Modeling DNA electrophoresis in confined geometries

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

Nabil Laachi

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor Of Philosophy

August, 2010

Acknowledgements

My deepest gratitude goes to my adviser Kevin Dorfman. Thank you for all your help. For the continual trust you put in me. For your unremitting availability. For everything I learned from you in the past five years. For reading my manuscripts. For the countless hours of discussions and prolonged conversations. You have been a true source of inspiration and motivation, both on a professional and a personal level.

Thanks to the David and Lucile Packard Foundation for the financial support, to the University of Minnesota for the support through a Doctoral Dissertation Fellowship and to the University of Minnesota Supercomputing Institute for the computing resources.

Thanks to the members of the thesis committee, V. Barocas, K. Dorfman, S. Kumar and V. Noireaux, for their service.

Thanks to Professor Satish Kumar and his former student Nazish Hoda for their help with Brownian dynamics simulations in nanofilters.

Thanks to Carmelo Declet and Christina Matson who contributed to some of the work on nanofilters during their summer internship in our group.

Thanks to Jaesol Cho, a former postdoc in our group who wrote the Brownian dynamics code for the post unhooking.

Thanks to the Dorfman group members. In particular, Martin who was more a dear friend than a colleague and who likes to argue for its own sake (he will probably challenge this description too!). Maggie, without whom, the whom would be a who. OJ, thanks

to whom I know more about China and World War II than Wikipedia can ever tell me. Dan, for all the quarters and the sports updates. And Doug for caring so much.

Thanks to my friends from the Douglas house (Aaron, Christine, Dan, Dave E. and Derek) and beyond (Alison, Ben, Dave R., Emilie, John B. and Will) for their friendship and all the great moments we spent studying, partying or simply hanging out. I owe them my most sincere gratitude.

Thanks to Vanessa for her love and support over the past year and a half.

Thanks to Dan for squash, philosophy, politics, the lakes, lunches, the bike rides, tennis, Frisbee, listening, and so much more. For just being there for the past five years. Thank you so much man.

Merci à Messan, pour avoir été là quand j'avais le plus besoin d'un ami ces deux dernières années. Pour le gari, le kom et le poulet togolais. Pour son soutien moral, sans lequel plusieurs épreuves auraient été encore plus dures qu'elles ne l'avaient été.

Merci à mes parents, pour leur confiance infaillible ainsi que leurs encouragements réguliers.

Merci à mes deux soeurs, Safaa et Meryem, pour l'amour et l'affection dont elles me témoignent incessament. Merci pour leur bonne humeur qu'elles partagent avec moi.

Mais aussi, merci à MC, sans laquelle rien de cela ne se serait jamais produit.

Dedication

A ma mère et à mon père.
Pour leur amour et leur soutien indéfectibles.

A mes deux soeurs, Safaa et Meryem.
Pour la joie de vivre et la complicité.

Abstract

Size-based DNA separation is at the heart of numerous biological applications. While gel electrophoresis remains widely utilized to fractionate DNA according to their size, the method has several shortcomings. Recent advances in micro- and nano-fabrication techniques engendered several microfabricated devices aimed at addressing some of the limitations of gel electrophoresis for DNA separations. In this thesis, we employ a combination of analytical and computational methods to characterize the electrophoretic motion of DNA molecules in microfabricated, confining geometries. In particular, we consider three situations: (i) the migration of short DNA in nanofilters (a succession of narrow slits, connecting deep wells) under a high electric field; (ii) the metastable unhooking of a long DNA chain wrapped around a cylindrical post; and (iii) the dynamics of long DNA chains in an array of spherical cavities connected by nanopores. We provide insights on the physical mechanisms underlying the transport of DNA in such geometries. Useful guidelines for the optimal design of new separation devices result from the fundamental understanding gained by the approach we propose.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The separation of biomolecules such as DNA or proteins by molecular weight is at the heart of several biotechnological applications, ranging from genetics to forensics. The early completion of the genome sequencing, for instance, was enabled by the emergence of novel and more efficient techniques, such as automated, multiplexed capillary electrophoresis, that reduced both the cost and the time of separations.

Other applications of DNA analysis will certainly benefit from further development in separation techniques. In electrophoresis, the advent of such new techniques is based on a more general scheme: **miniaturization**. The basic premise is that microfabrication (or even nanofabrication) enables the downsizing of the current lab equipment and processes onto single, miniaturized, highly integrated chips, the so-called lab-on-a-chip systems, which are reputed to be efficient and robust.

A fundamental understanding of the dynamics of DNA at the micro-scale will undoubtedly benefit the design of these micro-devices and, hence, improve the separations. And it is in this spirit that we propose to model the electrophoretic migration of DNA in confining micro-devices.

Without delving into complex biological details, we present in the first section of this introductory chapter some general information and data about DNA, as well as some applications that motivate DNA size-based separation. An overview of separation techniques will follow in the second section. The last section of this introduction will be dedicated to outline the rest of the manuscript.

1.1 General motivation

1.1.1 DNA: facts and data

The history of physical and natural sciences is punctuated with groundbreaking years that laid the foundations for future developments. The publication of *Philosphiae Naturalis Principia Mathematica* by Isaac Newton in 1687, or the three revolutionary papers on the photo-voltaic effect, Brownian motion and the theory of special relativity by Albert Einstein in 1905, are prefect examples of such landmark years. In biology, the year 1953 witnessed the discovery of the molecular structure of DeoxyriboNucleic Acid (or DNA) by James Watson and Francis Crick. Today, over fifty years after the discovery of the helical structure of DNA, it has become a common knowledge that it not only

contains all the information necessary to characterize different species and single individuals in these species, but also the genetic instructions essential to the development and growth of living individuals.

In nature, DNA molecules usually exist as tightly packed and twisted double strands of paired bases, the famous double helix. The backbone of each of the DNA strands forming the helix consists of alternating phosphate (in the form of acid) and sugar (deoxyribose) residues. Because of the former, DNA backbone is negatively charged in solution. The genetic information (or genes) encoded within a DNA molecule is sequentially stored using a set of four unit blocks, the nucleotides or nitric bases: adenine (A), cytosine (C), guanine (G) and thymine (T), that are attached to each sugar group. Single stranded RNA have a similar structure, but with a different type of sugar and the base uracil (U) replacing thymine. In a DNA molecule, the two long strands are held together by hydrogen bonds linking the two pairing bases. The base pair associations are not random or made of identical bases. Rather, each base of one strand bonds with one complementary counterpart on the other strand, and the combination determines what is called a base pair. Thus, (A) in one strand pairs with (T) on the other, and similarly, one (G) base pairs exclusively with (C) from the complementary strand. As a result of this complementarity, no additional information is contained in the second strand.

To get an idea of the length scales involved in the study of DNA, consider a human cell, for example, which has a size in the range of 1–10 μm. The backbone of DNA inside each cell is 2.2–2.4 nm wide and one base pair is 0.34 nm long [1]. Each DNA contains 46 chromosomes, 23 from each parent, and can encode about 20,000–30,000 genes. The average human chromosome can contain up to a hundred million base pairs, the largest chromosome being 222 million base pairs long [2]. As a consequence, although the building units of DNA are very small, the entire chain can be enormous. A fully stretched DNA molecule can extend to about a foot!

1.1.2 Why separate DNA fragments?

Major properties of genes can be obtained via an analysis of DNA fragments based on their respective size [3, 4]. As a result, separation of DNA and other biomolecules on a molecular weight or size basis is arguably the single most important tool in genetics

and several applications are undoubtedly sensitive to improvements in DNA separation. These include genome mapping [5], the development of biotechnology or diagnosis of genetic diseases, not to mention broader applications in genetic fingerprinting [6] or DNA profiling [7]. DNA separation also enabled the completion of the Human Genome Project [8]. The project aimed at the deciphering of the gene sequence buried in a DNA molecule and was a milestone towards the understanding of life, with new insights into the properties and functions of genes. The more recently initiated proteome analysis, *i.e.*, the study of sequences of proteins, will also prove very crucial for a progress in molecular biology and will certainly benefit from advances in separation techniques.

1.2 Overview of separation techniques

1.2.1 Gel electrophoresis

In biology and analytical chemistry labs, gel electrophoresis is arguably the most widely used method for DNA separation, owing its success to the "ergonomy" of the equipment and the simplicity of the procedure; all you need is a gel, a voltage source and a method of detection. The term electrophoresis refers to the motion of charged objects, and particularly polyelectrolytes (charged polymers such as DNA) under the action of an external electric field. In typical gel electrophoresis applications, DNA travels through a gel, usually agarose or polyacrylamide, when subject to an electric field on the order of 1–10 V/cm. This often constitutes an upper limit for the efficient use of the technique.

The gel is essential to the separation. The gel fibers represent a sieving matrix that affects the overall mobility of the polyelectrolyte through steric interactions. The mobility is otherwise size-independent in free-solution. The dynamics of polyelectrolytes through gel matrices are best described using the biased reptation model, reviewed in Chapter 2. The key result that contributed to the success of the theory is that it predicts the experimentally observed $1/M$ decrease of the mobility with the molecular weight M of the chain. However, the separation of large molecules ($\geq$ 30 kilo base pairs) fails using conventional gel electrophoresis [9]. Under a constant electric field, large molecules tend to orient their leading monomers in the direction of the field, leading to a mobility that saturates to a size-independent value. Nevertheless, the separation of large DNA molecules can still be achieved in gel electrophoresis when one resorts

to pulsed-field protocols [10]. While this extends the useful range of sizes, pulsed-field gel electrophoresis remains a time consuming technique, as the separation of one chromosomal DNA for example can require up to several days [11].

1.2.2 Capillary electrophoresis

Capillary Electrophoresis (CE) is a biomolecule analysis technique that attracted much attention for its efficiency and high resolution. The separation occurs in a narrow channel with an inner diameter between 20–100 μm and a total length of several centimeters (50 cm is common). To prevent problems such as clogging or gel breakage, the separation matrices used in CE are more fluid than gels. Entangled polymer solutions are the most widely and successfully used ones [9, 12].

The technique has proven very useful and efficient in operating fast separations with higher output than conventional gel electrophoresis. The main reason for the improved output and resolution is the use of high electric fields (typically 200-300 V/cm in CE, while less than 50 V/cm in conventional gel electrophoresis), made possible by the increased Joule heat dissipation in devices with large surface/volume ratio. CE devices also benefited from the level of automation they introduced; capillary-array automated sequencers (16, 96, even 384 capillaries) allow separations at least one order of magnitude faster than slab-gel electrophoresis [13]. In fact, the launching of these types of automated devices enabled the completion of the human genome sequencing project earlier than projected [8].

This rather flawless picture of capillary electrophoresis should be nonetheless moderated if we consider the problems raised by electro-osmosis or the electro-osmotic flow, *i.e.*, the fluid motion generated by the mobile counterions close to the charged surface of the channel walls. Heterogeneities that develop next to the walls result in a reduced reproducibility and a catastrophic dispersion, both affecting the separation. This problem is usually mitigated by coating the surface with neutral polymers that increase the local viscosity near the wall and reduce the effects of the electro-osmotic flow. It is also worth mentioning that, similar to gel electrophoresis, there exists a critical size above which DNA molecules cannot be separated using entangled polymers in capillaries. However, the loss of resolution is attributed to hydrodynamic instabilities [14, 15] rather than chain orientation.

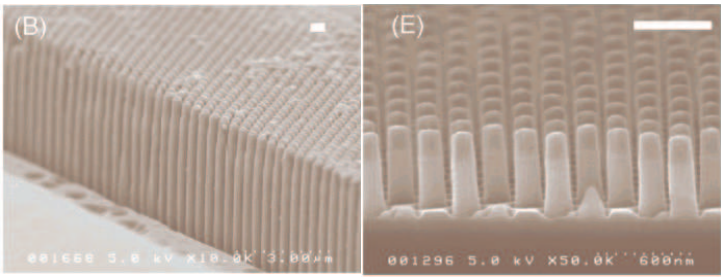

Figure 1.1: A microfabricated array of posts made of quartz. The white bars are 1 micron long. Reproduced from [23].

1.2.3 Rise of microfabrication

In a seminal paper in 1990, Manz introduced a new paradigm of electrophoresis systems (and analytical chemistry in genenal) as integrated components in a lab-on-a-chip or micro-total analysis system μTAS [16]. This original idea consisted in downscaling the existing equipment and separation techniques onto highly integrated devices [17, 18], somewhat similar to earlier developments in microelectronic fabrication. The use of such microfabricated devices was originally motivated by the increased quality of separation as the channel width becomes smaller; the number of theoretical plates, which gives a measure of the degree of separation, decreases quadratically with the width of the channel due to efficient Joule heat dissipation [19]. The reduced size of the injection plug also contributed to the early success of microfabricated devices since the resolution of the device is no longer affected by the minimum separation distance constraint from the injection plug [20]. If most of the original microfabricated devices were smaller duplicates of CE systems, the miniaturization of conventional techniques to the microscale (or even the nanoscale) was also accompanied by the emergence of a multitude of new separation devices whereupon the underlying physics was unknown on the macroscopic scale [21, 22].

Post arrays. Following the pioneering idea of Volkmuth and Austin [24], Doyle *et al.* [25] and Kaji *et al.* [23] engineered a well-defined regular array of fixed posts, with a diameter $\approx 1\mu$m, to replace the random array of fibers in a gel. A microfabricated array of posts is illustrated in Fig. 1.1. Experimental results [26] along with theoretical investigations of the mobility and the diffusivity of DNA based on a model of single-post

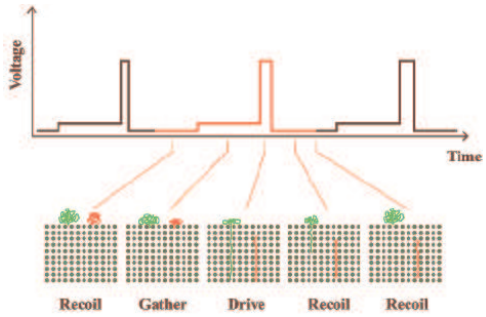

Figure 1.2: Schematic for separation based on entropic recoil, with an electric field pulse and responses of a small (red) and a large (green) molecule. Reproduced from [29].

collisions and random translation between the collisions [27,28], indicate that the resolution is independent of the electric field magnitude. In contrast to gel electrophoresis, separations of long DNA (50-150 kilo base pairs) within a few minutes were reported for an optimal post spacing on the order of 7 μm [26].

In a way to circumvent the complexities in the fabrication of the post arrays [21,22], Doyle *et al.* [25] proposed supermagnetic beads, with a diameter $\approx$ 560 nm, as an alternative to build cylindrical pillars. Under the action of a magnetic field ($\geq$ 10 mT), the homogeneous suspension of beads self-assemble to form columns, or pillars, with inter-spacing $\approx 5\,\mu$m. The pillars thus obtained act in a similar way to the array of posts.

Entropic-trapping. Further examples of microfabricated systems include entropic trapping-based devices, a concept that was introduced from simulations by Baumgärtner and Muthukumar in 1987 [30–32]. In contrast to posts arrays, the structure (or pore) size of the entropy-based systems is comparable to the radius of gyration (a measure of the spatial extent) of the molecules ($\approx 1\,\mu$m). As a consequence, internal conformational entropy effects can be effectively used to sort DNA molecules based on their size. In other words, DNA molecules can get trapped in high conformational entropy regions with an escape rate that is related to their size. A successful manifestation of entropy-based effects resulted from experiments by Cabodi *et al.*, members of the Craighead group at Cornell [29, 33, 34]. In these experiments, a dense array of nanopillars was used to trap DNA molecules. When the electric field is turned off, larger molecules exit the array by "entropic recoil" [29], disengaging towards the pillar-free entrance region

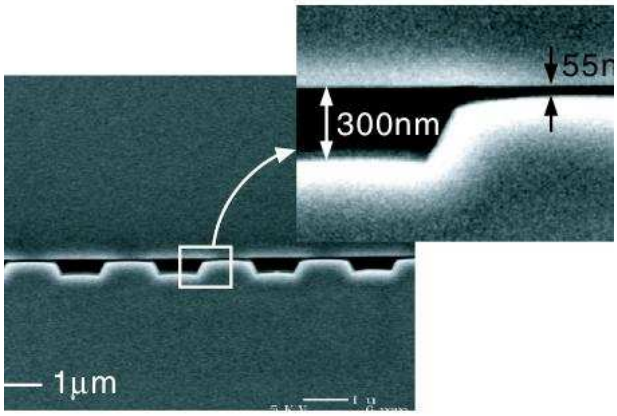

Figure 1.3: SEM micrographs of a nanofilter consisting of alternating deep and shallow regions. Reproduced from [35].

where the entropy is maximized. This principle is described in Fig. 1.2.

The nanofilter. This device, originally called an entropic trap, first appeared following the work of Han from the same Craighead group [36, 37] and consists of a succession of wells, with a depth of 300-500 nm, connected by narrow slits, with a size of 50-100 nm. An SEM image of a nanofilter is shown in Fig. 1.3. In the low field regime, the device ability to perform fast separations of large and small DNA molecules was demonstrated both on experimental [36–38] and theoretical [35, 39] grounds. The separation was attributed to the effects that originate from the loss of entropy when a molecule enters the shallow region. The rationale behind this thermodynamic argument is the necessity of operating the device close to equilibrium conditions where the electric field introduces a weak bias to the diffusion-driven random fluctuations. More generally, Giddings [40] reported that most chromatographic separations are near-equilibrium processes, where the motive force responsible for the transport of the solute molecules (fluid flow for example, in the case of chromatography), is slow enough compared to molecular diffusion. Indeed, the equilibrium requirement seems to be a *leitmotiv* in much of the separation science where the transport of solutes is dominated by a weakly biased thermal energy, introducing a trade-off between the throughput and the resolution of separations.

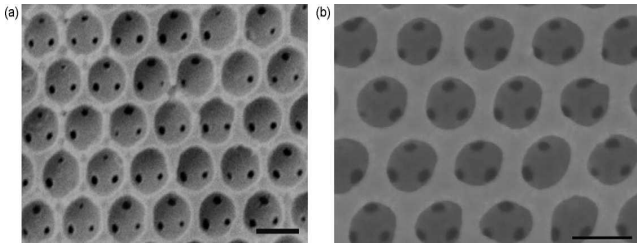

Figure 1.4: SEM images of nanoporous structures constructed from silica particles with diameter (a) 300 nm and (b) 570 nm. Reproduced from [41].

1.2.4 Self-assembly

As briefly mentioned above in the case of post arrays, one can take advantage of the self-assembly properties of supermagnetic beads to create pillars comparable to post arrays without the use of sophisticated lithographic techniques [25]. Based on the same self-assembly paradigm, numerous ordered sieving structures have been proposed for DNA separation. The idea is essentially to induce the self-assembly of colloidal particles (by evaporation [42] or electro-capillary forces [43, 44], for example), to construct a well-ordered nanostructure that can be used to sort DNA molecules. While earlier devices focused on the void spaces between the particles as a sieving matrix in the microfluidic system [42, 45–48], the inverse structures, consisting of well-ordered spherical cavities, connected by narrow pores, are also currently used for DNA separations [49], as well as the study of single-molecule behavior in confining geometries [41, 50, 51]. These periodic arrays of spheres, known as inverse opals, can be fabricated from the colloidal crystal template by first depositing a continuous phase in the gaps of the crystal and then removing the colloids by chemical etching [52]. An array fabricated in this manner is shown in Fig. 1.4. While the vast majority of applications for inverse opals are in photonics [52], if the continuous phase is a polymer [51] or photoresist [41], then the material is compatible with DNA electrophoresis. Depending on the size of the colloids used for the template, the cavity diameters are several hundred nanometers, while the pore diameters are tens of nanometers. Finally, it is worth mentioning that, while inverse opals currently used for DNA electrophoresis are face-centered cubic arrays of spherical cavities, other geometries of photonic crystals have been constructed [53, 54]. Examples

include simple cubic array [53, 54], diamond-like and gyroid-like structures [53].

1.3 Research outline

The broad goal of this thesis is to study the dynamics of DNA molecules in three different situations, the nanofilter, unhooking from a single post and inside an array of spheres. We ultimately seek to provide useful guidelines in the design of micro-devices for DNA size-based separations, focusing in our approach on a combination of theoretical and computational methods.

Chapter 2 of this document is dedicated to brief discussions of the physical concepts of which a well-contained understanding is necessary in order to pursue our goal. Some aspects of polymer physics, as well as the concept of nanopore translocation will therefore be introduced. Other topics, such as Brownian dynamics simulation or Monte Carlo simulation will appear, when necessary, with greater details in the appropriate chapters. The fundamental results of our research will be discussed in detail in Chapters 3-6. A brief summary is presented below. A conclusion and some final thoughts will follow in Chapter 7.

Brief summary of the research

Nanofilters for DNA Separation; A Non-equilibrium Study (Chap. 3)

We consider theoretically the dynamics of short duplex DNA during high-field electrophoresis through a periodic array of narrow slits and deep wells (a nanofilter), where the slit depth is less than the contour length of the essentially rigid DNA strand. In contrast with the known behavior under weak fields, we predict that the larger chains will elute first under strong electric fields via "torque-assisted escape" from the wells. This contradicts the maxim that separations must be performed close to equilibrium, and opens the way for enhanced nanofluidic separations of DNA based upon their out-of-equilibrium transport properties

DNA unhooking from a single post (Chap. 4)

Using stochastic methods developed for DNA translocation through nanopores, we study the unhooking of a long DNA chain from an isolated, stationary micro-post. Such methods quickly and efficiently furnish both the full probability distribution of the unhooking time and the ensuing moments for a wide range of chain and field parameters. The results compare favorably with more realistic but computationally intense Brownian dynamics simulations. For typical chain lengths and applied electric fields used in experiments, the unhooking process is effectively deterministic; diffusive fluctuations make a negligible contribution to the first and second moments of the unhooking time. This result lends credence to continuous-time random walk models of the overall transport process that treat the unhooking as a convective process.

Partition function of chains with fixed ends inside a spherical cavity (Chap. 5)

We compute the partition function of self-avoiding chains tethered inside a confining sphere using Monte Carlo simulations on a three-dimensional lattice. Two cases are considered: i) single-tethered (ST) chains, with one end anchored and one end free, and ii) double-tethered (DT) chains, where both ends are tethered at a distance equal to the diameter of the sphere. The self-avoidance, confinement and tethering constraints dramatically decrease the number of allowed configurations when compared to an unconstrained random coil, thereby affecting the sampling method used in the Monte Carlo procedure. The effect of an external applied force and the bias it introduces in the partition function are also investigated. Our method involves a decomposition of the partition function into the product of several terms that can be evaluated independently. For short chains, we demonstrate the validity of our approach through a direct evaluation of the partition function using an exact enumeration of the appropriate paths on the lattice. In the case of long chains, scaling laws for the behavior of the partition function are identified.

Transport of DNA in a array of spheres; A new lakes-straits model (Chap. 6)

We present a method to study the dynamics of long DNA molecules inside a cubic array of confining spheres, connected through narrow openings. Our method is based

on the coarse-grained, lakes-straits model of Zimm and is therefore much faster than Brownian dynamics simulations. In contrast to Zimm's approach, our method uses standard stochastic kinetic simulation to account for the mass transfer through the narrow straits and the formation of new lakes. The different rates, or propensities, of the reactions are obtained using first-passage time statistics and the Monte Carlo sampling of Chap. 5 to compute the total free energy of the chain. The total free energy takes into account the self-avoiding nature of the chain as well as confinement effect from the impenetrable spheres. The mobilities of various chains agree with biased reptation theory at low and high fields. At moderate fields, a new regime of reptation is revealed where the mobility is a linear function of molecular weight and the dispersion is minimal.

Chapter 2

Some background

2.1 Polyelectrolytes in free solution

A polyelectrolyte is a polymer (a long molecular chain) that carries an electric charge. The chain consists of a sequence of simple chemical units, called monomers, that are linked together by covalent bonds. In some cases (when dissolved in a polar solvent such as water), the monomer units can lose some ions, called counterions, that dissociate from the chain. As a result, the chain acquires an electric charge opposite to that of the counterions. Examples of polyelectrolytes include polyacrylic and polymethacrylic acid, which are synthetic polymers, and DNA or protein molecules that become electrically charged biopolymers when dissolved in water. Polyelectrolytes are thus a class of polymers, and any attempt to describe physical properties of polyelectrolyte should be preceded by an understanding of the physical aspects of neutral polymers. We restrain our discussions to single linear polymers, and omit other topologies such as branched polymers or cross-linked polymers. Although they involve some interesting properties, high density polymer solutions, such as entangled solution or melts, are also not considered below.

2.1.1 Neutral single chain: Static conformations

Flexibility

A linear polymer is a long chain of chemically linked repeat units. The number of repeat units or monomers N that form the chain, also called the degree of polymerization, ranges from 10^2 to 10^4 for synthetic polymers and can reach up to 10^9 or 10^{10} for biopolymers such as DNA. The length of the fully stretched chain, called the contour length L, is proportional to the degree of polymerization. For a human DNA for example, a single chain in a cell can contain about 1 billion base pairs. If each base pair is about 3 angstroms long, then the contour length of the entire chain is about 30 cm. Obviously, the chain would not fit into the cell if the chain were to be fully stretched. Some bending of the chain is thus necessary. Indeed, if we consider an unconstrained polymer molecule in a solution, the chain entangles to form a random, three-dimensional coil under the continuous bombardment from the solvent molecules[1] . Synthetic polymers

[1] A three-dimensional Gaussian coil of DNA is still too big to fit inside a human cell. Condensation to chromatin (by histones) is important in this case.

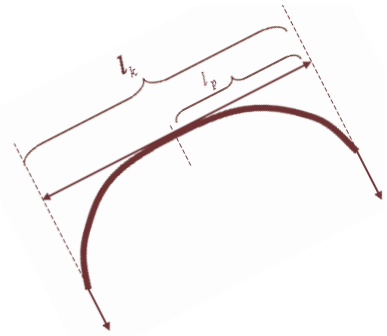

Figure 2.1: The persistence length is measured as the distance along the backbone, in a given direction, before the chain orientation changes by 90°. Since the memory of the orientation extends in two directions, the Kuhn length is defined as twice the persistence length.

will bend because parts of the chain will rotate around the single C−C bonds while a molecule like a double-stranded DNA, will bend by deformation of the angle between the bonds.

A quantitative description of the flexibility of the chain can be achieved by introducing the so-called persistence length l_p. For length scales shorter than the persistence length, the memory of chain direction is preserved; the chain appears to be rigid. In contrast, for length scales that exceed the persistence length, all memory of direction is forgotten and different portions of the chain can have different and uncorrelated directions; the chain bends. Since the chain memory extends in both directions, a related quantity to l_p is also commonly used; the Kuhn length l_k, defined as twice the persistence length, $l_k \equiv 2l_p$ [55]. The two lengths are schematically defined in Fig. 2.1.

Characteristic values of l_p can range from a few angstroms[2] for flexible polymers such as polyethylene and polystyrene to several hundred angstroms for semi-flexible polymers such as poly(*p*-phenylene). Double-stranded DNA in a solvent like salty water is also semi-flexible, with a persistence length of order 50-60 nm, *i.e.*, about 150 base pairs per persistence length.

Different models with different degrees of complexity, were introduced for the description of polymer flexibility, among which are the freely-jointed chain, the bead-spring

[2] One C−C bond is ≈ 1.54 Å

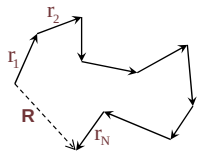

Figure 2.2: Freely jointed model. The orientation of segment is random and totally independent from that of the other segments.

model, the freely rotating model or the worm-like chain. Although based on different assumptions, these models are equivalent in predicting the universal properties of a polymer chain and their dependence on the polymer length or molecular weight. Therefore, in the following discussion, we adopt the simplest of these, the freely jointed chain.

Static conformations

We consider a freely jointed polymer chain consisting of N_k Kuhn segments and of contour length $L = N_k l_k$, as illustrated in Fig. 2.2. We also make the assumption that long range interactions between segments can be neglected, *i.e.*, segments can intersect and occupy simultaneously identical positions. Short range interactions between the monomers are allowed, but do not affect the general results if the Kuhn length l_k is replaced by an effective length l_{eff} [56]. As a result, in the following, short range interactions are also neglected. This is the model of an ideal chain. In these circumstances, the chain behaves as a random walk where the direction of each segment is chosen randomly and independently of other segments. A naive estimation of the size of the coil is given by taking the average of the end-to-end vector $\mathbf{R}_{N_k}$.

For a long chain where $N_k \gg 1$, the central limit theorem ensures that the conformation obeys Gaussian statistics, *i.e.*, the chain end-to-end vector $\mathbf{R}_{N_k}$ is determined according to a Gaussian probability density. In particular, if $\mathbf{r}_i$ is the end-to-end vector for the ith segment, then the end-to-end vector $\mathbf{R}_{N_k} = \sum_{i=1}^{N_k} \mathbf{r}_i$ has a zero average value

$$\langle \mathbf{R}_{N_k} \rangle = 0, \tag{2.1}$$

since each segment can point in one direction or its opposite with equal probabilities.

A more realistic estimation of the size of the coil is given by the mean square end-to-end distance $R^2 \equiv \langle \mathbf{R}_{N_k} \cdot \mathbf{R}_{N_k} \rangle$ According to the definition of $\mathbf{R}_{N_k}$, we can write

$$\mathbf{R}_{N_k} \cdot \mathbf{R}_{N_k} = \sum_{i=1}^{N_k} \mathbf{r}_i \cdot \mathbf{r}_i + \sum_{i=1}^{N_k} \sum_{j \neq i} \mathbf{r}_i \cdot \mathbf{r}_j. \tag{2.2}$$

The first term in the right-hand side of the previous equation is simply $N_k {l_k}^2$ since $\mathbf{r}_i \cdot \mathbf{r}_i = {l_k}^2$ for every segment i. The second term will be zero on average, since for a freely jointed chain, the directions are independent of each other, and thus the cosine of the angle between $\mathbf{r}_i$ and $\mathbf{r}_j$ takes all values between -1 and 1, with equal probability, for every $i \neq j$. Accordingly, the mean square end-to-end distance reduces to

$$R^2 = N_k {l_k}^2 = L l_k, \tag{2.3}$$

using the definition of the Kuhn segment.

The coil size thus estimated is generally hard to measure experimentally. A more convenient estimation of the size of the coil can be considered using the radius of gyration R_g [57]. It represents the average distance between the segments and the center of mass of the polymer. A straightforward calculation shows that [56]

$$R_g^2 = \frac{R^2}{6} = \frac{N_k l_k^2}{6}. \tag{2.4}$$

The square-root law ($R \sim R_g \sim N_k^{1/2}$) described above for ideal chains is not limited to the freely jointed model; it is rather a universal law. Other models, such as the freely rotating chain or the models including the effect of short range interactions lead to a similar scaling law. It is valid for all ideal chains and the 1/2 exponent does not depend on the chemical structure of the chain. The molecular details only intervene through a prefactor in Eqs. (2.3) and (2.4). Also, it is worth mentioning that this exponent appears in the Brownian motion of a small particle, where the displacement b of the particle in a time t, is proportional to the square-root of t. Analogous to a Brownian trajectory, an ideal chain coil exhibits a statistical self-similarity (or fractal) behavior, *i.e.*, any magnified small portion of the chain reveals the same properties, in a statistical sense, as the entire chain.

In physical situations, the chain segments cannot simultaneously occupy the same region in space. Ideal chain models take into account, at best, the short range interactions between segments located close to each other along the chain backbone. In real

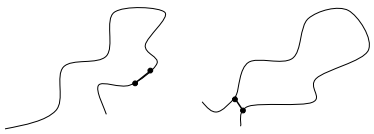

Figure 2.3: Short (left) and long (right) range interactions between different portions of the chain.

chains, repulsive interactions appear when widely separated segments along the chain get spatially closer to each other [see Fig. 2.3]. These long range interactions are also known as the excluded volume effect, where the excluded volume is the whole allowable volume for intersection with other segments [58]. In a good solvent, the excluded volume v_F consists of the "pancake" obtained by rotating the segment around its center of mass and is given by

$$v_F = l_k^2 d, \tag{2.5}$$

where d is the diameter of a Kuhn segment. As a result of these repulsive effects, the chain is no longer Gaussian, but rather swells and behaves as a self-avoiding walk. The end-to-end distance now scales as

$$R_F = l_k \left(\frac{v_F}{l_k^3} \right)^{2\nu-1} N_k^{\nu}, \tag{2.6}$$

where the Flory exponent ν is approximately 3/5 [59]. The scaling above remains valid as long as the Flory radius R_F for a swollen chain is larger than the average end-to-end distance for a Gaussian chain as given by Eq. (2.3). This condition translates into a critical chain size for the excluded volume effect to occur, obtained by setting $R_F = R$. It follows then, that the critical chain size is given by

$$N_k^{\text{crit}} = \left(\frac{l_k^3}{v_F} \right)^2 = \left(\frac{l_k}{d} \right)^2. \tag{2.7}$$

The previous results on the chain size and conformation are summarized in Table 2.1.

2.1.2 Neutral single chain: Dynamic properties

We examine in this section the dynamical properties of a polymeric chain, characterized by the interactions with solvent molecules and between different parts of the chain.

chain length	chain size	chain conformation
$N_k < 1$	$\sim N_k$	rod-like chain
$1 \ll N_k < N_k^{\text{crit}}$	$\sim N_k^{1/2}$	Gaussian chain
$N_k > N_k^{\text{crit}}$	$\sim N_k^{\nu}$	swollen or self-avoiding chain

Table 2.1: Summary of the size and conformation of a linear chain.

When no external force is applied, the chain is subject to constant bombardment from the solvent molecules and, as a consequence, inherits a random motion-type dynamics. Historically, two classical models are used to describe the random or Brownian dynamics of a single chain: the Rouse Model and the Zimm Model.

Rouse model

This model is due to P. E. Rouse who used a bead-spring representation of the chain. The beads, each with a friction coefficient ξ, are connected by linear harmonic springs. When the chain moves through the solvent, each bead experiences a drag force proportional to its velocity (independent of the motion of other beads) and a force from the two neighboring beads. In this case, the chain is said to be free-draining. A description of the dynamics follows from the solution of coupled linear harmonic oscillators [56], whereby the motion of each bead can be decomposed into several internal modes, p, each with a relaxation time τ_p, given by [56]

$$\tau_p = \frac{\tau_1}{p^2}. \tag{2.8}$$

The longest relaxation time τ_1, also known as Rouse relaxation time, is calculated as [56]

$$\tau_1 = \frac{\xi l_k^2}{3\pi^2 k_B T} N_k^2 \tag{2.9}$$

where $k_B T$ is the thermal energy.

In addition, a simple calculation using the internal modes and correlation functions lead to the following result for the diffusion coefficient of the chain [56]

$$D_c = \frac{k_B T}{N_k \xi}, \tag{2.10}$$

indicating that the total friction of the chain is simply $N_k\xi$, the sum of all the friction coefficients over the beads. Combining Eqs. (2.9) and (2.10), the following expression is readily obtained

$$D_c \sim \frac{N_k l_k^2}{\tau_1}, \tag{2.11}$$

where Rouse relaxation time can now be interpreted as the time it takes the chain to diffuse a distance equal to its own size.

The scaling relationships, $\tau_1 \sim N_k^2$ and $D_c \sim N_k^{-1}$, are unfortunately not in agreement with experimental results. Indeed, the following dependencies have been measured experimentally in θ-state solution [60], *i.e.*, when the chain behaves as an ideal chain

$$\tau_1 \sim N_k^{3/2} \quad \text{and} \quad D_c \sim N_k^{-1/2}. \tag{2.12}$$

The discrepancy lies in the very key assumption of the Rouse model. The velocity of each bead was supposed to be independent of the motion of the other beads. In practice, this is not true. When a bead moves through the solvent, the velocity field of the whole fluid is affected, which in turn, drags along other beads. This is called hydrodynamic interactions effect. If this effect is taken into consideration, we obtain the Zimm model.

Zimm model

The bead-spring model was adapted to take into account hydrodynamic effects [61]. This was accomplished by substituting the friction coefficient ξ by a mobility matrix H_{ij} that depends on the position of all the beads and that captures the velocity response of bead i to the motion of all other beads j.

Similar to Rouse model, the linearity of the resulting equation guarantees the use of the internal mode concept from which the new relaxation time τ_p of mode p and the center of mass diffusion coefficient can be calculated as [56]

$$\tau_p = \frac{1}{p^{3/2}} \frac{\eta(\sqrt{N_k} l_k)^3}{\sqrt{3\pi} k_B T} = \frac{\tau_1}{p^{3/2}}, \tag{2.13}$$

and

$$D_c = \frac{8 k_B T}{3\sqrt{6\pi^3}\eta\sqrt{N_k} l_k}, \tag{2.14}$$

where η is the solution viscosity. Note that in this case, the relaxation times decay as $p^{-3/2}$ instead of a quadratic decay. The longest relaxation time and diffusion coefficient

dependencies on the chain length are then

$$\tau_1 \sim N_k^{3/2} \quad \text{and} \quad D_c \sim N_k^{-1/2}, \tag{2.15}$$

in agreement with the experimental relations from Eq. (2.12).

After examining static and dynamic properties of neutral polymers, we address in the following section the modifications that arise in the case of charged polymers or polyelectrolytes.

