INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®

Université d'Ottawa · University of Ottawa

The Electrophoretic Properties of End-Labeled DNA Molecules in Gels, Polymer Solutions and Free-Solutions: A Theoretical and Experimental Study

by

© Claude Desruisseaux

Thesis submitted to the University of Ottawa

in partial fulfilment of the requirements for the

degree of Doctor of Philosophy in Physics

Departments of Physics,

University of Ottawa, Ottawa, Canada

September 1999.

I would like to dedicate this thesis
to Mr. Mini-Wheat for showing me
that it was possible to have
a sweet side and a serious side

"And having heard, or more probably read somewhere, in the days when I thought I would be well advised to educate myself, or amuse myself, or stupefy myself, or kill time, that when a man in a forest thinks he is going forward in a straight line, in reality he is going in a circle, I did my best to go in a circle, hoping in this way to go in a straight line. For I stopped being half-witted and became sly, whenever I took the trouble. And my head was a storehouse of useful knowledge. And if I did not go in a rigorous straight line, with my system of going in a circle, at least I did not go in circle, and that was something. And by going on doing this, day after day, and night after night, I looked forward to getting out of the forest, some day."

Samuel Beckett

Molloy

Sommaire

Nous présentons une étude détaillée de l'électrophorèse sur gel, en solution de polymères diluée et en solution libre, de molécules d'ADN auxquelles nous avons attaché un objet globulaire à un bout (complexe ADN-protéine). De nouveaux modèles théoriques pour le coefficient de friction des molécules hybrides ont été développés pour l'électrophorèse dans ces trois milieux. Des expériences ont aussi été faites dans le but de valider ces nouveaux modèles. L'étude des gradients de champ électrique qui apparaissent aux bouts des gels de polyacrylamide s'est avérée nécessaire pour en éliminer leurs effets néfastes (et ainsi obtenir des résultats expérimentaux reproductibles et utiles). En utilisant cette méthode, il a été possible de vérifier la validité de notre modèle de reptation biaisée. Finalement, une étude théorique, utilisant entre autre des simulations numériques, nous a permis de démontrer qu'il est possible d'avoir un système où deux molécules ayant la même charge (l'ADN et le complexe ADN-protéine) migrent dans des directions opposées quand on applique un champ électrique pulsé de type ZIFE ("zero-integrated field electrophoresis").

Summary

Many aspects of the electrophoretic properties of end-labeled DNA molecules were studied in free-solutions, in dilute polymer solutions as well as in gels. New theoretical models for the friction coefficient of the hybrid molecules (for all three cases) are presented in this thesis. Our experimental investigation allowed us to validate the theoretical approach. A study of the gel edge electric field gradients in denaturing polyacrylamide gels was also performed and a method to take this gradient into account (in order to obtain reproducible and useful results) is presented. Using that method, the effect of steric trapping in polyacrylamide gels could be investigated (for different electric fields and polyacrylamide concentrations) and the validity of our biased reptation model was experimentally verified. Finally, we demonstrated theoretically that it is possible to build a ratchet system where two molecules having the same charge (DNA and end-labeled DNA fragments) can move in opposite directions when electrophoresed in a zero-integrated pulsed electric field.

Acknowledgments

I would like to take the opportunity to thank my supervisors, Gary W. Slater and Guy Drouin, for providing me with the opportunity to do a Ph. D. in Physics and to work in a molecular biology lab (and then have the opportunity to thank them...). Their encouragements and the conversations I had with them in the last years made my life easier and very enjoyable.

I would also like to thank the postdocs, graduate students and honours students who worked in the laboratory (or in front of a computer screen for the physics group). They include Pascal Mayer, Jean Rousseau, Hongji Ren, Tarso Kist, Hong Yan Zhou, Anick DeMoors, Ghislaine Allard, Michael Ell, Mario Moniz de Sá, Jamie Nickerson, Banoo Malick, Robert Beiko, Julie Chapados, Marc Carrier, Shirine Eltaher, Jean-François Mercier, Sylvain Hubert, Grant Nixon, Kathy MacEachern, Claude Raymond. They made working in the lab a lot of fun.

I extend special thanks to Jennifer Jensen for proofreading and for helping me with the English language. I would also like to thank Marianne Fillion-Bergeron for performing some experiments for me in the summer of 1998 (the results of those experiments are presented in Chapter 6). I would also thank Didier Long who helped us to develop the new theories presented in the second and sixth chapters of my thesis. I also thank the University of Ottawa for giving me the SAD award and Perkin-Elmer for providing us the ABI PRISM™ 310 Genetic Analyzer.

Finally I would like to thank my friends and family for their love and for being there for me. I could not have made it without them.

Table of Contents

Chapter 1

Introduction

1.1 General Introduction

Essentially all the information necessary for the development of an organism is encoded in the DNA that constitutes its genome. The field of molecular biology was revolutionized twenty years ago with the possibility of sequencing DNA. Two new methods were then developed, namely the Maxam-Gilbert [1,2] and Sanger [3] methods (in fact, Sanger and Gilbert won the 1980 Nobel prize in Chemistry for this invention). These two methods are based on the fact that one can separate (on polyacrylamide gels) DNA fragments that differ by only one base (one monomer). The principle is quite simple: we load a mixture of DNA fragments at one end of a separation matrix (e.g., a polyacrylamide gel in a slab or capillary format) and we apply an electric field. The DNA molecules then move since they are negatively charged (DNA is an acid and it possesses a negative charge when in solution). Under the proper electrophoretic conditions, DNA molecules move at different speeds depending on their molecular size, such that single-base resolution leads to DNA sequencing (see Figure 1). Electrophoresis and molecular separations greatly interest physicists: Why do DNA molecules migrate at different speeds v? What new separation mechanisms can be used in order to have better/faster separation of DNA fragments? Can we automate the process? etc. This thesis

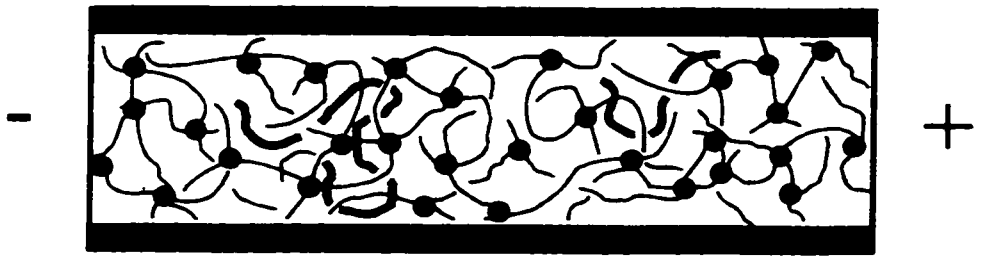


Figure 1: Schematic representation of the separation of DNA molecules (in red) during a gel electrophoresis experiment. The molecules are moving from the left to the right and the small molecules are faster than the large ones. The gel is made of crosslinked polymer chains (in black).

studies different fundamental and applied aspects of DNA electrophoresis in gels and free-solutions.

1.2 Trapping Electrophoresis

Over the last decade, a substantial amount of money has been invested by the DOE and the NIH (and various granting agencies in Canada and the other industrialized contries) to encourage scientists to improve the methods used to sequence DNA. Different strategies have been used in order to increase the sequencing power of these methods, but most of the time the basic principles are essentially the same as twenty years ago. Fortunately, some groups are now developing very original ways to separate DNA molecules. One such promising scheme was suggested by Ulanovsky, Drouin and Gilbert (UDG) in 1990 [4]. The idea is to attach streptavidin (a globular protein of radius $R_S \cong 2.5$ nm [5]) to one end of DNA fragments and to electrophorese these hybrid molecules in polyacrylamide gels. This new separation process is called "Trapping Electrophoresis" (TE) because the hybrid molecule (called the streptavidin-DNA complex, or S-DNA) may get sterically trapped when the unlabeled head of the DNA enters a pore whose radius (a) is smaller than that (R_S) of the streptavidin label (Figure 2). The electric force pulling on the S-DNA molecule keeps it trapped in this state until a thermally activated backward "jump" makes the leading head of the DNA disengage from the narrow pore and choose a wider path. Because the depth of the trap is related to the net electric force pulling on the molecule (the latter is proportional to the applied electric field E as well as to the charge Q which increases linearly with the molecular size M (in bases) of the DNA molecules), larger DNA fragments should be trapped more severely than smaller ones; therefore, we

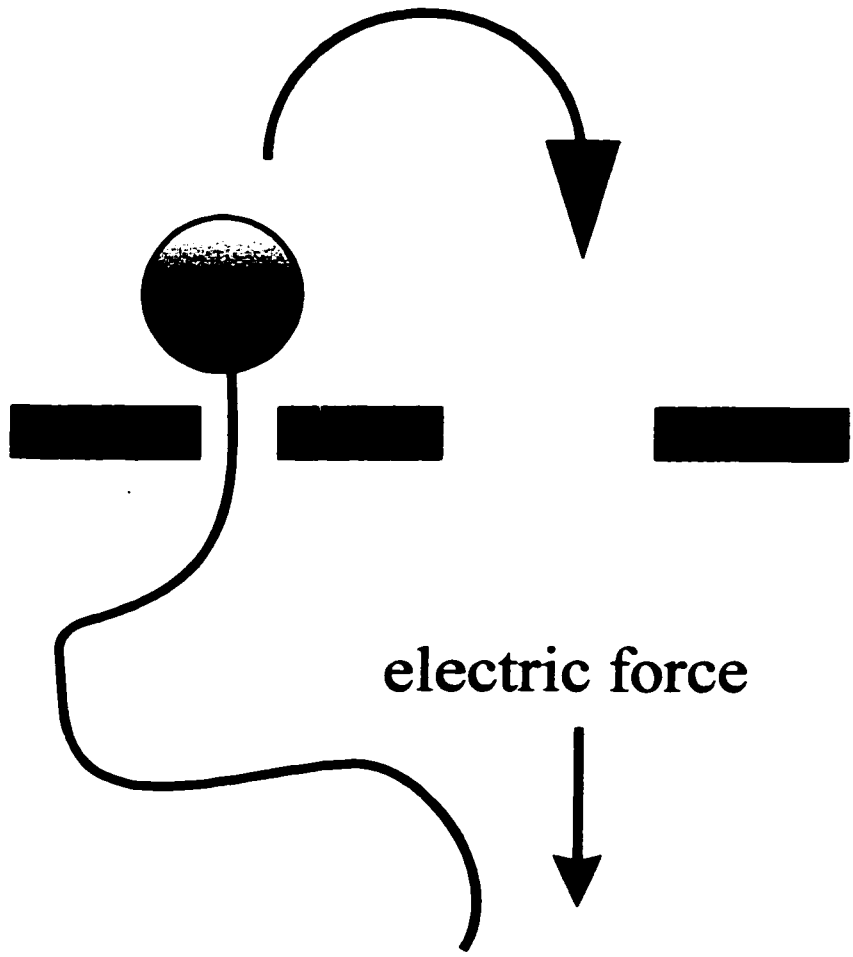


Figure 2: Schematic representation of an end-labeled DNA molecule in a trap. The molecule must move backward to choose a wider pore in order to continue its progression in the gel.

expect an abrupt decrease of the electrophoretic mobility $\mu=v/E$ beyond a certain critical molecular size M*. Experimentally, one does indeed observe a very abrupt (exponential) decrease in the mobility beyond a certain critical molecular size $M^* \sim E^{-2/3}$, as we shall see later.

When long naked DNA molecules are separated in gels, on the other hand, we observe either the reptation of random coil fragments (for small molecules $M<M_0(E)$) or the reptation of longer $M>M_0(E)$ fragments oriented along the field direction. In these two limits, the biased reptation model predicts that the mobility of a fragment of size M in a field E should scale like [6-10]:

$$\frac{\mu}{\mu_0} \sim \frac{1}{\min\{M,\ M_0(E)\}} \qquad (1)$$

where the critical size $M_0(E) \sim E^{-\delta}$, with $2 \geq \delta \geq 0$ and μ_0 is the mobility of DNA in free-solution (independent of the molecular size). For small sizes $M<M_0$, the 1/M scaling law provides excellent separations, in agreement with experimental data [8-10]. In the opposite limit $M>M_0$, size separation is impossible since the mobility plateaus at a field-dependent value $\mu \sim 1/M_0 \sim E^{\delta}$. Clearly, the molecular orientation that leads to the latter regime is a major nuisance. This effect has been studied extensively, and is due to the fact that the external field biases the direction taken by the DNA-snake as it migrates through the gel [11-13].

Remarkably, the TE critical molecular size for trapping is such that $M^*<M_0$ [14]. Consequently, we do not observe the technology-limiting mobility plateau for the end-labeled DNA fragments. This is huge advantage over ordinary DNA gel electrophoresis and this is what got everyone excited after the UDG article was published. We will show in Chapter 5 that even though

the mobility plateau is never reached and the peak spacing is increased with TE, there is in fact a loss of resolution due to the nature of the detrapping mechanism (i.e., the diffusion coefficient is greatly increased).

1.3 Charged Interfaces

The inner wall of a fused silica capillary is negatively charged, due to the dissociation of silanol groups, when in contact with standard buffers. The wall then attracts cations which form the so-called double-layer, a thin layer of cations of thickness $\lambda_D \approx 1\text{-}10$ nm, the Debye length. In the presence of an electric field, the diffuse part of this layer moves and drags the whole liquid towards the cathode: this is the electroosmotic flow, or EOF [15] (Figure 3). It is often preferable to suppress EOF for DNA applications. For example, EOF may not be constant along the capillary, and the resulting axial flow gradient may affect resolution. More importantly, however, one must take EOF into account in order to test theories since the apparent mobility is given by $\mu = \mu_{electrophoretic} + \mu_{EOF}$. The fixed (and negative) μ_{EOF} contribution can in principle be measured directly using, e.g., an uncharged marker. Several buffer additives and capillary wall coating agents (covalent or dynamic) have been proposed in order to eliminate EOF [16].

But what happens to DNA during free-solution electrophoresis? When the electric field E=0, collective hydrodynamic effects make the DNA coil act like an impermeable sphere (Figure 4A) with a radius of gyration $R_G \sim M^{1/2}$ and a friction coefficient $\xi \sim R_G \sim M^{1/2}$; this is the so-called Zimm regime [17]. However, much like the walls, DNA is negatively charged and attracts a cloud of counter-ions

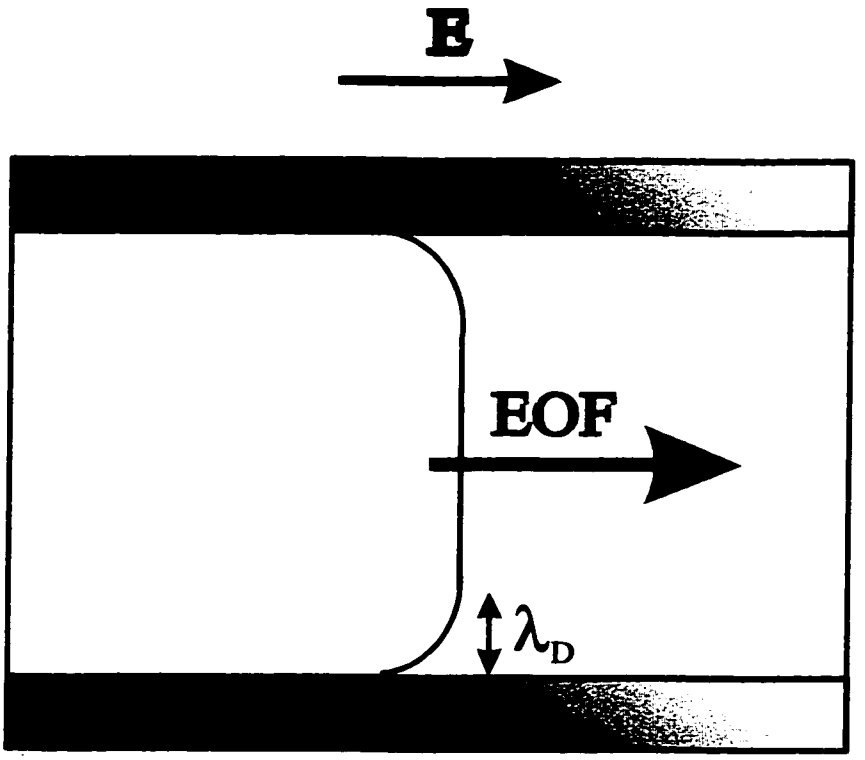


Figure 3: Schematic representation of the electroosmotic flow (EOF) in a capillary. The bulk of the fluid is dragged because of the electrophoretic movement of the cations near the capillary walls (the Debye layer λ_D).

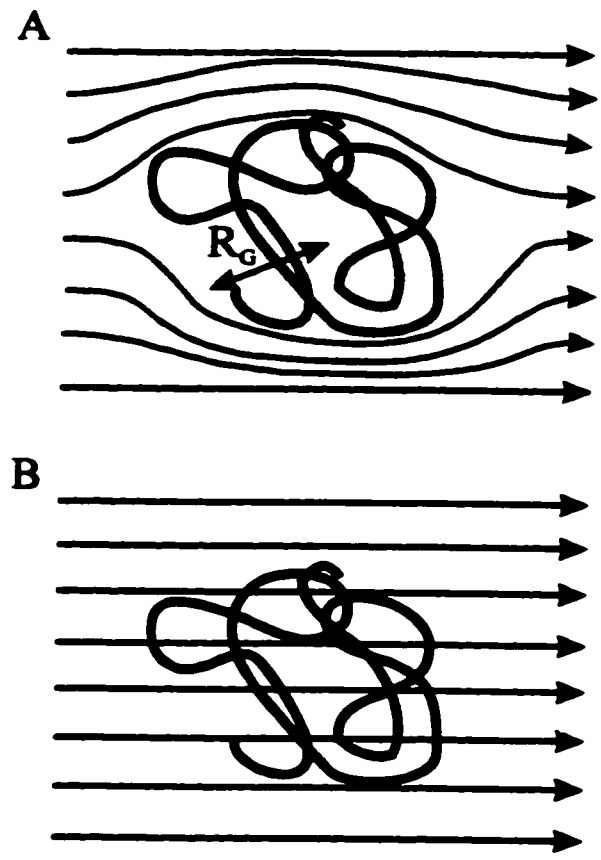


Figure 4: A random coil DNA molecule with a radius of gyration R_G is moving in a fluid. (A) In absence of an electric field, the hydrodynamic interactions between the different parts of the polymer make the coil move like an impermeable sphere of radius R_G. (B) During electrophoresis, the counter-ions screen the hydrodynamic interactions and the flow penetrates the random coil.

in its vicinity. When $E \neq 0$, the DNA and the cations move in opposite directions. Moreover, the hydrodynamic interactions between the different parts of the DNA molecule are then screened over distances larger than λ_D. This screening kills the collective (Zimm) effects inside the DNA coil, and the friction coefficient now scales like $\xi \sim M$. Since the mobility $\mu(M)=Q(M)/\xi(M)$, where $Q \sim M$ is the charge of the DNA molecule, the resulting mobility is independent of the DNA size M! This is the famous (an electrophoretically unfavorable) free-draining property of DNA (Figure 4B). Size-dependent mobilities are sometimes observed when the ionic strength is too weak to hinder hydrodynamic interactions (giving $\lambda_D > R_G$), but this is an extreme case of little practical value. The current use of sieving matrices is due to this microscopic phenomenon.

1.4 End-Labeled Free-Solution Electrophoresis

Because of the free-draining property of DNA described above, it is not possible to separate DNA molecules in free solution. A very interesting scheme was thus proposed by Mayer et al. (and later confirmed experimentally by Heller et al. [18] and Ren at al. [19]) in order to make separation possible without sieving media. The idea here is to label one end of the DNA fragments with a neutral object (e.g., streptavidin again!) which provides extra friction but no charge, hence the name End-Labeled Free-Solution Electrophoresis (ELFSE). Because this affects only the denominator of the ratio $\mu=Q/\xi$, the mobility μ becomes size-dependent and size separation becomes possible. In most cases, the mobility μ of the end-labeled DNA fragment is then given as:

$$\mu(M) = \mu_0 \times \frac{M}{M + \alpha} \qquad (2)$$

where μ_0 is the size-independent free mobility of DNA and α is the label's friction coefficient (relative to the friction coefficient of one DNA monomer). Although it was demonstrated that the ELFSE concept is experimentally valid and viable, the process itself has never been understood and studied at the microscopic level. Such a fundamental study is presented in Chapters 2 and 3 of this thesis.

1.5 The Properties of DNA and the Relevant Length Scales

DNA is obviously an important biological molecule. In fact, it may be the most important molecule of the universe! It would be easy to give the reader a lot of detailed information about the structure of the DNA molecule or the chemical structure of the four nucleotides. However, for a physicist (like me), the DNA molecule is simply a charged linear worm-like polymer (a polyelectrolyte). Chemical details do not affect the electrophoretic properties in most cases. This section will give the reader the information necessary to understand the physical and electrophoretic properties of DNA.

The four nucleic acids (Adenine (A), Cytosine (C), Guanine (G) and Thymine (T)) are the building blocks of DNA. For single-stranded DNA (i.e., each strand of denatured DNA; this is denoted ssDNA), the distance between two consecutive monomers is given by b=0.43 nm [20]. DNA can then be seen as a flexible cylinder (or worm) of contour length L=Mb (M is the number of monomers, often called the molecular size) and radius ≈b.

1.5.1 The Persistence Length of DNA

The radius of gyration of a polymer is an important quantity since it represents the space occupied by the polymer. The radius of gyration R_G is defined as:

$$R_G^2 \equiv \frac{1}{M} \sum_{i=1}^{M} < (\vec{R}_i - \vec{R}_{CM})^2 > \quad (3)$$

where R_{CM} is the position of the center of mass of the polymer. The DNA fragments adopt a random coil conformation that looks like the path adopted by a particle that moves by Brownian motion (i.e., a random walk conformation). In polymer science, the stiff/flexible properties of DNA are summarized by the Kratky-Porod equation [17]:

$$R_G^2(M) \approx \frac{pL}{3} \left[1 - 3\left(\frac{p}{L}\right) + 6\left(\frac{p}{L}\right)^2 - 6\left(\frac{p}{L}\right)^3 (1 - e^{-L/p}) \right] \quad (4)$$

where p is the so-called persistence length of DNA. A very small DNA fragment is a stiff rod; in this case, eq 4 correctly predicts that the radius of gyration is given by:

$$R_G(M) = \frac{L}{\sqrt{12}} \quad (5)$$

Now, let's look at two consecutive monomers (the j^{th} and the $j+1^{th}$ monomers) of a long DNA fragment. The directions of these two DNA fragments, $\mathbf{u}_j$ and $\mathbf{u}_{j+1}$, are strongly correlated so that $\langle \mathbf{u}_j \cdot \mathbf{u}_{j+1} \rangle = \langle \cos\theta_{j,j+1} \rangle \cong 1$. In fact, the correlation between the orientation of monomers that are separated by a contour length s is given by $\langle \cos\theta(s) \rangle = \exp(-s/p)$, where the persistence length p is a constant for

each polymer [17]. Therefore, p is the basic parameter that characterizes polymer flexibility. According to eq 4, the radius of gyration for DNA fragments of contour length L=Mb»p is given by

$$R_G^2 \approx \frac{p\,M\,b}{3} \tag{6}$$

as for a normal random-walk with a step size b_k=2p, the so-called Kuhn length. As we can see, the Kratky-Porod equation has the right limit for small (i.e., rigid or L<p) and large (flexible or L>2p) DNA fragments.

