the extension of the confined chain. Although our work was limited to a single value of l_p/w for chains confined in a fixed geometry (circular tubes), we expect our conclusions to be valid for other values of l_p/w and confinement in other geometries such as rectangular channels (Werner and Mehlig, 2015). Indeed, Odijk's theory is superior for chains with a higher value of l_p/w as it allows for more than 3 parallel segments at a certain cross section in the tube, which is likely to be the case for very thin chains, while the maximum number of backfolds in the cooperativity model is fixed. Furthermore, the boundaries for the backfolded Odijk regime are well defined by the scaling theory, both for the confinement size ($\xi < 1$) and the contour length ($L > g\xi^{-2/3}$). In contrast, the boundaries for the cooperativity model are estimated as $l_p \leq D \leq 2l_p$.

Chapter 7

Backfolding in rectangular channels: Test of Odijk's theory

7.1 Introduction

* The confinement of a wormlike chain in a channel smaller than its persistence length is commonly referred to as the Odijk regime (Odijk, 1983; Reisner et al., 2012). For short chains in the Odijk regime, the frequency of backfolding is governed by a balance between the enthalpic penalty due to sharply bending the chain and the entropy gain due to the larger configurational space of a folded chain (Odijk, 2006). For long chains, excluded volume interactions between backfolded segments of the chain become important as well (Odijk, 2008). Based on these two ideas, Odijk proposed that long wormlike chains, such as DNA, confined to channels smaller than the persistence length can exist in two different regimes: (i) a "classic" Odijk regime, where the excluded volume effect is strong and the chain is strongly extended; and (ii) a "backfolded" Odijk regime, where the excluded volume is weak and the chain is able to backfold (Odijk, 2008). We have recently provided simulation evidence that strongly supports Odijk's theory for confinement in square nanochannels (Muralidhar et al., 2014a) and circular tubes (Muralidhar and Dorfman, 2016). In the present contribution, we provide similar data for rectangular channels. Our work thus completes the validation of the backfolded Odijk regime.

Odijk's theory (Odijk, 2008) for rectangular channels has two important implications for experiments on confined DNA. From a fundamental standpoint, rectangular

*This chapter is based on A. Muralidhar, M. J. Quevillon, and K. D. Dorfman, "The backfolded Odijk regime for wormlike chains confined in rectangular nanochannels", Polymers (In review) (2016)

nanochannels are the key to an efficient experimental strategy, known as confinement spectroscopy, that uses devices with a fixed channel depth and a variable channel width (Persson et al., 2009; Gupta et al., 2014; Gupta et al., 2015). Such devices allow one to measure the chain properties over a range of channel sizes on a single device, which is a cost-effective strategy. Perhaps more importantly, these devices allow one to observe the same molecule in many different channel sizes, making them truly single-molecule experiments (Gupta et al., 2015). From a practical standpoint, rectangular channels often result from limitations in the device fabrication. Lithography and etching have a finite precision, and this effect is enhanced as the channel size decreases. In other words, a 200 nm wide by 200 nm deep channel that ends up being 210 nm wide is less of an issue than a 40 nm wide by 40 nm deep channel that ends up being 50 nm wide. Moreover, during device fabrication, it is important that the channels be smooth to avoid DNA adsorption. In practice, this often leads to a sacrifice in the exact value of the channel width in exchange for smooth channels.

In the present contribution, we used pruned-enriched Rosenbluth method (PERM) simulations (Grassberger, 1997; Prellberg and Krawczyk, 2004) of an off-lattice, discrete wormlike chain model (Tree et al., 2013a) to test Odijk's theory (Odijk, 2008) for DNA confined to a rectangular nanochannel. PERM is a biased chain growth method that avoids the attrition problem for self-avoiding random walks by suitably pruning and enriching configurations, thus enabling simulations of extraordinarily long chains (Grassberger, 1997). Our implementation of off-lattice PERM, described in detail elsewhere (Tree et al., 2013a; Tree et al., 2013b), allows us to grow chains out to very high molecular weights (circa 10^5 beads) with high resolution (at least 10 beads per persistence length). Such high molecular weights are key to our analysis. First, they allow us to compute the global persistence length characterizing the bending of ideal chains (Muralidhar et al., 2014a), which is a required input for Odijk's theory (Odijk, 2008). Second, high molecular weights are required to obtain asymptotically long chains in rectangular channels, since the chain must experience the full effect of confinement from each dimension. We use the simulation data to test Odijk's predictions (Odijk, 2008) for the confinement free energy, average chain extension, and variance about the average extension. At the conclusion of our analysis, we discuss an intriguiging implication of our results for genome mapping technology, as well as the perspectives for using simulations such as ours to test predictions for the blob regimes of confinement in rectangular channels (Odijk, 2008; Werner and Mehlig, 2015).

7.2 Odijk Scaling Theory

The problem at hand consists of a wormlike chain of persistence length l_p, effective width w, and contour length L confined in an infinitely long, rectangular channel of height D and width A. For definiteness, the channel height is the smaller dimension, i.e. $D \leq A$. Odijk (2008) describes such channels as nanoslits. This term is used differently in different papers, so it is worthwhile to clarify its definition here. We use the most common nomenclature (Dorfman et al., 2013), where a rectangular nanochannel corresponds to values of A where the chain feels the effect of confinement of the walls in that direction and a nanoslit corresponds to values of A so large that confinement is effectively between parallel plates. In what follows, we only consider rectangular nanochannels and chains with contour length L sufficiently long so that the average extension, variance about that extension, and confinement free energy become extensive in contour length. We also only consider channel sizes that correspond to the Odijk regime. Odijk proposed that $D \lesssim \pi l_p$ should be the restriction on the channel size (Odijk, 2008); detailed simulations (Dai et al., 2014) suggest the requirement is closer to $D < 4l_p$. Here, we use a more conservative condition, $D \leq 2l_p$. In the nomenclature of Werner and Mehlig (2015), the portions of the phase space we consider corrrespond to regimes IIIa and IIIb for confinement of a semiflexible polymer in a rectangular channel. Since regime IIIa is the classic Odijk regime, which has been studied extensively in rectangular channels (Burkhardt, 1997; Burkhardt et al., 2010), our focus will be on the backfolded Odijk regime (regime IIIb).

For channels where $D \lesssim l_p$, Odijk's theory (Odijk, 2008) replaces the original wormlike chain with a "chain of deflection segments" (Muralidhar et al., 2014a). Thus, the chain now consists of $N_\lambda = L/\lambda$ segments, where $\lambda \cong D^{2/3} l_p^{1/3}$ is the contour length of each deflection segment (Odijk, 1983). Note that the deflection segment length is governed only by the smallest dimension in the channel. The bending of the chain of deflection segments along the axis of the channel is characterized by a global persistence length, g, which we will discuss in more detail below. For the moment, we simply point out that g is defined in the absence of excluded volume for exactly the same reason that the persistence length l_p of a wormlike chain is defined in the absence of excluded volume (Odijk, 2006; Muralidhar et al., 2014a).

The free energy of confinement, F, of a real chain can be written as

$$F = F^\circ + F^{\text{ex}}, \tag{7.1}$$

where F° represents the energetic cost to convert a coiled ideal chain in free solution into deflection segments, when the chain is confined between the four walls of the channel (Odijk, 1983; Muralidhar et al., 2014a). The term, F^{ex}, accounts for the excess energetic cost involved in confining a real chain. Using a Flory-like approach, the confinement free energy can be written as (Odijk, 2008)

$$\frac{F}{k_{\text{B}}T} \cong \frac{F^\circ}{k_{\text{B}}T} + \frac{X^2}{gL} + \frac{N_\lambda^2 v_{\text{ex}}}{XAD}, \tag{7.2}$$

where $k_{\text{B}}T$ is the thermal energy. The excess free energy, F^{ex}, is split into two parts. The first part, X^2/gL, represents the entropic cost of stretching an ideal chain of deflection segments to a distance X in the nanochannel. The stretching contribution to the free energy is standard Flory theory applied to a chain of deflection segments, which is the reason why the wormlike chain persistence length l_{p} is replaced by the global persistence length g. The second part of the excess free energy is the excluded volume arising from N_λ deflection segments confined inside a volume $V = XAD$. Odijk computed the excluded volume as (Odijk, 2008)

$$v_{\text{ex}} = \lambda^2 w \langle |\sin\delta| \rangle, \tag{7.3}$$

where $\lambda^2 w$ is the excluded volume between two deflection segments and $\langle |\sin\delta| \rangle$ accounts for the orientation angle δ between two deflection segments.

If $D \leq A \leq 2l_{\text{p}}$, the leading order contribution to the first term in Eq. 7.2 is given by the free energy of an ideal chain confined in a rectangular channel in the classic Odijk regime (Burkhardt et al., 2010),

$$\frac{F^\circ}{k_{\text{B}}T} = \frac{c_\square L}{l_{\text{p}}^{1/3}} \left(\frac{1}{D^{2/3}} + \frac{1}{A^{2/3}} \right), \tag{7.4}$$

where $c_\square = 1.1036$ is a universal constant (Burkhardt, 1997; Burkhardt et al., 2010). In the case of a square channel with size D, the right hand side of Eq. 7.4 reduces to $2c_\square N_\lambda$. Note that this part of the free energy was not included in the original Flory theory by Odijk (2008), as it has no effect on the average chain extension, but its inclusion here is useful to test the predictions for the excess free energy of confinement. For a rectangular

channel with $D \leq A \leq 2l_{\mathrm{p}}$, this orientation factor is approximately (Odijk, 2008)

$$\langle |\sin\delta| \rangle \approx \left(\frac{A}{l_{\mathrm{p}}}\right)^{1/3}. \tag{7.5}$$

Odijk (Odijk, 2008) discussed the error inherent in this approximation; the prefactor of $[1 + (D/A)^{2/3}]^{1/2}$ on the right hand side arising in this limit proves to be a small correction to the scaling theory (see supplementary material).

The average extension of the chain is obtained by minimizing Eq. (7.2) with respect to X, leading to (Odijk, 2008)

$$\langle X \rangle \simeq L\xi^{1/3}, \tag{7.6}$$

where

$$\xi \equiv \frac{g v_{\mathrm{ex}}}{\lambda^2 AD} \cong \frac{gw}{A^{2/3} D l_{\mathrm{p}}^{1/3}}. \tag{7.7}$$

Note that there is a typographical error in Eq. (19) of Ref. (Odijk, 2008) that is corrected here. The confinement free energy F_{c} corresponds to evaluating Eq. (7.2) at $\langle X \rangle$, which produces

$$\frac{F_{\mathrm{c}}}{k_B T} \simeq \frac{F^{\circ}}{k_{\mathrm{B}} T} + c\frac{L\xi^{2/3}}{g}, \tag{7.8}$$

where c is a prefactor for the scaling of the excess free energy. Accordingly, a restatement of the free energy scaling predicted by Flory theory is

$$\frac{F^{\mathrm{ex}}}{k_{\mathrm{B}} T} \simeq \frac{L\xi^{2/3}}{g}. \tag{7.9}$$

Finally, the effective spring constant is obtained from the second derivative of Eq. (7.2) evaluated at $\langle X \rangle$,

$$k_{\mathrm{eff}} = \left(\frac{\partial^2 F}{\partial X^2}\right)_{X=\langle X \rangle}. \tag{7.10}$$

From equipartition of energy for a harmonic oscillator, the spring constant is inversely related to the variance of the chain extension about its mean,

$$k_{\mathrm{eff}} = \frac{k_B T}{\delta X^2}. \tag{7.11}$$

Combining the latter pair of equations with Eq. (7.2) furnishes the variance in chain

extension,

$$\delta X^2 \cong Lg. \tag{7.12}$$

The scaling laws derived thus far are under the assumption that excluded volume interactions between deflection segments of the chain are weak enough to permit backfolding and hairpin formation. This condition is expressed in scaling form as $(g/\lambda)^2 v_{\text{ex}} < gAD$, which translates to $\xi < 1$ (Odijk, 2008; Muralidhar et al., 2014b). The key results in Eqs. (7.6), (7.8) and (7.12) are seemingly identical to those for a square channel in Odijk's theory (Odijk, 2008; Muralidhar et al., 2014a). The differences are (i) the length scale $D^{-5/3}$ appearing in ξ for a square channel of size D is partitioned into $A^{-2/3}D^{-1}$ in a rectangular channel and (ii) the global persistence length g is different in a square and a rectangular channel. The latter effect is easily seen at a qualitative level. As the aspect ratio increases for a fixed value of D, it becomes easier for the chain to bend along the wider direction because the radius of curvature of the hairpin increases. As a result,

$$\left(\frac{\partial g}{\partial A}\right)_D < 0. \tag{7.13}$$

Odijk proposed a mechanical theory for the global persistence length in a rectangular channel, leading to the expression (Odijk, 2006; Odijk, 2008)

$$g = 3.3082\bar{r}\exp\left(\frac{F_{\text{mc}}}{k_B T}\right), \tag{7.14}$$

where

$$\bar{r} = \frac{1.5071 l_{\text{p}} A}{A + 3.0142 l_{\text{p}}} \tag{7.15}$$

is the size of the hairpin that minimizes the mechanical energy and

$$\frac{F_{\text{mc}}}{k_B T} = \frac{1.5071 l_{\text{p}}}{\bar{r}} - \ln\left[\left(\frac{A - 2\bar{r}}{A}\right)\frac{D}{\pi\bar{r}}\right] + 1 \tag{7.16}$$

is the free energy required to bend the hairpin in the mechanical limit (Odijk, 2006). Note that Eqs. (7.14)-(7.16) correspond to those in the more recent paper by Odijk (Odijk, 2008).

While our primary objective is to asses Odijk's scaling theory results in Eqs. (7.6), (7.8) and (7.12), a necessary first step is to test Odijk's mechanical theory for the global persistence length in Eq. (7.14). As was the case with our study of Odijk's scaling theory in square channels (Muralidhar et al., 2014a), the validity (or lack thereof) of

Odijk's theory for the global persistence length is not a prerequisite for assessing the scaling theory. We will compute the values of g from simulations of channel-confined ideal chains, which will then be used to analyze the simulations of channel-confined real chains.

7.3 Simulation Method

The simulation method used here is identical to our previous study of the backfolded Odijk regime in square nanochannels (Muralidhar et al., 2014a) with the obvious exception that the channels here are rectangular. In order to verify that our results for rectangular channels reduce that of squares when the aspect ratio tends to unity, we perform simulations in square channels as well. Briefly, the chain is modeled by $N + 1$ beads connected by rods of bond length a, corresponding to a contour length of $L = Na$. To incorporate excluded volume effects, we impose a hard-core width of w for each bead. For ideal chain simulations, $w = 0$ and for real chain simulations, $w = a$. We make use of PERM, a chain growth method that natively produces information as a function of contour length. The wormlike nature of the chain is enforced by a bending energy between each trio of contiguous beads, and the beads experience hard-core excluded volume interactions with other beads and the channel walls. Additional technical information about our implementation of PERM (Tree et al., 2013a; Tree et al., 2013b) and its application to confinement in the Odijk regime (Muralidhar et al., 2014a; Tree et al., 2014) are available elsewhere.

To measure the global persistence length for a particular value of D and A, we simulate chains with zero excluded volume. We then measure the projection of the mean squared end-to-end distance R_x^2 of these ideal chains along the channel axis as a function of contour length (Muralidhar et al., 2014a). The global persistence length is extracted by fitting the data with the one-dimensional wormlike chain result (Spakowitz and Wang, 2003; Tkachenko and Rabin, 1995)

$$R_x^2 = \frac{1}{3}(1 + 2m)[2gL - 2g^2(1 - \exp(-L/g))] \tag{7.17}$$

where

$$m = \frac{1}{2}(3\langle\cos^2\theta\rangle - 1) \tag{7.18}$$

is the orientational order parameter (Wagner et al., 2007; Spakowitz and Wang, 2003;

Chen, 2013a). In the latter, θ is the angle formed between the tangent to the chain backbone and the channel axis.

In the analysis of real chains, we account for the excluded volume between the chain and the wall by defining an effective channel depth $D_{\text{eff}} = D - w$ and effective channel width $A_{\text{eff}} = A - w$. As a result, the data presented in Section 7.4.2 correspond to

$$\xi \cong \frac{gw}{A_{\text{eff}}^{2/3} D_{\text{eff}} l_{\text{p}}^{1/3}} \tag{7.19}$$

rather than Eq. (7.7). Likewise, the values of g used to characterize the properties of real chains correspond to the effective channel dimensions $g = g(D_{\text{eff}}, A_{\text{eff}})$. To test the prediction of Odijk's theory for the scaling of the average extension and its variance, we use the mean span,

$$X \equiv \left\langle \max_i (x_i) - \min_i (x_i) \right\rangle, \tag{7.20}$$

as the metric of extension as this is the quantity that is measured in experimental studies of DNA in nanochannels (Reisner et al., 2012; Wang et al., 2011). In the latter equation, x_i is the x coordinate of the ith bead, where the bead index i is an integer such that $i \in [1, N]$.

We carried out a large number of simulations for both ideal and real chains for a variety of persistence lengths and channel sizes, A and D. The details are summarized in Table 7.1.

Table 7.1: Summary of parameters used in PERM simulations. The quantity $[a, b]_n$ in the table indicates a sequence of n values logarithmically spaced between a and b, including a and b. The maximum molecular weight in each simulation was set based on the values of l_p, w, D and A so as to reach the limit where the theromdynamic quantities are extensive in chain length. Depending on these parameters, the maximum molecular weight in our simulations ranges from 10,000 to 100,000 beads.

		Confined ideal chains		
	l_p/a	D/l_p	A/D	# Tours ($\times 10^5$)
Set 1a	(10, 15, 20)	$[0.5, 1.94]_8$	(1,2,3,4)	30
Set 1b	(10, 15, 20)	$[0.5, 1.32]_6$	1.5	30
Set 1c	(10, 15, 20)	$[2.36, 20]_{12}$	2	5
Set 1d	(10, 15, 20)	$[2.36, 6.23]_6$	3	5
Set 1e	(10, 15, 20)	$[2.36, 4.23]_4$	4	5
Set 2a	(12.5, 17.5)	$[0.5, 2.0]_7$	1	60
Set 2b	(12.5, 17.5)	$[0.5, 1.0]_5$	1.5	30
Set 2c	(12.5, 17.5)	$[0.5, 1.0]_5$	2	60
Set 2d	(12.5, 17.5)	$[0.5, 0.67]_3$	3	60
Set 2e	(12.5, 17.5)	0.5	4	60
		Confined real chains		
	l_p/w	$D_\mathrm{eff}/l_\mathrm{p}$	$A_\mathrm{eff}/D_\mathrm{eff}$	# Tours ($\times 10^5$)
Set 3a	(10, 15, 20)	$[0.5, 1.94]_8$	1	10
Set 3b	(10, 15, 20)	$[0.5, 1.32]_6$	1.5	10
Set 3c	(10, 15, 20)	$[0.5, 0.90]_4$	2	10
Set 3d	(10, 15, 20)	$(0.5, 0.61)$	3	10
Set 3e	(10, 15, 20)	0.5	4	10
Set 4a	(12.5, 17.5)	$[0.5, 2.0]_7$	1	30
Set 4b	(12.5, 17.5)	$[0.5, 1.0]_5$	1.5	10
Set 4c	(12.5, 17.5)	$[0.5, 1.0]_5$	2	30
Set 4d	(12.5, 17.5)	$[0.5, 0.67]_3$	3	30
Set 4e	(12.5, 17.5)	0.5	4	30

7.4 Results

7.4.1 Global Persistence Length

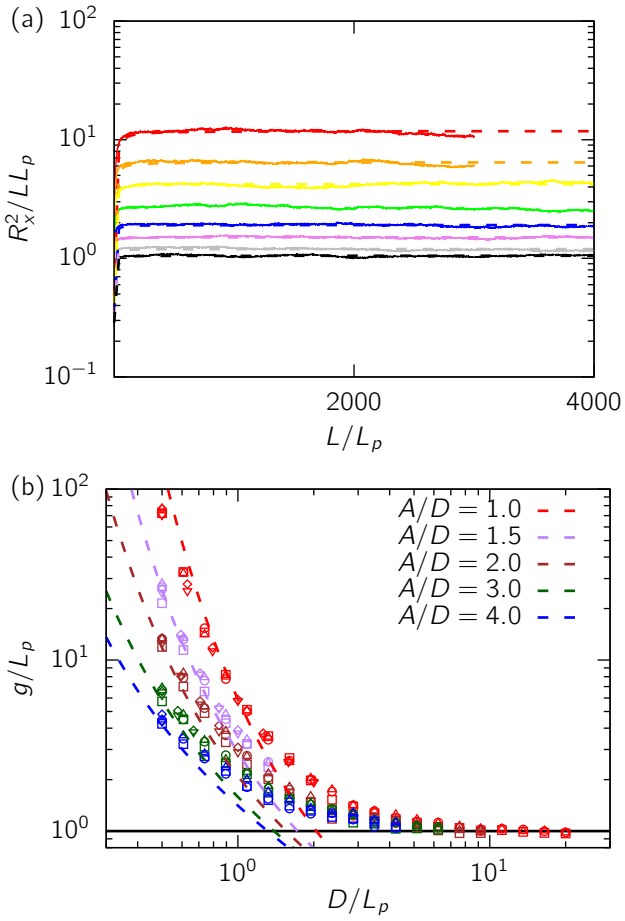


Figure 7.1. (a) Illustration of our method for obtaining the global persistence length. The curves here correspond to a subset of Set 1a (see Table 7.1) of our data for $l_p/a = 10$ and a fixed aspect ratio of $A/D = 2$. The solid curves are our simulation data and the dashed lines are the best fits for Eq. 7.17. The red curve corresponds to the smallest channel size $D/l_p = 0.5$ and the black curve corresponds to the biggest channel size of $D/l_p = 1.94$. The channel size increases from top to bottom. (b) The global persistence length thus obtained for all our data in Table 7.1 against dimensionless channel size. The five colors represent the five aspect ratios considered here. The different point types correspond to $l_p/a = 10$ (□), $l_p/a = 12.5$ (▽), $l_p/a = 15$ (◯), $l_p/a = 17.5$ (◇) and $l_p/a = 20$ (△). The dashed lines are from Eq. 7.21. The horizontal black line indicates $g = l_p$, which should be the limiting value of g for $D \gg l_p$.

Figure 7.1 shows our data for the global persistence length for the entire range of parameters in Set 1 and 2. Figure 7.1(a) contains a typical plot of the end-to-end distance along the axis of the channel against the contour length. By using Eq. 7.17, we obtain the global persistence length as a fitting parameter to these curves, as shown in Figure 7.1(a). Figure 7.1(b) plots the normalized global persistence length against dimensionless channel size. We observe a collapse of our data for the global persistence length for a fixed aspect ratio of the channel, irrepective of the discretization used in the DWLC models for ideal chains ($l_p/a = 10, \ldots, 20$). As expected, for a given value of D/l_p, g decreases with increase in A in accordance with Eq. 7.13. Furthermore, all the curves in Figure 7.1(b) limit to the persistence length ($g \to l_p$) for $D \gg l_p$, as expected.

To quantitatively assess the dependence of g on A and D, we compare our data to Odijk's mechanical theory (Odijk, 2006). Odijk derived Eqs. 7.14-7.16 in the mechanical limit, neglecting fluctuations of hairpins in the channel. Moreover, the theory makes certain approximations about the shape of the hairpin by considering only the leading order terms to estimate its shape. Cognizant of such approximations, Odijk (2006) proposed that there should be corrections to the free energy of hairpins in Eq. 7.16, which can be lumped into an additional term as $F = F_{mc} + H$. Our comparison with Eqs. 7.14-7.16 suggests that the mechanical theory overestimates g by about 2 orders of magnitude (see supplementary material). We find that

$$g = 3.3082\bar{r} \exp\left(F_{mc}/k_B T - 5.01\right) \tag{7.21}$$

roughly captures the dependence of g on the channel size for all four A/D ratios used in our work. In our previous work on square nanochannels, we found that such a correction leads to good agreement with the simulation data given that $D/l_p < 1$, which is roughly the range of channel sizes where Odijk's mechanical theory should be valid. Here, the agreement between this correction for the free energy and its prediction of g increasingly deviates from our data as the aspect ratio A/D increases. Note that for a given channel height, D/l_p, the value of A/l_p increases with increasing aspect ratio. Accordingly, the deviation of g from the correction to Odijk's theory for high aspect ratio channels can be attributed to the approximations in Odijk's mechanical model, which become increasingly worse for larger channels.

7.4.2 Scaling Theory

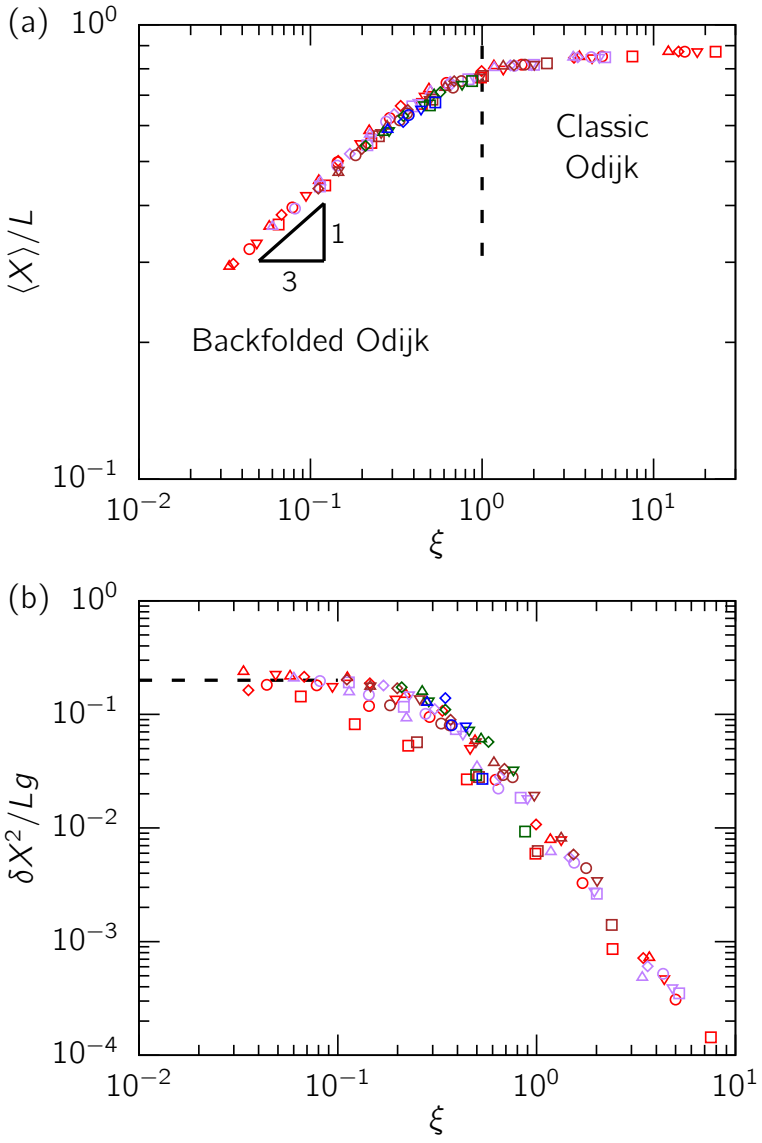


Figure 7.2. (a) Fractional extension against the scaling variable ξ. The vertical dashed line shows the boundary between the classic and backfolded Odijk regimes according to the scaling theory, $\xi = 1$. A power law fit to our data for $\xi < 0.3$ yields an exponent of 0.33. (b) Normalized variance of extension versus ξ. The horizontal dashed line corresponds to $\delta X^2/Lg = 0.2$. In both the panels, the l_p/w values shown are $l_p/w = 10$ (□), $l_p/w = 12.5$ (▽), $l_p/w = 15$ (○), $l_p/w = 17.5$ (◇) and $l_p/w = 20$ (△). Different colors indicate different A_{eff}/D_{eff} ratios: red (1), purple (1.5), brown (2), green (3) and blue (4). All the data points satisfy the condition $D_{eff} \leq A_{eff} \leq 2l_p$.