2.1.3 Polyelectrolytes at rest in a solution

Size and conformation

In general, the static conformation of a uniformly charged, linear polyelectrolyte is determined based on relative values of a number of parameters, such as the polyelectrolyte diameter d, the Debye screening length κ^{-1}, the contour length L, the Kuhn length l_k and the charge density σ [62]. In the limit of an ion-free solution, *i.e.*, zero salt concentration and a "sufficiently" charged polymer, repulsive electrostatic interactions between similarly charged monomers cause the chain to swell relative to a Gaussian coil.

Due to Onsager-Mannig condensation [3] , most polyelectrolytes used in electrophoresis have a "sufficiently" high charge density in the order of $\sigma \sim e/l_B$, where the l_B is the Bjerrum length [4] and e, the elementary charge [9]. However, polyelectrolytes such as DNA only dissolve in solutions with non-zero salt concentration [9], whereby the electrostatic interactions decay exponentially and are screened within a distance of order $\kappa^{-1} \sim 1-3$ nm for most situations encountered in electrophoresis [9]. As a result, most polyelectrolytes in electrophoresis experiments can be treated as their neutral counterparts. This means that their size obeys Eqs. (2.3)-(2.7), in which κ^{-1} must replace the diameter d whenever $\kappa^{-1} > d$ [9].

[3] When counterions dissociate from the backbone of the chain, a fraction tends to condense back onto the chain due to electrostatic interactions, thereby lowering the effective charge of the chain.

[4] The Bjerrum length represents the distance at which the electrostatic energy of interactions between two particles with a charge e is exactly balanced by the thermal energy. In water at room temperature, $l_B \approx 0.7$ nm [9].

2.1.4 Electrophoresis of polyelectrolytes

In this section, we examine the electrophoretic mobility of a polyelectrolyte in the presence of an electric field. We consider a polyelectrolyte placed in a polar solvent (salty water), subject to an external electric force E. Similar to the Brownian dynamics of a neutral chain, both Rouse and Zimm models can be extended to the present case. Let us start first with a freely draining, Rouse-like chain (bead-spring model). The total friction coefficient of the chain is simply the sum of all the friction coefficients ξ of the beads. Likewise, the total electric charge of the chain is $Q = N_k q_k$ where q_k is the charge per bead and N_k the number of beads or Kuhn steps. Balancing the electric force $F_e = QE = N_k q_k E$ and the friction force $F_f = N_k \xi v$, results in the following center of mass velocity [22]

$$v = \frac{N_k q_k E}{N_k \xi} = \frac{q_k}{\xi} E \tag{2.16}$$

or, equivalently, the electrophoretic mobility

$$\mu \equiv \frac{v}{E} = \frac{q_k}{\xi}. \tag{2.17}$$

This expression indicates that the mobility is independent of the length or the molecular weight of the chain.

When hydrodynamic interactions are considered, *i.e.*, Zimm effect, the effective friction coefficient of each bead is changed due to the additional drag from the opposite motion of the counterions in the Debye layer [22]. Also, the total friction coefficient is modified, not only because the individual friction coefficient of each bead is modified, but also because of the hydrodynamic coupling between the motions of the beads. In this case, the effective friction coefficient of the chain in a solvent of viscosity η is given by [22, 63]

$$\xi_{\text{eff}} = 3\pi\eta N_k l_k \left[\ln(\kappa l_k)\right]^{-1} \tag{2.18}$$

which indicates that, similar to the Rouse chain, the friction coefficient scales linearly with N_k, although the prefactor is different. The electrophoretic mobility is consequently

$$\mu = \frac{q_k}{3\pi\eta l_k} \ln(\kappa l_k), \tag{2.19}$$

which is again independent of the length or molecular weight of the polyelectrolyte.

The persistence of this size-independent behavior is due to the screening of hydrodynamic interactions that decay beyond the Debye layer. As a result, even when hydrodynamic coupling is taken into account, polyelectrolytes are often called "free-draining" when subject to an electric force [22]. As a consequence, size-based separations of long polyelectrolytes cannot be realized in free solution. A sieving medium that induces a non-electric steric force is therefore necessary for size-based separations. The study of the dynamics of polyelectrolytes in such media (gels) is the subject of the next section.

2.2 Reptation concept

After the static and dynamic properties of a single polymer chain when placed in a free solution, we examine the situation where the chain is inserted a sieving matrix, such as a melt or a gel.

The basic concept for understanding such behavior is the tube model, introduced by Edwards [64] and used by de Gennes [65] to describe the properties of a chain moving through randomly distributed, fixed obstacles. The main idea is the following: since the trapped chain cannot go through the obstacles, it is constrained in a tube formed by the neighboring obstacles (gels fibers, other entangled chains) and is therefore impeded in its motion by steric interactions with the "walls" of the tube. Note that for the case of entangled solutions or melts, the surrounding chains will also move, and some of the constraints forming the tube progressively disappear. This effect, however, can be neglected since it occurs much more slowly than the escape of the chain [59, 65]. One can treat the chain as if it were moving through a frozen network of obstacles. In the absence of external forces, two mechanisms contribute to the purely diffusive escape of the chain, summarized in Fig. 2.4:

- A snakelike motion, or, as de Gennes called this mechanism, "reptation" from the Latin *reptare*, "to creep". In this case, the chain can crawl along the tube contour in either direction with equal probability, causing the creation or destruction of tube sections at the end of the chain.

- The escape can also occur by the formation of loops, or "hernias" where the chain has to complete multiple crosses of the tube walls. Nevertheless, in the absence of

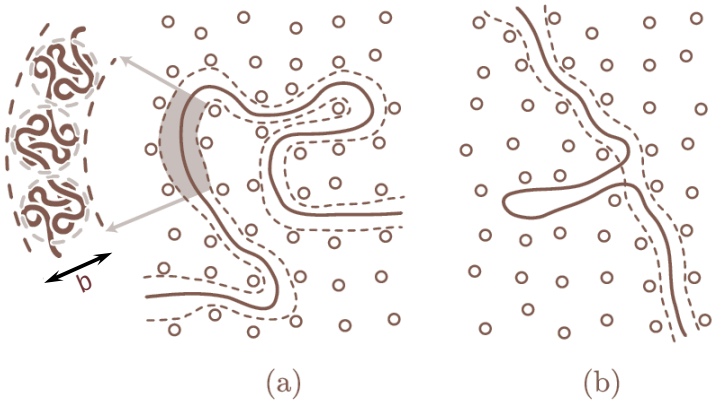

Figure 2.4: Motion of a chain inside the tube through a fixed network of obstacles. (a) Chain reptation along the tube axis. The onset depicts several blobs. (b) Hernia formation. Reproduced from [22].

an external force, such conformation is accompanied by an entropy loss and thus the chain leakage across the walls of the tube is highly improbable.

Reptation is thus the main mechanism of motion of flexible chains in a confining medium. In other words, the motion of the chain can be described by a curvilinear one-dimensional Brownian diffusion along the tube axis through a frozen network of obstacles [59, 65].

Interest lies here in long chains, with a radius of gyration R_g larger than the uniform pore size b. The case where the molecule size is much smaller than the pore size involves a different physical mechanism called entropic trapping and is discussed in the next chapter. The polymer is modeled as a flexible, linear chain that consists of N_k Kuhn segments of length l_k and friction coefficient ξ_k. We make the assumption that the chain is flexible on the scale of one pore and therefore, we define a "blob" as the section of the chain that is contained in one pore. This "local" flexibility assumption is reasonable in the case of single-stranded DNA and flexible synthetic polyelectrolytes in the most widely used gels in gel electrophoresis experiments, polyacrylamide or agarose, or duplex DNA in agarose (the Kuhn length of duplex DNA is about 100 nm and the average pore size in agarose is 200-500 nm).

Following Edwards and de Gennes, the tube can be viewed as a random walk of N_b

connected pores, inside which the chain constantly drifts back and forth. Furthermore, since excluded volume interactions are screened on the scale of one pore, each blob can be considered as Gaussian at all scales larger than l_k. The general situation can then schematically be described as a random walk of Kuhn segments crawling along a random walk of pores as in Fig. 2.4. The number of blobs N_b can be expressed in terms of the chain parameters N_k and l_k and the pore size b as

$$N_b = \frac{N_k\, l_k^2}{b^2}, \tag{2.20}$$

which simply states that the average square end-to-end distance of the polymer is unchanged, irrespective of the random-walk considered, or equivalently, that there are b^2/l_k^2 segments in one blob.

Two parameters are of great relevance in the description of the dynamical properties of the chain, namely, the translational diffusion coefficient D_t and the terminal time τ_t. We first start with the terminal time τ_t. It is the time that the reptating chain takes to completely leave and renew its original tube. Let us estimate its dependence on the length of the chain or, equivalently, on the number of blobs N_b. When the chain moves along the tube, each segment experiences a frictional force that is, according to Stokes law, proportional to the velocity and inversely proportional to the friction coefficient ξ. In addition, the frictional forces acting on each segment are totally independent. Thus, the total frictional force acting on each blob, ξ_b, and the entire chain, ξ_c, are simply the corresponding sum over the appropriate segments

$$\xi_c = N_k \xi_k, \tag{2.21}$$

and

$$\xi_b = \xi_c\ N_b = (b^2/l_k^2)\xi_k. \tag{2.22}$$

Using this expression, the curvilinear diffusion coefficient of the chain is readily given by the Einstein relationship

$$D_c = \frac{k_B T}{\xi_c} = \frac{D_b}{N_b}, \tag{2.23}$$

where $D_b = k_B T/\xi_b$ is the diffusion coefficient of a single blob. The terminal time τ_t, defined as the time it takes the chain to diffuse along the tube by a distance equal to the length of the tube axis, $L = N_b b$, is then estimated by

$$\tau_t \sim \frac{N_b^2 b^2}{D_c}. \tag{2.24}$$

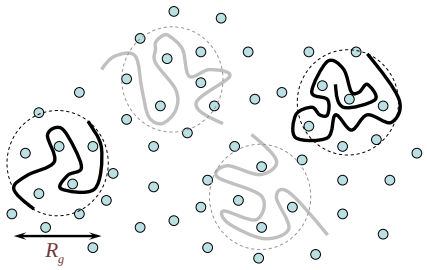

Figure 2.5: The tube and chain conformation are renewed after each terminal time τ_t during which the chain has moved a distance comparable to its radius of gyration. The newly formed tube is totally independent from previous ones.

Inserting the expression of D_c from Eq. (2.23) and using the expressions from Eq. (2.20) and Eq. (2.22), yields

$$\tau_t \sim N_b^3 \frac{\xi_k b^4}{l_k^2 k_B T} = N_b^3 \frac{\xi_b b^2}{k_B T} = N_b^3 \tau_b, \tag{2.25}$$

where we introduced the blob time τ_b, which represents an estimation of the time required for a blob to diffuse a distance equal to its own length b. We see from the expression in Eq. (2.25) that τ_t is proportional to N_b^3. A more accurate evaluation [56] of τ_t that takes into account the fluctuations of the chain inside the tube, *i.e.*, the extensional motion, leads to a slightly higher power $\tau_b \sim N_b^{3.4}$, very close to measurements from elastic experiments [66].

The translational diffusion coefficient can in turn be estimated if we consider that during the time τ_t, the chain has moved in space by a distance on the order of its radius of gyration $R_g \sim l_k N_k^{1/2} \sim b N_b^2$. After each time interval of order τ_t, a new tube is created and no memory of the the original conformation is maintained. This situation is illustrated in Fig. 2.5. Time intervals of length τ_t are then statistically independent, leading to the following estimation of the translational diffusion coefficient

$$D_t \sim \frac{R_g^2}{\tau_t} \sim \frac{D_b}{N_b^2}. \tag{2.26}$$

The diffusion coefficient is very small and decreases rapidly with the length of the polymer chain $D_t \sim N^{-2} \sim N_k^{-2}$, in agreement with experimental data [67]. Note that this scaling law is very different from the case of a polymer coil diffusing in a free solution of viscosity η. In the latter case, if we consider that the coil behaves roughly as

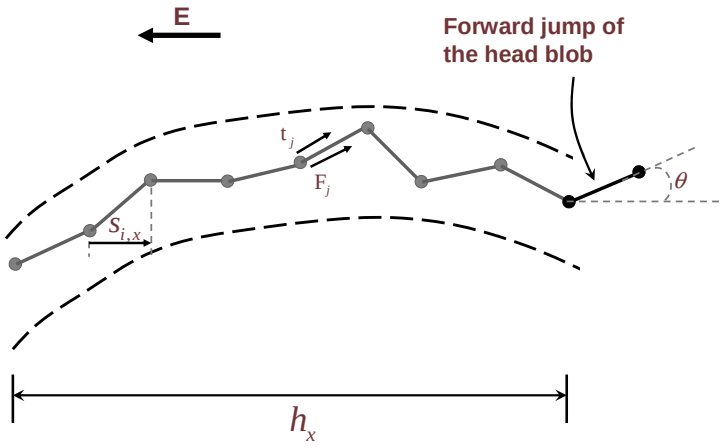

Figure 2.6: Chain reptating inside its tube.

a sphere of radius R_h, where the hydrodynamic radius is of order R_g, then the diffusion coefficient $D = k_B T/6\pi\eta R_h$ scales as $D \sim N_k^{-1/2}$, much bigger than D_t, as expected in the absence of retarding obstacles.

2.3 Reptation in a weak field: Biased Reptation Model

2.3.1 General scaling laws

After presenting the purely diffusive reptation model, we consider next a fundamental extension of this model to situations where the polyelectrolyte is subject to an external electric field. This corresponds to gel electrophoresis used in most applications of DNA and denatured protein separations.

The general scaling laws for the mobility of a polymeric chain as a function of its size and the applied electric field can be formulated using an early approach by Lerman and Frish [68] and Lumpkin and Zimm [69]. Their approach relies on two fundamental assumptions: (i) The electric energy at the scale of a blob is weak enough compared to thermal energy; no hernias are allowed, and (ii) the tube length remains constant, or weakly fluctuating at best.

The first step is to determine the total tangential electric force acting on the chain. If each blob j, with a charge q_b, is oriented in the direction of the unit vector $\mathbf{t}_j$, as in

Fig. 2.6, the electric force $\mathbf{F}_j$, acting on the blob is

$$\mathbf{F}_j = q_b \left(\mathbf{E} \cdot \mathbf{t}_j\right) \mathbf{t}_j. \tag{2.27}$$

The total tangential electric force on the entire chain is therefore

$$F_{el} = \sum_j \frac{q_b}{b} E s_j = \frac{QE}{L} \sum_j s_{j,x} = \frac{QE}{L} h_x, \tag{2.28}$$

where $Q = N_b q_b$ and $L = N_b b$ are respectively, the total charge and contour length of the chain, $s_{j,x}$ is the blob orientation parallel to the electric field and h_x is the end-to-end distance projected onto the direction of the field [see Fig. 2.6]. The balance between the electric and the drag forces gives the equation of motion along the chain

$$\xi_c \dot{s} = QE \frac{h_x}{L}, \tag{2.29}$$

where the friction coefficient of the chain ξ_c is given by Eq. (2.21), s and $\dot{s}$ are respectively the curvilinear coordinate and mean velocity along the chain. From the assumption of reptative motion and the definition of the center of mass, the translational velocity $\dot{X}_{CM}$ of the center of mass of the chain in the x-direction can be related to the curvilinear velocity $\dot{s}$ by [69]

$$\dot{X}_{CM} = \dot{s} \frac{h_x}{L}. \tag{2.30}$$

This equation simply indicates that during the time $t = L/\dot{s}$ for the tube renewal, *i.e.*, for the last blob to leave the tube, the chain is displaced a distance h_x in the x-direction, with a velocity $\dot{X}_{CM} = h_x/t$. Consequently, plugging Eq. (2.30) into Eq. (2.29), averaging over all the tube conformations and using the assumption of weak tube length fluctuations give the electrophoretic mobility

$$\mu \equiv \frac{\dot{X}_{CM}}{E} = \frac{Q}{\xi_c} \left\langle \frac{h_x^2}{L^2} \right\rangle = \mu_0 \left\langle \frac{h_x^2}{L^2} \right\rangle = \mu_0 \frac{\left\langle h_x^2 \right\rangle}{L^2}, \tag{2.31}$$

where $\mu_0 = Q/\xi_c$ is the size-independent mobility in free solution, as calculated in Eq. (2.17).

The next step is to determine the conformational average $\left\langle h_x^2 \right\rangle$ which measures the degree of orientation of the chain as a function of the external electric field. Two limiting cases emerge.

Biased reptation without orientation. When the electric energy of a blob is weaker than the thermal energy, the chain conformation retains the Gaussian properties discussed in the previous chapter. Consequently, the mean-square end-to-end distance $\langle h_x^2 \rangle \sim L \sim N_b$. In this case, the mobility in Eq. (2.31) becomes $\mu \sim 1/L \sim 1/N_b$ [69]. This is the biased reptation without orientation regime. Note that DNA fragments can be separated in this regime as a result of their size-dependent mobility.

Biased reptation with orientation. In contrast, when the chain is subject to a strong electric field, it tends to orient in the direction of the field with an end-to-end distance $\langle h_x^2 \rangle \sim L^2 \sim N_b^2$. The mobility in Eq. (2.31) thus becomes independent of the chain length or molecular weight $\mu \sim L^0 \sim N_b^0$. This is the case of biased reptation with orientation, where the size-independent mobility prevents DNA separation.

The scaling laws presented above for the mobility expression, were considered more carefully in two distinct approaches (LDZ and SN models) that we discuss below.

2.3.2 Quantitative models

Lumpkin, Déjardin and Zimm (LDZ) model

In the Lumpkin-Déjardin-Zimm (LDZ) model [70], the motion of the chain is determined by that of the front blob, the "head" that leads the rest of the chain along the tube. The orientation of the head, biased by the electric field, determines the orientation of all other blobs which, accordingly, inherit the bias from the head. In the weak field limit, the electric energy is smaller that the thermal energy, so that

$$\epsilon \equiv \frac{q_b E b}{k_B T} \ll 1. \tag{2.32}$$

Based on a calculation of the average orientation of the leading blob, the electrophoretic mobility of the chain becomes [70]

$$\mu = \frac{\mu_0}{3}\left[\frac{1}{N_b} + \frac{\epsilon^2}{12}\right]. \tag{2.33}$$

Slater and Noolandi (SN) model

In contrast to the LDZ model, the Slater-Noolandi (SN) model [71–74] takes into account the Brownian motion of the chain inside the tube. Here, both end blobs can drive the

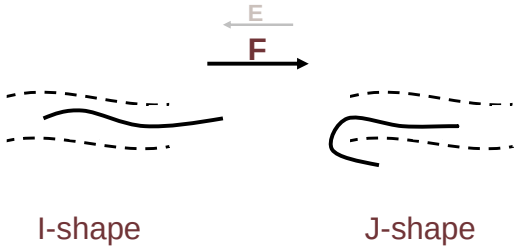

Figure 2.7: I-shape conformation corresponding to forward jumps (LDZ model). When backward jumps are allowed (SN model), the chain can also adopt a J-shape conformation.

chain out of its tube. The "head" blob, corresponding to a "forward" jump, is favored by the electric field while a "rear" blob, or backward jump, will also tend to align in the direction of the field [74]. This additional freedom in the motion of the chain leads to occasional J-shape conformations, as opposed to I-shape conformations in the case of the LDZ model, where the chain can only renew the tube through motion of the "head" blob. These two situations are depicted in Fig. 2.7. The electrophoretic mobility of the chain in this case has the form [73]

$$\mu = \frac{\mu_0}{3}\left(\frac{1}{N_b} + \frac{\epsilon^2}{9}\right). \tag{2.34}$$

Both models presented above predict the existence of two regimes in the reptation of the chain: a non-oriented regime for short chains where the separation is possible and a size-independent oriented regime for long chains where the mobility depends on the applied field. The transition between the regimes occurs when both terms in Eq. (2.33) or Eq. (2.34) are of the same order of magnitude, which translates to

$$N_b^* \sim \epsilon^{-2}. \tag{2.35}$$

For molecules shorter than N_b^*, the mobility is inversely proportional to the length of the molecule, while for longer molecules, the mobility saturates to a size-independent value, scaling like $\mu \sim \epsilon^2 N_b^0$.

The latter scaling was never observed experimentally and constitutes a major shortcoming of the biased reptation model (BRM). It was addressed in a new model, biased reptation with fluctuations (BRF), which we briefly introduce below.

2.3.3 Effect of fluctuations: Biased Reptation with Fluctuations

The predictions of the biased reptation model fail to agree with the experimentally observed scaling with the electric field for long flexible polymers in the saturated regime (in BRM, $\mu \sim N_b^0 \epsilon^2$ for large molecules) [9]. This shortcoming is a direct result of the main assumption in BRM, namely, the chain orientation calculations neglected the fluctuations of the chain length inside the tube. In particular, when the electric field is weak, *i.e.*, $\epsilon \ll 1$, rapid fluctuations in the positions of the end blobs will occur due to their Rouse-like motion [56] and internal modes can no longer be ignored.

A more rigorous approach, biased reptation with fluctuations (BRF), which takes into account the internal modes, was proposed by Semenov, Duke and Viovy [75–77]. The following scaling of the mobility was obtained

$$\frac{\mu}{\mu_0} = \frac{1}{3N_b} \quad \text{if} \quad \epsilon N_b < 1, \tag{2.36}$$

$$\frac{\mu}{\mu_0} = \frac{1}{2}\epsilon \quad \text{if} \quad \epsilon N_b > 1. \tag{2.37}$$

A convenient interpolation formula between small and long chains, valid at low fields, was later proposed by Barkema [78]

$$\mu = \mu_0 \left[\frac{1}{9N_b^2} + \frac{4\epsilon^2}{25} \right]^{1/2}. \tag{2.38}$$

Similar to the BRM case, two reptation regimes appear in the BRF formulation. A regime without orientation for $N_b < N_b^{**}$ where size-based separation is possible and a reptation with orientation regime where the mobility is size-independent but rather exhibits a dependence on the electric field. The cross-over between the two regimes occurs for N_b^{**} such that

$$N_b^{**} \sim \epsilon^{-1}, \tag{2.39}$$

which contrasts with the prediction of the BRM in Eq. (2.35). Finally, the field-dependent mobility scales like $\mu \sim N_b^0 \epsilon^1$ in the case $N_b > N_b^{**}$, a scaling that was experimentally confirmed for large chains [79, 80].

2.4 Nanopore Translocation and First-Passage Processes

The purpose of this section is to introduce the reader to the concept of nanopore translocation. As a promising prospect for single-molecule real-time sequencing [81], the topic of nanopore translocation has received tremendous attention both from experimental and theoretical perspectives [82–93], and is beyond the scope of the present research. However, both the unhooking dynamics of the chain and the molecular motion in an array of spheres share the main feature of a translocating chain, namely, the action of being pulled in a single file. The link between these situations will be made more explicit in Chapters 4 and 6.

In Sec. 2.4.1, we briefly overview some the aspects of nanopore translocation. Our approach is rather elementary and pedagogical. Some of the theoretical results on first-passage time statistics, applied in the context of translocation, will be presented in Sec. 2.4.2.

2.4.1 Nanopore translocation

Nanopore translocation represents one of the most promising avenues for real-time sequencing of a single molecule. The ultimate goal of the procedure is to achieve the complete sequencing of an entire DNA molecule, containing millions of bases, in a few minutes. For comparison, the Human Genome Project, initiated in 2003 by the US Department of Energy Joint Genome Institute, could only sequence 1.5 billion bases per month [8].

The basic premise of nanopore sequencing relies on the possible ability to read the different bases of a given DNA sequence while the molecule is linearly threading through a small opening in a membrane [94,95]. Nanopore translocation thus refers to the process of molecular transfer, during which, fragments of DNA change location from one side of a membrane (the *cis* side) to the other side (the *trans* side) through a narrow opening that only permits a linear threading. In fact, several exchange mechanisms occurring at the cellular level are of the translocation-type. During the viral infection of a cell, for example, viruses inject their genomic DNA into the cell through nanopores across the cell's membrane [96]. Other examples include the translocation of proteins through a sub-cellular membrane [97], or the leakage of a messenger RNA through the cell's

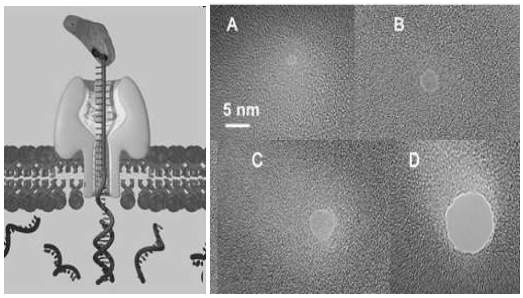

Figure 2.8: Two examples of nanoscopic pores. An α-hemolysin protein channel (left) and a solid-state nanopore (right). Reproduced from [98].

nuclear membrane [96].

The rationale behind nanopore translocation for DNA sequencing dates to the KBBD proposal, coined after John Kasianowicz, Eric Brandin, Daniel Branton and David Deamer, who pioneered the technique in a seminal paper published in 1996 [82]. The basic principle is rather simple. DNA molecules are placed in one of two adjacent chambers containing an ionic solution and separated by a membrane. The membrane has a pore with a diameter that is commensurate with that of the DNA strand so that the molecule can only move in a single file. If an electric voltage is applied between the two chambers, the migration of the ions through the pore can be detected via a measurement of the current generated by the flow of ions. In the absence of DNA molecules inside the pores, the ions will move unhindered between the chambers and the measured current will be maximal. In contrast, as a DNA molecule passes through the pore, it impedes the flow of the ions, thereby lowering the measured current, or conductance. If each nucleotide imparts a unique signature on the blockade, or obstruction of the current (due to their different sizes, chemical compositions, etc.), then the translocation of the molecule will enable a sequential reading of DNA with a single-base resolution. While the paper by Kasianowicz *et al.* showed that the blockade of the ionic current informs about the molecular length of DNA, single-base resolution remains the *Holy Grail* of DNA nanopore sequencing, with indubitable advantages to the current technology, both in terms of sequencing speed and range of molecules to be sequenced [81, 94]. The technique could also lead to a substantial cost reduction since it is reagent free and does not require cloning or amplifying preprocessing.

Before we move on to the next section, a few words about the aperture in the membrane. Figure 2.8 shows the two types of nanoscopic pores used in nanopore translocation experiments. They can either be a natural biological pore, such as an α-hemolysin protein channels [82], or an artificial solid-state one [99, 100]. While natural pores can be chemically designed to change their properties, their fixed size and limited stability (over pH, temperature, etc.) still represent a constraint during manipulation. Solid-state nanopores can be produced instead by irradiating a silicon membrane with a tightly focused high energy electron beam. The pores obtained in this manner are very stable, with higly tunable sizes, both in diameter and length.

2.4.2 First-Passage Processes

A first-passage process is a stochastic process in which events are triggered when a given variable reaches a specific value for the first time [101]. Fluorescence quenching or the triggering of stock options are typical examples of first-passage processes. In the former, light emission by a fluorescent molecule stops when it reacts with a quencher [101], while in the latter, orders of sell/buy are executed when a stock price reaches a threshold for the first time [101]. During translocation, the polymeric chain diffuses back and forth between the *cis*- and *trans*-side of the membrane. When the total mass of the chain is entirely located on either side, the translocation is complete and the chain "exits" the system. The escape of a polymer during translocation thus clearly falls within the definition of a first-passage process. In fact, the translocation of a polymer through a membrane can be viewed as the first-passage of a one-dimensional Brownian particle, trapped between two absorbing walls.

Given the analogy above, some of the results on the first-passage time statistics of a one-dimensional Brownian particle in an interval are recalled below. These are well-established results, and for simplicity, no proofs will be provided in this presentation. Detailed demonstrations can be found in the excellent monograph by Sidney Redner [101].

Field-free results

Consider the case of a one-dimensional Brownian particle placed initially between two absorbing walls. The (–) wall is located at the origin while the (+) wall is a distance

L from the origin. The particle, initially at a distance x_0, $0 \leq x_0, L$, from the origin and is free to diffuse in the interval between the two walls. Once the particle reaches either wall, it is removed from the system. The motion of the particle is governed by a Fokker-Planck equation for the time-dependent probability distribution function of the position of the particle, subject to two boundary conditions at the absorbing walls. In this situation, the total probability that the particle exits the interval at the (+) wall is simply proportional to its distance from the wall, and

$$P_+(x_0) = \frac{x_0}{L}. \tag{2.40}$$

Likewise, the complimentary probability for the exit point to be at the (–) wall is

$$P_-(x_0) = 1 - P_+(x_0) = \frac{L - x_0}{L}. \tag{2.41}$$

One might also inquire about the average time it takes the particle to reach either wall. This is given by the mean first passage time (MFPT) for each exit wall. We have then

$$\langle \tau_+(x_0) \rangle = \frac{2Lx_0 - x_0^2}{6D}, \tag{2.42}$$

for an exit at the (+) wall and

$$\langle \tau_-(x_0) \rangle = \frac{L^2 - x_0^2}{6D}, \tag{2.43}$$

to exit at the (–) wall. Higher moments of the first-passage time are similarly obtained from the solution of the Fokker-Planck equation.

Diffusion with a drift

The results stated above can further be extended to the case where the particle is subject to a force which is constant everywhere in the interval. The force introduces a uniform bias, or drift, quantified by a velocity v. In this case, the total probability for an exit at the (+) wall is modified to

$$P_+(x_0) = \frac{1 - e^{vx_0/D}}{1 - e^{vL/D}}, \tag{2.44}$$

while, $P_-(x_0) = 1 - P_+(x_0)$. The calculation of the modified mean first-passage time follows in a similar way, and

$$\langle \tau_+(x_0) \rangle = \frac{L}{v}\left[\frac{1 + e^{-vL/D}}{1 - e^{vL)/D}}\right] - \frac{x_0}{v}\left[\frac{1 + e^{-vx_0/D}}{1 - e^{-vx_0/D}}\right], \tag{2.45}$$

and

$$\langle \tau_-(x_0) \rangle = \frac{x_0}{v} \left[\frac{1 + e^{v(x_0-L)/D}}{1 - e^{v(x_0-L)/D}} \right] - \frac{2L/v}{1 - e^{v(x_0-L)/D}} \left[\frac{e^{vx_0/D} - 1}{e^{vL/D} - 1} \right], \tag{2.46}$$

for an exit at the (–) wall.

In the particular case of a particle initially located half-way between the two walls, *i.e.*, $x_0 = L/2$, the quantities above reduce to [101, 102]

$$P_\pm = \frac{1}{1 + e^{\mp vL/2D}}, \tag{2.47}$$

and

$$\langle \tau_\pm \rangle = \frac{L}{2v} \tanh \frac{vL}{4D}. \tag{2.48}$$

Note here the peculiar result where the MFPT does not depend on the direction of the bias and is the same for exits at both walls.

Another situation of interest arises when the absorbing wall at the origin is replaced by a reflecting wall. In this case, the governing Fokker-Planck equation remains unchanged, while the boundary conditions are appropriately modified to account for the reflective boundary. If the particle starts its journey very close to the reflecting wall, the mean first-passage time to reach the absorbing wall at L, subject to a drift with a velocity v, is [101]

$$\langle \tau \rangle = \frac{L}{v} \left[1 - \frac{1 - e^{-vL/D}}{vL/D} \right]. \tag{2.49}$$

The relevancy of these results, in particular Eqs. (2.47), (2.48) and (2.49), to the present thesis will become more apparent later in Chapters 4 and 6. Indeed, the unhooking from a stationary post, treated in Chapter 4, is essentially a first-passage process. Likewise, first-passage time statistics will prove crucial in the study of the dynamics inside an array of spheres, as will be made explicit in Chapter 6.

Chapter 3

Nanofilters for DNA Separation; A Non-equilibrium Study

This chapter is based on the publication:
N. Laachi, C. Declet, C. Matson, and K. D. Dorfman
Phys. Rev. Lett., **98**, 098106, (2007).

3.1 Introduction

Nanofilters are a promising new class of microfluidic devices [37] that have proven very successful in rapid separations (a few minutes) of both long and short biomolecules such as DNA and proteins. The device is fabricated by standard etching on a silicon wafer [36] to produce a series of narrow slits alternating with deep wells that can be used to trap the migrating molecules. The separation is then enabled by the size-dependent elution process of the various molecules.

In early applications of nanofilters, low-field conditions were exploited to allow for an entropy-based relaxation of the molecules inside the well. Experiments, supported by computer simulations involving large DNA molecules, with a radius of gyration comparable to the slit diameter, led to separations where longer molecules travel faster than smaller ones [36, 103]. This behavior was attributed to entropic trapping effects. In contrast, experimental and theoretical results on rod-like, short molecules achieved separations where smaller molecules migrate more quickly than their larger counterparts due to an Ogston sieving-like mechanism [35, 38].

Although very useful, the preceding results are only valid in the low-field regime. Indeed, the introduction of a thermodynamic concept such as entropy requires that the device operate in conditions close to equilibrium. In this regime, the molecules have sufficient time to relax and explore all their allowable conformational (in the case of large molecules), or, configurational (in the case of short molecules) space.

More generally, it is a long held tenet that separation processes (chromatography, distillation, adsorption, etc.) are typically performed close to equilibrium [40], where the driving force (electric field, a fluid flow, etc.) is rather a weak perturbation, or a bias, to the randomizing diffusion. An increase of the motive force will obviously lead to an increase in the throughput, but in parallel, will take the system away from equilibrium and reduce the separation resolution. Thus, there appears to be an intrinsic trade-off between the throughput and the resolution which limits any desired increase in the motive force.

Our goal in this chapter is to show that nanofilters are not subject to this trade-off, but can rather operate away from equilibrium when high fields are used. This extension not only increases the range of application but also enhances the performance of the

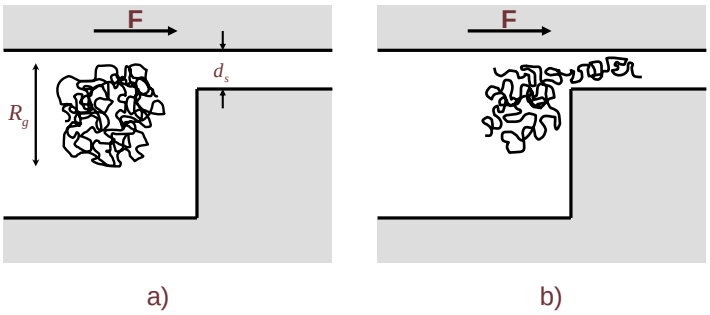

Figure 3.1: a) Relaxation of a large polymeric chain inside the well. b) The chain sequentially escapes the well by stretching its leading monomers into the slit.

device, owing to the increased throughput attained at high fields.

The chapter is organized as the following. We recall in the first section the essential results on the separation of long (entropic trapping) and short (Ogston sieving) DNA molecules in nanofilters under weak fields. We proceed afterwards with the extension of nanofilters to out-of equilibrium conditions in the case of short molecules using scaling arguments in Sec. 3.3.1 and computer simulations in Sec. 3.3.2. The final section summarizes the results. The general approach and results for the non-equilibrium regime are mostly based on a previously published paper [104].

3.2 "Equilibrium transport" in nanofilters

3.2.1 Long molecules: Entropic trapping

The motion of long DNA molecules (several kilo base pairs) in the nanofilter is largely dominated, under weak fields, by entropic-trapping effects. As depicted in Fig. 3.1, DNA molecules with a radius of gyration $R_g \sim 1\,\mu$m are hindered from entering the slit of diameter $d_s \sim 100$ nm. Such motion will be accompanied by an internal conformational entropy loss; the slit entrance thus constitutes an entropic barrier. As a consequence, at low fields, long DNA molecules are trapped in the well with a trapping lifetime τ_{trap} before they can escape. In the low field regime, a simple kinetic model [36, 39], corroborated by Monte-Carlo simulations [103] suggest that the mean trapping time

should scale like

$$\tau_{trap} \sim \frac{1}{\sqrt{M}} \exp\left(\frac{\alpha}{E_s k_B T}\right), \tag{3.1}$$

where M is the molecular weight of the DNA (proportional to the chain length), E_s is the field magnitude inside the slit and α, a fitting constant.

To overcome the entropic barrier, long DNA molecules must stretch their leading monomers into the constriction. Once the leading "squad" is formed, the rest of the polymer chain becomes stretched and enters the narrow region sequentially as can be seen in Fig. 3.1. This process is easily triggered for longer molecules since they have a higher probability that some part of the chain is facing the slit entrance. As a result, the trapping time τ_{trap} decreases with the size of the chain and longer DNA polymers, counterintuitively, migrate faster.

3.2.2 Short molecules: Ogston sieving

When the molecules to be separated are small enough so that their radius of gyration is comparable to the constriction size, the previous conformational entropy argument is no longer valid. This is true because DNA molecules of size ~ 100 nm are not flexible enough to deform and coil inside the deep region; they are rigid, rod-like molecules instead. Short DNA migration through the nanofilter is then an electric field-driven partitioning process, where the entropy, of configurational nature, is governed by the rod allowable configurational space inside the well and the slit as in Fig. 3.2.

The difference between the high entropy in the deep region and the reduced entropy in the narrow region creates a configurational entropic barrier at the interface. Similar to Ogston sieving [105, 106], the entropy barrier is characterized by the ratio of the partition coefficient in the well, K_w and in the slit K_s, as $-T\Delta S \sim -k_B T \ln(K)$, $K \equiv K_s/K_w$. For rod-like molecules, the partition coefficients K_w and K_s can be calculated analytically [107].