In most relevant cases, the persistence length of ssDNA is p≈1-10 nm. The persistence length of DNA depends on the ion concentration of the buffer solution. The electrostatic repulsion due to the charges on the backbone of the DNA molecule tends to make the DNA molecule quite stiff. When we add ions to the buffer, however, there is some screening of the electrostatic repulsion and we observe a reduction of the persistence length. The recent theory of Ha and Thirumalai [21] predicts that

$$p \approx p_0 + \frac{A}{I} \tag{7}$$

where A is a constant, I is the ionic strength of the buffer solution (related to the quantity of ions in solution) and p_0 is the intrinsic persistence length of DNA (the persistence length when I→∞).

1.5.2 The Structure of Polyacrylamide Gels

Traditionally, DNA electrophoretic sequencing takes place in polyacrylamide gels. Polyacrylamide gels are chemical gels made of a mixture of two monomers: acrylamide and N,N'-methylenebisacrylamide (or bisacrylamide). The first one constitutes the gel fibers, while the second one is the crosslinker (see Figure 1 and 5). The space between the polyacrylamide fibers is called the "pore". When DNA migrates into the polyacrylamide gel, it migrates through the "loops" formed by the polyacrylamide network. The mean pore size (a) was studied by Rousseau et al. [22] and is approximately given by $a \approx 29.5\ \text{nm}/(\%T)^{0.75}$ (with $3<\%T<12$), where %T is the concentration of the polyacrylamide gel (in %w/v). As expected, denser gels have smaller pore sizes.

1.5.3 The Streptavidin Label

The protein (the label) that we attach to one end of our ssDNA molecules is streptavidin. This protein is essentially a sphere of radius $R_S \approx 2.5$ nm [5]. The size of this label is not much smaller than the size of the polyacrylamide mean pore size a. This explains why there is steric trapping when the S-DNA molecules migrate in these gels (Figure 2). Moreover, the size of the label is very close to the persistence length p of ssDNA.

1.6 Motivation for this Work

Why study the migration of a hybrid molecule composed of a ssDNA fragment and a protein? The electrophoretic properties of this molecule are very different from the properties of naked DNA, hence the possibility of exploiting these differences to design more performant sequencing methods.

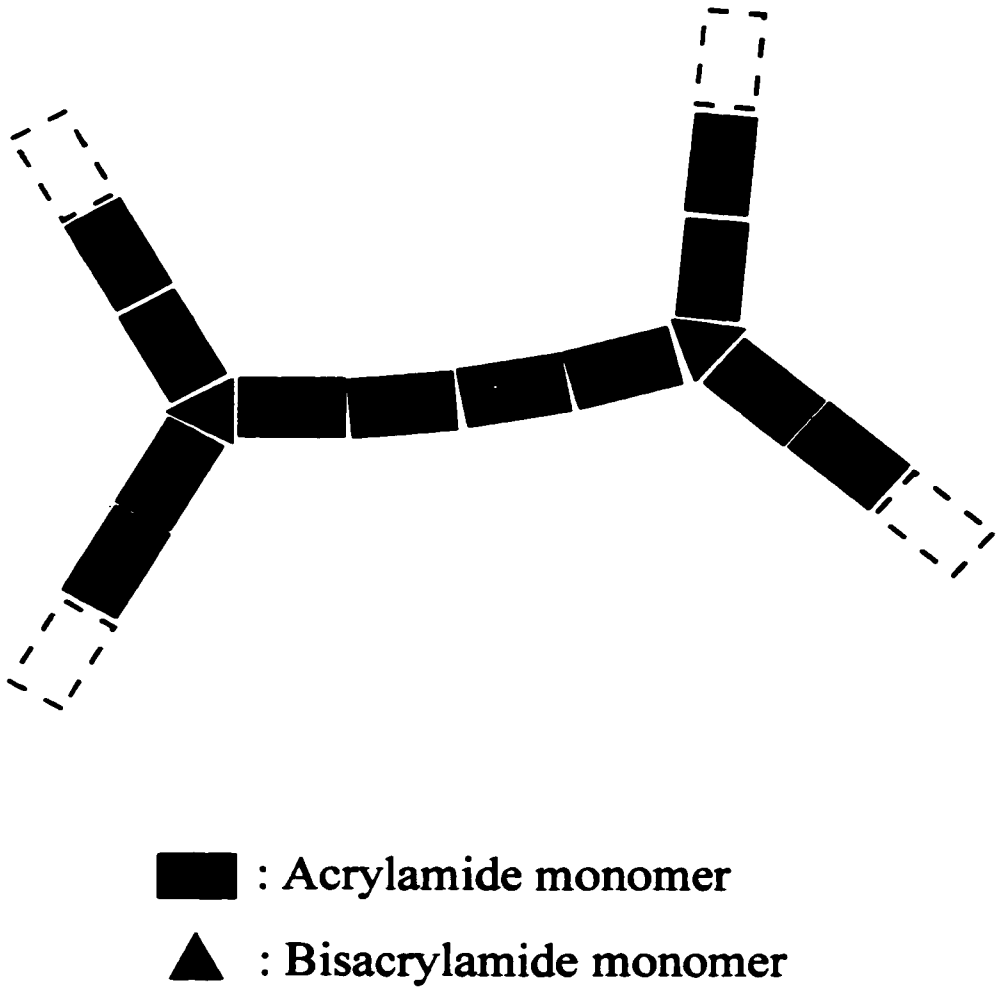


Figure 5: Schematic representation of crosslinked polyacrylamide. The rectangles represent acrylamide monomers while the triangles represent bisacrylamide (the crosslinker).

It was a challenge to study the electrophoretic properties of these molecules since very little was known about these issues when I started my Ph. D. program. It also represented a great opportunity to develop theories and obtain experimental confirmation in the laboratory myself. It is unfortunately very rare nowadays that researchers can enjoy the freedom to do the kind of work that I did over the last four years, i.e. both experimental and theoretical work using the tools of physics, analytical chemistry and molecular biology. This is mostly due to the existence of a somewhat arbitrary segregation between theoricians and experimentalists. Thus, this thesis is a mixture of theoretical and experimental studies of the dynamical properties of the streptavidin-DNA complex in a variety of electrophoretic systems.

1.7 References:

[1] Maxam, A. M.; Gilbert, W.; *Proc. Nat. Acad. Sci. USA*, **74**, 560, 1977.

[2] Maxam, A. M.; Gilbert, W.; *Methods Enzymol.*, **65**, 499, 1980.

[3] Sanger, F.; *Science*, **214**, 1205, 1981.

[4] Ulanovsky, L.; Drouin, G.; Gilbert, W.; *Nature*, **343**, 190, 1990.

[5] Hendrickson, W. A.; Pähler, A.; Smith, J. L.; Satow, Y.; Merritt, E. A.; Phizackerley, R. P.; *Proc. Nat. Acad. Sci. USA*, **86**, 2190, 1989.

[6] Duke, T. A. J.; Viovy, J. L.; Semenov, A. N., *Biopolymers*, **34**, 239, 1994.

[7] Semenov, A. N.; Duke, T. A. J.; Viovy, J. L.; *Phys. Rev. E*, **51**, 1520, 1995.

[8] Heller, C.; Duke, T. A. J.; Viovy, J. L.; *Biopolymers*, **34**, 249, 1994.

[9] Viovy, J. L.; Mechanisms of polyelectrolyte gel electrophoresis. (Submitted for publication to *Rev. Mod. Phys.*), 1999.

[10] Slater, G. W.; Electrophoresis theories, in *Analysis of nucleic acids by capillary electrophoresis* (Heller, C., ed.), Vieweg & Son, Wiesbaden, pp.24-66, 1997.

[11] Barron, A. E.; Blanch, H. W.; *Separation and Purification Methods*, **24**, 1, 1995.

[12] Mayer, P.; Slater, G. W.; Drouin, G.; *Appl. Theor. Electroph.*, **3**, 147, 1993.

[13] Lumpkin, O. J.; Déjardin, P.; Zimm, B. H.; *Biopolymers*, **24**, 1573, 1985.

[14] Slater, G. W.; Villeneuve, C.; *J. Polym. Sci. B:Polym. Phys.*, **30**, 1451, 1992.

[15] Grossman, P. D.; Factors affecting the performance of capillary electrophoresis separations: Joule heating, electroosmosis, and zone dispersion, in *Capillary Electrophoresis Theory &*

Practice: (Grossman, P. D. and Colburn, J. C., eds.), Academic Press, San Diego, pp.3-43, 1992.

[16] Chiari, M.; Nesi, M.; Righetti, P. G.; Surface Modification of Silica Walls: A Review of Different Methodologies, in *Capillary Electrophoresis in Analytical Biotechnology*, (Righetti, P. G., ed.), CRC Press, Boca Raton, 1996.

[17] Doi, M.; Edwards, S. F.; The Theory of Polymer Dynamics; Clarendon Press, Oxford, 1986.

[18] Heller, C.; Slater, G. W.; Mayer, P.; Dovichi, N. J.; Pinto, D.; Viovy, J. L.; Drouin, G.; *J. Chromatogr, A*, **806**, 113, 1998.

[19] Ren, H.; Karger, A. E.; Oaks, F.; Menchen, S.; Slater, G. W.; Drouin, G.; *Electrophoresis*, **20**, 2501, 1999.

[20] Record, M. T., Jr; Anderson, C. F.; Lohman, T. M.; *Quart. Rev. Biophys.*, **11**, 103, 1978.

[21] Ha, B.-Y.; Thiramalai, D.; *J. Chem. Phys.*, **110**, 7533, 1999.

[22] Rousseau, J.; Slater, G. W.; Drouin, G.; *Phys. Rev. Lett.*, **79**, 1945, 1997.

Chapter 2

Free-Solution Electrophoresis of Composite Molecular Objects: the Relation between Friction, Charge and Ionic Strength*

In this chapter, we study the effect of the persistence length of single-stranded DNA (ssDNA) on the migration of end-labeled DNA fragments in free solution capillary electrophoresis (CE). Our data show that the relative electrophoretic friction coefficient of our (streptavidin) label depends on the buffer's ionic strength because it is related to the persistence length of the DNA fragment. Therefore, the ionic strength affects not only the free-solution mobility of naked DNA, but also the effective electrophoretic frictional drag generated by the globular label. Using a blob model of the hybrid molecule, we can extract the ionic strength dependence of the DNA persistence length from our mobility measurements. Our results are in agreement with recent polyelectrolyte theories. The data can be explained without using the concept of an effective DNA charge, thus showing that the concept of charge screening is not relevant in electrophoresis.

---

* The results presented in this chapter will be submitted for publication before the end of the current year.

2.1 Introduction

DNA is a free-draining polyelectrolyte whose electrophoretic mobility does not normally depend upon its contour length [1]. This is why DNA separation cannot be obtained in free-solution electrophoresis. However, two methods to separate DNA using free-solution have recently been proposed. The first method requires uncharged labels [2] whereas the second one uses the electroosmotic flow (EOF) [3]. Here, we examine the first method, called End-Labeled Free-Solution Electrophoresis (ELFSE), and the physics underlying ELFSE separations [4,5]. The first paper about ELFSE [2] made predictions about the resolution of this new method using a very simple analytical model with empirical parameters that were rather ill-defined. The present chapter bridges the gap between the original ELFSE concept, a new theoretical model of electrophoresis [6-10], recent data [5], and our new experimental results.

The charge Q of a DNA fragment is directly proportional to its molecular size because all monomers are equally charged. In most practical cases, the electrophoretic drag of the DNA coil is essentially that of a cylinder having the same contour length as the DNA fragment. Since the molecular size dependence of the friction coefficient is then also directly proportional to the molecular size of the DNA fragment, the net electrophoretic mobility is size-independent. Labeling the DNA fragments with a different type of molecule, the key idea of ELFSE, can break this charge/friction symmetry and make separation possible. In this chapter, we will limit ourselves to the case of the streptavidin-DNA complexes used by Heller et al. [4] and Ren et al. [5]. We will consider that the residual charge of the streptavidin label is exactly zero in order to facilitate data analysis (recent ELFSE results indicate that it has a very small charge [Ren et al., to be submitted]).

We will also deliberately avoid discussing labels that can be deformed during electrophoresis since this leads to quite different (and exciting!) situations.

The feasibility of ssDNA sequencing using ELFSE, first discussed in 1994 [2], was recently demonstrated by Ren et al. [5]. In the original theoretical treatment of this problem, the limitations of ELFSE as a sequencing technique were estimated assuming the absence of electrostatic and hydrodynamic interactions between the DNA and the friction generating label, and using a local model of free-flow electrophoresis. In the light of the new theoretical approach for the electrophoretic migration of composite molecular objects developed by Long et al. [6-10], the theory of ELFSE must be revisited. This chapter is not intended to explain every aspect of the new theory, even though we did include enough details to make the Chapter self-contained; instead, our aim here is to describe new experimental results that can be used to test this new theory. The concepts of hydrodynamic friction coefficients, DNA stiffness and ionic strength will be discussed and we will show how they are intimately related. One cannot understand (or optimize) the ELFSE separation process without a better and more fundamental understanding of those key concepts.

2.2 Theory

2.2.1 DNA Flexibility, Debye Length and Friction Coefficients

A DNA fragment is a charged polymer chain which adopts a random coil conformation whose size (its radius-of-gyration R_G) depends on its persistence length p. The ionic strength (I) of the buffer solution affects the persistence length since screening of the electrostatic repulsion is more efficient at high ionic strengths [11-13]. This explains why DNA is more rigid in a low ionic strength

buffer. The stiffer the DNA, the more volume it occupies. In the absence of an electric field, the Zimm friction coefficient $\xi_{DNA}(M)$ of a M-base long DNA molecule is given by [14]:

$$\xi_{DNA}(M) = 6 \pi \eta R_H(M) \tag{1}$$

where η is the buffer viscosity and $R_H(M)$ is the hydrodynamic radius of the molecule. For a Gaussian coil one has $R_H \approx \frac{2}{3} R_G$ [14], and $R_H(M)$ can then be estimated using the Kratky-Porod equation [14]:

$$R_H^2(M) = \left(\frac{2}{3}\right)^2 R_G^2(M) \approx \left(\frac{2}{3}\right)^2 \frac{pL}{3}\left[1 - 3\left(\frac{p}{L}\right) + 6\left(\frac{p}{L}\right)^2 - 6\left(\frac{p}{L}\right)^3\left(1 - e^{-L/p}\right)\right] \tag{2}$$

where L=Mb is the contour length of the DNA molecule and b is the size of the monomer (b=0.43nm for ssDNA). This equation has the right scaling for low ($R_H \sim L$) and high ($R_H \sim L^{1/2}$) molecular weights, but it does not take into account the excluded volume effects and the logarithmic correction factors in the limit of stiff molecules. For ssDNA sequencing, the molecular size of the DNA fragments is generally moderate (i.e., $L>p$ but $L \not> p$); excluded volume interactions can then be neglected and the utilization of the Kratky-Porod equation is justified [15].

The negatively charged DNA monomers are surrounded by a cloud of cations. The thickness of this cloud is called the Debye length and it is given by [16]:

$$\lambda_D = \sqrt{\frac{\epsilon_b \epsilon_0 k_B T}{2 e^2 I}} \tag{3}$$

where ϵ_b is the dielectric constant of the solvent, ϵ_0 is the permittivity of the vacuum, k_B is the Boltzmann constant, T is the absolute temperature, e is the charge of an electron and I, the ionic strength of the buffer, is defined as:

$$I = \frac{1}{2} \sum_k z_k^2 C_k \tag{4}$$

(the sum is over the ionic species k, with z_k and C_k being the ion's valence and concentration (mol/L), respectively).

When an electric field E is applied, the Debye layer of cations directly affects the friction coefficient of the DNA because the DNA and its counter ions move in opposite directions generating extra friction [6]. There are in fact four important length scales in this system: the monomer size b, the Debye length λ_D, the persistence length p and the DNA radius-of-gyration $R_G(M)$. The contour length L=Mb is related to $R_G(M)$ and p via the Kratky-Porod relation (eq 2). For most practical cases, we have $R_G(M)>p>b$, i.e. the DNA molecule is a moderately long chain with a stiff backbone. We then have four possible regimes depending on the relative value of λ_D and the other length scales.

When the Debye length is larger that the radius of gyration of DNA, $\lambda_D>R_G(M)$ (Figure 1A), the electrophoretic friction coefficient is still given by the Stoke relation $\xi_{DNA}(M)=6\pi\eta R_H(M)$ and the resulting electrophoretic mobility $\mu=Q/\xi$ is given by (note that Q=Me):

$$\mu_{DNA}(M) = \frac{M\ e}{6\ \pi\ \eta\ R_H(M)} \tag{5}$$

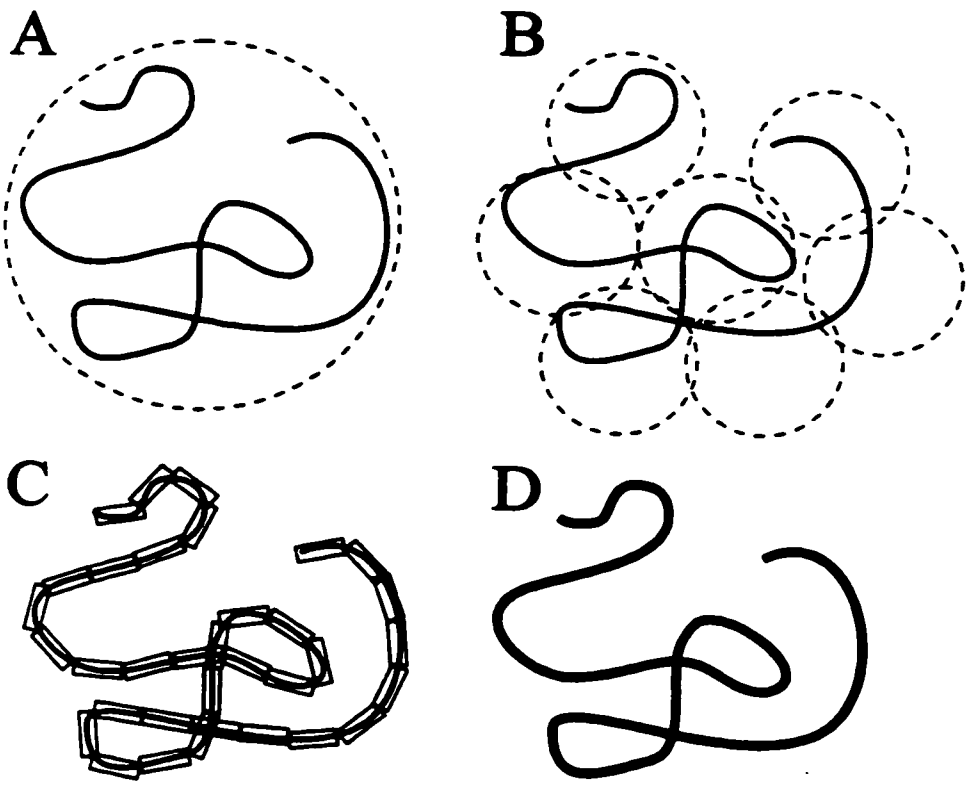


Figure 1 Schematic representation of the concept of the Debye length λ_D for a polyelectrolyte. The four different regimes shown here are: A) $\lambda_D > R_G > p > b$; B) $R_G > \lambda_D > p > b$; C) $R_G > p > \lambda_D > b$; and D) $R_G > p > b \approx \lambda_D$. In each case, the independent hydrodynamic blobs of size λ_D are shown.

In this limit (very low ionic strengths and/or small DNA molecules), it is sometimes possible to separate DNA fragments [17]. Unfortunately, such conditions are rather extreme and quite far from standard DNA electrophoresis conditions.

In the three other cases (Figures 1B, 1C and 1D), a DNA molecule has the same mobility as a DNA blob (or cylinder) of size λ_D because these blobs are hydrodynamically independent (collective effect are screened out over distances larger than λ_D). Therefore the mobility becomes independent of the DNA size M. In the case where $p<\lambda_D<R_G(M)$ (Figure 1B), for instance, the mobility is given by:

$$\mu_{DNA}(M) = \mu_0 \approx \frac{e\,\lambda_D}{2\,\pi\,\eta\,b\,p} \qquad (6)$$

since the charge of each of the R_H^2/λ_D^2 blobs is $Q_D \approx 3e\lambda_D^2/pb$ while their electrophoretic friction coefficient is $\xi_D=6\pi\eta\lambda_D$. This equation is valid only if λ_D is large enough to have Gaussian blobs.

When the DNA is rigid vs. the blob size ($p>\lambda_D$, Figure 1C), the net mobility is that of a cylinder of length λ_D and radius $\approx b$. The charge of those effective DNA segments is $Q_D \approx e\,\lambda_D/b$ while their electrophoretic friction is $\xi_D \approx 3\pi\eta\,\lambda_D/\ln(\lambda_D/b)$ so that the electrophoretic mobility becomes:

$$\mu_{DNA}(M) = \mu_0 \approx \frac{e}{3\,\pi\,\eta\,b}\,\ln\left(\frac{\lambda_D}{b}\right) \qquad (7)$$

This can be expressed as a function of the ionic strength:

$$\mu_{DNA}(M) = \mu_0 \approx \frac{e}{3\,\pi\,\eta\,b} \ln\left(\sqrt{\frac{I_0}{I}}\right) \qquad (8)$$

where the parameter $I_0=\epsilon_b\epsilon_0 k_B T/2e^2b^2$ has units of mol/L. We note that the dependence upon the ionic strength I is a lot weaker than in eq 6; in fact, it comes entirely from the logarithmic correction factor to the friction coefficient of a rod-like object [18]. The extreme case $\lambda_D<b$ (Figure 1D) should lead to a mobility that is independent of the ionic strength since the Debye length ceases to play a role in this limit.

The Debye length for DNA in an aqueous monovalent buffer with ionic strength I is $\lambda_D \simeq 3.0\text{Å}\ I^{-1/2}$ at 30°C [16]. This value is small compared to the size of a DNA coil so that, for most practical cases, DNA is surrounded by a thin Debye layer and all DNA molecules have the same free mobility μ_0 (given by either eq 6 or 8).

2.2.2 Composite Objects, DNA Conformations and Hydrodynamics

DNA normally adopts a random coil conformation in free solution, even during electrophoresis, because of the local cancellation of the electric and frictional forces [6]. When streptavidin is attached at one end of a DNA fragment, the electrophoretic velocity is reduced by a factor $\Delta v=(\mu_{DNA}-\mu_{S\text{-}DNA})\times E$, where E is the electric field intensity and the subscript S-DNA refers to the streptavidin-DNA composite molecule. Long et al. [8] demonstrated that the conformation of the DNA fragment should then be the same as if this naked fragment were pinned at the streptavidin

end and subjected to a solvent flow of velocity Δv (much like in the optical tweezer experiments of Perkins et al. [19]). The friction coefficient of the DNA fragment, ξ_{DNA}, is then a function of the counter flow Δv, as described by Brochard-Wyart [20], because the DNA coil may deform.

The mobility of composite objects such as the streptavidin-DNA hybrid molecules can be calculated using the formalism of Long et al. [6-10]. These authors showed that the net mobility $\mu_{S\text{-}DNA}$ is not given by the total charge of the object divided by the total friction coefficient, but rather by an average of the free-solution mobilities of the different elements weighted by their friction coefficients (the subscript S refers to the streptavidin label):

$$\mu_{S\text{-}DNA} \approx \frac{\mu_{DNA}\,\xi_{DNA} + \mu_S\,\xi_S}{\xi_{S\text{-}DNA}} \tag{9}$$

Note that $\xi_{S\text{-}DNA}=\xi_{DNA}+\xi_S$ only in the limit where there is no hydrodynamic interaction between the DNA and the label (i.e., if there is physical segregation of the two components) [9].