Although the mechanical theory of Odijk (Odijk, 2006) does not quantitatively explain the variation of g, this inconsistency does not pose a problem in testing Odijk's scaling theory (Odijk, 2008) as we calculated the value of g numerically for a wide range of parameters. Here, we compare our data for real chains to the predictions of the scaling theory by using the numerically obtained values of g for equivalent ideal chains, as discussed in Section 7.4.1. In what follows, we only consider long chains with $L \gg g$, where quantities such as the average extension, its variance and the confinement free energy are extensive in the contour length.

We begin our test of the scaling theory by examining the dependence of the average extension of the confined molecule on the scaling variable ξ in Figure 7.2(a). This plot was generated by numerically calculating the value of ξ from Eq. 7.19 from our ideal chain simulations for equivalent channel sizes A_{eff} and D_{eff}, similar to our procedure in Refs. (Muralidhar et al., 2014a; Muralidhar and Dorfman, 2016). We observe that our data collapse onto a single curve irrespective of the l_{p}/w and A/D ratios. Indeed, the data exhibit the scaling $\langle X \rangle / L \sim \xi^{1/3}$ for $\xi \ll 1$, in accordance with Odijk's scaling theory (Eq. 7.6). Furthermore, our data transition to fractional extensions close to unity predicted for the classic Odijk regime as ξ approaches 1, consistent with our previous simulations in square channels (Muralidhar et al., 2014a).

We consider the variance of extension in Figure 7.2(b). Although the agreement of the variance data with the scaling theory (Eq. 7.12) is not as striking as that for the extension, we indeed see that the data collapse onto the scaling prediction, $\delta X^2 \sim Lg$ (Eq. 7.12). Similar to our findings for square channels (Muralidhar et al., 2014a), this scaling is limited to the parameter range where $\xi < 0.1$. In comparison, the range of data for the mean extension that agree with the scaling theory (Eq. 7.6) is wider, as shown in Figure 7.2(a). The data for the variance appear to follow $\delta X^2 \approx 0.20Lg$, a slightly different prefactor from the value 0.25 that we obtained from our simulations for square channels (Muralidhar et al., 2014a). Considering the error in measuring the variance of extension, and the statistical errors associated with estimating g, the agreement observed in Figure 7.2(b) is acceptable.

We now turn our attention to the confinement free energy. In order to verify if Flory theory correctly predicts the scaling of the free energy as per Eq. 7.9, we take a two step approach. First, we calculate the confinement free energy F° of ideal chains with various l_{p}/a ratios confined in rectangular channels of various sizes and aspect ratios, as summarized in Table 7.1. In all our calculations, the free energy is computed with

respect to an ideal unconfined chain with the same l_p/a ratio as the reference state. Second, we compute the confinement free energy of equivalent real chains, F_c, confined in channels with the same values of A_{eff} and D_{eff} used for the ideal chain simulations, again using a reference state of an ideal unconfined chain. This allows us to compute the excess free energy as a difference of the two quantities as $F^{ex} = F_c - F^\circ$.

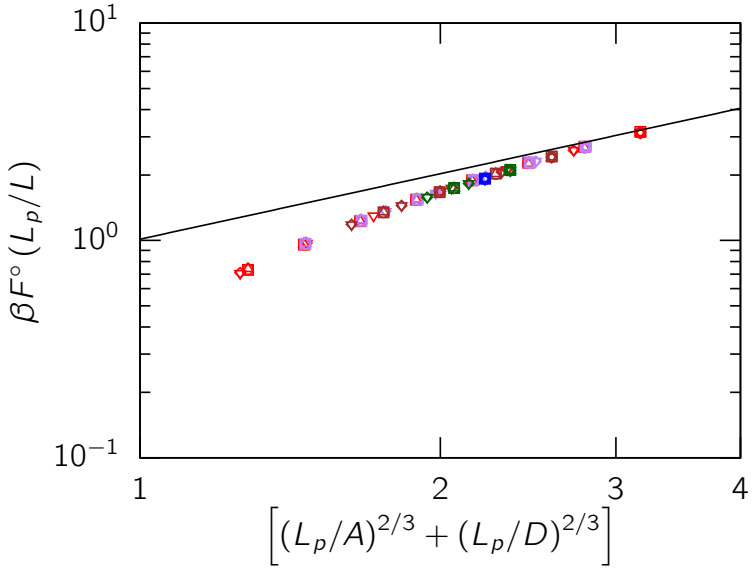


Figure 7.3. Confinement free energy of ideal chains in various channel sizes. The black solid line is the confinement free energy in the classic Odijk regime valid in the limit $D \ll l_p$ and $A \ll l_p$ (Eq. 7.4). The l_p/a values for ideal chains shown here are $l_p/w = 10$ (□), $l_p/w = 12.5$ (▽), $l_p/w = 15$ (◯), $l_p/w = 17.5$ (◇) and $l_p/w = 20$ (△). Different colors indicate different A/D ratios: red (1), purple (1.5), brown (2), green (3) and blue (4). All the data points satisfy the condition $D \leq A \leq 2l_p$.

Figure 7.3 shows the confinement free energy of ideal chains confined in rectangular channels, F°, against a form of the dimensionless channel size suggested by Eq. 7.4. We observe a collapse of our data points irrespective of the aspect ratio of the channel and the channel size indicating that the dimensionless quantity $(l_p/A)^{2/3} + (l_p/D)^{2/3} \gg 1$ determines the behavior of the free energy in this range of channel sizes, in accordance with Eq. 7.4. Furthermore, in the limit $(l_p/A)^{2/3} + (l_p/D)^{2/3} \gg 1$ corresponding to the classic Odijk regime, our data collapse onto to the exact solution (Eq. 7.4) from Burkhardt et al. (2010), which was obtained by numerically solving a Fokker-Planck type differential equation for the confined chain. As expected, for larger channels, for which $(l_p/A)^{2/3} + (l_p/D)^{2/3} \approx 1$, the rescaled free energy increasingly deviate as the sum

of the rescaled channel sizes approaches 1.

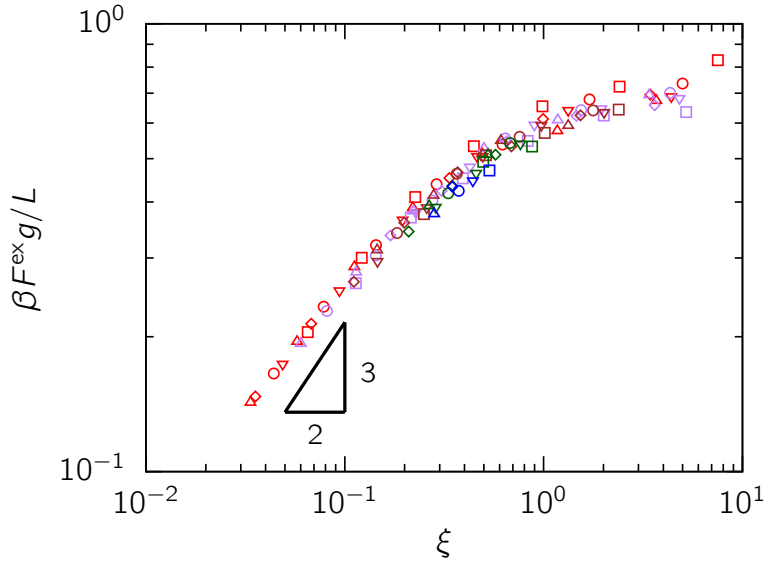


Figure 7.4. The excess free energy of real chains confined in rectangular channels. The l_p/w values shown are l_p/w = 10 (□), l_p/w = 12.5 (▽), l_p/w = 15 (◯), l_p/w = 17.5 (◇) and l_p/w = 20 (△). Different colors indicate different A_{eff}/D_{eff} ratios: red (1), purple (1.5), brown (2), green (3) and blue (4). A power law fit of our data for $\xi < 0.1$ reveals an exponent of 0.56. All the data points satisfy the condition $D_{eff} \leq A_{eff} \leq 2l_p$.

To test the scaling prediction of Odijk's Flory theory in Eq. 7.9, we calculated F^{ex} by subtracting the ideal-chain contribution from the real chain free energy. Figure 7.4 shows the rescaled excess free energy plotted against ξ. For $\xi \ll 1$, we observe a collapse of the data and a scaling of $F^{ex} \sim \xi^{0.56}$ close to $F^{ex} \sim \xi^{2/3}$ predicted by the scaling theory of Odijk (Odijk, 2008). In light of the known drawbacks of the Flory free energy terms for interaction energy and entropy (des Cloizeaux, 1976; de Gennes, 1979; des Cloizeaux, 1980; Schäfer, 2012), the proximity of the observed scaling to the predicted scaling of the free energy is remarkable. Indeed, Flory theory has been consistently known to incorrectly predict the free energy of real chains in free solution, although the cancelation of errors in the individual terms leads to a good prediction of the scaling of the size of a swollen coil (des Cloizeaux, 1976; de Gennes, 1979; des Cloizeaux, 1980; Schäfer, 2012). We suspect that this consistency between Eq. 7.9 and our simulation results is due to the accuracy of the mean field assumptions inherent in Flory theory in the limit $\xi \ll 1$, owing to the weakness of excluded volume in the backfolded Odijk

regime.

7.5 Discussion

7.5.1 Implication for Genome Mapping

In nanochannel-based genome mapping technology, DNA are decorated with sequence-specific probes and then injected into a nanochannel. The physical distance between the probes can be converted into a genomic distance provided that (i) the probes are linearly ordered in the channel, so that the order of the probes in the channel is the same as their order along the DNA backbone, and (ii) the stretching of the chain is uniform, so that there is a simple proportionality constant between the physical distance (in nanometers) and the genomic distance (in base pairs). These criteria are met by confining the DNA into a nanochannel that is smaller than the chain's persistence length, thereby suppressing bending due to thermal fluctuations.

One of the most intriguing aspects of Odijk's theory is the linear dependence of the variance of the chain extension on the global persistence length in Eq. (7.12). The global persistence length decreases with increasing channel width, but the tight confinement in the channel height maintains the deflection segment behavior. At first glance, it would seem that nanochannel genome mapping could be improved by moving from square nanochannels (Lam et al., 2012) to rectangular channels of the same depth. For instance, moving from a 50 nm square channel to a 50 nm by 100 nm rectangular channel, the change in extension is only around 30%. This is because the extension would be reduced by $(D/A)^{2/9}(g_{\text{rectangle}}/g_{\text{square}})^{1/3}$. However, the variance in chain extension is almost halved. Provided that the chain is still sufficiently extended to linearly order the sequence-specific probes on the genome, anisotropic channels may be preferable to their square counterparts by reducing the variance in chain extension.

There is a precedence for using rectangular channels. Indeed, the earliest published data on genome mapping in a nanochannel (Jo et al., 2007) used 100 nm × 1000 nm slits in a low ionic strength buffer such that the persistence length was similar to the channel depth. The use of wide channels in the latter experiments was motivated primarily by simplifying the fabrication, since such channels are readily fabricated by optical lithography and they can be replicated in poly(dimethyl) siloxane for ease of use in the lab. The introduction of commercially viable devices with 45 nm × 45 nm cross-sections (Lam et al., 2012) has marginalized the fabrication advantages of the wider slits. Nevertheless,

Odijk's theory suggests rectangular channels may have additional advantages for reducing the error in the genome mapping experiment. This potential advantage needs to be balanced against the reduction in throughput (or, equivalently, the increased device footprint) introduced by wider channels. Moreover, the image analysis may be more complicated in rectangular channels because the DNA can deflect more easily in the wider direction.

7.5.2 Backfolded Odijk regime for $A > 2l_\mathrm{p}$

In our work, we only considered channels with $D \leq A \leq 2l_\mathrm{p}$. However, when $D \leq 2l_\mathrm{p}$ and $A > 2l_\mathrm{p}$, the chain is tightly confined along the side of length D, but is free to bend along the other dimension in the channel. Depending on the relative sizes of the sides of the channel, g can still be greater than l_p. One can then derive the scaling of the average extension, the variance of extension and the confinement free energy for such a case using Flory theory along the lines of Eq. 7.2 (Odijk, 2008). The confinement free energy of an ideal chain in this case is dominated by collisions of the chain with the walls along the smallest dimension, D, of the channel and can be written as

$$\frac{F^\circ}{k_\mathrm{B}T} = \frac{c_\square L}{l_\mathrm{p}^{1/3}}\left(\frac{1}{D^{2/3}}\right). \tag{7.22}$$

Note that there is a higher order term arising from the collisions of the chain along the larger dimension that scales as A^{-2}, analogous to the scaling of the confinement free energy in the blob regimes ($A \gg 2l_\mathrm{p}$) (de Gennes, 1979; Tree et al., 2013a; Dai et al., 2014).

The second term in Eq. 7.2, X^2/gL, remains unchanged in this case, where the global persistence length, g, is again a function of A and D. However, the weak confinement along the larger dimension causes the deflection segments to be more or less isotropically aligned on an average ($\langle|\sin\delta|\rangle \approx 1$), as in a slit (Odijk, 2008; Tree et al., 2014). Therefore, $v_\mathrm{ex} \simeq \lambda^2 w$ (see supplementary material). Minimizing the free energy again with respect to the extension yields the average extension

$$\langle X\rangle \simeq L\psi^{1/3}, \tag{7.23}$$

where

$$\psi \cong \frac{gw}{AD}. \tag{7.24}$$

The scaling of the variance of extension is unaffected by the change in A,

$$\delta X^2 \simeq Lg \tag{7.25}$$

and the excess free energy of confinement scales as

$$\frac{F^{\mathrm{ex}}}{k_{\mathrm{B}}T} \simeq \frac{L\psi^{2/3}}{g}. \tag{7.26}$$

The simulation results presented in this paper correspond to relatively modest values of the channel anistropy, ranging from a square channel to a rectangle with $A = 4D$. Unfortunately, testing the scaling laws embodied in Eqs. 7.23, 7.25 and 7.26 would require simulations in channels with very large aspect ratios ($A/D > 4$) and is computationally challenging. In order to reach the limits where the thermodynamic properties of the confined chain become extensive in contour length, we require chains that are sufficiently long to explore the larger length scale A. At the same time, we require that the discretization of the model be sufficiently fine to resolve the chain contour in the smaller length scale D. Simultaneously satisfying both constraints necessitates a very large number of beads and thus expensive simulations. While our computational resources would permit somewhat larger values of A, it would be very challenging to reach the values $A = 10D$ used in the highly anisotropic channels used for some early genome mapping experiments (Jo et al., 2007) or recent confinement spectroscopy experiments (Gupta et al., 2014). From a practical perspective, the aspect ratios we studied here are well within expectations for state-of-the-art nanochannel genomic mapping systems (Lam et al., 2012; Hastie et al., 2013; Cao et al., 2014) and thus relevant to the design and operation of these devices.

7.6 Conclusions

By performing simulations of asymptotically long chains, we demonstrated that Odijk's Flory theory correctly captures the mean extension, the variance of extension and free energy for chains confined in rectangular channels for channel sizes where both dimensions are smaller than the persistence length. Our simulations of confined semiflexible polymers in channels with 5 different aspect ratios show that Odijk's mechanical theory overestimates the value of the global persistence length by about 2 orders of magnitude. Nonetheless, the numerical estimation of g allowed us to verify the scaling laws over a

decade in the scaling variable, ξ. The prefactors we obtained for the scaling laws associated with the mean extension, its variance and the confinement free energy can used to predict these quantities for the backfolded Odijk regime in rectangular channels.

Chapter 8

Kirkwood diffusivity of long semiflexible chains in nanochannel confinement

8.1 Introduction

* The dynamic properties of polymers in dilute solution are strongly coupled to their static properties due to long-ranged intramolecular hydrodynamic interactions (HI) (de Gennes, 1979; Rubinstein and Colby, 2003). For a polymer chain confined in a channel, the range of these hydrodynamic interactions is greatly restricted, due to the screening of HI induced by the walls (Diamant, 2009). However, the presence of walls complicates accurate calculation of the frictional forces on the confined chain. The walls not only alter the distribution of monomers of the confined chain (de Gennes, 1979), but also affect the motion of the solvent in the channel (Diamant, 2009). The dynamic properties of confined chains, therefore, are a function of the coupling between the monomer conformations and the motion of the solvent, both of which depend on the cross-sectional dimensions of the channel (Graham, 2011).

The dependence of the near-equilibrium diffusivity of channel-confined single molecules on the confinement size has been the subject of various studies (Harden and Doi, 1992; Tree et al., 2012; Jendrejack et al., 2003) since the seminal work of de Gennes and cowork-

*This chapter is based on A. Muralidhar and K. D. Dorfman, "Kirkwood Diffusivity of Long Semiflexible Chains in Nanochannel Confinement", Macromolecules **48**, 2829 (2015)

ers almost 35 years ago (de Gennes, 1979; Brochard and de Gennes, 1977). The results of these studies, however, were interpreted in the context of the knowledge of the static properties at the time. Renewed interest in the problem of nanochannel-confined DNA (Dorfman et al., 2013) has brought new insights into the static properties of semiflexible polymers in channel confinement (Odijk, 2008; Wang et al., 2011). For a semiflexible polymer of persistence length, l_p, effective width, w, and contour length, L, confined in a square channel of size, H, the dependence of static properties such as the extension of the molecule, the variance of extension, and the confinement free energy on the channel size is now reasonably well understood (Odijk, 2008; Reisner et al., 2012; Wang et al., 2011; Muralidhar et al., 2014a; Dai et al., 2014; Werner and Mehlig, 2014). Indeed, the behavior of the static properties can be categorized into various regimes depending on the value of the channel size, H, relative to the persistence length, l_p and width, w.

For a wide channel, whose channel size is much bigger than the persistence length ($H \gg l_\mathrm{p}$), the static properties of confined semiflexible polymers can be explained by blob theory (de Gennes, 1979; Daoud and de Gennes, 1977; Odijk, 2008). When $H \gtrsim 2l_\mathrm{p}$, de Gennes' blob theory has been shown to correctly predict the scaling of the average extension (Wang et al., 2011; Dai et al., 2014; Cifra, 2009; Muralidhar et al., 2014a). However, other properties including the variance of extension and confinement free energy exhibit a scaling different from de Gennes' blob theory for $2l_\mathrm{p} \lesssim H \lesssim l_\mathrm{p}^2/w$, while the blob theory scaling is recovered for $H \gtrsim l_\mathrm{p}^2/w$ (Dai et al., 2014). A modified blob theory that accounts for the effect of stiffness (Odijk, 2008), however, can be applied to explain this deviation from de Gennes' theory, which does not account for local stiffness of the chain (de Gennes, 1979). Accordingly, the range of channel sizes where de Gennes' theory is satisfied on all counts ($H \gtrsim l_\mathrm{p}^2/w$) is termed the 'de Gennes regime' or the 'classic de Gennes regime', while the region where the modified blob theory is applicable is called the 'extended de Gennes regime' (Dai et al., 2014; Werner and Mehlig, 2014).

In contrast, if $H \ll l_\mathrm{p}$, Odijk suggested that the confined chain forms repeating units of deflection segments, which are linearly ordered, giving rise to an extension close to the contour length (Odijk, 1983). The scaling of static properties in this 'classic Odijk regime' has been extensively studied, so much so that the prefactors in these scaling relations are known to multiple decimal places (Burkhardt et al., 2010; Chen, 2013a). Our recent work confirmed the existence of a universal backfolded Odijk regime, wherein the deflection segments can form hairpins inside the channel (Odijk, 2008; Muralidhar et al., 2014a). Indeed, our calculations of the static properties in the backfolded Odijk regime(Muralidhar

et al., 2014a) showed excellent agreement with Odijk's theory (Odijk, 2008).

The lack of experimental data for single molecule diffusivity in channel confinement means much of our understanding of the scaling of the diffusivity comes from theory and simulation (de Gennes, 1979; Brochard and de Gennes, 1977; Harden and Doi, 1992; Jendrejack et al., 2003; Tree et al., 2012). Although this problem has been the subject of attention for several decades, there are three key open questions surrounding the diffusivity of channel-confined semiflexible polymers.

The first of these questions concerns the scaling of the diffusivity in the classic de Gennes regime. Employing blob theory, Brochard and de Gennes (1977) predicted that the diffusivity of a polymer parallel to the axis of the channel, D, scales with the channel size as $D \sim H^{2/3}$ for $H \gg l_{\mathrm{p}}$. Using the contemporary value of the Flory exponent (Clisby, 2010), this scaling law would translate to $D \sim H^{0.701}$. Harden and Doi (1992) argued that the diffusivity has a slightly weaker scaling, $D \sim H^{0.61}$ (Harden and Doi, 1992). They attribute this deviation from the scaling theory prediction to a transverse correlation length that arises in their variational calculation, which is based on self-consistent field theory. More recently, Brownian dynamics simulations of a bead-spring model by Jendrejack et al. (2003) found that $D \sim H^{0.5}$ in the de Gennes regime. Subsequent simulations of a touching bead model by Tree et al. (2012) appear to show a scaling that is closer to the one predicted by Harden and Doi (1992), albeit with limited amount of data. While the scaling of the static properties of channel-confined polymers in the classic de Gennes regime has been well established for many decades, a definitive answer for the scaling of the diffusivity still remains elusive.

A related question pertains to the scaling of the diffusivity in the extended de Gennes regime. Several authors have proposed that blob theory can also be applied to estimate the diffusivity in the extended de Gennes regime, which would result in a predicted scaling of $D \sim H^{0.701}$ (Reisner et al., 2012; Tree et al., 2012). However, recent work (Dai et al., 2013) on DNA diffusion in slits suggests otherwise. To summarize briefly, early experiments on DNA confined in slits (Balducci et al., 2006) suggested that the scaling of the in-plane diffusivity is much weaker ($D \sim H^{0.5}$) than the scaling predicted by blob theory ($D \sim H^{2/3}$) (Brochard and de Gennes, 1977). Interestingly, this anamalous behavior was observed in the range of channel sizes where DNA exhibits the blob theory prediction of the extension parallel to the slit walls, $R_{\parallel} \sim H^{1/4}$. Subsequent Monte Carlo simulations demonstrated that this deviation in the scaling of the diffusivity from the blob theory prediction arises from the semiflexible nature of DNA, even though

semiflexibility does not appreciably affect the scaling in extension (Dai et al., 2013). Given the similarity of the behavior of properties of confined semiflexible polymers in slits and channels (Tree et al., 2014; Muralidhar et al., 2014a), one would expect stiffness to affect the diffusivity in the extended de Gennes regime for channel confinement. The behavior of the diffusivity in the extended de Gennes regime remains an open question, primarily due to lack of sufficient data.

Finally, a preliminary simulation study from our group recently by Tree et al. (2012) revealed a "Rouse-like" diffusion regime in between the Odijk and the extended de Gennes regimes, where the diffusivity is inversely proportional to the contour length of the chain but independent of its characteristic size (i.e., the extension of the chain). The diffusivity in this regime not only increases with stiffness, but the range of fractional extension spanned by this Rouse-like regime widens with increasing monomer anisotropy, l_p/w (Tree et al., 2012). These conclusions need to be treated as tentative, as the results from these simulations correspond to only 3 values of l_p/w and are limited to relatively low molecular weights. Although at the time of the study (Tree et al., 2012) the range of channel sizes for such Rouse-like behavior was thought to be a sharp transition between the Odijk and the blob regimes (Tree et al., 2012), it is now clear that this range of channel sizes is characterized by the backfolded Odijk regime, particularly for molecules with large values of l_p/w (Muralidhar et al., 2014a). The behavior of the static properties of a chain in the backfolded Odijk regime remain unconnected to these observations about chain diffusivity, in particular why the diffusivity appears to increase with stiffness.

Despite the fact that simulations have contributed to much of the progress in understanding the physics of static properties, the available simulation data on diffusivity of channel-confined chains are limited and indeed conflicting (Jendrejack et al., 2003; Tree et al., 2012), owing to the challenges involved in simulation methodology arising from the disparity in the relevant length scales of the molecule such as contour length, persistence length and width, and juxtaposition of other length scales such as the confinement size (Dorfman et al., 2014). Often, bead-spring models have been used to simulate the dynamics of such polymers in confinement via Brownian Dynamics (BD) simulations, both with (Jendrejack et al., 2003; Chen et al., 2007; Chen, 2013b) and without hydrodynamic interactions (HI) (Arnold et al., 2007). The coarse-grained nature of these models allows simulation of experimentally relevant molecular weights with a small number of beads for a sufficiently long time (Larson, 2005). Because such bead-spring models do not have the necessary resolution when the confinement size is of the order of the persis-

tence length of the polymer, finer models have been used recently for Metropolis Monte Carlo (Tree et al., 2012; Dai et al., 2013) and lattice Boltzmann-based simulations (Chen et al., 2014) of semiflexible chains in confinement. Models with adequate resolution used in conjunction with these methods, however, require hundreds or even thousands of beads to reach the length scales relevant to scaling theories and experiments, making such simulations infeasible.

In this contribution, we compute axial diffusivity of long semiflexible polymers over the entire range of channel confinement from the classic de Gennes regime to the classic Odijk regime. We achieve this by combining pruned-enriched Rosenbluth method (PERM) simulations (Prellberg and Krawczyk, 2004; Tree et al., 2013b) of a discrete wormlike chain model with computational fluid dynamics (CFD) calculations of the hydrodynamic tensor. We account for the effect of HI through the pre-averaging approximation in Kirkwood-Riseman theory (Kirkwood and Riseman, 1948), which has recently been shown to provide reliable estimates of diffusivity in channel confinement (Jain and Dorfman, 2015). By computing the diffusivity for asymptotically long chains for several values of l_p/w, we show that the blob theory proposed by Brochard and de Gennes (1977) is valid in the classic de Gennes regime. We further develop a modified blob theory inspired by the work of Dai et al. (2013) for the scaling of diffusivity in the extended de Gennes regime, and demonstrate that this theory captures the deviation of our simulation data from blob theory in this regime. Finally, we propose a simple model for the backfolded Odijk regime, which is consistent with the Rouse-like regime proposed by Tree et al. (2012). Our simulations and corresponding theoretical analysis address the three key open questions surrounding the diffusion of semiflexible polymers in channels.