The loss of entropy that occurs upon entering the slit is compensated by the drop in the electrical potential energy during the translation of DNA along the field direction. The total free energy of the molecule is then at least of order $k_B T$, which permits the molecule to overcome the entropic barrier with a characteristic mean time τ_{trap}. An

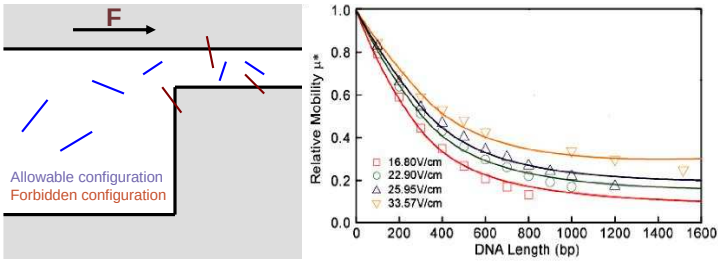

Figure 3.2: a) Partitioning of the rod-like molecules inside the well and the slit. b) Plot of the relative mobility as a function of the molecule size for different values of the electric field. Solid lines are obtained using Eq. (3.2) and the symbols are experimental data. Adapted from [35].

estimation of the trapping time using Kramers theory [108, 109] reveals that [35]

$$\tau_{trap} \sim \frac{\alpha M}{E_{av}^2 K} \exp(-\varepsilon), \tag{3.2}$$

where E_{av} is the average field, ε a reduced electric field and α another fitting constant. The trapping time is linearly proportional to the size of the molecule, which indicates that, in contrast to the previous case, shorter molecules migrate faster. Finally, as can be seen from Fig. 3.2, the theoretical calculation [35] leads to a relative mobility that agrees well with the experimental data, especially in the case of low fields (near-equilibrium conditions) and small DNA molecules where the rod-like assumption is enhanced.

3.3 Non-equilibrium migration in nanofilters

3.3.1 Scaling arguments

We focus in the following on recent applications of the nanofilters for separating short DNA molecules with a contour length L on the order of the slit size $d_s \sim 100$nm [35, 38]. This limitation ensures that DNA of interest here can be treated as rigid rod-like particles, hindered in their motion by interactions with the nanofilter walls. In the case of an unhindered (free-solution) migration, the electrophoretic mobility μ_0 is at best a weak function of the molecular weight for the range of DNA considered [110,111].

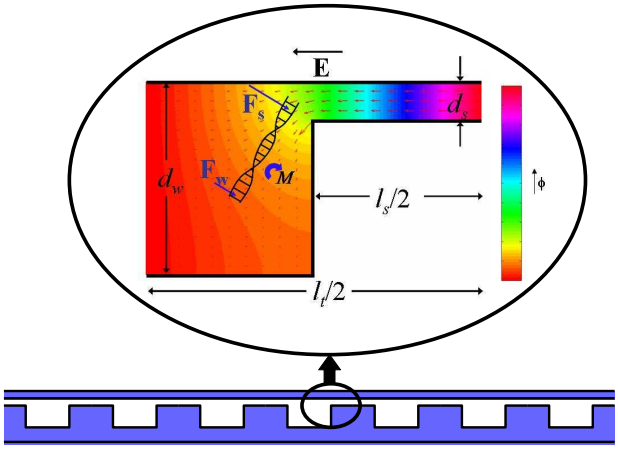

Figure 3.3: Schematic of a periodically constricted nanofilter for separating short DNA. The colors represent the values of the electric potential and several field vectors are indicated near the entrance of the slit. A representative short DNA (not to scale)is also depicted. The molecule is subject to a torque near the slit entrance that tends to orient it towards the slit. The torque is a result of the non-uniformity of the electric field in the nanofilter. Reproduced from [104].

For the purpose of the scaling arguments below, it is considered to be constant. The notations and symbols used below are summarized in Fig. 3.3.

The DNA migrate through the nanofilter under the action of an applied electric field E_{av}. The specific geometry of the device introduces a non-uniformity in the properties of the local electric field. Nevertheless, to leading order, the values of the electric field in the slit, E_s, and the well, E_w can be estimated from a simple model of resistance in series. Thus, for a symmetric nanofilter, the two field strengths obey the relationship

$$\frac{E_w}{E_s} = \frac{d_s}{d_w} \equiv \delta, \tag{3.3}$$

together with the definition of E_{av}

$$E_{av} = \frac{E_w + E_s}{2}, \tag{3.4}$$

yield

$$E_w = \frac{2\delta}{1+\delta} E_{av} \quad \text{and} \quad E_s = \frac{2}{1+\delta} E_{av}. \tag{3.5}$$

As a result, DNA that move entirely unhindered through the nanofilter would migrate with an average velocity $\bar{v} = \mu_{max} E_{av}$ [35], where

$$\mu_{max} = \frac{4\delta}{(1+\delta)^2}\mu_0, \tag{3.6}$$

is the maximum possible mobility in the nanofilter [112]. We also define the translational Péclet number,

$$\mathrm{Pe}_t = \frac{\mu_0 E_w d_w^2}{D l_w}, \tag{3.7}$$

ratio of the time for translational diffusion over the depth of the well

$$\tau_t \sim \frac{d_w^2}{D}, \tag{3.8}$$

and the time for convection across the well

$$\tau_c \sim \frac{l_w}{\mu_0 E_w}. \tag{3.9}$$

In the equilibrium limit previously described in Sec. 3.2.2, the electric field-induced convection is a small perturbation to diffusion, which translates to $\mathrm{Pe}_t \ll 1$. In this limit, as mentioned earlier, larger molecules tend to partition into the well, leading to an Ogston sieving-type mechanism where the smaller DNA exit first [35, 38]. As Pe_t is increased, the electrical potential energy drop is commensurate with the entropic penalty, in which case it has been shown experimentally that the separation resolution decreases, with a compressed band regime around 100 V/cm [38]. Strong electric fields which would enhance the separation speed, seem to result in a catastrophic loss of resolution. Our goal here is to test the validity of this statement.

The apparent loss of resolution when $E_{av} \sim 100$ V/cm does not result from the absence of trapping. This would be true for point-like particles that do not have sufficient time to diffuse down the well, but rather are freely convected by the electric force from one slit to another. However, this criterion is not sufficient for a rod-like particle and trapping should not vanish for all $\mathrm{Pe}_t \gg 1$ since the latter must also arrive at the slit with a favorable orientation in order to avoid trapping. This situation can be translated in terms of a condition on the maximum allowable angle between the rod and the net electric field. At the slit entrance, if the center of mass of the molecule is located at a

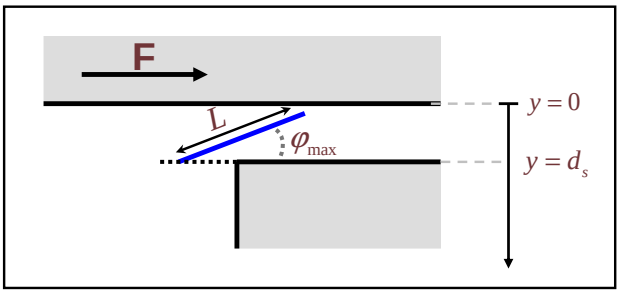

Figure 3.4: Calculation of the maximum allowable angle for a rod of length L with a center of mass located here at $y \geq d_s/2$. The same calculation applies for $y \leq d_s/2$, leading to Eq. (3.10).

distance y from the top of the nanofilter, then the maximum allowable angle φ_{max} is estimated from a simple geometric calculation by [see Fig. 3.4]

$$\sin(\varphi_{max}) = (d_s + |2y - d_s|)/L \quad \text{for} \quad 0 \leq y \leq d_s. \tag{3.10}$$

From this equation, the y-dependent function φ_{max} clearly has a maximum at the slit center where $y = d_s/2$. Accordingly, the condition

$$\sin \varphi_{max}^{(0)} = d_s/L, \tag{3.11}$$

implies that the highest probability for entering the slit will occur if the center of mass of the chain is at the center of the slit entrance. Once the chain leaves a slit, it is convected through the well with a characteristic time τ_c given in Eq. (3.9). The angle of rotational diffusion is then characterized by

$$\varphi_{rot}^2 \sim D_r \tau_c = \frac{D_r \, l_w}{\mu_0 \, E_w}, \tag{3.12}$$

in which D_r is the rotational diffusion coefficient of a chain of length L, approximated from one of a slender ellipsoid of length L and diameter b. In a solution of viscosity η [113]

$$D_r = k_B T \frac{\ln(2L/b) - 1/2}{\pi \eta \, L^3}. \tag{3.13}$$

The condition for a chain to be trapped inside the well at the slit entrance can thus be expressed as

$$\varphi_{max}^{(0)} \leq \varphi_{rot}, \tag{3.14}$$

which, together with Eq. (3.5), imply that there exits a critical average electric field

$$E^*_{av} = \frac{1+\delta}{2\delta}\frac{D_r l_w}{\mu_0[\sin^{-1}(d_s/L)]^2}, \tag{3.15}$$

above which the DNA will tend to possess the correct orientation to engage the slit, unaffected by the slow rotational diffusion. Notice that a more elaborate evaluation of the trapping condition will provide values higher than the one from Eq. (3.15) since the center of mass of the chain is not necessarily located at $y = d_s/2$; the chain will have a distribution of y-values and may be located very close to the bottom of the slit.

To estimate E^*_{av} from Eq. (3.15), we use the following experimental values: a double-stranded DNA 300 bp long, which, according to the Kratky-Porod model corresponds to a rod-like molecule of contour length $L \sim 100$ nm [114]. In water at room temperature, using the DNA backbone width of 2 nm, the rotational diffusion coefficient $D_r \sim 10^4$ s^{-1} [113]. Also, considering a well length $l_w \sim 1\mu$m, a depth ratio $\delta \sim 0.1$, an electrophoretic mobility $\mu_0 \sim 10^{-4}$ cm^2/Vs [110,111], together with the approximation $\sin^{-1}(d_s/L) \sim 1$, we obtain a critical average field $E^*_{av} \sim 50,000$ V/cm, that is much larger than any practical value used in separation experiments.

It appears thus, that even for high values of the electric field (within the range of realistic values) the DNA molecules are not convected, without restraint, from slit to slit. Rather, DNA will suffer some trapping in the well due to rotational diffusion.

In order to escape and enter the slit, DNA must lose the forbidden conformation and "undo" the rotational diffusion by reorienting themselves. The reorientation of the DNA is governed by two physical transport processes:

- Rotational diffusion, which not only results in the trapping, but also maintains random variations of the orientation of DNA molecule.
- Electric-field induced torque (or electro-rotation), which originates from the non-uniformity of the electric field.

In order to compare the relative importance of rotational diffusion and electro-rotation, we introduce the rotational Péclet number, ratio of the characteristic times of the two mechanisms. The characteristic time for rotation of a rod around its center of mass by diffusion is given by

$$\tau_d \sim D_r^{-1}, \tag{3.16}$$

where D_r is the rotational diffusion coefficient from Eq. (3.13). The equivalent time for reorientation under a torque or moment $\mathcal{M}$ is

$$\tau_e \sim \xi_r/\mathcal{M}, \tag{3.17}$$

where ξ_r is the frictional rotational coefficient. Following the definition of the rotational Péclet number along with the Einstein relationship $\xi_r D_r = k_B T$, we obtain

$$\mathrm{Pe}_r \equiv \tau_d/\tau_e = \mathcal{M}/k_B T. \tag{3.18}$$

When $\mathrm{Pe_r} \ll 1$, the reorientation is governed by thermal diffusion, and conversely, when $\mathrm{Pe}_r \gg 1$, the reorientation is governed by the torque.

The characteristic torque $\mathcal{M}$ is estimated by noting that the electric field gradient acting on a DNA near the slit entry is proportional to the difference in the electric field in the slit, E_s, and the well, E_w; $\nabla E \sim (E_s - E_w)/L$. The torque on a molecule of length L is thereby estimated as

$$\mathcal{M} \sim \hat{q}(E_s - E_w)L^2, \tag{3.19}$$

where $\hat{q}$ is the charge per unit length. Combining the latter along with the resistances in series model in Eq. (3.5) to approximate $E_s - E_w$, we arrive at the rotational Péclet number

$$\mathrm{Pe}_r = \left(\frac{1-\delta}{1+\delta}\right)\left(\frac{\hat{q}E_{av}L^2}{k_B T}\right). \tag{3.20}$$

For low values of Pe_r, rotational diffusion dominates the torque in the escape process. In this limit, inasmuch as both Ogston sieving and rotational diffusion favor short molecules ($D_r \sim [\log(L)/L^3$ from Eq. (3.13)], we would expect the shorter DNA to elute first due to either mechanism.

As the rotational Péclet number passes through values of (1), we would expect a transition from rotational diffusion to "torque-assisted" escape. Moreover, electro-rotation mechanism favors longer molecules as the torque $\mathcal{M}$ scales like L^2. As a consequence, in a fully developed torque-assisted regime, a band-inversion is predicted, *i.e.*, longer molecules exit first at strong fields. Note that the existence of this regime does not contradict any of the results previously established in [35, 38], since they solely involved weak to moderate values of E_{av} (up to 100 V/cm).

3.3.2 Simulation description

The results predicted using the scaling arguments above, namely, separation and band-inversion beyond 100 V/cm, are now confronted with simulation data. The simulation features and model are described in the following.

Geometry: The nanofilter is modeled as a succession of m 2D unit cells of periodicity $l_t = 1\mu$m. The respective lengths of the well and the slit are l_w and l_s. The unit cell is considered to be symmetric, *i.e.*, $l_w = l_s = l_t/2$. The well and the slit depths are $d_w = 300$ nm and $d_s = 60$ nm, thus a depth ratio of $\delta = 1/5$. This particular geometry is similar to nanofilters experimentally considered in [35,38]. However, in the simulation, all variables and parameters are made dimensionless by a suitable choice of characteristic values made explicit below.

Electric field: The DNA are animated by an applied voltage V that leads to an effective electric field of magnitude $E_{av} = V/ml_t$. For each of the simulations performed, the value of E_{av} is fixed. However, the specific geometry of the nanofilter and the insulating walls of the device cause the electric field lines to curve, as indicated in Fig. 3.3, thus introducing a non uniform local electric field $\mathbf{E} \equiv \mathbf{E}(x, y) = -\nabla\phi$ where $\phi \equiv \phi(x, y)$ is the position-dependent electric potential. Moreover, both the electric field and potential inherit the periodicity property imposed by the geometry of the nanofilter. This property of $\mathbf{E}$ and ϕ restricts the computation to a unit cell, the rest being inferred by periodicity.

In the absence of a charge source or sink in the nanofilter, the electric potential ϕ is governed by the Laplace equation

$$\nabla^2\phi = 0, \tag{3.21}$$

together with the no-flux boundary condition on the insulating walls with a normal unit vector $\mathbf{n}$

$$\frac{\partial\phi}{\partial\mathbf{n}} = 0 \tag{3.22}$$

and the periodic voltage drop

$$\phi(0, y) - \phi(l_t, y) = -E_{av}l_t. \tag{3.23}$$

The governing equation Eq. (3.21) is numerically integrated using the boundary conditions Eqs. (3.22) and (3.23) to compute the potential distribution inside the unit

cell. The local electric field value is then obtained upon numerical differentiation from $\mathbf{E}(x, y) = -\nabla\phi$.

DNA: The short DNA molecule is modeled as a rigid dumbbell that consists of two charged beads connected by a non-deformable rod so that no internal degrees of freedom are allowed. To fully characterize the dumbbell, we need to prescribe the values of the length of the rod and the size and charge of the bead. The rod length is simply considered to be equal to the size L of the DNA molecule. Values of the bead radius R_{bead} and charge q_{bead} are less straightforward. They are rather obtained from preliminary simulations where the dumbbell is placed in an unbounded domain (free solution).

We start first with the bead radius where the electric field is turned off so that the dumbbell moves under the sole action of molecular diffusion. The basic idea is to choose a bead radius so that the translational diffusion coefficient D_{simu} calculated from the mean square displacement of the center of mass of the dumbbell by

$$D_{simu} = \lim_{t\to\infty} \frac{\langle r^2 \rangle}{4t}, \tag{3.24}$$

matches the scaling for a slender ellipsoid given by [113]

$$D_{ellipsoid} = k_B T \frac{\ln(2L/b)}{3\pi\eta L}. \tag{3.25}$$

Explicitly, the friction that results from the diffusive motion of the dumbbell, depends on the friction coefficient of each bead, which itself is a function of the bead radius by virtue of Stokes law. Ultimately, the diffusion coefficient D_{simu} is a function of the bead radius. Note that the previous argument to evaluate the bead radius using the translational diffusion coefficient of a slender ellipsoid, overestimates the rotational diffusion coefficient of the dumbbell relative to the same ellipsoid.

Following the evaluation of the bead radius, the electric field is now turned on to determine the charge carried by each bead. Likewise, the bead charge is obtained by trial-and-error runs until the calculated free solution mobility is equal to the nominal value $\mu_0 = 2 \times 10^{-4}$ cm^2 /Vs [110, 111]. This condition ensures that, in a free solution, all DNA would move at the same speed. The separation occurs only when a sieving medium, in this case, the nanofilter, is used.

Forces: We discuss in the following the forces acting on each bead. A characteristic time scale of the motion of DNA in the unit cell is of order $10^{-5}s$ [115], much shorter than laboratory or experiment time scales of interest ($\sim$ 1 minute). Inertial effects (acceleration of the beads) can then be neglected. Also, hydrodynamic interactions between the beads are excluded, which is a reasonable assumption in 2D [116, 117]. Hydrodynamic interactions with the walls are also neglected. Under these assumptions, the balance between the forces acting on each bead is described by the Langevin equation

$$\mathbf{F}_i^h + \mathbf{F}_i^b + \mathbf{F}_i^e + \mathbf{F}_i^c = 0, \quad i = 1, 2 \tag{3.26}$$

a stochastic differential equation that governs the motion of bead i.

The stochastic nature of Eq. (3.26) is exclusively due to the random Brownian force $\mathbf{F}_i^b$. During the simulation, it is generated by

$$\mathbf{F}_i^b = \sqrt{\frac{4k_BT\xi}{\Delta t}}\mathbf{n}_i, \tag{3.27}$$

where ξ is the friction coefficient of the bead, calculated using the bead radius. In this expression, Δt is the simulation time step, $\mathbf{n}_i$, a unit vector of which the components are uniformly distributed between -1 and 1. Note that this distribution of the vector $\mathbf{n}_i$ leads to the classical statistics of the Brownian force

$$\left\langle \mathbf{F}_i^{b,\alpha} \right\rangle = 0, \tag{3.28}$$

and

$$\left\langle \mathbf{F}_i^{b,\alpha}(t) \cdot \mathbf{F}_j^{b,\beta}(t + \Delta t) \right\rangle = 2\delta_{ij}\delta_{\alpha\beta}\frac{k_BT\xi}{\Delta t}, \tag{3.29}$$

for each bead $i, j = 1, 2$ and each component $\alpha, \beta = x, y$, while δ_{ij} is the Kronecker delta symbol.

Let's proceed now with the hydrodynamic force. Explicitly, each of the beads, while moving through the stationary solvent, experiences a viscous drag that opposes this motion. Moreover, when the velocities of the beads are not too high, the viscous force acting on each bead is fairly assumed to be linearly proportional to the velocity of the corresponding bead. We can thus write for each bead $i = 1, 2$

$$\mathbf{F}_i^h = -\xi\mathbf{v}_i = -6\pi\eta R_{bead}\dot{\mathbf{r}}_i, \tag{3.30}$$

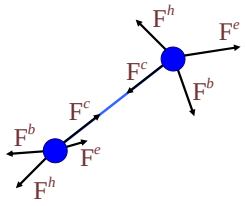

Figure 3.5: Schematic of the forces exerted on each bead.

in which $\mathbf{r}_i \equiv (x_i, y_i)$ is the position vector of the center of mass of bead i and $\dot{\mathbf{r}}_i = \mathbf{v}_i$ is the bead velocity. Note that Eqs. (3.27) and (3.30) both represent the action of the solvent. The collisions with the solvent molecules lead to a diffusive motion of the bead, described by the fluctuating Brownian force. However, the bead motion is slowed down by these same collisions because the bead has to find its way among the surrounding solvent molecules. The response is linear and is described by the viscous or friction force $\mathbf{F}_i^h$. This is the fluctuation-dissipation theorem.

Using the local electric field and the bead charge previously calculated, the electric force acting on each bead is simply

$$\mathbf{F}_i^e = q_{bead}\mathbf{E}(x_i, y_i). \tag{3.31}$$

In addition to the action of the solvent and the electric field, the inextensible tether connecting the beads exerts a constraining force, acting along the rod [115] as depicted in Fig. 3.5, so that

$$\mathbf{F}_1^c = \mathrm{T}\,(\mathbf{r}_2 - \mathbf{r}_1) \quad \text{and} \quad \mathbf{F}_2^c = -\mathbf{F}_1^c. \tag{3.32}$$

The parameter T is simply a tension that maintains the distance between the beads constant, equal to the molecule length L, i.e $\parallel \mathbf{r}_2 - \mathbf{r}_1 \parallel = L$. Subtracting the equations obtained by taking the dot product of $(\mathbf{r}_2 - \mathbf{r}_1)$ and each of the governing equations from Eq. (3.26) yields

$$\mathrm{T} = \frac{1}{2L^2}[(\mathbf{F}_2^b + \mathbf{F}_2^e) - (\mathbf{F}_1^b + \mathbf{F}_1^e)] \cdot [\mathbf{r}_2 - \mathbf{r}_1], \tag{3.33}$$

where the hydrodynamic force terms cancel out since the constant rod length condition implies

$$\xi(\mathbf{r}_2 - \mathbf{r}_1) \cdot (\dot{\mathbf{r}}_2 - \dot{\mathbf{r}}_1) = \frac{\xi}{2}\frac{d \parallel \mathbf{r}_2 - \mathbf{r}_1 \parallel^2}{dt} = 0. \tag{3.34}$$

The boundary conditions are specified by the device geometry so that the beads do not cross the walls of the channel. In the simulation, they are implemented by rejecting position updates that result in the beads crossing the domain boundaries. Also, at each time step, the time counter is incremented regardless of the status of the update (rejected or accepted), *i.e.*, a Monte-Carlo like move [118].

Upon insertion of the expressions from Eqs. (3.27), (3.30), (3.31) and (3.33) into Eq. (3.26), the latter is made dimensionless by introducing the Kuhn length $l_k = 100$ nm, the relaxation time $t_r = \xi l_k^2/k_BT$ and k_BT/l_k as characteristic length, time and force scale respectively. The resulting dimensionless differential equation is then integrated numerically via a mid-point algorithm to obtain updated values of the position of the beads. The integration mid-point algorithm is detailed in [119, 120].

The simulation was performed for various values of the electric field and three different DNA sizes (300 bp, 500 bp, 766 bp) over $m = 1000$ cells to allow for a steady value of the total migration time to be reached. The mobility is then calculated from the applied field, the total migration time and the total distance traveled in the direction of the field. The results are presented below.

3.4 Results and discussion

Simulations results are summarized in Fig. 3.6, which depicts the relative mobility μ/μ_{max} as a function of a reduced electric field for different DNA lengths. The hydrodynamically reduced field ϵ_h is defined such that

$$\epsilon_h \equiv \frac{\eta\mu_0 E_{av} l_w^2}{2k_BT}. \tag{3.35}$$

The figure indicates that the mobility increases monotonically with the electric field, irrespective of the size of the molecule, until it approaches the limiting value μ_{max} imposed by the nanofilter geometry at values of the field such that $\epsilon_h \gg 40$. Inserting numerical values in Eq. (3.35), *i.e.*, assuming a viscosity close to that of water $\eta = 1$ cp, $l_w = 0.5\,\mu$m, the mobility saturates for $E_{av} \gg 700$ V/cm which is in good agreement with the critical value in Eq. (3.15) at which all molecules would move from slit to slit without much rotational diffusion inside the well.

Also, as can be seen in the figure, the simulations produce a band compression regime

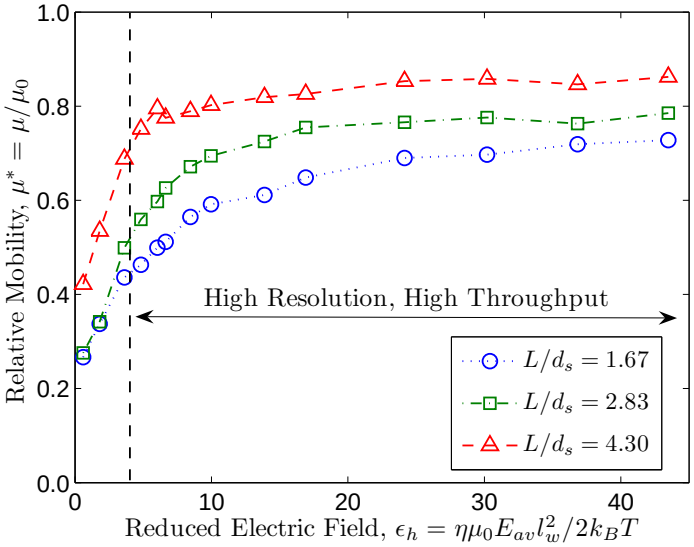

Figure 3.6: Plot of the relative mobility as a function of a reduced electric field for three molecules of respective lengths 300, 500 and 766 bp. The vertical line corresponds to the point where $\mathrm{Pe}_r > 0.1$ for all three molecules.

between the two shortest DNA at low values of the electric field $\epsilon_h < 5$. Using Eq. (3.35), this corresponds to an electric field $E_{av} \sim 100$V/cm, which is in good agreement with the loss of resolution value observed experimentally in [35]. Note that the largest DNA is predicted to move faster than the compressed band. This can be understood if we calculate its corresponding rotational Péclet number. Indeed, for the largest molecule, $\mathrm{Pe}_r = 0.6$ is not so small, so that the electric field-induced torque is in the same order of magnitude of rotational diffusion. For the two shortest molecules, the torque largely dominates the diffusion. In general, the transition into torque-assisted escape depends on the molecular weight [see Eq. (3.20)]. As a result, the model presented above does not predict a single region where all of the DNA would move at the same speed.

The simulation results also suggest that a band inversion regime follows the band compression one, as expected from the scaling arguments above. The transition occurs at strong field $E_{av} \sim 100$ V/cm where all of the DNA have passed through $\mathrm{Pe}_r \sim 0.1$, indicating that the torque-assisted escape is the mechanism responsible for such behavior. In other words, the DNA molecules are driven to the next slit while rotating (via

rotational diffusion) and then rotate back to an acceptable conformation (via torque-assisted escape) while near the slit entrance without much center-of-mass motion. Since the torque favors longer molecules, they escape faster, causing the band inversion.

The scaling arguments and simulation results presented above show that nanofilters need not necessarily operate near equilibrium in order to separate short rigid DNA molecules. The loss of separation observed in the experiments in [38] is probably not the upper limit of operation of the device. Rather, the model presented above predicts that strong field conditions will still lead to a separation with a band inversion; longer molecules elute first. This prediction contrasts with the prevailing idea that equilibrium conditions are essential to chromatographic separation processes [40].

Recently, our predictions were experimentally tested by increasing the electric field beyond the loss of separation limit [121,122]. In fact, recent experimental evidence confirmed our theoretical predictions on the extended range of application of the nanofilter and the concomitant performance increase [122]. The transition to the torque-assisted escape regime was observed experimentally to occur around values of the rotational Péclet number $\mathrm{Pe}_r \sim 1$, consistent with our results. Our findings thus proved useful for enhancing the separation speed and the performance of this promising class of DNA separation devices.

Chapter 4

DNA unhooking from a single post

This chapter is based on the publication:
N. Laachi, J. Cho, and K. D. Dorfman
Phys. Rev. E, **79**, 031928, (2009).

4.1 Introduction

The emergence of microfabrication has opened a number of promising avenues towards addressing the limitations of gel electrophoresis for DNA separations [9, 21, 123, 124]. Using standard micro-lithography etching techniques, Volkmuth and Austin engraved an array of cylindrical obstacles and utilized it to separate DNA molecules in the tens to hundreds of kilobase pair (kbp) range [24]. In subsequent experiments, Baba and coworkers successfully separated λ-DNA (48.5 kbp) and T4-DNA (168.9 kbp) within 10 seconds in a quartz chip containing micro-pillars of diameter and spacing 500 nm without using a pulsed field technique [23]. Similarly, Viovy and coworkers reported rapid separations of various DNA sizes in magnetic matrices with bead columns of size 1-5 μm and spacing 3-10 μm [25, 26].

The failure of the reptation concept to explain the dynamics of DNA chains in sparse arrays of posts launched several theoretical and numerical studies aimed at characterizing such motion [23, 25–27, 125–132]. As a result, it was established that the motion of the DNA in post arrays can be treated as a repetitive cycle of three elementary steps shown in Fig. 4.1: i) the chain collides with a post followed by an extension into two arms, ii) the chain disengages from the post in a way similar to rope over a pulley, and then iii) the chain translates a number of rows before it collides with the next post [27, 133].

The simple mechanical description of the post disengagement, or unhooking, by a rope-over-pulley motion, first proposed by Volkmuth *et al.* [126], has been used extensively [27, 127, 132, 134, 135]. In this approach, a deterministic force balance leads to a logarithmic size dependence of the unhooking time of the chain [27, 126, 132]. In the case of large obstacles, Saville *et al.* [128–130] showed that the unhooking time is independent of the chain size, but rather depends on the so-called impact parameter, while Randall *et al.* [131, 132] showed in meticulous and extensive experiments that the collisions depend on the Deborah number characterizing the extent of stretching due to the electric field gradient near the post. More recently, Holleran and Larson [136] employed a bead-spring model to simulate the interaction between the polymer chain and a thin post over a wide range of chain lengths and field magnitudes. They established the existence of four distinct regimes, dominated by either convection or diffusion,

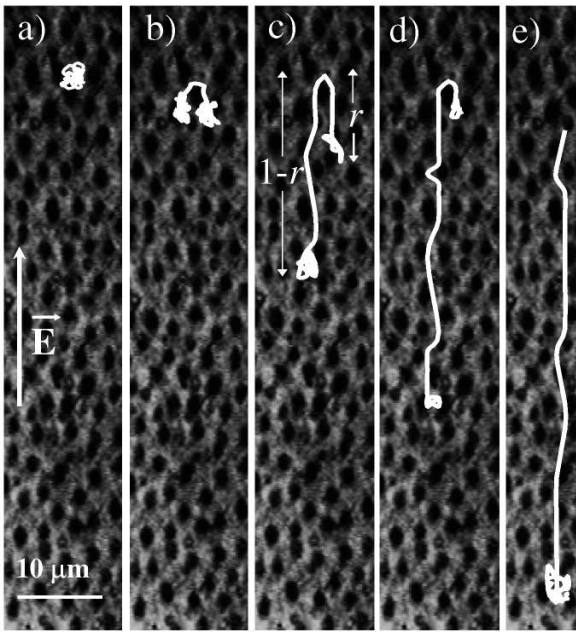

Figure 4.1: Schematic of the cyclic motion of DNA in a post array [26]: (a) After collision with a post, DNA begins step (i). (b) DNA unravels and extends into two arms. (c) DNA begins step (ii) after extension. (d) The DNA disengages by a rope-over-pulley mechanism.(e) Step (iii) begins after the unhooking from the post and ends with the next collision. Reproduced from [27].

that describe the entanglement governing the delayed transport of the chains. Inspired by the geometration model of DNA migration through stationary networks of obstacles [133, 137–139], Minc *et al.* [27] expanded upon the simple rope-over-pulley picture to build a somewhat more complex microscopic model of DNA motion in arrays of posts. Their application of Continuous-Time Random Walk (CTRW) theory to the model led to analytical expressions for the mean velocity through the array and the dispersion characterizing the broadening of the bands [27]. The model was subsequently modified to take into account moderate magnitudes of the electric field and their effect on chain extension during the collision [140–142].

While this CTRW model captures the qualitative aspects of the transport properties of DNA through the post array, developing a quantitatively accurate model of the effective diffusion coefficient (dispersion) of long DNA requires a more fundamental

understanding of the elementary problem of collision and unhooking from a single post. Importantly, the CTRW models [27, 140, 142] assume that the dispersion arises solely from the random offset between the two arms and the random distance between collisions. The deterministic rope-over-pulley model used to quantify the unhooking time possesses a singularity (divergence of the total hold-up time) when the two arms are of equal length after unraveling around the post. If one includes diffusive fluctuations in the unhooking model, the singularity can be removed [126,127]. However, incorporating such fluctuations into the extant CTRW theories [27] for migration through an array is not trivial. The goal here is to study these fluctuations and determine whether they play an important role in experiments and therefore need to be included in the CTRW model.

The present chapter thus focuses on the first-passage time statistics of the unhooking of a long DNA molecule from a cylindrical post in a DC electric field. Two objectives motivate our approach: (i) Demonstrate that simple and efficient translocation models provide excellent agreement with results obtained using the Fokker-Planck equation and Brownian dynamics simulations. (ii) Take advantage of this computational efficiency to study the long chains used in experiments over a wide range of electric fields and initial configurations. Taken together, these two results will allow us to provide support for CTRW models of the transport in an array that incorporate a deterministic rope-over-pulley unhooking model. The rest of the chapter is organized as follows. In Sec. 4.2, we show how the unhooking of a DNA chain from a cylindrical post can be mapped to a translocation problem. The methods employed to study such motion are described in Sec. 4.3. A discussion of the results we obtain is the subject of Sec. 4.4 and some concluding remarks will follow in Sec. 4.5.

4.2 From translocation to unhooking

4.2.1 General considerations

We investigate in the following the unhooking of a single DNA chain wrapped around an isolated, fixed, cylindrical obstacle. The equally important problems of chain collision and translation in an array of posts are not considered in this thesis; these problems are often treated separately [27,133]. We only consider U and J-shaped collisions where the

unraveling and unhooking processes can be decoupled. This decoupling is essential to the partially separable probability distribution function invoked to describe the motion of the chain in an array [27]. Previous results from Randall and Doyle [132] indicate that unraveling and unhooking of the chain are sometimes coupled (the so-called X-collision). However, an effective J-model with a shortened chain adequately describes the X-collisions observed in experiments [132].

We model the DNA as a polymeric chain of total charge Q and contour length $L = N\,l_p$, where N is the number of persistence lengths of size l_p. The parameters N and l_p are chosen to reflect typical experimental situations. For double-stranded DNA, $l_p \sim 53$ nm (or 150 base pairs) and typical experiments use DNA of size $N \sim$ 50-500. Upon collision with the stationary post, two well-defined "arms" are formed on either side of the post. The applied electric field of magnitude E leads to an external tension (which favors stretching) that exceeds the entropic energy (which favors coiling) [143]. As a result of the excess external tension, the coiled arms of the chain extend on either side of the post to near to their full contour length, giving rise to U and J-shaped conformations. This simple picture is supported by experimental evidence on DNA molecules a tens of microns long, which can be extended under the action of electric fields as low as 1 V/cm [126]. In our analysis, we focus on electric fields $E \sim$ 5-50 V/cm. In terms of the dimensionless Péclet number (to be defined below), Pe $\sim$ 0.1-1.

Our starting point is a chain enveloping a post with fully extended arms and an initial configuration defined by an offset $-L \leq x_0 \leq L$ between the two arms. An offset of $x_0 = 0$ corresponds to a chain with equal arms, while an offset of $x_0 = \pm L$ represents a chain fully located on one side of the post. The diffusive motion, combined with the convection due to the electric field, causes the chain to sway back and forth around the post in a rope-over-pulley manner until the longer arm eventually engulfs its smaller counterpart. A full description of the motion of the polymer can be obtained via the probability distribution function of the exit (or unhooking) time, i.e., the time necessary for the entire mass of the chain to reside on one side of the post. This time is a function of the initial offset, the field and the chain properties. In the case of a purely convective disengagement, the deterministic unhooking time is

$$t(x_0) = \frac{L}{2U} \ln\left(\frac{L}{|x_0|}\right), \tag{4.1}$$

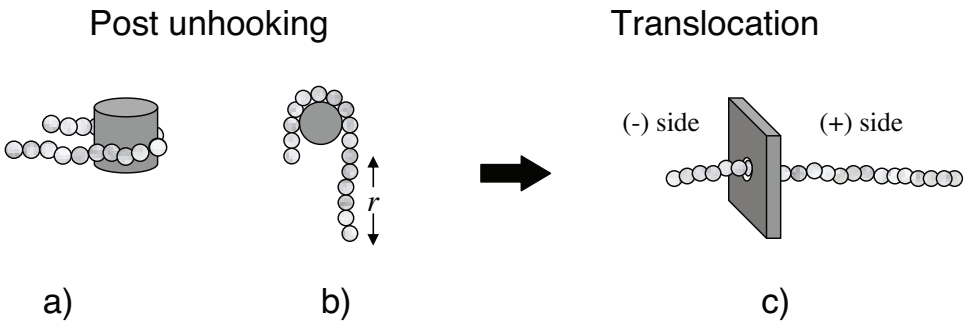

Figure 4.2: a) Side view of DNA chain with extended arms unhooking from a cylindrical obstacle. b) Top view of the same chain. c) Translocation of the polymer through a narrow hole in an impenetrable membrane. The parameter r represents the instantaneous offset between the lengths of the two arms.

where U is the curvilinear velocity of the chain during unhooking. As mentioned in the introduction of this chapter, it becomes now clear that when the chain starts with equal arms on both sides of the post, the deterministic unhooking time is logarithmically divergent. However, if x_0 is modeled as a uniformly distributed random variable, the logarithmic divergence of Eq. (4.1) at $x_0 = 0$ is integrable [127].