In the limit of small "flow" velocities Δv, the conformation of the complex is an undeformed coil as if there were no electric field (Figure 2A). In this case, the streptavidin is part of the coil (i.e., there is no segregation). We can deal with this molecule as if it were a polyampholyte containing N-1 DNA blobs and one streptavidin blob. The blobs are defined in such a way that they all have the same friction coefficient ξ_S. If each blob contains α bases of DNA, the mobility of the S-DNA complex is given by:

$$\mu_{S\text{-}DNA}(M) \approx \frac{\mu_0}{1 + \frac{\alpha}{M}} \tag{10}$$

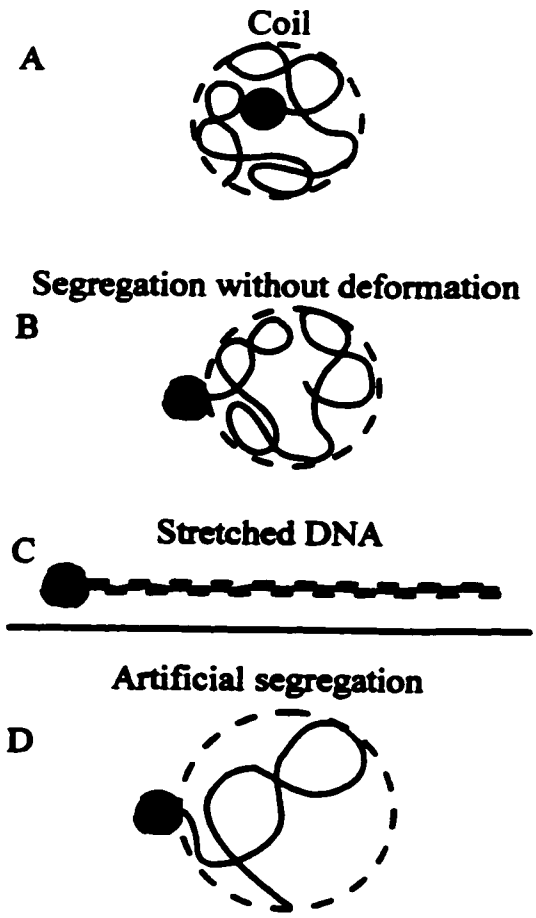


Figure 2 Schematic representation of the possible hydrodynamic conformations of the streptavidin-DNA complex. A) The streptavidin and the DNA form one hydrodynamic entity, and the streptavidin label is "hidden" inside the DNA coil. B) There is segregation of the DNA and streptavidin components, but the DNA is still not deformed. C) Segregation and DNA stretching occur. D) Artificial segregation may be obtained for short enough DNA molecules when the persistence length is large compared to the size of the label.

where μ_0 is the free mobility of DNA (α=0). Since α is the molecular size of a DNA blob of hydrodynamic radius $R_H(\alpha \text{ bases})=R_S$, the value of the parameter α is directly related to the persistence length of DNA (p) via the Kratky-Porod equation:

$$R_s^2 = \left(\frac{2}{3}\right)^2 \frac{b\,\alpha\,p}{3}\left[1 - 3\left(\frac{p}{b\,\alpha}\right) + 6\left(\frac{p}{b\,\alpha}\right)^2 - 6\left(\frac{p}{b\,\alpha}\right)^3\left(1 - e^{-b\,\alpha/p}\right)\right] \qquad (11)$$

where the radius of the streptavidin label is R_S=2.5 nm. Note that for the streptavidin-DNA complex, there is in fact no valid solution to this equation (for the persistence length p) when α<30.21. This critical value is obtained when the DNA backbone is so rigid that $p>R_S$.

If there is hydrodynamic segregation of the two components of our complex, we can use the relation:

$$\mu_{S\text{-}DNA} = \frac{\mu_{DNA}\,\xi_{DNA} + \mu_S\,\xi_S}{\xi_{DNA} + \xi_S} \qquad (12)$$

In the limit where we have segregation but no DNA deformation (Figure 2B), we have $\xi_{DNA} \approx 6\pi\eta R_H(M)$ and $\xi_S \approx 6\pi\eta R_S$ and eq 12 gives:

$$\mu_{S\text{-}DNA}(M) \approx \frac{\mu_0}{1 + \dfrac{R_S}{R_H(M)}} \approx \frac{\mu_0}{1 + \sqrt{\dfrac{\alpha'}{M}}} \qquad (13)$$

where $\alpha' \approx 27R_S^2/4pb$ since $\alpha'=MR_S^2/R_H^2$ and $R_H^2 \approx 4pMb/27$. Finally, in the extreme limit where DNA becomes completely stretched (Figure 2C) because of the high counter flow velocity Δv, the friction coefficient of the DNA rod is given by $\xi_{DNA}=2\pi\eta L/[\ln(L/b)-1/2]=2\pi\eta Mb/(\ln M-1/2)$ and eq

12 predicts:

$$\mu_{S\text{-}DNA} \approx \frac{\mu_0}{1 + \frac{3\,R_S}{M\,b}\,(\ln M - 1/2)} \tag{14}$$

It is interesting to note that the molecular size dependence of the mobility is predicted to be (accidentally) almost the same in the high (eq 14) and low (eq 10) velocity regimes since the (ln M - 1/2) term is a weak function of M (it is a number of order unity). We can then write:

$$\mu_{S\text{-}DNA} \approx \frac{\mu_0}{1 + \frac{\alpha''}{M}} \tag{15}$$

where $\alpha''=3R_S(\ln M - 1/2)/b$. However, the friction parameters α and α'' are quite different; for example, α does depend on the persistence length p of the DNA fragment (and hence on the ionic strength I), while α'' is independent of both p and I.

2.2.3 The Critical Electric Field E_0

The critical electric field E_0 necessary to have hydrodynamic segregation between the DNA coil and the streptavidin label corresponds to applying a counter-flow force $F_0 \geq k_B T/R_G(M)$ on the DNA, where $R_G(M)$ is the radius-of-gyration of the DNA in its coil state and $F_0=6\pi\eta R_H(M)\Delta v$ [9]. Using eq 10, the critical field E_0 can thus be written as:

$$E_0 = \frac{k_B\,T}{4\,\pi\,\eta\,R_G^2(M)\,\mu_0}\,\frac{M + \alpha}{\alpha} \tag{16}$$

For a DNA fragment that is large enough to look like a random coil (i.e., L»p), we can use $R_G \approx (pMb/3)^{1/2}$ to obtain

$$E_0 \approx \frac{3\ k_B\ T}{4\ \pi\ \eta\ b\ \mu_0} \times \frac{M + \alpha}{M\ p\ \alpha} \qquad (17)$$

Since T=30°C, $\eta \approx 1$cP and $\mu_0 \leq 3\times10^{-4}$ cm²/Vs in our experiments, we get $E_0 \geq 0.077(M+\alpha)/M\alpha p$ [V/cm] (where p is in cm), which is much larger than the maximum electric field we could reach. When we use reasonable values such as M=300, p=3 nm and α=30, for instance, we get E_0=9.4 kV/cm! In order to attain such high fields, one would need a very special set-up indeed! Therefore, we expect to be in the segregation-free (and stretching-free) regime and α should be a function of the persistence length p of DNA (except perhaps at very low ionic strengths).

2.2.4 Artificial Segregation and Persistence Length

Even if the electric field is not high enough to obtain segregation (i.e., $E<E_0$), it is possible to see some effective segregation when the persistence length is comparable to the size of the label and the DNA molecules are short (Figure 2D). If $p \geq R_S$, it may take a very long DNA fragment in order to completely "wrap" around the streptavidin label and make the latter look like a regular blob. We then predict three different sub-regimes. Very small DNA fragments (oligomers with a contour length L<p) will look "stretched" and their mobility should be given by eq 14. Intermediate size DNA fragments that are long enough to look like a small random coil but not long enough to surround the label (Figure 3D) should have a mobility given by eq 13. Finally, the longest fragments

should satisfy eq 10. We thus expect to observe this artificial segregation for low ionic strength buffers for which p may exceed R_S.

2.3 Materials and Methods

2.3.1 Electrophoresis Solutions

Solutions containing 3M of urea (BDH), various concentrations (ionic strengths) of TAPS buffer (N-tris[Hydroxymethyl]methyl-3-aminopropane-sulfonic acid) (Sigma) and 0.04% (w/v) of poly-N,N-dimethylacrylamide polymer (POP polymer; Perkin Elmer) were injected into the capillaries to perform the separation. The POP polymer, which is used to dynamically coat the capillary walls [5], has a molecular weight of 10^6 g/mol, a polydispersity index PI≈3, and an entanglement concentration $c^* \approx 1.5\%$ w/v. This solution was buffered to pH=8.4 (the pK_a of TAPS) with NaOH. A solution containing 1.0 mol/L of TAPS thus has an ionic strength of I=0.5mol/L.

2.3.2 DNA and S-DNA Samples

We used Pharmacia's 100 basepair ladder for our experiments. The two ends of these double stranded DNA fragments are different. This allowed us to label only one of the strands. One end has a 5′ GGCT 3′ overhang whereas the other end has a 5′ AGCC 3′ overhang. The first step of the labeling procedure consisted in blocking the GG site of the first strand by adding deoxycytosine. We mixed 5 μL (5 μg) of 100bp ladder, 5 μL of 5×Amplitaq buffer, 5 μL dCTP (5 pmol/μL), 0.5 μL Amplitaq FS, and 9.5 μL dH_20, and we allowed the filling reaction to proceed at 60°C for 20 minutes. Unincorporated nucleotides were then removed using Centrisep columns (Princeton

Separations). In the second step, we mixed the modified 100bp ladder (17 μL) obtained in step one with 8 μL of 5×Amplitaq buffer, 2 μL of C-term big dye (21.5 pmoles), 4.5 μL of biotin-dUTP (148.5 pm), 0.5 μL Amplitaq FS and 8 μL dH_20 (for a total volume of 40 μL), and we allowed the reaction to proceed at 60°C for 20 minutes. Centrisep columns were then used to remove the unincorporated nucleotides.

The solution thus obtained was diluted 10 times with distilled water. The final DNA samples were produced by mixing 3μL of this dilute solution with 10μL of formamide and 6μL of dH_2O. The streptavidin-labeled DNA (S-DNA) samples were obtained by adding an appropriate quantity of purified streptavidin to this solution [5].

2.3.3 Electrophoresis Conditions

The running temperature (30°C) was controlled through the thermostatic plate of the ABI PRISM™ 310 Genetic Analyser capillary electrophoresis instrument. Experiments were performed on 45 cm-long capillaries and the applied voltage was 15 kV (E=333V/cm). Sample injection was performed at 15 kV for 5 seconds. The distance between the injection end of the capillary and the detector was 34 cm.

2.4 Results

2.4.1 DNA Free Mobility and Debye Length

Figure 3 shows a plot of the free (naked) DNA mobility μ_0 (in cm^2/Vs) as a function of the ionic strength I of our TAPS buffer solutions (in mol/L). We see that μ_0 decreases smoothly when

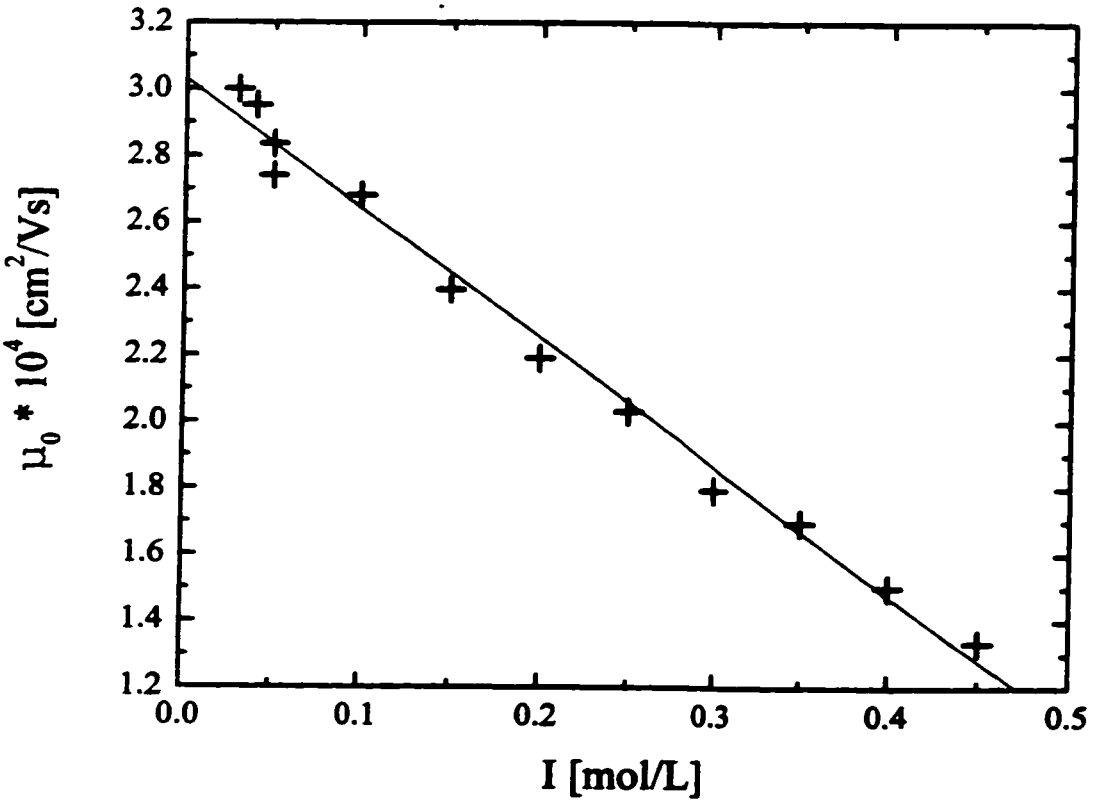


Figure 3 Free DNA mobility, $\mu_0 \times 10^4$ [cm^2/Vs], vs. the ionic strength I [mol/L]. An approximate linear behavior is observed with $\mu_0(I)=(3.03-3.88\ I\ [mol/L])\times 10^{-4}$ cm^2/Vs.

I increases, as expected. The mobility in the zero ionic strength limit (I=0) is given roughly by the extrapolated value $\lim_{I\to 0}\mu_0 \cong 3\times10^{-4}$ cm^2/Vs, as shown. The curve was fitted using the empirical linear relation $\mu_0=(3.03-3.88\ I)\times10^{-4}$ cm^2/Vs since it provided a good fit over this range. It was not possible to observe DNA separation in our experiments showing that we clearly remained in a regime for which $\lambda_D<R_G(M)$ for all molecular sizes M (i.e., eq 5 does not yet apply); in other words, this value of μ_0 does not correspond to true I→0 conditions.

Figure 4 replots this data as μ_0 vs. $\log(I^{-1/2})$. We observe a reasonably linear behavior for values of $\log(I^{-1/2})\lesssim0.8$, indicating that we are then in the regime where $p>\lambda_D$ (eq 7). The fit gives us a value of I_0=1.68 mol/L. The estimated value for the monomer size, using this value of I_0 and the fact that $b=(\epsilon_b\epsilon_0k_BT/2e^2I_0)^{1/2}$ according to our definition of I_0, is b=0.24 nm which is not too far from the actual value of b=0.43 nm given the semi-quantitative nature of our theory.

2.4.2 S-DNA Conformations and Ionic Strength

In the limit where the two parts of the S-DNA molecule do not segregate, eq 10 predicts that $\mu_0/\mu_{S\text{-}DNA}-1=\alpha/M$. Figure 5 shows a log-log plot of $\mu_0/\mu_{S\text{-}DNA}-1$ as a function of the molecular size M for I=0.45 mol/L. In agreement with the model, we obtain a slope of −1, and the fit gives the value of α=57.2 for the relative friction coefficient of the S-label. This is smaller than the value of $\alpha''\approx100$ (using M=500) predicted for stretched S-DNA molecules. As discussed previously, we are quite far from the stretching limit since our electric field (E=333V/cm) is much smaller than $E_0\cong9.4$kV/cm, the electric field required for DNA/streptavidin segregation.

Figure 6 shows a log-log plot of $\mu_0/\mu_{S\text{-}DNA}-1$ as a function of the molecular size M for a low

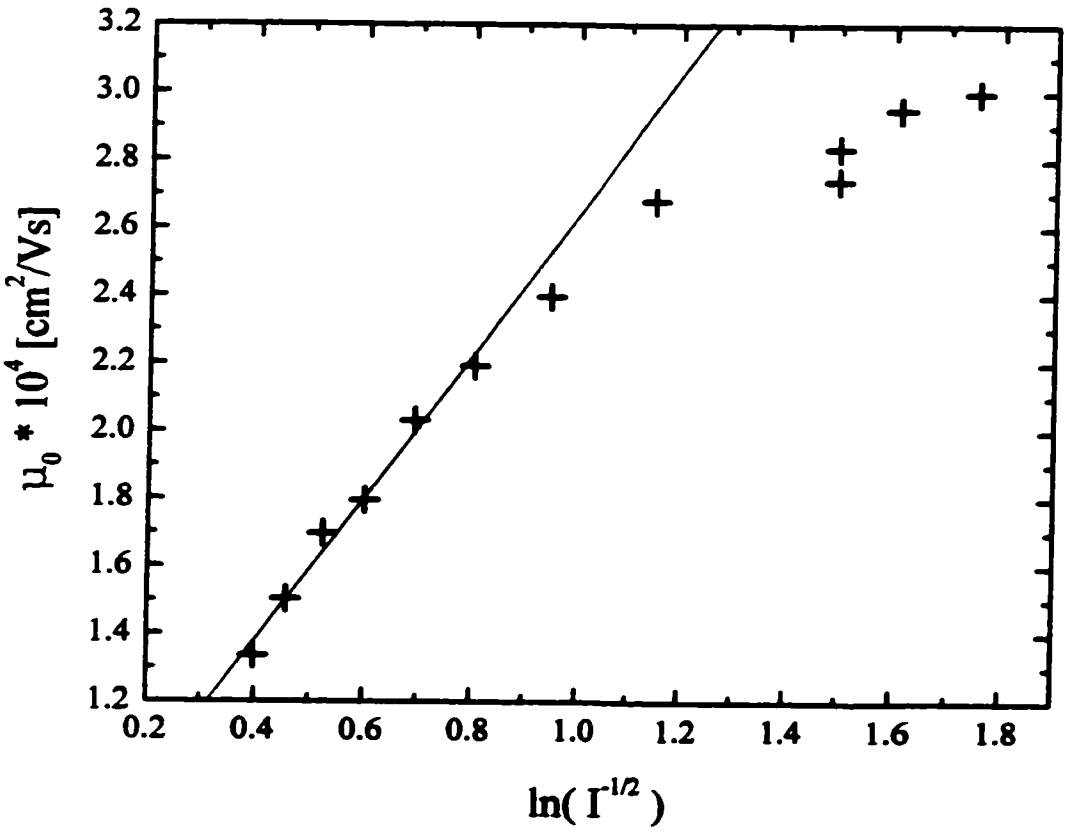


Figure 4 Free DNA mobility $\mu_0 \times 10^4$ [cm^2/Vs] vs. ln($I^{-1/2}$). The straight line $0.54+\ln(1.68/I)^{1/2}$ corresponds to a linear fit of the first six data points. The value of this fit leads us to a value of I_0=1.68.

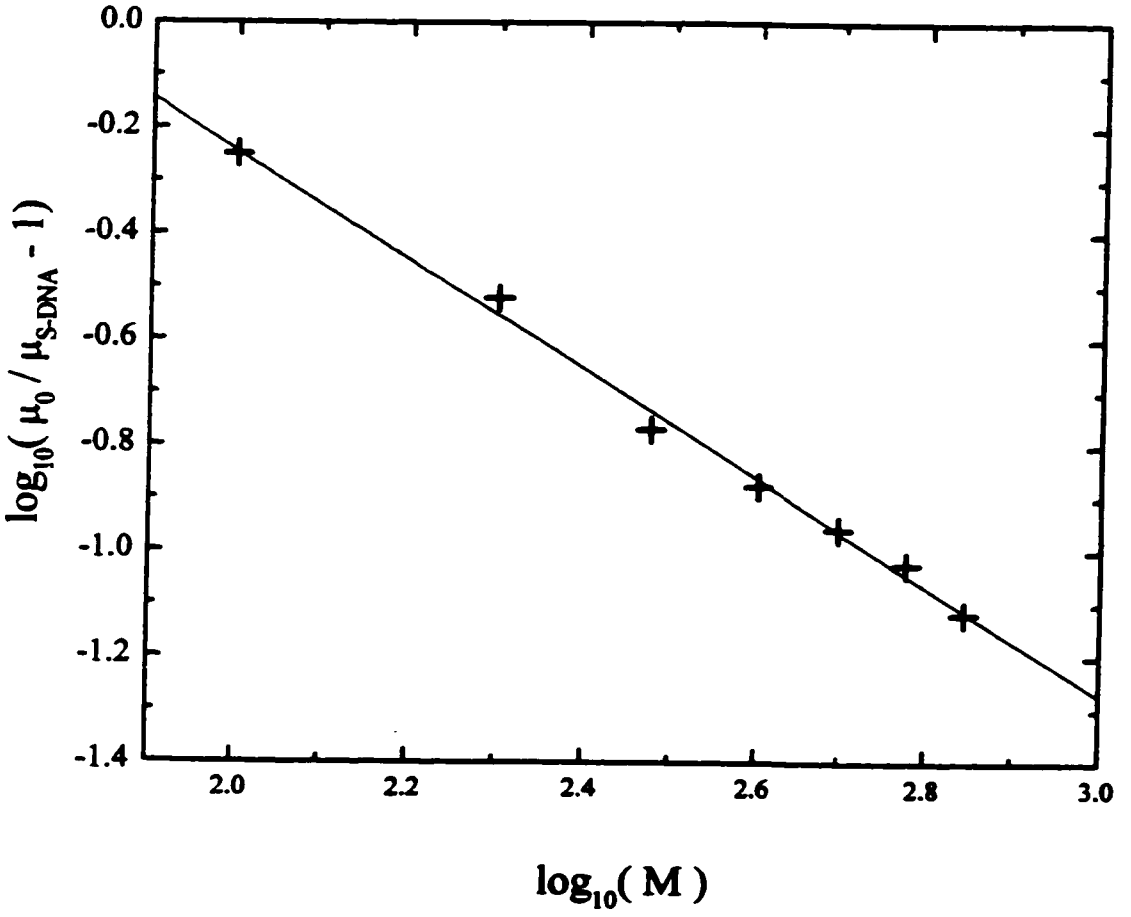


Figure 5 Log-log plot of $\mu_0/\mu_{S\text{-}DNA}$-1 vs. DNA size M for a high salt content buffer solution (I=0.45 mol/L). The fit gives a slope ≈-1.