8.2 Methods

Our objective here is to calculate the axial diffusivity or mobility of the center of mass of semiflexible polymers confined in square channels. The mobility tensor, $\boldsymbol{\mu}_\mathrm{CM}$, of the center of mass of a chain under an external force $\mathbf{F}_\mathrm{CM}$, with a resultant drift velocity of the center of mass $\mathbf{v}_\mathrm{CM}$, is defined by

$$\mathbf{v}_\mathrm{CM} = \boldsymbol{\mu}_\mathrm{CM} \cdot \mathbf{F}_\mathrm{CM}. \tag{8.1}$$

Note that the diffusivity tensor, $\mathbf{D}_\mathrm{CM}$, is related to the mobility tensor by the Einstein-Smoluchowski relation, $\mathbf{D}_\mathrm{CM} = \boldsymbol{\mu}_\mathrm{CM} k_\mathrm{B} T$, where k_B is the Boltzmann constant and T

is the absolute temperature. For our case of a polymer confined in a square channel, we cast the problem as a one-dimensional analogue along the axis of the channel. We are interested in the motion of the confined chain parallel to the x-axis, which points in the direction of the axis of the channel. Accordingly, the velocity of the chain in the x direction, v_x, due to a force parallel to the x axis, F_x, is determined by the xx component of the mobility tensor, μ, as

$$v_x = \mu F_x. \tag{8.2}$$

Since we are only concerned with center-of-mass diffusion, we drop the subscript 'CM' from this point forward. As μ and the xx component of the diffusion tensor, D, differ only by a factor of $k_\mathrm{B}T$, we use μ and D interchangeably in the rest of the paper. Additionally, in the subsequent discussion we confine ourselves to the diffusion of polymer chains near equilibrium. In other words, we assume that the conformations of the molecule are unperturbed on an average by the forces exerted by the solvent.

8.2.1 Model

To calculate the diffusivity and other thermodynamic properties of confined chains in PERM simulations, we use a variant of the discrete wormlike chain model that we employed previously for diffusivity calculations of DNA in free solution (Tree et al., 2012). In our model, the chain is discretized as a string of N_b touching beads of diameter d. This length scale corresponds to the local hydrodynamic diameter, d. The corresponding radius is a ($d = 2a$), which measures the distance between the center of the backbone and the surface of shear on the molecule. Our model, therefore, is essentially a more flexible analogue of the the 'shish-kebab' model that is often used to model rod-like polymers (Doi and Edwards, 1986). A bending potential quantified by

$$\beta U_\mathrm{bend} = -\kappa(1 - \cos\theta_i) \tag{8.3}$$

is imposed on all contiguous triplets of beads along the chain, where κ is the bending constant of the chain. The quantity $\cos\theta_i = (\mathbf{r}_{i+2} - \mathbf{r}_{i+1}) \cdot (\mathbf{r}_{i+1} - \mathbf{r}_i)$ for $i = 1, \ldots, N_b - 2$. From the bending potential, we can derive the average of $\cos\theta$ as $\langle\cos\theta\rangle = \coth(\kappa) - 1/\kappa$. Therefore the persistence length, $l_\mathrm{p} = d/(1 - \langle\cos\theta\rangle)$ (Yamakawa, 1971; Hiemenz and

Lodge, 2007), is related to the bending constant κ by

$$\frac{l_{\mathrm{p}}}{d} = \frac{\kappa}{\kappa - \kappa\coth(\kappa) + 1} \tag{8.4}$$

We associate a hard-core width, w, commensurate with the excluded volume strength of the segments of the chain. Furthermore, we assume that the width for intramolecular interactions is the same as the width for the polymer-wall interaction. This renders the size of the channel accessible to the center of the backbone to be $H_{\mathrm{eff}} = H - w$. For simplicity, we assume $d = w$. However, the two quantities are not necessarily identical as w corresponds to the length scale for excluded volume, while d represents the length scale for the friction due to the solvent. For instance, we recently found that a value of $d = 2.94\,\mathrm{nm}$, which is smaller than the value $w = 4.6\,\mathrm{nm}$ from Stigter's theory for DNA (Stigter, 1977), yields the best agreement between simulation results and experimental data for diffusivity of DNA in bulk solution (Tree et al., 2013b).

8.2.2 Kirkwood mobility

We adopt the Kirkwood preaveraging approximation (Kirkwood and Riseman, 1948; Yamakawa, 1971; Dorfman et al., 2014) to estimate the effect of HI on the diffusivity of confined molecules. In this approximation, the center of mass mobility of a chain of N_b beads is (Rodríguez Schmidt et al., 2012)

$$\mu_{N_b} = \frac{1}{N_b^2}\left\langle \sum_{i=1}^{N_b}\sum_{j=1}^{N_b}\left(\frac{\delta_{ij}}{6\pi\eta a} + \Omega_{ij}^{xx}(\mathbf{r}_i,\mathbf{r}_j)\right)\right\rangle, \tag{8.5}$$

where $\langle\ldots\rangle$ denotes an ensemble average over an equilibrium ensemble of chain configurations. The second term in Eq. 8.5 is the xx component of the hydrodynamic interaction tensor in confinement,

$$\mathbf{v}_i = \boldsymbol{\Omega}_{ij}(\mathbf{r}_i,\mathbf{r}_j)\cdot\mathbf{F}_j, \tag{8.6}$$

which measures the velocity perturbation in the solvent, $\mathbf{v}_i$, at a point $\mathbf{r}_i$ due to the point force, $\mathbf{F}_j$, applied on the solvent by a bead at $\mathbf{r}_j$. Note that the first term on the right hand side of Eq. 8.5, which accounts for self diffusivity, is required to capture the diffusivity at low molecular weights. However, the HI term dominates as $N_b \to \infty$, making the long-chain diffusivity insensitive to the choice of a (Zimm, 1980).

The Kirkwood approximation is known to introduce errors in the estimates of the

diffusivity for polymers in dilute solution, as it neglects intramolecular dynamic correlations (Wang et al., 1986; Liu and Dünweg, 2003; Mansfield et al., 2007; Rodríguez Schmidt et al., 2012). For this reason, the Kirkwood diffusivity is also sometimes referred to as the "short-time" diffusivity. However, recent work from our group indicates that the deviation in the value of the Kirkwood diffusivity from the "long-time" diffusivity given by Einstein's theory is suppressed in channel confinement (Jain and Dorfman, 2015). This result is encouraging in hindsight, as the Kirkwood approximation has been used extensively in scaling theories for polymer diffusion in confinement (Brochard and de Gennes, 1977; de Gennes, 1979; Tree et al., 2012; Dai et al., 2013).

8.2.3 CFD methodology

While our PERM simulations generate the configurations needed to compute the ensemble average in Eq. 8.5, we also need to evaluate the hydrodynamic tensor $\mathbf{\Omega}_{ij}$, which we carry out via a CFD solution of the Stokes equation for the solvent in the channel. To solve the Stokes equations, we use a method similar to that of Jendrejack et al. (2003).

We employ a three-step process to obtain the hydrodynamic tensor in the channel. First, we construct a three-dimensional, uniformly-spaced, staggered grid in the channel, consisting of pressure, x-velocity, y-velocity and z-velocity nodes, thus circumventing the need to explicitly specify the boundary conditions for pressure (Pozrikidis, 2009). The pressure grid points also serve as grid points to compute the velocity perturbations due to point forces in the channel. A unit force applied at one of these grid points $\mathbf{r}_j = (x_j, y_j, z_j)$ generates a velocity perturbation at another grid point $\mathbf{r}_i = (x_i, y_i, z_i)$, which can be quantified by the hydrodynamic tensor $\mathbf{\Omega}_{ij}$. In free solution, the hydrodynamic tensor is the Oseen tensor,

$$\mathbf{\Omega}_{\text{free}}(\mathbf{r}_i - \mathbf{r}_j) = \frac{1}{8\pi\eta}\left(\frac{\mathbf{I}}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{(\mathbf{r}_i - \mathbf{r}_j)(\mathbf{r}_i - \mathbf{r}_j)}{|\mathbf{r}_i - \mathbf{r}_j|^3}\right), \tag{8.7}$$

which can be used to evaluate the velocity perturbation, $\mathbf{u}_{\text{free}}(x_i, y_i, z_i)$. To compute the velocity perturbation in the channel, $\mathbf{u}(x_i, y_i, z_i)$, the calculation method (Jendrejack et al., 2003) exploits the linearity of the Stokes equations by splitting the velocity field as $\mathbf{u} = \mathbf{u}_{\text{free}} + \mathbf{u}_{\text{walls}}$. Here, $\mathbf{u}_{\text{walls}}(x_i, y_i, z_i)$ is the correction to the free solution velocity field that arises due to the presence of the channel walls. We obtain $\mathbf{u}_{\text{walls}}$ by numerically

solving the Stokes equations

$$\begin{aligned} -\nabla p + \eta \nabla^2 \mathbf{u}_{\text{walls}} &= 0, \\ \nabla \cdot \mathbf{u}_{\text{walls}} &= 0, \end{aligned} \tag{8.8}$$

subject to no-slip boundary conditions at the walls, i.e.,

$$\mathbf{u}_{\text{free}} + \mathbf{u}_{\text{walls}} = 0 \quad \text{at the walls.} \tag{8.9}$$

We compute the boundary velocities at grid points on the boundaries by calculating the flux of $-\mathbf{u}_{\text{free}}$ over each cell at the boundaries and then dividing by the total surface area of the cell. This method ensures mass conservation, which is not necessarily satisfied if one naively computes $-\mathbf{u}_{\text{free}}$ exactly at the boundary grid points.

Due to translational symmetry along the channel axis, the wall correction $\mathbf{\Omega}_{\text{walls}}(\mathbf{r}_i, \mathbf{r}_j) = \mathbf{\Omega}_{\text{walls}}(|x_i - x_j|, y_i, y_j, z_i, z_j)$. Therefore, the hydrodynamic tensor that determines the velocity perturbation at $\mathbf{r}_i$ due to a point force at $\mathbf{r}_j$ is given by

$$\mathbf{\Omega}_{ij}(|x_i - x_j|, y_i, y_j, z_i, z_j) = \mathbf{\Omega}_{\text{walls}}(|x_i - x_j|, y_i, y_j, z_i, z_j) + (1 - \delta_{ij})\mathbf{\Omega}_{\text{free}}(\mathbf{r}_i - \mathbf{r}_j) \tag{8.10}$$

We determine $\mathbf{\Omega}_{ij}$ for all pairs of grid points by successively solving Eq. 8.8 for point forces at each grid point $\mathbf{r}_i$, using a finite-difference method (FDM) with boundary conditions obtained from Eq. 8.9. Our FDM formulation uses forward differences for first derivatives and second-order central differences for second derivatives in Eq. 8.8 (Pozrikidis, 2009). Although in principle the channel is infinitely long along its axis, the length of the channel was fixed at $10H$ in these calculations (Jendrejack et al., 2003). This box size is sufficient to obtain all the relevant information about $\mathbf{\Omega}_{ij}$ because of the rapid decay of the magnitude of the hydrodynamic tensor $\mathbf{\Omega}_{ij}$ with $|x_i - x_j|$. In all our calculations, the domain is divided such that there are 72 grid points in the x direction and 18 grid points each in the y and z directions, regardless of the channel size. As a result, the total number of grid points in the domain is 23 328. The resulting set of linear equations generates a sparse matrix, which we solve using GMRES via an incomplete LU factorization. Thereafter, we compute each element of $\mathbf{\Omega}_{\text{walls}}$ following the method outlined in previous work (Jendrejack et al., 2003; Dorfman et al., 2014).

With the knowledge of $\mathbf{\Omega}_{\text{walls}}$ and $\mathbf{\Omega}_{\text{free}}$, we compute $\mathbf{\Omega}_{ij}$ for all combinations of $\mathbf{r}_i$ and $\mathbf{r}_j$ on our grid. We store the computed elements of the tensor $\mathbf{\Omega}_{ij}$ in a lookup table,

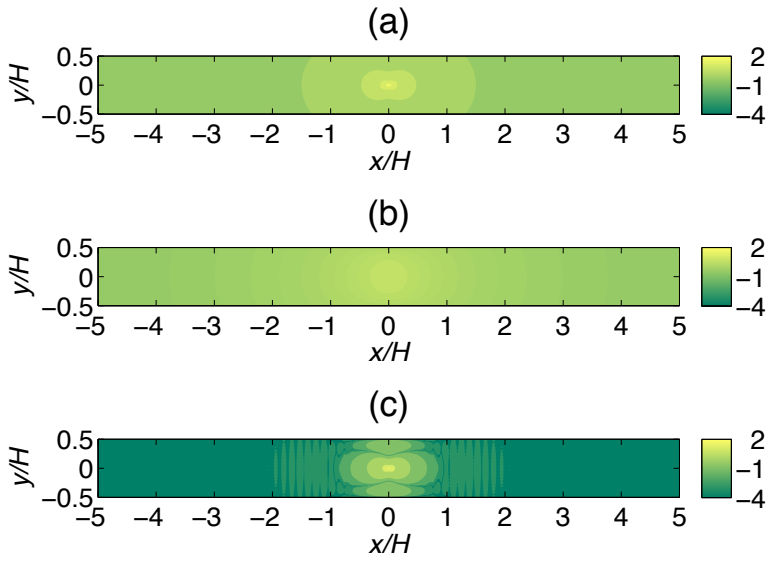


Figure 8.1. Magnitude of the xx component of the dimensionless hydrodynamic tensor due to a point force in the x-direction at $(x, y, z) = 0$ on the plane $z = 0$ in logarithmic scale: (a) free solution term, $6\pi\eta H|\Omega^{xx}_{\text{free}}|$, (b) wall correction, $6\pi\eta H|\Omega^{xx}_{\text{walls}}|$ and (c) resultant channel confinement term, $6\pi\eta H|\Omega^{xx}|$.

which is in turn used in our PERM simulations. Figure 8.1 shows a typical result from our CFD calculation. In particular, it illustrates the behavior of the magnitude of the xx component of the tensors, $\Omega^{xx}_{\text{free}}$, $\Omega^{xx}_{\text{walls}}$ and Ω^{xx} for a point force applied at $(0, 0, 0)$. The relatively low magnitude of Ω^{xx} in comparison to $\Omega^{xx}_{\text{free}}$ and $\Omega^{xx}_{\text{walls}}$, particularly for $|x/H| > 1$, points to the screening of HI as one traverses away from the perturbing force along the channel (Diamant, 2009).

8.2.4 Integrating CFD with PERM

We employ PERM, a biased chain growth method, to generate configurations of confined chains (Tree et al., 2013a; Muralidhar et al., 2014b; Muralidhar et al., 2014a). We start with a single bead placed randomly inside the channel and grow chains by adding a bead at a time, suitably enriching (branching) or pruning the growth as per Prellberg's algorithm (Prellberg and Krawczyk, 2004). Because we bias the chain growth to achieve growth of long chains, each configuration at each growth step is associated with a statistical weight based on the amount of bias. To compute a chain property, such as the Kirkwood diffusivity in Eq. 8.5, the bias is undone by weighting each configuration

appropriately. Unlike Grassberger's original algorithm (Grassberger, 1997), we do not need to set upper and lower threshold weights to guide the chain growth towards long chains. Prellberg's algorithm does not use any such parameters; the target weights for chain growth are obtained from performing preliminary PERM simulations. We use exactly the same algorithm as in previous work from our group and we refer the reader to earlier publications from our group for more details on PERM methodology (Tree et al., 2013a; Tree et al., 2014; Muralidhar et al., 2014b; Muralidhar et al., 2014a). Starting from a single bead, we grow the chain until it reaches a preset maximum number of beads or gets pruned with no more branches to grow. Such an excursion that starts from a single bead and grows till the end is known as a tour. In this work, we used PERM to grow chains up to 10 000 beads for a total of 50 000 tours in each simulation for a specific combination of (l_p, w, H) for calculations of diffusivity.

We compute the center-of-mass axial mobility from a given configuration by

$$\Omega^{xx}_{\mathrm{CM},N_b} = \frac{1}{N_b^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\frac{\delta_{ij}}{6\pi\eta a} + \Omega^{xx}_{ij}(|x_i - x_j|, y_i, y_j, z_i, z_j)\right). \tag{8.11}$$

The average axial mobility of the chain over the total number of tours, N_tours, is calculated on the fly during PERM chain growth as

$$\mu_{N_b} = \langle\Omega^{xx}_{\mathrm{CM},N_b}\rangle = \frac{\sum_{k=1}^{N_\mathrm{tours}} W_{N_b,k}\Omega^{xx}_{\mathrm{CM},N_b,k}}{\sum_{k=1}^{N_\mathrm{tours}} W_{N_b,k}}, \tag{8.12}$$

where $W_{N_b,k}$ is the weight of a chain with N_b beads in the kth tour. For the purpose of comparing our simulation results for diffusivity with scaling theory, employing the Kirkwood approximation does not imply any loss of generality, as the theoretical predictions themselves are based on the preaveraging approximation. Since the values of $\Omega^{xx}_{ij}(|x_i - x_j|, y_i, y_j, z_i, z_j)$ computed from our CFD simulations are tabulated only for grid points, we use multi-linear interpolation to estimate Ω^{xx}_{ij} for arbitrary values of $\mathbf{r}_i$ and $\mathbf{r}_j$. Given that we use the same number of grid points for all channel sizes, we restrict the channel sizes in our calculations to values between $H/d = 5$ and $H/d = 100$. The largest channel size, $H/d = 100$, is large enough that we can reach the classic de Gennes regime for chains with sufficiently low l_p/w, but does not introduce significant interpolation errors between the grid points. The double-sum in Eq. 8.11 makes our method $O(N_b^2)$ in computational time. We are able to run simulations for a confined chain of 10 000 beads for 50 000 tours on 60 processors within a span of 3 days. We only

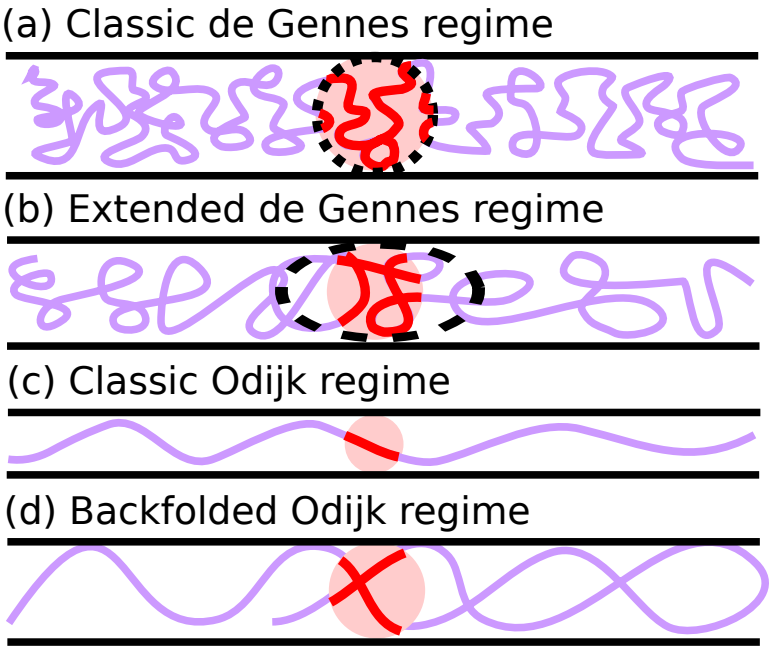


Figure 8.2. Schematic showing the conformation of the chain in different regimes. The red part of the chain in each of the channels represents the sub-chain within a hydrodynamic screening volume, which is shown as a red sphere. The black dashed lines are a representation of compression blobs in the blob regimes.

need to calculate the pairwise HI tensor for monomers within a cutoff length of $10H$ along the channel, which significantly reduces the computational time.

8.3 Summary of existing theories

As we discussed at the outset, our goal is to compare the results for diffusivity obtained from our PERM simulations to existing theories. Here, we briefly summarize these theories to facilitate our subsequent discussion.

8.3.1 Blob theory

Blob theory was originally proposed to describe the properties of confined polymers in the classic de Gennes regime (Brochard and de Gennes, 1977; de Gennes, 1979). In the classic de Gennes regime, the chain conformation consists of a series of linearly ordered compression blobs, each of which has a size proportional to the channel size, H. Figure 8.2(a) shows a typical conformation in the classic de Gennes regime. One can use the blob argument and write the scaling of the average extension in the classic de

Gennes regime as (Tree et al., 2013a; Muralidhar et al., 2014a)

$$\frac{X}{L} \simeq \left(\frac{H}{l_\mathrm{p}}\right)^{1-1/\nu} \left(\frac{w}{l_\mathrm{p}}\right)^{2-1/\nu}, \tag{8.13}$$

where $\nu = 0.587\ 597(7)$ is the Flory exponent (Clisby, 2010). Because the the intersegmental HI in confinement is screened beyond a length scale of H inside the channel (Diamant, 2009), one can estimate the axial mobility via the Kirkwood preaveraging approximation as (Brochard and de Gennes, 1977; de Gennes, 1979)

$$\mu \simeq \frac{1}{L} \int_0^{H/2} \Gamma(r) \Omega^{xx}(r) dr, \tag{8.14}$$

where $\Omega^{xx}(r)$ is the angle-average of the xx component of the Oseen tensor. Note that $\Omega^{xx}(r) = 1/6\pi\eta r$ is the identical to the angle-average of the whole Oseen tensor, $\Omega(r)$. The function $\Gamma(r) = 4\pi r^2 l_\mathrm{p} g(r)$ is a dimensionless form of the pair correlation function, $g(r)$, of Kuhn segments of the chain. In other words, the function $\Gamma(r)$ measures the average contour length of the chain present in a shell enclosed by radii r and $r + \Delta r$ as $\Delta r \to 0$. It turns out that the length scale for screening of HI, the hydrodynamic blob, also scales linearly with the channel size H (Figure 8.1). Note that the hydrodynamic blob and the compression blob are length scales that arise from different physical phenomena. In the classic de Gennes regime, however, both these length scales are proportional to H as shown in Figure 8.2(a).

In blob theory, the pair correlation function is estimated as if there is a gas of monomers within each blob, neglecting correlations along the backbone of the chain. In particular, we assume that the blob size is such that the local stiffness of the chain has very little effect on the pair correlation function inside a blob. Therefore, the pair correlation function is simply the concentration of monomers inside a blob, which is given by (Tree et al., 2012; Tree et al., 2013a)

$$g(r) \sim r^{1/\nu-3} l_\mathrm{p}^{2-2/\nu} w^{1/\nu-2} \tag{8.15}$$

Using eqs 8.14 and 8.15, the mobility in the classic de Gennes regime can be written as

$$\mu \simeq \frac{1}{6\pi\eta L} l_\mathrm{p}^{3-2/\nu} w^{1/\nu-2} H^{1/\nu-1}. \tag{8.16}$$

In terms of the mean extension, the scaling of μ becomes

$$\mu \simeq \frac{1}{6\pi\eta X} \text{ or } 6\pi\eta L\mu \simeq \left(\frac{X}{L}\right)^{-1}, \tag{8.17}$$

where we have used the well established scaling relation for the mean extension in the de Gennes regime given in Eq. 8.13. If we use the classical value of the Flory exponent, Eq. 8.16 assumes the form $6\pi\eta L\mu \simeq (H/l_\mathrm{p})^{2/3}\,(l_\mathrm{p}/w)^{1/3}$.

An equivalent way to determine the scaling law in Eq. 8.17 is by estimating the friction constant of the confined chain (Reisner et al., 2012). The use of Eq. 8.15 for the pair correlation implicitly assumes that sub-chain inside the blob is a swollen coil. Given that the subchain inside a blob is non-draining (Eq. 8.14), the friction of a blob scales with blob size as $\zeta_\mathrm{blob} \simeq 6\pi\eta H$. Due to screening of HI, the friction of the chain is the product of the friction of each blob and the number of blobs: $\zeta_\mathrm{chain} \sim 6\pi\eta H(X/H)$ or $\zeta_\mathrm{chain} \sim 6\pi\eta X$, where X/H is the number of blobs in the confined chain. The resultant scaling of the mobility, which is the inverse of the friction on the chain, is indistinguishable from that in Eq. 8.17.

It has been proposed that the blob theory outlined above can also be applied to the extended de Gennes regime (Tree et al., 2012; Reisner et al., 2012). Here, the static properties of the chain are described by a series of anisometric compression blobs each having a length proportional to $(Hl_\mathrm{p})^{2/3}\,w^{-1/3}$, as shown in Figure 8.2(b). The other two orthogonal dimensions of the blob are still proportional to H, because of the constraints imposed by the cross section of the channel. Accordingly, the extension scales exactly like Eq. 8.13 (Odijk, 2008), with a recent theory (Werner and Mehlig, 2014) suggesting that $\nu = 3/5$ gives the right scaling in the extended de Gennes regime. Nonetheless, one can distinguish between the two blob regimes using the fact that they exhibit dissimilar scaling in properties such as the variance of extension and the confinement free energy (Dai et al., 2014; Muralidhar et al., 2014a).

The scaling of the concentration of the monomers in the extended de Gennes regime is identical to that in the classic de Gennes regime owing to the similarity in the scaling of extension. Naively, we can then write the pair correlation in the extended de Gennes regime as the concentration of monomers within a blob, giving rise to $g(r) \sim r^{1/\nu-3} l_\mathrm{p}^{2-2/\nu} w^{1/\nu-2}$, which is again indistinguishable from Eq. 8.15. Figure 8.2(b) indicates that while the compression blob in the extended de Gennes regime is anisotropic, the size of the hydrodynamic blob is unchanged, as it only depends on the boundary

conditions on the solvent. Application of blob theory now leads to the scaling for the axial mobility given by eqs 8.16 and 8.17. We thus see that blob theory for the extended de Gennes regime predicts the same diffusivity as the classic de Gennes regime, i.e. that the diffusivity scaling is "extended" from large channels into the smaller ones characterizing the extended de Gennes regime.