4.2.2 Unhooking from a post as a translocation process

In our analysis, we would like to avoid computational issues inherent in the Fokker-Planck model [126] and Brownian dynamics [127] simulations used in previous studies. As illustrated in Fig. 4.2, if one neglects the size of the insulating post and the curvature of the electric field lines it induces, the two extended arms can be envisioned as being situated on two sides of an impenetrable "membrane." In Sec. 4.4.2, we will assess the validity of these assumptions. Transfer of mass between arms occurs only through a narrow opening in the "membrane" under the action of a well-defined force that depends on the difference in length between the arms. When viewed in this light, the unhooking process can be analyzed using the plethora of methods developed for DNA translocation through nanopores [87, 88, 144, 145].

At this point in our discussion, we would like to stress that the analogy between the translocation of a chain through a nanopore and the unhooking from a post lies in

the methods utilized to characterize the two motions rather than the actual physical mechanisms underlying these processes. Chains translocating through a nanopore are pulled via the combination of entropic elasticity and an external force (an electric field or a chemical potential difference), while the chain disengages from the post under the action of an electric field. Nevertheless, both processes can be described as the biased motion of 1-D Brownian particle in a well-defined potential field, governed by the free energy of the chain on either side of the "membrane." In characterizing such motion, the value of the total force driving the chain will prove more relevant than its actual physical nature. Taking advantage of the parallel between the unhooking problem and the translocation of a chain, the probability distribution function and the moments of the exit times can be computed using the methods outlined below.

4.3 Analysis methods

4.3.1 Fokker-Planck (FP) approach

The first attempt to include fluctuations in modeling the rope-over-pulley collision involved a Fokker-Planck-type equation [126]. In this approach, the two arms of the chain around the post exhibit an offset x at time t with a probability $p \equiv p(x,t)$, governed by a convection-diffusion equation of the form

$$\frac{\partial p}{\partial t} = 4D\frac{\partial}{\partial x}\left[\frac{\partial p}{\partial x} + \frac{1}{k_B T}\frac{\partial U(x)}{\partial x}p\right]. \tag{4.2}$$

In the latter, D is the free-solution diffusion coefficient of the chain and, for the post collision, $U(x) = -QEx^2/2L$ is the inverted harmonic electric potential engendered by the offset between the two arms [126]. The factor 4 arises because, when describing the translocation of a chain through a membrane, the Fokker-Planck equation is usually written in terms of the coordinate number n, corresponding to the number of segments on one side of the membrane, rather than the offset x [87]. A change of variables is then necessary to obtain the appropriate equation in terms of the offset. The governing equation for p is supplemented with the absorbing boundary conditions at $x = \pm L$

$$p(\pm L, t) = 0, \tag{4.3}$$

as well as the initial condition

$$p(x,0) = \delta(x - x_0). \tag{4.4}$$

The following dimensionless parameters are introduced for the rest of this chapter:

$$\mathrm{Pe} \equiv \frac{qEl_p}{k_BT}, \quad \tau \equiv \frac{t}{t_p}, \quad r \equiv \frac{x}{l_p}, \quad r_0 \equiv \frac{x_0}{l_p}, \tag{4.5}$$

where Pe is the Péclet number, k_BT is the thermal energy, $q = Q/N$ is the charge per persistence length and $t_p = l_p^2/2D_0$ is the diffusion time of a persistence length with a diffusion coefficient of $D_0 = ND$. The latter scaling implies Rouse dynamics for the diffusion of this highly stretched chain, consistent with the conclusions of André *et al.* [146]. The governing equation for p in Eq. (4.2) becomes

$$\frac{\partial p}{\partial \tau} = \frac{2}{N}\frac{\partial}{\partial r}\left(\frac{\partial p}{\partial r} - \frac{\mathrm{Pe}}{2}\, r\, p\right), \tag{4.6}$$

subject to the boundary conditions

$$p(-N,\tau) = p(N,\tau) = 0 \tag{4.7}$$

and the initial condition

$$p(r,0) = 2\delta(r - r_0). \tag{4.8}$$

The probability that the chain still remains hooked at time τ, or survival probability, is given by

$$S(\tau) = \int_{-N}^{N} p(r,\tau)\, dr, \tag{4.9}$$

while the probability distribution $\Pi \equiv \Pi(\tau)$ for observing an exit event before time τ can be expressed as

$$\Pi(\tau) = -\frac{\mathrm{d}S}{\mathrm{d}\tau}. \tag{4.10}$$

Given an initial offset r_0, the n^{th} moment of the exit time, $\langle \tau_e^{(n)} \rangle$, is given from the probability distribution by

$$\langle \tau_e^{(n)} \rangle = \int_0^\infty \tau^n\, \Pi(\tau)\, \mathrm{d}\tau. \tag{4.11}$$

Using the above notation and electric potential, the average exit time (or mean first passage time), $\langle \tau_e \rangle \equiv \langle \tau_e^{(1)} \rangle$, can be explicitly calculated by [126, 147]

$$\langle \tau_e(r_0) \rangle = \frac{2N}{\mathrm{Pe}} \int_0^\infty \frac{\mathrm{e}^{-4y^2}}{y} \sin[\lambda y(1 - r_0)] \sin[\lambda y(1 + r_0)]\, \mathrm{d}y, \tag{4.12}$$

where $\lambda \equiv N\sqrt{\mathrm{Pe}}$. The equivalent deterministic time in Eq. (4.1) becomes in dimensionless notation

$$\tau_d(r_0) = -\frac{N}{\mathrm{Pe}} \ln |r_0|. \tag{4.13}$$

When the two arms are initially of equal length, *i.e.*, $r_0 = 0$, the expression in Eq. (4.12) does not entail any singularities . Indeed, if $r_0 = 0$, the net force on the chain is zero; it is in an unstable state, sitting on top of the inverted harmonic potential. As thermal fluctuations will inevitably push the chain, even slightly, away from that position, the resulting non-zero electrical force leads to the chain unhooking in a finite time.

Using this Fokker-Planck approach, Volkmuth *et al.* [126] compared the average exit time given by the above expression with experimental results on the unhooking of a chain from a single post. While the closed-form expression of the average exit time in Eq. (4.12) makes the Fokker-Planck approach attractive at first sight, no closed-form expression for Π or its higher moments is available. Thus, numerical manipulation of Eqs. (4.6), (4.9) and (4.10) becomes necessary to obtain information about the variance of the unhooking times. As soon as the number of persistence lengths N exceeds a few hundred, accurate solutions necessitate smaller and smaller discretization steps to numerically integrate Eq. (4.6). Although feasible, the computationally intense numerical integration constitutes a major limitation of this method. In the discussion below, we show how the entire probability distribution of the exit time and its moments can be calculated without solving for $p(r, \tau)$.

4.3.2 Monte Carlo (MC) simulations

Gauthier and Slater [145] devised a simple calculation method to compute the probability distribution Π of the exit time τ_e of a chain translocating through a nanopore, which we will adapt to the problem at hand. This method is of particular interest here because it is simple, fast and efficient, as one need not resort to a numerical integration of the partial differential equation in Eq. (4.6) and its solution $p(r, \tau)$ to obtain Π. In this algorithm, the motion of the chain during the translocation process is modeled as a series of discrete jumps between the + and – side of the "membrane" used to model the steric effect of the post. The chain proceeds accordingly from one side to the other by

diffusion and electric field-induced convection until all the segments reside on the same side. If $n_\pm$ is the number of segments on the $\pm$ side of the post depicted in Fig. (4.2), the chain diffuses in an electric potential

$$\frac{U}{k_B T} = -\frac{1}{2}\,\mathrm{Pe}\,(n_+ - n_-)^2. \tag{4.14}$$

The force F acting upon the chain is thus

$$F = -\frac{1}{l_p}\frac{\partial U}{\partial n_+}, \tag{4.15}$$

or, in dimensionless form,

$$\epsilon \equiv \frac{F l_p}{2k_B T} = \frac{\mathrm{Pe}}{2}(n_+ - n_-) = \frac{\mathrm{Pe}}{2}(2n_+ - N). \tag{4.16}$$

At every time step of the translocation process, the chain motion is that of a 1-D Brownian particle in the potential U between two absorbing walls. Starting halfway between the walls, the particle undergoes a biased random walk until it is absorbed by one of the two walls after a finite time Δt. In other words, the chain will "jump" either to the right (+) or to the left (–) configuration with a probability p_+ or p_- in an average time Δt. However, this simple description is incomplete. In order to retrieve the correct mobility and diffusion coefficient without introducing higher moments of the exit time, one must also consider cases where the chain does not perform a jump during Δt, but rather remains in the same configuration [102]. Time steps without a jump are necessary to correctly simulate the biased random walk using constant time steps and calculate its transport properties. In this particular case, the jump probabilities in Eq. (2.47) must be appropriately modified to account for temporal fluctuations without the need for higher moments of the unhooking time. We consider then a chain initially distributed such that $n_+ - n_- = r_0$. Under the action of the force ϵ, the chain jumps to the +/– configuration with the probabilities [145]

$$p_\pm = \frac{1 - s(\epsilon)}{1 + \mathrm{e}^{\mp 2\epsilon}} \tag{4.17}$$

or remains at rest with the probability [145]

$$s(\epsilon) = \frac{\coth \epsilon}{\epsilon} - \mathrm{csch}^2\,\epsilon \tag{4.18}$$

during the dimensionless time [145]

$$\Delta t(\epsilon) = [1 - s(\epsilon)]\frac{\tanh \epsilon}{\epsilon} t_b'. \tag{4.19}$$

In the latter, $t_b' = N$ is a dimensionless rescaling of the Brownian time that takes into account the frictional effects during the motion of the chain [145].

The algorithm proceeds and time is updated until $n_+ = 0$ or N, hence obtaining one realization of the exit time τ_e. The probability distribution Π can thus be generated using multiple runs of the random walk simulation starting at the same initial offset.

4.3.3 Master equation (ME) approach

If one is interested in obtaining moments of the exit time for different initial configurations, one must generate ensembles of Monte Carlo calculations for each of the desired initial configurations. While certainly feasible, this is not computationally efficient. Such difficulty can be circumvented if one resorts to a master equation formulation approach [144], which was also developed for translocation by Gauthier and Slater [145]. Here, the chain, modeled as a series of N segments subject to the force ϵ, can be found at each particular time step in any of $N+1$ possible configurations. The configurations are defined by the excess (or deficit) number of segments on the + side of the membrane. Let P_+^i and P_-^i denote the probabilities to escape on the + and – side, respectively, given that the chain starts at some configuration $1 \leq i \leq N+1$; $i \equiv n_+ + 1$ such that $i = 1$ corresponds to a chain escaping on the – side, and $i = N+1$ to a chain escaping on the + side.

Using a master equation approach, a chain in configuration $2 \leq i \leq N$ at a particular point in time will assume a configuration $(i-1, i, i+1)$ at the next time step. The corresponding transition probabilities are (p_-, s, p_+). The escape probabilities $P_\pm^i$ can thus be calculated using the tridiagonal system [145]

$$P_\pm^i = p_+^i P_\pm^{i+1} + s^i P_\pm^i + p_-^i P_\pm^{i-1}, \tag{4.20}$$

where $p_\pm^i$ and s are defined in Eq. (4.17) for each configuration i. The absorbing boundary conditions when the polymer is entirely located on either side ensure that $P_+^1 = 0, P_+^{N+1} = 1$ and $P_-^1 = 1, P_-^{N+1} = 0$. Note that since all chains ultimately

exit through one of the two sides for any given initial configuration i, we must have $P_+^i + P_-^i = 1$.

The mean exit times $\tau_\pm^i$ through the $\pm$ side are obtained by solving well-defined tridiagonal systems [145]. However, interest lies here in the average time $\langle \tau_e^i \rangle = P_+^i \tau_+^i + P_-^i \tau_-^i$ to exit on either side. Combining the equation for $\tau_\pm$ and $P_\pm$, one arrives at the following system for $\langle \tau_e^i \rangle$:

$$\langle \tau_e^i \rangle = p_+^i \langle \tau_e^{i+1} \rangle + s^i \langle \tau_e^i \rangle + p_-^i \langle \tau_e^{i-1} \rangle + \Delta t^i. \tag{4.21}$$

The average exit time of a chain with a configuration i at a particular step is thus equal to the weighted average exit time to escape starting at adjacent configurations $(i-1, i, i+1)$ at the next step, augmented by the time elapsed during the step, Δt^i. The previous system is supplemented by the boundary conditions $\langle \tau_e^1 \rangle = \langle \tau_e^{N+1} \rangle = 0$.

One can repeat this procedure to compute higher moments of τ_e. For the sake of completeness, we provide the tridiagonal system to be solved for the second moment $\langle \tau_e^{i^{(2)}} \rangle$:

$$\langle \tau_e^{i^{(2)}} \rangle = p_+^i \langle \tau_e^{i+1^{(2)}} \rangle + s^i \langle \tau_e^{i^{(2)}} \rangle + p_-^i \langle \tau_e^{i-1^{(2)}} \rangle + \langle \tau_e^i \rangle^2 - (\langle \tau_e^i \rangle - \Delta t^i)^2, \tag{4.22}$$

supplemented by $\langle \tau_e^{1^{(2)}} \rangle = \langle \tau_e^{N+1^{(2)}} \rangle = 0$. The standard deviation σ of the exit time is thus given by

$$\sigma = \sqrt{\langle \tau_e^{i^{(2)}} \rangle - \langle \tau_e^i \rangle^2}. \tag{4.23}$$

The enthusiasm raised by this rapid procedure of calculating the moments of the exit time for all possible initial configurations in one single calculation should nevertheless be tempered by the fact that one can only obtain the moments of the exit time, rather than the full probability distribution Π of the exit time. However, several techniques [148] can be used to build an approximation of Π from its moments. We do not engage in such a task though, since the exact distribution is readily available from the previously outlined Monte Carlo algorithm.

4.3.4 Brownian dynamics (BD) simulations

In order to compare results from MC and ME calculations, a Brownian dynamics (BD) simulation code was also implemented[1] . The BD simulations employ a bead-spring

[1] The BD code for this study was written by Jaesol Cho.

model with a chain of contour length L represented by N_b beads connected by $N_s = N_b - 1$ springs. Each spring contains $\nu = l_s/l_p$ persistence lengths, where l_s is the maximum extension of a spring and l_p is the persistence length of DNA. The total number of persistence lengths (or, equivalently, MC beads) is therefore $N = \nu N_b$.

The position of each bead, $\mathbf{r}_i$, obeys the following equation of motion [149]

$$\frac{d\mathbf{r}_i}{dt} = -\mu_0 \mathbf{E}(\mathbf{r}_i) + \frac{1}{\zeta}(\mathbf{F}_i^B + \mathbf{F}_i^{EV} + \mathbf{F}_i^T), \tag{4.24}$$

where μ_0 is the bead electrophoretic mobility and ζ is the bead drag coefficient. Exact expressions for the components of the electric field, $\mathbf{E}$, as a function of the bead position can be obtained analytically by conformal mapping [131]. The Brownian force, $\mathbf{F}^B$, is computed from a Gaussian distribution with zero mean and correlation

$$\langle \mathbf{F}_i^B(t_1)\mathbf{F}_j^B(t_2)\rangle = \frac{2k_BT\zeta\delta_{ij}}{\delta t}\mathbf{I}, \tag{4.25}$$

where δ_{ij} is the Kronecker delta function, δt is the time step and $\mathbf{I}$ is the unit tensor. The excluded volume force, $\mathbf{F}^{EV}$, is derived from the gradient of the soft potential developed by Jendrejack *et al.* [150],

$$U^E(r_{ji}) = \frac{1}{8}v^{\mathrm{EV,p}}k_BT\nu^2\left(\frac{9\nu}{4\pi l_s^2}\right)^{3/2}\exp\left(-\frac{9\nu r_{ji}^2}{4l_s^2}\right), \tag{4.26}$$

where $r_{ji} = |\mathbf{r}_i - \mathbf{r}_j|$ and $v^{\mathrm{EV,p}} = 0.0004\ \mu\mathrm{m}^3$ [149] is the excluded volume of a bead. Each bead is also subject to a tension force, $\mathbf{F}^T$, that results from the flexibility of the adjacent springs. For a worm-like chain, the force, $\mathbf{f}_{i,j}^{\mathrm{s}}$, acting on bead i from the spring connecting adjacent beads i and $j = i + 1$ or $j = i - 1$ obeys the Marko-Siggia spring force [151],

$$\mathbf{f}_{i,j}^{\mathrm{s}} = \frac{k_{\mathrm{B}}T}{A_{\mathrm{eff}}}\left[\frac{r_{ji}}{l_s} - \frac{1}{4} + \frac{1}{4(1_s - r_{ji}/l_s)^2}\right]\frac{\mathbf{r}_j - \mathbf{r}_i}{r_{ji}}. \tag{4.27}$$

The parameter A_{eff} denotes the effective persistence length, which is determined by a worm-like chain model with the Underhill-Doyle low-force criteria [152]. The tension force, $\mathbf{F}_i^T$, on each inner bead is thus the sum of the two adjacent spring forces, whereas, for the two end beads, it is given by the force from the one adjacent spring.

In the simulations, the force balance given by Eq. (4.24) is non-dimensionalized using appropriate length, time and force scales for a BD model. Since the motion of the chain can only be resolved down to the spring length scale in BD, we choose the maximum

extension of the spring, $l_s = \nu l_p$, as a characteristic length. Likewise, if $D_s = D_0/\nu$ is the diffusion coefficient of a spring composed of ν persistence lengths of diffusion coefficient D_0, then the diffusion time of a spring, $t_s \equiv l_s^2/2D_s = \nu^3 t_p$, is a proper characteristic time scale for the motion of the chain. Following Kim and Doyle [149], the electric force can be expressed in a dimensionless form as $F_s \equiv \mu_0 E l_s/D_s = \nu^2 \mathrm{Pe}$.

Note that these definitions of the motion characteristic scales differ from those introduced previously in Eq. (4.5) for MC and ME calculations. In the former, the spring constitutes the elementary unit of the chain, while in the latter, all quantities are defined relative to a persistence length, considered to be the elementary unit.

The equation of motion in (4.24) is integrated using an adaptive method [149]. The interaction between the beads and the post (non-penetrability) is enforced using the Heyes-Melrose algorithm [149,153]. Due to the presence of $\mathcal{O}(1)$ gradients of the electric field near the post, small time steps are required to achieve an adequate resolution of the motion of the beads there, resulting in a high computational cost.

In our simulations in Sec. 4.4.2, we represent a chain $L = 10.25$ microns long, with $N_b = 29$ Brownian beads (or $N \sim 194$ persistence lengths), leading to a maximum spring extension $l_s = 366.1$ nm. This chain length represents half the contour length of stained λ-DNA. To initialize the chain, the 15^{th} bead is placed at the center of the backside of an infinitely extended cylindrical obstacle of diameter $d = 1\mu$m. The other beads are distributed symmetrically around the obstacle with a spacing equal to the equilibrium spring length, $\sqrt{\nu} l_p$. Upon reaching an angle of $\pm\pi/2$ along the obstacle contour, any remaining beads are placed with the same spacing $\sqrt{\nu} l_p$ in the direction of the nominal electric field. We found that simulations starting from this symmetric state are very sensitive to the number of beads used to represent the chain, but a choice of 29 beads yields satisfactory results.

4.4 Results and discussion

4.4.1 Influence of diffusion

We first demonstrate the consistency of the unhooking probability distribution and moments of that distribution obtained using the Fokker-Planck, Monte Carlo and master equation methods. For this purpose, we consider a chain composed of 50 beads, thus

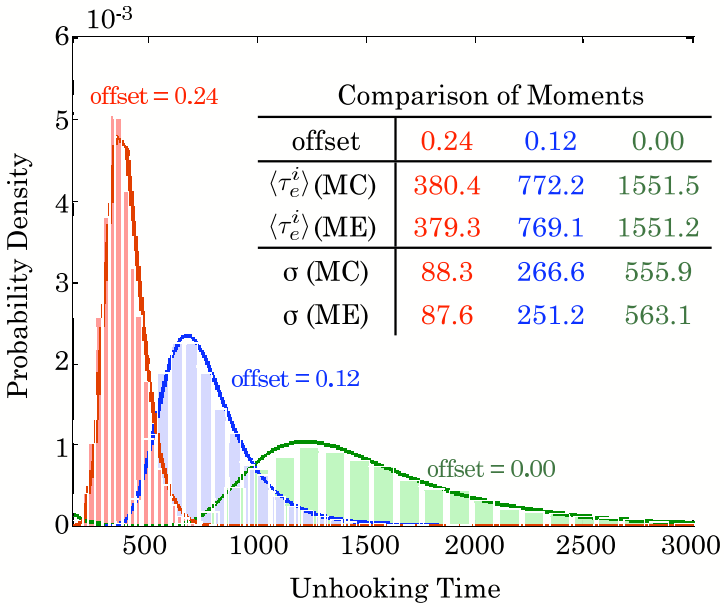

Figure 4.3: Probability distribution of the exit time for several fractional initial offsets. An initial offset of zero corresponds to two equal arms and an offset of one represents a chain fully located on one side of the post. The histograms are obtained from the Monte Carlo (MC) algorithm and the solid lines from the Fokker-Planck equation in Eq. (4.10). The inset shows comparisons of the average $\langle \tau_e^i \rangle$ and standard deviation σ of the exit time calculated using the MC and ME methods.

corresponding to a total contour length of about 7.5 kbp. The chain, wrapped around the post, can unhook through the combined action of diffusion and the electric force. The Péclet number, Pe = 0.1, corresponds to $E \sim 5$ V/cm. Figure 4.3 presents the probability distribution of the exit time for various initial offsets between the two extended arms of the chain. The histograms are obtained from 50,000 realizations of the exit time using the Monte-Carlo algorithm described above, while the solid lines correspond to the numerical integration of the Fokker-Planck approach. As can be seen from the figure, the discrete MC approach yields a distribution in accordance with the continuous FP approach. The table in the inset of Fig. 4.3 compares the average and standard deviation of the exit time computed using the MC and ME methods. Here too, the agreement between the ensemble averages of the MC calculations and the exact ME method is very good for all initial offsets. Thus, we will proceed with the ME approach to compute the moments, owing to its computational efficiency. For subsequent comparisons between the present approach and unhooking time distributions obtained by Brownian dynamics simulations, we will use the MC approach.

In addition to illustrating the concordance between the methods, Fig. 4.3 illuminates the role of the initial offset on the exit time. As the initial offset decreases, the peak of the exit time probability distribution is shifted towards increasing values of time, meaning a slower unhooking process. This shift is accompanied by a broadening of the probability distribution, indicating an increase in the exit time fluctuations.

Both effects are understood by recalling that the driving force is proportional to the difference between the length of the two arms, as given by Eq. (4.16). Accordingly, for small offsets, diffusive effects (responsible for broadening of the distribution) become important since the convective driving force (responsible for the location of the peak) is reduced. The critical offset value, N_c, at which the two effects are comparable in magnitude can be estimated if we consider the point at which convection and diffusion are balanced [134]. The characteristic time for a chain of size N to diffuse over a distance $N_c a$ is

$$t_d \sim \frac{{N_c}^2 a^2}{D_0/N} \sim N {N_c}^2, \tag{4.28}$$

while the convective time scale for the same chain subject to a force $F \sim N_c \mathrm{Pe}$ is

$$t_c \sim \frac{N_c a}{F/\xi} \sim \frac{N_c}{N_c \mathrm{Pe}/N} \sim \frac{N}{\mathrm{Pe}}. \tag{4.29}$$

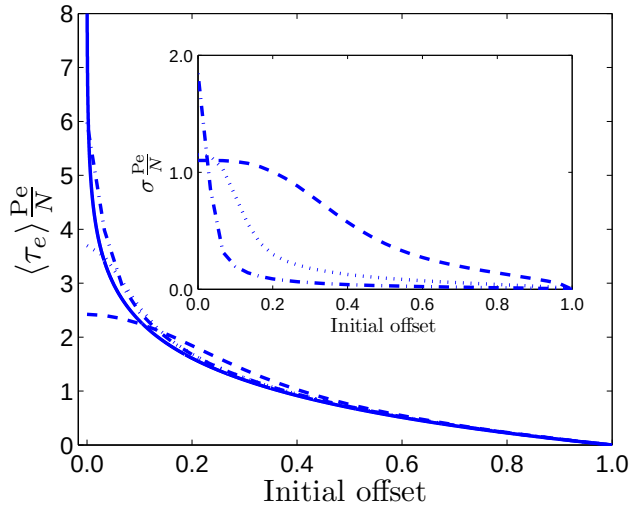

Figure 4.4: Plot of the rescaled average exit time $\langle \tau_e \rangle$ Pe$/N$ as a function of the initial offset between the two arms for various Péclet numbers, Pe $= 0.02$ $(--)$, Pe $= 0.2$ $(\cdots)$ and Pe $= 2$ $(- \cdot -)$ and a value of $N = 60$. The average exit time was obtained using a master equation (ME) calculation and compared to the deterministic rope-over-pulley (solid line). The inset shows the variations of the standard deviation of the exit time, also scaled with Pe$/N$.

Balancing the two time scales, we arrive at the condition

$$\lambda^2 = {N_c}^2 \mathrm{Pe} \sim 1. \tag{4.30}$$

This situation is reminiscent of a fluid boundary layer near a flat plate where viscous (diffusive) effects are non-negligible while the flow beyond the boundary layer is purely inertial (convective). This analogy can be extended to the thickness δ of the boundary layer as a function of the Reynolds number (or driving force), $\delta \sim 1/\sqrt{\mathrm{Re}}$, or its equivalent in our case, $N_c \sim 1/\sqrt{\mathrm{Pe}}$.

The parameter $\lambda \equiv N\sqrt{\mathrm{Pe}}$ is thus critical in that it distinguishes between two regimes: for $\lambda^2 \ll 1$, the chain unhooking is governed by diffusive effects; for $\lambda^2 \gg 1$, the chain moves deterministically under the action of the electric field. This distinction is further highlighted in Fig. 4.4. Here, we plot the rescaled average and standard deviation of the exit time as a function of the initial offset for different Pe. The plot also contains

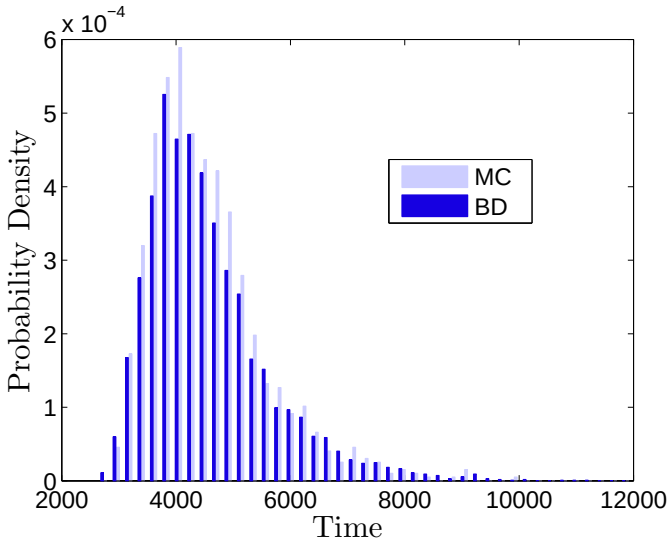

Figure 4.5: Probability distribution of the exit time of a chain with $N = 194$ Monte Carlo beads, using an electric field Pe $= 0.21$. The chain has initially a zero offset between the arms. The histograms are the results of Brownian dynamics (BD) simulations and Monte Carlo (MC) calculations. The two distributions share the same features, namely, the shape, the peak position and the width of the distribution.

the deterministic exit time obtained from the logarithmic rope-over-pulley model in Eq. (4.13). Note that the standard deviation in the deterministic case is identically zero. As expected, the impact of diffusive fluctuations is strongest when the initial offset is zero. The purely deterministic model incurs a singularity at this point, since it does not include a mechanism to break the symmetry. As anticipated from the argument above, both the mean unhooking time and its standard deviation decrease as the initial offset increases because the process becomes more and more deterministic. Importantly, as the offset increases, the rescaled average exit time $\langle \tau_e \rangle N/\mathrm{Pe}$ lines collapse onto the deterministic, purely convective curve. Note that as Pe is increased, the collapse occurs for smaller values of initial offset, as dictated by the condition $N_c \sim 1/\sqrt{\mathrm{Pe}}$. We tried to verify the scaling in Eq. (4.30) by plotting a critical initial offset N_c as a function of Pe. Unfortunately, the value of N_c proved very sensitive to the criterion defining the transition zone.

4.4.2 Comparison with Brownian dynamics

While we have demonstrated the consistency between the Fokker-Planck, Monte Carlo and master equation stochastic methods outlined above, one might legitimately inquire about the validity of the assumptions underlying the transport model used in these calculations. More specifically, the results we presented above rely on three key assumptions: (i) The chain flexibility was ignored. Rather, both arms were assumed to be completely stretched upon collision with the post; (ii) The post has a vanishing diameter. The experiments cited at the outset [23–26] typically use post diameters commensurate with the radius of gyration of the chain; and (iii) The electric field was assumed to be uniform. In typical experiments, the insulating post generates field non-uniformities, both in magnitude and direction, in the vicinity of the post. These field gradients can cause lateral displacements of the chain that are not taken into account in our models.

To test the validity of our assumptions, more realistic Brownian dynamics simulations were performed to study the chain unhooking problem. As mentioned earlier, our comparison with the Brownian dynamics simulations corresponds to the unhooking of a 10.25 μm chain (or $N \sim 194$ persistence lengths), using a nominal electric field of 10 V/cm (or Pe ~ 0.21). Figure 4.5 compares the probability distribution of the exit time obtained from the BD simulations and the MC calculation. The BD histograms are the result of 1000 runs of the simulation and the MC bars are obtained from 5000 realizations of the exit time.

The agreement between the two methods is very good given the limited number of computationally expensive BD simulation runs. Qualitatively, the two distributions exhibit identical features (shape, peak position and width of the distribution). To further investigate the quantitative agreement, we plot in Fig. 4.6 the average, $\langle \tau_0 \rangle \equiv \langle \tau_e \rangle_{r_0=0}$, and standard deviation, $\sigma_0 \equiv \sigma_{r_0=0}$, of the exit time for the same chain as a function of the Péclet number. The BD data in this figure correspond to 100-300 simulations at each value of the Péclet number. As can be seen from the figure, the agreement between the fast ME calculation and the much slower BD simulation is very satisfying, in particular for the average exit time. The discrepencies of the standard deviation can be improved with more simulation runs, albeit at a high computational cost. The previous results from Figs. 4.5 and 4.6 validate our assumptions in modeling the chain. The flexibility of the chain barely affects the escape process and one can neglect the electric

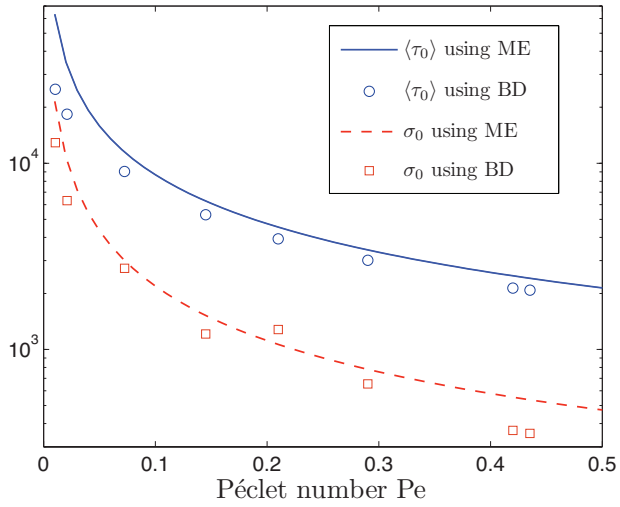

Figure 4.6: Plots of the average $\langle \tau_0 \rangle$ and standard deviation σ_0 of the exit time of a chain ($N = 194$) starting with equal arms. Both quantities are computed from Brownian dynamics (BD) simulations and a master equation (ME) calculation.

field non-uniformities with minimal loss of accuracy. When evaluating the transport coefficients of the unhooking chain, one can thus resort to fast, easily implemented, stochastic methods that perform as well as time-consuming, complex BD simulations.

4.4.3 Offset-averaged behavior

Having now established the consistency and accuracy of these translocation methods, we would like to use them to gain insights into experimentally relevant parameter ranges. To aid in the analysis, let us introduce now the random variable $\bar{\tau}_e$ defined as the "mean exit time" or the average of the exit times corresponding to uniformly distributed initial offsets. Nixon and Slater [127] previously considered both uniform and Gaussian distributions of initial offsets in a simpler model and concluded that the choice of distribution has little effect on the scaling behavior of the exit times [127]. If M_0 is the number of possible offsets, then the n^{th} moment of $\bar{\tau}_e$ can thus be obtained by

$$\langle \bar{\tau}_e{}^{(n)} \rangle = \frac{1}{M_0} \sum_{i=1}^{M_0} \langle \tau_e^{i(n)} \rangle. \tag{4.31}$$

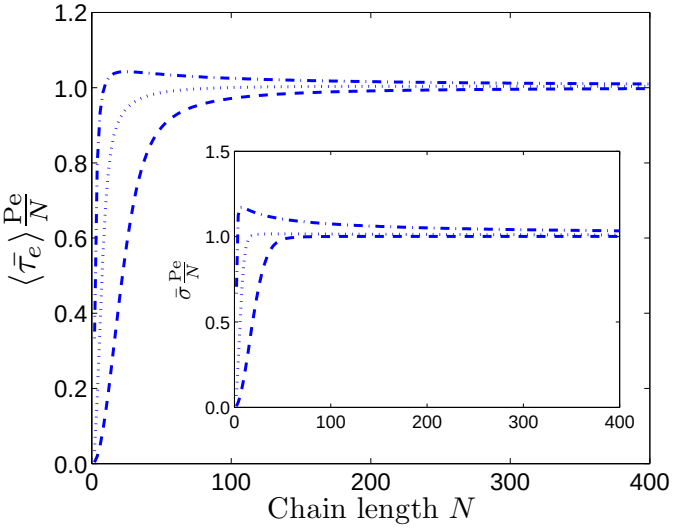

Figure 4.7: Plot of the rescaled average "mean exit time" $\langle \bar{\tau}_e \rangle$ Pe$/N$ as a function of the chain length N for various Péclet numbers, Pe = 0.01 $(--)$, Pe = 0.1 $(\cdots)$ and Pe = 1 $(- \cdot -)$. The "mean exit time" is obtained by averaging the exit time τ over all possible initial offsets. The inset shows the same calculation for the standard deviation of the "mean exit time."

We can therefore write for the average of $\bar{\tau}$

$$\langle \bar{\tau}_e \rangle = \frac{1}{M_0} \sum_{i=1}^{M_0} \langle \tau_e^i \rangle, \tag{4.32}$$

and standard deviation

$$\bar{\sigma} = \sqrt{\langle \bar{\tau}_e^{(2)} \rangle - \langle \bar{\tau}_e \rangle^2}. \tag{4.33}$$

The rescaled $\langle \bar{\tau}_e \rangle$ and $\bar{\sigma}$ are plotted in Fig. 4.7 as a function of chain length N for various values of Pe. Both $\langle \bar{\tau}_e \rangle$ and $\bar{\sigma}$ saturate to the deterministic value N/Pe [26] after a "boundary layer" zone of size $N_c \sim 1/\sqrt{\mathrm{Pe}}$, explaining why the saturation occurs earlier for high values of Pe. Also, both the average and standard deviation of the "mean exit time" overshoot their saturation value for small values of N. This should not appear very surprising if one recalls that the zero offset exit time dominates the summation for $\bar{\tau}_e$. For small values of N and a zero initial offset, diffusive fluctuations dominate the early stages of the unhooking, thus leading to an averaged exit time greater than

its saturation (deterministic) value. This overshoot was already apparent from Fig. 4.4, where $\langle\tau_0\rangle N/\mathrm{Pe} \approx 6$ and $\sigma_0 N/\mathrm{Pe} \approx 2$ for a value of $\mathrm{Pe} = 2$. Note that the average and standard deviation of the exit time over all uniformly distributed initial offsets for the deterministic rope-over-pulley model is N/Pe (the singularity at zero initial offset is integrable [127]), independent of N or Pe [26]. In other words, in the high Péclet regime, the diffusive fluctuations of the chain are negligible and the chain behaves deterministically. In the language of fluid mechanics, the flow is uniform everywhere with no boundary layer.

In experiments, one typically deals with chains several microns long ($N \sim$ 200-500) animated by an electric field of about 10 V/cm ($\mathrm{Pe} \sim 0.2$). Figure 4.7 indicates that, for this range of N and Pe, the transport properties of the chain when it disengages from the post are dominated by convective effects; both the mean and standard deviation of the exit time saturate to their deterministic, diffusion-free, limits well before the experimental regime is reached. The result above lends credence to one of the main assumptions underlying the CTRW model discussed at the outset. In this model, the diffusive fluctuations of the chain upon collision are neglected. The dispersivity of the chain mainly arises due to differences in the initial offset between the arms, as well as variations in the distance traveled by the chain between collisions. Including the diffusive fluctuations in the CTRW model should, therefore, only affect the transport coefficients of the chain for a limited range of N and Pe, corresponding to the condition $\lambda^2 \ll 1$. Consequently, treating the unhooking process as a purely convective process is a valid assumption.