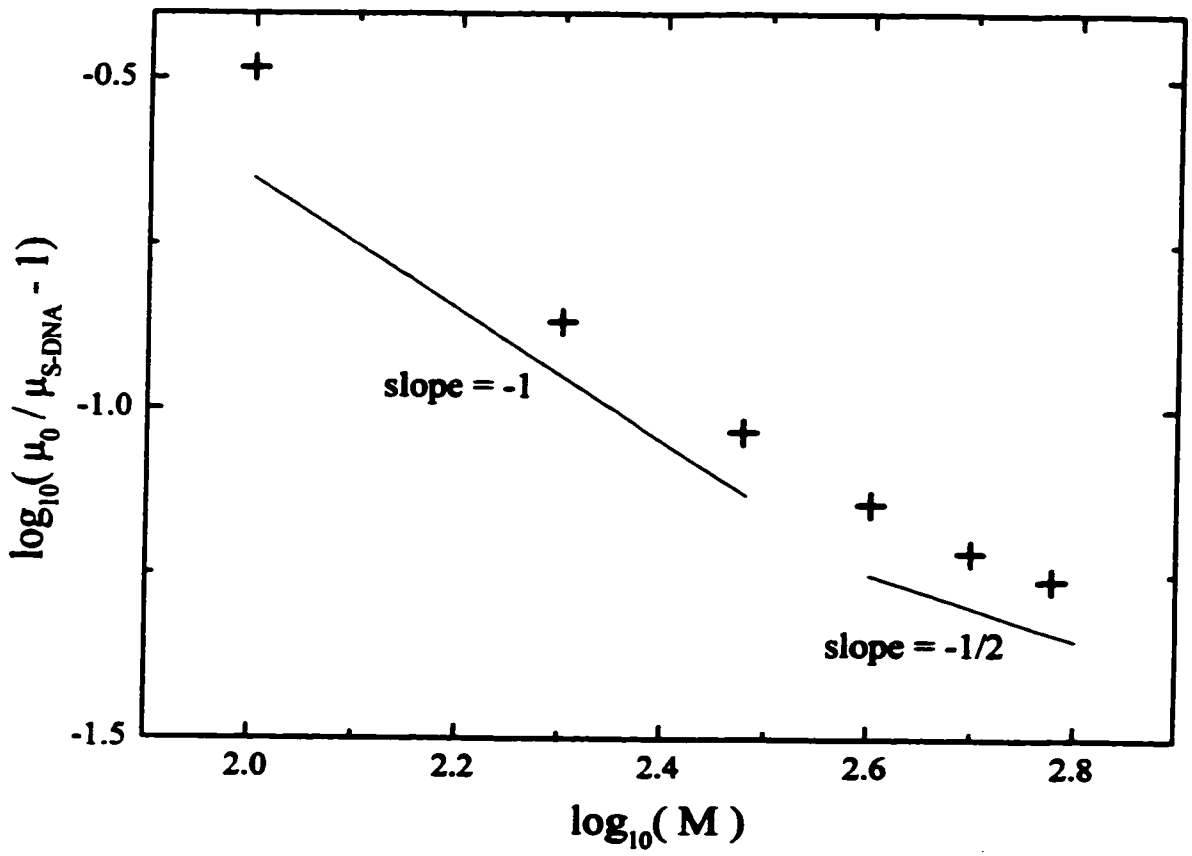


Figure 6 Log-log plot of $\mu_0/\mu_{S\text{-}DNA}$-1 vs. DNA size M for a low salt content buffer solution (I=0.04 mol/L).

ionic strength case where I=0.04 mol/L. In the presence of artificial segregation, we would expect three different regimes depending on the molecular size of the DNA fragment. For the smaller fragments, we should observe a slope of -1 since eq 15 should apply. For intermediate molecular sizes, eq 13 predicts a slope of -½. Finally, a slope of -1, as predicted by eq 10, should be recovered for large enough DNA fragments. We apparently see the transition between the first two of these regimes in Figure 6, but the last transition was not reached because our fragments were too small. With such a low ionic strength, we are clearly in the large persistence length limit and some artificial DNA/streptavidin segregation takes place. We conclude that it takes a ssDNA molecular size M>400 bases at that ionic strength to get a situation like the one shown in Figure 2B.

2.4.3 The Relative Friction Coefficient α of Streptavidin and the DNA Persistence Length p

Figure 7 shows a plot of α as a function of the ionic strength I (α was found using eq 10 to fit the data, as shown in Figure 5). We see that α increases smoothly when we increase the concentration of salt in our buffer solution. This is in agreement with the fact that the persistence length of DNA decreases with increasing ionic strength. We need more DNA bases to form a blob that has a hydrodynamic radius $R_H=R_S$ when we increase the flexibility of DNA (or decrease the persistence length).

From the value of α obtained in the region where the Kratky-Porod equation has a solution (α>30.21), it is possible to estimate the persistence length p of DNA using eq 11; the results are shown in Figure 8. A fit of the form $p=p_0+A/I$ was found to be appropriate; here, p_0 is the intrinsic

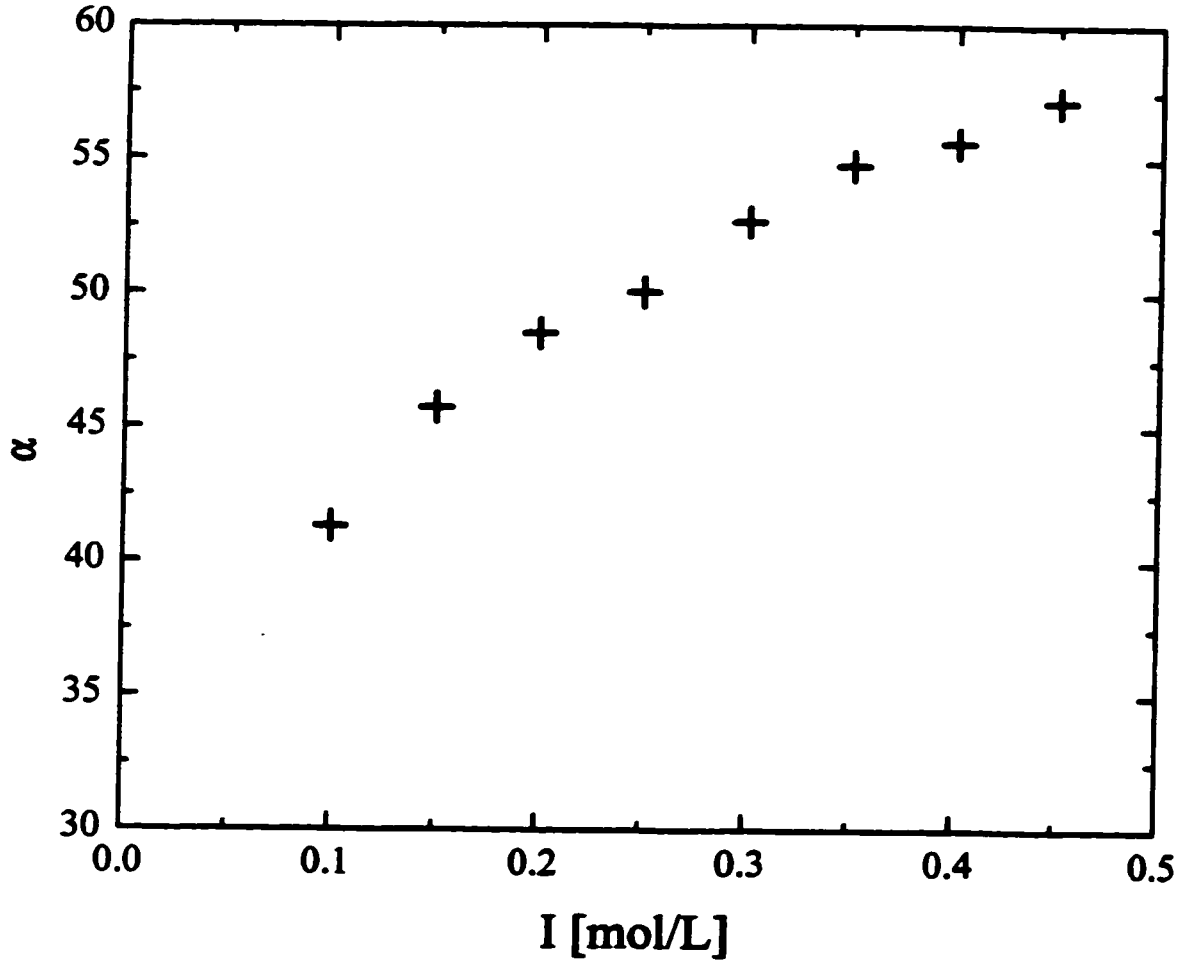


Figure 7 Relative friction coefficient of the streptavidin label (α) vs. the ionic strength I of the buffer [in mol/L].

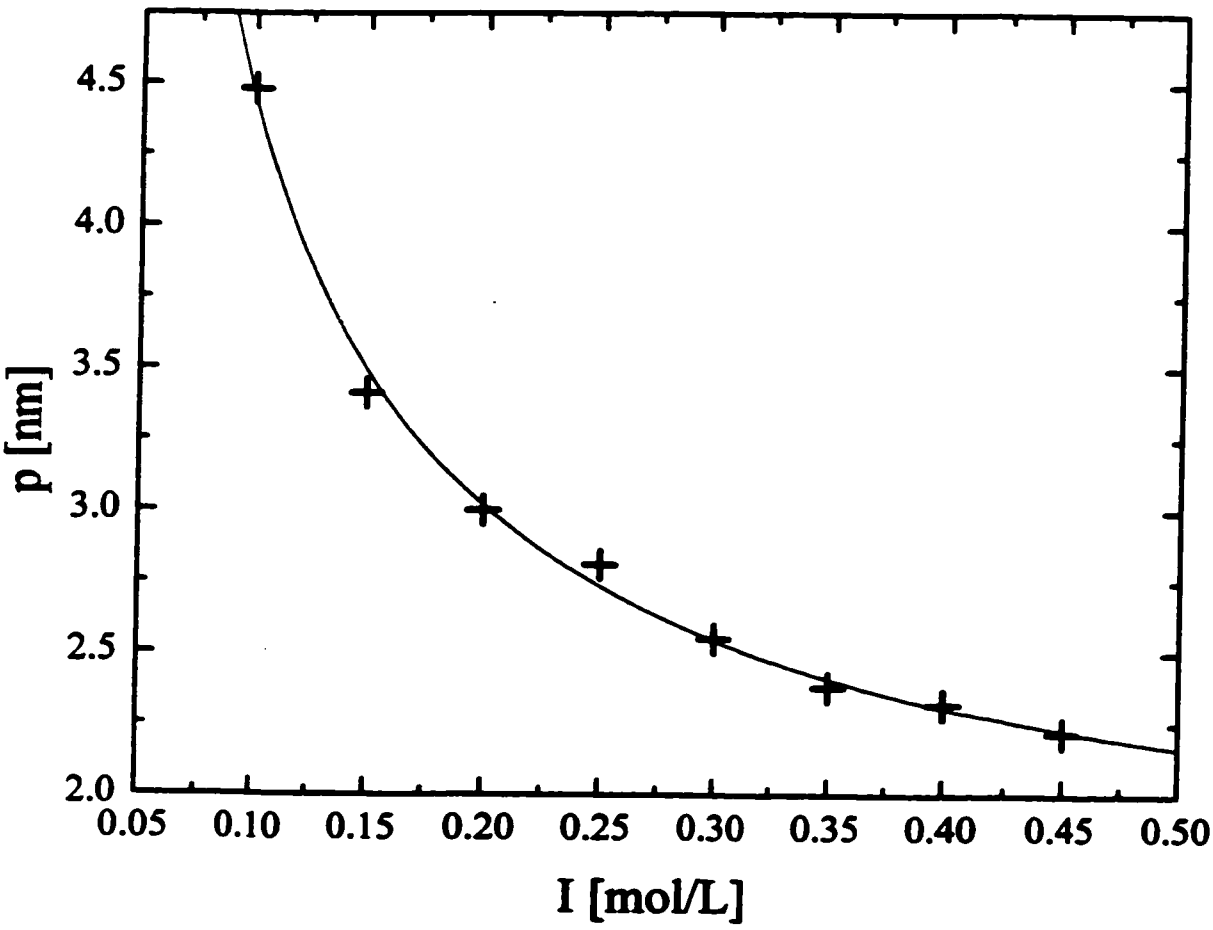


Figure 8 ssDNA persistence length p [in nm] vs. the ionic strength I of the buffer [in mol/L]. The best fit gives p=(1.60+0.28/I [mol/L]) nm.

or elastic persistence length of DNA (the persistence length DNA would have if it were an uncharged molecule). This is in agreement with the prediction of Ha and Thirumalai [13]. The intrinsic persistence length of ssDNA, p_0=1.6 nm, is slightly higher than the one estimated by Tinland et al. [20], 0.8nm<p_0<1.3 nm, using diffusion measurements. Our fit p=(1.6+0.28/I) nm gives values that are higher than the ones estimated by Tinland over the whole range of ionic strengths. This can be due in part to the fact that we have neglected excluded volume effects and assumed that the relation between the hydrodynamic radius and the radius of gyration is always R_H/R_G ~2/3; in addition, we used a different buffer.

2.5 Discussion

In this Chapter, we showed that the ionic strength plays a very important role during an ELFSE experiment because it affects several key parameters. A study of the free-solution mobility of DNA and streptavidin-labeled ssDNA fragments showed that even if a ssDNA fragment is globally in a random coil conformation, one must describe this molecule as a collection of blobs on the relevant length scale(s) in order to explain its static and dynamic properties (e.g. the Debye length, the persistence length or the size of the label). Analysis and theoretical modeling must be done very carefully.

In most theories, DNA (and other linear polymer chains) is considered to be a random coil on length scales larger than its persistence length p. However, the situation is far more complex when one tries to understand the electrophoretic properties of DNA or of hybrid molecules such as our streptavidin-ssDNA complexes. In order to fully understand the physics underlying ELFSE, for

instance, one must actually consider 4 different length scales: the radius of gyration of the DNA molecule $R_G(M)$, its persistence length p, the Debye length λ_D and the size of the globular label R_S. Our analysis showed that although DNA can look globally flexible at one length scale, its microscopic rigidity is an important element, especially when p and R_S are comparable.

Since we have $R_G \gg \lambda_D$ and $R_G > p$ over a wide range of buffer ionic strengths, DNA is a "flexible" molecule with the local dynamical properties of a "rigid" one. When no electric field is applied, the collective hydrodynamic effects make the DNA fragment move like a solid sphere of radius $R_H \sim R_G$. In the presence of an electric field E, however, the hydrodynamic interactions vanish over a distance λ_D (because of the motion of the cations surrounding DNA), and the DNA blobs of size λ_D are independent objects of electrophoretic mobility μ_{BLOB}. Since those independent blobs all move at the same velocity, the size of the molecule is of little importance and the mobility of the undeformed DNA random coil molecule is also given by μ_{BLOB}. This is why it is generally impossible to separate naked DNA fragments in free solution.

In our experiments, we observe two distinct regimes for the free mobility of DNA, μ_0. The free mobility decreases with increasing ionic strength like $\mu_0 \propto \ln(I^{-1/2})$ (eq 8) for values of I>0.2mol/L (the linear part of Figure 4), but it is smaller than predicted by eq 8 for smaller values of I (the curvature on Figure 4). This corresponds to the transition from eq 8 to eq 6, and it occurs at a value of p that corresponds to the size of the hydrodynamically independent DNA blobs (or $p \approx \lambda_D$).

It was shown that the addition of an undeformable globular label at one end of the DNA fragments makes free-solution electrophoretic separation possible [4]. The first theoretical model of ELFSE made the assumption that there was no hydrodynamic interaction between the DNA and

streptavidin components [2]. We showed that this assumption is only valid for unrealistically high electric fields (note, however, that longer DNA molecules and larger labels could easily satisfy this condition; see eqs 16 and 17). The hydrodynamic interactions then lead to three distinct regimes.

At "low" electric field, the DNA-streptavidin complex can be considered as one hydrodynamic unit and the mobility is given by $\mu=\mu_0/(1+\alpha/M)$, where α is closely related to the persistence length of DNA and the size of the label. In fact, α is the molecular size of a blob of DNA having the same friction coefficient as streptavidin ($R_H(\alpha)=R_S$). Since the number of DNA bases in such a blob (α) decreases with increasing persistence lengths (p), there is a strong correlation between the observed value of α and the ionic strength I of the running buffer. From our values of $\alpha(I)$, we were able to estimate the persistence length p(I) of the ssDNA fragments. Interestingly, attaching a label to DNA seems to be a good method to estimate the persistence length of the latter! For all practical purposes, the label becomes a "persistence length probe". However, streptavidin may not be the optimal label for this purpose because it is not very large compared to p; this limits our ability at determining p via the Kratky-Porod equation. Nevertheless, our estimated persistence lengths are in good agreement with the predicted theoretical form $p=p_0+A/I$. Our results indicate that $p\approx(1.60+0.28/I)$nm. The value of the intrinsic persistence length $p_0=1.60$nm is comparable to the value obtained using a method based on the measurement of diffusion coefficients [21], although our persistence lengths are systematically higher than those of Tinland [21].

In the second hydrodynamic regime, the higher electric fields induce the segregation of the complex. In that limit, however, the field is not yet high enough to stretch DNA and the deformation of the DNA strand can be neglected. The total friction coefficient $\xi_{S\text{-}DNA}$ is simply the sum of the

individual component ξ_S and ξ_{DNA}, and the net mobility is given by $\mu=\mu_0/(1+R_S/R_H(M))$. We were unable to observe this regime since the required electric field intensities are much higher than what can be achieved with commercial capillary electrophoresis instruments. However, observations at very low ionic strengths seem to indicate that this regime can also be observed at low fields when the persistence length is not much smaller than the radius of gyration R_G. We call this situation "artificial segregation" since the segregation between the streptavidin and the DNA fragment is not due, in this case, to the effect of the field driven drag force.

The last regime is characterized by DNA stretching. The mobility equation is somehow similar to the equation in the low field regime, but with $\alpha=\alpha''=3R_S(\ln M-1/2)/b$ (this gives, e.g., $\alpha''\sim 91$ for M=300). The value of α would then be larger than in the low field regime, thus increasing the quality of the separation. Moreover, there is a major benefit in using very high fields when we consider the effect of diffusion. The diffusion coefficient at low field is given by the Stoke-Einstein relation $D_M=k_BT/(6\pi\eta R_H(M))$. As far as hydrodynamics is concerned, a S-DNA molecule containing M bases is roughly equivalent to a DNA molecule with $M+\alpha$ bases. Since in general, we have $M\gg\alpha>p$, we expect the scaling $D_M\sim 1/M^{1/2}$. In the opposite limit where DNA is stretched, one should have $D_M\sim 1/M$. One thus expects much sharper bands in the latter case. The optimal ELFSE strategy remains to be established.

In conclusion, our new ELFSE theory can explain the experimental data reported in this chapter. It is worth mentioning that the concept of effective charge is not necessary to explain our experimental results. Therefore the concept of effective charge is not relevant in electrophoresis, a point recently stressed by Long et al. [8].

2.6 References

[1] Olivera, B. M.; Baine, P.; Davidson, N.; *Biopolymers*, **2**, 245, 1964.

[2] Mayer, P.; Slater, G. W.; Drouin, G.; *Anal. Chem.*, **66**, 1777, 1994.

[3] Iki, N.; Kim, Y.; Yeung, E. S.; *Anal. Chem.*, **68**, 4321, 1996.

[4] Heller, C.; Slater, G. W.; Mayer, P.; Dovichi, N. J.; Pinto, D.; Viovy, J. L.; Drouin, G.; *J. Chromatogr, A*, **806**, 113, 1998.

[5] Ren, H.; Karger, A. E.; Oaks, F.; Menchen, S.; Slater, G. W.; Drouin, G.; *Electrophoresis*, **20**, 2501, 1999.

[6] Long, D.; Viovy, J. L.; Ajdari, A.; *Phys. Rev. Lett.*, **76**, 3858, 1996.

[7] Long, D.; Viovy, J. L.; Ajdari, A.; *J. Phys.: Condens. Matter*, **8**, 9471, 1996.

[8] Long, D.; Viovy, J. L.; Ajdari, A.; *Biopolymers*, **39**, 755, 1996.

[9] Long, D.; Ajdari, A.; *Electrophoresis*, **17**, 1161, 1996.

[10] Long, D.; Dobrynin, A. V.; Rubinstein, M.; Ajdari, A.; *J. Chem. Phys.*, **108**, 1234, 1998.

[11] Odijk, T.; *J. Polym. Sci.*, **15**, 477, 1977.

[12] Barrat, J. L.; Joanny, J. F.; *Europhys. Lett.*, **24**, 333, 1993.

[13] Ha, B.-Y.; Thiramalai, D.; *J. Chem. Phys.*, **110**, 7533, 1999.

[14] Doi, M.; Edwards, S. F.; The Theory of Polymer Dynamics; Clarendon Press, Oxford, 1986.

[15] Viovy, J. L.; Heller, C.; Principles of Sized-Based Separations in Polymer Solutions, in *Capillary Electrophoresis in Analytical Biotechnology*, (Righetti, P. G., ed.), CRC Press, Boca Raton, 1996.

[16] Israelachvili, J.; Intermolecular & Surface Forces; Academic Press, San Diego, 1992.

[17] Hoagland, D. A.; Arvanitidou, E. S.; *Polymer Preprints*, **34**, 1059, 1993.

[18] Berg, H. C.; Random Walks in Biology, Princeton University Press, Princeton, 1993.

[19] Perkins, T. T.; Quake, S. R.; Smith, D. E.; Chu, S.; *Science*, **264**, 822, 1994.

[20] Brochard-Wyard, F.; *Europhys. Lett.*, **30** 387, 1995.

[21] Tinland, B.; Pluen, A.; Sturm, J.; Weill, G.; *Macromolecules*, **30**, 5763, 1997.

Chapter 3

Competition between Trapping and Frictional Effects during Capillary Electrophoresis of Streptavidin-DNA Molecules in Polymer Solutions*

It was previously shown that labeling one of the ends of single-stranded DNA molecules with a neutral label like streptavidin increases the interband separation of these hybrid DNA molecules when they are electrophoresed in denaturing cross linked polyacrylamide gels. This process is called Trapping Electrophoresis (TE) because the dynamical properties of the analytes are strongly affected by steric trapping in the gel (Chapter 5 and 6). End-labeled ssDNA molecules can also be electrophoretically separated in free-solution: this process is called End-Labeled Free Solution Electrophoresis (ELFSE) and the separation is then due to the extra hydrodynamic friction of the streptavidin (Chapter 2). In this chapter, we present a study of end-labeled DNA capillary electrophoresis in the presence of non-cross linked polymer solutions ranging from entangled to ultra-dilute conditions. In the entangled limit, steric and sieving interactions dominate and we observe that small molecules move faster than larger ones. In the ultra-dilute limit, we recover free-solution like separations where small molecules are more retarded than larger ones. We also observe a fascinating transition between these regimes for intermediate polymer concentrations and DNA sizes where the competition between trapping, sieving and frictional effects leads to non-monotonic mobility-size relationships.

---

* The results presented in this chapter will be submitted for publication before the end of the current year.

3.1 Introduction

3.1.1 Dilute Polymer Solutions

In free buffer solutions, it is impossible to separate DNA fragments (either single-stranded (ssDNA) or double-stranded (dsDNA)) of different sizes because of the famous free draining properties of DNA (see Chapter 2). We also know that in dense gels, DNA molecules must deform in order to migrate through the fixed matrix. In matrices like polyacrylamide or agarose gels, a flexible DNA fragment can find its way through the gel like a snake through thick grass; in other words, it is reptating in a reptation tube as originally proposed by De Gennes and Edwards for polymer melts [1,2]. In this situation, the net electrophoretic velocity decreases with the molecular size of the DNA fragment, and separation (e.g. sequencing) becomes possible [3-7]. Very large DNA molecules (e.g., beyond 40 kbp in agarose gels), however, tend to orient in the field direction and separation then requires the use of pulsed fields [8-10].