8.3.2 Classic Odijk regime

In the classic Odijk regime, the chain consists of rod-like deflection segments linearly ordered in the channel (Muralidhar et al., 2014a). To estimate the total friction on the chain, we need to estimate the friction of the sub-chain in the screening volume H^3, which is shown in red in Figure 8.2(c). To do so, recall that the translational friction constant of a rod of length l and hydrodynamic diameter d for motion parallel to its axis is (Doi and Edwards, 1986)

$$\zeta_{\parallel}^{\text{rod}} = \frac{2\pi\eta l}{\ln(l/d)}. \tag{8.18}$$

Owing to tight confinement in the Odijk regime, the motion of the polymer is roughly that of a string of rods moving parallel to their axes. A scaling law for the friction constant for the chain can be obtained by summing up the friction on each of these rods of length H contained in the screening volume. Accordingly, the friction is (Morse, 1998)

$$\zeta_{\parallel} \simeq \frac{2\pi\eta X}{\ln(H/d)}. \tag{8.19}$$

Therefore, the axial diffusion coefficient in the Odijk regime scales as

$$\mu \simeq \frac{\ln(H/d)}{2\pi\eta X}. \tag{8.20}$$

Note that L is often used instead of X in eqs 8.19 and 8.20 (Reisner et al., 2012; Tree et al., 2012). This does not affect the scaling relations as X and L are related to each other by a factor of almost unity (Dorfman et al., 2014).

8.3.3 Backfolded Odijk regime

Recently, we showed (Muralidhar et al., 2014a) that the "transition regime" that bridges the Odijk and the de Gennes blob regimes is the backfolded Odijk regime predicted by Odijk (Odijk, 2008). Here, the chain contains deflection segments that can form

hairpins as long as the chain is sufficiently long to overcome the bending penalty, and excluded volume interactions between deflections segments are weak enough to allow for backfolding, as shown in Figure 8.2(d).

Our recent work also indicated that the upper limit in channel size for the backfolded Odijk regime is $H/l_\mathrm{p} \approx 2$ (Muralidhar et al., 2014a). While this limit in H is fixed, the lower limit in fractional extension corresponding to this channel size depends on l_p, w, and H. Specifically, the average fractional extension in this regime is given by

$$\frac{X}{L} \simeq c\xi^{1/3}, \tag{8.21}$$

where the scaling variable ξ is defined as

$$\xi \simeq \frac{g}{l_\mathrm{p}}\left(\frac{l_\mathrm{p}}{H}\right)^{5/3}\left(\frac{w}{l_\mathrm{p}}\right). \tag{8.22}$$

and $c = 0.92$ is a prefactor obtained from simulations (Muralidhar et al., 2014a). Here, g is the global persistence length, which is a measure of the average contour length between two hairpin turns along the chain (Muralidhar et al., 2014a; Odijk, 2006). The upper limit in fractional extension for this regime is determined by $\xi \approx 1$, which translates to a fixed upper limit in fractional extension, $(X/L)_\mathrm{max} \approx 0.92$ (Eq. 8.21), irrespective of the value of l_p/w. However, because g/l_p is a function of H/l_p only (Muralidhar et al., 2014a), the lower limit in fractional extension corresponding to $H/l_\mathrm{p} = 2$ scales as $(X/L)_\mathrm{min} \sim (w/l_\mathrm{p})^{1/3}$. In other words, the lower limit in X/L for the backfolded Odijk regime decreases with increasing l_p/w, although the upper limit is more or less fixed regardless of the value of l_p/w. Consequently, the range in fractional extension occupied by the backfolded Odijk regime increases with monomer anisotropy.

Previous work from our group by Tree et al. (2012) revealed a Rouse-like scaling for the axial diffusion coefficient in this regime. The premise behind the model was that the number of monomers in this regime is not high enough for substantial HI within the screening volume (Tree et al., 2012). This Rouse-like behavior inside the screening volume leads to a Rouse-like behavior for the whole chain, with a diffusion coefficient that scales as

$$\mu \sim \frac{1}{\eta L}. \tag{8.23}$$

An alternate way to write Eq. 8.23 is (Tree et al., 2012)

$$\mu\eta L \sim \left(\frac{X}{L}\right)^0, \tag{8.24}$$

which explicitly expresses that the mobility is independent of the extension of the molecule, This is of course contrasting behavior to that of the blob regimes, where the mobility is inversely proportional to the average extension (Eq. 8.17).

8.4 Results and Discussion

All of the theories discussed above apply to long chains. PERM natively produces data as a function of chain length via Eq. 8.12. A necessary, but not especially interesting, first step in our analysis is to confirm that the data we are analyzing indeed correspond to the long-chain limit. In the Supporting Information, we provide data and explanatory text confirming that the data presented here correspond to $\mu \sim 1/L$ regardless of the regime of confinement for long enough chains. In what follows, we restrict ourselves to examining the scaling of the asymptotic, long-chain diffusivity with respect to other quantities such as the channel size and fractional extension.

8.4.1 Blob regimes

Figure 8.3(a) shows the dimensionless mobility plotted against the fractional extension, X/L, which is also obtained from simulations, for channel sizes $H_{\text{eff}}/l_{\text{p}} > 3$. This is the range of channel sizes where we observe agreement with the extension scaling from blob theory (Eq. 8.13) (Muralidhar et al., 2014a). Note that each data point for each value of l_{p}/w corresponds to a distinct channel size. For the plot of normalized mobility against dimensionless channel size, we refer the reader to the Supporting Information. Our data in Figure 8.3(a) show good agreement with the prediction from blob theory (Eq. 8.17) for $X/L < 0.1$ for the range of l_{p}/w values used in our simulations. However, for higher fractional extension values (narrower channels), we observe a deviation from the predicted -1 slope. For a given value of fractional extension, this deviation grows with increasing monomer anisotropy, suggesting that blob theory fails to explain the diffusivity for relatively stiff chains. Indeed, the departure of our data from the blob theory prediction for stiff chains is consistent with diffusivity studies of DNA confined in slits (Dai et al., 2013; Balducci et al., 2006).

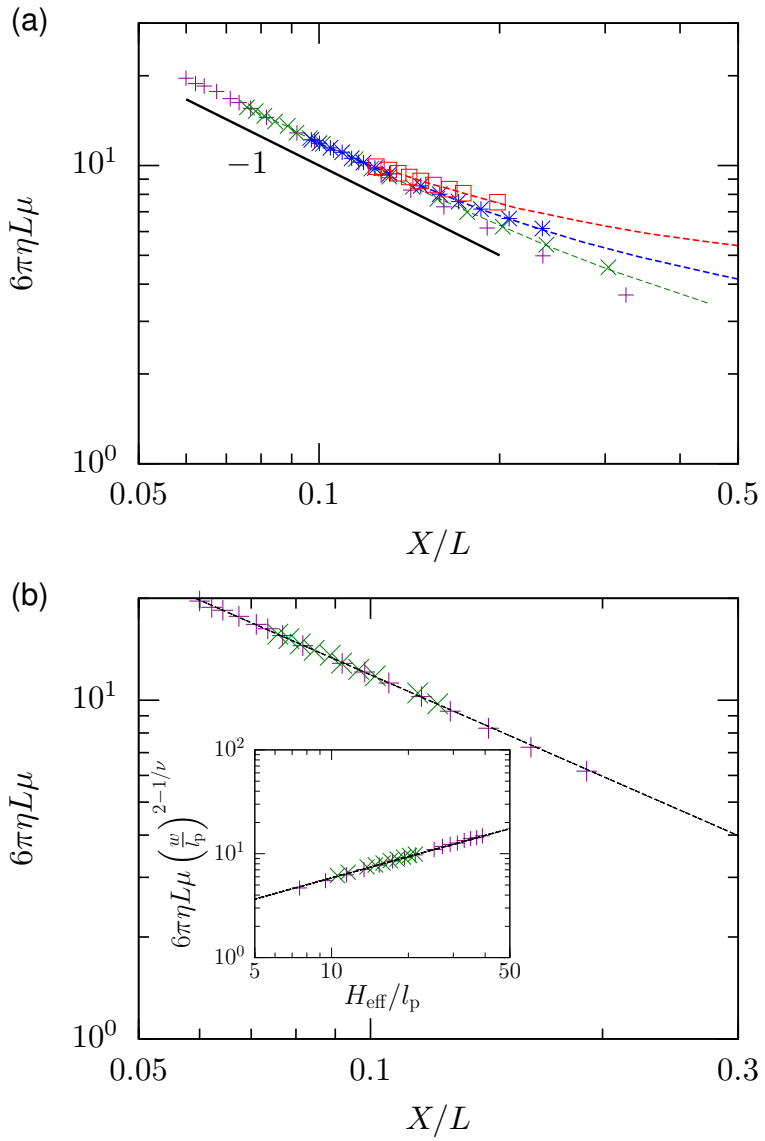


Figure 8.3. Mobility data in the blob regimes versus fractional extension. The l_p/w values of the molecules shown here are 2.54 (purple +), 4.65 (green ×), 8.95 (blue ∗), 17.70 (red □). The standard error of the mean for all the data points is smaller than the symbol size (see Supporting Information). (a) Comparison to the slope predicted from blob theory and modified blob theory to account for stiffness. The dashed lines are best fit curves using Eq. 8.27 with $\beta = 0.54$, $\delta = 1.12$ (green curve); $\beta = 0.62$, $\delta = 1.01$ (blue curve); $\beta = 0.71$, $\delta = 0.87$ (red curve). There is no purple dashed line because all of the data for $l_p/w = 2.54$ correspond to the classic de Gennes regime. (b) Power law fits to the data points belonging to the classic de Gennes regime. Data in the main panel can be fit to the expression, $6\pi\eta\mu L = 1.2(X/L)^{-0.995}$. The inset shows the same data against H_{eff}/l_p, for which a fit yields $6\pi\eta\mu L(w/l_p)^{-0.298} = 1.23(H_{\mathrm{eff}}/l_p)^{0.68}$.

To ascertain whether blob theory produces the correct prediction in the classic de Gennes regime, we plot the subset of our data that correspond to the scaling of the classic de Gennes regime with respect to static properties such as extension and confinement free energy. To identify this subset, we use the condition furnished by Dai et al. (2014), $H_{\text{eff}} > 2.24 l_{\text{p}}^2/w$. Figure 8.3(b) shows the collapse of mobility when plotted against fractional extension in the classic de Gennes regime. A fit indicates an exponent of −0.995, which is extraordinarily close to the exponent from blob theory (Eq. 8.17). Instead, if we plot the same data against $H_{\text{eff}}/l_{\text{p}}$ as shown in the inset of Figure 8.3(b), we obtain a fit that suggests $\mu\eta L \sim (H_{\text{eff}}/l_{\text{p}})^{0.68}$, again showing agreement with the 0.70 exponent from blob theory (Eq. 8.16).

Our results for diffusivity show excellent agreement with blob theory provided that the confinement corresponds to the classic de Gennes regime. This means that the data points inconsistent with blob theory in Figure 8.3(a) correspond to the extended de Gennes regime, as we have plotted only data points with $H_{\text{eff}} > 3l_{\text{p}}$. To explain the the behavior of mobility in the extended de Gennes regime, we propose an alternate theory along the lines of the modified blob theory for DNA in slit confinement (Dai et al., 2013). This modification to blob theory (Dai et al., 2013) supposes that the difference between the mobility and the prediction from the classic blob theory arises from the neglect of the stiffness of semiflexible polymers in describing the pair correlation function in Eq. 8.15 for short distances along the contour length. We use the modified form of the dimensionless pair correlation function proposed by Dai et al. (2013)

$$\Gamma(r) = \begin{cases} 2 & \text{if } r \le l_{\text{p}}/2, \\ \alpha r^{1/\nu-1} l_{\text{p}}^{3-2/\nu} w^{1/\nu-2} & \text{if } l_{\text{p}}/2 < r < H_{\text{eff}}/2, \end{cases} \tag{8.25}$$

where α is the proportionality constant for the scaling of the pair correlation function. The factor of 2 ensures that the contour length inside a sphere of radius $r = l_{\text{p}}/2$ is $\int_0^{l_{\text{p}}/2} 2dr = l_{\text{p}}$. For channel sizes $H_{\text{eff}} > l_{\text{p}}$, the modified blob theory now yields

$$\mu \simeq \frac{1}{L}\left[\int_0^{l_{\text{p}}/2} \frac{2}{6\pi\eta r} dr + \int_{l_{\text{p}}/2}^{H_{\text{eff}}/2} \frac{\alpha r^{1/\nu-1} l_{\text{p}}^{3-2/\nu} w^{1/\nu-2}}{6\pi\eta r} dr\right]. \tag{8.26}$$

In the latter, we have assumed the concentration of monomers as an approximation to the pair correlation function, and employed the angle-averaged Oseen tensor to account for HI. This is acceptable because the orientational order parameter, which measures

the alignment of the chain along the axis of the channel, is nearly zero for channels with $H \gtrsim l_\mathrm{p}$ (Muralidhar et al., 2014a). Furthermore, as noted by Dai et al. (2013), this modification to the pair correlation does not affect the scaling of extension significantly (Muralidhar et al., 2014a), but leads to an appreciable change in the scaling of diffusivity. This behavior of the diffusivity can be explained by the factor of $1/r$ in the integral in Eq. 8.26, which amplifies the effect of correlations arising from stiffness at short distances along the chain (Dai et al., 2013).

Computing the integrals in Eq. 8.26 results in

$$\mu = \frac{1}{6\pi\eta L}\left[2\beta\ln(l_\mathrm{p}/a) + \delta\left(l_\mathrm{p}^{3-2/\nu}w^{1/\nu-2}H_\mathrm{eff}^{1/\nu-1} - l_\mathrm{p}^{2-1/\nu}w^{1/\nu-2}\right)\right], \tag{8.27}$$

where we have regularized the first integral in Eq. 8.26 (Dai et al., 2013). Notably, Dai et al. (2013) used a single constant δ in their work on slit confinement. Considering the approximations involved, it is not obvious that the modified blob theory leads to an exact prefactor in the logarithmic term in Eq. 8.27. Accordingly, we use the two constants β and δ as fitting parameters to correct for the appoximations such as the sharp crossover of $\Gamma(r)$ and the use of the free-solution Oseen tensor. As a result, these constants are not universal and are functions of l_p/w. Note that modified blob theory for slits (Dai et al., 2013) was only developed for a single value of l_p/w corresponding to DNA, so the universality (or lack thereof) of the constants was speculated upon but not addressed in the latter work. Figure 8.3(a) shows that Eq. 8.27 satisfactorily captures the essential physics giving rise to the deviation of mobility away from blob theory for $l_\mathrm{p}/w = 4.65, 8.95, 17.70$. The second term in the sum present in Eq. 8.27 is identical to the result from blob theory, and it is the only term dependent on H (Eq. 8.16). In contrast, the other two terms in Eq. 8.27 originate from the stiffness of the chain; the first term, $2\beta\ln(l_\mathrm{p}/a)$, is greater than the third term, $\delta l_\mathrm{p}^{2-1/\nu}w^{1/\nu-2}$. The relative values of these terms elucidate why chains with higher l_p/w exhibit bigger deviation from the blob theory value. Furthermore, Eq. 8.27 also indicates that we eventually recover the scaling expected from blob theory for sufficiently wide channels (i.e., in the classic de Gennes regime), as the second term dominates for $H \gg l_\mathrm{p}$.

Having explained the deviation from the blob theory prediction in the extended de Gennes regime, let us turn our attention again to the scaling in the classic de Gennes regime and compare our results to previous studies of chain mobility in this regime. In

Figure 8.3(b), we find that

$$6\pi\eta\mu = 1.2/X \tag{8.28}$$

is a good fit for our simulation data. The prefactor furnished by our simulations allows us to compare with the prefactor one would expect from scaling theory. Recall that the diffusivity for real flexible chains in the bulk, D_{bulk}, is

$$D_{\text{bulk}} = \frac{1}{12.07}\frac{k_{\text{B}}T}{\eta R_{\text{g}}}, \tag{8.29}$$

where R_{g} is the radius of gyration in free solution (Oono and Kohmoto, 1983). Therefore, the friction constant in bulk is given by $\zeta_{\text{bulk}} = 12.07\eta R_{\text{g}}$. de Gennes' blob theory suggests that the sub-chain inside the blob behaves as if it is a chain in bulk, unperturbed by the presence of walls. If this theory was quantitatively correct, the friction constant for a blob is $\zeta_{\text{blob}} = 12.07\eta(H_{\text{eff}}/2)$. Now the total friction of the confined chain can be written as

$$\zeta = \frac{12.067\eta X}{2} = 6.03\eta X, \tag{8.30}$$

by multiplying the friction of the blob by the number of blobs. Therefore, the rescaled mobility of the chain from Eq. 8.30 is $\mu\eta X = 0.17$. In comparison to the fit obtained in our simulations in Eq. 8.28, $\mu\eta X = 0.064$, the prefactor estimated from bulk behavior is remarkably close. Indeed, if we write the actual friction of the chain as $\zeta = c(6.03\eta X)$, c is just a $O(1)$ prefactor that presumably accounts for the errors due to chain-wall friction and the approximations in the bulk diffusivity calculation by Oono and Kohmoto (1983). The existence of these $O(1)$ prefactors for both dynamic and static properties (Tree et al., 2013a; Muralidhar et al., 2014b), points to the extraordinary accuracy of de Gennes' blob theory applied to confinement of flexible chains in the classic de Gennes regime.

The inset of Figure 8.3(b) reveals a dependence on H_{eff} given by $D\eta L \sim H^{0.68}$. However, previous studies have questioned this scaling of D with H citing the effect of polymer-wall interactions. Using a coarse-grained bead-spring model, Jendrejack et al. (2003) found that $D \sim H^{1/2}$, although their data for mean extension reveals that $X \sim H^{-2/3}$, in agreement with blob theory. This discrepancy was attributed to the finite value of the ratio a/Δ, where a is the segment hydrodynamic radius and Δ is the length-scale for polymer-wall interactions. Their claim was further substantiated by a comparison of their results with that of Harden and Doi (1992), who suggested that $D \sim H^{0.61}$, a slightly weaker scaling as opposed to de Gennes' blob theory (Eq. 8.16). Harden and

Doi (1992) obtained this scaling by combining the Kirkwood pre-averaging approach to estimate transport properties, with the self-consistent field method to calculate the monomer distribution in the presence of excluded volume (Harden and Doi, 1992).

In our work, which produced mobility data for very long chains with very high spatial resolution of the chain configurations and without any self-consistent field approximations, we find no evidence supporting the conclusions of Jendrejack et al. (2003) or Harden and Doi (1992) concerning the scaling of D with H. In the case of the work of Jendrejack et al. (2003) on diffusivity of DNA, one could argue that the exponent of 1/2 observed in their simulations might indicate an apparent exponent for the extended de Gennes regime. Yet it is unlikely that their model captures the physics of the extended de Gennes regime as the springs in their bead-spring model are about 40 persistence lengths long (Jendrejack et al., 2003). Rather, this points at a possible shortcoming in the resolution of such bead-spring models especially when they are employed in problems where stiffness effects are important. With respect to the results from Harden and Doi (1992), we note that the applicability of self-consistent field theory (SCFT) for obtaining monomer distributions in channel confinement has been called into question recently (Werner et al., 2013). Thus, we believe that the recent objections on the application of SCFT (Werner et al., 2013) and a superior high-resolution model used in our work in comparison to the highly coarse-grained model of Jendrejack et al. (2003), explains the aforementioned discrepancies with respect to the scaling of diffusivity in the classic de Gennes regime.

8.4.2 Backfolded and classic Odijk regimes

We now examine the the behavior of diffusivity outside the blob regimes. Figure E.3(a) shows dimensionless mobility plotted against fractional extension. To begin, we observe that the data points in Figure E.3(a) approach the blob theory prediction for $X/L \lesssim 0.1$, as we discussed in section 8.4.1. However, the points on the extreme right of Figure E.3(a) for $X/L \gtrsim 0.8$ correspond to the classic Odijk regime (Muralidhar et al., 2014a), as the curve generated using Eq. 8.20 passes right through the data points in this region. Our results thus confirm the the scaling relation in the classic Odijk regime given by Eq. 8.20. Our simulations also allow us to study the contour length dependence of mobility, in addition to the the calculation of the prefactor in Eq. 8.20 (see Supporting Information).

For large values of l_p/w, the region between the extended de Gennes regime and the

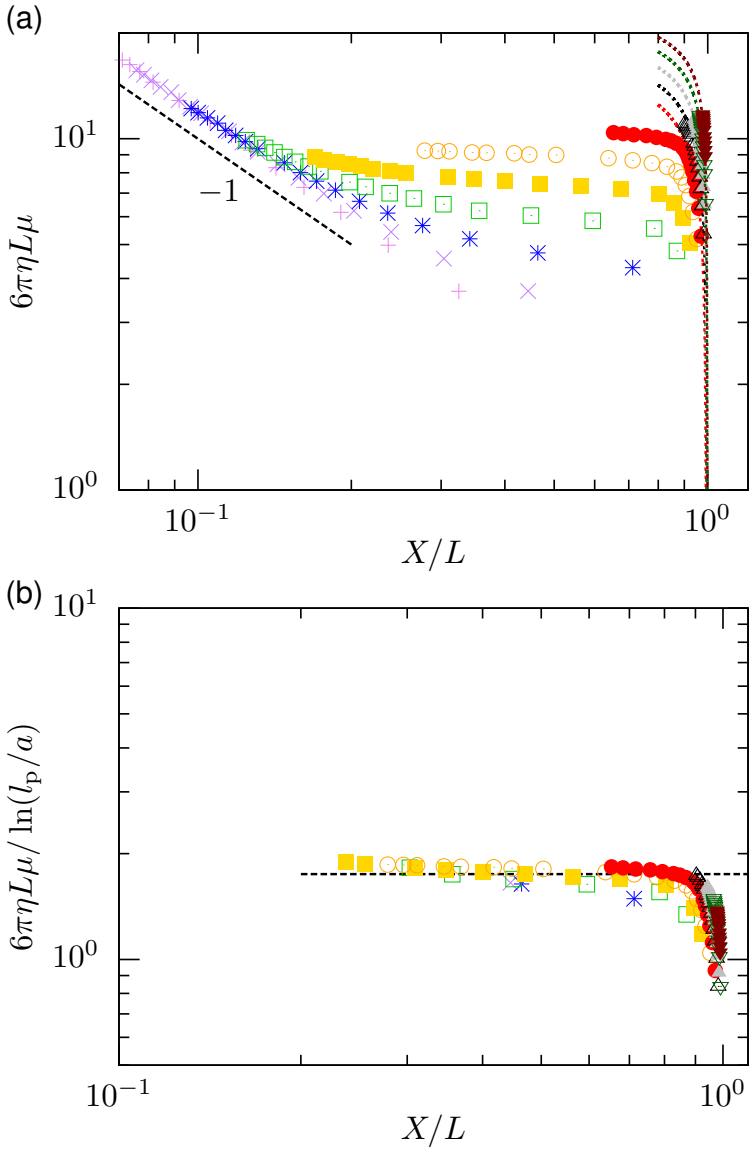


Figure 8.4. Normalized mobility as a function of fractional extension. The l_p/w values are 2.54 (violet +), 4.65 (purple ×), 8.95 (blue), 17.70 (green □), 35.54 (gold ■), 71.85 (orange ○), 145.8 (red •), 296.5 (black △) and 603.2 (gray ▲), 1228 (dark-green ▽), 2500 (brown ▼). The standard error of the mean for all the data points is smaller than the symbol size (see Supporting Information). (a) The dotted lines show the expected behavior in the classic Odijk regime. The dashed line shows the -1 slope expected in the classic de Gennes regime. (b) Subset of the data for channels with $H_{\text{eff}} < 2l_p$ showing collapse of mobility data in the backfolded Odijk regime. The black dashed line represents $6\pi\eta\mu L = 1.75(X/L)^0$.

classic Odijk regime in Figure E.3(a) is characterized by a value of the dimensionless mobility that is almost independent of fractional extension (or channel size). The dependence of the mobility on channel size is shown in the Supporting Information. Our results are therefore consistent with the observation of Tree et al. (2012). However, the crossover point in X/L from the de Gennes scaling depends on the l_p/w ratio. Chains with low l_p/w values (flexible chains) exhibit the de Gennes scaling out to higher values of X/L, while chains with high l_p/w transition into the $\mu\eta L \sim (X/L)^0$ scaling quickly. This insight was not possible using the data from Tree et al. (2012), as they only considered 3 relatively flexible values of l_p/w.

We posit that the backfolded Odijk regime (Odijk, 2008; Muralidhar et al., 2014a) can explain the insensitivity of the data to channel size for this range of channel sizes. Noting that the confined chain in the backfolded Odijk regime consists of a one-dimensional random walk of deflection segments of length $\lambda \simeq H^{2/3} l_p^{1/3} \approx l_p$ for $H \approx l_p$, we can write the diffusivity of the chain as

$$\mu \simeq \frac{1}{L} \int_0^{H/2} \Gamma(r) \Omega^{xx}(r) dr \approx \int_0^{l_p/2} \frac{2}{6\pi\eta r} dr \approx \frac{1}{3\pi\eta L} \ln(l_p/a), \tag{8.31}$$

along the same lines of the modified blob theory for the extended de Gennes regime. Note that using the angle-averaged Oseen tensor is still justified as the aligment in minimal when $H \approx l_p$, indicated by the relatively modest value of the order parameter (Muralidhar et al., 2014a). Indeed, the region around $H \approx l_p$ corresponds to the range of channel sizes where there is a steep change in fractional extension from about 0.2 to 0.8 for the l_p/w values considered here (Muralidhar et al., 2014a). However, we should bear in mind that the angle-averaged Oseen tensor will become a poor approximation as we approach channels with $H \ll l_p$. We know from our previous work that chains with a moderate value of l_p/w, such as DNA ($l_p/w \approx 10$), exhibit backfolded Odijk behavior for a very narrow range of channel sizes between $1 \lesssim H_{\text{eff}}/l_p \lesssim 2$ (Muralidhar et al., 2014a). Only chains with extremely high l_p/w values, such as actin and microtubules (l_p/w = 2375 and 9600 respectively) could backfold substantially for $H_{\text{eff}} \ll l_p$, if they are available at a sufficiently high molecular weight (Muralidhar et al., 2014a). As a result, the expression in Eq. 8.31 is not a universal scaling law for the backfolded Odijk regime, as chains with very high values of l_p/w still show behavior characteristic of this regime even for $H \ll l_p$. In such cases, the mobility will scale more like the classic Odijk regime (Eq. 8.20), as the alignment of the chain along the channel dictates the hydrodynamics. However, for

readily available semiflexible polymers such as DNA and actin, for which the backfolded Odijk regime corresponds to channels with $H \approx l_p$, we expect Eq. 8.31 to be a reasonable explanation for the behavior of diffusivity.