4.5 Conclusion

We have explored the problem of chain unhooking from a post using stochastic methods developed for DNA translocation through a nanopore. We have shown how these methods can be implemented to obtain the probability distribution of the unhooking time and its moments in a fast and efficient way, thereby establishing the consistency between the various methods. The Monte Carlo method provides the full probability distribution for a given initial offset, while the master equation approach leads to a single, direct calculation of the moments of the exit time for all possible initial offsets.

We have also confirmed that the simplified transport model used in the translocation calculation captures the behavior exhibited during more costly Brownian dynamics simulations. The results thus obtained indicate that the unhooking process is dominated by two regimes: (i) a purely convective regime, compatible with the deterministic model; and (ii) a "boundary layer" diffusion-dominated regime. The transition between the two regimes occurs at a critical offset of $N_c \sim 1/\sqrt{\mathrm{Pe}}$. When typical experimental values are considered, i.e., $N \sim 300$ and $\mathrm{Pe} \sim 0.1$, convective effects dominate and the chain unhooking process can be modeled in a deterministic way.

Chapter 5

Partition function of chains tethered inside a spherical cavity

This chapter is based on the publication:
N. Laachi and K. D. Dorfman
J. Chem. Phys., **132**, 084108, (2010).

5.1 Introduction

This chapter is dedicated to a description of the calculation of the free energy (or partition function) of a chain inside a spherical cavity. The partition function is at the basis for evaluating the forces at play when chains exchange segments between confining spheres [154, 155]. A similar situation is encountered in models of translocation [87, 88], where the forces are derived from differences in the free energy of parts of the chain on each side of a membrane [145]. Also of interest is the situation where a long flexible chain can simultaneously occupy multiple confining spheres connected through narrow openings, which we envisage in the next chapter.

The confinement-induced properties of long flexible polymers are the subject of ongoing research interest, extending beyond the study of polymer physics to numerous biological and physico-chemical applications [156, 157]. In the context of statistical physics, interest lies particularly in the restricted number of allowed configurations due to steric hindrances from the confining medium or membrane. In the case of planar and tubular confinements, the excess free energy, relative to an unrestricted chain, varies as $\Delta F \sim {R_G}^{1/\nu} \sim L$, where L is the contour length of the chain and $R_G \sim L^\nu$ its radius of gyration, with $\nu = 1/2$ for ideal chains and $\nu \approx 3/5$ for self-avoiding chains [59,158,159]. The situation is less clear in the case of confinement inside a spherical or cubical cavity. On one hand, Casassa [158] and Edwards *et al.* [159] analytically demonstrated the linear scaling of the free energy for ideal chains in a spherical or cubical cavity. Also, Monte Carlo simulations by Jaeckel *et al.* [160] and Abadie *et al.* [161] numerically confirmed this linear scaling for ideal chains and suggested a similar scaling for self-avoiding chains inside confining spheres or cubes. On the other hand, Grosberg *et al.* [162] argued that the monomer concentration increases inside an impermeable sphere upon confinement, resulting in a stronger dependence of the confinement energy on the size of the chain. As a consequence, the extensive picture of confinement breaks down and $\Delta F \sim L^{3\nu/(3\nu-1)}$ [162, 163]. This scaling was also confirmed by recent Monte Carlo simulations [157].

In addition to steric interactions that prevent the chain from crossing the interface, a few segments of the chain can also adsorb to the interface, thus essentially behaving as grafted or tethered sites (also called sticky feet) [164, 165]. Specific sites, like the

end segments, can be selected through a preferential chemical treatment of the chain to enhance their adsorption affinity [164]. The resulting tethered chains have even fewer allowed configurations than their merely confined counterparts. In the absence of confinement, single-tethered (ST) chains and, to a lesser extent, double-tethered (DT) chains have received considerable attention, owing to their role in modeling various polymeric structures, such as brushes or diblock copolymers (see Jones *et al.* [165] and the references therein).

Interest in polymeric chains attached to the surface of a confining medium arises in the modeling of DNA dynamics in various confining geometries [154,155,166–168]. In all these examples, the motion of DNA molecules is the consequence of segmental exchange between confined portions of the chain. The entropic contribution to the force driving DNA originates from differences in the free energy (or the partition function) between different parts of the chain. While several models of the free energy landscape have been introduced in previous publications, the combined effect of confinement, self-avoidance and tethering has not been explored.

The statistical properties of confined self-avoiding walks with fixed ends were first studied mathematically by Abbott *et al.* [169] and later by Whittington *et al.* [170]. They showed, among other things, that the average number of walks starting at a corner of a square and ending at the opposite corner scales like g^{a^2}, where a is the length of the square side and g is a lattice-dependent constant [171]. The latter scaling was further generalized to hypercubes of any dimension [172]. However, the direct enumeration [170–173] and transfer matrix methods [174–176] used to determine the exact number of confined self-avoiding walks are limited to short chains. Long chains are not amenable to such analytical methods. One needs to consider other treatments such as Monte Carlo simulations [160,161,177,178] to study the statistical properties of long, confined chains. Further, little has been reported on the combined effect of confinement and tethering on the conformation of the chain in the presence of an external potential field, although a fugacity (or bias) associated with each step of the walk is readily incorporated into exact enumeration methods.

We employ here a systematic approach to compute the partition function of self-avoiding chains with one or two ends attached to the surface of a confining sphere and subject to an additional bias from a external force. The method is based on a

decomposition of the partition function and is described in Sec. 5.2. In Sec. 5.3, the results obtained for chains in different conditions are presented, followed in Sec. 5.4 by some applications and extensions of the results derived here.

5.2 Analysis methods

We compute the partition function of a single and double tethered self-avoiding polymeric chain under spherical confinement through Monte Carlo simulations. The chain is composed of N beads connected by $N-1$ bonds, where each bond represents a basic unit of length, l, of the chain (a monomer or a Kuhn segment, for example). For simplicity, we take $l = 1$ in what follows. To keep the calculations tractable, the conformation of the chain is constrained to a cubic lattice of step size equal to l, whereby each bead occupies a single lattice site. An end bead is placed at the origin of the lattice while the other is free (for a ST chain) or anchored to a specified site in the lattice (for a DT chain). To allow for an anchoring point on the surface of the sphere, relevant values of the radius of the sphere, ρ, are integer multiples of l. The corresponding two-dimensional situation is depicted in Fig. 5.1(a).

In addition to the confinement constraint, we consider the more general case where each base unit of the chain is subject to a constant force $\mathbf{F} = F\mathbf{u}$, of magnitude F and a direction given by the unit vector $\mathbf{u}$. Then

$$E_i = -\mathbf{F} \cdot (\mathbf{R}_i - \mathbf{R}_1) = -F\,\mathbf{u} \cdot (\mathbf{R}_i - \mathbf{R}_1) \tag{5.1}$$

is an external potential field (a fluid flow or electric potential in the case of charged polymers), where $\mathbf{R}_i$ denotes the position of the i^{th} bead. The total energy of a chain with a configuration j, defined by the set of position vectors $\{\mathbf{R}_i\}^{(j)}_{1\le i\le N}$, is

$$E^{(j)} = \sum_{i=1}^{N} E_i^{(j)}. \tag{5.2}$$

The corresponding partition function of the chain is therefore

$$Z \equiv Z(N, \beta, F) = \sum_{(j)} \exp(-\beta E^{(j)}), \tag{5.3}$$

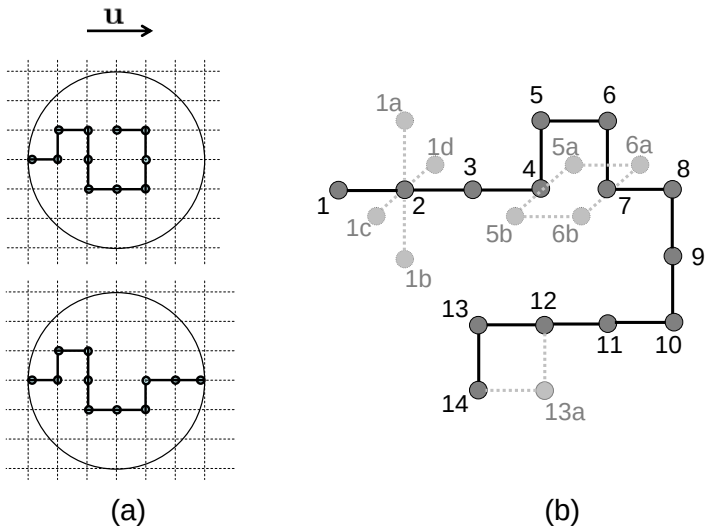

Figure 5.1: (a) Two-dimensional lattice polymer chain with 11 beads (10 bonds) inside a circle. The chain can be attached to the circle at one end (top) or two ends (bottom) and is subject to a constant force with direction **u**. (b) Three-dimensional lattice polymer chain with 14 beads. Examples of allowed moves during each Monte Carlo cycle are shown in gray: end bead 1 can rotate to any of 1a, 1b, 1c or 1d, beads (5,6) can simultaneously flip to (5a,6a) or (5b,6b) by a crankshaft-like rotation and corner beads like 13 can only move to 13a. Colinear beads like 2 or 11 are not allowed to move. Moves are attempted only if they satisfy self-avoidance, tethering and boundary conditions.

where $\beta \equiv 1/k_B T$, with k_B the Boltzmann constant and T the absolute temperature. In the previous expression, the summation runs over all possible configurations (or self-avoiding paths), j, that start at the origin and end anywhere inside the sphere (for ST chains) or at an anchoring point (for DT chains). A bounded spherical domain might suggest, at first, a naive attempt to enumerate all the possible relevant walks on the lattice. While this task is possible (and carried out later to validate our algorithm) for small values of N ($N \leq 12$), a sampling of the configuration space becomes necessary as N becomes large [177,178]. Furthermore, the partition function cannot be directly estimated in the form of Eq. (5.3) [179]. Rather, Z is evaluated by differentiating Eq. (5.3) with respect to the inverse temperature β and reintegrating, arriving at [179]

$$Z(N, \beta, F) = \mathcal{N} \exp\left(- \int_0^{\beta} \langle E \rangle \, d\beta \right) = \mathcal{N} \, Q(N, \beta, F). \tag{5.4}$$

The average energy of the chain

$$\langle E \rangle = \frac{1}{Z} \sum_{(j)} E^{(j)} \exp(-\beta E^{(j)}), \tag{5.5}$$

can be estimated from a generated ensemble of chain configurations. The parameter $\mathcal{N}$ is an integration constant which needs to be determined as well.

With this result in hand, we describe in the following the procedures to compute the terms Q and $\mathcal{N}$ in the case of ST and DT confined chains subject to an external force.

5.2.1 The exponential term Q

For the moment, let us focus on the calculation of the exponential term $Q \equiv Q(N, \beta, F)$. As indicated in Eqs. (5.4) and (5.5), an estimation of the Q term is more complicated than evaluating the average energy $\langle E \rangle$. For a specified set of N, β and F, it involves calculating the average energy over a range of inverse temperatures from 0 up to β. In addition, the numerical integration requires a sufficiently small temperature step to obtain accurate results for Q. If the dependence of Q on different values of the external force is also of interest, one should expect these calculations to require substantial computing time.

Fortunately, the need for this wealth of simulation time can be mitigated if one combines the inverse temperature and the applied force into a single dimensionless

variable, the Péclet number

$$\mathrm{Pe} \equiv \beta F l. \tag{5.6}$$

Incorporating the definition of the Péclet number into Eqs. (5.1) and (5.2), the dimensionless energy, $U^{(j)}$, of configuration j can be written as

$$U^{(j)} = -\mathrm{Pe}\,\mathbf{u} \cdot \sum_{i=1}^{N} \mathbf{X}_i^{(j)} \tag{5.7}$$

where $\mathbf{X}_i = \mathbf{R}_i/l$ is a dimensionless position of bead i. In our definition of Eq. (5.7), recall that $\mathbf{R}_1 = \mathbf{0}$. By means of the above change of variables, one can readily verify that Q becomes

$$Q = \exp\left(-\int_0^{\mathrm{Pe}} \frac{\langle U \rangle}{\mathrm{Pe}}\, d\mathrm{Pe}\right). \tag{5.8}$$

In the latter expression, $\langle U \rangle \equiv \beta \langle E \rangle$ is a dimensionless average energy and $Q = Q(N, \mathrm{Pe})$ is now a function of the single parameter Pe. As a result, when evaluating Q, different values of the Péclet number can be regarded as different values of β *or* different values of the applied force F. This procedure has the advantage of limiting the number of necessary runs; runs for various inverse temperatures can actually count as runs for various forces, as changes in β amount to changes in Pe.

We now describe the general procedure to generate a Markovian set of configurations for the ST and DT chains with an energy given by Eq. (5.5). Several sampling methods can be invoked to build a statistically meaningful set of chain configurations [177, 178, 180, 181]. In the context of the present study, successive configurations of the chain are dynamically obtained through a well-defined motion of one or multiple beads [182]. At the end of each generation step, the boundary conditions and self-avoidance are automatically satisfied, making this class of methods very suitable to compute static properties of chains under various topological and geometrical constraints.

Consider then a self-avoiding chain with N beads placed inside a confining sphere, with end beads subject to ST and DT constraints. Starting with such a configuration, we sample the phase space using the kink-jump algorithm briefly outlined here [182–184]. During each Monte Carlo cycle, a bead is selected randomly and allowed to move according to the following set of rules. End beads can only move through a 90° rotation of the corresponding bond. When an inner bead is selected, if the two bonds connecting it to its neighboring beads are colinear, the move is rejected and a new bead is selected.

Otherwise, the bead attempts a kink or crankshaft move as described in Fig. 5.1(b). A Metropolis test dictates the acceptance or rejection of the moves, *i.e.* moves that lower the energy of the chain are accepted with a probability of 1 while others that increase the energy are accepted with a probability $P_{acc} = \exp\left[-(U^{new} - U^{old})\right]$. Naturally, moves are attempted only if they satisfy self-avoidance and boundary constraints. The bias introduced from the initial configuration is eliminated in an equilibration run under zero force. After this preliminary run, non-zero forces are introduced to generate successive configurations of the chains. Regardless of the success or failure of the corresponding moves, the resulting chain configurations will contribute to statistical averages with equal weights. Thus, if M configurations with energies ${U^{(j)}}_{1\leq j\leq M}$ are available, the dimensionless average energy $\langle U\rangle$ in Eq. (5.8) can be estimated by

$$\langle U\rangle \approx \frac{1}{M}\sum_{j=1}^{M} U^{(j)}. \tag{5.9}$$

The average energy for various Pe numbers is then numerically integrated to produce the exponential term Q.

5.2.2 The prefactor $\mathcal{N}$

We now turn our attention to evaluating the prefactor $\mathcal{N}$ in the expression of the partition function. With the change of variable introduced in Eq. (5.6), the physical significance of $\mathcal{N}$ becomes more apparent. According to Eqs. (5.4) and (5.8), $\mathcal{N}$ can be regarded as the total number of self-avoiding walks inside the sphere that satisfy the ST or DT constraints. In other words, it corresponds to the partition function at zero force or zero Péclet number. With this interpretation, we can decompose $\mathcal{N}$ into

$$\mathcal{N} = 6^{N-1} P_{\mathrm{RW}}\, r_{\mathrm{SAW/RW}} \tag{5.10}$$

where $r_{\mathrm{SAW/RW}}$ is the fraction of self-avoiding walks among all random walks (including self-crossing) that stay within the sphere and satisfy the tethering and boundary conditions. The factor P_{RW} represents the probability that a random walk, starting at the origin, ends anywhere inside the sphere (for a ST chain) or at a specified site (for a DT chain) in N steps without ever crossing the surface of the sphere. (The factor $6^{N-1}P_{\mathrm{RW}}$ is thus the total number of such walks.) We outline below how both terms

can be evaluated. For convenience, the subscripts SAW/RW and RW are dropped in what follows.

To evaluate r, we use a similar approach to the kink jump algorithm described above in the calculation of the exponential term Q. Starting with a chain inside the spherical domain and that satisfies the boundary conditions, multiple realizations of the chain are visited by successive Monte Carlo moves. However, contrary to the Q calculation case, we do not enforce the self-avoidance condition. Rather, beads along the chain are allowed to simultaneously occupy the same lattice site. More precisely, during each of the total number of Monte Carlo cycles, n_{cycle}, we distinguish four possible outcomes following a bead selection: i) When a colinear bead is selected, the Monte Carlo move is not attempted and a new bead is selected. This has the effect of reducing the effective number of cycles from n_{cycle} to some n_{eff}, ii) a number n_{BC} of these effective moves will be rejected because they do not satisfy the boundary conditions, iii) among the remaining $n_{\text{eff}} - n_{\text{BC}}$ cycles, a majority, n_{ns}, of cycles will result in non-self-avoiding chains with multiple beads occupying the same site and finally, iv) only a portion $n_s = n_{\text{eff}} - n_{\text{BC}} - n_{ns}$ will result in a self-avoiding random walk. Note that regardless of the completed move, the energy of the chain remains constant since r is evaluated at zero force. By monitoring the evolution of the previous numbers as a function of the total number of cycles, the fraction r can be estimated by

$$r \approx \frac{n_s}{n_s + n_{ns}}, \tag{5.11}$$

provided the total number of cycles is sufficiently large for the fraction to converge to a value independent of n_{cycle}.

We address now the calculation of the probability term P in Eq. (5.10). As mentioned above, this term represents the probability that a three-dimensional random walk starting at the origin ends somewhere inside the confining sphere without ever crossing the surface of the sphere. In principle, one can arrive at such probability by solving the Edwards equation for the probability density function $G \equiv G(\mathbf{R}, \mathbf{R}'; N)$ of a walk starting at $\mathbf{R}'$ and ending at $\mathbf{R}$ in N steps of length l,

$$\frac{\partial G}{\partial N} - \frac{l^2}{6}\nabla^2 G = \delta(N)\delta(\mathbf{R} - \mathbf{R}'), \tag{5.12}$$

where δ is the Dirac delta function. The latter is solved subject to the boundary

condition at the surface of the sphere

$$G(\mathbf{R}, \mathbf{R}'; N) = 0, \tag{5.13}$$

for every $N > 0$. For the case of a chain with a single tether, the probability P is computed by integrating the Green's function G over the entire spherical domain. When the angular degrees of freedom are averaged out, the remaining radial component in Eq. (5.12) leads to a simplified equation with an analytical solution for P [166, 167]. For the case of double tethered chains, the angular dependence of G and P is essential and is governed by the location of the second tether. Therefore, no analytical solution for Eq. (5.12) is attainable and one must content with a numerical integration to produce appropriate values of P. While seemingly attractive at first sight, this approach is only justified for long chains ($N \gg 1$). In fact, the continuous nature of the equation above does not take into account the finite extensibility of the chain for small N and predicts non-zero probabilities for tethering at distances larger than the chain contour length.

Consequently, a discrete description of P becomes imperative. Let $W(i, j, k, N)$ denote the probability that a walk will end at a point (i, j, k) in N steps and let p denote the transition probability of each step. (On a three-dimensional cubic lattice, $p = 1/6$ for an unbiased random walk.) The master equation for the growing process is

$$\begin{aligned} W(i,j,k,N) \;=\; & p \times [\, W(i-1,j,k,N-1) + W(i+1,j,k,N-1) + \\ & W(i,j-1,k,N-1) + W(i,j+1,k,N-1) + \\ & W(i,j,k-1,N-1) + W(i,j,k+1,N-1)\,]. \end{aligned} \tag{5.14}$$

The boundary condition can be expressed now by

$$W(i, j, k, N) = 0, \tag{5.15}$$

for every $N > 0$ when the point (i, j, k) falls outside the domain of interest. The initial condition anchoring the first bead at the origin implies

$$W(i, j, k, 1) = \delta_{i,0}\delta_{j,0}\delta_{k,0}\,, \tag{5.16}$$

where $\delta_{i,j}$ is the Kronecker delta function. For finite values of i, j and k in the cube embedding the sphere, the previous equation can be solved iteratively to compute

$$P^{(\mathrm{ST})} = \sum_{i,j,k} W(i, j, k, N)\,, \tag{5.17}$$

Tether	ρ	$\log Q$	λ	μ	a	b
Single Tether						
	3	$N\,\mathrm{Pe}\,\rho$	0.65	0.86	N/A	
	4		0.71	0.91		
Double Tether						
	3	$N\,\mathrm{Pe}\,\rho$	0.65	N/A	25.31	0.16
	4		0.71		61.89	0.11

Table 5.1: Summary of the scaling laws for the various terms in the normalized partition function $Z^* = Q\,r\,P$ as a function of the chain size N, the dimensionless force Pe and the radius ρ. For both ST and DT chains, the exponential term Q depends on the single reduced parameter $N\,\mathrm{Pe}\,\rho$. For large N, the fraction r of SAW to RW scales exponentially with N such that $r \sim \lambda^N$. Finally, the probability P that a RW ends inside the sphere is such that $P^{(\mathrm{ST})} \sim \mu^N$ for ST chains and $P^{(\mathrm{DT})} \sim \exp\left(-a/N - b\,N\right)$ for DT chains.

for a single-tethered chain, and

$$P^{(\mathrm{DT})} = W(i_0, j_0, k_0, N), \tag{5.18}$$

for a chain with a second tether located at (i_0, j_0, k_0).

5.3 Results

We present in this section the results obtained by implementing the methods described above. The chains used in our simulations contain up to 45 beads for a confining sphere with a radius $\rho = 3$ and up to 65 beads for a sphere with $\rho = 4$. In both cases, the center of the sphere is located at $(\rho, 0, 0)$, such that the first bead is at the origin and the last bead is on the surface of the sphere at $(2\rho, 0, 0)$ for a DT chain. The external force acts favorably along $\mathbf{u} = \mathbf{i}_x$ with a dimensionless magnitude $0 \leq \mathrm{Pe} \leq 2$. The scaling laws for Q, r and P are summarized in Table 5.1.

Continuing in the same order as the previous section, we start with the exponential term Q. The dependence of the exponential term Q on the chain size N and the dimensionless force Pe is presented in Fig. 5.2. There, $\log Q$ is plotted as a function

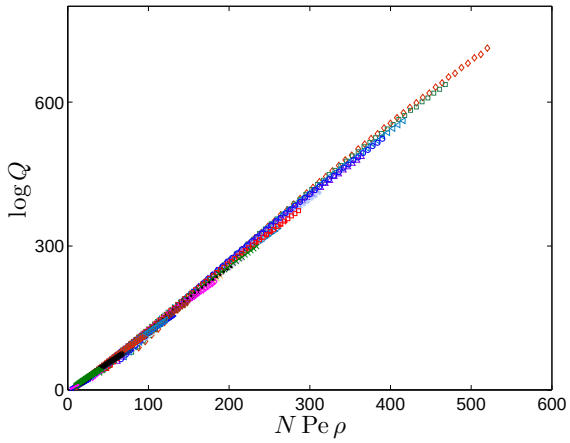

Figure 5.2: Data of $\log Q$ for various ST and DT chain sizes (N), Péclet numbers (Pe) and sphere radii (ρ), collapse onto a single linear curve when plotted as a function of the reduced parameter $N\,\mathrm{Pe}\,\rho$.

of the reduced parameter $N\,\mathrm{Pe}\,\rho$ for different values of N, Pe and ρ. As can be seen from the figure, all values of $\log Q$ merge onto a single linear curve as a function of the parameter $N\,\mathrm{Pe}\,\rho$. One thus finds

$$\log Q \sim N\,\mathrm{Pe}\,\rho, \tag{5.19}$$

for both ST and DT chains, similar to the potential energy of a spherical ball of radius ρ and subject to a total force NPe.

The exponential behavior of Q can be retrieved from Eq. (5.8). If the average position of a bead i inside the sphere is of order ρ , *i.e.* $\langle \mathbf{u} \cdot \mathbf{X}_i \rangle \sim \rho$, then the average energy of the chain in Eq. (5.7) can be approximated by $\langle U \rangle \sim -N\mathrm{Pe}\rho$. Inserting this expression in Eq. (5.8) for Q readily leads to the scaling relationship in Eq. (5.19).

The strictly monotonic behavior of the latter scaling implies that Q never reaches a maximum independent of N. However, when one considers the partition function Z, one might expect that entropic effects due to confinement and self-avoidance should set an upper limit on the number of beads a sphere can allow. The entropic consequence of increasing N is embodied in the prefactor $\mathcal{N}$, which we discuss now.

As described in Sec. 5.2, the probability (or total number) of a ST or DT self-avoiding walk can be decomposed into a product of a ratio r and a probability P,

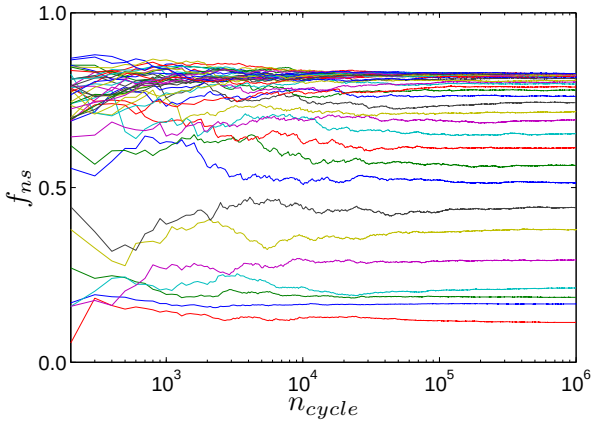

Figure 5.3: Plot of the fraction f_{ns} of cycles resulting in self-crossing walks as a function of the total number of cycles for various sizes of a ST chain inside a sphere of radius $\rho = 3$. The convergence of f_{ns} is established for all sizes after $\approx 10^5$ cycles. A similar behavior for DT chains was obtained.

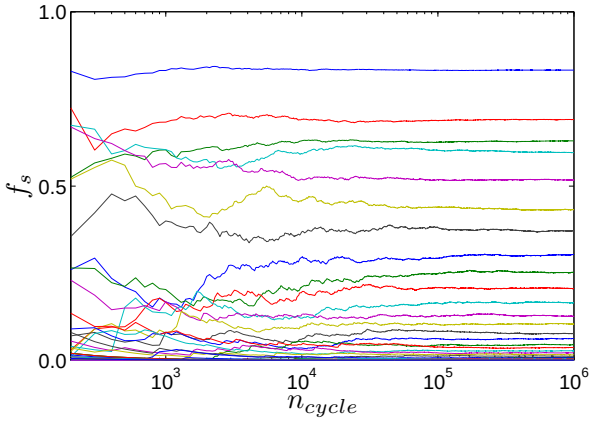

Figure 5.4: Same as Fig. 5.3 for the fraction f_s of cycles resulting in self-avoiding walks.

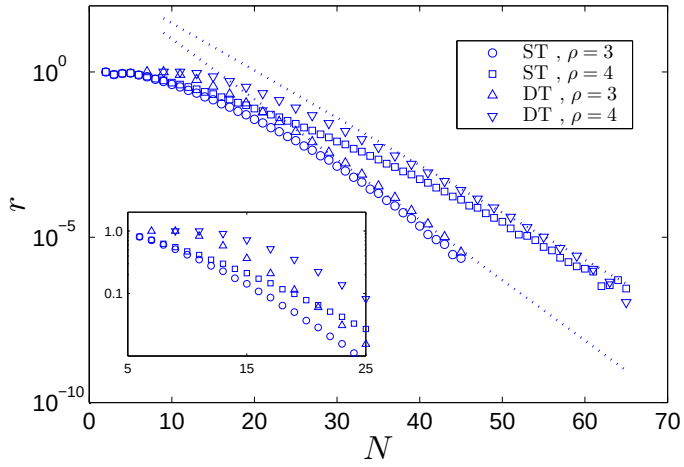

Figure 5.5: Plots of the fraction r of SAW to RW inside a sphere as a function of chain size N for ST and DT chains. The dotted lines represent a linear convergence of r at large N. The slopes are $\log \lambda_3 \approx -0.42$ for $\rho = 3$ and $\log \lambda_4 \approx -0.33$ for $\rho = 4$. The inset shows r for small values of N, indicating that $r^{(\mathrm{ST})} < r^{(\mathrm{DT})}$.

thus simplifying the procedure to calculate $\mathcal{N}$. The ratio r is estimated by counting the number of Monte Carlo cycles that result in self-avoiding chains. The accuracy of this approach, however, relies on the condition that the fractions $f_s = n_s/n_{cycle}$ and $f_{ns} = n_{ns}/n_{cycle}$ do not depend on the total number of cycles. We found that f_s and f_{ns} converge after $n_{cycle} \approx 10^6$. To illustrate the convergence of the two fractions, the plots in Fig. 5.3 and Fig.5.4 represent the evolution of f_{ns} and f_s for ST chains as a function of n_{cycle} up to 10^6. As shown in the two figures, the convergence of f_{ns} and f_s is demonstrated for all chain sizes. In all of our simulation runs, we used a total number of cycles $n_{cycle} = 10^8$, resulting in $n_s + n_{ns} \approx 0.8\, n_{cycle}$ runs to build the ensemble used to compute r.

As can be seen from Fig. 5.5, the ratio of SAW to RW decreases exponentially with increasing size of the chain for large N. For both ST and DT chains

$$r \sim {\lambda_\rho}^N. \tag{5.20}$$

where $\lambda_3 \approx 0.65$ and $\lambda_4 \approx 0.71$ for $\rho = 3$ and 4 respectively.

This scaling for relatively small spherical cavities is approaching the result from Hammersley and Morton on the attrition of the number of non-confined, end-free self-avoiding walks [185]. In the latter, $r \sim {\lambda_\infty}^N$ and $\lambda_\infty \approx 0.78$ for a cubic lattice [178, 185]. In the case of long ST or DT confined chains, the total number of configurations (both self-avoiding and self-crossing) is reduced compared to a free non-confined chain. Moreover, this reduction is accompanied by an even greater decrease in the number of self avoiding chains; self-avoiding chains are more affected by confinement than their self-crossing counterparts. This attrition in the number of walks is higher with increasing degree of confinement and expectedly, $\lambda_3 < \lambda_4 < \lambda_\infty$.

The difference between ST and DT chains becomes more apparent for small values of the chain size N. In this case, the inset in Fig. 5.5 shows that the ratio r is smaller for ST chains than DT chains, $r^{(\mathrm{ST})} < r^{(\mathrm{DT})}$, indicating a greater attrition for ST chains. When a small chain is attached at both ends to the surface of the sphere, it is in a nearly stretched state that favors self-avoiding configurations. On the other hand, contrary to a DT chain, a ST chain enjoys more freedom to explore more self-crossing configurations, thus leading to more rejection or attrition in the number of relevant self-avoiding walks.

We consider next the behavior of the probability P for a random walk to start at the origin and end somewhere inside a confining sphere. For a ST chain, Fig. 5.6(a) shows that P is a strictly monotonic function of the chain length. Starting at the origin, the walk is less likely to stay within the spherical domain as it proceeds step after step. For large values of N, the exponential decrease of P is in accordance with the result obtained by Park *et al.* [167] and Muthukumar [166] using the continuous Edward's model to represent the chain. Solving Eq. (5.12) subject to the appropriate boundary conditions, they showed that (in our dimensionless notation)

$$P^{(\mathrm{Ed})} \sim \sum_{k=1}^{\infty} \exp\left(-\frac{\pi^2 N}{6\rho^2} k^2\right). \tag{5.21}$$

For $\rho \sim \mathcal{O}(1)$, the above series can be truncated at first order, so that

$$P^{(\mathrm{Ed})} \sim \exp\left(-\frac{\pi^2 N}{6\rho^2}\right) = {\mu_\rho}^N \tag{5.22}$$

Using the values of $\rho = 3$ and 4, we obtain

$$\mu_3^{(\mathrm{Ed})} \approx 0.83 \quad \text{and} \quad \mu_4^{(\mathrm{Ed})} \approx 0.90\,, \tag{5.23}$$

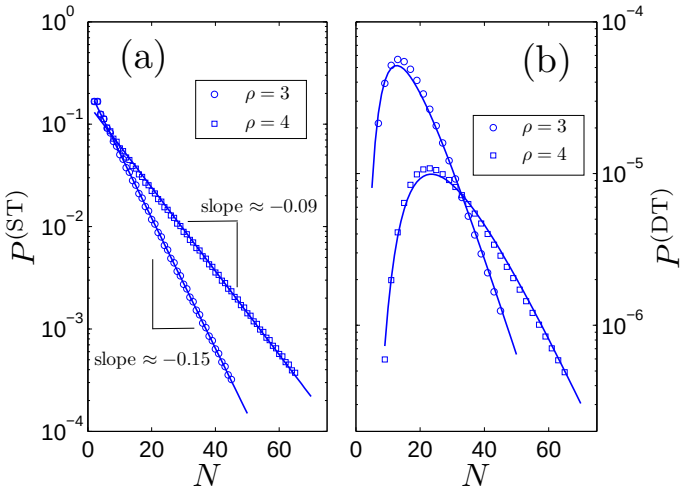

Figure 5.6: Plots of the probability P that a random walk with one end attached (a) or two ends attached (b) stay within a confining sphere as a function of the walk size N. For chains with a single tether, $\log P$ is linear with N with slopes $\log \mu_3^{(\mathrm{ST})} \approx -0.15$ for $\rho = 3$ and $\log \mu_4^{(\mathrm{ST})} \approx -0.09$. Note the 3-order magnitude difference in P between ST and DT chains, indicative of the stronger attrition for DT chains.

in very good agreement with the values

$$\mu_3^{(\mathrm{ST})} \approx 0.86 \quad \text{and} \quad \mu_4^{(\mathrm{ST})} \approx 0.91, \tag{5.24}$$

obtained using Eq. (5.14).

If we consider a DT chain, the probability is no longer a monotonic function of N. Rather, P reaches a maximum at some critical size N^*, as shown in Fig. 5.6(b). For values of $N < N^*$, the two anchored ends constrain the chain to a nearly stretched regime, thus limiting the volume of phase-space the chain can explore. The addition of segments allows the flexible chain to "breathe", thereby increasing the number of allowable configurations and the concomitant probability. On the other hand, when the chain size exceeds the critical size $N > N^*$, confinement effects dominate the elasticity of the chain and the resulting probability P decreases. Using our discrete model, we found that $N^* = 11$ for $\rho = 3$ and $N^* = 21$ for $\rho = 4$. When the elasticity and

confinement effects are combined, we obtain the scaling law

$$P^{(\mathrm{DT})} \sim \exp\left[-\frac{a_\rho}{N} - b_\rho\, N\right], \tag{5.25}$$

where a_ρ and b_ρ are two positive coefficients that govern the magnitude of the elasticity and confinement effects, respectively. The numerical values of a_ρ and b_ρ for $\rho = 3$ and 4 are reported in Table 5.1.

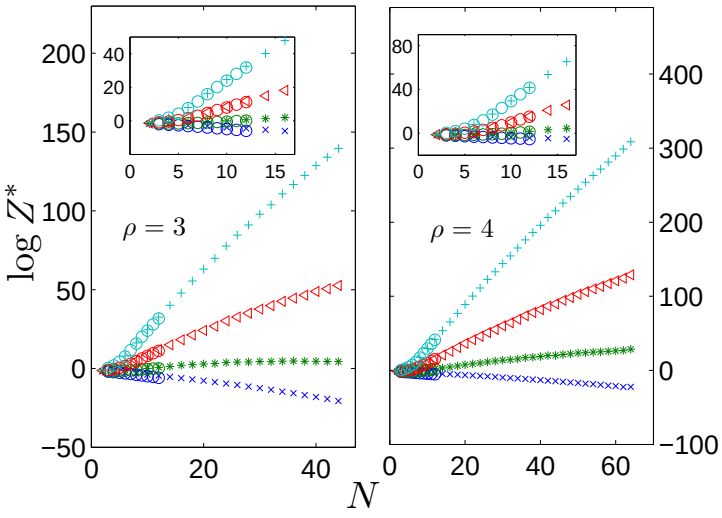

Figure 5.7: Monte Carlo results for the normalized partition function Z^* as a function of size N in the case of ST chains inside a sphere of radius ρ. The different plots correspond to different values of the dimensionless force Pe: Pe = 0 ($\times$), Pe = 0.2 ($*$), Pe = 0.5 ($\triangleleft$) and Pe = 1 ($+$). The circles (o) represent direct calculations of Z^* using an exact enumeration of the allowed paths on the lattice. The agreement between the exact enumeration method and the Monte Carlo calculation is shown in the inset for small values of N.