The introduction of capillary electrophoresis (CE) allows the use of entangled polymer solutions as sieving matrices. Such matrices cannot be called gels because of the absence of crosslinkers; we then have a somewhat "dynamic gel" where each "pore" has a finite lifetime [11]. The DNA migration mechanisms are believed to be essentially the same in concentrated polymer solutions and crosslinked gels as long as the DNA residence time in a "pore" is short compared to the lifetime of the latter. The critical advantage of polymer solutions over chemical gels is the possibility of injecting a fresh polymer solution after every run (gels degrade with time). Moreover, this process can be fully automated and as many as 100 runs can be carried out in a single capillary. Barron et al. [12] discovered that even dilute polymer solutions (whose concentration c is smaller

than the entanglement threshold c*) can provide separation of dsDNA fragments. For example, these authors showed that DNA separation occurs with hydroxyethylcellulose (HEC) concentrations as low as 1% of the entanglement limit (the matrix certainly does not look like a gel in that limit). The separation process, suggested by Barron et al. [12] and modeled by Hubert et al. [13], is a dynamic capture/release mechanism. While the velocity of the small fragments is not really affected by the presence of the HEC polymer (the probability of colliding with the polymers and the subsequent contact time are both very small), the velocity loss experienced by the large DNA fragments is sufficient to observe the separation of restriction fragments 2-45 kbp in size. Hubert et al. [13] proposed a clever analytical model to explain these results where it is assumed that DNA molecules collide and temporarily capture HEC molecules during their migration. It is in fact the increase in the total friction coefficient of the DNA-polymer complex that makes separation possible. Assuming that both the DNA and the HEC polymer slide around their common point of contact, they calculated the mean lifetime of this DNA-HEC association. This theoretical model is in fair agreement with the experimental results of Barron et al. [12] for short HEC polymers.

3.1.2 Trapping Electrophoresis

In 1990, Ulanovsky, Drouin and Gilbert (UDG) [14] published the first paper about Trapping Electrophoresis (TE). UDG's idea was to attach streptavidin, a neutral globular protein, at one end of the DNA to be electrophoresed in order to eliminate the non-separative regime observed for large oriented DNA molecules (the latter co-migrate during gel electrophoresis). This end of the DNA molecule thus has an excess of friction due to the presence of the streptavidin label while the other

end is left unmodified. The extra friction due to the presence of streptavidin makes the total friction coefficient of this hybrid streptavidin-DNA (or S-DNA) molecule larger than the friction coefficient of the corresponding naked DNA. Because of this extra friction, S-DNA molecules move slower than naked DNAs. While the bulky label has almost no effect on small molecules (apart from the fact that these S-DNA molecules have a lower velocity due to this extra hydrodynamic friction), molecules larger than a critical molecular size M* experience an abrupt velocity drop which improves their separation [15]. This dramatic effect can be explained by the fact that when a S-DNA molecule moves through a pore smaller than the size of the label, the molecule stops its migration since the label cannot cross this topological constraint. In fact, the molecule needs to backtrack out of this small pore in order to continue its progression into the gel (see Figure 2 of Chapter 1). The pores smaller than the size of the label are thus called traps. Traps have a larger effect on larger S-DNA molecules since the electric force pulling the S-DNA inside the trap is proportional to the total charge of the molecule [16-18].

3.1.3 End-Labeled Free Solution Electrophoresis (ELFSE)

The electrophoretic properties of S-DNA molecules have not only been studied in gels. This label can actually make separation of DNA fragments possible even in free solution. While it is impossible to separate naked DNA in free solution because of the well-known size-independence of its charge over friction ratio, the streptavidin label makes separation possible by creating an asymmetry between force and friction. This separation mechanism, first discussed quantitatively by Mayer et al. [19], is called End-Labeled Free Solution Electrophoresis (ELFSE) and can be used for

DNA sequencing [20]. This new method presents many advantages, including the fact that free solution separations make the injection into the capillary very easy (no viscous polymers). Moreover, the velocity is often higher because of the absence of a sieving matrix. Another interesting fact is that the resolution is better at high electric field intensity, a property that can also help speed up DNA sequencing.

This new separation process is rather similar to the one observed by Barron et al. [12] in dilute polymer solutions. As mentioned previously, it is the extra friction (or the dragging force) due to the contact between the DNA and the polymer that makes separation possible in dilute polymer solutions. In ELFSE, on the other hand, it is the extra friction due to the presence of streptavidin that makes separation possible. In that sense, we can say that separation in dilute polymer solutions is a stochastic version of the ELFSE process. Even though the reason for separation is similar in both cases, the velocity increases with molecular size in ELFSE but decreases in dilute polymer solutions. This is simply due to the fact that the extra friction is the same for every molecule in ELFSE (one streptavidin per DNA fragment) while the number of HEC polymers in contact with the DNA molecule is a function of the molecular size of the latter: the longer the DNA molecule, the more likely it will drag HEC polymers.

3.1.4 This Study

The goal of this study is to bridge the gap between the experimental results obtained in the ELFSE regime [20] and those characteristic of the dilute polymer solutions [12]. We are interested in understanding how the presence of a small quantity of neutral polymers, in conjunction with the

frictional properties of the label, will modify the dynamics of ssDNA molecules. It is important to realize that the electrophoretic mobility would decrease with the DNA molecular size M if we only added the polymers to the solution. Similarly, we would have the opposite situation if we only added the label to the DNA. Obviously, the electrophoretic process will be more complicated when we both label the DNA and use a dilute polymer solution. A model of this new regime, based on the model of Hubert et al. [13] and on our theory of ELFSE (Chapter 2), is presented here. We then compare our theory to our experimental results.

3.2 Theory

3.2.1 Trapping Electrophoresis

A streptavidin-DNA molecule that moves in a gel has two possible states: the molecule is either inside a trap (velocity $V_{S\text{-}DNA,\,in}$ is zero) or free to move with a velocity $V_{S\text{-}DNA,\,out}>0$ given by:

$$V_{S\text{-}DNA,\,out}(M) = V_{DNA}(M + \alpha)\ \frac{M}{M + \alpha} \qquad (1)$$

The derivation of this simple relation is presented in Chapter 6. It is worth mentioning that the interpretation of the added friction α is exactly the same as in free solution (Chapter 2).

Because of trapping, there are two critical times in our system: τ_0, the time it takes to get inside a trap, and τ_d, the time spent (or "lost") inside a trap. The equation for the net velocity of the streptavidin-DNA complex is then given by:

$$V_{S\text{-}DNA}(M) = V_{DNA}(M + \alpha) \times \frac{M}{M + \alpha} \times \frac{\tau_0}{\tau_0 + \tau_d} \qquad (2)$$

where $\tau_0/(\tau_0+\tau_d)$ simply represents the fraction of the time the molecule is actually moving. Since detrapping is a thermally activated process, the equation for the detrapping time follows an Arrhenius form (see Chapter 5). That is, the detrapping time increases exponentially with the electric field (E) and the DNA molecular size (M), explaining the dramatic decrease of the velocity for molecular sizes $M>M^*(E)$ [14-18].

In the weak trapping regime, the time spent inside a trap is negligible compared to the time spent outside the trap and $\tau_0/(\tau_0+\tau_d) \approx 1$. The velocity of the streptavidin-DNA complex is then related to the velocity of naked DNA by the relation $V_{S\text{-}DNA}(M) \approx V_{S\text{-}DNA,\,out}(M)$ and eq 1.

3.2.2 End-Labeled Free Solution Electrophoresis

We have seen that the streptavidin label plays two roles in a gel: it increases the friction retarding the DNA, and it generates a steric trapping phenomenon. In free solution, it is solely the increase of the total friction which makes separation possible. Figure 1 shows a schematic representation of the main elements of the capillary electrophoresis experiments we carried out. The buffer solution contains TAPS ions (N-tris[Hydroxymethyl]methyl-3-aminopropane-sulfonic acid) and enough urea (3M) to keep the DNA denatured; streptavidin-ssDNA molecules were prepared as explained later; and finally a dilute solution of poly-N,N-dimethylacrylamide (Performance Optimized Polymer, Perkin-Elmer) is added to the buffer. This polymer is slightly hydrophobic and sticks to the wall of the capillary, thus killing the electroosmotic flow [20,21] (this process is

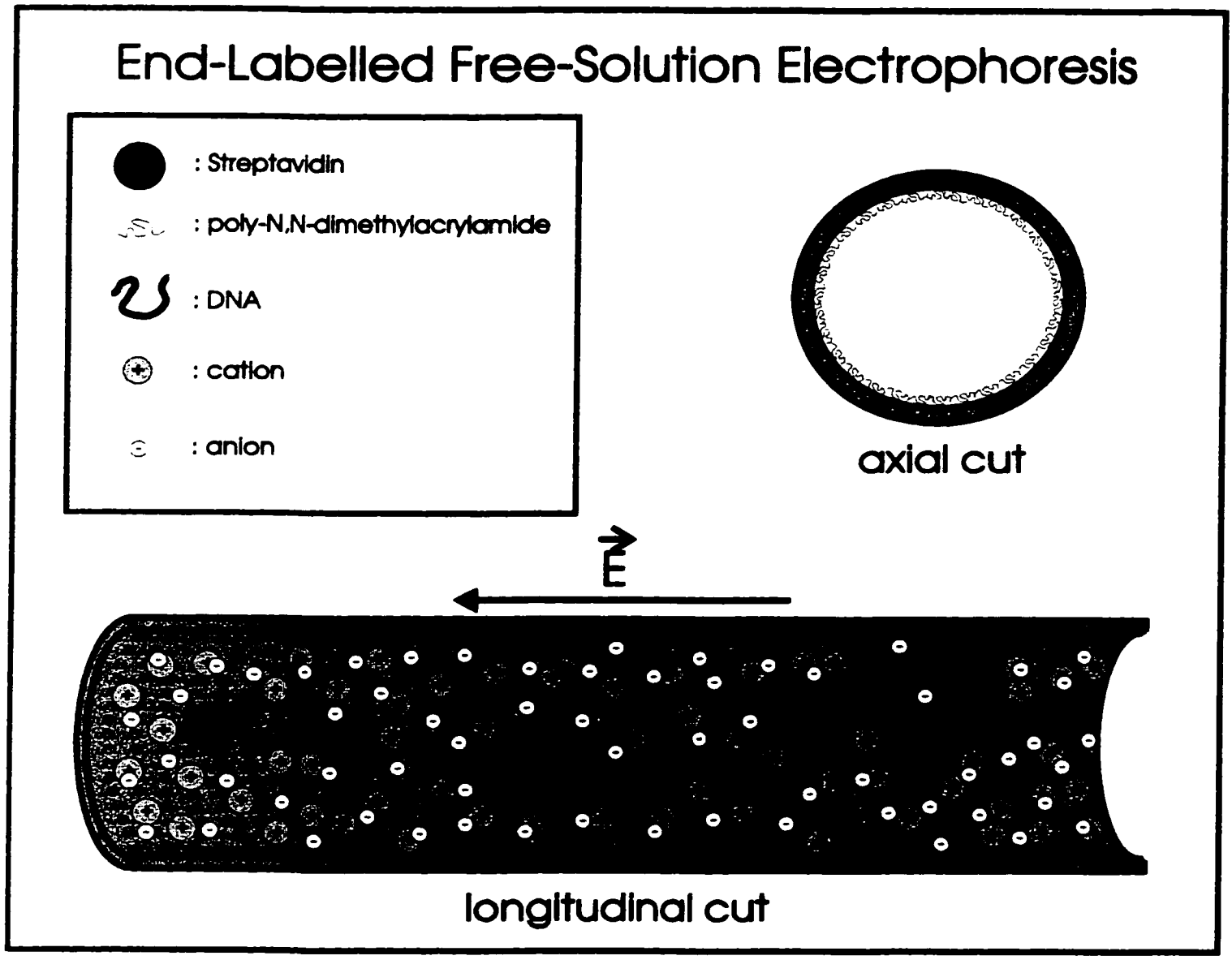


Figure 1: Schematic representation of the migration of the end-labeled DNA molecules in free-solution. The poly-N,N-dimethylacrylamide is present in order to kill the electroosmotic flow.

commonly called dynamic coating). As discussed in Chapter 2, the relation between the velocity of naked DNA and the velocity of the corresponding streptavidin-DNA complex is then given by:

$$V_{S\text{-}DNA}(M) = V_{DNA}(M) \times \frac{M}{M + \alpha} \qquad (3)$$

Since $V_{DNA}(M) \equiv V_0$ is independent of the DNA molecular size M, this relationship can also be rewritten as $V_{S\text{-}DNA}(M)=V_{DNA}(M+\alpha)\times M/(M+\alpha)$. Surprisingly enough, this relation between the DNA and S-DNA velocities is the one found in the weak trapping regime in the previous section (eq 1; see also Chapter 6)! The value of α is now the size (in number of DNA bases) of a DNA coil having the same hydrodynamic radius as the streptavidin label (i.e., $R_H(\alpha)=R_s$). In other words, α is a relative friction coefficient.

3.2.3 Dilute Polymer Solutions

We have seen that the relation between the velocities of naked DNA fragments and streptavidin-DNA complexes is the same in a gel (in the weak trapping regime) and in free solution. We will now be interested in the behavior of DNA and S-DNA in a dilute (non-entangled) polymer solution. Collisions with the free polymer chains slow down the DNA molecules and make separation possible in this case [12,13]. The addition of streptavidin will make collisions more likely to occur since collisions with the extra label are also possible. Let's assume that the steady-state average number (n) of polymers dragged by the analyte is the same for a S-DNA fragment containing M bases and a naked DNA fragment containing M+β bases. The mobility of the

corresponding naked DNA fragment would be given by [13]:

$$\mu_{DNA}\,(M + \beta) = \frac{Q\,E - n\,F_{drag}}{(M + \beta)\,\xi} \tag{4}$$

where $Q \sim M+\beta$ is the charge of the DNA, E is the electric field intensity, $F_{drag}=\mu_{DNA}(M+\beta)E\,\xi_p$ is the average frictional force acting on the DNA when the latter drags one polymer chain, and ξ_p is the polymer friction coefficient. On the other hand, the mobility of the S-DNA fragment would be given by:

$$\mu_{S\text{-}DNA}(M) = \frac{Q'\,E - n\,F'_{drag}}{(M + \alpha)\,\xi} \tag{5}$$

where $Q' \sim M$ takes into account the fact that the charge is smaller since the DNA and S-DNA molecules do not have the same lengths. Note that $F'_{drag}=\mu_{S\text{-}DNA}(M)E\xi_p \neq F_{drag}$. From eqs (4) and (5), it is easy to show that:

$$\frac{M + \beta}{\mu_{DNA}(M + \beta)} = \frac{M}{\mu_{S\text{-}DNA}(M)} + \frac{\beta - \alpha}{\mu_0} \tag{6}$$

where $\mu_0=Q'E/M\xi$ is the free-flow mobility of a free-draining DNA molecule. Since $M \gg \beta-\alpha$ and $\mu_0>\mu_{DNA}>\mu_{S\text{-}DNA}$ in all realistic cases, this equation can be simplified to $\mu_{S\text{-}DNA}(M) \approx \mu_{DNA}(M+\beta)\times M/(M+\beta)$ which is exactly the same as eq (1) if $\alpha=\beta$. As we will see later, our experimental results actually suggest that $\alpha=\beta$. According to the theory of Hubert et al. theory [13],

the parameter n is only proportional to the collision cross-section between the analyte and the free polymers. Since the streptavidin label is (hydrodynamically) equivalent to adding a blob of α extra nucleotides, one would indeed expect that $\alpha \approx \beta$. Therefore, we can again write that:

$$\mu_{S\text{-}DNA}(M) \approx \mu_{DNA}(M + \alpha) \times \frac{M}{M + \alpha} \tag{7}$$

To a good approximation, this universal relation should thus apply to all three regimes (free solution, dilute polymer solution and gel)!

As discussed in section 2.2.3, we actually have good reasons to believe that our S-DNA complexes are not in a stretched conformation at an electric field of E=333V/cm (the highest field obtained with our apparatus). This field is much smaller that the critical electric field required to stretch S-DNA molecules. Therefore, we do not have to worry about such effects here.

3.2.4 From ELFSE to TE

Figure 2 provides a schematic diagram with every possible electrophoresis situation with naked DNA and S-DNA. It is assumed that the relation $V_{S\text{-}DNA}(M)=V_{DNA}(M+\alpha)\times M/(M+\alpha)$ is always valid. We have:

A) DNA in a crosslinked gel: The DNA molecule reptates in the gel. The velocity decreases like $V_{DNA} \propto 1/M$ and then reaches a plateau where the velocity is molecular-size independent.

B) S-DNA in a crosslinked gel: We now have trapping electrophoresis. The velocity of the streptavidin DNA complex decreases like $V_{S\text{-}DNA} \propto M/(M+\alpha)^2$ for small molecules in the weak

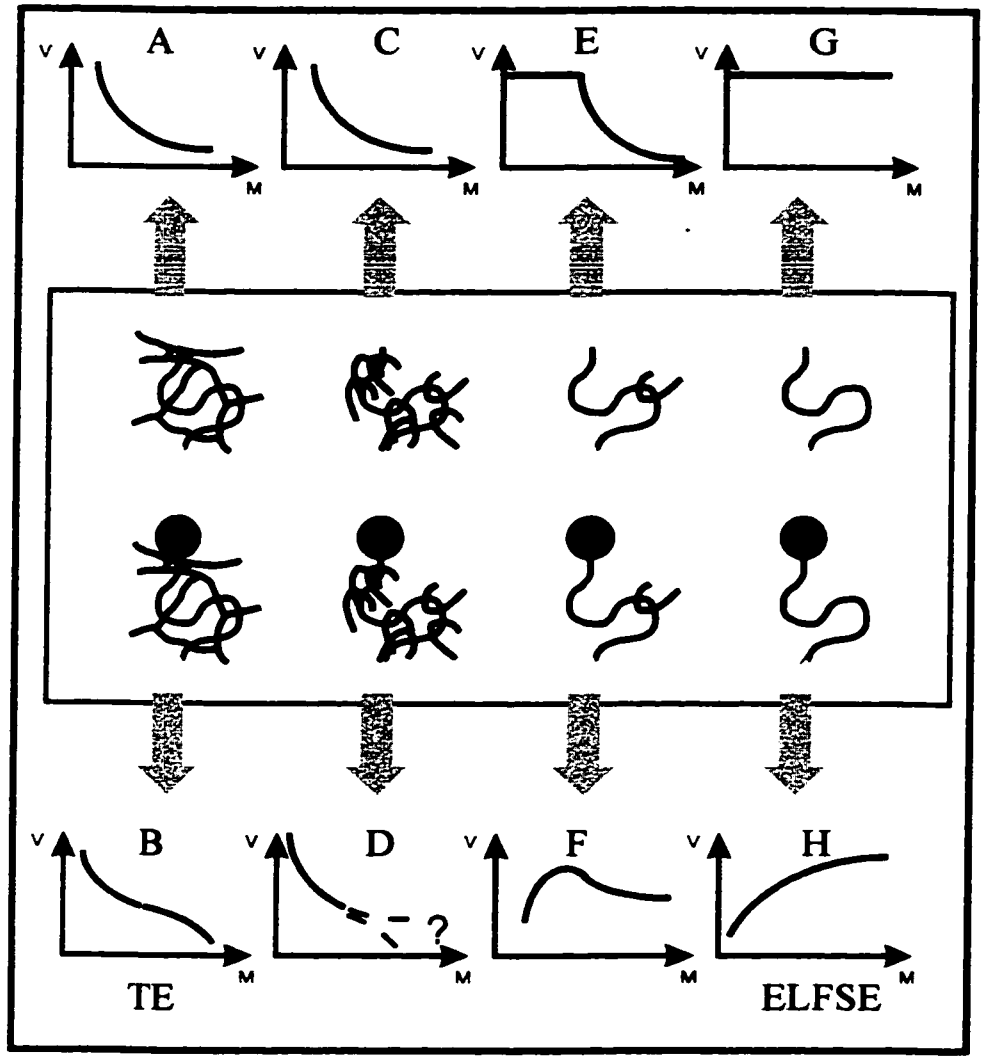


Figure 2: Schematic representation of the different electrophoretic separation mechanisms for DNA and S-DNA molecules.

trapping regime. This is followed by an abrupt velocity drop for large trapped molecules (M>M*).

C) DNA in a tight (entangled) polymer solution: The dynamics in a tight polymer solution is similar to the one observed in a crosslinked gel and $V_{DNA} \propto 1/M$ (see A).

D) S-DNA in a tight (entangled) polymer solution: The velocity decreases like $V_{S\text{-}DNA} \propto M/(M+\alpha)^2$ as in B for weak trapping. However, the existence of strong trapping can be questioned since small pores have a finite lifetime and hence there are no truly permanent topological constraints.

E) DNA in a dilute polymer solution: Very small DNA fragments do not see the polymers at all (they are too small to drag free polymers) and their velocity is size independent, much like in free solution (see G), while larger fragments are retarded by the collisions with free polymers.

F) S-DNA in a dilute polymer solution: We should observe a situation like that in H below for DNA fragments too small to entangle with free polymers, and a situation like E above for larger DNAs. We thus predict what is usually referred to as band inversion (or peak inversion in a capillary), although this situation is unlike any other reported previously.

G) DNA in free solution: We do not observe any separation (i.e., all molecules move at velocity V_0) since the charge over friction ratio is constant.

H) S-DNA in free solution (or ELFSE): The velocity of the S-DNA molecules increases monotonically like $V_{S\text{-}DNA}(M)=V_{DNA}(M+\alpha)\times M/(M+\alpha)=V_0\times M/(M+\alpha)$.

3.3 Materials and Methods

3.3.1 Electrophoresis Solutions

Solutions containing 3M of urea (BDH), 0.1 mol/L of TAPS (N-tris[Hydroxymethyl]methyl-3-aminopropane-sulfonic acid) (Sigma) and various concentration (%w/v) of poly-N,N-dimethylacrylamide polymer (POP polymer; Perkin Elmer) were used as buffers for the separations. The POP polymer, which is also used to dynamically coat the capillary walls [20,21], has a molecular weight of MW=10^6 g/mol, a polydispersity index PI≈3, and an entanglement concentration $c^* \approx 1.5\%$ w/v. This solution was buffered to pH=8.4 (the pK_a of TAPS) with NaOH.

3.3.2 DNA and S-DNA Samples

We used Pharmacia's 100 basepair ladder for our experiments. The two ends of these double stranded DNA fragments are different. This allowed us to label only one of the strands. One end has a 5′ GGCT 3′ overhang whereas the other end has a 5′ AGCC 3′ overhang. The first step of the labeling procedure consisted in blocking the GG site of the first strand by adding deoxycytosine. We mixed 5 μL (5 μg) of 100bp ladder, 5 μL of 5×Amplitaq buffer, 5 μL dCTP (5 pmol/μL), 0.5 μL Amplitaq FS, and 9.5 μL dH_20, and we allowed the filling reaction to proceed at 60°C for 20 minutes. Unincorporated nucleotides were then removed using Centrisep columns (Princeton Separations). In the second step, we mixed the modified 100bp ladder (17 μL) obtained in step one with 8 μL of 5×Amplitaq buffer, 2 μL of C-term big dye (21.5 pmoles), 4.5 μL of biotin-dUTP (148.5 pm), 0.5 μL Amplitaq FS and 8 μL dH_20 (for a total volume of 40 μL), and we allowed the reaction to proceed at 60°C for 20 minutes. Centrisep columns were then used to remove the

unincorporated nucleotides.