We verify our prediction in Eq. 8.31 for channels with $H_{eff} \approx l_p$, by plotting the subset of our data for $H_{eff} < 2l_p$ in Figure E.3(b) as this roughly represents the upper limit for the backfolded Odijk regime (Muralidhar et al., 2014a). Note that the lower limit in channel size for the backfolded Odijk regime corresponds to $\xi \approx 1$ (Eq. 8.21), whereupon there is a transition to the classic Odijk regime for narrower channels. To make this clear, we also plot the data from the classic Odijk regime in Figure E.3(b) to show the crossover between the two Odijk regimes. We observe that data points for chains with different l_p/w ratios fall essentially on the same horizontal line for $0.2 \lesssim X/L \lesssim 0.8$. Additionally, our data in this region indicate that

$$6\mu\pi\eta L \approx 1.75 \ln (l_p/a) \tag{8.32}$$

suggesting that the simple scaling arguments used to derive Eq. 8.31 give rise to nearly accurate prefactors. Because of the approximate nature of Eq. 8.31, we do not expect absolute collapse of the data in Figure E.3(b), in contrast to the classic de Gennes regime (Figure 8.3). Along the same lines, the plateau in Figure E.3(b) is not perfectly horizontal. But the model in Eq. 8.31 is a simple and useful explanation of diffusivity for highly stiff chains when $H \approx l_p$.

We further observe that range of X/L spanned by the plateau of the normalized mobility in Figure E.3 widens with increase in l_p/w. Recall from section 8.3.3 that the range of fractional extension spanned by the backfolded Odijk regime widens with increasing monomer anisotropy. Thus the window of fractional extension spanned by the mobility plateau in Figure E.3 is consistent with what we would expect in the backfolded Odijk regime.

8.5 Conclusion

By computing the Kirkwood diffusivity of semiflexible chains for molecules over many decades of l_p/w confined in channels spanning a wide range of sizes, we mapped out the diffusion behavior from the Odijk regime of strong confinement to the classic de Gennes regime of weak confinement as summarized in Table 8.1. We were able to show conclusively that blob theory correctly predicts the diffusivity in the classic de Gennes

Table 8.1: Summary of the behavior of axial diffusivity of semiflexible polymers in channel confinement. Prefactors from simulations are also included.

Regime	$6\pi\eta\mu L$
Classic de Gennes ($H \gtrsim l_\mathrm{p}^2/w$)	$\approx 1.2\left(\frac{H}{l_\mathrm{p}}\right)^{1/\nu-1}\left(\frac{w}{l_\mathrm{p}}\right)^{1/\nu-2}$
Extended de Gennes ($H \lesssim l_\mathrm{p}^2/w$)	$\approx 2\beta\ln(l_\mathrm{p}/a) + \delta\left(l_\mathrm{p}^{3-2/\nu} w^{1/\nu-2} H^{1/\nu-1} - l_\mathrm{p}^{2-1/\nu} w^{1/\nu-2}\right)$
Backfolded Odijk ($H \approx l_\mathrm{p}$ and $\xi \ll 1$)	$\approx 1.75\ln\left(\frac{l_\mathrm{p}}{a}\right)$
Classic Odijk ($H \ll l_\mathrm{p}$ and $\xi \gtrsim 1$)	$\approx 2.44\left(\frac{L}{X}\right)\ln\left(\frac{H}{d}\right)$

regime ($H \gtrsim l_\mathrm{p}^2/w$). However, we find that blob theory needs to be modified to account for local stiffness of the chain in order to explain the scaling of the diffusivity in the extended de Gennes regime. Thus, the behavior of diffusivity in the blob regimes mirrors the behavior in slits for commesurate confinement sizes.

Likewise, we also demonstrated that a similar theory predicts the insensitivity of the diffusivity to channel size (or fractional extension) in the backfolded Odijk regime when $H \approx l_\mathrm{p}$. Although this regime spans a very narrow range in channel size for most practical circumstances like channel-confined DNA, it is extremely important in practice due the steep change in fractional extension from about 0.2 to 0.8 for molecules such as DNA. Our previous work has shown that this steep change in extension is due to the exponential dependence of the global persistence length on the channel size (Muralidhar et al., 2014a). Based on the successful analogies between diffusion in the extended de Gennes regime in channels and the corresponding slit heights in slits (Dai et al., 2013), we expect to observe similar Rouse-like diffusion behavior for slit heights $H \approx l_\mathrm{p}$, where the diffusivity is insensitive to the channel size. Nonetheless, we do not expect this region to be very important in terms of the range of fractional extension spanned for slits. Explicitly, the range of fractional extensions when $H \approx l_\mathrm{p}$ should be extremely narrow, because unlike the case of channels, the global persistence length in a slit does not vary exponentially with slit height and has an upper bound of $2l_\mathrm{p}$ in slits (Tree et al., 2014; Muralidhar et al., 2014a). Taken together, we believe our work will prove useful in further understanding the dynamics of confined semiflexible polymers and serve as a valuable reference when single molecule experimental data become available for DNA diffusivity in nanochannels.

Chapter 9

Conclusions and Perspective

As stated at the outset, our goal was to gain a better understanding of the fundamental physics underlying confinement of semiflexible polymers, such as DNA. Much of this disseration has focused on the problem of backfolding and hairpin formation in nanochannel-confined semiflexible polymers. Hairpin formation is highly undesirable in nanochannel-based genome mapping technology, as it scrambles genomic information. Knowledge of the parameters governing the backfolded Odijk regime can be used to engineer conditions such as channel size and ionic strength to prevent hairpin formation. Our comprehensive study of the near-equilibrium dynamics of the chain confirmed classic theories, such that of Brochard and de Gennes (1977), while clarifying some of the discrepancies and confusions surrounding the diffusivity of channel-confined polymers that existed in literature.

Perhaps the most significant by-product of our work, which mostly focused on long-chain behavior, are our findings on the molecular weight dependence on various physical properties of the confined chain. Unlike traditional methods such as molecular dynamics, Brownian dynamics and Metropolis Monte Carlo simulations, PERM allows us to calculated properties as a function of molecular weight out to enormous length scales in a single simulation. We were able to achieve this by carefully parallelizing and optimizing different parts of the PERM code. We also employed customized neighbor lists and growth strategies tailor-made for simulating the off-lattice wormlike chain model. We hope others borrow our methodology to solve other similar problems involving semiflexible polymers.

Our findings in Chapter 4 remind us to be wary of finite chain length artifacts as we study the physics of semiflexible polymers in confinement. Although long-range ex-

cluded volume effects are suppressed by the walls of the channels (Tree et al., 2013b), one has to be cautious in interpreting results from simulations and experiments, especially when dealing with short chain lengths. Dynamic properties such as the diffusion coefficient (Tree et al., 2012), and the relaxation time in confinement (Tree et al., 2013c) are also affected by finite chain length (Tree et al., 2013b). Likewise, short length-scale effects arising from the semiflexible nature of DNA have been shown to affect diffusivity in both slit-confinement (Dai et al., 2013) and channel-confinement (Tree et al., 2012). Recognizing the need to carefully analyze the finite-length effects in the diffusivity of the confined chain, Chapter 8 and Appendix E are devoted to addressing the effect of molecular weight on diffusivity, while also examining long-chain behavior.

While we feel that the Odijk theory (Odijk, 2008) is was validated in Chapter 5, this is not the first time that the problem of channel-confined wormlike chains has been declared nearly "solved." Naturally, the Odijk theory still requires an experimental test, and this may not be an easy task. Aside from the challenges we highlighted already concerning the narrowness of the backfolded Odijk regime and the need to use very high molecular weight chains, most experimental systems use rectangular channels due to fabrication limitations or the desire to increase experimental throughput via tapered nanochannels (Persson et al., 2009). We recently found that such channels can exhibit mixing of regimes, where the two channel dimensions lie in different regimes and thus affect the chain statistics differently (Gupta et al., 2014). We expect that the behavior of g in this case will be a non-trivial function of the aspect ratio of the channel, and the narrowness of the backfolded Odijk regime will make it difficult to isolate this regime. Thus, while simulation and theory now seem to be in accord, substantial challenges remain in the experimental area.

Although our results in Chapter 6 support the use of Odijk's theory for predicting the extension in general or a simple polynomial fit of the data to model the extension for a particular data set, the ideas used in the S-loop model such as nucleation and propagation of hairpins may be useful to describe the kinetics of folding and unfolding of DNA (Levy et al., 2008; Alizadehheidari et al., 2015) in the backfolded Odijk regime. Finally, despite the experimental difficulties associated with testing Odijk's Flory theory, due to the narrow range of channel sizes and the high monomer anisotropy required for the backfolded Odijk regime, we hope more experimental data becomes availabile for relatively stiffer chains such as RecA-DNA (Alizadehheidari et al., 2015) or actin (Köster et al., 2007).

The simulation approach we used in Chapter 7 could be extended to test the phase diagram proposed by Werner and Mehlig (2015) for confinement in rectangular channels in the blob regimes ($D > 4l_{\mathrm{p}}$). They proposed the existence of 7 different regimes of behavior depending on both the smallest length scale D and the aspect ratio A/D (Werner and Mehlig, 2015). While simulating the polymer statistics over this range of channel sizes seems feasible, we again may require very long chains in order to satisfy the strong inequalities that distinguish the different scaling laws. In this case, it may be preferable to switch from our touching-bead model to a bead-rod model, as the latter model has proven effective for simulating wormlike chain confinement in the blob regimes (Dai et al., 2014). Our results here complete the verification of the phase diagram for rectangular channels in the backfolded Odijk regime (and others (Burkhardt et al., 2010) have done equivalent work for the classic Odijk regime), and we are optimistic that the remainder of the phase diagram for confinement in rectangular channels (Werner and Mehlig, 2015) will be verified via simulations in the near future.

Chapter 8 employs a combination of PERM and CFD simulations to calculate the Kirkwood diffusivity of confined chains for the entire range of confinement. From a methodological standpoint, combining PERM simulations with CFD calculations of the hydrodynamic tensor demonstrates the capability of PERM in estimating the diffusivity of polymers in confined geometries. Due to the general nature of our method, it can extended to compute the dynamic properties of single molecules confined in other geometries such as nanoslits (Dai et al., 2013) and triangular channels (Reinhart et al., 2013). Certainly, another area of interest would be to adopt this method in rectangular channels (Werner and Mehlig, 2015), where the 'regime mixing' observed with respect to static properties (Gupta et al., 2014), in combination with the asymmetric nature of hydrodynamics, will lead to complicated behavior in the dynamic properties. Furthermore, the asymptotic friction factor obtained in our simulations can be used in conjunction with the variance of extension to obtain estimates of the longest relaxation time of confined molecules using the dumbbell model (Tree et al., 2013c).

We would like to point out that our work considered a neutral polymer model for DNA. DNA is a polyelectrolyte, and the electrostatic interactions between charged walls, the solvent and the DNA molecule can dramatically alter the behavior of a confined chain, especially in narrow channels whose size is less than 100 nm. Proper treatment of electrostatics would require solution of the full non-linear Poisson-Boltzmann equation or consideration of explicit ions in the simulation. These methods are extremely expensive

and would not allow us to reach molecular weights of genomic DNA in simulations. While we have gained significant insights with a neutral polymer model, substantial challenges remain in modeling and simulation of polyelectrolytes in similar environments.

Bibliography

Alizadehheidari, M., E. Werner, C. Noble, M. Reiter-Schad, L. K. Nyberg, J. Fritzsche, B. Mehlig, J. O. Tegenfeldt, T. Ambjörnsson, F. Persson, and F. Westerlund, "Nanoconfined Circular and Linear DNA: Equilibrium Conformations and Unfolding Kinetics", Macromolecules **48**, 871 (2015).

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

Allison, S., S. S. Sorlie, and R. Pecora, "Brownian Dynamics Simulations of Wormlike Chains: Dynamic Light Scattering from a 2311 Base Pair DNA Fragment", Macromolecules **23**, 1110 (1990).

Arnold, A., B. Bozorgui, D. Frenkel, B.-Y. Ha, and S. Jun, "Unexpected relaxation dynamics of a self-avoiding polymer in cylindrical confinement", J. Chem. Phys. **127**, 164903 (2007).

Bakajin, O. B., T. A. J. Duke, C. F. Chou, S. S. Chan, R. H. Austin, and E. C. Cox, "Electrohydrodynamic Stretching of DNA in Confined Environments", Phys. Rev. Lett. **80**, 2737 (1998).

Balducci, A., C.-C. Hsieh, and P. S. Doyle, "Relaxation of Stretched DNA in Slitlike Confinement", Phys. Rev. Lett. **99**, 238102 (2007).

Balducci, A., P. Mao, J. Han, and P. S. Doyle, "Double-Stranded DNA Diffusion in Slitlike Nanochannels", Macromolecules **39**, 6273 (2006).

Bamshad, M. J., S. B. Ng, A. W. Bigham, H. K. Tabor, M. J. Emond, D. A. Nickerson, and J. Shendure, "Exome sequencing as a tool for Mendelian disease gene discovery", Nat. Rev. Genet. **12**, 745 (2011).

Baumann, C. G., S. B. Smith, V. A. Bloomfield, and C. Bustamante, "Ionic effects on the elasticity of single DNA molecules", Proc. Natl. Acad. Sci. U.S.A. **94**, 6185 (1997).

Benková, Z. and P. Cifra, "Simulation of Semiflexible Cyclic and Linear Chains Moderately and Strongly Confined in Nanochannels", Macromolecules **45**, 2597 (2012).

Bennink, M. L., O. D. Schärer, R. Kanaar, K. Sakata-Sogawa, J. M. Schins, J. S. Kanger, B. G. de Grooth, and J. Greve, "Single-molecule manipulation of double-stranded DNA using optical tweezers: Interaction studies of DNA with RecA and YOYO-1", Cytometry **36**, 200 (1999).

Bensimon, A, A Simon, A Chiffaudel, V Croquette, F Heslot, and D Bensimon, "Alignment and sensitive detection of DNA by a moving interface", Science **265**, 2096 (1994).

Binder, K., *Introduction: Theory and "technical" aspects of Monte Carlo simulations* (Springer).

Binder, K., ed. (1995), *Monte Carlo and Molecular Dynamics Simulations in Polymer Science* (Oxford University Press).

Brochard-Wyart, F., T. Tanaka, N. Borghi, and P.-G. de Gennes, "Semiflexible Polymers Confined in Soft Tubes", Langmuir **21**, 4144 (2005).

Brochard, F. and P. G. de Gennes, "Dynamics of confined polymer chains", J. Chem. Phys. **67**, 52 (1977).

Burkhardt, T., "Free energy of a semiflexible polymer in a tube and statistics of a randomly-accelerated particle", J. Phys. A: Math. Nucl. Gen. **30**, 167 (1997).

Burkhardt, T. W., Y. Yang, and G. Gompper, "Fluctuations of a long, semiflexible polymer in a narrow channel", Phys. Rev. E **82**, 041801 (2010).

Bustamante, C, J. Marko, E. Siggia, and S Smith, "Entropic elasticity of lambda-phage DNA", Science **265**, 1599 (1994).

Cao, H., A. R. Hastie, D. Cao, E. T. Lam, Y. Sun, H. Huang, X. Liu, L. Lin, W. Andrews, S. Chan, et al., "Rapid detection of structural variation in a human genome using nanochannel-based genome mapping technology", GigaScience **3**, 34 (2014).

Cazalet, C., C. Rusniok, H. Bruggemann, N. Zidane, A. Magnier, L. Ma, M. Tichit, S. Jarraud, C. Bouchier, F. Vandenesch, F. Kunst, J. Etienne, P. Glaser, and C. Buchrieser, "Evidence in the Legionella pneumophila genome for exploitation of host cell functions and high genome plasticity", Nat. Genet. **36**, 1165 (2004).

Chan, E. Y., N. M. Goncalves, R. A. Haeusler, A. J. Hatch, J. W. Larson, A. M. Maletta, G. R. Yantz, E. D. Carstea, M. Fuchs, G. G. Wong, et al., "DNA mapping using microfluidic stretching and single-molecule detection of fluorescent site-specific tags", Genome research **14**, 1137 (2004).

Chen, J. Z. Y., "Free Energy and Extension of a Wormlike Chain in Tube Confinement", Macromolecules **46**, 9837 (2013).

Chen, J. Z. Y. and D. E. Sullivan, "Free Energy of a Wormlike Polymer Chain Confined in a Slit: Crossover between Two Scaling Regimes", Macromolecules **39**, 7769 (2006).

Chen, Y.-L., H. Ma, M. D. Graham, and J. J. de Pablo, "Modeling DNA in Confinement: A Comparison between the Brownian Dynamics and Lattice Boltzmann Method", Macromolecules **40**, 5978 (2007).

Chen, Y.-L., "Electro-entropic excluded volume effects on DNA looping and relaxation in nanochannels", Biomicrofluidics **7**, 054119, 054119 (2013).

Chen, Y.-L., Y.-H. Lin, J.-F. Chang, and P.-k. Lin, "Dynamics and Conformation of Semiflexible Polymers in Strong Quasi-1D and -2D Confinement", Macromolecules **47**, 1199 (2014).

Cifra, P., "Channel confinement of flexible and semiflexible macromolecules", J. Chem. Phys. **131**, 224903 (2009).

Cifra, P., "Weak-to-strong confinement transition of semi-flexible macromolecules in slit and in channel", J. Chem. Phys. **136**, 024902 (2012).

Cifra, P. and T. Bleha, "Detection of chain backfolding in simulation of DNA in nanofluidic channels", Soft Matter **8**, 9022 (2012).

Cifra, P., Z. Benková, and T. Bleha, "Persistence Lengths and Structure Factors of Wormlike Polymers under Confinement", J. Phys. Chem. B **112**, 1367 (2008).

Cifra, P., Z. Benková, and T. Bleha, "Chain Extension of DNA Confined in Channels", J. Phys. Chem. B **113**, 1843 (2009).

Cifra, P., Z. Benková, and T. Bleha, "Persistence length of DNA molecules confined in nanochannels", Phys. Chem. Chem. Phys. **12**, 8934 (2010).

Clisby, N., "Accurate Estimate of the Critical Exponent ν for Self-Avoiding Walks via a Fast Implementation of the Pivot Algorithm", Phys. Rev. Lett. **104**, 055702 (2010).

Crick, F., "Central dogma of molecular biology", Nature **227**, 561 (1970).

Cui, T., J. Ding, and J. Z. Y. Chen, "Dynamics of a self-avoiding polymer chain in slit, tube, and cube confinements", Phys. Rev. E **78**, 061802 (2008).

Dai, L. and P. S. Doyle, "Comparisons of a Polymer in Confinement versus Applied Force", Macromolecules **46**, 6336 (2013).

Dai, L., S. Y. Ng, P. S. Doyle, and J. R. C. van der Maarel, "Conformation Model of Back-Folding and Looping of a Single DNA Molecule Confined Inside a Nanochannel", ACS Macro Lett. **1**, 1046 (2012).

Dai, L., D. R. Tree, J. R. C. van der Maarel, K. D. Dorfman, and P. S. Doyle, "Revisiting Blob Theory for DNA Diffusivity in Slitlike Confinement", Phys. Rev. Lett. **110**, 168105 (2013).

Dai, L., J. van der Maarel, and P. S. Doyle, "Extended de Gennes Regime of DNA Confined in a Nanochannel", Macromolecules **47**, 2445 (2014).

Daniels, H. E. and F. Smithies, "The probability distribution of the extent of a random chain", Math. Proc. Cambridge Philos. Soc. **37**, 244 (1941).

Daoud, M. and P. G. de Gennes, "Statistics of macromolecular solutions trapped in small pores", J. Phys.-Paris **38**, 85 (1977).

Das, S. K., M. D. Austin, M. C. Akana, P. Deshpande, H. Cao, and M. Xiao, "Single molecule linear analysis of DNA in nano-channel labeled with sequence specific fluorescent probes", Nucleic Acids Res. **38**, e177 (2010).

de Gennes, P.-G., *Scaling Concepts in Polymer Physics* (Cornell University Press).

de Gennes, P.-G., in *Polymer liquid crystals*, edited by R. B. M. A. Ciferri, W. R. Krigbaum (Academic Press), pp. 115–131.

de Gennes, P.-G., in *Polymers in confined environments*, Vol. 138, edited by S. Granick, K. Binder, P.-G. Gennes, E. Giannelis, G. Grest, H. Hervet, R. Krishnamoorti, L. Léger, E. Manias, E. Raphaël, and S.-Q. Wang, Advances in Polymer Science (Springer Berlin Heidelberg), pp. 91–105.

des Cloizeaux, J, "On the absence of Flory terms in the energy and in the entropy of a polymer chain", J. Phys.-Paris **37**, 431 (1976).

des Cloizeaux, J, "Short range correlation between elements of a long polymer in a good solvent", J. Phys.-Paris **41**, 223 (1980).

Diamant, H., "Hydrodynamic Interaction in Confined Geometries", J. Phys. Soc. Jpn. **78**, 041002 (2009).

Dierssen, M., Y. Herault, and X. Estivill, "Aneuploidy: From a Physiological Mechanism of Variance to Down Syndrome", Physiol. Rev. **89**, 887 (2009).

Dobrynin, A. V., "Effect of Counterion Condensation on Rigidity of Semiflexible Polyelectrolytes", Macromolecules **39**, 9519 (2006).

Doi, M and S. Edwards, *The Theory of Polymer Dynamics* (Clarendon Press Oxford).

Dorfman, K. D., D Gupta, A Jain, A Muralidhar, and D. R. Tree, "Hydrodynamics of DNA confined in nanoslits and nanochannels", Eur. Phys. J. Spec. Top. **223**, 3179 (2014).

Dorfman, K. D., "DNA electrophoresis in microfabricated devices", Rev. Mod. Phys. **82**, 2903 (2010).

Dorfman, K. D., S. B. King, D. W. Olson, J. D. P. Thomas, and D. R. Tree, "Beyond Gel Electrophoresis: Microfluidic Separations, Fluorescence Burst Analysis, and DNA Stretching", Chem. Rev. **113**, 2584 (2013).

Edmund A. DiMarzio, "Proper Accounting of Conformations of a Polymer Near a Surface", J. Chem. Phys. **42**, 2101 (1965).

Eid, J. et al., "Real-Time DNA Sequencing from Single Polymerase Molecules", Science **323**, 133 (2009).

Ermak, D. L. and J. A. McCammon, "Brownian dynamics with hydrodynamic interactions", J. Chem. Phys. **69**, 1352 (1978).

Fakhri, N., F. C. MacKintosh, B. Lounis, L. Cognet, and M. Pasquali, "Brownian Motion of Stiff Filaments in a Crowded Environment", Science **330**, 1804 (2010).

Feuk, L., A. R. Carson, and S. W. Scherer, "Structural variation in the human genome", Nat. Rev. Genet. **7**, 85 (2006).

França, L. T. C., E. Carrilho, and T. B. L. Kist, "A review of DNA sequencing techniques", Q. Rev. Biophys. **35**, 169 (2002).

Fredrickson, G. H., *The Equilibrium Theory of Inhomogeneous Polymers* (Clarendon Press Oxford).

Freed, K. F., *Renormalization Group Theory of Macromolecules* (Wiley New York et al.).

Frenkel, D. and B. Smit, *Understanding Molecular Simulation, Second Edition: From Algorithms to Applications (Computational Science)*, 2nd ed. (Academic Press).

Frykholm, K., M. Alizadehheidari, J. Fritzsche, J. Wigenius, M. Modesti, F. Persson, and F. Westerlund, "Probing Physical Properties of a DNA-Protein Complex Using Nanofluidic Channels", Small **10**, 884 (2014).

Gao, J., P. Tang, Y. Yang, and J. Z. Y. Chen, "Free energy of a long semiflexible polymer confined in a spherical cavity", Soft Matter **10**, 4674 (2014).

Godfrey, J. E., "The flexibility of low molecular weight double-stranded DNA as a function of length: I. Isolation and physical characterization of seven fractions", Biophys. Chem. **5**, 285 (1976).

Godfrey, J. E. and H. Eisenberg, "The flexibility of low molecular weight double-stranded DNA as a function of length: II. Light scattering measurements and the estimation

of persistence lengths from light scattering, sedimentation and viscosity", Biophys. Chem. **5**, 301 (1976).

Gorbunov, A. A. and A. M. Skvortsov, "Statistical properties of confined macromolecules", Adv. Colloid Interface Sci. **62**, 31 (1995).

Graham, M. D., "Fluid Dynamics of Dissolved Polymer Molecules in Confined Geometries", Annu. Rev. Fluid. Mech. **43**, 273 (2011).

Grassberger, P., "Pruned-enriched Rosenbluth method: Simulations of θ polymers of chain length up to 1 000 000", Phys. Rev. E **56**, 3682 (1997).

Grassberger, P. and R. Hegger, "Simulations of three-dimensional θ polymers", J. Chem. Phys. **102**, 6881 (1995).

Gray Jr., H. B. and J. E. Hearst, "Flexibility of native DNA from the sedimentation behavior as a fun ction of molecular weight and temperature", J. Mol. Biol. **35**, 111 (1968).

Grosberg, A. and A. Khokhlov, *Statistical Physics of Macromolecules* (AIP, New York).

Grunwald, A., M. Dahan, A. Giesbertz, A. Nilsson, L. K. Nyberg, E. Weinhold, T. Ambjörnsson, F. Westerlund, and Y. Ebenstein, "Bacteriophage strain typing by rapid single molecule analysis", Nucleic Acids Res., gkv563 (2015).

Günther, K., M. Mertig, and R. Seidel, "Mechanical and structural properties of YOYO-1 complexed DNA", Nucleic Acids Res. **38**, 6526 (2010).

Gupta, D., J. Sheats, A. Muralidhar, J. J. Miller, D. E. Huang, S. Mahshid, K. D. Dorfman, and W. Reisner, "Mixed confinement regimes during equilibrium confinement spectroscopy of DNA", J. Chem. Phys. **140**, 214901 (2014).

Gupta, D., J. J. Miller, A. Muralidhar, S. Mahshid, W. Reisner, and K. D. Dorfman, "Experimental Evidence of Weak Excluded Volume Effects for Nanochannel Confined DNA", ACS Macro Lett. **4**, 759 (2015).

Harden, J. L. and M. Doi, "Diffusion of macromolecules in narrow capillaries", J. Phys. Chem. **96**, 4046 (1992).

Harpst, J., "Analysis of low angle light scattering results from T7 DNA", Biophys. Chem. **11**, 295 (1980).

Hastie, A. R., L. Dong, A. Smith, J. Finklestein, E. T. Lam, N. Huo, H. Cao, P.-Y. Kwok, K. R. Deal, J. Dvorak, M.-C. Luo, Y. Gu, and M. Xiao, "Rapid Genome Mapping in Nanochannel Arrays for Highly Complete and Accurate *De Novo* Sequence Assembly of the Complex *Aegilops tauschii* Genome", PLoS ONE **8**, e55864 (2013).

Hiemenz, P. C. and T. Lodge, *Polymer Chemistry* (CRC Press, Boca Raton).

Hirst, M. and M. A. Marra, "Next generation sequencing based approaches to epigenomics", Brief. Funct. Genomics **9**, 455 (2010).

Hsieh, C.-C. and P. S. Doyle, "Studying confined polymers using single-molecule DNA experiments", Korea-Australia Rheology Journal **20**, 127 (2008).