Using the previous results for Q, r and P, an estimation of the normalized partition function $Z^* = Z/6^{N-1} = Q \times r \times P$ is readily available. For small values $N \leq 12$, a direct enumeration of the ST and DT paths inside a sphere is possible, rendering the exact evaluation of Z^* immediate through Eq. (5.3). As illustrated in Figs. 5.7 and 5.8, the agreement between our Monte Carlo calculations and the exact evaluation is

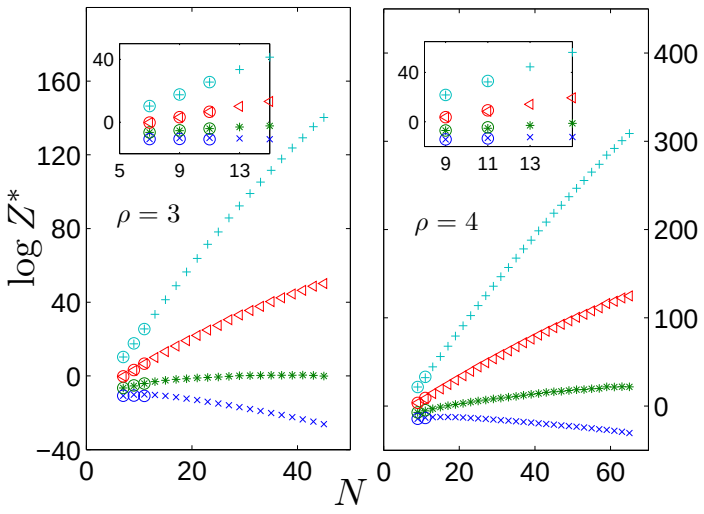

Figure 5.8: Same as Fig. 5.7 for DT chains.

quite remarkable, thereby validating our global approach. Furthermore, as the size of the chain increases, the behavior of the partition function depends on the applied force. In the absence of a force, Pe = 0, entropic effects due to self-avoidance and confinement result in a decreasing partition function. For higher values of the Péclet number, a balance between the enthalpic contribution to Z^* and the entropic penalty sets a maximum value on the partition function at some size $N_{\max}$. For values of $N < N_{\max}$, the favorable bias from the applied force dominates the effect of the constraints and Z^* increases with N. When N exceeds $N_{\max}$, the concave partition function begins to decrease. This transition can be seen in Figs. 5.7 and 5.8 for $\rho = 3$ and Pe = 0.2 around $N = 40$. The transition should also occur in the other cases, as suggested by the shape of the corresponding curves. However, as Pe and/or ρ are increased, the transition to a confinement-dominated regime occurs for higher values of N, falling outside the range used here. Finally, the difference between ST and DT chains is limited to weak forces, in which case, the second tether reduces the total number of configurations and the resulting partition function. As the force is increased, the difference between ST chains and DT chains is blurred since the biasing force will greatly favor configurations that end near the second tether.

5.4 Conclusion

In summary, we have presented a Monte Carlo procedure to compute the partition function of a linear chain, subject to several constraints. The self-avoiding chain is anchored to the inner surface of a confining sphere and subject to an external force. Compared to a free self-avoiding chain, the introduction of confinement and tethering constraints translates into a dramatic increase in the attrition of the number of successful walks. Our method entails a decomposition of the partition function into several independent terms that can be evaluated separately through Markovian samples of chain configurations and the numerical solution of a master equation. For short chains, our method does an excellent job of capturing exact results available from a direct enumeration of the total set of allowed paths. In this chapter, we restricted our discussion to forces acting along the same direction $\mathbf{u} = \mathbf{i}_x$, albeit with different magnitudes. Likewise, in the case of chains with both ends tethered to the surface of the sphere, only diametrically opposite tethers were considered. However, the general approach can be easily extended to different directions of the applied force as well as different locations of the second point on the surface, or even different confining geometries. To illustrate the variety of situations that arise, one can consider, for example, a "turning" chain where the second tether is located on the surface, 90° away from the first tether. Further cases can be explored by varying the direction of the external force. These include neutral or unfavorable directions relative to the "natural" direction of the chain. For example, in the case of DT chains, a natural direction can be defined by the two tether points while any neutral force has a direction perpendicular to the natural direction. On the other hand, ST chains we considered here will inherit an unfavorable bias when the direction of the applied force is reversed while maintaining the zero energy reference at the tether point. In fact, the different situations described above arise when investigating the transport of a long chain inside an array of spheres under the action of an external electric field. A knowledge of the corresponding free energies is therefore essential to modeling such motion. The study of the dynamics of chains inside arrays of spheres is the subject of the next chapter.

Chapter 6

DNA transport in an array of spheres: A new Lakes-Straits model

The chapter is based on the manuscript:
N. Laachi, and K. D. Dorfman
submitted to *J. Chem. Phys.*, (July 01, 2010).

6.1 Introduction

Taking advantage of the increased knowledge and control of the physics of fluids at the micro- and nano-scale, several microfluidic devices have been proposed in the past two decades for the separation of DNA molecules based on their size [124, 186, 187]. These devices aim to challenge the dominance of gels as a separation medium and address the limitations of conventional gel electrophoresis [9,79]. Different geometries have emerged, including arrays of cylindrical posts [23–25], introduced in Chapter 4, and planar wells connected by narrow slits [37], considered in Chapter 3. While these devices undeniably outperform gel electrophoresis for long DNA separations, their daily use is still limited by fabrication constraints.

The self-assembly of colloidal particles offers an ergonomic alternative to microfabrication for DNA separation devices. By inducing the self-assembly of the particles, colloidal crystals (opals) can be constructed and used as a sieving matrix to sort DNA molecules [48, 49, 188–190]. The inverse structures consisting of periodic arrays of spheres, known as inverse opals, have also been proposed for DNA electrophoresis [41,50,51]. These structures can be fabricated from the colloidal crystal template by first depositing a continuous phase in the gaps of the crystal and then removing the colloids by chemical etching [52]. While the vast majority of applications for inverse opals are in photonics [52], if the continuous phase is a polymer [51] or photoresist [41], then the material is compatible with DNA electrophoresis. For example, Nykypanchuk *et al.* observed the Brownian motion of various DNA molecules inside a two-dimensional array of cavities with a diameter $\approx 0.9\mu$m, connected by circular holes with a diameter of $\approx 0.1\mu$m. They noted that, while short molecules tend to localize inside a single cavity and move by overcoming entropic barriers [32,36] at the hole entrance, larger DNA can span three to four cavities and move by extension or contraction from the end cavities. They found that, in this case, the linear chain rarely forms new configurations through side excursions, or "hernias", out of the inner cavities [50]. The motion of large DNA was attributed to an onset of the cross-over from entropic barriers to reptation [50]. Likewise, Shiu *et al.* studied experimentally the motion of DNA inside a face-centered cubic array of spheres with a diameter of several hundreds of nanometers, connected by narrow pores with a diameter of tens of nanometers [41]. They showed that M13mp18,

a short DNA molecule with a size of 7.25 kilobase pairs (kbp), coils up inside a single sphere, while λ-DNA, a larger molecule with a size of 48.5 kbp, occupies multiple spheres instead and moves through the constrictions like an unraveling thread [41]. They also found that the mobilities of the two molecules depend on their respective size and the applied electric field. These experimental results thus establish not only that a DNA molecule can migrate inside periodic arrays of spheres, but that a separation based on their molecular weight is also possible.

We focus on a numerical approach to understand the electrophoretic motion of long DNA inside a cubic array of hollow spheres with a diameter of several hundred nanometers, connected by narrow openings with a diameter of tens of nanometers. This situation is typified in Fig. 6.1. These length scales guarantee that a long DNA molecule cannot entirely coil up inside a single cavity; entropic trapping-like motion is ruled out [32, 36]. Rather, the chain is distributed over multiple cavities, suggesting a type of motion similar to that of a reptating chain in a gel [9, 50] where the length of the tube embedding the chain constantly fluctuates. The analogy can be extended even further and one can regard these nano-structures as ideal, highly ordered gels. During reptation in a gel, the renewal of the tube is governed by a continuous order parameter, $\langle \cos \theta \rangle$, that defines the average orientation of the leading/trailing blob with respect to the electric field [77]. In contrast, the ordered geometry of arrays of cavities offers only limited connectivity between the confining spheres. As a consequence, the order parameter during reptation in the array is quantized, with only a discrete, finite set of possible orientations.

In order to characterize the electrophoretic migration of DNA molecules, one might consider, at first, a straightforward Brownian dynamics (BD) simulation. Indeed, the problem at hand entails well-defined obstacles and boundary conditions and, therefore, may appear to be well suited for a Langevin-type description. Despite their attractive features, the geometrical and molecular length scales make bead-spring/bead-rod models inefficient in the present case; the chains we consider are comprised of several hundred Kuhn segments while threading through narrow holes with a diameter of $\approx$ 10 nm. Resolving of the dynamics inside the confining spheres, near the walls, and inside the narrow pores will necessitate very long simulation times. A treatment with a higher level of coarse-graining is therefore necessary.

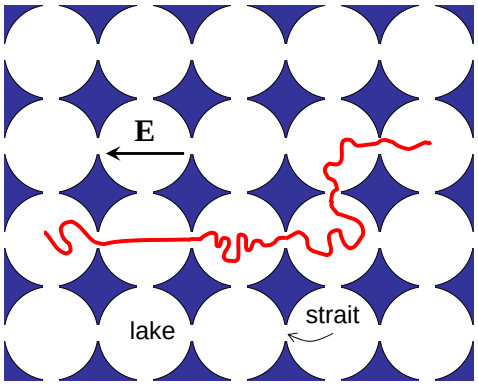

Figure 6.1: Example of a chain configuration. The chain occupies 6 lakes. Each lake is defined by its position, the number of segments it contains and the orientation variable c. Note that in this example, lake 2 contains exactly $n_{\min}$ segments, and a sliding move is necessary for a forward jump to lake 3.

In an attempt to bridge the gap between BD simulations and direct analytical approaches, such as biased reptation (BRM) or biased reptation with fluctuations (BRF), several simulation-based reptation models have been proposed [9]. Among these, tube models with internal modes, such as the lakes-straits model [154], the 1-D Langevin approach [191] or the biased repton model [192], include length fluctuations of the tube embedding the chain which proved essential to understanding gel electrophoresis at relatively high fields [9]. These three models rely on a coarse-graining of the chain whereby the details of the dynamics inside a single pore are simply ignored.

Given the geometry of interest, the lakes-straits model appears to be the most suitable choice in the present situation. The coarse-grained lakes-straits (LS) model was first introduced by Zimm to study the mobility minimum during field inversion gel electrophoresis [154, 155]. In this model, the chain occupies a fluctuating number of lakes connected by narrow straits, through which the chain is pulled in a single file. As a consequence, the number of degrees of freedom necessary to describe the dynamics of the chain is greatly reduced compared to a BD approach. In the original model, the random structure of the gel is manifest when new lakes are created; the directions of new lakes are randomly selected, though biased by the external electric field [155]. Furthermore, the entropic force in the LS model only accounts for the chain elasticity and

is therefore purely Gaussian. In contrast, in highly ordered geometries, the confining spheres occupy specific sites on a three-dimensional lattice. Likewise, the free energy can no longer be Gaussian as one needs to include self-exclusion, as well as confinement effects from the solid walls of the structure.

Our goal is, therefore, to modify the lakes-straits model to account for the limited connectivity and confinement inside the array of spheres. In our attempt to describe the transport of DNA, we prefer the terminology of Zimm and we shall refer to the confining spheres as lakes, and the narrow openings as straits. The rest of this chapter is organized as follows: In Sec. 6.2, after some general considerations, we present a detailed description of the algorithm we used to simulate the motion of the chain. A discussion of the results we obtained is the subject of Sec. 6.3, while concluding remarks and some perspectives are presented in Sec. 6.4.

6.2 Description of the algorithm

6.2.1 General considerations

The dynamics of long DNA molecules inside an array of spheres is closely related to that of a translocating chain through a membrane [87, 88]. In the latter case, the chain is pulled as a single thread through a narrow hole in a membrane. In the present case, however, the chain occupies several spheres at the same time and motion occurs through mass transfer between the spheres via the narrow openings. In essence, one can identify the migration of DNA as a multi-translocation event during which multiple segmental transfers between non-empty spheres take place.

Let us then consider a long DNA chain composed of N Kuhn segments, each of size b such that the total contour length of the chain $\mathcal{L}_c = Nb$. The chain is subject to an external electric field $\mathbf{E} = -\mathbf{i}_x E$ and is initially placed into a cubic array of lakes of radius R connected through narrow straits of diameter d. On one hand, the opening diameter is smaller than the Kuhn length b so that the penalty associated with bending the chain at the strait entrance results in a single file motion. On the other hand, the radius R is greater than the Kuhn length b, but much smaller than the contour length $\mathcal{L}_c$. For the purpose of this study, we take the diameter of the lake to be an integer multiple of the Kuhn length, such that $2R = n_{\min} b$.

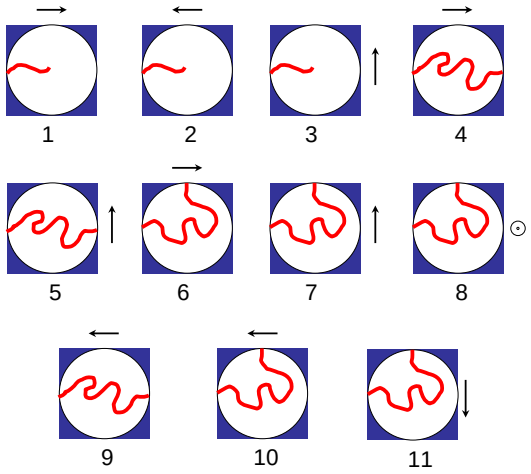

Figure 6.2: Schematic of the 11 possible orientations of a subchain inside a given lake, relative to the applied electric field (represented by the arrow). For $c = 8$, the dot points in the direction perpendicular to the current plane. The partial free energy of a subchain is calculated for each conformation, with a reference at the entry point (considered here at the left-most point). All orientations fall within one of the situations above. Since the electric field acts along a constant direction $-\mathbf{i}_x$, the lakes and subchains depicted above are appropriately rotated so that all entry and exit points for all situations are projected onto the same plane.

Let A_i denote the amount of mass, or subchain, inside lake i with $1 \leq i \leq m$ and $m \equiv m(t)$ the total number of lakes simultaneously spanned by the chain at each time t. For each portion A_i, one can specify three attributes: i) the vector position $\mathbf{x}_i$ of the center of the lake containing A_i, ii) the number of segments n_i composing A_i, and finally, iii) a number $1 \leq c_i \leq 11$, describing the relative direction of the electric field with respect to the first and last segments of A_i. The definition of c_i is made explicit in Fig. 6.2. The finite extensibility of the chain results in the constraint $n_i \geq n_{\min}$ for inner lakes, while $n_i \geq 1$ for end lakes. The set of these attributes for each portion, $\{\mathbf{x}_i, n_i, c_i\}_{1 \leq i \leq m}$, defines the configuration of the chain, such as the one depicted in Fig. 6.1.

As the chain is pulled, the number of lakes it occupies will vary and the configuration of the chain will change accordingly. The conservation of the total number of segments

implies that

$$N = \sum_{i=1}^{m} n_i, \tag{6.1}$$

and, to a first approximation, one can define the position of the center of mass of the chain by [154, 155]

$$\mathbf{x}_{cm} \equiv \frac{\sum_{i=1}^{m} n_i \, \mathbf{x}_i}{\sum_{i=1}^{m} n_i} = \frac{1}{N} \sum_{i=1}^{m} n_i \, \mathbf{x}_i, \tag{6.2}$$

although the center of mass of each portion A_i does not necessarily coincide with the center of lake i. Similarly, we introduce the total free energy of the chain, which, as an extensive property, can be decomposed into the sum

$$\mathcal{F} = \sum_{i=1}^{m} \mathcal{F}_i, \tag{6.3}$$

where $\mathcal{F}_i \equiv \mathcal{F}_i(\mathbf{x}_i, n_i, c_i)$ is the partial free energy of the subchain A_i.

6.2.2 From dynamics to kinetics

In the LS model, the time-dependent trajectories of the chain center of mass in Eq. (6.2) are obtained from appropriate mass balances at each strait [155]. The corresponding set of differential equations is then numerically integrated with an adaptive time step to produce successive updates of chain configurations. We propose here a more efficient method to generate new configurations that does not entail fractional exchanges of segments, while rigorously accounting for confinement and self-exclusion effects inside the lakes. In our method, we model the segmental exchanges between various lakes as a set of discrete, stochastic "chemical" reactions, that can be simulated to update the configuration of the chain. In particular, the continuous motion of the chain is decomposed into a set of three discrete, elementary moves, or reactions: translocation, sliding and nucleation.

The translocation between two adjacent lakes i and $i+1$ can be viewed as two reactions where

$$A_i \rightarrow A_{i+1}, \tag{6.4}$$

is the forward reaction indicating a transfer of one segment from lake i to lake $i+1$, and

$$A_{i+1} \rightarrow A_i, \tag{6.5}$$

for the backward motion or reverse reaction. For a chain occupying m lakes, one can introduce $2(m-1)$ translocation moves, or first-neighbor reactions, to describe transfers of mass between neighboring lakes. Translocations are the most common type of moves and will, accordingly, account for the majority of reactions in our simulations. When lake i contains $n_{\min}$ segments, a forward (or backward) transfer to lake $i+1$ (or $i-1$) is impossible because of the finite extensibility of the chain. In this case, the forward (or backward) reaction is replaced by a sliding reaction that enables the chain to move like a taut string. For example, if $i-j$ is the nearest lake to lake i with sufficient mass, *i.e.*, if $n_{i-j} > n_{\min}$ and $n_k = n_{\min}$ for $i-j+1 \leq k \leq i-1$, then the translocation move in Eq. (6.4) is substituted with

$$A_{i-j} + A_{i-j+1} + \ldots + A_i \rightarrow A_{i-j+1} + \ldots + A_i + A_{i+1}. \tag{6.6}$$

A similar sliding reaction is introduced when the backward move in Eq. (6.5) is forbidden, *i.e.*, $n_{i+1} = n_{\min}$. The introduction of a sliding mode is also necessary to release the chain from long-lived configurations. If one only allows first-neighbor moves, the escape of a chain trapped in a U-shaped configuration can only be completed through successive single-segment moves, against the electric force direction, from one end of the chain to the other [75, 138]. Sliding moves thus guarantee that the escape of the chain occurs in physically realistic time intervals [75].

In addition to the previous moves, two nucleation reactions are introduced to ensure the overall translation of the chain in space. When an end lake contains at least $n_{\min}+1$ segments, the overflow of mass into a new lake can be represented by

$$A_1 \rightarrow A_0^* \quad \text{or} \quad A_m \rightarrow A_{m+1}^*, \tag{6.7}$$

where A_0^* and A_{m+1}^* are newly created lakes containing one segment.

This set of three elementary modes of motion proves sufficient for a physical description of the motion of the chain. Additional refinements, or modes of transport, can be introduced and one can build an entire hierarchy of reactions to describe the jump process of segments between lakes. Each level of this hierarchy introduces an additional discrete mode of dynamics involving transfers between a higher number of lakes. Exchange between adjacent lakes is the lowest level of this hierarchy, exemplified by the two first-neighbor reactions (6.4) and (6.5). Examples of a higher mode include

$$A_i + A_{i+2} \rightarrow 2\,A_{i+1}, \tag{6.8}$$

for a second-neighbor reaction where a segment from lake i and a segment from lake $i+2$ are simultaneously pulled to increase the population of lake $i+1$. The unnecessary profusion of reactions thus generated can be prevented if one restricts the model to the most kinetically relevant reactions, *i.e.*, translocation, sliding and nucleation moves. For a chain occupying m lakes, this assumption leads to a total of $\mathcal{M} = 2m$ reactions that must be simulated.

We have then at hand a total of $\mathcal{M}$ stochastic reactions, describing transport between lakes and overflow of the end lakes. Given a configuration of the chain, each reaction ν occurs with a rate, or propensity, a_ν, the calculation of which is the subject of the next section. This representation has the advantage of reducing the dynamics of the chain to a stochastic kinetics problem. In other words, the motion of the chain in a particular configuration can be characterized from the behavior of a system of discrete, stochastic reactions occurring at different rates.

Several algorithms, of various degrees of complexity and sophistication, have been proposed to simulate the behavior of a system of discrete, stochastic reactions [see the review by Kaznessis [193] and the references therein]. Since the propensities depend on time only through the changing chain configuration, Gillespie's first-reaction method was our method of choice [194]. In the following, we briefly outline the iteration process we implemented to simulate the $\mathcal{M}$ stochastic reactions. Our description is straightforward, akin to the first-reaction method itself. Ample details on the method can however be found elsewhere [194].

Starting with an initial configuration of the chain, the probability that a reaction ν occurs during the interval between t and $t + dt$, if no other reaction occurred first, is given by [194]

$$P_\nu(t)\, dt = \exp(-a_\nu t) \cdot a_\nu\, dt \tag{6.9}$$

where the propensity a_ν does not depend explicitly on time. In the first-reaction method, the tentative time τ_ν for the occurrence of reaction ν is calculated at each step of the simulation from the probability density $P_\nu(t)$ as [194]

$$\tau_\nu = -\frac{\log \mathrm{URN}}{a_\nu}, \tag{6.10}$$

where URN is a uniform random number between 0 and 1, and the time counter increment is chosen as [194]

$$\tau_\mu = \min_\nu \tau_\nu. \tag{6.11}$$

At the end of the step of duration τ_μ, the corresponding reaction μ is enforced and the configuration of the chain is updated. The new position of the center of mass is then calculated using Eq. (6.2). The chain configuration and the corresponding propensities are updated and a new step is attempted until the chain has traveled far enough for accurate statistical averaging.

6.2.3 Propensities

We turn our attention now to the calculation of the configuration-dependent propensities a_ν, requisite for the kinetic Monte Carlo simulation. As described above, translocation moves entail exchange of segments between adjacent lakes. The basic idea is to treat the two subchains translocating through the strait as a single particle diffusing between two absorbing walls separated by $2b$ [145], similar to the first-passage process introduced in Chapter 2. The occurrence of either reaction, or translocation in either direction, is equivalent to the absorption of the particle by either wall.

If $\mathbf{u}_i$ is the unit vector joining lake i to lake $i+1$, *i.e.*

$$\mathbf{u}_i = \frac{\mathbf{x}_{i+1} - \mathbf{x}_i}{\parallel \mathbf{x}_{i+1} - \mathbf{x}_i \parallel}, \tag{6.12}$$

then, the total force, $\mathbf{F}_i$, acting on the particle describing subchains A_i and A_{i+1} can be written as [145]

$$\mathbf{F}_i = \frac{1}{b}\left(\frac{\partial \mathcal{F}_i}{\partial n_i} - \frac{\partial \mathcal{F}_{i+1}}{\partial n_{i+1}}\right)\mathbf{u}_i \tag{6.13}$$

where the calculation of the partial free energy $\mathcal{F}_i$ will be discussed below. Let $P_\pm$ and $\Delta t_\pm$ respectively denote the total probability and average time of the translocation in the $\pm\mathbf{u}_i$ direction. Using first-passage time statistics of a particle initially located between two absorbing walls [101,195], the total probability is calculated from Eq. (2.47) as

$$P_\pm = \frac{1}{1 + e^{\mp\epsilon}}, \tag{6.14}$$

while the mean first-passage time is computed from Eq. (2.48) as

$$\Delta t_\pm = \frac{\tanh \epsilon/2}{\epsilon/2}\,\tau_b. \tag{6.15}$$

In the equations above, the dimensionless force ϵ and the particle diffusion time τ_b are

$$\epsilon \equiv \frac{F\, b}{k_B T} \tag{6.16}$$

and

$$\tau_b = \frac{b^2}{2D_0}, \tag{6.17}$$

where k_B is Boltzmann constant, T is the absolute temperature, D_0 is the diffusion coefficient of the particle [86,145] and $F = \mathbf{F}_i \cdot \mathbf{u}_i$ is the projection of the force introducing a drift in the Brownian motion of the particle. If one approximates the conditional probability that the translocation occurs in an interval of length dt by $dt/\Delta t_\pm$, then the mean probability, $a_\pm dt$, for the translocation reactions can be estimated by [75]

$$a_\pm dt \equiv P_\pm \times \frac{dt}{\Delta t_\pm}. \tag{6.18}$$

yielding the propensities

$$a_\pm = \frac{P_\pm}{\Delta t_\pm}. \tag{6.19}$$

of the forward reaction (+) in Eq. (6.4) and the backward reaction (–) in Eq. (6.5).

These expressions for the propensities can no longer hold when sliding reactions are necessary. For simplicity, we consider the case where lake i contains $n_{\min}$ segments so that the forward reaction in Eq. (6.4) must be replaced by the sliding reaction in Eq. (6.6). The case where lake $i+1$ contains $n_{\min}$ segments and a backward sliding reaction is treated in a similar way.

The goal is then to calculate the propensities for the pair of backward reactions in Eq. (6.5) and forward sliding reaction in Eq. (6.6). Note that both reactions are irreversible, and reflective boundary conditions must be used in the first-passage time analysis. For the backward reaction, the particle describing A_i and A_{i+1} is now placed between an absorbing and a reflective wall, subject to a total force given by Eq. (6.13). Similar to translocation moves, the propensity can be estimated by

$$a_s \equiv \frac{1}{\Delta t_s}. \tag{6.20}$$

The average time Δt_s to reach the absorbing wall a distance b from the reflecting barrier, is computed from Eq. (2.49) as

$$\Delta t_s = 2 \left(\frac{e^{-\epsilon} + \epsilon - 1}{\epsilon^2} \right) \tau_b, \tag{6.21}$$

where ϵ and τ_b are given by Eqs. (6.16) and (6.17), respectively. This expression for the propensity is also valid for sliding reactions, provided the appropriate force acting on the particle is used. For the sliding reaction in Eq. (6.6), the particle is now composed of subchains $\{A_k\}_{i-j\leq k\leq i+1}$ whereupon the total force can be expressed as

$$\mathbf{F}_{i,j} = \frac{1}{b}\left(\frac{\partial \mathcal{F}_{i-j}}{\partial n_{i-j}} - \frac{\partial \mathcal{F}_{i+1}}{\partial n_{i+1}}\right)\mathbf{u}_{i,j} + \left(\sum_{k=i-j+1}^{i} n_k\right) q\mathbf{E}, \tag{6.22}$$

since the conformational entropy of the taut subchains $\{A_k\}_{i-j+1\leq k\leq i}$ remains constant during the sliding motion of the chain. In the expression above, $\mathbf{u}_{i,j}$ is the unit vector from lake $i-j$ to lake $i+1$.

Finally, the propensity of nucleation reactions can be evaluated using classical nucleation theory [196]. If $\mathcal{F}^{\text{new}}$ and $\mathcal{F}^{\text{old}}$ respectively denote the free energy of the chain after and before the overflow of one segment into a new lake, then the propensity can be estimated using the success rate of the overflow attempts by [196, 197]

$$a_n = k_0 \exp\left[-(\mathcal{F}^{\text{new}} - \mathcal{F}^{\text{old}})\right], \tag{6.23}$$

where k_0 is a rate constant accounting for the local friction during the overflow [196,197]. In the dimensionless units of the simulation, $k_0 = 1$.

6.2.4 Free energies

The calculation of the propensities outlined above follows from an evaluation of the partial free energies in Eq. (6.3). As discussed by Kumar and Muthukumar in the context of translocations though a pore, the calculation of free energy barrier will be dominated by the conformational entropy of the polyelectrolyte chain [198]. The distribution of ionic clouds and solvent molecules have a minor effect and the calculation of the free energy for the polyelectrolyte chain is quantitatively equivalent to that of an uncharged self-avoiding walk [198]. As a result, we only consider conformational changes of the subchain inside the sphere. We used a Monte Carlo simulation [199] that allows us to compute the free energy of each subchain as a function of its size n_i and its orientation with respect to the field, described by the variable c_i [see Fig. 6.2]. The overall method, described in detail in Chapter 5, provides the free energy of self-avoiding, single- or double-tethered chains, under spherical confinement and subject to an external force.

To keep the calculations tractable, all the appropriate paths corresponding to a subchain conformation inside the sphere lie on a cubic lattice. The algorithm relies on a decomposition of the partition function (or free energy) into the product of several independent terms that can be evaluated separately using a Monte Carlo sampling and/or the numerical solution of a master equation [199].

6.2.5 Constants and parameters

The procedure presented above and summarized in Table 6.1 was implemented to compute the mobility and diffusivity of various DNA molecules embedded in an array of spheres, each of diameter $a = 600$ nm. For an electric force acting along the $\mathbf{i}_x$ direction, the x-component of the velocity v is calculated from an ensemble of time-dependent trajectories of the chain center of mass $x \equiv x_{\mathrm{cm}}(t)$ by [200]

$$v = \lim_{t\to\infty} \frac{\mathrm{d}\,\langle x\rangle}{\mathrm{d}t}, \tag{6.24}$$

and the diffusivity D is

$$D = \frac{1}{2}\lim_{t\to\infty} \frac{\mathrm{d}}{\mathrm{d}t}[\langle x^2\rangle - \langle x\rangle^2], \tag{6.25}$$

while the mobility is, by definition, $\mu \equiv v/E$.

In our simulations, all the position and time variables are expressed in terms of the length $b \approx 100$ nm and Brownian time $\tau_k \approx 10^{-3}$ s of a Kuhn segment [201]. Likewise, the mobility and diffusivity of the chain are respectively written in units of the mobility of *ds*DNA in free solution $\mu_0 \approx 2\times 10^{-4}$ (cm^2 V^{-1} s^{-1}) [9] and the diffusion coefficient of a Kuhn segment $D_k \equiv b^2/\tau_k \approx 10^{-11}$ m^2/s. The strength of the electric field is measured relative to Brownian forces by the Péclet number

$$\mathrm{Pe} = \frac{qEb}{k_BT}, \tag{6.26}$$

where q is the charge of a Kuhn segment. Using the charge $q = 50e$ for a duplex DNA Kuhn segment [202], we found that $\mathrm{Pe} \approx 0.1$ translates to $E \approx 5$ V/cm. Moreover, the molecules of interest here range in size from 100 to 1000 Kuhn segments (or 30 to 300 kilobase pairs) and are driven by several electric fields from $\mathrm{Pe} = 0.05$ to 2 (or 2.5 to 100 V/cm). Note that these values extend well beyond typical values one usually encounters in conventional DC gel electrophoresis, namely lengths up to up

1. Start with an initial configuration of the chain embedded in the array of spheres, *i.e.*, specify the triplet $\{n_i, x_i, c_i\}$ for each subchain A_i.

2. At each step of the simulation,

   (a) For every translocation, or sliding move, if any:

      i. Calculate the total force from the partial free energies using Eqs. (6.13) and (6.22).

      ii. Calculate the propensities using Eq. (6.19) or (6.20).

   (b) For every nucleation reaction, if any:

      i. Calculate the propensity from the chain free energy using Eq. (6.23).

   (c) Using propensities of all reactions:

      i. Calculate tentative time steps using Eq. (6.10).

      ii. Select reaction with the smallest tentative time.

      iii. Update the configuration of the chain by enforcing the selected reaction.

   (d) Calculate the new position of the center of mass using Eq. (6.2).

3. Repeat step 2 until the asymptotic limits in Eqs. (6.24) and (6.25) are verified.

Table 6.1: Summary of the new lakes-straits algorithm describing the dynamics of a long DNA in an array of spheres.

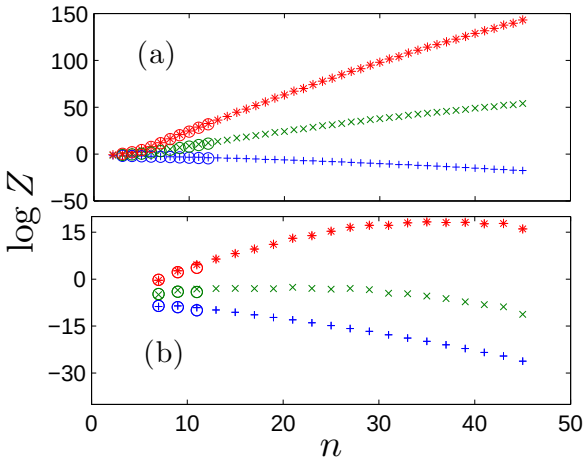

Figure 6.3: Plots of the partition function of a subchain as a function of its size n for various Péclet numbers, Pe = 0.05 (+), Pe = 0.5 (×), and Pe = 1.0 (∗). The circles (∘) show results from direct calculations of the partition function using an exact enumeration of the relevant paths on the lattice for $n \leq 12$. The agreement between the Monte Carlo sampling and the exact result is excellent. We provide here examples of: (a) a single-tethered subchain with $c = 1$ from Fig. 6.2, and (b) a double-tethered subchain with $c = 7$.

to 30 segments ($\approx$ 10 kilobase pairs) under fields $\approx 1 - 5$ V/cm [9]. Finally, it is worth mentioning that the complete set of results presented below (except the free energy calculations) was generated using two to three hours of computing on a shared workstation. In comparison to standard BD simulations, this represents an improvement of several orders of magnitude.

6.3 Results

6.3.1 Free energy calculation

Before we embark on a description of the transport coefficients of DNA, we present sample calculations of the partial free energies. As mentioned above, the free energy $\mathcal{F}_i$ of each subchain inside lake i is estimated using a Monte Carlo sampling of the relevant walks inside a confining sphere. To illustrate such calculations, we report in

c	α_{00}	α_{10}	α_{20}	α_{01}	α_{11}	α_{21}	α_{02}	α_{12}	α_{22}
1	-0.5497	-0.4225	-0.0038	-13.6722	4.2856	-0.0091	-0.2732	0.2885	-0.0016
2	-0.5829	-0.3248	-0.0036	2.2125	-1.5152	-0.0146	-0.6617	0.2503	-0.0012
3	-0.6782	-0.3647	-0.0040	-5.4850	1.5901	-0.0155	-0.1551	0.2213	-0.0004
4	-7.6043	-0.2707	-0.0062	-10.9239	4.3240	-0.0098	-2.1435	0.3672	-0.0027
5	-7.9646	-0.1926	-0.0073	-13.4775	1.9245	-0.0191	-0.8245	0.2284	-0.0006
6	-5.9393	-0.3852	-0.0044	-20.2588	4.5252	-0.0113	-1.2283	0.3322	-0.0025
7	-5.9010	-0.3460	-0.0044	-2.6463	1.5847	-0.0149	-1.2762	0.2765	-0.0016
8	-5.7872	-0.3481	-0.0043	-12.5079	1.8690	-0.0182	-1.2763	0.2616	-0.0012
9	-7.9840	-0.1293	-0.0071	-11.5176	-1.5539	-0.0110	-1.8888	0.3158	-0.0022
10	-5.6250	-0.3155	-0.0037	-2.4116	-1.4165	-0.0143	-1.2066	0.2571	-0.0013
11	-5.9244	-0.3243	-0.0045	-20.5464	1.5748	-0.0116	-1.0437	0.3034	-0.0022

Table 6.2: Coefficients of fitting polynomials of $\log Z$. Note that fitting polynomials are dominated by linear terms in Pe and n for each orientation c.

the present contribution plots of the partition function Z, or, $\log Z = -\mathcal{F}/k_BT$ as a function of the occupancy n for two subchains with relative orientations $c = 1$ and $c = 7$ depicted in Fig. 6.2. The plots corresponding to the other orientations are included, for completeness, in Appendix A. As shown in Fig. 6.3, the estimated partition function from the MC sampling is in excellent agreement with its exact value using a direct enumeration of the relevant walks for small n, thereby demonstrating the validity of the sampling approach. For a fixed Pe, the partition function exhibits a maximum at a critical size n^*. For $n < n^*$, the electric field contribution dominates the confinement effects and Z increases with n. In contrast, when $n > n^*$, the inclusion of additional segments inside the lake is more dependent on confinement and self-exclusion effects. Under this condition, Z is a decreasing function of the n. The reader can refer to Chapter 5 for a detailed discussion of the behavior of the partition function. To incorporate the free energies thus calculated into our algorithm, the curves of the partition function for each orientation c were fitted using polynomials of n and Pe

$$\log Z \approx Q_c(n, \mathrm{Pe}) = \sum_{i=0}^{2} \sum_{j=0}^{2} \alpha_{ij} n^i \mathrm{Pe}^j, \tag{6.27}$$

where the c-dependent coefficients are listed in Table 6.2.

6.3.2 Time trajectories

To ensure that the transport coefficients result from fully converging ensemble averages, the two expressions in Eqs. (6.24) and (6.25) indicate that both $\log \langle x \rangle$ and $\log[\langle x^2 \rangle - \langle x \rangle^2]$ must be linear functions of $\log t$, for long time t, with a slope of 1. From the plots of the first two moments of the center of mass position in Fig. 6.4, we can see that these two conditions are satisfied, provided a sufficiently long running time to eliminate the bias from the initial configuration of the chain. The plots shown are for a molecule with $N = 500$, under fields of magnitude $\mathrm{Pe} = 0.5$ and $\mathrm{Pe} = 2.0$. Similar results were obtained for all other sizes and Péclet numbers.