The solution thus obtained was diluted 10 times with distilled water. The final DNA samples were produced by mixing 3μL of this dilute solution with 10μL of formamide and 6μL of dH_2O. The streptavidin-labeled DNA (S-DNA) samples were obtained by adding an appropriate quantity of purified streptavidin to this solution [20].

3.3.3 Electrophoresis Conditions

The running temperature (30°C) was controlled through the thermostatic plate of the ABI PRISM™ 310 Genetic Analyser. Experiments were performed on 45 cm-long capillaries and the applied voltage was 15 kV (E=333V/cm). Sample injection was performed at 15 kV for 5 seconds. The distance between the injection end of the capillary and the detector was 34 cm.

3.4 Results

3.4.1 The Transition Region

If we plot the velocity of S-DNA fragments of different molecular sizes vs. the POP polymer concentration (Figure 3) we observe a remarkable transition (we actually plotted the difference between the velocity of three different molecular sizes (100, 300 and 1100 bases) and the velocity of a 18-base fragment in order to eliminate fluctuations due to the variations of the residual electroosmotic flow between runs). For concentration ranging from 0.01% to 0.32%, the relative velocity of the 1100 base fragments (stars) is almost constant, but it decreases quickly for larger concentrations. This is due to the fact that the probability of encountering a POP polymer is too

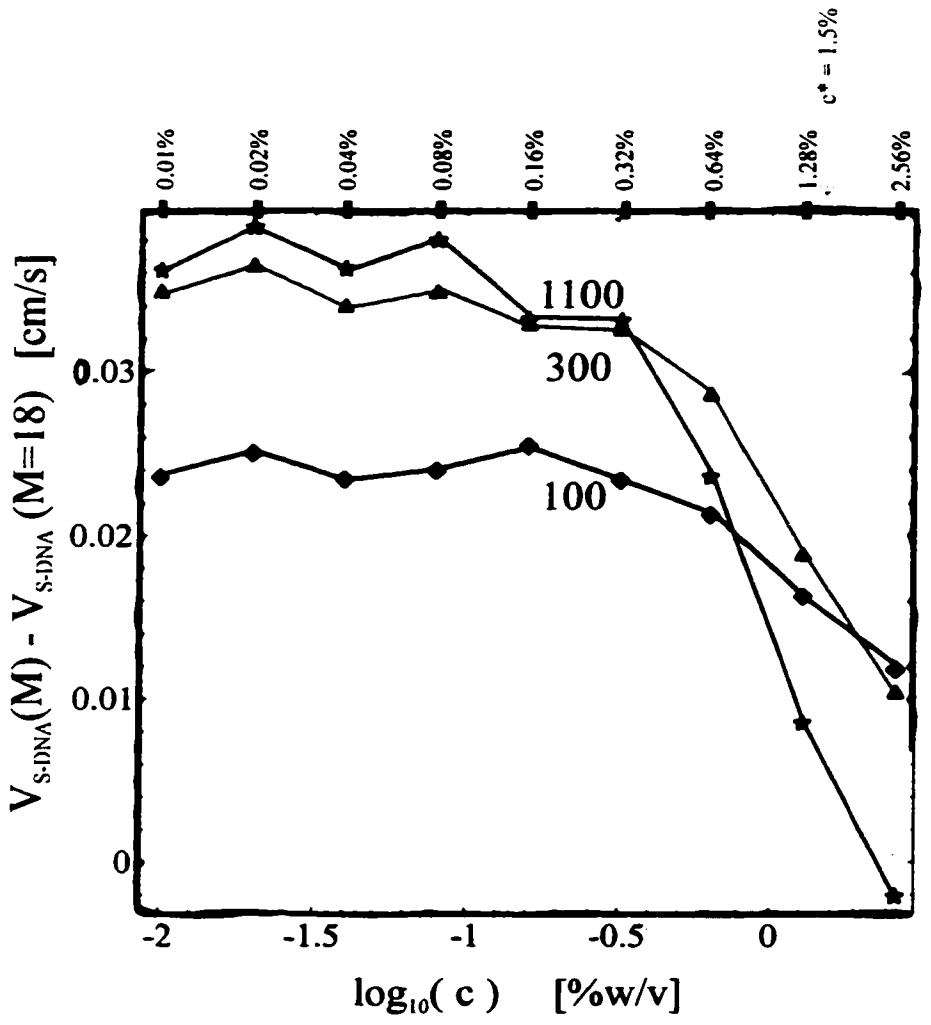


Figure 3: Velocity of three S-DNA molecules (100 bases, blue diamonds; 300 bases, red triangles; and 1100 bases, green stars) minus the velocity of a 18-base S-DNA fragment as a function of the concentration of POP polymer. We observe the transition from ELFSE to reptation.

small to affect the dynamics for $c \leq 0.32\%$. At higher concentrations, we have a transition from ELFSE to reptation in entangled polymer solutions. A similar transition is observed for the other molecular sizes. Indeed, we see that the large molecules move faster than the small ones ($V_{1100}>V_{300}>V_{100}$) for very small polymer concentrations (e.g., 0.01%), typical of the ELFSE regime, while we have the opposite situation ($V_{100}>V_{300}>V_{1100}$) at high concentrations $c \geq 2.56\%$, indicating the presence of reptation sieving. Therefore, there is a remarkable intermediate region where band inversion (peak inversion) occurs ($V_{300}>V_{100}>V_{1100}$). This was one qualitative prediction of the relation $V_{S\text{-}DNA}(M) \approx V_{DNA}(M+\alpha)\times M/(M+\alpha)$ (see Figure 2F). Clearly, such a regime is to be avoided in practice.

3.4.2 ELFSE Separations

Figure 4A shows the electropherogram of our naked DNA sample at c=0.04% POP. This concentration is much lower than the entanglement concentration $c^* \approx 1.5\%$. We observe no separation, i.e. DNA fragments of various sizes all co-migrate and reach the detector at the same time. Figure 4B shows the electropherogram for the S-DNA sample in the same condition. We can see that small S-DNA molecules move slower than larger ones. Figure 4C presents a plot of velocity vs. molecular size for both naked and labeled DNAs. We observe an horizontal line for naked DNA while the velocity increases with molecular size for S-DNA.

We will now examine whether the relation $V_{S\text{-}DNA}(M)=V_{DNA}(M)\times M/(M+\alpha)$ holds in the ELFSE regime. First, we note that this relation can be rewritten as:

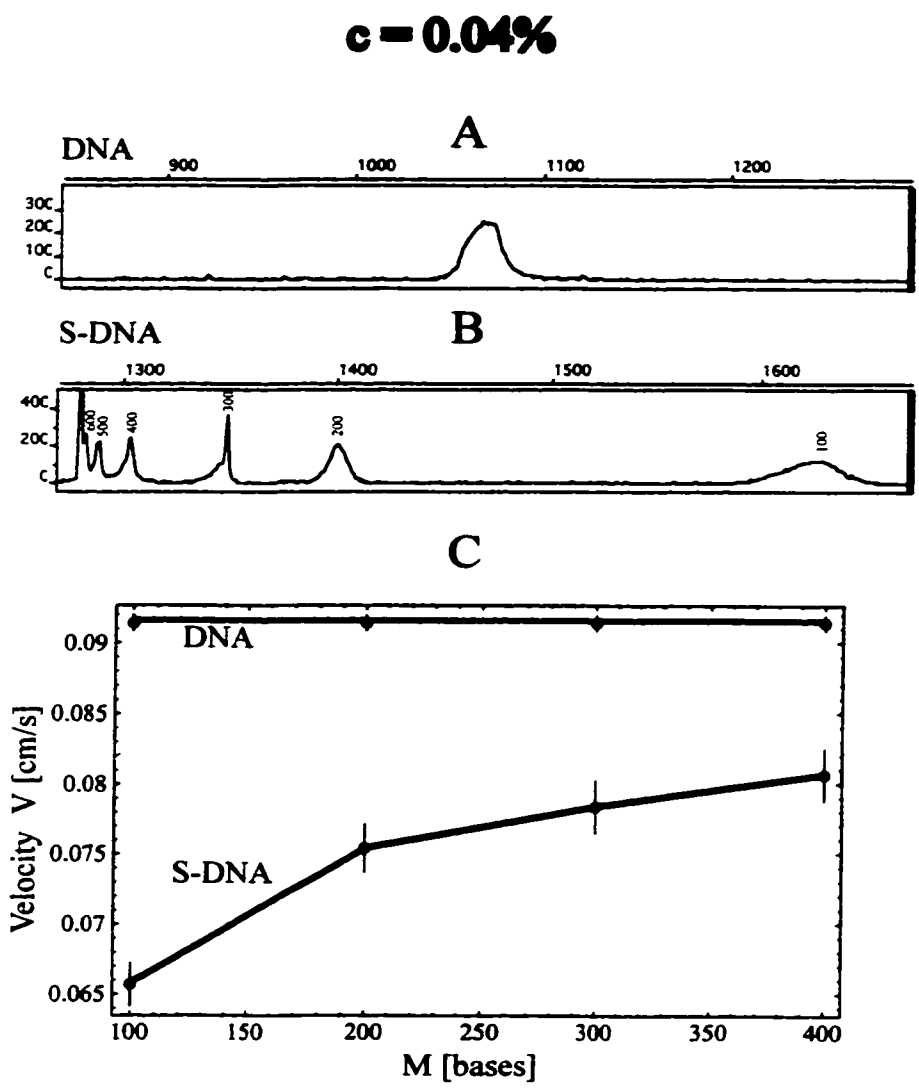


Figure 4: Electropherogram (the y-axis is the intensity of the signal detected while the x-axis represents time in the units used by the Perkin-Elmer software (note that $t=(30+t_{Perkin-Elmer}/3.1022)s$) for A) DNA and B) S-DNA in a low POP concentration buffer (c=0.04%). Figure 4C is a plot of the velocity V vs. the molecular size M for DNA (red) and S-DNA (blue) molecules.

$$\frac{V_{DNA}(M)}{V_{S\text{-}DNA}(M)} = 1 + \frac{\alpha}{M} \tag{8}$$

Therefore, if we plot $V_{DNA}/V_{S\text{-}DNA}$ vs. 1/M, we should get a straight line with a slope α that crosses the origin at 1. We can see on Figure 5 that the data satisfy this relation. The value of α thus obtained is $\cong$34.2, and the y-intercept is 1.05 (which is very close to the expected value of 1).

At an intermediate concentration c=0.64%$\approx$c*/2, separation also occurs for naked DNA (Figure 6). As we can see in Figure 6A, the small DNA fragments exit the capillary first. In other words (Figure 6C), the velocity decreases smoothly for all DNA sizes. The situation is rather complex for the S-DNA molecules, and the electropherogram (Figure 6B) looks very strange. The 100-base fragment is one of the slowest fragments. The velocity increases for small molecules (M<300 bases) and reaches a maximum for M$\approx$300 bases, as shown in Figure 6C. The velocity then starts to decrease much like for naked DNA. In order to test if the relation $V_{S\text{-}DNA}(M) \cong V_{DNA}(M+\alpha)\times M/(M+\alpha)$ is valid in this regime, we plotted in Figure 7 the ratios of the velocities at c=0.64% and c=0.04% for both S-DNA and DNA as a function of the effective molecular size (M for DNA and M+α for S-DNA). Our results indicate that, in agreement with eq 7, we do have:

$$\frac{V_{S\text{-}DNA}(M, c_1)}{V_{S\text{-}DNA}(M, c_2)} \approx \frac{V_{DNA}(M + \alpha, c_1)}{V_{DNA}(M + \alpha, c_2)} \tag{9}$$

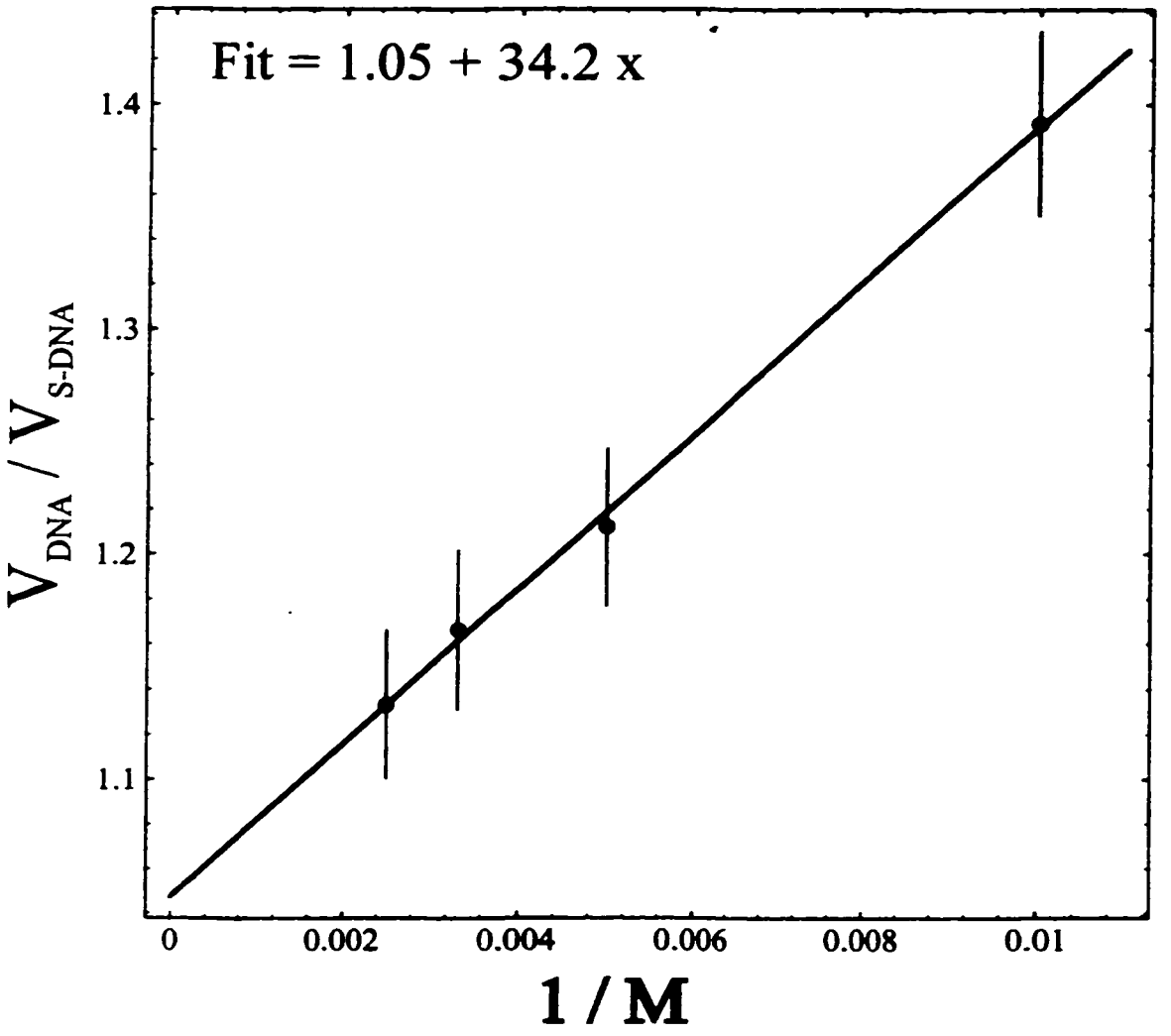


Figure 5: $V_{DNA}/V_{S\text{-}DNA}$ vs. 1/M for a low concentration buffer (c=0.04%). The linear behavior demonstrates that the relation $V_{S\text{-}DNA}(M)=V_{DNA}(M)\times M/(M+\alpha)$ is verified.

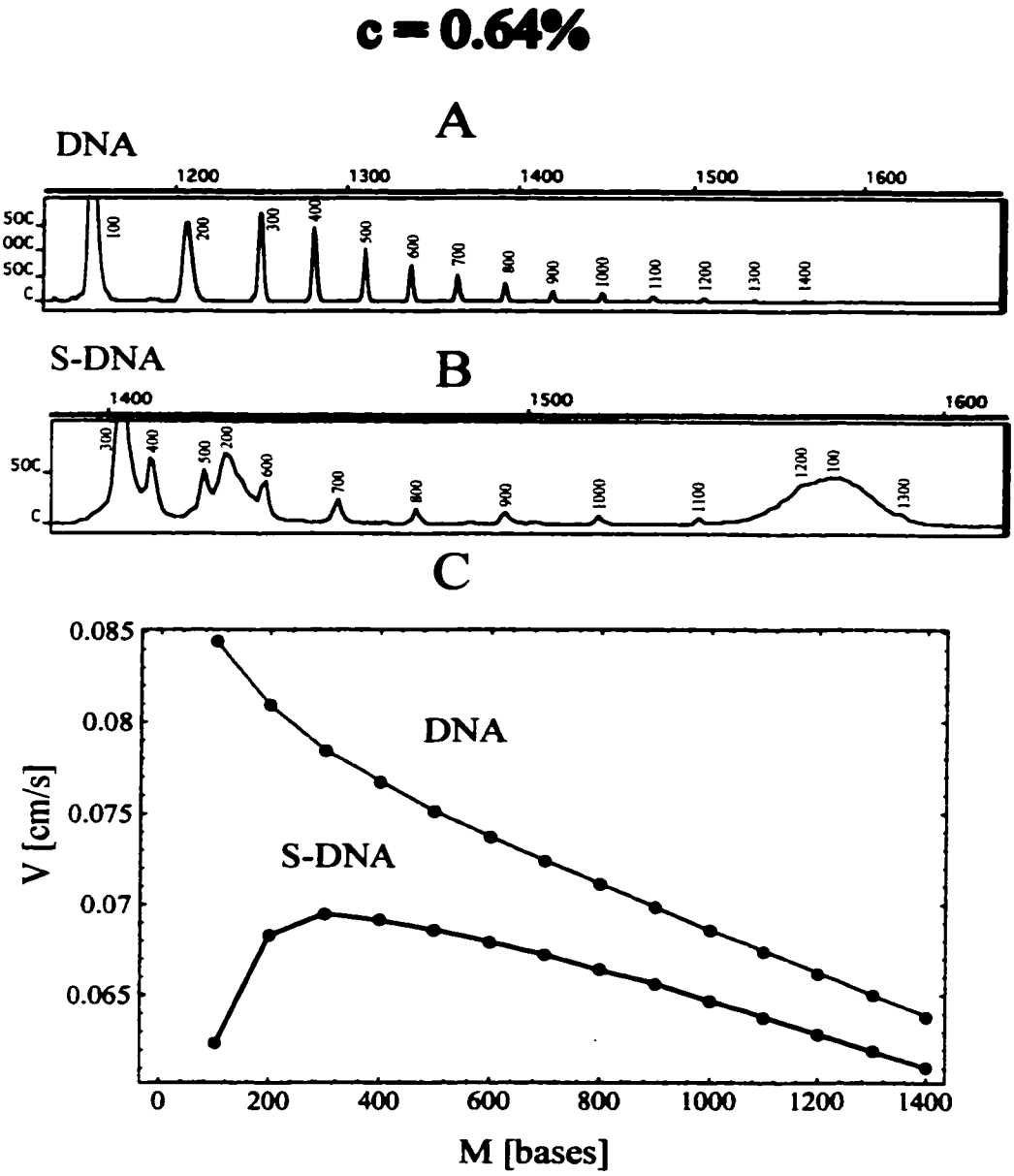


Figure 6: Electropherogram of A) DNA and B) S-DNA in an intermediate POP concentration buffer (c=0.64%). Figure 4C is a plot of the velocity V vs. the molecular size M for DNA (red) and S-DNA (blue) molecules. The predicted band inversion is observed for S-DNA molecules.

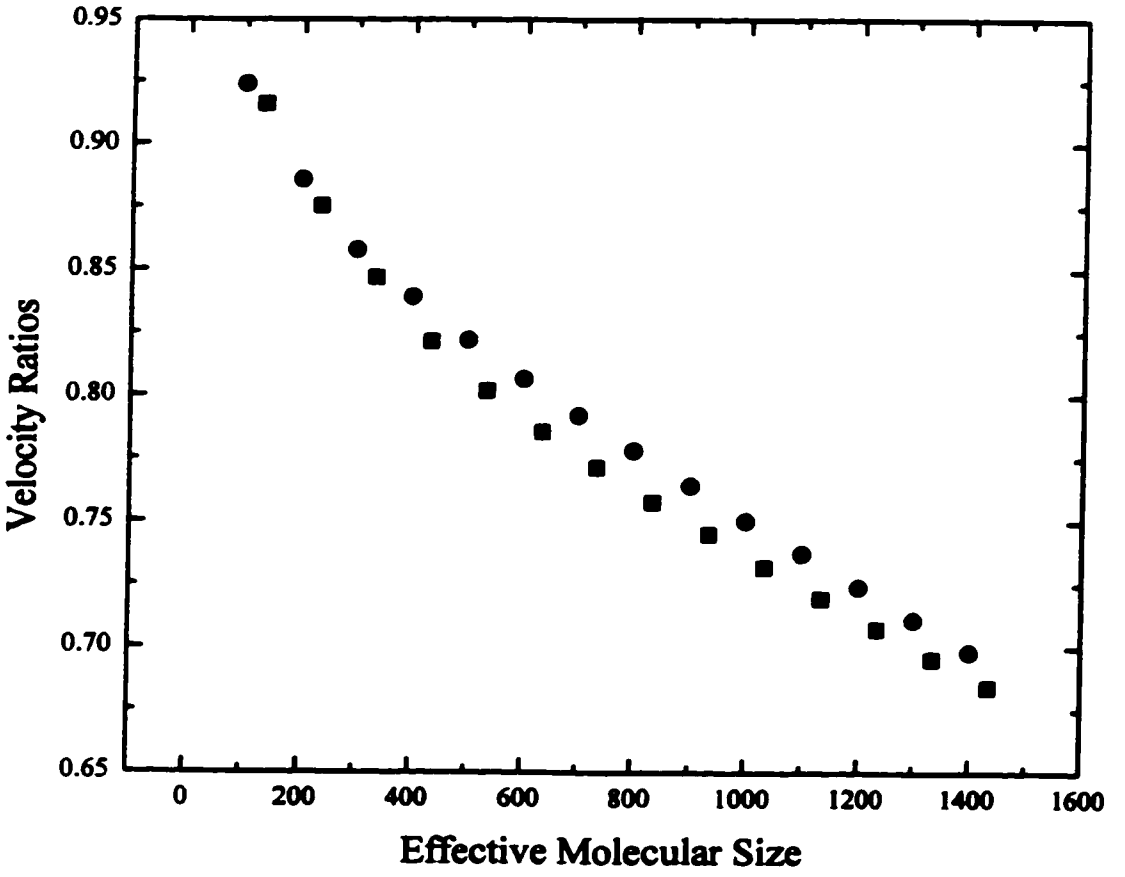


Figure 7: Velocity ratio (velocity at 0.64% POP/velocity at 0.04% POP) for DNA (●) and S-DNA (■) vs. the effective molecular size (i.e. M for DNA and M+α for S-DNA, with α=34.2).