Hsieh, C.-C., A. Balducci, and P. S. Doyle, "Ionic Effects on the Equilibrium Dynamics of DNA Confined in Nanoslits", Nano Lett. **8**, 1683 (2008).

Hsu, H.-P. and P. Grassberger, "2-Dimensional polymers confined in a strip", Eur. Phys. J. B **36**, 209 (2003).

Hsu, H., W. Paul, and K. Binder, "Standard definitions of persistence length do not describe the local âăijintrinsicâăi stiffness of real polymer chains", Macromolecules **43**, 3094 (2010).

Hsu, H.-P. and K. Binder, "Semi-flexible polymer chains in quasi-one-dimensional confinement: a Monte Carlo study on the square lattice", Soft Matter **9**, 10512 (2013).

Hsu, H.-P. and P. Grassberger, "Polymers confined between two parallel plane walls", J. Chem. Phys. **120**, 2034 (2004).

Hsu, H.-P. and P. Grassberger, "A Review of Monte Carlo Simulations of Polymers with PERM", J. Stat. Phys. **144**, 597 (2011).

Hsu, H.-P., W. Paul, and K. Binder, "Polymer chain stiffness vs. excluded volume: A Monte Carlo study of the crossover towards the worm-like chain model", Europhys. Lett. **92**, 28003 (2010).

Iafrate, A. J., L. Feuk, M. N. Rivera, M. L. Listewnik, P. K. Donahoe, Y. Qi, S. W. Scherer, and C. Lee, "Detection of large-scale variation in the human genome", Nat. Genet. **36**, 949 (2004).

Iarko, V., E. Werner, L. K. Nyberg, V. Müller, J. Fritzsche, T. Ambjörnsson, J. P. Beech, J. O. Tegenfeldt, K. Mehlig, F. Westerlund, and B. Mehlig, "Extension of nanoconfined DNA: Quantitative comparison between experiment and theory", Phys. Rev. E **92**, 062701 (2015).

Illumina, *An Introduction to Next-Generation Sequencing Technology*, Illumina, Inc. (San Diego, CA).

Jain, A. and K. D. Dorfman, "Evaluation of the Kirkwood approximation for the diffusivity of channel-confined DNA chains in the de Gennes regime", Biomicrofluidics **9**, 024112 (2015).

Jendrejack, R. M., D. C. Schwartz, M. D. Graham, and J. J. de Pablo, "Effect of confinement on DNA dynamics in microfluidic devices", J. Chem. Phys. **119**, 1165 (2003).

Jing, J., J. Reed, J. Huang, X. Hu, V. Clarke, J. Edington, D. Housman, T. S. Anantharaman, E. J. Huff, B. Mishra, et al., "Automated high resolution optical mapping using arrayed, fluid-fixed DNA molecules", Proc. Natl. Acad. Sci. U.S.A. **95**, 8046 (1998).

Jo, K., D. Dhingra, T. Odijk, J. De Pablo, M. Graham, R. Runnheim, D. Forrest, and D. Schwartz, "A single-molecule barcoding system using nanoslits for DNA analysis", Proc. Natl. Acad. Sci. U.S.A. **104**, 2673 (2007).

Jolly, D. and H. Eisenberg, "Photon correlation spectroscopy, total intensity light scattering with laser radiation, and hydrodynamic studies of a well fractionated DNA sample", Biopolymers **15**, 61 (1976).

Kam, Z., N. Borochov, and H. Eisenberg, "Dependence of laser light scattering of DNA on NaCl concentration", Biopolymers **20**, 2671 (1981).

Karger, B. L. and A. Guttman, "DNA sequencing by CE", Electrophoresis **30**, S196 (2009).

Khorshid, A., P. Zimny, D. Tétreault-La Roche, G. Massarelli, T. Sakaue, and W. Reisner, "Dynamic Compression of Single Nanochannel Confined DNA via a Nanodozer Assay", Phys. Rev. Lett. **113**, 268104 (2014).

Kim, Y., K. S. Kim, K. L. Kounovsky, R. Chang, G. Y. Jung, J. J. dePablo, K. Jo, and D. C. Schwartz, "Nanochannel confinement: DNA stretch approaching full contour length", Lab. Chip **11**, 1721 (2011).

Kirkwood, J. G. and J. Riseman, "The Intrinsic Viscosities and Diffusion Constants of Flexible Macromolecules in Solution", J. Chem. Phys. **16**, 565 (1948).

Klotz, A. R., "DNA Polymer Physics in Complex Nanofluidic Environments", PhD thesis (McGill University, Montreal, Quebec).

Klotz, A. R., L. Duong, M. Mamaev, H. W. de Haan, J. Z. Y. Chen, and W. W. Reisner, "Measuring the Confinement Free Energy and Effective Width of Single Polymer Chains via Single-Molecule Tetris", Macromolecules **48**, 5028 (2015).

Koren, S., G. P. Harhay, T. P. Smith, J. L. Bono, D. M. Harhay, S. D. Mcvey, D. Radune, N. H. Bergman, and A. M. Phillippy, "Reducing assembly complexity of microbial genomes with single-molecule sequencing", Genome Biology **14**, R101 (2013).

Köster, S., D. Steinhauser, and T. Pfohl, "Brownian motion of actin filaments in confining microchannels", J. Phys.: Condens. Matter **17**, S4091 (2005).

Köster, S., H. Stark, T. Pfohl, and J. Kierfeld, "Fluctuations of single confined actin filaments", Biophys. Rev. Lett. **2**, 155 (2007).

Kundukad, B., J. Yan, and P. S. Doyle, "Effect of YOYO-1 on the mechanical properties of DNA", Soft matter **10**, 9721 (2014).

Lam, E. T., A. Hastie, C. Lin, D. Ehrlich, S. K. Das, M. D. Austin, P. Deshpande, H. Cao, N. Nagarajan, M. Xiao, et al., "Genome mapping on nanochannel arrays for structural variation analysis and sequence assembly", Nat. Biotechnol. **30**, 771 (2012).

Larson, R. G., "The rheology of dilute solutions of flexible polymers: Progress and problems", J. Rheol. **49**, 1 (2005).

Latinwo, F. and C. M. Schroeder, "Model systems for single molecule polymer dynamics", Soft Matter **7**, 7907 (2011).

Latreille, P., S. Norton, B. S. Goldman, J. Henkhaus, N. Miller, B. Barbazuk, H. B. Bode, C. Darby, Z. Du, S. Forst, et al., "Optical mapping as a routine tool for bacterial genome sequence finishing", BMC Genomics **8**, 321 (2007).

Lederer, H., R. P. May, J. K. Kjems, G. Baeῖ, and H. Heumann, "Solution structure of a short DNA fragment studied by neutron scattering", Eur. J. Biochem. **161**, 191 (1986).

Levin, J. Z., M. F. Berger, X. Adiconis, P. Rogov, A. Melnikov, T. Fennell, C. Nusbaum, L. A. Garraway, A. Gnirke, et al., "Targeted next-generation sequencing of a cancer transcriptome enhances detection of sequence variants and novel fusion transcripts", Genome Biol. **10**, R115 (2009).

Levy, S. L. and H. G. Craighead, "DNA manipulation, sorting, and mapping in nanofluidic systems", Chem. Soc. Rev. **39**, 1133 (2010).

Levy, S. L., J. T. Mannion, J. Cheng, C. H. Reccius, and H. G. Craighead, "Entropic Unfolding of DNA Molecules in Nanofluidic Channels", Nano Lett. **8**, 3839 (2008).

Li, B., N. Madras, and A. Sokal, "Critical exponents, hyperscaling, and universal amplitude ratios for two- and three-dimensional self-avoiding walks", English, J. Stat. Phys. **80**, 661 (1995).

Liu, B. and B. Dünweg, "Translational diffusion of polymer chains with excluded volume and hydrodynamic interactions by Brownian dynamics simulation", J. Chem. Phys. **118**, 8061 (2003).

Liu, L., Y. Li, S. Li, N. Hu, Y. He, R. Pong, D. Lin, L. Lu, and M. Law, "Comparison of next-generation sequencing systems", J. BioMed. Biotechnol. **2012**, 251364 (2012).

Madras, N. N. and G. Slade, *The self-avoiding walk* (Birkhäuser).

Madras, N. and A. D. Sokal, "The pivot algorithm: A highly efficient Monte Carlo method for the self-avoiding walk", J. Stat. Phys. **50**, 109 (1988).

Mak, A. C., Y. Y. Lai, E. T. Lam, T.-P. Kwok, A. K. Leung, A. Poon, Y. Mostovoy, A. R. Hastie, W. Stedman, T. Anantharaman, et al., "Genome-wide structural variation detection by genome mapping on nanochannel arrays", Genetics **202**, 351 (2016).

Mannion, J., C. Reccius, J. Cross, and H. Craighead, "Conformational analysis of single DNA molecules undergoing entropically induced motion in nanochannels", Biophys. J. **90**, 4538 (2006).

Mansfield, M. L. and J. F. Douglas, "Is duplex DNA a swollen random coil?", Soft Matter **9**, 8914 (2013).

Mansfield, M. L., J. F. Douglas, S. Irfan, and E.-H. Kang, "Comparison of Approximate Methods for Calculating the Friction Coefficient and Intrinsic Viscosity of Nanoparticles and Macromolecules", Macromolecules **40**, 2575 (2007).

Marenduzzo, D, C Micheletti, and E Orlandini, "Biopolymer organization upon confinement", J. Phys.: Condens. Matter **22**, 283102 (2010).

Metzker, M. L., "Sequencing technologies — the next generation", Nat. Rev. Genet. **11**, 31 (2010).

Morse, D. C., "Viscoelasticity of Concentrated Isotropic Solutions of Semiflexible Polymers. 2. Linear Response", Macromolecules **31**, 7044 (1998).

Muralidhar, A. and K. D. Dorfman, "Kirkwood Diffusivity of Long Semiflexible Chains in Nanochannel Confinement", Macromolecules **48**, 2829 (2015).

Muralidhar, A. and K. D. Dorfman, "Backfolding of DNA Confined in Nanotubes: Flory Theory versus the Two-State Cooperativity Model", Macromolecules (in press) (2016) `10.1021/acs.macromol.5b02556)`.

Muralidhar, A., D. R. Tree, and K. D. Dorfman, "Backfolding of Wormlike Chains Confined in Nanochannels", Macromolecules **47**, 8446 (2014).

Muralidhar, A., D. R. Tree, Y. Wang, and K. D. Dorfman, "Interplay between chain stiffness and excluded volume of semiflexible polymers confined in nanochannels", J. Chem. Phys. **140**, 084905 (2014).

Muralidhar, A., M. J. Quevillon, and K. D. Dorfman, "The backfolded Odijk regime for wormlike chains confined in rectangular nanochannels", Polymers (In review) (2016).

Ng, P. C. and E. F. Kirkness, in *Genetic variation*, Vol. 628, edited by M. R. Barnes and G. Breen, Methods in Molecular Biology (Humana Press), pp. 215–226.

Niedringhaus, T. P., D. Milanova, M. B. Kerby, M. P. Snyder, and A. E. Barron, "Landscape of Next-Generation Sequencing Technologies", Anal. Chem. **83**, 4327 (2011).

Nöding, B. and S. Köster, "Intermediate Filaments in Small Configuration Spaces", Phys. Rev. Lett. **108**, 088101 (2012).

Nyberg, L., F. Persson, B. Åkerman, and F. Westerlund, "Heterogeneous staining: a tool for studies of how fluorescent dyes affect the physical properties of DNA", Nucleic Acids Res. **41**, e184 (2013).

Odijk, T., "The statistics and dynamics of confined or entangled stiff polymers", Macromolecules **16**, 1340 (1983).

Odijk, T., "DNA in a liquid-crystalline environment: Tight bends, rings, supercoils", J. Chem. Phys. **105**, 1270 (1996).

Odijk, T., "DNA confined in nanochannels: Hairpin tightening by entropic depletion", J. Chem. Phys. **125**, 204904 (2006).

Odijk, T., "Scaling theory of DNA confined in nanochannels and nanoslits", Phys. Rev. E **77**, 060901 (2008).

Oono, Y. and M. Kohmoto, "Renormalization group theory of transport properties of polymer solutions. I. Dilute solutions", J. Chem. Phys. **78**, 520 (1983).

Pendleton, M. et al., "Assembly and diploid architecture of an individual human genome via single-molecule technologies", Nat. Meth. **12**, 780 (2015).

Perkins, T., Quake, D. Smith, and S Chu, "Relaxation of a single DNA molecule observed by optical microscopy", Science **264**, 822 (1994).

Perkins, T., D. Smith, R. Larson, and S Chu, "Stretching of a single tethered polymer in a uniform flow", Science **268**, 83 (1995).

Perkins, T. T., D. E. Smith, and S. Chu, "Single Polymer Dynamics in an Elongational Flow", Science **276**, 2016 (1997).

Persson, F. and J. O. Tegenfeldt, "DNA in nanochannels-directly visualizing genomic information", Chem. Soc. Rev. **39**, 985 (2010).

Persson, F., P. Utko, W. Reisner, N. B. Larsen, and A. Kristensen, "Confinement Spectroscopy: Probing Single DNA Molecules with Tapered Nanochannels", Nano Lett. **9**, 1382 (2009).

Pozrikidis, C., *Fluid dynamics: theory, computation, and numerical simulation* (Springer).

Prellberg, T., "Rare event sampling with stochastic growth algorithms", EPJ Web of Conferences **44**, 01001 (2013).

Prellberg, T. and J. Krawczyk, "Flat Histogram Version of the Pruned and Enriched Rosenbluth Method", Phys. Rev. Lett. **92**, 120602 (2004).

Quail, M. A., M. Smith, P. Coupland, T. D. Otto, S. R. Harris, T. R. Connor, A. Bertoni, H. P. Swerdlow, and Y. Gu, "A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers", BMC genomics **13**, 341 (2012).

Raphael, E. and P. Pincus, "Scaling theory of polymer solutions trapped in small pores : the θ-solvent case", J. Phys. II **2**, 1341 (1992).

Reinhart, W. F., D. R. Tree, and K. D. Dorfman, "Entropic depletion of DNA in triangular nanochannels", Biomicrofluidics **7**, 024102 (2013).

Reisner, W., K. J. Morton, R. Riehn, Y. M. Wang, Z. Yu, M. Rosen, J. C. Sturm, S. Y. Chou, E. Frey, and R. H. Austin, "Statics and Dynamics of Single DNA Molecules Confined in Nanochannels", Phys. Rev. Lett. **94**, 196101 (2005).

Reisner, W., J. P. Beech, N. B. Larsen, H. Flyvbjerg, A. Kristensen, and J. O. Tegenfeldt, "Nanoconfinement-Enhanced Conformational Response of Single DNA Molecules to Changes in Ionic Environment", Phys. Rev. Lett. **99**, 058302 (2007).

Reisner, W., J. N. Pedersen, and R. H. Austin, "DNA confinement in nanochannels: physics and biological applications", Rep. Prog. Phys. **75**, 106601 (2012).

Riehn, R., W. Reisner, J. O. Tegenfeldt, Y. M. Wang, C.-K. Tung, S.-F. Lim, E. Cox, J. C. Sturm, K. Morton, S. Y. Chou, and R. H. Austin, in *Integrated Biochips for DNA Analysis*, edited by R. Liu and A. Lee, Biotechnology Intelligence Unit (Springer New York), pp. 151–186.

Robertson, R. M. and D. E. Smith, "Direct Measurement of the Intermolecular Forces Confining a Single Molecule in an Entangled Polymer Solution", Phys. Rev. Lett. **99**, 126001 (2007).

Rodríguez Schmidt, R., J. Hernández Cifre, and J. García de la Torre, "Translational diffusion coefficients of macromolecules", Eur. Phys. J. E **35**, 1 (2012).

Rosenbluth, M. N. and A. W. Rosenbluth, "Monte Carlo Calculation of the Average Extension of Molecular Chains", J. Chem. Phys. **23**, 356 (1955).

Rouse, P. E., "A theory of the linear viscoelastic properties of dilute solutions of coiling polymers", J. Chem. Phys. **21**, 1272 (1953).

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

Schaefer, D. W., J. F. Joanny, and P. Pincus, "Dynamics of Semiflexible Polymers in Solution", Macromolecules **13**, 1280 (1980).

Schäfer, L., *Excluded Volume Effects in Polymer Solutions: as Explained by the Renormalization Group* (Springer Science & Business Media).

Schellman, J., "Flexibility of DNA", Biopolymers **13**, 217 (1974).

Schellman, J. A. and D. Stigter, "Electrical double layer, zeta potential, and electrophoretic charge of double-stranded DNA", Biopolymers **16**, 1415 (1977).

Schmid, C. W., F. P. Rinehart, and J. E. Hearst, "Statistical length of DNA from light scattering", Biopolymers **10**, 883 (1971).

Schroeder, C. M., H. P. Babcock, E. S. G. Shaqfeh, and S. Chu, "Observation of Polymer Conformation Hysteresis in Extensional Flow", Science **301**, 1515 (2003).

Seils, J. and T. Dorfmüller, "Internal dynamics of linear and superhelical DNA as studied by photon correlation spectroscopy", Biopolymers **31**, 813 (1991).

Shaqfeh, E. S., "The dynamics of single-molecule DNA in flow", J. Non-Newtonian Fluid Mech. **130**, 1 (2005).

Sheats, J., J. G. Reifenberger, H. Cao, and K. D. Dorfman, "Measurements of DNA barcode label separations in nanochannels from time-series data", Biomicrofluidics **9**, 064119 (2015).

Skolnick, J. and M. Fixman, "Electrostatic Persistence Length of a Wormlike Polyelectrolyte", Macromolecules **10**, 944 (1977).

Smith, D. E., T. T. Perkins, and S. Chu, "Dynamical Scaling of DNA Diffusion Coefficients", Macromolecules **29**, 1372 (1996).

Smithe, T. S. C., V. Iarko, A. Muralidhar, E. Werner, K. D. Dorfman, and B. Mehlig, "Finite-size corrections for confined polymers in the extended de Gennes regime", Phys. Rev. E **92**, 062601 (2015).

Sorlie, S. S. and R. Pecora, "A dynamic light scattering study of four DNA restriction fragments", Macromolecules **23**, 487 (1990).

Spakowitz, A. J. and Z.-G. Wang, "Semiflexible polymer solutions. I. Phase behavior and single-chain statistics", J. Chem. Phys. **119**, 13113 (2003).

Steinberg, K. M., V. A. Schneider, T. A. Graves-Lindsay, R. S. Fulton, R. Agarwala, J. Huddleston, S. A. Shiryev, A. Morgulis, U. Surti, W. C. Warren, et al., "Single haplotype assembly of the human genome from a hydatidiform mole", Genome Res. **24**, 2066 (2014).

Stigter, D., "The Charged Colloidal Cylinder with a Gouy Double Layer", J. Colloid Interface Sci. **53**, 296 (1975).

Stigter, D., "Interactions of highly charged colloidal cylinders with applications to double-stranded DNA", Biopolymers **16**, 1435 (1977).

Su, T., S. K. Das, M. Xiao, and P. K. Purohit, "Transition between two regimes describing internal fluctuation of dna in a nanochannel", PLoS One **6**, e16890 (2011).

Tang, J., S. L. Levy, D. W. Trahan, J. J. Jones, H. G. Craighead, and P. S. Doyle, "Revisiting the Conformation and Dynamics of DNA in Slitlike Confinement", Macromolecules **43**, 7368 (2010).

Tegenfeldt, J. O., C. Prinz, H. Cao, S. Chou, W. W. Reisner, R. Riehn, Y. M. Wang, E. C. Cox, J. C. Sturm, P. Silberzan, and R. H. Austin, "The dynamics of genomic-length DNA molecules in 100-nm channels", Proc. Natl. Acad. Sci. U.S.A. **101**, 10979 (2004).

Tkachenko, A. and Y. Rabin, "Coupling between Thermodynamics and Conformations in Wormlike Polymer Nematics", Macromolecules **28**, 8646 (1995).

Tree, D. R., "DNA Confined in Nanochannels and Nanoslits", PhD thesis (University of Minnesota, Minneapolis, Minnesota).

Tree, D. R., Y. Wang, and K. D. Dorfman, "Mobility of a Semiflexible Chain Confined in a Nanochannel", Phys. Rev. Lett. **108**, 228105 (2012).

Tree, D. R., Y. Wang, and K. D. Dorfman, "Extension of DNA in a Nanochannel as a Rod-to-Coil Transition", Phys. Rev. Lett. **110**, 208103 (2013).

Tree, D. R., A. Muralidhar, P. S. Doyle, and K. D. Dorfman, "Is DNA a Good Model Polymer?", Macromolecules **46**, 8369 (2013).

Tree, D. R., Y. Wang, and K. D. Dorfman, "Modeling the relaxation time of DNA confined in a nanochannel", Biomicrofluidics **7**, 054118 (2013).

Tree, D. R., W. F. Reinhart, and K. D. Dorfman, "The Odijk Regime in Slits", Macromolecules **47**, 3672 (2014).

Van den Heuvel, M. G. L., S. Bolhuis, and C. Dekker, "Persistence Length Measurements from Stochastic Single-Microtubule Trajectories", Nano Lett. **7**, 3138 (2007).

van Rensburg, E. J. J., "Monte Carlo methods for the self-avoiding walk", J. Phys. A: Math. Theor. **42**, 323001 (2009).

Voordouw, G., Z. Kam, N. Borochov, and H. Eisenberg, "Isolation and physical studies of the intact supercoiled: The open circular and the linear forms of CoIE1-plasmid DNA", Biophys. Chem. **8**, 171 (1978).

Vroege, G. J. and T. Odijk, "Induced chain rigidity, splay modulus and other properties of nematic polymer liquid crystals", Macromolecules **21**, 2848 (1988).

Wagner, F., G. Lattanzi, and E. Frey, "Conformations of confined biopolymers", Phys. Rev. E **75**, 050902 (2007).

Wang, J. and H. Gao, "A generalized bead-rod model for Brownian dynamics simulations of wormlike chains under strong confinement", J. Chem. Phys. **123**, 084906 (2005).

Wang, S.-Q., J. F. Douglas, and K. F. Freed, "Corrections to preaveraging approximation within the Kirkwood–Riseman model for flexible polymers: Calculations to second order in epsilon with both hydrodynamic and excluded volume interactions", J. Chem. Phys. **85**, 3674 (1986).

Wang, Y., G. H. Peters, F. Y. Hansen, and O. Hassager, "Equilibrium partitioning of macromolecules in confining geometries: Improved universality with a new molecular size parameter", J. Chem. Phys. **128**, 124904 (2008).

Wang, Y., I. Teraoka, F. Y. Hansen, G. H. Peters, and O. Hassager, "A Theoretical Study of the Separation Principle in Size Exclusion Chromatography", Macromolecules **43**, 1651 (2010).

Wang, Y., D. R. Tree, and K. D. Dorfman, "Simulation of DNA Extension in Nanochannels", Macromolecules **44**, 6594 (2011).

Wang, Y., W. F. Reinhart, D. R. Tree, and K. D. Dorfman, "Resolution limit for DNA barcodes in the Odijk regime", Biomicrofluidics **6**, 014101 (2012).

Wang, Z., M. Gerstein, and M. Snyder, "RNA-Seq: a revolutionary tool for transcriptomics", Nat. Rev. Genet. **10**, 57 (2009).

Watson, J. D. and F. H. C. Crick, "Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid", Nature **171**, 737 (1953).

Werner, E. and B. Mehlig, "Confined polymers in the extended de Gennes regime", Phys. Rev. E **90**, 062602 (2014).

Werner, E. and B. Mehlig, "Scaling regimes of a semiflexible polymer in a rectangular channel", Phys. Rev. E **91**, 050601 (2015).

Werner, E., F. Persson, F. Westerlund, J. O. Tegenfeldt, and B. Mehlig, "Orientational correlations in confined DNA", Phys. Rev. E **86**, 041802 (2012).

Werner, E., F. Westerlund, J. O. Tegenfeldt, and B. Mehlig, "Monomer Distributions and Intrachain Collisions of a Polymer Confined to a Channel", Macromolecules **46**, 6644 (2013).

Werner, E., "Equilibrium statistics of channel-confined DNA", PhD thesis (University of Gothenburg, Gothenburg, Sweden).

Yamakawa, H., *Modern Theory of Polymer Solutions* (Interscience, New York).

Yamakawa, H. and M. Fujii, "Statistical mechanics of helical wormlike chains. I. Differential equations and moments", J. Chem. Phys. **64**, 5222 (1976).

Yang, Y., T. W. Burkhardt, and G. Gompper, "Free energy and extension of a semiflexible polymer in cylindrical confining geometries", Phys. Rev. E **76**, 011804 (2007).

Yoshikawa, K., M. Takahashi, V. V. Vasilevskaya, and A. R. Khokhlov, "Large Discrete Transition in a Single DNA Molecule Appears Continuous in the Ensemble", Phys. Rev. Lett. **76**, 3029 (1996).

Zhang, C., F. Zhang, J. A. van Kan, and J. R. C. van der Maarel, "Effects of electrostatic screening on the conformation of single DNA molecules confined in a nanochannel", J. Chem. Phys. **128**, 225109 (2008).

Zimm, B. H. and J. K. Bragg, "Theory of the Phase Transition between Helix and Random Coil in Polypeptide Chains", J. Chem. Phys. **31**, 526 (1959).

Zimm, B. H., "Dynamics of Polymer Molecules in Dilute Solution: Viscoelasticity, Flow Birefringence and Dielectric Loss", J. Chem. Phys. **24**, 269 (1956).

Zimm, B. H., "Chain Molecule Hydrodynamics by the Monte-Carlo Method and the Validity of the Kirkwood-Riseman Approximation", Macromolecules **13**, 592 (1980).

Appendix A

Note on the definition of persistence length and Kuhn length

There are several definitions of persistence length and Kuhn length in the polymer physics literature. Often, it can be confusing as to which is the right definition and under what conditions the definitions are valid. I'm going to use the most general definition that does not make any assumptions about the model used to describe the polymer. The persistence length, l_{p}, is defined as the sum of the average projection of the unit vector pointing in the direction of the first bond vector, $\vec{\ell}_0$, of a long *ideal* chain on the end-to-end vector of the chain, $\vec{R}$, i.e.

$$l_{\mathrm{p}} = \frac{1}{\ell}\left\langle \vec{\ell}_0 \cdot \vec{R} \right\rangle = \frac{1}{\ell}\left\langle \vec{\ell}_0 \cdot \sum_{i=1}^{\infty} \vec{\ell}_i \right\rangle, \tag{A.1}$$

where ℓ is the bond length in the chain (Hiemenz and Lodge, 2007). There are other definitions such as the one based on the decay of the correlation of bond vectors along the chain (Rubinstein and Colby, 2003). However, this defintion is consistent with Eq. A.1 only in the limit of a wormlike chain (stiff chain) and when the bond length goes to 0 ($\ell \to 0$ or continuum approximation).