6.3.3 Mobility

Let us focus now on the mobility of the chain. In Fig. 6.5, we plot the dimensionless mobility of several molecules as a function of their lengths for various Péclet numbers. As can be seen from the figure, the mobility of DNA crosses three distinct regimes:

i) For low values of the Péclet number, $\mathrm{Pe} \leq 0.15$, the small bias introduced by the electric field barely affects the direction of the newly created lakes. The chain of lakes behaves as a random coil (albeit on a lattice) and, as a result, the mobility of the DNA scales like $\mu \sim N^{-1}$. This result is consistent with the scaling from biased reptation theory in the non-oriented regime [9] and ensures a size-based fractionation in the low field regime of molecules up to $N \approx 400$ in arrays of spheres. Note that this range of separation extends beyond that in DC electrophoresis in gels under normal concentrations (1 wt% agarose), where the separations are limited to molecules shorter than 30 Kuhn segments driven by dimensionless fields $\mathrm{Pe} \sim 0.1$ [9, 79]. Although early experiments by Fangman using dilute agarose gels (0.1 % agarose) successfully separated long molecules $N \approx 150 - 2000$ under an electric field of 2.5 V/cm ($\mathrm{Pe} \approx 0.05$), the overall structure of arrays of spheres is more robust and more mechanically stable than dilute gels. They should, therefore, be easier to manipulate than dilute agarose gels.

ii) At high values of the Péclet number, $\mathrm{Pe} \geq 1$, the direction of the newly created lakes is strongly biased by the electric field. If θ is the angle between the newly created lake and the direction of the force, then Fig. 6.6 indicates that $\langle \cos\theta \rangle \approx 1$ for high values of Pe. The alignment of the chain of lakes in the direction of the force leads to a

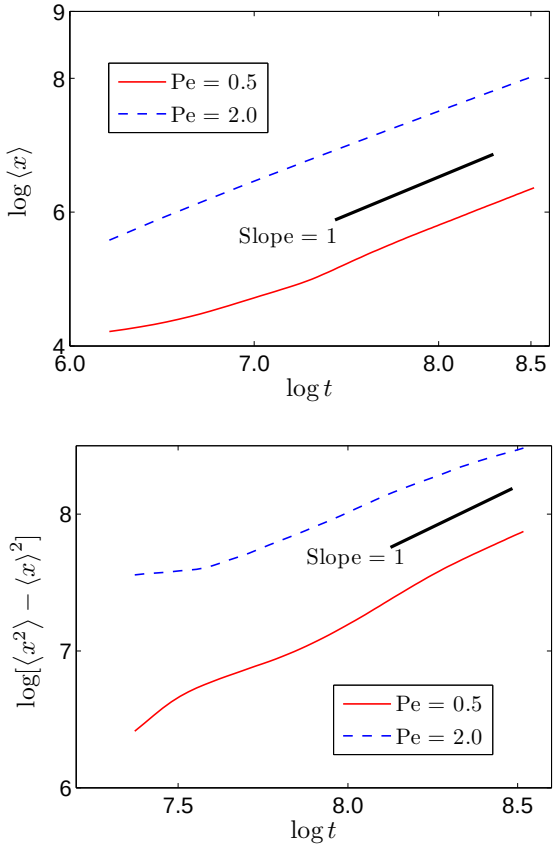

Figure 6.4: Plots of the time trajectory of the first moment, $\langle x \rangle$ (top), and the second centered moment $\langle x^2 \rangle - \langle x \rangle^2$ (bottom), of the center of mass position, for a molecule with $N = 500$ under two electric fields, Pe = 0.5 and Pe = 2.0. The behavior of both moments is linear at long times with a slope of 1.

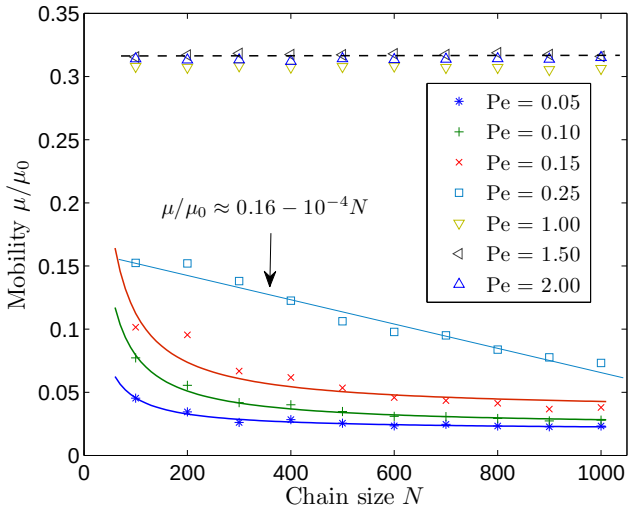

Figure 6.5: Plots of the dimensionless mobility for various chains lengths and Péclet numbers. As the electric field is increased, the mobility crosses three different regimes. At low electric fields, the mobility $\mu \sim N^{-1}$, while $\mu \sim N^0$ at high electric fields. At moderate electric fields, the mobility decreases linearly with the Péclet number. Note the saturation value at high Pe, where $\mu/\mu_0 \approx 0.32$.

mobility $\mu \sim N^0$ and a loss of separation. This result also agrees with biased reptation, but now in the oriented regime [9, 79].

The saturation value of the mobility can further be justified in terms of biased reptation where the mobility follows from Zimm's result [155]

$$\mu = \mu_0 \left\langle \frac{h_x^2}{\mathcal{L}_c \mathcal{L}} \right\rangle, \tag{6.28}$$

which distinguishes between the dimensionless contour length of the DNA, $\mathcal{L}_c = N$, and the contour length $\mathcal{L}$ of the chain of lakes containing the molecule. In the above expression, $h_x \simeq \mathcal{L}$ represents the projection of the end-to-end vector along the force direction $\mathbf{i}_x$ in the oriented regime. If $\langle n \rangle$ denotes the average number of segments, or occupancy of a lake, then

$$\langle \mathcal{L} \rangle = \frac{N}{\langle n \rangle} \left(\frac{a}{b} \right), \tag{6.29}$$

whereupon the dimensionless mobility of the chain becomes

$$\frac{\mu}{\mu_0} = \frac{1}{\langle n \rangle_\infty} \left(\frac{a}{b} \right), \tag{6.30}$$

where $\langle n \rangle_\infty$ is the average occupancy of a lake at high fields. The dependence of the average occupancy on the field is plotted in Fig. 6.7. When the field is increased, the chain extends over a larger number of lakes. As a result, the average occupancy decreases with the Péclet number until it reaches a limiting behavior at high fields whereby $\langle n \rangle_\infty \approx 19$, independent of the chain length. Inserting the ratio $a/b = 6$ into Eq. (6.30), the mobility of DNA in inverse opals saturates to a value $\mu/\mu_0 \approx 0.32$, in excellent agreement with the limiting value from Fig. 6.5.

iii) For intermediate values of the Péclet number, the mobility displays a strong dependence on the molecular weight of DNA, indicating that a size-based separation is possible over our entire range of lengths. In this new regime, we found that the decrease of the mobility is best approximated by a linear fit, a new scaling that was not previously observed in conventional DC gel electrophoresis. To further investigate the origin of such behavior, we repeated our simulations using purely elastic free energies [155], thus completely neglecting the effects from the spherical confinement. We found that, by doing so, the mobilities became independent of molecular weight, consistent with the original LS model [155] and other models of reptation [9] for which confinement

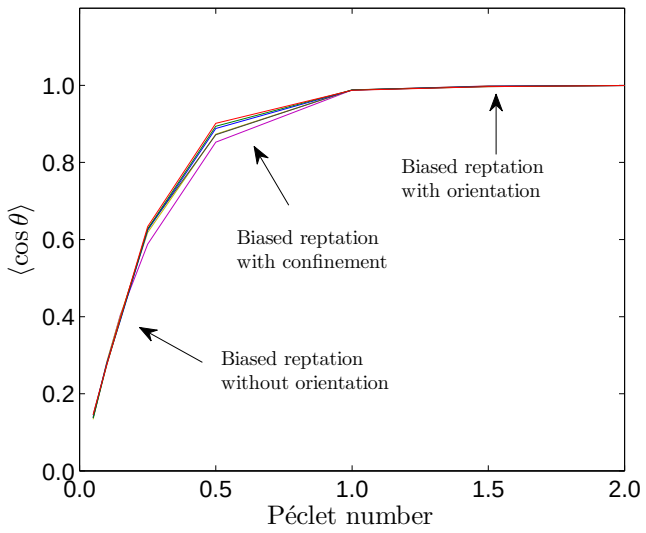

Figure 6.6: Plot of the average orientation of newly created lakes as a function of the Péclet number for various chain lengths. The average orientation is independent of chain length. However, as the electric field is increased, $\langle \cos\theta \rangle$ approaches a limiting value of 1, indicating that each new lake is created in the direction of the external field.

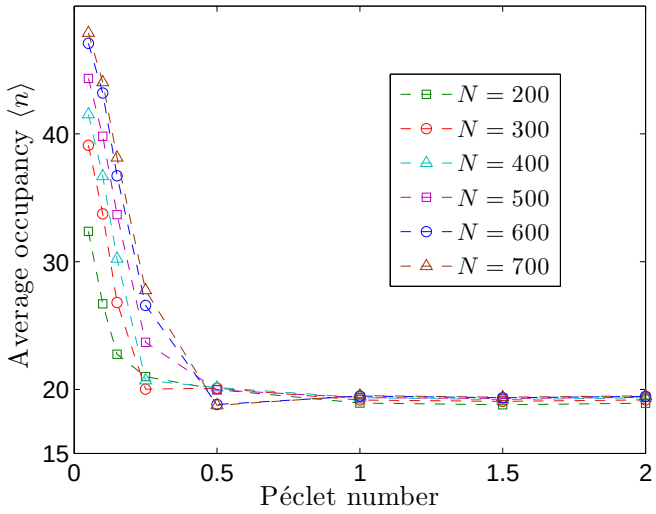

Figure 6.7: Plot of the average occupancy of a lake as a function of the Péclet number for various molecular weights. The average occupancy of a lake decreases with the electric field and saturates at high fields, where $\langle n \rangle_\infty \approx 19$.

considerations are absent. As a result, the emergence of this new regime at intermediate fields is primarily induced by the well-defined confinement from the solid spheres. We call it, accordingly, biased reptation with confinement (BRC).

6.3.4 Diffusivity

The quality of DNA separations in arrays of spheres also depends on the diffusivity of the molecule. It is an essential transport property that governs the broadening of the bands during the separation. The variations of the dimensionless diffusivity as a function of the chain size and the applied field are shown in Fig. 6.8. As is apparent from the figure, the diffusivity is an increasing function of the Péclet number for a molecule of fixed length. This can be understood from the following argument. In the rare events when a molecule is trapped in long lived configurations (for example, delayed overflows), its instantaneous velocity greatly differs from the average velocity. The random deviation from the average velocity is more pronounced with increasing field, resulting in an increased dispersion or diffusivity. The figure also indicates that the diffusivity of large

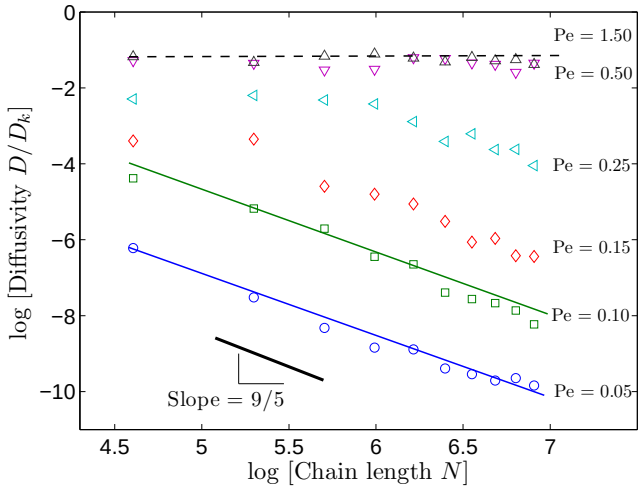

Figure 6.8: Plots of the dimensionless diffusivity for various chains lengths and Péclet numbers.

molecules is exponentially reduced compared to their shorter counterparts at low fields. At small Péclet numbers, all molecules are essentially in "equilibrium reptation," a term first coined by Slater [203]. Equivalently, they are driven by an electric field that does not affect their coarse-grained, self-avoiding walk conformation. A linear fit of the log-log plot of D as a function of N results in the scaling $D \sim N^{-1.7}$, while, according to biased reptation theory, the diffusivity $D \sim N^{-9/5}$ for self-avoiding chains in the low field regime [203]. However, the scaling law for the diffusivity in arrays of spheres is satisfied for a wider range of molecular weights, contrasting with conventional gels where the sharp decrease of the diffusivity in the "equilibrium reptation" regime is only valid for molecules with $N < 30$ [203]. In fact, larger molecules in gels undergo a less steep decline, followed by an incline, of their diffusivity, with a direct consequence on the overall resolution of the separation [203]. As the Péclet number is increased, differences in the diffusivity between molecules of different lengths subside until all diffusivities approach a size-independent value at high Pe, in agreement with biased reptation theory in the oriented regime [203].

6.3.5 Resolution in the BRC regime

The resolution of separations in array of spheres in the BRC regime can be estimated as follows. Two molecules that respectively traveled distances d_1 and d_2 during a time interval t are separated by a distance

$$\Delta d \equiv |d_2 - d_1| = \mu_0 E |\alpha(N_2 - N_1)| t \tag{6.31}$$

where $\alpha \approx -10^{-4}$ is the slope of a linear fit of the dimensionless mobility at $E = 12.5$ V/cm (or, Pe $= 0.25$). For a rise in lengths of $N_2 - N_1 = 100$, we find that $\Delta d \approx 0.25$ mm after an interval $t \approx 1000$ seconds, or a migration over a total average distance of $d \approx 3.75$ mm. Furthermore, we can evaluate the impact of dispersion in this regime from the bandwidth $\sigma = \sqrt{2Dt}$. From Fig. 6.8, the diffusivity $D \approx 5 \times 10^{-2} D_k$ at $E = 12.5$ V/cm, where the diffusion coefficient of a Kuhn segment $D_k \approx 10^{-11}$ m^2/s was estimated above. After a migration time of $t = 1000$ s, the estimated bandwidth $\sigma \approx 0.03$ mm is much less than the separation distance $\Delta d \approx 0.25$ mm, indicating a sharp separation.

Similarly, the resolution between two molecules for a separation length L is [40]

$$R_s = \frac{\Delta U}{\langle U \rangle} \sqrt{\frac{\langle \mathcal{N} \rangle}{16}}, \tag{6.32}$$

where $\Delta U \equiv |U_1 - U_2|$ and $\langle U \rangle \equiv (U_1 + U_2)/2$ are calculated from the average velocities, U_i, of each molecule $i = 1, 2$. The second term in Eq. (6.32), $\langle \mathcal{N} \rangle \equiv (\mathcal{N}_1 + \mathcal{N}_2)/2$, is calculated from the number of theoretical plates,

$$\mathcal{N}_i = \frac{U_i L}{2D_i}, \tag{6.33}$$

for each molecule $i = 1, 2$.

Using the linear fit of the mobility at $E = 12.5$ V/cm (or, Pe $=0.25$) from Fig. 6.5 and the diffusivity $D \approx 5 \times 10^{-2} D_k$ from Fig. 6.8, the resolution between two molecules $N_1 = 500$ and $N_2 = 600$ is

$$R_s \approx 0.03 \sqrt{\frac{L}{a}}. \tag{6.34}$$

For a cavity diameter $a = 600$ nm, we find that the minimum distance for a baseline separation, *i.e.* $R_s = 1.5$ [40], is $L_{\min} \approx 1.3$ mm. For a separation distance of 1 cm, the resolution increases to $R_s \approx 3.8$. Our results suggest that, compared to conventional

gel electrophoresis, a much wider range of molecular weights can be separated in arrays of spheres in the BRC regime with a minimal dispersion.

6.4 Conclusion

We studied the electrophoretic motion of a long DNA molecule inside a periodic array of confining spheres using a modified lakes-straits model that accounts for the limited connectivity of the array and the confinement free energy of the chain. Our method relies on coarse-graining of the chain at the scale of the confining sphere that results in much faster running times than Brownian dynamics simulations. It is worthwhile to point out here that, since the interconnecting pores have a diameter of tens of nanometers, the energy cost associated with bending the *ds*DNA near a pore entrance prevents lateral excursions of the chain out of internal lakes. As a result, we did not incorporate the possibility of hernias formation in our simulations. Likewise, we did not allow for a looping of the chain back on itself, *i.e.*, the one-dimensional chain of lakes is considered self-avoiding.

The mobilities and diffusivities from our calculations are consistent with biased reptation theory in both the low and high field regime. At moderate electric fields, our simulations show that the mobility is a strong function of chain length suggesting that separations of DNA are possible in arrays of spheres for a wider range of sizes than conventional gel electrophoresis. Our results also reveal a new regime of DNA reptation at moderate fields that we call biased reptation with confinement.

While we carried all of our simulations for three-dimensional cubic lattices, the algorithm we described above can be further applied to one- and two-dimensional lattices, as well as different types of lattices with the appropriate change of entry and exit points of the lakes. In fact, an ordered array is not even necessary as long as the chain can be pulled between the confining cells in a single file. In addition to spherical confinement, other cell geometries (cubes for examples, or spheres with a varying diameter) can also be considered, provided a calculation of the new confinement free energies is carried out.

Chapter 7

Conclusion

In this thesis, we employed several computational approaches to investigate the motion of DNA molecules in three different geometries. We also developed a Monte Carlo method for an efficient calculation of the partition function of a self-avoiding polymeric chain subject to confinement and tethering constraints. We showed how coarse-grained models can accurately inform about the dynamics of DNA without resorting to time-consuming molecular dynamics simulations or Brownian dynamics simulations.

Short DNA in a nanofilter

The motion of short DNA in a nanofilter was first approached using scaling arguments. The relevant physical transport processes, namely, the electric field induced convective translation between slits and rotational diffusion, were defined. The rates (or time scales) at which these processes occur depend on the geometry of the device, the applied electric field and the size of DNA. The faster rate determines the dominant transport and the ensuing dynamics of the chain. The predictions from the scaling argument were compared to Brownian dynamics simulations of the motion. Here the rod-like DNA molecule was modeled as an inextensible dumbbell that consists of two charged beads connected by an undeformable rod. The motion of each bead is governed by a Langevin equation, a stochastic differential equation that results from the balance of the forces acting on each bead. Upon integration of the differential equation with the appropriate boundary conditions, the time-dependent trajectory of each bead, and therefore the dumbbell, were obtained. Our results show that nanofilters need not necessarily operate under low electric fields (near equilibrium) in order to separate short rigid DNA molecules. The loss of separation observed in the experiments is not the upper limit of device operation [38]. Rather, our model predicts that strong field conditions will still lead to a separation with a band inversion; longer molecules elute first, in contrast to low-field separations where shorter molecules escape faster than their larger counterparts [35, 38]. Recent experimental evidence confirmed our theoretical predictions on the extended range of application of the nanofilter and the concomitant performance increase [121, 122].

DNA unhookning from a single post

In our study of the unhooking of a chain wrapped around a cylindrical post, we used stochastic methods (Monte Carlo and Master Equation) introduced in the context of translocation of a chain through a membrane [145]. Under a set of well-defined assumptions, we established an analogy between translocation of a chain through a small pore and unhooking from a post. In particular, both translocation through a membrane and the unhooking of a chain from a post can be modeled as the one-dimensional random motion of a particle subject to an external potential field. We therefore used a Monte Carlo (MC) algorithm based on elementary jumps of the chain to obtain the probability distribution of the exit time [102]. A direct calculation of the moments of the exit time (average and standard deviation) resulted from solving a master equation (ME) of the moments of the exit times for all initial configurations of the chain around the post. To assess the validity of our model and the assumptions underlying it, more realistic Brownian dynamics simulations were performed in our group. We obtained very good agreement between the two models, thus corroborating the results from the MC/ME model. By implementing the MC/ME algorithm, we were able to obtain statistical properties of the unhooking time of the chain in a fast and efficient way. Our results indicate that the unhooking nature depends on a dimensionless number that measures the relative effect of thermal fluctuations to that of the electric field. For typical chain lengths and applied electric fields used in experiments, thermal fluctuations are rather negligible; the unhooking process is effectively deterministic. This lends credence to models in the literature that treat the unhooking as a purely convective (deterministic) process [27].

DNA in an array of spheres

The motion of long DNA in a network of spherical cavities connected by narrow openings can be envisioned as a series of simultaneous mass transfers between adjacent spheres. The continuous threading of the chain through the narrow pores was modeled as set of three elementary moves, or stochastic reactions. The reaction rates are calculated using first-passage time statistics and nucleation theory while the total free energy of the chain is computed via a Monte Carlo sampling. Using a standard stochastic kinetic

algorithm, the trajectories of the chain center of mass are simulated for various chain lengths and field strengths. In the limits of low and high fields, our results agree with standard biased reptation theory [9]. At moderate fields, the strong confinement inside the cavities leads to a new regime of DNA electrophoresis that we term biased reptation with confinement. The mobility in this regime scales linearly with molecular weight and there is minimal dispersion.

Final thoughts

By using coarse-grained approaches to model the motion of DNA, we were able to save a great deal of computer time; most of our calculations required a few minutes to a two-three hours to run on a shared workstation. In comparison, several hours of computer time were necessary to run the Brownian dynamics simulation of the unhooking of the chain. That being said, the inclusion of additional refinements into the description will ultimately require some Langevin-like approach. For instance, in insulating geometries, the value of the electric field is no longer uniform and a detailed knowledge of its spatial distribution becomes imperative. Also, the methods we used in this work are very flexible and can be adapted to a variety of situations where a fine-grained description of the motion of DNA is not necessary. For example, multi-post unhooking, arising in long DNA dynamics in post-arrays, can easily be approached using the translocation analogy. Moreover, the motion of DNA in inverse opals under pulsed fields can be studied using the new lakes-straits algorithm we developed. In the long term, the separation tools we studied in our research, in conjunction with other analysis tasks (sequencing, PCR, etc.), can be integrated onto a highly precise, single lab-on-a-chip, with vastly reduced processing time and sample volumes. The increased knowledge gained from fundamental studies, such as the work in this thesis, will certainly contribute to enhancing the performance of microfluidic devices for DNA separation.

Appendix A

Plots of the partition function

The plots of the partition function of a subchain confined inside a spherical cavity are presented in this appendix. The plots correspond to the 11 possible orientations depicted in Fig. 6.2.

- Orientation $c = 1$

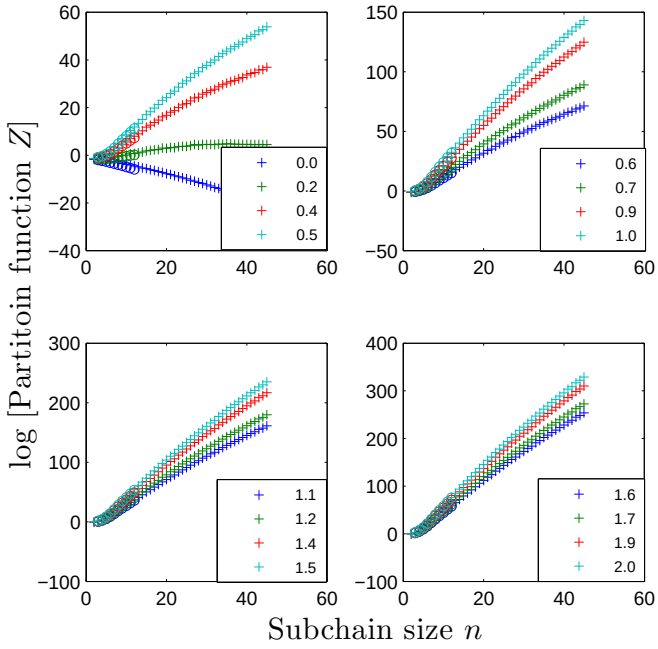

Figure A.1: Plots of the logarithm of the partition function of several subchains with orientation $c = 1$ for different Péclet numbers.

- Orientation $c = 2$

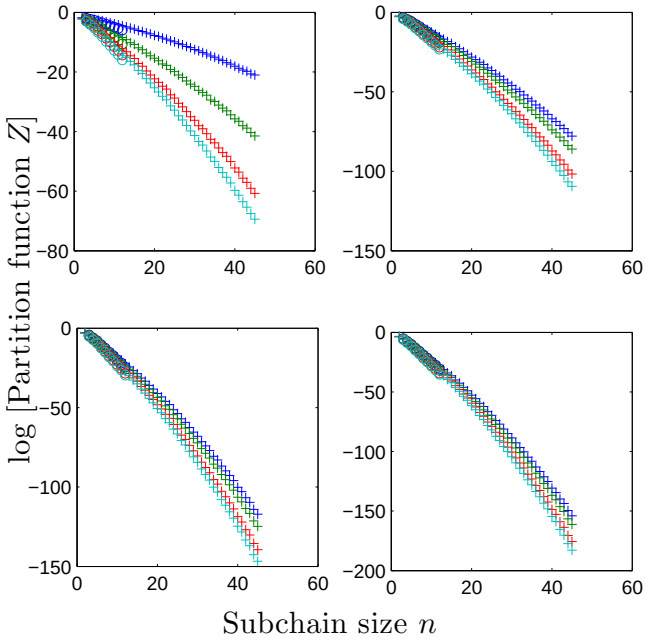

Figure A.2: Same as Fig. A.1 for orientation $c = 2$.

- Orientation $c = 3$

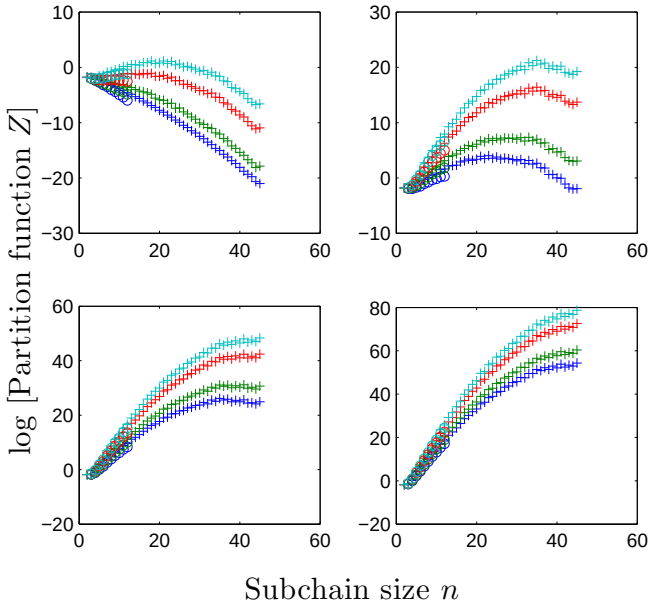

Figure A.3: Same as Fig. A.1 for orientation $c = 3$.

- Orientation $c = 4$

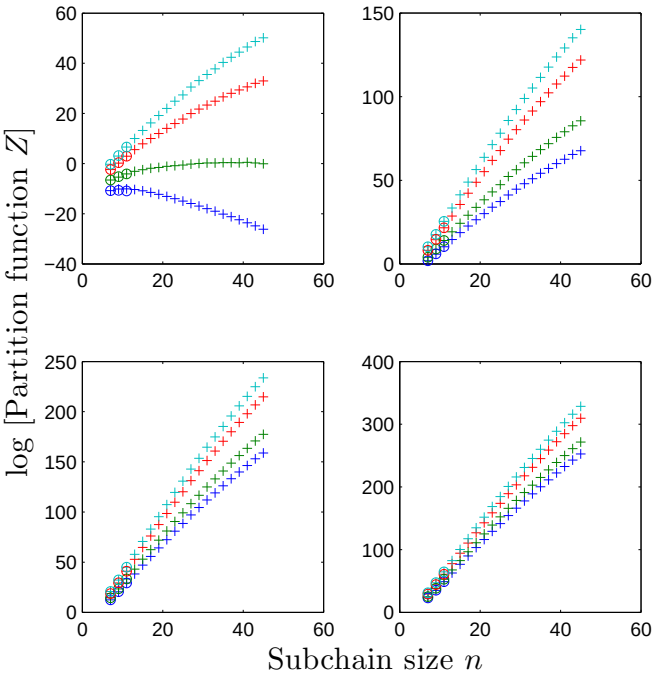

Figure A.4: Same as Fig. A.1 for orientation $c = 4$.

- Orientation $c = 5$

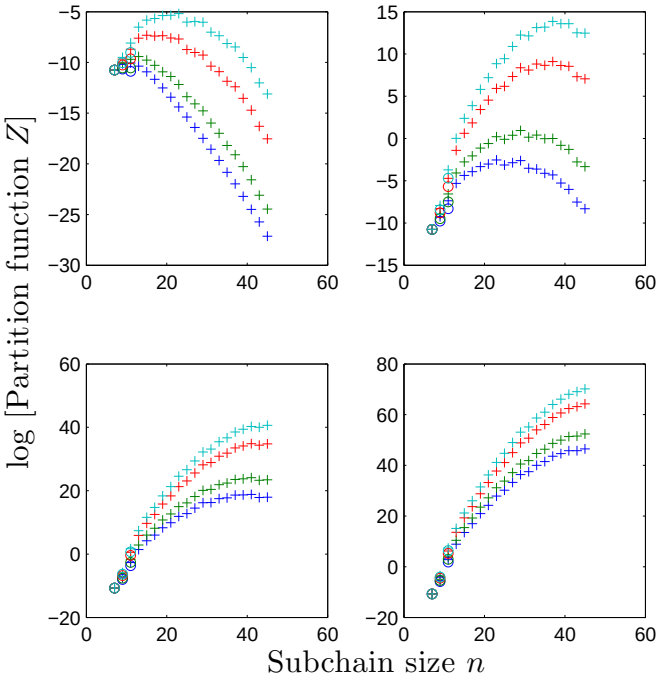

Figure A.5: Same as Fig. A.1 for orientation $c = 5$.

- Orientation $c = 6$

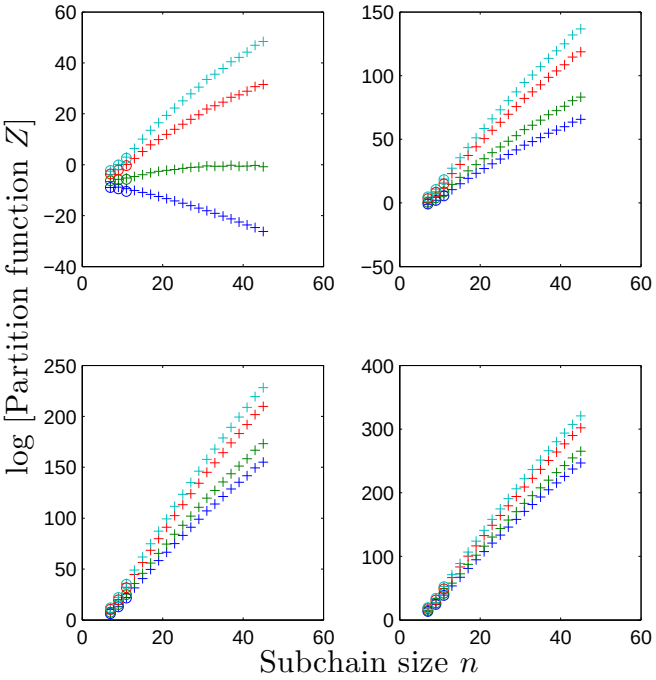

Figure A.6: Same as Fig. A.1 for orientation $c = 6$.

- Orientation $c = 7$

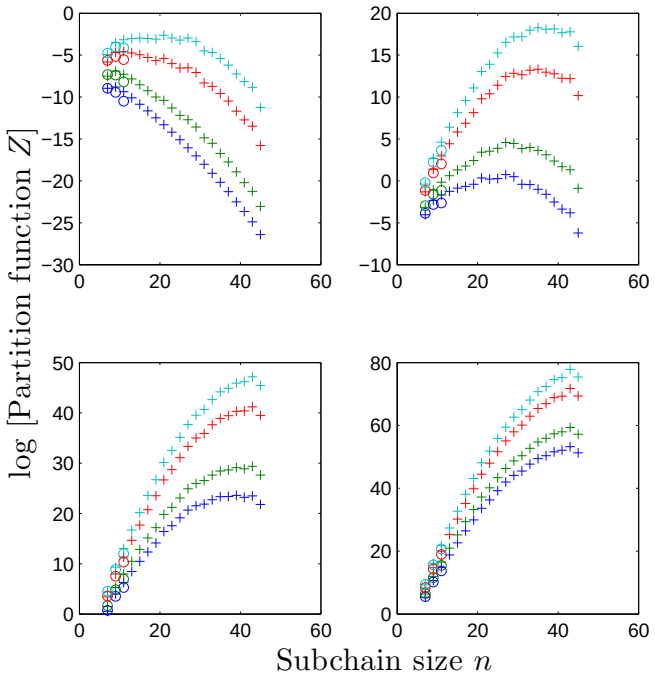

Figure A.7: Same as Fig. A.1 for orientation $c = 7$.

- Orientation $c = 8$

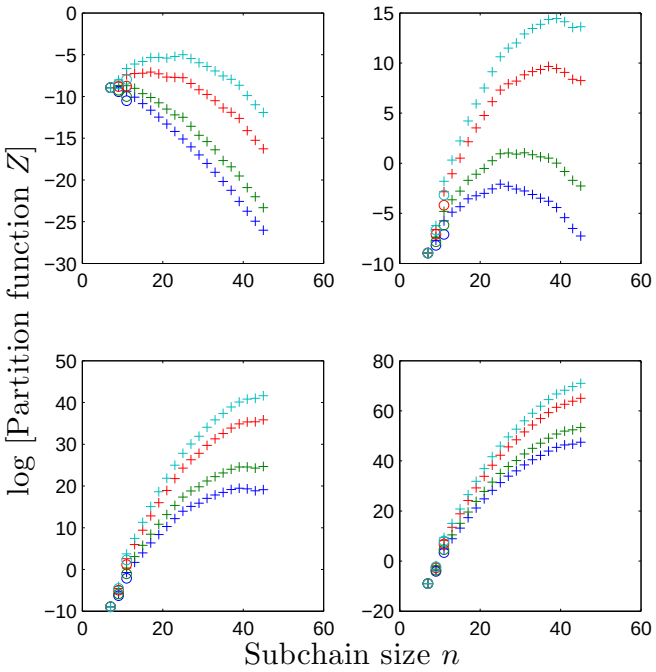

Figure A.8: Same as Fig. A.1 for orientation $c = 8$.

- Orientation $c = 9$

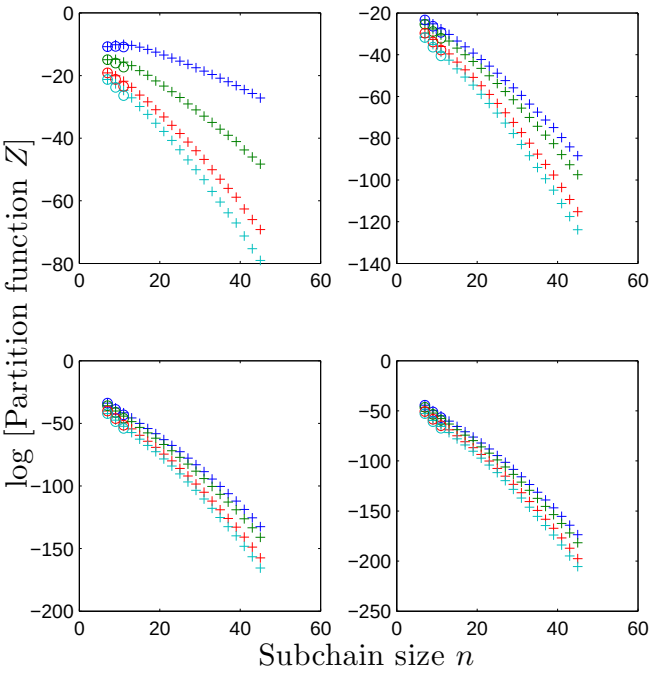

Figure A.9: Same as Fig. A.1 for orientation $c = 9$.

- Orientation $c = 10$

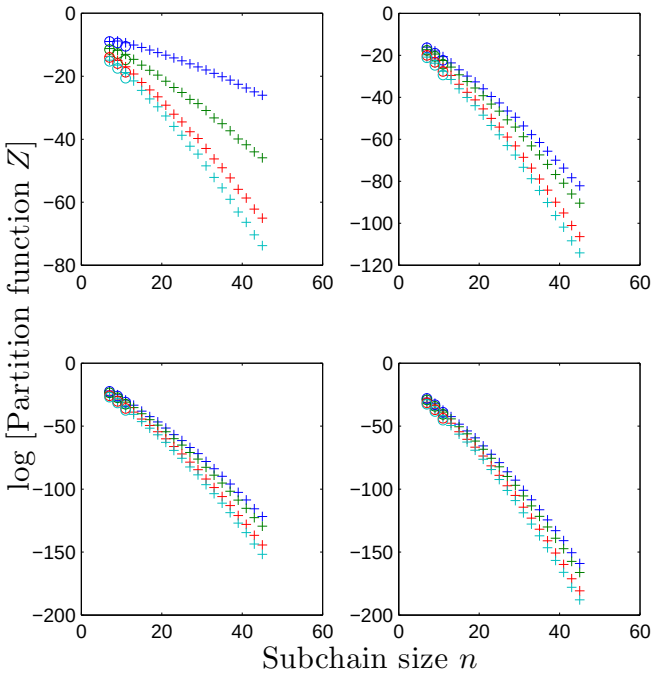

Figure A.10: Same as Fig. A.1 for orientation $c = 10$.