Note that there was no fitting involve in obtaining Figure 7 (we simply use the value α=34.2 obtained for c=0.04%). Figure 8 shows the same graph for c=1.28%≈c* (using α=34.2 again); we see that even for a concentration close to c*, eqs 7 and 9 are satisfied.

3.5 Discussion

In this chapter, a theoretical model for the migration of end-labeled DNA molecules in dilute polymer solutions was developed. This model generalizes that of Hubert et al. [13] which applied to the migration of naked DNA fragments. Our new theoretical model leads to a surprisingly simple relation between the velocities of DNA and S-DNA molecules. It turns out that the relation is the same as the one obtained in free-solution and in gels. This is a very surprising result since the separation mechanisms are very different in these three cases.

We also obtained experimental confirmation of this universal relationship. Our experiments used concentrations of poly-N,N-dimethylacrylamide ranging from very dilute (c=0.01%) to entangled (c=2.56%). We were able to see the transition from free-solution to gel electrophoresis. At low polymer concentrations, the large S-DNA molecules are faster than the small ones while we observed the opposite situation at high polymer concentrations. At intermediate polymer concentrations (c of the order of 0.64%), the small molecules are in a ELFSE-like separation regime while the large molecules are in a gel electrophoresis-like regime. This fascinating regime was correctly predicted by our theoretical model.

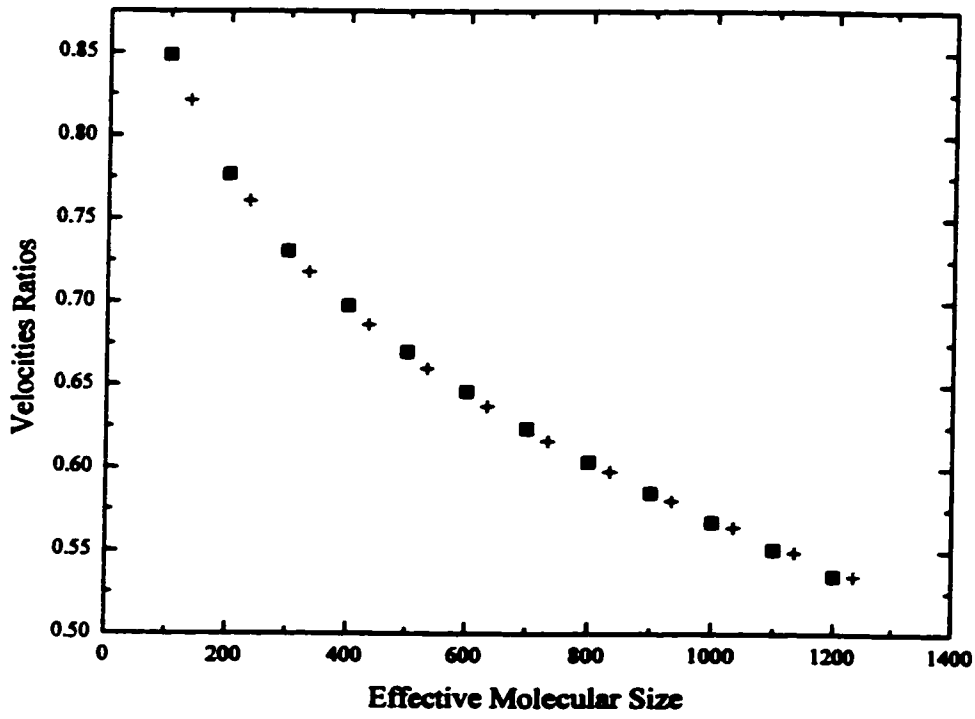


Figure 8: Velocity ratio (velocity at 1.28% POP/velocity at 0.04% POP) for DNA (■) and S-DNA (✚) vs. the effective molecular size (i.e. M for DNA and M+α for S-DNA, with α=34.2).

Our results actually demonstrate a crucial point about the conformation of the streptavidin-DNA complexes during ELFSE experiments. In the first paper about ELFSE [19], it was assumed that S-DNA molecules were stretched when electrophoresed in free solution. This assumption is rather appealing since one does indeed expect the two ends of a streptavidin-DNA complex to have different natural velocities, and hence to lead to DNA stretching. However this argument neglects the very large internal entropic elasticity of the DNA coil and the hydrodynamic coupling between the DNA and its label. In light of the theories developed by Long et al. [22-26], the stretching of S-DNA was questioned. According to these theories, the electric field necessary to stretch S-DNA is very large compared to the electric field used in our experiments (see Chapter 2). Our results can be used to clarify the situation. If the streptavidin-DNA complex is in its random coil conformation (like naked DNA), the effect of the collisions with the free POP should be the same for both species. On the other hand, if the S-DNA molecules alone were stretched, the probability of colliding with a free polymer chain would be much lower (because of the much reduced collision cross-section). Our results indicate that the POP polymers have the same effect on DNA and S-DNA molecules since eq 9 was verified experimentally. This means that those molecules adopt the same conformation, thus establishing that S-DNA is not stretched under these conditions (as mentionned in Chapter 2). This is an important conclusion of this chapter.

3.6 References

[1] de Gennes, P. G.; Scaling Concepts in Polymer Physics. Cornell University Press, NY, 1979.

[2] Doi, M.; Edwards, S. F.; The Theory of Polymer Dynamics, Oxford University Press, NY, 1986.

[3] Duke, T. A. J.; Viovy, J. L.; Semenov, A. N., *Biopolymers*, **34**, 239, 1994.

[4] Semenov, A. N.; Duke, T. A. J.; Viovy, J. L.; *Phys. Rev. E*, **51**, 1520, 1995.

[5] Heller, C.; Duke, T. A. J.; Viovy, J. L.; *Biopolymers*, **34**, 249, 1994.

[6] Viovy, J. L.; Mechanisms of polyelectrolyte gel electrophoresis. (Submitted for publication), 1999.

[7] Slater, G. W.; Electrophoresis theories, in *Analysis of nucleic acids by capillary electrophoresis* (Heller, C., ed.), Vieweg & Son, Wiesbaden, pp.24-66, 1997.

[8] Barron, A. E.; Blanch, H. W.; *Separation and Purification Methods*, **24**, 1, 1995.

[9] *Pulsed-Filed Gel Electrophoresis Protocols, Methods, and Theories* (Burmeister, M.; Ulanovsky, L., eds.), Humana Press, Totowa.

[10] Schwartz, D. C.; Cantor, C. R.; *Cell*, **37**, 67, 1984.

[11] Cottet, H.; Gareil, P.; Viovy, J. L.; *Electrophoresis*, **19**, 2151, 1998.

[12] Barron, A. E.; Blanch, H. W.; Soane, D. S.; *Electrophoresis*, **15**, 597, 1994.

[13] Hubert, S. J.; Slater, G. W.; Viovy, J. L; *Macromolecules*, **29**, 1006, 1996.

[14] Ulanovsky, L.; Drouin, G.; Gilbert, W.; *Nature*, 1990, **343**, 190.

[15] Desruisseaux, C.; Slater, G. W.; Drouin, G.; *Macromolecules*, **19**, 6499, 1998.

[16] Slater, G. W.; Desruisseaux, C.; Villeneuve, C.; Guo, H. L.; Drouin, G.; *Electrophoresis*, **16**,

704, 1995.

[17] Défontaines, A.-D.; Viovy, J. L.; *Electrophoresis*, **14**, 8, 1993.

[18] Défontaines, A.-D.; Viovy, J. L.; *Electrophoresis*, **15**, 111, 1994.

[19] Mayer, P.; Slater, G. W.; Drouin, G.; *Anal. Chem.*, **66**, 1777, 1994.

[20] Ren, H.; Karger, A. E.; Oaks, F.; Menchen, S.; Slater, G. W.; Drouin, G.; *Electrophoresis*, **20**, 2501, 1999.

[21] Madabhushi, R. S.; *Electrophoresis*, **18**, 2393, 1997.

[22] Long, D.; Viovy, J. L.; Ajdari, A.; *Phys. Rev. Lett.*, **76**, 3858, 1996.

[23] Long, D.; Viovy, J. L.; Ajdari, A.; *J. Phys.: Condens. Matter*, **8**, 9471, 1996.

[24] Long, D.; Viovy, J. L.; Ajdari, A.; *Biopolymers*, **39**, 755, 1996.

[25] Long, D.; Ajdari, A.; *Electrophoresis*, **17**, 1161, 1996.

[26] Long, D.; Dobrynin, A. V.; Rubinstein, M.; Ajdari, A.; *J. Chem. Phys.*, **108**, 1234, 1998.

Chapter 4

The Gel Edge Electric Field Gradients in Denaturing Polyacrylamide Gel Electrophoresis*

It has previously been shown [1-3] that zones of altered salt concentration form close to the gel ends during denaturing polyacrylamide gel electrophoresis. Here we show that the field can reach up to three times its normal mean value a few cm in front of the loading wells when 0.5×TBE is used as the gel buffer. We also demonstrate that this electric field is mostly due to the difference in ion transference numbers at the gel/buffer interface caused by the high viscosity of the urea solution contained in the gel. These field gradients lead to increased band width and force us to redefine both the electrophoretic mobility and the mean field intensity. We discuss some methods that can be used to minimize the effects of this gradient.

---

*The results presented here have been published in *Electrophoresis*, **19**, 627, 1998.

4.1 Introduction

It is well-known that the electric current often decreases with time during a gel electrophoresis (see, e.g., [4]). Indeed it takes a fairly long time before the electric current reaches a time-independent, steady-state value. During that period of time a zone of altered salt concentration is created, resulting in an electric field gradient along the gel. This phenomenon has been studied recently in the context of DNA sequencing in denaturing polyacrylamide gels [1,3]. Using multiple DNA loadings, Mayer, Slater and Drouin [1,3] noticed that the velocity of DNA fragments is substantially larger in the first few cm of a standard 4% polyacrylamide gel with 8M urea and a 0.5×TBE buffer, but essentially uniform in the middle of the gel. Figeys, Renborg and Dovichi [2] observed a larger electric field intensity in the first few cm of a capillary filled with non-crosslinked denaturing polyacrylamide. Both the initial current drop and the electric field gradients along the direction of migration of DNA are due to the depletion of ions in the region of the gel near the negative electrode [5-7]. A direct measure of the electric field gradient was performed by Houri et al. [8] with platinum electrode probes inserted into a polyacrylamide gel containing a Tris-glycine buffer. They observed that the electric field profile can change significantly with time. Interestingly enough, a prerun of opposite polarity was observed to greatly decrease the field gradient at the loading-end of the gel [1]. It is important to note that since the gradients take long to grow and reach their steady-state shape and amplitude, the field gradients may be strongly history-dependent if proper preruns are not used. This may make reproducibility a serious issue for quantitative studies of DNA sequencing gel electrophoresis (e.g., studies aimed at testing specific theories).

Spencer explained why ion concentration gradients develop near a buffer/gel interface for two- [5] and three- [6] ion systems in a series of papers that, unfortunately, are not well known. Consequently, these effects have been largely overlooked in the literature. In order to make this introduction more complete, we thus offer below a simple presentation of the physics of two-ion systems (mostly based on the work of Spencer). Three (or more) ion systems are more complex to understand but lead to qualitatively similar predictions. It is not the goal of this chapter to develop a model describing the complexity of the TBE buffer [9,10]. However, because the ion/field gradients are common but misunderstood (when not completely ignored), we believe that the simple ideal example presented below is both a useful framework to understand the rest of the chapter and an accessible introduction to the ionic gradient problems studied by Spencer.

4.1.1 Ionic Concentrations near an Interface

In order to illustrate the mechanisms responsible for an accumulation (or depletion) of ions at the boundary between two media (the entrance of the sequencing gel in our case), we start by studying the currents through two imaginary surfaces, one on each side of the interface. We first calculate the number of positive ions (per unit time) that accumulate in (or leave) the volume limited by those two surfaces. Since in our gels the positive ions are moving from the gel to the buffer, this number is simply given by $(I_{+,gel}-I_{+,buffer})/F$, where $I_{+,a}$ is the current due to the positive ions (not the total current) in medium a, and F is Faraday's constant (F simply converts the charge into the number of moles). The fraction of the total current I (where we choose I>0 for simplicity) that is transported by the positive ions in the less retarding medium (the buffer) is given by the transference

number $T=\lambda_+/(\lambda_+ + \lambda_-)$, where λ_+ and λ_- are the equivalent ion (+ or -) conductivities (proportional to mobilities). Therefore, $I_{+,buffer}=I\times T$. Similarly, $I_{+,gel}=I\times T_r$ for the most retarding medium (the gel), where the transference number is now given by $T_r=R_+\lambda_+/(R_+\lambda_+ + R_-\lambda_-)$, with $R_\pm\lambda_\pm$ being the equivalent ion conductivities in the gel and with $R_\pm \leq 1$ being the retardation factor for the two ions. The number of positive ions accumulating in this imaginary volume is then simply given by $I\times(T_r-T)/F$ per unit time, while the accumulation of negative ions is $I\times[(1-T)-(1-T_r)]/F=I\times(T_r-T)/F$, respecting electroneutrality. The number of charges accumulated (or lost) after a certain time t in the vicinity of the interface is thus given by:

$$\Delta N_{ions} = \int_{\Delta x}[c(x) - c_0]\,dx = \frac{[T_r - T]}{F}\int_0^t I(t')\,dt' \qquad (1)$$

where c_0 and $c(x)$ are the initial (uniform) and altered ion concentrations (number of mole of ions per unit length), respectively, and the x-integral is over the width Δx of the interface region. Clearly, there will be an accumulation of ions at the buffer/gel interface if $\Delta T \equiv T_r - T > 0$ and a depletion if $\Delta T < 0$. Obviously, we will see the opposite situation at the other end of the gel. Coming back to the definition of T and T_r, we note that $\Delta T = T_r - T > 0$ when $R_- < R_+$ (the anions are more retarded than the cations) while $\Delta T < 0$ when $R_+ < R_-$ (the cations are more retarded than the anions). So the inequality for the retardation factors R_+ and R_- determines how the system will evolve (depletion or accumulation of ions at the interfaces).

4.1.2 The Local Electric Field

Rather than dealing with local ionic concentrations, it is more convenient to look at the local electric field. The solution conductivity σ is related to the ion concentrations via the relation $\sigma=\sum_i \Lambda_i c_i$, where Λ_i is the molar conductance and c_i is the concentration of the i-th ionic species. Since the resistance R is inversely proportional to σ, it will be higher (lower) where there is a depletion (accumulation) of ions. The current I being uniform along the direction of migration in the steady-state, the local potential drop across a width Δx is given by $\Delta V(x)=IR(x) \sim 1/c(x)$. Hence, a local depletion (accumulation) of ions leads to an increased (decreased) local electric field intensity $E(x)=\Delta V(x)/\Delta x$. Therefore, a zone of increased local field intensity E(x) may be observed near the loading wells (together with a zone of lower fields near the other end of the gel) if the cations are more retarded than the anions.

4.1.3 Evolution of the Ionic Profile

The ion concentration gradient builds up at the gel/buffer interface in response to the difference ΔT in transference numbers across this interface. This gradient eventually eliminates the difference T_r-T at the gel/buffer interface because $T_r=T_r(c(x))$ is a function of the ion concentration c(x) (the system evolves towards a local equilibrium situation). After the new transference numbers match at the interface, a new mismatch in transference numbers will be present inside the gel, at the end of the depletion (or accumulation) region. In this case, the mismatch is due to the inhomogeneous ion concentration profile c(x) inside the gel. The system will further evolve in order to eliminate this new interface. Consequently, there will be a net migration of the ionic profile c(x).

Kohlrausch [11] and von Laue [12] derived differential equations to account for the velocity v_x of this interface for a general case. These equations lead to

$$v_x = \left(\frac{\partial x}{\partial t}\right)_c = \frac{I}{F}\left(\frac{dT}{dc}\right) \qquad (2)$$

when the system is restricted to a uniform cross-section containing only two ionic species. If the transference number T is not a function of the concentration of ions c, we obtain $v_x=0$ and the concentration profile does not change with time. When the transference number T is a function of c, however, we must have a drift of the profile. The direction of the drift will be determined by the sign of dT/dc and is the same for the two gradients (at the beginning and at the end of the gel). This implies that one of the gradients will tend to disappear (move out of the gel) while the other will move inside the gel.

4.1.4 Effects of Ionic Gradients on Gel Electrophoresis

The electric field inhomogeneities resulting from these fundamental effects have practical consequences. First, one cannot define the electrophoretic velocity of an analyte as v=x/t, where x=x(t) is the total distance migrated in a time t, since the instantaneous velocity is itself position dependent (v=v(x)). Second, because the field gradient stabilizes only after a long period of time [1-3], mobility measurements can be very dependent upon the history of the system if they are performed during the redistribution of ionic concentrations in the system. Third, band broadening is strongly affected by these inhomogeneities because of the band focusing and anti-focusing effects of negative and positive field gradients, respectively [13-15]. This latter effect also means that one

cannot estimate the diffusion coefficient from the width of the bands after a fixed migration time. In fact, the band widths are also very dependent upon the history of the system.

In this chapter, we present a study of the effect of the field gradients present near the gel edges on the migration of a small DNA molecule through a polyacrylamide gel under typical denaturing sequencing conditions. For this study, we use the multiple loading approach described previously [1-3]. Since the mobility of a small molecule (a small DNA fragment in our case) is strictly field independent, its instantaneous velocity v(t)=v(x(t)) during the electrophoretic migration can be related to the local field intensity E(x) along the gel. Moreover, from the shape of the band we are able to directly observe the effect of the local field inhomogeneities on band broadening. We also show that the high urea concentration of our denaturing gel is the major cause for the observed gradients and that the polyacrylamide fibers only play a minor role. The main goal of this chapter is to demonstrate the negative impact of intrinsic field gradients on the analysis and optimization of DNA sequencing systems.

4.2 Materials and Methods

4.2.1 Polyacrylamide Gels

Four percent polyacrylamide gels with a 19:1 acrylamide (BDH) to bisacrylamide (BDH) ratio were used. They contained 8M urea (BDH) and a solution of 0.5×TBE [44.5 mM Tris (BDH), 44.5 mM boric acid (BDH), 1 mM EDTA (BDH)] was used as the running buffer. The 0.4 mm thick gels were prepared on a Pharmacia-LKB Macrophor apparatus. Polymerization was initiated with the addition of 0.10%w/v ammonium persulfate (BDH) and 0.10% v/v TEMED (BDH), and the gels

were left to polymerize at room temperature for at least 12 hours.

4.2.2 DNA Samples

A radiolabeled (^{32}P) 21 base long single-stranded oligonucleotide was used to map the velocity field along the direction of net electrophoretic migration. It was prepared by annealing 5 pmoles (1 μL) of the 17 base long T7 oligonucleotide primer (Stratagene) to 5 pmoles (25 μL) of pBluescript SK+ single-stranded DNA in the presence of 8μL of 5× Sequenase buffer (200mM Tris-HCl, pH 7.5, 100mM $MgCl_2$, 250mM NaCl) and 3.5μL of sterile distilled water. Labeling was performed by adding 0.5μL of ^{32}P dCTP (3000 Ci/mmoles, Amersham), 1μL of 100mM dGTP (Pharmacia) and 1μL of Klenow (Pharmacia) and incubating at room temperature for 15 minutes. The reaction was stopped by adding 1μL of 0.5M EDTA. The salts were removed from the samples (in order to reduce electrodispersion) using two dyalises on a 0.025μm nylon membrane (Millipore) floating on 250mL of 0.5×TBE at room temperature.

4.2.3 Electrophoresis Conditions

The running temperature (55°C) was controlled through the thermostatic plate of the Pharmacia-LKB apparatus (note that this system does not allow recirculation of the gel tank buffer). The temperature on the surface of both plates was measured with a temperature probe and the fluctuations did not exceed ±3°C. The temperature between the glass plate was not measured directly. Experiments with voltages ranging from 500 to 4000 volts were performed on 52.3-cm-long gels using a Fisher Biotech FB600 and a Pharmacia-LKB power supplies. The gels were prerun at

the selected voltage for at least 8000 volts×hours, e.g. for 2 h at 4000 V or 16 h at 500 V [1,3]. After this prerun, the electric current had fallen to about 40% of its initial value, and remained constant during the whole electrophoresis (less than 2% variation). Such a long prerun is necessary in order to obtain reliable results that do not change significantly with time [1], i.e. to remove the history dependence of the results. In order to measure the local electrophoretic velocity v(x) at position x along the migration path, we loaded one DNA sample every Δt minutes (typically Δt=0.5 to 4 minutes) in consecutive wells. The distance Δx between two consecutive bands thus directly gives the velocity v(x), as described in the next section. This straightforward method was shown to give excellent results [1] in the presence of non-uniform fields.

4.2.4 Measurement of the Distance Migrated

Gels were transferred to chromatography paper (3MM Whatman), dried, and exposed to X-rays films. Only undistorded gels were kept for analysis. The latter were scanned with a desktop 400 dpi scanner (HP Scanjet IIc). The distance of migration of each band was then measured on these scans with a C++ graphical software written by P. Mayer. The center of the optical density peaks corresponding to the bands and to the loading well were located visually with a cursor. The uncertainty of the measurements were approximately ±0.12 mm which is about 5 times better than what one usually achieves using a ruler directly on sequencing autoradiograms.

4.2.5 Data Analysis

In general, one may have to define a number of local electrophoretic conditions at position

x, such as the field intensity E(x), the temperature T(x), the buffer viscosity $\eta(x)$, the ionic concentration c(x), etc. Our results strongly suggest that the gel temperature $T(x)=T_0$ and the buffer viscosity $\eta(x)=\eta_0$ are essentially uniform in the gel because the magnitude of the velocity effect we observed (e.g. a 3-fold increase of the velocity a few cm away from the gel end) could not be explained by any realistic temperature (i.e., the gel would need to boil to explain a 3-fold increase) or viscosity gradients (small gradients may exist, of course, but they are not relevant to our problem). Another concern here is that of urea diffusion outside the gel. To study this effect, we examined gels that had gone through long preruns. Although most of the gel was slightly opaque, the region in contact with the buffer were very clear up to 5 mm (indicating the absence of urea in that part of the gel). Obviously, this urea gradient cannot explain the very large velocity gradient observed up to 5 cm into the gel (see below). In conclusion, none of the trivial physical inhomogeneities that always exist in such a sequencing instrument can explain the enormous velocity gradients observed here. Therefore, we conclude that the velocity gradients, observed at the edges of the gel, are directly related to the local field gradients.

If V is the voltage across the gel (the voltage drop in the reservoir has been measured and is negligible), the space-averaged (often called "applied") electric field is given by

$$\langle E \rangle_x = \frac{1}{L} \int_0^L E(x)\, dx = \frac{V}{L} \equiv E_0 \qquad (3)$$

where L is the length of the gel. This is the standard definition because it is normally assumed that the field is uniform along the gel ($E(x)=E_0$). Note that this is averaged over space and not time,

hence the subscript x. The electrophoretic mobility μ of a very small DNA molecule is field-independent. However, it can be position-dependent since the local conditions (T, η, etc.) may vary. In this case, we must write

$$v(V,x) = \mu(x)\ E(V,x) \tag{4}$$

In the presence of gradients, it thus becomes difficult to define and measure a "mean" mobility in order to test theoretical predictions. In fact, the mean mobility is meaningless.