Similarly the Kuhn length, b, is defined as the projection of a bond vector in the interior of the chain on the bond vectors on either side,

$$b = \frac{1}{\ell}\left\langle \vec{\ell}_0 \cdot \sum_{i=-\infty}^{\infty} \vec{\ell}_i \right\rangle. \tag{A.2}$$

Eq. A.1 can be used to show that

$$l_{\mathrm{p}} = (C_\infty + 1)\frac{\ell}{2}, \tag{A.3}$$

where $C_\infty = (1 + \cos\theta)/(1 - \cos\theta)$, a function of average of cosine of the angle between two bond vectors in the chain, θ (Hiemenz and Lodge, 2007). Along the same lines, the Kuhn length can be written as

$$b = C_\infty \ell \tag{A.4}$$

However, for most stiff polymers, $\cos\theta \approx 1$ and therefore $C_\infty \gg 1$, which results in an approximation to the expression for the persistence length given by

$$l_{\mathrm{p}} \approx C_\infty \frac{\ell}{2}. \tag{A.5}$$

It turns out that Eq. A.5 is widely used in literature primarily with respect to considerably stiff chains. Surely, when $C_\infty \gg 1$, Eq. A.3 and Eq. A.5 are equivalent. It is also clear that the traditional definition, $b = 2l_{\mathrm{p}}$, is valid only in the limit, $C_\infty \gg 1$. There is a problem, however, when $C_\infty \approx 1$, which is the case for relatively flexible chains. In fact, we have used Eq. A.5 instead of the more general Eq. A.3 as the definition of the persistence length in some of our earlier papers (Tree et al., 2013a; Tree et al., 2013b; Tree et al., 2014; Muralidhar et al., 2014b; Muralidhar et al., 2014a). However, for most of the latter publications, we dealt with relatively stiff chains and therefore the difference between the results of Eq. A.1 and Eq. A.5.

I will end with a note on our model (discrete wormlike chain model) and the mapping of the parameters used in the model to the persistence length of the molecule. The bending potential for the angle between two contiguous bonds used in our model is

$$\beta U_{\mathrm{bend}} = \kappa(1 - \cos\theta), \tag{A.6}$$

where κ is the bending constant and $\beta = 1/kT$. An equivalent potential that is often used in literature is

$$\beta U_{\mathrm{bend}} = \frac{\kappa}{2}(\hat{\ell}_i - \hat{\ell}_{i+1}) \cdot (\hat{\ell}_i - \hat{\ell}_{i+1}), \tag{A.7}$$

where the hat denotes unit bond vectors (Wang et al., 2011). In order to estimate how this bending potential translates to the value of the persistence length, we first need to

calculate the average of $\cos\theta$ as

$$\langle\cos\theta\rangle = \frac{\int_0^\pi \cos\theta \exp(-\kappa(1-\cos\theta))\sin\theta \ \mathrm{d}\theta}{\int_0^\pi \exp(-\kappa(1-\cos\theta))\sin\theta \ \mathrm{d}\theta}. \tag{A.8}$$

Note that the factor of $\sin\theta$ arises because we are integrating in a spherical coordinate system. The integration in Eq. A.8 can be performed analytically and plugged into the relation of $C_\infty = (1+\langle\cos\theta\rangle)/(1-\langle\cos\theta\rangle)$ and combined with Eq. A.3 to obtain

$$l_\mathrm{p} = \frac{\kappa}{\kappa - \kappa\coth(\kappa) + 1}\ell, \tag{A.9}$$

which is a general relation for l_p as a function of κ. A similar expression for b in terms of κ is

$$b = \left(\frac{\kappa - 1 + \kappa\coth\kappa}{\kappa + 1 - \kappa\coth\kappa}\right)\ell \tag{A.10}$$

We have used the approximation $l_\mathrm{p} \approx b/2$ in our older papers (Tree et al., 2013b; Tree et al., 2014; Muralidhar et al., 2014b; Muralidhar et al., 2014a)

$$l_\mathrm{p} \approx \left(\frac{\kappa - 1 + \kappa\coth\kappa}{\kappa + 1 - \kappa\coth\kappa}\right)\frac{\ell}{2}. \tag{A.11}$$

There is also another approximation, $l_\mathrm{p}/\ell = \kappa$, used widely in literature with respect to the model in Eqs. A.6 and A.7 including papers from our group (Wang et al., 2011). Again, this definition is valid for relatively stiff polymers ($\kappa > 10$), such as DNA. In fact, if we take the limit of $\kappa \to \infty$ Eq. A.9 yields $l_\mathrm{p}/\ell = \kappa$. However, in practice this is valid for $\kappa \gtrsim 10$ as $\coth(\kappa)$ quickly goes to 1.

In summary, one has to be careful in using the correct definition of persistence length that is valid both for stiff and flexible polymers. Additionally, one has to map the parameters of the model to obtain the correct value of the persistence length using such a defintion.

Appendix B

Supporting Information to Chapter 5

B.1 PERM simulations

We used the discrete wormlike chain model that we have used in our previous work to run all our PERM simulations (Tree et al., 2013a; Muralidhar et al., 2014b; Tree et al., 2014). Our PERM simulations were carried out in a DELL Linux cluster using a Fortran 95 code parallelized via Message Passing Interface (MPI). For different data sets correspoding to different values of (l_{p}, w, D) we used a minimum of 250,000 tours, while the maximum was 500,000 tours. The number of beads in our simulations spanned a range starting from 1601 to 50,001 ensuring that our chains reached the asymptotic limit of the property under consideration. To calculate the global persistence length, g for comparision with Odijk's mechanical theory, we turned off the excluded volume potential, thereby employing a special case of our model for ideal wormlike chains. However, the estimation of g for real chains was performed using ideal chains confined in channels with identical $D_{\mathrm{eff}}/l_{\mathrm{p}}$ ($D_{\mathrm{eff}} = D - w$) value as that of real chains.

B.2 Order parameter

Figure B.1 shows the behavior of the order parameter

$$m \equiv \langle P_2(\cos\theta)\rangle = \frac{1}{2}\left(3\left\langle\cos^2\theta\right\rangle - 1\right), \tag{B.1}$$

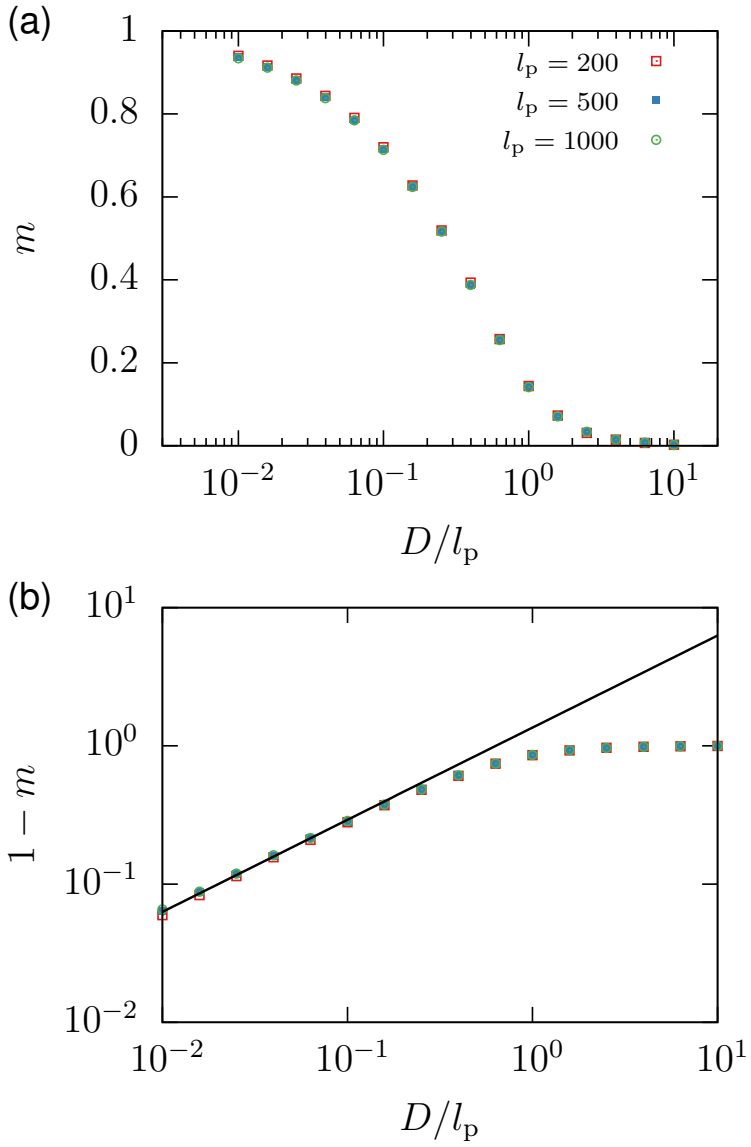


Figure B.1. (a) The orientational order parameter as a function of dimensionless channel size. (b) The scaling of $1-m$ as $D/l_\mathrm{p} \to 0$.

for ideal wormlike chains as a function of channel size. We notice in Figure B.1(a) that, irrespective of the value of l_p employed in the model, our data for m collapse when plotted against D/l_p, demonstrating as expected that m is a unique function of D/l_p. Furthermore, it can be clearly seen that $m \to 1$ (perfect alignment) as $D/l_p \to 0$ and $m \to 0$ (completely isotropic) as $D/l_p \to \infty$ for long chains. In the Odijk regime, as $D/l_p \to 0$, we find that $1 - m = 1.36(D/l_p)^{2/3}$ (Figure B.1(b)), in agreement with previous predictions (Jo et al., 2007; Odijk, 2008; Chen, 2013b). Note that a similar scaling of the order parameter is observed for wormlike liquid crystalline polymers at high field strength (Odijk, 2006; Spakowitz and Wang, 2003).

B.3 Calculation of global persistence length

Global persistence length was calculated by fitting our PERM data for the x projection of the mean square end-to-end distance R_x^2 with the following expression:

$$R_x^2 = \frac{1}{3}(1+2m)\left[2gL - 2g^2\left(1 - \exp\left(-L/g\right)\right)\right], \tag{B.2}$$

where g is the global persistence length and L is the contour length of the molecules. For $D \gg l_p$, $m \to 0$ and $g \to l_p$ giving rise to the free solution behavior of R_x^2 given by,

$$R_x^2 = \frac{1}{3}\left[2Ll_p - 2{l_p}^2\left(1 - \exp\left(-L/l_p\right)\right)\right], \tag{B.3}$$

as shown in Figure C.2.

B.4 Comparison between Odijk's mechanical theory and simulations

In order to determine the effect of fluctuations that were not included in Odijk's mechanical theory, we obtained the quantity H numerically via simulations. In Figure B.3, we show that H is a weak function of channel size for $0.5 < D/l_p < 1$, which is the range over which we have simulation data for g. The data for $g_{\text{Odijk}}/g_{\text{sim}}$ yields a fitted value of 135.365.

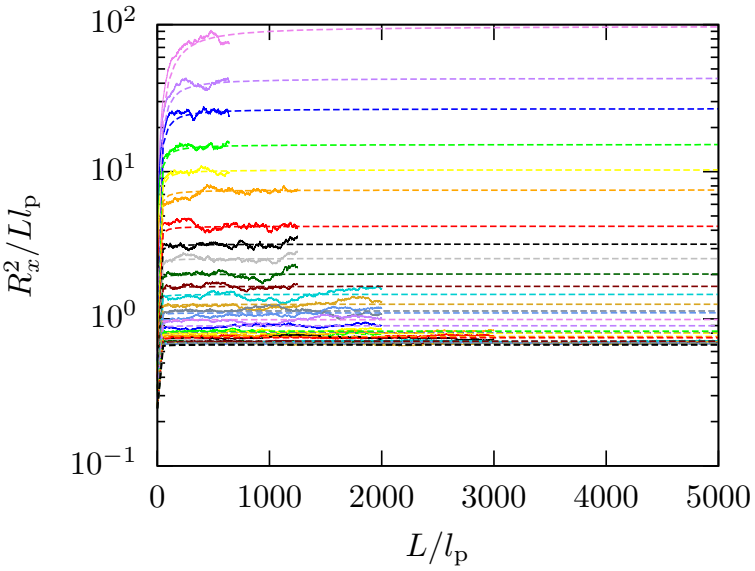


Figure B.2. Data for Figure 2 of the manuscript. RMS end-to-end distance against the contour length normalized by the native persistence length for ideal wormlike chains with $l_\mathrm{p}/a = 10$. The solid lines show raw data from PERM simulations. The colored dashed lines are the results from the fitted value of g. The channel sizes from top to bottom correspond logarithmically spaced values between $D/a = 5.0$ to $D/a = 200.0$. As D increases all the curves converge to the black dashed line that corresponds to Eq. C.2. Eq. C.1 was fit only to data satisfying $L/l_\mathrm{p} \geq 300$.

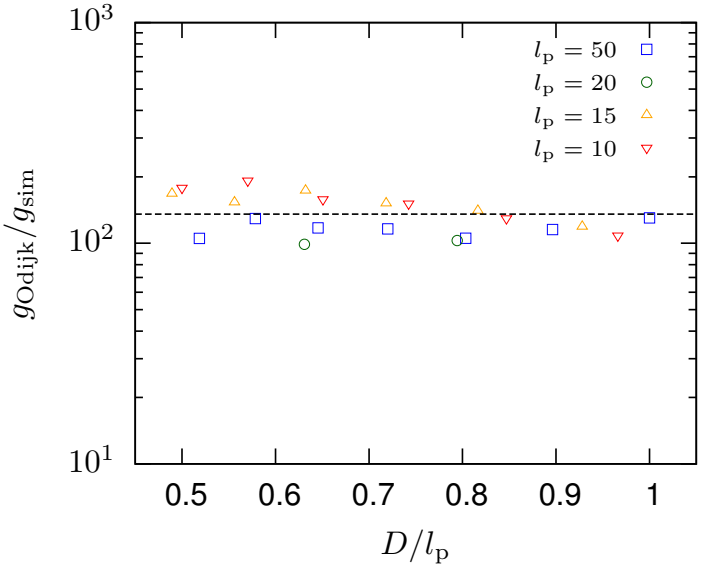


Figure B.3. Correction to the free energy of hairpin formation, H, as a function of channel size D/l_p). The dashed line is $g_\mathrm{Odijk}/g_\mathrm{sim} = 135.365$.

B.5 Contour length dependence on averages

Figure B.4 shows the contour length dependence of the mean extension and variance of the mean extension for a typical wormlike chain. It can be clearly seen that the mean extension reaches the asymptotic limit much quicker than the variance. Furthermore, the data for variance is much more noisy as it is the second moment of extension. The data for variance eventually reaches a plateau, albeit with considerble error. For all the simulations we ran for this work, the contour lengths used in our simulations were good enough to show asymptotic behavior in the mean extension. However, not all chains reached the limiting value for the variance. In this case, we only used the values that reached the asymptotic limit.

B.6 Failure to collapse data in the blob regimes

Figure 5 of Chapter 5 shows the collapse of the data for $\xi \ll 1$ and $D_{\text{eff}} < 2l_{\text{p}}$, which should correspond to the backfolded Odijk regime. Figure B.5 shows all of the remaining extension data plotted as a function of ξ. It is apparent that the gray data, corresponding to channels where $D_{\text{eff}} > 2l_{\text{p}}$, fail to follow the Odijk theory because they are in the blob (de Gennes) regimes. Note that the variable ξ accounts for the strength of excluded volume between deflection segments in a hairpin. The value of ξ loses its physical significance outside the Odijk regimes, because the properties of the confined chain are not dictated by deflection segments any more.

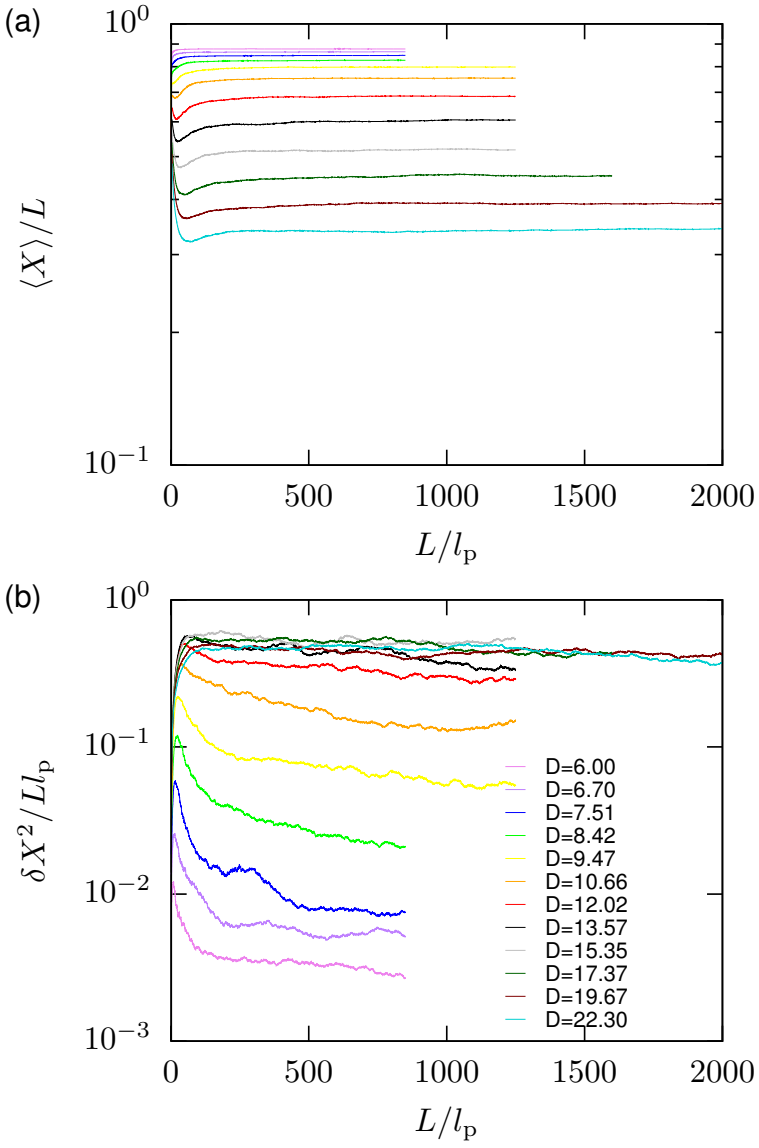


Figure B.4. (a) Normalized extension as a function of dimensionless contour length for real chains with l_p/w = 10. (b) Normalized variance versus contour length. The channel sizes and contour lengths in both panels are exactly the same. The channel size D is in units of w.

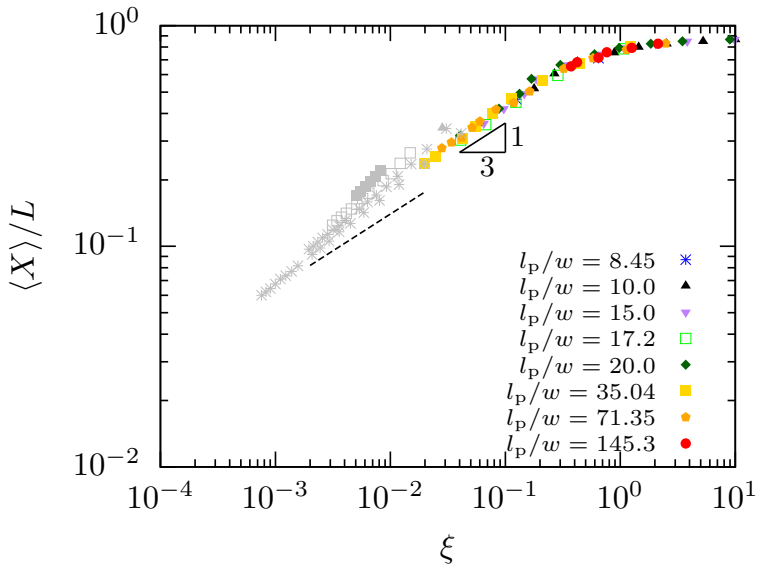


Figure B.5. Fractional extension as a function of ξ for data in all the regimes. The gray data points are for channels where $D_{\text{eff}} > 2l_p$. The black dashed line alongside the gray points shows the supposed $1/3$ scaling of the backfolded Odijk regime.

Appendix C

Supporting Information to Chapter 6

C.1 Global persistence length calculation

We calculated the global persistence length by fitting our PERM data for the x-projection of the mean square end-to-end distance R_x^2 of ideal chains with the following expression (Spakowitz and Wang, 2003):

$$R_x^2 = \frac{1}{3}(1+2m)\left[2gL - 2g^2\left(1-\exp\left(-L/g\right)\right)\right], \tag{C.1}$$

where g is the global persistence length and L is the contour length. The free solution behavior of R_x^2 is given by $m \to 0$ and $g \to l_\mathrm{p}$ leading to the one-dimensional Kratky-Porod result

$$R_x^2 = \frac{1}{3}\left[2Ll_\mathrm{p} - 2l_\mathrm{p}^2\left(1-\exp\left(-L/l_\mathrm{p}\right)\right)\right]. \tag{C.2}$$

Note that m, the orientational order parameter, was calculated in the same simulations. Figure C.1 shows our data for the order parameter in the range of tube sizes used here. Figure C.2(a) shows our simulation data alongside the fitting curves estimating the global persistence length.

Odijk employed a mechanical model to estimate the global persistence length for ideal chains confined in circular tubes (Odijk, 2006). Odijk proposed that

$$g = \alpha\bar{r}\exp\left(\bar{F}_\mathrm{mc}/kT\right), \tag{C.3}$$

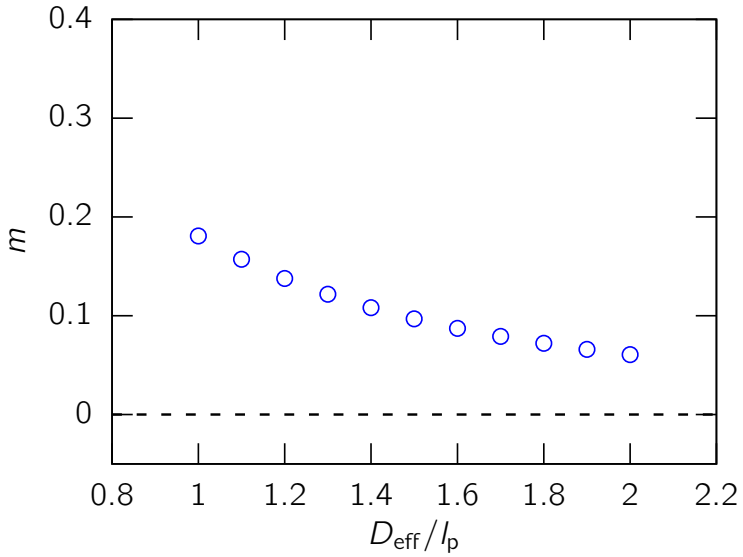


Figure C.1. Order parameter for ideal chains with $l_{\mathrm{p}}/a = 10$, where a is the bond length of the discretized model for the range of channel sizes in the main text. The horizontal dashed line is the order parameter ($m = 0$) for isotropic chains, which is expected for $D_{\mathrm{eff}} \gg l_{\mathrm{p}}$.

where $\alpha = 3.3082$,

$$\frac{\bar{F}_{\mathrm{mc}}}{kT} = \frac{E_m l_{\mathrm{p}}}{\bar{r}} - \frac{3}{2}\ln\left(\frac{D-\bar{r}}{D}\right) - \ln\left(\frac{8\sqrt{2}}{3\pi}\right) \tag{C.4}$$

and

$$\bar{r} = \frac{1}{3}\left[\left(E_m^2 l_{\mathrm{p}}^2 + 6E_m l_{\mathrm{p}} D\right)^{1/2} - E_m l_{\mathrm{p}}\right]. \tag{C.5}$$

Here $E_m = 1.5071$ (Odijk, 2006). Not long ago, we proposed that neglecting fluctuations led Odijk to overestimate the global persistence length by about 2 orders of magnitude for confinement in square channels (Muralidhar et al., 2014a). We observe a similar occurence here even for confinement in circular tubes in Figure C.2(b). However, the effect of fluctuations appears to be smaller for circular tubes compared to that of squares, presumably due to the larger surface area available to the hairpin in a square cross-section as it aligns along the diagonals. We find that the relation

$$g = \alpha\bar{r}\exp\left(\bar{F}_{\mathrm{mc}}/kT - 2.7\right), \tag{C.6}$$

leads to reasonable agreement with our simulation data in Figure C.2(b).

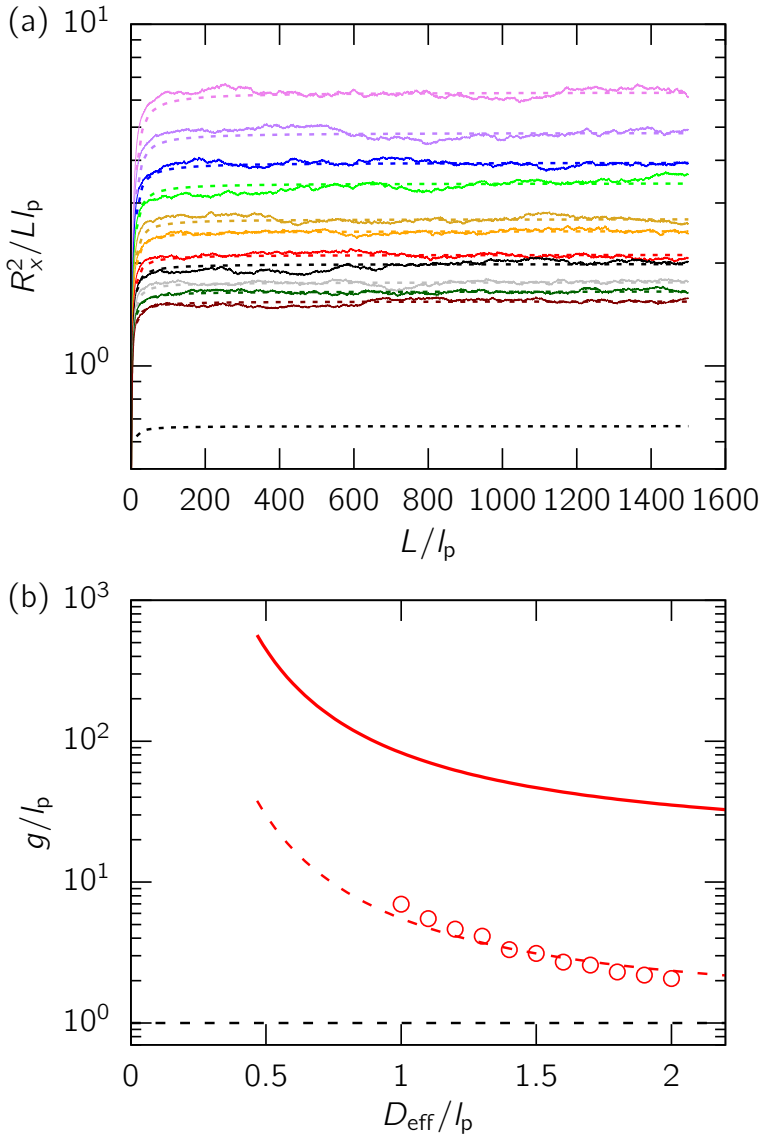


Figure C.2. Data used for Figure 3 of the main text. (a) Square of the x-projection of the end-to-end distance as a function of contour length. The results are for an ideal chain with $l_p/a = 10$, where a is the bond length in the discretized model, confined in a channel with following D_{eff}/l_p values: 1 (violet), 1.1 (purple), 1.2 (blue), 1.3 (green), 1.4 (gold), 1.5 (orange), 1.6 (red), 1.7 (black), 1.8 (gray), 1.9 (dark green), 2.0 (brown). The solid lines indicate data from PERM simulations. The colored dashed lines are the results from the fitted value of g. The black dashed line corresponds to free solution (Eq. C.2). Eq. C.1 was fitted only for data with contour lengths $L/l_p > 350$. (b) Resulting values of the global persistence length calculated from simulations. The solid red line was plotted from Eq. D.15 and the dashed line from Eq. C.6. The horizontal black line represents the relation of $g = l_p$ when $D \gg l_p$.