- Orientation $c = 11$

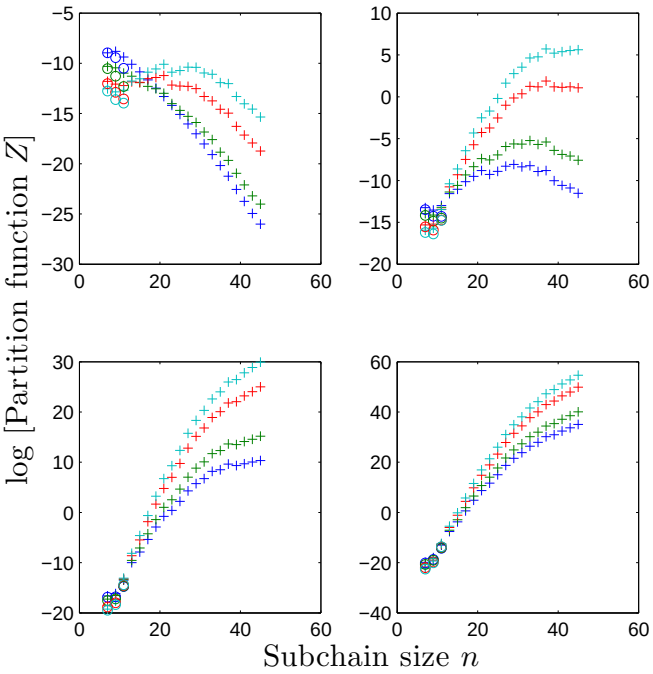

Figure A.11: Same as Fig. A.1 for orientation $c = 11$.

References

[1] M. Mandelkern, J. G. Elias, D. Eden, and D. M. Crothers. The dimensions of DNA in solution. *J. Mol. Biol.*, 152:153–161, 1981.

[2] S. G Gregory and *et al.* The DNA sequence and biological annotation of human chromosome 1. *Nature*, 441:315–321, 2006.

[3] F. Sanger and A. R. Coulson. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.*, 94:441–448, 1975.

[4] F. Sanger, S. Nicklen, and A. R. Coulson. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.*, 74:5463–5467, 1977.

[5] M. Delseny, J. Salses, R. Cooke, C. Sallaud, F. Regad, P. Lagoda, E. Guiderdoni, M. Ventelon, C. Brugidou, and A. Ghesquiere. Rice genomics: Present and future. *Plant Physiol. Biochem.*, 39:323–334, 2001.

[6] D. R. Soll. The ins and outs of DNA fingerprinting the infectious fungi. *Clin. Microbiol. Rev.*, 13:332–370, 2000.

[7] A. Collins and N. E. Morton. Likelihood ratios for DNA identification. *Proc. Natl. Acad. Sci.*, 91:6007–6011, 1994.

[8] Human Genome Project **http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml**.

[9] J.-L. Viovy. Electrophoresis of DNA and other polyelectrolytes: Physical mechanisms. *Rev. Mod. Phys.*, 72:813–872, 2000.

[10] D. C. Schwartz and C. R. Cantor. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell*, 37:67–75, 1984.

[11] C. L. Smith and C. R. Cantor. Purification, specific fragmentation, and separation of large DNA molecules. *Meth. Enzymol.*, 155:449–467, 1987.

[12] J. Weber, V. Barbier, S. Pages-Berhouet, V. Caux-Moncoutier, D. Stoppa-Lyonnet, and J.-L. Viovy. A High-throughput mutation detection method based on heteroduplex analysis using graft copolymer matrixes: Application to Brca1 and Brca2 analysis. *Anal. Chem.*, 76:4839–4848, 2004.

[13] C. A. Emrich, H. J. Tian, I. L. Medintz, J. Tom, and R. A. Mathies. Ultra-high-throughput genotyping on a 384-lane microfabricated capillary array electrophoresis device. *Anal. Chem.*, 74:5076–5083, 2002.

[14] S. Magnusdottir, H. Isambert, C. Heller, and J.-L. Viovy. Electrohydrodynamically induced aggregation during constant and pulsed field capillary electrophoresis of DNA. *Biopolymers*, 49:385–401, 1999.

[15] L. Mitnik, C. Heller, J. Prost, and J.-L. Viovy. Segregation in DNA solutions induced by electric fields. *Science*, 267:219–222, 1995.

[16] A. Manz, N. Graber, and H. M. Widmer. Miniaturized total chemical analysis systems: a novel concept for chemical sensing. *Sensor Actuat. B-Chem.*, 1:244–248, 1990.

[17] M. A. Burns, B. N. Johnson, S. N. Brahmasandra, K. Handique, J. R. Webster, M. Krishnan, T. S. Sammarco, P. M. Man, D. Jones, D. Heldsinger, et al. An integrated nanoliter DNA analysis device. *Science*, 282:484–487, 1998.

[18] J. C. McDonald, D. C. Duffy, J. R. Anderson, D. T. Chiu, H. Wu, O. J. Schueller, and G. M. Whitesides. Fabrication of microfluidic systems in poly (dimethylsiloxane). *Electrophoresis*, 21:27–40, 2000.

[19] H. Becker and L. E. Locascio. Polymer microfluidic devices. *Talanta*, 56:267–287, 2002.

[20] S. C. Jacobson, R. Hergenroder, L. B. Koutny, R. J. Warmack, and J. M. Ramsey. Effects of injection schemes and column geometry on the performance of microchip electrophoresis devices. *Anal. Chem.*, 66:1107–1113, 1994.

[21] G. W. Slater, S. Guillouzic, M. G. Gauthier, J. F. Mercier, M. Kenward, L. C. McCormick, and F. Tessier. Theory of DNA electrophoresis(1999-2002 1/2). *Electrophoresis*, 23:3791–3816, 2002.

[22] G. W. Slater, K. D. Dorfman, and J.-L. Viovy. *Understanding the electrophoresis of biological and synthetic macromolecules*. (Unpublished notes).

[23] N. Kaji, Y. Tezuka, Y. Takamura, M. Ueda, T. Nishimoto, H. Nakanishi, Y. Horiike, and Y. Baba. Separation of long DNA molecules by quartz nanopillar chips under a direct current electric field. *Anal. Chem.*, 76:15–22, 2004.

[24] W. D. Volkmuth and R. H. Austin. DNA electrophoresis in microlithographic arrays. *Nature*, 358:600–602, 1992.

[25] P. S. Doyle, J. Bibette, A. Bancaud, and J. L. Viovy. Self-assembled magnetic matrices for DNA separation chips. *Science*, 295:2237, 2002.

[26] N. Minc, C. Fütterer, K. D. Dorfman, A. Bancaud, C. Gosse, C. Goubault, and J. L. Viovy. Quantitative microfluidic separation of DNA in self-assembled magnetic matrices. *Anal. Chem.*, 76:3770–3776, 2004.

[27] N. Minc, J.-L. Viovy, and K. D. Dorfman. Non-Markovian transport of DNA in microfluidic post arrays. *Phys. Rev. Lett.*, 94:198105–198108, 2005.

[28] K. D. Dorfman and J.-L. Viovy. Semiphenomenological model for the dispersion of DNA during electrophoresis in a microfluidic array of posts. *Phys. Rev. E*, 69:11901, 2004.

[29] M. Cabodi, S. W. P. Turner, and H. G. Craighead. Entropic recoil separation of long DNA molecules. *Anal. Chem.*, 74:5169–5174, 2002.

[30] A. Baumgärtner and M. Muthukumar. A trapped polymer chain in random porous media. *J. Chem. Phys.*, 87:3082–3088, 1987.

[31] M. Muthukumar and A. Baumgärtner. Diffusion of a polymer chain in random media. *Macromolecules*, 22:1941–1946, 1989.

[32] M. Muthukumar and A. Baumgärtner. Effects of entropic barriers on polymer dynamics. *Macromolecules*, 22:1937–1941, 1989.

[33] M. Cabodi, Y. F. Chen, S. W. P. Turner, H. G. Craighead, and R. H. Austin. Continuous separation of biomolecules by the laterally asymmetric diffusion array with out-of-plane sample injection. *Electrophoresis*, 23:3496–3503, 2002.

[34] S. W. P. Turner, M. Cabodi, and H. G. Craighead. Confinement-induced entropic recoil of single DNA molecules in a nanofluidic structure. *Phys. Rev. Lett.*, 88:128103, 2002.

[35] J. Fu, J. Yoo, and J. Han. Molecular sieving in periodic free-energy landscapes created by patterned nanofilter arrays. *Phys. Rev. Lett.*, 97:18103–18107, 2006.

[36] J. Han, S. W. Turner, and H. G. Craighead. Entropic trapping and escape of long dna molecules at submicron size constriction. *Phys. Rev. Lett.*, 83:1688–1691, 1999.

[37] J. Han and H. G. Craighead. Separation of long dna molecules in a microfabricated entropic trap array. *Science*, 288:1026, 2000.

[38] J. Fu, P. Mao, and J. Han. Nanofilter array chip for fast gel-free biomolecule separation. *Appl. Phys. Lett.*, 87:263902–263902, 2005.

[39] J. Han and H. G. Craighead. Characterization and optimization of an entropic trap for DNA separation. *Anal. Chem.*, 74:394–401, 2002.

[40] J. C. Giddings. *Unified separation science*. Wiley New York, 1991.

[41] J. Y. Shiu, W. T. Whang, and P. Chen. Behavior of single DNA molecules in the well-ordered nanopores. *J. Chromatogr. A*, 1206(1):72–76, 2008.

[42] S. M. Yang and G. A. Ozin. Opal chips: vectorial growth of colloidal crystal patterns inside silicon wafers. *Chem. Commun.*, 2000:2507–2508, 2000.

[43] J. Y. Shiu, C. W. Kuo, and P. Chen. Actively controlled self-assembly of colloidal crystals in microfluidic networks by electrocapillary forces. *J. Am. Chem. Soc*, 126:8096–8097, 2004.

[44] J. Y. Shiu and P. Chen. Active patterning using an addressable microfluidic network. *Adv. Mat.*, 17:1866–1868, 2005.

[45] P. Jiang, J. F. Bertone, K. S. Hwang, and V. L. Colvin. Single-crystal colloidal multilayers of controlled thickness. *Chem. Mater*, 11:2132–2140, 1999.

[46] S. M. Yang, H. Miguez, and G. A. Ozin. Opal circuits of light-planarized microphotonic crystal chips. *Adv. Func. Mat.*, 12:425–431, 2002.

[47] G. A. Ozin and S. M. Yang. The race for the photonic chip: colloidal crystal assembly in silicon wafers. *Adv. Func. Mat.*, 11:95–104, 2001.

[48] M. Tabuchi, M. Ueda, N. Kaji, Y. Yamasaki, Y. Nagasaki, K. Yoshikawa, K. Kataoka, and Y. Baba. Nanospheres for DNA separation chips. *Nat. Biotechnol.*, 22:337–340, 2004.

[49] Y. Zeng and D. J. Harrison. Self-assembled colloidal arrays as three-dimensional nanofluidic sieves for separation of biomolecules on microchips. *Anal. Chem.*, 79:2289–2295, 2007.

[50] D. Nykypanchuk, H. H. Strey, and D. A. Hoagland. Brownian motion of DNA confined within a two-dimensional array. *Science*, 297:987–990, 2002.

[51] Y. Zeng and D. J. Harrison. Confinement effects on electromigration of long DNA molecules in an ordered cavity array. *Electrophoresis*, 27:3747–3752, 2006.

[52] C. López. Materials aspects of photonic crystals. *Adv. Mat.*, 15:1679–1704, 2003.

[53] C. K. Ullal, M. Maldovan, E. L. Thomas, G. Chen, Y. J. Han, and S. Yang. Photonic crystals through holographic lithography: Simple cubic, diamond-like, and gyroid-like structures. *Appl. Phys. Lett.*, 84:5434, 2004.

[54] M. Wada, Y. Doi, K. Inoue, J. W. Haus, and Z. Yuan. A simple-cubic photonic lattice in silicon. *Appl. Phys. Lett.*, 70:2966, 1997.

[55] A. Y. Grosberg and A. R. Khokhlov. *Giant Molecules, here, there, and everywhere*. Academic Press, San Diego, 1997.

[56] M. Doi. *Introduction to polymer physics*. Oxford Univerisity Press, Oxford, 1996.

[57] P. C. Hiemenz and T. P. Lodge. *Polymer Chemistry, 2nd edition*. Marcel Dekker, 2007.

[58] P. J. Flory. *Principles of polymer chemistry*. Cornell Univerisity Press, Ithaca and London, 1953.

[59] P.-G. de Gennes. *Scaling concepts in polymer physics*. Cornell University Press, 1979.

[60] M. Matsumoto, T. Sakaguti, H. Kimura, M. Doi, Y. Matsuzawa, K. Minagawa, and K. Yoshikawa. Direct observation of Brownian motion of macromolecules by fluorescence microscope. *J. Polym. Sci.*, 30:779–783, 1991.

[61] B. Zimm. Dynamics of polymer molecules in dilute solution: viscoelasticity, flow birefringence and dielectric loss. *J. Chem. Phys.*, 24:269–278, 1956.

[62] J. L. Barrat and J. F. Joanny. Theory of polyelectrolyte solutions. *Adv. Chem. Phys.*, 96:1–66, 1996.

[63] G. S. Manning. Limiting laws and counterion condensation in polyelectrolyte solution. 7. electrophoretic mobility and conductance. *J. Phys. Chem.*, 85:1506–1515, 1981.

[64] M. Doi and S. F. Edwards. Dynamics of concentrated polymer systems. *J. Chem. Soc., Faraday Trans*, 274:1789–1832, 1978.

[65] P.-G. de Gennes. Reptation of a polymer chain in the presence of fixed obstacles. *J. Chem. Phys.*, 55:572–579, 1971.

[66] K. Osaki. *Molecular Conformation and Dynamics of Macromolecules in Condensed Systems; Nagasawa, M., Ed.* Elsevier, Amsterdam, 1988.

[67] D. S. Pearson, G. Ver Strate, E. Von Meerwall, and F. C. Schilling. Viscosity and self-diffusion coefficient of linear polyethylene. *Macromolecules*, 20:1133–1141, 1987.

[68] L. S. Lerman and H. L. Frisch. Why does the electrophoretic mobility of DNA in gels vary with the length of the molecule? *Biopolymers*, 21:995–997, 1982.

[69] O. J. Lumpkin and B. H. Zimm. Mobility of DNA in gel electrophoresis. *Biopolymers*, 21:2315–2316, 1982.

[70] O. J. Lumpkin, P. Déjardin, and B. H. Zimm. Theory of gel electrophoresis of DNA. *Biopolymers*, 24:1573–1593, 1985.

[71] G. W. Slater and J. Noolandi. New biased-reptation model for charged polymers. *Phys. Rev. Lett.*, 55:1579–1582, 1985.

[72] G. W. Slater and J. Noolandi. Prediction of chain elongation in the reptation theory of DNA gel electrophoresis. *Biopolymers*, 24:2181–2184, 1985.

[73] G. W. Slater and J. Noolandi. On the reptation theory of gel electrophoresis. *Biopolymers*, 25:431–454, 1986.

[74] G. W. Slater and J. Rousseau, J. Noolandi. On the stretching of DNA in the reptation theories of gel electrophoresis. *Biopolymers*, 26:863–872, 1987.

[75] T. A. J. Duke and J.-L. Viovy. Motion of megabase deoxyribonucleic acid during field-inversion gel electrophoresis: Investigation by nonlocal Monte Carlo. *J. Chem. Phys.*, 96:8552, 1992.

[76] T. A. J. Duke, Viovy J.-L., and A. N. Semenov. Electrophoretic mobility of DNA in gels. I. New biased reptation theory including fluctuations. *Biopolymers*, 34:239–247, 1994.

[77] A. N. Semenov, T. A. J. Duke, and J.-L. Viovy. Gel electrophoresis of DNA in moderate fields: The effect of fluctuations. *Phys. Rev. E*, 51:1520–1537, 1995.

[78] G. T. Barkema, J. F. Marko, and B. Widom. Electrophoresis of charged polymers: Simulation and scaling in a lattice model of reptation. *Phys. Rev. E*, 49:5303–5309, 1994.

[79] C. Heller, T. Duke, and J-.L. Viovy. Electrophoretic mobility of DNA in gels. II: Systematic experimental study in agarose gels. *Biopolymers*, 34:249–259, 1994.

[80] G. W. Slater and G. Drouin. Why can we not sequence thousands of DNA bases on a polyacrylamide gel? *Electrophoresis*, 13:574–82, 1992.

[81] J. Shendure, R.D. Mitra, C. Varma, and G.M. Church. Advanced sequencing technologies: methods and goals. *Nature Reviews Genetics*, 5:335–344, 2004.

[82] J. J. Kasianowicz, E. Brandin, D. Branton, and D. W. Deamer. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Aca. Sci.*, 93:13770, 1996.

[83] A. Meller, L. Nivon, E. Brandin, J. Golovchenko, and D. Branton. Rapid nanopore discrimination between single polynucleotide molecules. *Proc. Natl. Aca. Sci.*, 97:1079, 2000.

[84] J. Li, D. Stein, C. McMullan, D. Branton, M. J. Aziz, and J. A. Golovchenko. Ion-beam sculpting at nanometre length scales. *Nature*, 412:166–169, 2001.

[85] J. Li, M. Gershow, D. Stein, E. Brandin, and J. A. Golovchenko. DNA molecules and configurations in a solid-state nanopore microscope. *Nature Mat.*, 2:611–615, 2003.

[86] A. J. Storm, C. Storm, J. Chen, H. Zandbergen, J.-F. Joanny, and C. Dekker. Fast DNA translocation through a solid-state nanopore. *Nano. Lett.*, 5:1193–1197, 2005.

[87] W. Sung and P. J. Park. Polymer translocation through a pore in a membrane. *Phys. Rev. Lett.*, 77:783–786, 1996.

[88] M. Muthukumar. Polymer translocation through a hole. *J. Chem. Phys.*, 111:10371, 1999.

[89] J. Chuang, Y. Kantor, and M. Kardar. Anomalous dynamics of translocation. *Phys. Rev. E*, 65:11802, 2001.

[90] H. C. Loebl, R. Randel, S. P. Goodwin, and C. C. Matthai. Simulation studies of polymer translocation through a channel. *Phys. Rev. E*, 67:41913, 2003.

[91] R. Zandi, D. Reguera, J. Rudnick, and W. M. Gelbart. What drives the translocation of stiff chains? *Proc. Natl. Aca. Sci*, 100:8649, 2003.

[92] I. Huopaniemi, K. Luo, T. Ala-Nissila, and S. C. Ying. Langevin dynamics simulations of polymer translocation through nanopores. *J. Chem. Phys.*, 125:124901, 2006.

[93] A. Baumgärtner and J. Skolnick. Spontaneous translocation of a polymer across a curved membrane. *Phys. Rev. Lett.*, 74:2142–2145, 1995.

[94] H. Bayley. Sequencing single molecules of DNA. *Current opinion in chemical biology*, 10:628–637, 2006.

[95] D. R. Meldrum. Sequencing genomes and beyond. *Science*, 292:515, 2001.

[96] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. Molecular biology of the cell. *Garland, New York*, 1994.

[97] W. Wickner and R. Schekman. Protein translocation across biological membranes. *science*, 310:1452, 2005.

[98] Dr Andre Marziali's research web page **www.physics.ubc.ca/abl/pages/research/nanoporesynthetic.html**.

[99] S. M. Iqbal, D. Akin, and R. Bashir. Solid-state nanopore channels with DNA selectivity. *Nature Nanotech.*, 2:243–248, 2007.

[100] C. Dekker. Solid-state nanopores. *Nature Nanotech.*, 2:209–215, 2007.

[101] S. Redner. *A guide to first-passage processes*. Cambridge University Press, 2001.

[102] M. G. Gauthier and G. W. Slater. Building reliable lattice Monte Carlo models for real drift and diffusion problems. *Phys. Rev. E*, 70:15103, 2004.

[103] F. Tessier, J. Labrie, and G. W. Slater. Electrophoretic separation of long polyelectrolytes in submolecular-size constrictions: a Monte Carlo study. *Macromolecules*, 35:4791–4800, 2002.

[104] N. Laachi, C. Declet, C. Matson, and K. D. Dorfman. Nonequilibrium transport of rigid macromolecules in periodically constricted geometries. *Phys. Rev. Lett.*, 98:098106, 2007.

[105] A. G. Ogston. The spaces in a uniform random suspension of fibres. *Trans. Faraday Soc.*, 54:1754–1757, 1958.

[106] D. Rodbard and A. Chrambach. Unified theory for gel electrophoresis and gel filtration. *Proc. Natl. Acad. Sci.*, 65:970–977, 1970.

[107] J. C. Giddings, E. Kucera, C. P. Russell, and M. N. Myers. Statistical theory for the equilibrium distribution of rigid molecules in inert porous networks. Exclusion chromatography. *J. Phys. Chem.*, 72:4397–4408, 1968.

[108] W. H. Stockmayer. In "molecular fluids". *Balian, R., Weill, G*, 1976.

[109] A. Ajdari and J. Prost. Free-flow electrophoresis with trapping by a transverse inhomogeneous field. *Proc. Natl. Acad. Sci.*, 88:4468–4471, 1991.

[110] N. C. Stellwagen, C. Gelfi, and P. G. Righetti. The free solution mobility. *Biopolymers*, 42:687–703, 1997.

[111] D. A. Hoagland, E. Arvanitidou, and C. Welch. Capillary electrophoresis measurements of the free solution mobility for several model polyelectrolyte systems. *Macromolecules*, 32:6180–6190, 1999.

[112] M. Streek, F. Schmid, T. T. Duong, and A. Ros. Mechanisms of DNA separation in entropic trap arrays: a Brownian dynamics simulation. *J. Biotechnol.*, 112:79–89, 2004.

[113] H. C. Berg. *Random Walks in Biology.* Princeton University Press, 1993.

[114] T. R. Strick, M. N. Dessinges, G. Charvin, N. H. Dekker, J. F. Allemand, D. Bensimon, and V. Croquette. Stretching of macromolecules and proteins. *Rep. Prog. Phys.*, 66:1–45, 2003.

[115] S. H. Kim, A. S. Panwar, S. Kumar, K. H. Ahn, and S. J. Lee. Electrophoresis of a bead-rod chain through a narrow slit: A Brownian dynamics study. *J. Chem. Phys.*, 121:9116–9122, 2004.

[116] O. B. Bakajin, 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.*, 80:2737–2740, 1998.

[117] P. D. Patel and E. S. G. Shaqfeh. A computational study of DNA separations in sparse disordered and periodic arrays of posts. *J. Chem. Phys.*, 118:2941–2951, 2003.

[118] A. S. Panwar and S. Kumar. Time Scales in Polymer Electrophoresis through Narrow Constrictions: A Brownian Dynamics Study. *Macromolecules*, 39:1279–1289, 2006.

[119] E. J. Hinch. Brownian motion with stiff bonds and rigid constraints. *J. Fluid Mech.*, 271:219–234, 2006.

[120] P. Grassia and E. J. Hinch. Computer simulations of polymer chain relaxation via Brownian motion. *J. Fluid Mech.*, 308:255–288, 2006.

[121] A. Strychalski, W. Lau, L. Archer, and H. Craighead. Length-based separation of short DNA using nanoslit arrays. *Twelfth International Conference on Miniaturized Systems for Chemistry and Life Sciences, San Diego, California, USA*, page 856, 2008.

[122] E. A. Strychalski, H. W. Lau, and L. A. Archer. Nonequilibrium separation of short DNA using nanoslit arrays. *J. Appl. Phys.*, 106:024915–024915, 2009.

[123] T. M. Squires and S. R. Quake. Microfluidics: Fluid physics at the nanoliter scale. *Rev. Mod. Phys.*, 77:977–1026, 2005.

[124] R. B. Schoch, J. Han, and P. Renaud. Transport phenomena in nanofluidics. *Rev. Mod. Phys.*, 80:839–883, 2008.

[125] P. Mayer, J. Bibette, A. Bancaud, and J. L. Viovy. Separation of DNA using ferrofluid array electrophoresis. *Mater. Res. Soc. Symp. Proc.*, 463:57–67, 1997.

[126] W. D. Volkmuth, T. A. J. Duke, M. C. Wu, R. H. Austin, and A. Szabo. DNA Electrodiffusion in a 2D Array of Posts. *Phys. Rev. Lett.*, 72:2117–2120, 1994.

[127] G. I. Nixon and G. W. Slater. DNA electrophoretic collisions with single obstacles. *Phys. Rev. E*, 50:5033–5038, 1994.

[128] E. M. Sevick and D. R. M. Williams. Motion of a polyelectrolyte chain hooked around a post. *Phys. Rev. E*, 50:R3357–R3360, 1994.

[129] E. M. Sevick and D. R. M. Williams. Collision of a field-driven polymer with a post: Electrophoresis in microlithographic arrays. *Phys. Rev. Lett.*, 76:2595–2598, 1996.

[130] P. M. Saville and E. M. Sevick. Collision of a field-driven polymer with a finite-sized obstacle: A Brownian dynamics simulation. *Macromolecules*, 32:892–899, 1999.

[131] G. C. Randall and P. S. Doyle. Electrophoretic collision of a DNA molecule with an insulating post. *Phys. Rev. Lett.*, 93:58102–58105, 2004.

[132] G. C. Randall and P. S. Doyle. DNA deformation in electric fields: DNA driven past a cylindrical obstacle. *Macromolecules*, 38:2410–2418, 2005.

[133] S. Popelka, Z. Kabatek, J.-L. Viovy, and B. Gas. Peak dispersion due to geometration motion in gel electrophoresis of macromolecules. *J. Chromatogr. A*, 838:45–53, 1999.

[134] M. Kenward and G. W. Slater. Molecular-dynamics simulations with explicit hydrodynamics ii: On the collision of polymers with molecular obstacles. *Eur. Phys. J. E*, 20:125–141, 2006.

[135] E. M. Sevick and D. R. M. Williams. Long-lived states in electrophoresis: Collision of a polymer chain with two or more obstacles. *Eur. Phys. Lett.*, 56:529–535, 2001.

[136] S. P. Holleran and R. G. Larson. Multiple regimes of collisions of an electrophoretically translating polymer chain against a thin post. *Macromolecules*, 41:5042–5054, 2008.

[137] J. M. Deutsch. Theoretical studies of DNA during gel electrophoresis. *Science*, 240:922–924, 1988.

[138] J. M. Deutsch and T. L. Madden. Theoretical studies of DNA during gel electrophoresis. *J. Chem. Phys.*, 90:2476, 1989.

[139] J. M. Deutsch. Explanation of anomalous mobility and birefringence measurements found in pulsed field electrophoresis. *J. Chem. Phys.*, 90:7436–7441, 1989.

[140] K. D. Dorfman. DNA electrophoresis in microfluidic post arrays under moderate electric fields. *Phys. Rev. E*, 73:061922–061931, 2006.

[141] K. D. Dorfman. Erratum: DNA electrophoresis in microfluidic post arrays under moderate electric fields. *Phys. Rev. E*, 77:E019901–019902, 2008.

[142] A. Mohan and P. S. Doyle. Stochastic modeling and simulation of DNA electrophoretic separation in a microfluidic obstacle array. *Macromolecules*, 40:8794–8806, 2007.

[143] J. M. Schurr and S. B. Smith. Theory for the extension of a linear polyelectrolyte attached at one end in an electric field. *Biopolymers*, 29:1161–1165, 1990.

[144] O. Flomenbom and J. Klafter. Single stranded DNA translocation through a nanopore: A master equation approach. *Phys. Rev. E*, 68:041910–041916, 2003.

[145] M. G. Gauthier and G. W. Slater. A Monte Carlo algorithm to study polymer translocation through nanopores. I. Theory and numerical approach. *J. Chem. Phys.*, 128:065103, 2008.

[146] P. André, D. Long, and A. Ajdari. Polyelectrolyte/post collisions during electrophoresis: Influence of hydrodynamic interactions. *Eur. Phys. J. B*, 4:307–312, 1998.

[147] A. Szabo, K. Schulten, and Z. Schulten. First passage time approach to diffusion controlled reactions. *J. Chem. Phys.*, 72:4350–4357, 1980.

[148] R. M. Mnatsakanov. Hausdorff moment problem: Reconstruction of probability density functions. *Stat. Prob. Lett.*, 78:1869–1877, 2008.

[149] J. M. Kim and P. S. Doyle. A Brownian dynamics-finite element method for simulating DNA electrophoresis in nonhomogeneous electric fields. *J. Chem. Phys.*, 125:074906, 2006.

[150] R. M. Jendrejack, J. J. de Pablo, and M. D. Graham. Stochastic simulations of DNA in flow: Dynamics and the effects of hydrodynamic interactions. *J. Chem. Phys.*, 116:7752–7759, 2002.

[151] J. F. Marko and E. D. Siggia. Stretching DNA. *Macromolecules*, 28:8759–8770, 1995.

[152] P. T. Underhill and P. S. Doyle. On the coarse-graining of polymers into bead-spring chains. *J. Non-Newtonian Fluid Mech.*, 122:3–31, 2004.

[153] D. M. Heyes and J. R. Melrose. Brownian dynamics simulations of model hard-sphere suspensions. *J. Non-Newtonian Fluid Mech.*, 46:1–28, 1993.

[154] B. H. Zimm. Size fluctuations can explain anomalous mobility in field-inversion electrophoresis of DNA. *Phys. Rev. Lett.*, 61:2965–2968, 1988.

[155] B. H. Zimm. "Lakes–straits" model of field-inversion gel electrophoresis of DNA. *J. Chem. Phys.*, 94:2187, 1991.

[156] J. O. Tegenfeldt, C. Prinz, H. Cao, R. L. Huang, R. H. Austin, S. Y. Chou, E. C. Cox, and J. C. Sturm. Micro-and nanofluidics for DNA analysis. *Anal. Bioanal. Chem.*, 378:1678–1692, 2004.

[157] A. Cacciuto and E. Luijten. Self-avoiding flexible polymers under spherical confinement. *Nano. Lett.*, 6:901–905, 2006.

[158] E. F. Casassa. Equilibrium distribution of flexible polymer chains between a macroscopic solution phase and small voids. *J. Polym. Sci., Part B: Polym. Lett.*, 5:773–778, 1967.

[159] S. F. Edwards and K. F. Freed. The entropy of a confined polymer. I. *J. Phys. A*, 2:145–150, 1969.

[160] A. Jaeckel and J. Dayantis. Statistics of confined self-avoiding walks. II. Entropy and pressure of confinement. *J. Phys. A*, 27:7719–7731, 1994.

[161] M. R. L. Abadie and J. Dayantis. Monte Carlo simulations of chains confined inside a cube. *Macromol. Theo. Simul.*, 5:93–105, 1996.

[162] A. Grosberg, A. R. Khokhlov, and Y. A. Atanov. *Statistical physics of macromolecules*. American Institute of Physics, New York, 1994.

[163] T. Sakaue and E. Raphael. Polymer chains in confined spaces and flow-injection problems: some remarks. *Macromolecules*, 39:2621–2628, 2006.

[164] J. M. Stouffer and T. J. McCarthy. Polymer monolayers prepared by the spontaneous adsorption of sulfur-functionalized polystyrene on gold surfaces. *Macromolecules*, 21:1204–1208, 1988.

[165] R. L. Jones and R. J. Spontak. Conformational analysis of double-tethered chain molecules at an impenetrable interface: A Monte Carlo study. *J. Chem. Phys.*, 101:5179, 1994.

[166] M. Muthukumar. Polymer escape through a nanopore. *J. Chem. Phys.*, 118:5174, 2003.

[167] P. J. Park and W. Sung. Polymer release out of a spherical vesicle through a pore. *Phys. Rev. E*, 57:730–734, 1998.

[168] D. Stein, F. H. J. van der Heyden, W. J. A. Koopmans, and C. Dekker. Pressure-driven transport of confined DNA polymers in fluidic channels. *Proc. Natl. Acad. Sci.*, 103:15853, 2006.

[169] H. L. Abbott and D. Hanson. A Lattice Path Problem. *Ars Combinatoria*, 6:163–178, 1978.

[170] S. G. Whittington and A. J. Guttmann. Self-avoiding walks which cross a square. *J. Phys. A*, 23:5601–5609, 1990.

[171] M. Bousquet-Mélou, A. J. Guttmann, and I. Jensen. Self-avoiding walks crossing a square. *J. Phys. A*, 38:9159–9181, 2005.

[172] N. Madras. Critical behaviour of self-avoiding walks: that cross a square. *J. Phys. A*, 28:1535–1547, 1995.

[173] M. Lautout-Magat. Contribution to the study of self-avoiding random walks confined to strips and capillaries. *J. Polym. Sci., Polym. Chem. Ed.*, 20:2705–2713, 1982.

[174] T. W. Burkhardt and I. Guim. Self-avoiding walks that cross a square. *J. Phys. A*, 24:L1221–L1228, 1991.

[175] M. T. Batchelor and C. M. Yung. Exact transfer-matrix enumeration and critical behaviour of self-avoiding walks across finite strips. *J. Phys. A*, 27:4055–4067, 1994.

[176] L. Livadaru and H. J. Kreuzer. Confinement of a polymer chain in a tube. *New J. Phys.*, 5:95, 2003.

[177] J. Baschnagel, J. P. Wittmer, and H. Meyer. *Monte Carlo simulation of polymers: coarse-grained models*, volume 23. N. Attig *et al.*, 2004.

[178] K. Kremer and K. Binder. Monte Carlo simulation of lattice models for macromolecules. *Comp. Phys. Rep.*, 7:259–310, 1988.

[179] M. E. J. Newman and G. T. Barkema. *Monte Carlo methods in statistical physics*. Oxford University Press, USA, 1999.

[180] A. van Heukelum and G. T. Barkema. Lattice models of DNA electrophoresis. *Electrophoresis*, 23:2562–2568, 2002.

[181] K. Shida, K. Ohno, and Y. Kawazoe. An improved chain extension algorithm and its application for various branched polymers. *J. Chem. Phys.*, 116:10938, 2002.

[182] P. H. Verdier and W. H. Stockmayer. Monte Carlo calculations on the dynamics of polymers in dilute solution. *J. Chem. Phys.*, 36:227, 1962.

[183] H. J. Hilhorst and J. M. Deutch. Analysis of Monte Carlo results on the kinetics of lattice polymer chains with excluded volume. *J. Chem. Phys.*, 63:5153, 1975.

[184] M. Lax and C. Brender. Monte Carlo study of lattice polymer dynamics. *J. Chem. Phys.*, 67:1785–1787, 1977.

[185] J. M. Hammersley and K. W. Morton. Poor man's Monte Carlo. *J. Roy. Stat. Soc. B*, 16:23–38, 1954.

[186] G. B. Salieb-Beugelaar, K. D. Dorfman, A. van den Berg, and J. C. T. Eijkel. Electrophoretic separation of DNA in gels and nanostructures. *Lab Chip*, 9:2508–2523, 2009.

[187] S. L. Levy and H. G. Craighead. DNA manipulation, sorting, and mapping in nanofluidic systems. *Chem. Soc. Rev.*, 39:1133–1152, 2010.

[188] L. Meistermann and B. Tinland. DNA electrophoresis in a monodisperse porous medium. *Phys. Rev. E*, 62:4014–4017, 2000.

[189] Y. Zeng, M. He, and D. J. Harrison. Microfluidic self-patterning of large-scale crystalline nanoarrays for high-throughput continuous DNA fractionation. *Angew. Chem. Int. Ed.*, 47:6388–6391, 2008.

[190] H. Zhang and M. J. Wirth. Electromigration of single molecules of DNA in a crystalline array of 300-nm silica colloids. *Anal. Chem*, 77:1237–1242, 2005.

[191] J. Noolandi, G. W. Slater, H. A. Lim, and J. L. Viovy. Generalized tube model of biased reptation for gel electrophoresis of DNA. *Science*, 243:1456, 1989.

[192] T. A. J. Duke. Tube model of field-inversion electrophoresis. *Phys. Rev. Lett.*, 62:2877–2880, 1989.

[193] Y. N. Kaznessis. Multi-scale models for gene network engineering. *Chem. Eng. Sci.*, 61:940–953, 2006.

[194] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.*, 22:403–434, 1976.

[195] C. W. Gardiner. *Handbook of stochastic methods*. Springer Berlin, 1985.

[196] E. M. Lifshitz and L. P. Pitaevskii. *Physical Kinetics*. Pergamon Press Oxford, 1981.

[197] M. Muthukumar. Translocation of a confined polymer through a hole. *Phys. Rev. Lett.*, 86:3188–3191, 2001.

[198] R. Kumar and M. Muthukumar. Origin of translocation barriers for polyelectrolyte chains. *J. Chem. Phys.*, 131:194903, 2009.

[199] N. Laachi and K. D. Dorfman. Statistics of tethered self-avoiding chains under spherical confinement and an external force. *J. Chem. Phys.*, 132:084108, 2010.

[200] S. J. Hubert, M. Krzywinski, I. L'Heureux, and G. W. Slater. Reptation dynamics with random local interactions. *Macromolecules*, 31:181–192, 1998.

[201] T. A. J. Duke, R. H. Austin, E. C. Cox, and S. S. Chan. Pulsed-field electrophoresis in microlithographic arrays. *Electrophoresis*, 17:1075–1079, 1996.

[202] S. B. Smith and A. J. Bendich. Electrophoretic charge density and persistence length of DNA as measured by fluorescence microscopy. *Biopolymers*, 29:1167–1173, 1990.

[203] G. W. Slater. Theory of band broadening for DNA gel electrophoresis and sequencing. *Electrophoresis*, 14:1–7, 1993.