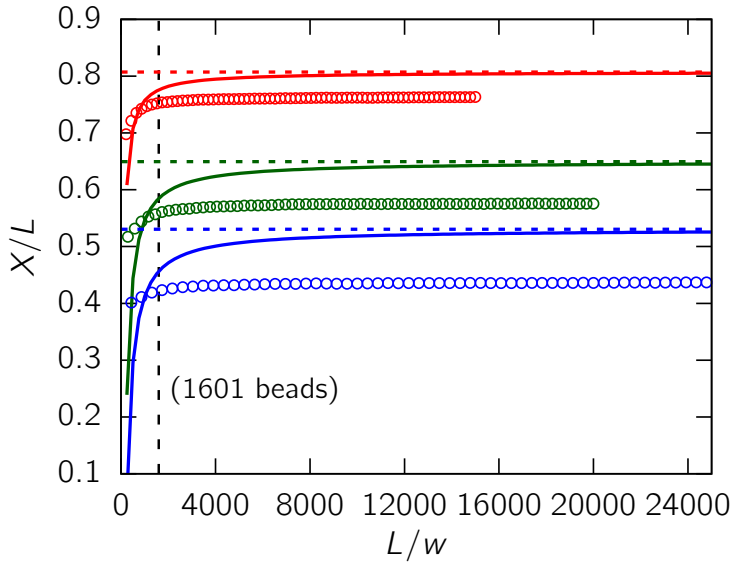


Figure C.3. Fractional extension plotted against contour length for 3 tube sizes: $D_{\text{eff}}/l_{\text{p}} = 1$ (red), $D_{\text{eff}}/l_{\text{p}} = 1.4$ (green) and $D_{\text{eff}}/l_{\text{p}} = 1.8$ (blue). Note the similarity to Figure 5 of the main text, which corresponds to $D_{\text{eff}}/l_{\text{p}} = 2$. The circles are PERM simulation data, the dashed lines represent the S-loop model and the solid lines represent the prediction of the C-loop model. The vertical dashed line corresponds to the contour length for which the C-loop model was parameterized (1601 beads).

C.2 Contour length dependence of fractional extension

Figure C.3 shows the fractional extension as a function of contour length for 3 other tube sizes in the range $1 \leq D_{\text{eff}}/l_{\text{p}} \leq 2$ along the lines of Figure 5 in the main text. We again observe that the C-loop curve intersects our simulation data around the contour lengths for which the model was parameterized (1601 beads). In addition, the fractional extension in the long-chain limit systematically deviate from the predictions of the S-loop model. This suggests that the trend observed in Figure 5 of the main text extends to the other channel sizes in the range of valididty of the C-loop model (Dai et al., 2012).

Appendix D

Supporting Information to Chapter 7

D.1 Dependence of the scaling of extension on the aspect ratio of the channel

The deflection segment length is governed by the smallest channel dimension D and is given by

$$\lambda \simeq D^{2/3} l_{\mathrm{p}}^{1/3}. \tag{D.1}$$

The excluded volume between deflection segments is

$$v_{\mathrm{ex}} \simeq \lambda^2 w \langle |\sin\delta| \rangle, \tag{D.2}$$

where δ is the angle is between two deflection rodlets. Odijk (Odijk, 2008) estimated the scaling of $\langle |\sin\delta| \rangle$ by calculating the average orientation of deflection segments assuming that the confined chain is in a harmonic well with an appropriate potential. The resultant scaling is

$$\langle |\sin\delta| \rangle \simeq \left[\frac{G_D + G_A + 1}{(G_D + 1)(G_A + 1)} \right]^{1/2}. \tag{D.3}$$

If $D \leq A \leq 2l_{\mathrm{p}}$, the two terms in the latter equation are

$$G_D \simeq \left(\frac{l_{\mathrm{p}}}{D} \right)^{2/3} \tag{D.4}$$

and

$$G_A \simeq \left(\frac{l_\mathrm{p}}{A}\right)^{2/3}. \tag{D.5}$$

For a square, we have $G_D = G_A$. We further assume that the condition for the Odijk regime leads to $G_D \gg 1$. As a result, the leading-order scaling of $\langle|\sin\delta|\rangle$ for a square channel of size D is given by

$$\langle|\sin\delta|\rangle \cong \left[\frac{2G_D}{G_D^2}\right]^{1/2} \cong \left[\frac{1}{G_D}\right]^{1/2} = \left[\left(\frac{D}{l_\mathrm{p}}\right)^{2/3}\right]^{1/2} = \left(\frac{D}{l_\mathrm{p}}\right)^{1/3}, \tag{D.6}$$

which is Eq. 5 of Odijk (2008). For a high aspect ratio channel where $A \gg D$, we have $G_A \ll G_D$. Accordingly, the leading-order behavior is

$$\langle|\sin\delta|\rangle \cong \left[\frac{G_D}{G_D G_A}\right]^{1/2} = G_A^{-1/2} = \left(\frac{A}{l_\mathrm{p}}\right)^{1/3}. \tag{D.7}$$

This approximation is used to obtain Eq. 5 of Chapter 7 and the resultant scaling of extension in Eq. 6.

We now compute the errors arising from the approximations in Eq. D.7. Assume that $A = nD$. This means that

$$G_D \simeq \left(\frac{nl_\mathrm{p}}{A}\right)^{2/3} = n^{2/3} G_A. \tag{D.8}$$

If we continue to consider only the leading-order terms by assuming that the channels are small compared to the persistence length in all cases, then we have

$$\langle|\sin\delta|\rangle \simeq \left[\frac{G_D + G_A}{G_D G_A}\right]^{1/2}. \tag{D.9}$$

Substituting for G_D gives

$$\langle|\sin\delta|\rangle \simeq \left[\frac{(n^{2/3}+1)G_A}{n^{2/3}G_A^2}\right]^{1/2}. \tag{D.10}$$

We see that $n \gg 1$ is just the large-channel result we derived earlier,

$$\langle|\sin\delta|\rangle \simeq \left[\frac{n^{2/3}G_A}{n^{2/3}G_A^2}\right]^{1/2} = G_A^{-1/2} \simeq \left(\frac{A}{l_\mathrm{p}}\right)^{1/3} \tag{D.11}$$

If n is not too big, then we have

$$\langle|\sin\delta|\rangle \simeq \left(\sqrt{1+n^{-2/3}}\right) G_A^{-1/2} \tag{D.12}$$

This prefactor is not an issue at the scaling level — for the simulations we have the errors in the table below:

$n = A/D$	$\sqrt{1+n^{-2/3}}$
1	1.414
1.5	1.328
2	1.277
3	1.217
4	1.182

Therefore, the approximation

$$\langle|\sin\delta|\rangle \cong \left(\frac{A}{l_\mathrm{p}}\right)^{1/3} \tag{D.13}$$

has very little effect on the prefactor associated with the scaling of extension in Eq. 6 of Chapter 7.

Finally, if $D \leq 2l_\mathrm{p}$ and $A > 2l_\mathrm{p}$, $G_A \approx 0$. Therefore,

$$\langle|\sin\delta|\rangle \simeq \left[\frac{G_D+1}{G_D+1}\right] \approx 1, \tag{D.14}$$

which results in the scaling of extension given in Eq. 23 of Chapter 7.

D.2 Comparison with Odijk's theory for global persistence length

Odijk employed a mechanical model to estimate the global persistence length for confinement of ideal wormlike chains in channels with different kinds of cross sections: (i) square (ii) rectangle and (iii) circle (Odijk, 2006). For rectangular channels, the results of Odijk's mechanical model are outlined in Eqs. 14-16 of Chapter 7 (Odijk, 2006). In Figure D.1, we observe that predictions of the mechanical model (Eqs. 14-16 of Chapter 7) overestimates g by about 2 orders of magnitude. This is consistent with what

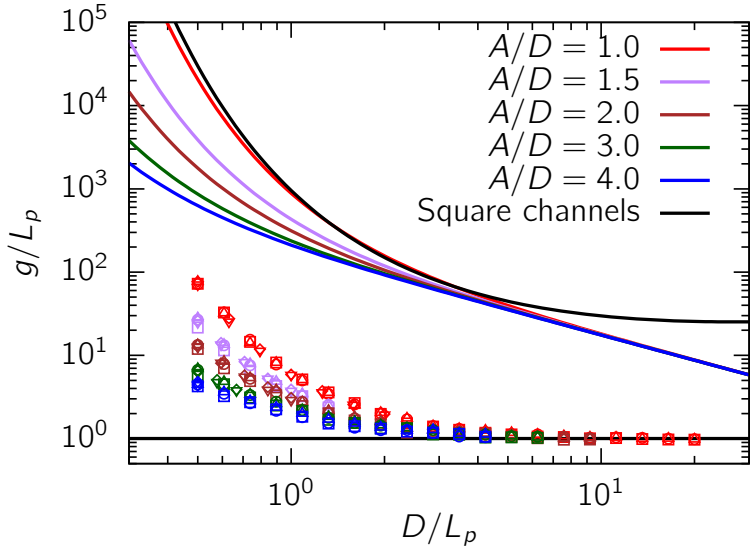


Figure D.1. Global persistence length normalized with the native persistence length of the molecule against dimensionless channel size. The five colors (red, purple, brown, green, blue) represent the five aspect ratios considered here. The different point types correspond to $l_p/a = 10$ (□), $l_p/a = 12.5$ (▽), $l_p/a = 15$ (○), $l_p/a = 17.5$ (◇) and $l_p/a = 20$ (△). The colored lines are from Eqs. 14-16 of Chapter 7. The black curve is the prediction for square channels (Eqs D.15-D.17). The horizontal black line indicates $g = l_p$, which should be the limiting value of g for $D \gg l_p$.

we observed in relation to the mechanical model for confinement in square channels (Muralidhar et al., 2014a) and circular tubes (Muralidhar and Dorfman, 2016).

Odijk derived a similar set of equations that govern the dependence of g on the channel size, D, for confinement in *square* channels as shown below.

$$g = 3.3082\bar{r}\exp\left(F_{\rm mc}/kT\right), \tag{D.15}$$

$$\frac{F_{\rm mc}}{kT} = \frac{1.5071 l_{\rm p}}{\bar{r}} - 3\ln\left(\frac{D-\bar{r}\sqrt{2}}{D}\right) - \ln\left(\frac{8}{3\pi}\right) \tag{D.16}$$

and

$$\bar{r} = \frac{1}{6}\left[\left(2.271 l_{\rm p}^2 + 9.0426\sqrt{2} l_{\rm p} D\right)^{1/2} - 1.5071 l_{\rm p}\right]. \tag{D.17}$$

One would expect that the estimate of g in rectangular channels with aspect ratio of unity should reduce to the estimate of g for square channels. To test if this is indeed true, we compare the estimates of g for rectangles with $A/D = 1$ and that for square

channels from Eq. D.15-D.17. Although it is not obvious from the equations, we observe that the predictions of the mechanical model for g in square channels and rectangular channels with $A/D = 1$ are almost identical in Figure D.1. The two curves for square and rectangle with $A/D = 1$ deviate from each other for $D/l_p > 4$. However, the mechanical model is only valid for $D < l_p$. Therefore, the apparent disagreement between the results for square and rectangular channels for $D > 4l_p$ is unimportant.

Appendix E

Supporting Information to Chapter 8

E.1 Molecular weight dependence in the classic de Gennes regime

Fig. E.1 shows the dependence of the mobility of a confined chain as a function of molecular weight in the classic de Gennes regime. Note that the x-axis in Fig. E.1, X/H_{eff}, is a measure of the number of blobs in the confined chain, assuming that the blob size scales as H_{eff} with a proportionality factor of unity (Daoud and de Gennes, 1977; de Gennes, 1979), where $H_{\text{eff}} = H - w$. Since X, the average extension in confinement, is a monotonic function of the contour length, L, X (or X/H_{eff}) is also a measure of the molecular weight. At very low molecular weights, the channel walls have very little effect on the chain mobility, which is indicated by the proximity of the confined-chain mobility with the free solution value. The curves deviate from one another with increasing molecular weight as the chain begins to feel the effect of the friction imposed by the walls. Finally, when the chain contour length is high enough to coil into the size of a blob, the mobility (scaled with contour length in the y-axis of Fig. E.1) reaches an asymptotic value, in agreement with blob theory (Brochard and de Gennes, 1977). Although the number of beads corresponding to a blob ($X/H_{\text{eff}} \approx 1$) is approximately 2000 for this value of l_{p}/w, the transition to the asymptotic limit is rather quick in terms of the number of blobs. Just as the case with static properties (Muralidhar et al., 2014b), we observe that 4 to 5 blobs are adequate to ensure long-chain limit in the mobility.

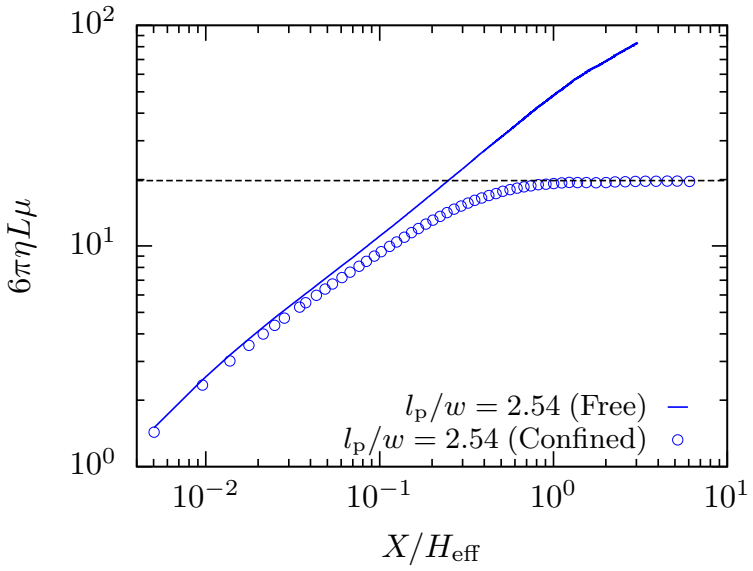


Figure E.1. PERM simulation data of chain mobility as a function of molecular weight for a chain in the classic de Gennes regime in comparison to that in free solution. The black dashed line shows the asymptotic value in the mobility. The channel size here is $H = 100w$.

E.2 Molecular weight dependence in the classic Odijk regime

As opposed to Fig. E.1, where we have used $X/H_{\rm eff}$ as the x-axis for the de Gennes regime, for the Odijk regime we plot the mobility against L/λ in Fig. E.2. In other words, we have switched the x-axis from the number of blobs to the number of deflection segments. Although the length scale for decay of HI is $H_{\rm eff}$, we observe that μ is far from the asymptotic limit when $L = H_{\rm eff}$. However, we notice that a few deflection segments ($L \gtrsim \lambda$) are adequate to ensure the asymptotic limit in μ. This is because the static properties of the molecule, such as the extension, reach the asymptotic limit only for $L > \lambda$. In other words, the molecular weight dependence of μ can be explained by the existence of the following sub-regimes of behavior: (i) the chain behaves like a freely rotating rod for $L < H_{\rm eff}$; (ii) $H_{\rm eff} < L < \lambda$ corresponds to the orientation regime, where the chain orients along the axis of the channel; and (iii) for $L > \lambda$, the chain exhibits long-chain properties of the Odijk regime (Tree et al., 2014). Because the asymptotic scaling of μ (scaled with L as shown in Fig. E.2) depends on linear ordering of deflection segments of length λ along the axis of the channel, we see this scaling for $L > \lambda$, even though HI is screened over a length scale of $H_{\rm eff} < \lambda$.

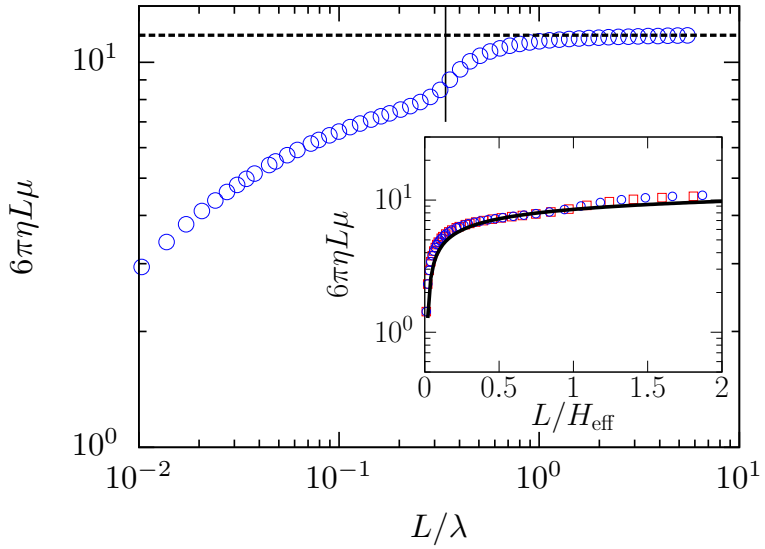


Figure E.2. Dependence of chain length on the axial mobility in the Odijk regime for a chain with $l_p/w = 2500$ confined in a channel with $H = 100w$. The horizonal black dashed line shows the long-chain limit. The truncated vertical black line corresponds to the chain length when $L = H_{\text{eff}}$. The inset shows the molecular weight dependence for short chains before they transition to the asymptotic limit. The solid black line in the inset corresponds to the relation $\mu = \ln(L/d)/(3.24\pi\eta L)$. The blue and red data points correspond to $l_p/w = 2500$ and $l_p/w = 1228$ respectively.

For lower molecular weights, we observe proof for the existence of sub-regimes as a function of molecular weight, akin to the behavior of static properties (Tree et al., 2014). For $L < H_{\text{eff}}$, the chain is weakly confined and therefore behaves like a freely rotating rod of length L (Tree et al., 2014). The inset of Fig. E.2 shows the molecular weight dependence of μ for $L < H_{\text{eff}}$. The two curves for $l_{\text{p}}/w = 2500$ and $l_{\text{p}}/w = 1228$ confined in a channel of size $H = 100w$ fall on top of each other. This clearly indicates that both chains behave like a rod of length L, regardless of the value of l_{p} as long as $L < H_{\text{eff}}$ and $H_{\text{eff}} \ll l_{\text{p}}$. The diffusivity of such a freely rotating rod of length L in bulk solution is given by (Doi and Edwards, 1986),

$$3\pi\mu\eta L \simeq \ln(L/d). \tag{E.1}$$

For short chains with $L < H$, the scaling in Eq. E.1 explains the data in the inset of Fig. E.2, with a prefactor that is remarkably close to 3. However, for longer chains with $H_{\text{eff}} < L < \lambda$, the chain begins to orient in the direction of the channel. This region is observed as a transition between the diffusion behavior of a freely rotating rod ($L < H_{\text{eff}}$) and the diffusion behavior of a long-chain ($L > \lambda$) in the asymptotic limit.

E.3 Diffusivity in the classic Odijk regime

We plot the long chain diffusivity as a function of channel size in the classic Odijk regime in Fig. E.3. For $H_{\text{eff}} \ll l_{\text{p}}$, our data agree with the scaling proposed for the classic Odijk regime (Reisner et al., 2012; Tree et al., 2012). Our data nicely fit the expression

$$\mu\pi\eta X = \frac{\ln\left(H_{\text{eff}}/d\right)}{2.46}. \tag{E.2}$$

A denominator of 2.46, which is slightly greater than the value of 2 that one would expect for the diffusivity of a rod moving parallel to itself in free solution (Doi and Edwards, 1986), indicates the higher friction imposed on the polymer by the solvent due to the boundary conditions at the walls. As long as $H \ll l_{\text{p}}$, replacing the factor of X in Eq. E.2 by L does not affect the scaling (Reisner et al., 2012), but offsets the relation by a prefactor, which is nearly unity (Yang et al., 2007).

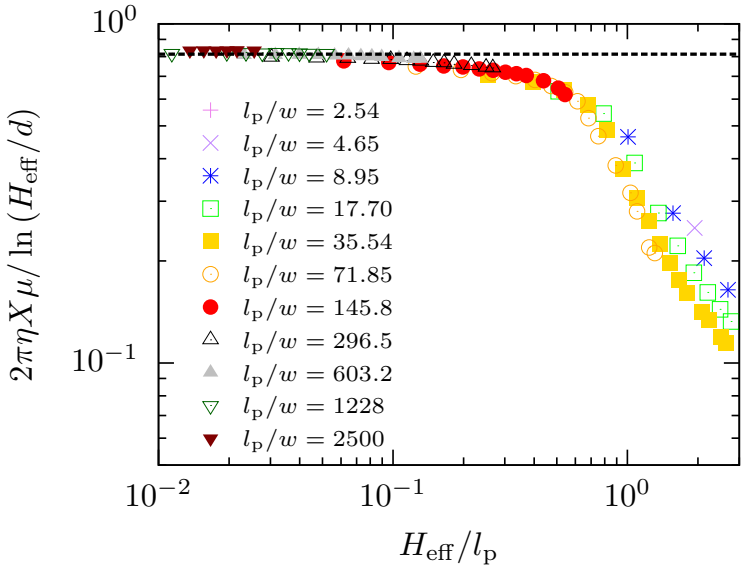


Figure E.3. Long-chain diffusivity in the Odijk regime. The black dashed line is a fit to the data for $H_{\rm eff}/l_{\rm p} < 0.05$ corresponding to a value of $\mu\pi\eta X = \ln(H_{\rm eff}/d)/2.46$.

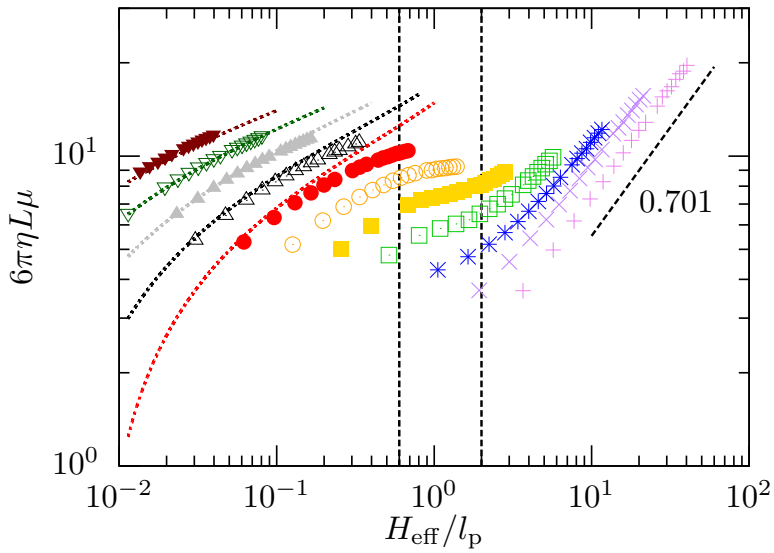


Figure E.4. Normalized mobility as a function of dimensionless channel size. The $l_{\rm p}/w$ values are 2.54 (violet +), 4.65 (purple ×), 8.95 (blue ∗), 17.70 (green □), 35.54 (gold ■), 71.85 (orange ○), 145.8 (red •), 296.5 (black △) and 603.2 (gray ▲), 1228 (dark-green ▽), 2500 (brown ▼). The colored dotted lines correspond to Eq. E.2. The dashed line on the right shows the expected 0.701 slope in the classic de Gennes regime. The two vertical lines indicate the approximate window of channel sizes where we observe relative insensitivity of $\mu\eta L$.

Table E.1: Example of data for the dimensionless mobility, $6\pi\eta\mu L$, from 5 independent data sets for a specific combination of H, l_p and w in each regime of confinement. The parameters used here are: classic Odijk (l_p/w = 2500.5, H_{eff}/d = 34.0 & H_{eff}/l_p = 0.0136), backfolded Odijk (l_p/w = 35.54, H_{eff}/d = 34.0 & H_{eff}/l_p = 0.9568), extended de Gennes (l_p/w = 17.70, H_{eff}/d = 99.0 & H_{eff}/l_p = 5.593) and classic de Gennes (l_p/w = 2.54, H_{eff}/d = 99.0 & H_{eff}/l_p = 38.98). Note that $d = w$.

Set number	Classic Odijk	Backfolded Odijk	Extended de Gennes	Classic de Gennes
Set 1	8.91570	7.32499	9.90826	18.933
Set 2	8.92542	7.34169	9.90610	20.316
Set 3	8.94148	7.39928	9.93035	19.750
Set 4	8.92901	7.26063	9.85246	19.460
Set 5	8.93359	7.35499	9.94458	19.660
Standard error	0.00428	0.02258	0.01569	0.2236
Mean	8.92904	7.33632	9.90835	19.624
Relative standard error(%)	0.048	0.308	0.158	1.139

E.4 Mobility as a function of channel size

Figure E.4 shows the normalized mobility as a function of channel size for the data points in Figure 4(a). We notice that the range of channel sizes for the region with the mobility plateau is a narrow window around H_{eff}/l_p = 1, although the fractional extension ranges from about 0.2 to 0.8. This steep change in the fractional extension in this narrow range of channel sizes is consistent with what we observe in the backfolded Odijk regime (Muralidhar et al., 2014a).

E.5 Estimation of error in mobility

For our calculations of the mobility, we divide our data set of 50 000 tours into 5 subsets (set 1 to set 5) containing 10 000 tours each. Since each tour is independent of the others, the subsets themselves are independent of each other. Note that the values of μ in the long-chain limit reported in the paper correspond to the average mobility for the longest contour length available from our simulations. We compute the standard error

of the mean from the mean mobility of each subset as illustrated in Table E.1. Although the standard error appears to be relatively high in the classic de Gennes regime, the error for all the regimes is still smaller than the symbol size used in our plots.