We obtained the position-dependent velocity v(x) of the 21 base DNA fragment using the differential method described in ref. [1]. If the sample in lane i (with i= 1, 2...) is at position x_i after having migrated for a time $t_i=i\Delta t$, where Δt is the time difference between consecutive loadings, we define the velocity $v_{i+½}=(x_{i+1}-x_i)/\Delta t$ as the instantaneous (or local) velocity of the molecule at time $t=t_i+½\Delta t=(i+½)\Delta t$ and position $x=½(x_i+x_{i+1})$. This is a straightforward procedure which does not assume that the field is uniform but only that Δt is small enough. The standard alternative, which defines $v_i=x_i/t_i$, is clearly invalid if the velocity is not uniform and can lead to serious errors [1]. The space-averaged velocity is then defined as

$$\langle v \rangle_x = \frac{1}{L} \int_0^L v(x)\ dx \tag{5}$$

Note that this is not equal to the time-average-velocity $\langle v \rangle_t=\int v(t)dt/\tau=L/\tau$, where τ is the time required for the molecule to migrate through the entire gel of length L (the elution time). There is often confusion in the literature because many authors use the space-averaged field given in eq 3 together with a time-averaged velocity; this is valid only if the field is uniform. In fact, it is easy to

show that $\langle v \rangle_t \leq \langle v \rangle_x$, with the equality sign being valid for uniform fields only. Finally, the elution time τ (a useful concept to draw conclusions about capillary systems) is given by

$$\tau = \int_0^L \frac{dx}{v(x)} \tag{6}$$

If the system is uniform, this gives simply $\tau=L/v=L/\mu E_0$. If the field in non-uniform, however, the elution time is always larger [13].

4.3 Results

4.3.1 The Existence of the Velocity Gradients

Figure 1 shows a typical gel pattern obtained using our multiple-loading method. We have drawn a time axis along the series of loading wells and a position axis along the direction of migration. The gel picture is thus a x(t) vs. t diagram. The instantaneous velocity v(t) is given by the slope of the curve at time t. It is clear that this velocity is small between the first two points (the molecules start at x(0)=0, as shown by the symbolic point at the origin of the diagram), then much larger for a few minutes (over a few cm), before it decreases to reach a constant value in the middle section (≈43 cm) of the gel. The velocity then decreases slightly in the last 2-3 cm of gel. We also note that the band broadens very quickly in the first 3 cm, then sharpens between 3 and 6 cm, and sharpens again near the end of the gel. In fact, we observe band sharpening when the velocity decreases, and a sudden band broadening when the field increases. This is consistent with our

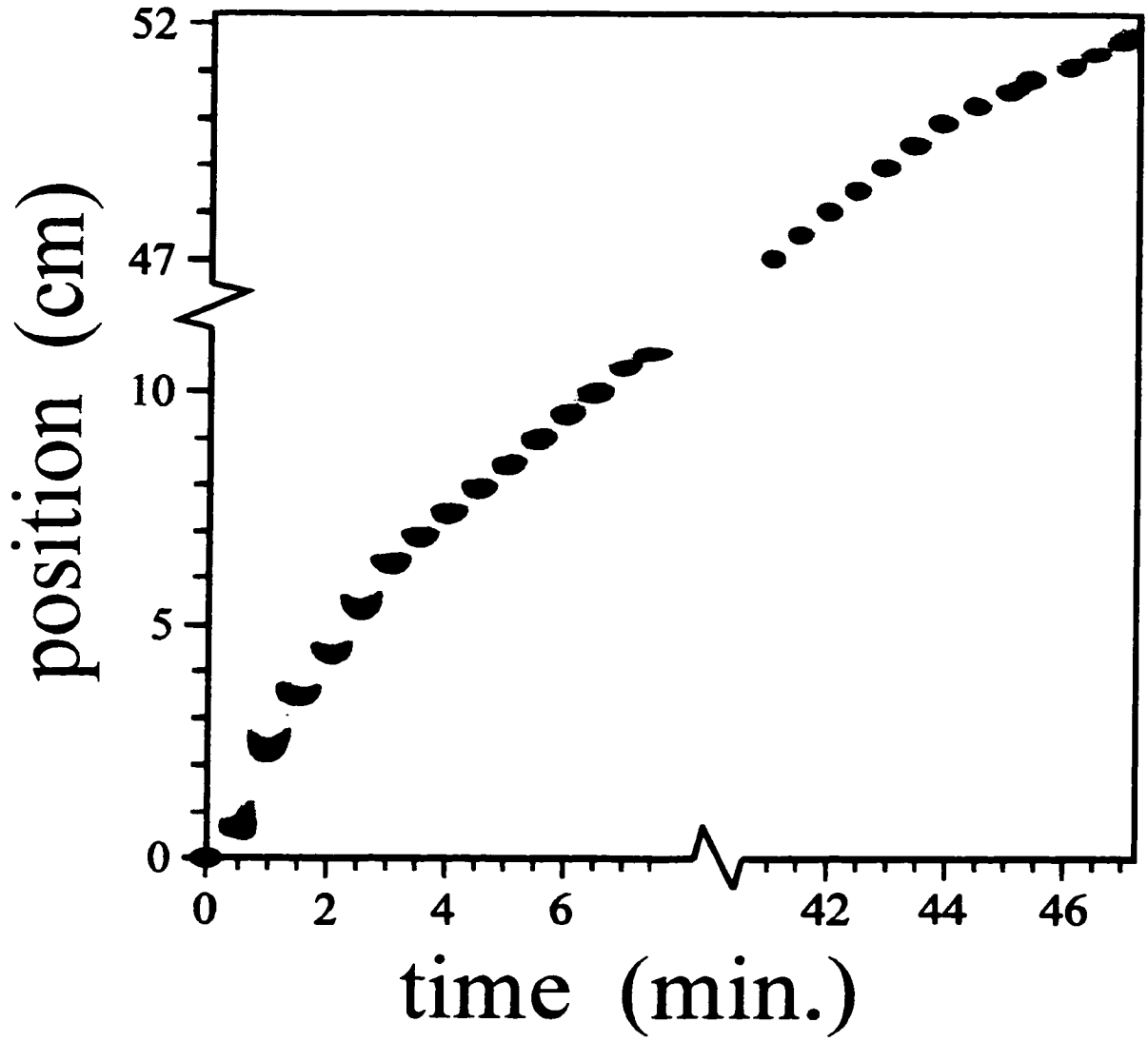


Figure 1 Gel picture showing the position x (in cm) of the band formed by a 21-base DNA fragments after a migration time t (in min). Each lane was loaded every 30 seconds. A voltage of 4000 Volts was applied on a 4% polyacrylamide gel containing 8M urea and 0.5×TBE, while the other conditions are described in the text. We clearly see that the electrophoretic velocity is not uniform.

understanding of band focusing (and antifocusing) by field gradients [13-15]. This figure nicely demonstrates that the electrophoretic conditions are not uniform along the sequencing gel, a factor that has been largely overlooked previously.

4.3.2 Effect of the Voltage

Figure 2 shows the local velocity v(x) vs. position x along the gel for three different voltages. Data points from many experiments are shown (the point are joined in order to avoid confusion). The velocity v(x) was not measured in the middle of the gel because it is essentially uniform in this region; instead, a straight line representing a linear continuation between the two ends of the gel is drawn. The velocity first reaches a maximum at a distance ≈25 mm from the loading wells. It then decreases and stays constant in the middle of the gel. There is also a small decrease at the opposite end of the gel. Clearly, the two velocity gradients are larger for higher voltages. The bands become wider in the region where the velocity increases whereas they becomes narrower in the two regions where the velocity decreases (see Figure 1). Note that the position of the different gradients appears to be somewhat independent of the voltage applied, while their amplitude increases with the voltage.

Figure 3 shows the space-averaged velocity $\langle v \rangle_x$ vs. the applied voltage V as obtained from our data. There is an approximately linear relationship between the two parameters. Given that the velocity of the 21-base-long DNA fragment is linearly dependent on the local field intensity (small DNA molecules do not exhibit non-linear effects like larger ones [2]), this result suggests that the main source of the velocity gradients is the existence of local field gradients. For instance, if we assume that the mobility is constant because there is no other gradient (such as temperature gradient,

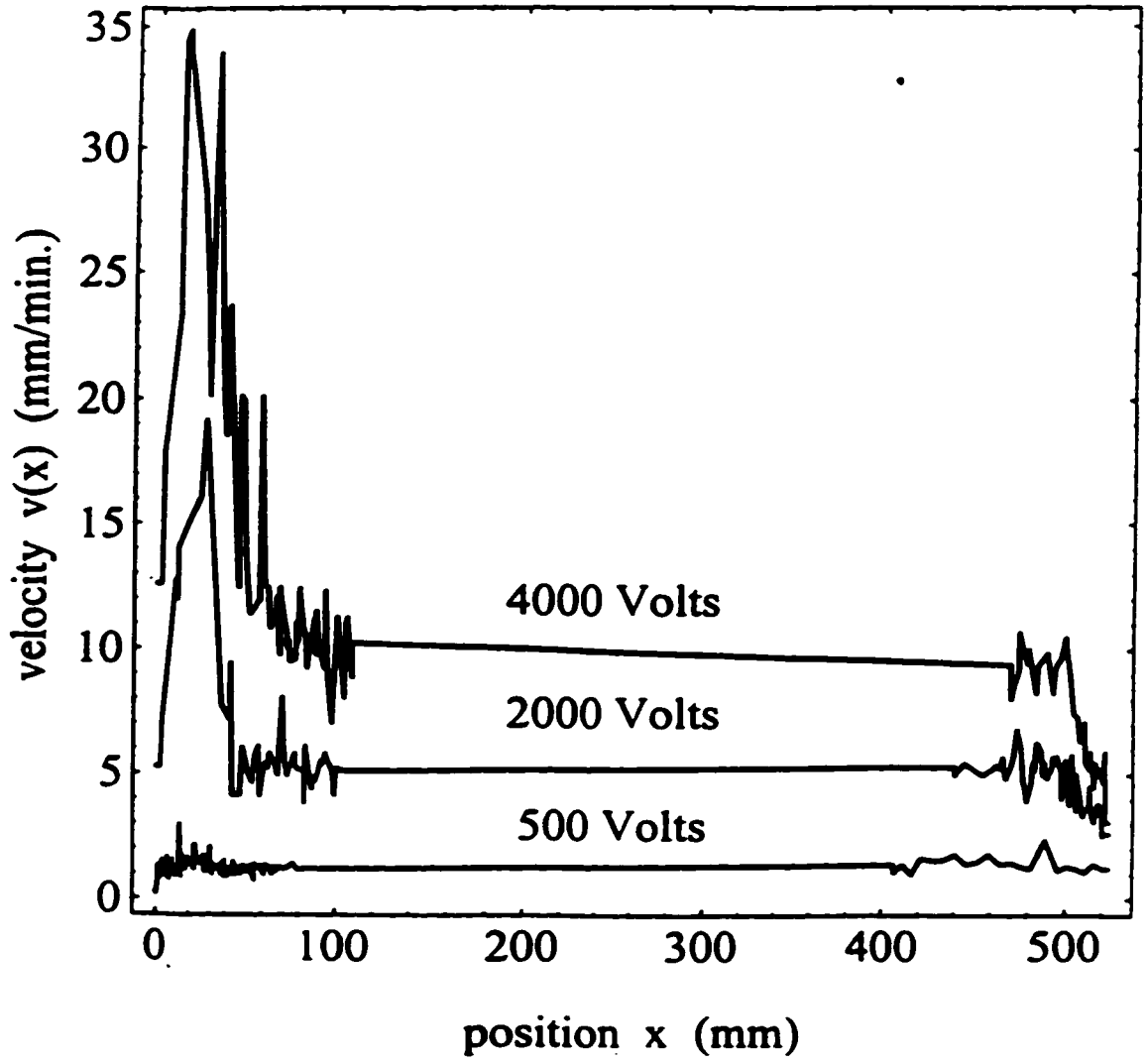


Figure 2 Local velocity v(x) vs. position x along the gel for three different voltages, as shown. The data from different experiments have been superimposed. The points have been joined to avoid confusion between different voltages. There is no data point for the middle of the gel; instead, we have drawn straight lines joining the data on both sides.

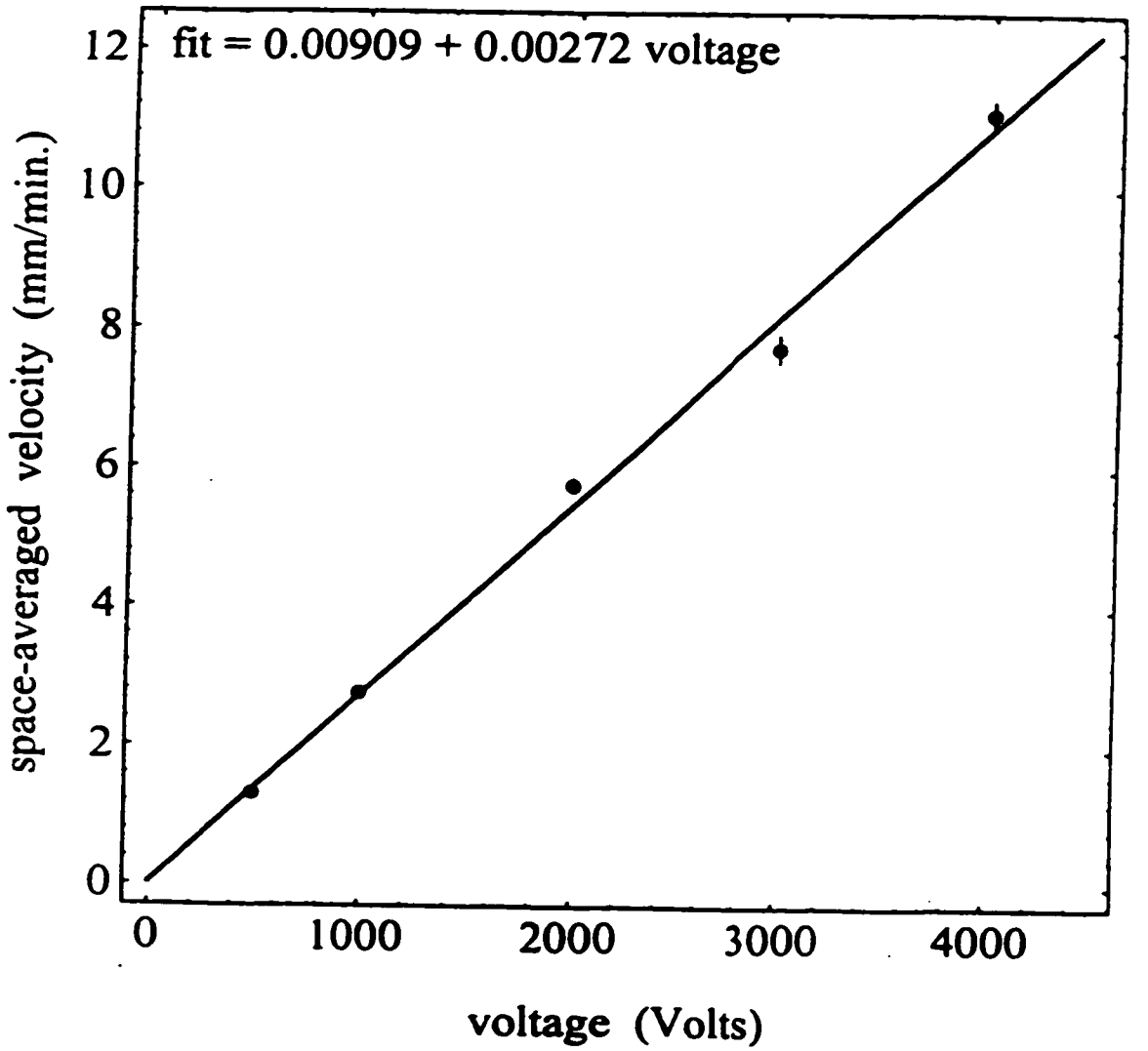


Figure 3 **Space-averaged velocity $\langle v \rangle_x$ vs. applied voltage V.**

viscosity gradient, etc.) in the system, substitution of Eqs. (3) and (4) in Eq.(5) gives directly $\langle v\rangle_x=\mu V/L$, showing the linear dependence observed here. The slope of the linear fit gives $\mu/L=0.00272$ mm/(V min), as shown, from which the electrophoretic mobility is found to be $\mu=2.37\times10^{-4}$ cm^2/(Vs). This is consistent with the free-solution mobility $\mu_0=3.8\times10^{-4}$ cm^2/(Vs) (as found using the standard method which consists in measuring the mobility of a small molecule in decreasing gel concentrations and extrapolating the results to zero concentration [16].

From our data, we can also plot the velocity in the middle of the gel, v(x=L/2), vs. the space-averaged velocity $\langle v\rangle_x$ (Fig. 4). In the gradient-free situation, one would obviously obtain $v(x=L/2)=\langle v\rangle_x$ (dashed line in Fig. 4). All the points fall below this line. The best fit to the data (solid line) gives a slope of 0.9. In other words, the velocity in the middle of the gel is approximately 10% lower than the spaced-averaged velocity. Assuming that the temperature is uniform, we can conclude that the field E(x=L/2) is 10% lower than $\langle E\rangle_x=V/L$. Consequently, the total migration (or elution) time τ is larger than it would be for a uniform field.

4.3.3 Effect of the Urea Concentration U and of the Gel Concentration %T

The viscosity of the buffer varies substantially if we change the urea concentration U. Therefore, we studied how the ratio v(max)/v(x=L/2) (the maximum velocity observed a few cm ahead of the loading zone divided by the velocity in the middle of the gel) changes with U in 4% gels at 4000 V (Figure 5). Remarkably, the velocity gradient essentially disappears when U=0 M and is essentially inexistant when U≤2M. An 8 M urea denaturing solution is roughly (15±2)% more viscous than water (results not shown). This larger viscosity generates a difference in transference

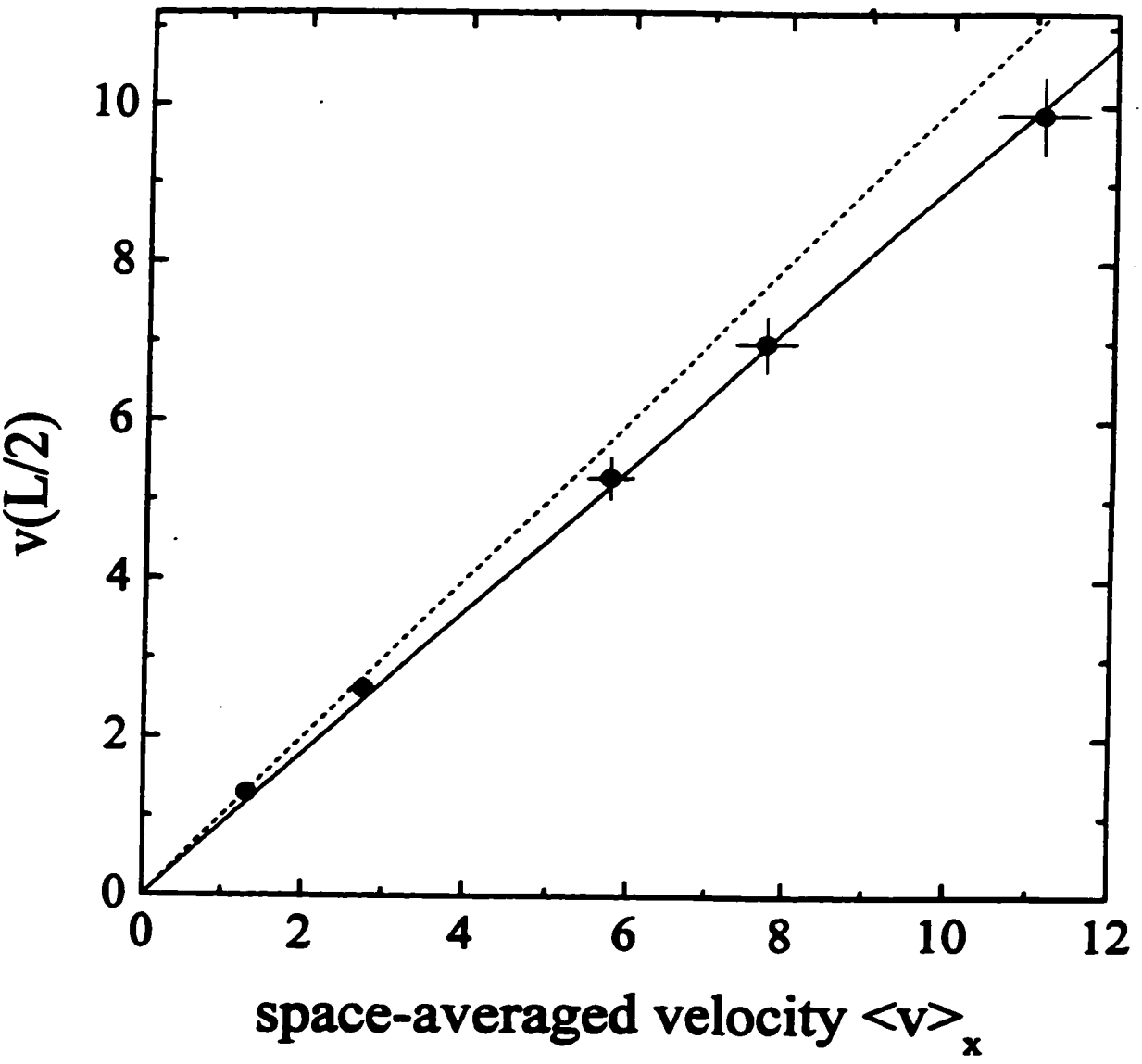


Figure 4 Velocity in the middle of the gel v(x=L/2) vs. space-averaged velocity $\langle v \rangle_x$. The solid line represents a fit of the data and the dashed line shows the situation for a gel without gradients. The data of experiments performed at 500, 1000, 2000, 3000 and 4000 Volts are shown from left to right.

numbers at the gel interfaces, causing most of the observed electric field gradient. In fact, adding 8 M urea to the top tank buffer eliminated the velocity gradient (result not shown), a clear demonstration of the role of the interface in this phenomenon. We also observed that the position where the velocity reaches its maximum value was deeper inside the gel for lower urea concentrations (results not shown). Furthermore, the inset of Fig. 5, which shows v(max)/v(x=L/2) vs. %T (percentage of acrylamide in the gels) for V=4000 V and U=8 M, clearly indicates that the gel concentration does not contribute significantly to the formation of gel edge velocity gradients over this range of concentrations.

4.3.4 Other Buffers

The presence of a velocity gradient was also tested in 4% acrylamide gels containing 8 M urea and buffers other that 0.5×TBE. Experiments were done using 1×TBE (89 mM Tris, 89 mM boric acid, 2 mM EDTA), 0.2×TAE (0.16 mM Tris, 0.16 mM glacial acetic acid, 0.003 mM EDTA, pH 8), Tris-HCl (100 mM Tris, adjusted pH 8.0 using Hcl) and 0.5×TB (TBE without EDTA). The velocity gradient was observed with all of these buffers (not shown). Interestingly enough, we observed a monotonic gradient (the velocity is a maximum at x=0) for the Tris-HCl (which is a two-ion system if we neglect the H^+ and OH^- ions). The other buffers showed a velocity profile very similar to the one plotted in Fig. 2 for 0.5×TBE (results not shown). The v(max)/v(x=L/2) ratio was equal to 6.5 and 17 for Tris-HCl and 0.2×TAE, respectively, showing that ions with different transference numbers will produce gradients of different amplitudes. Note that the experiments with Tris-HCl and TAE had to be preformed at lower temperature (i.e., 40°C as opposed to 55°C) in order

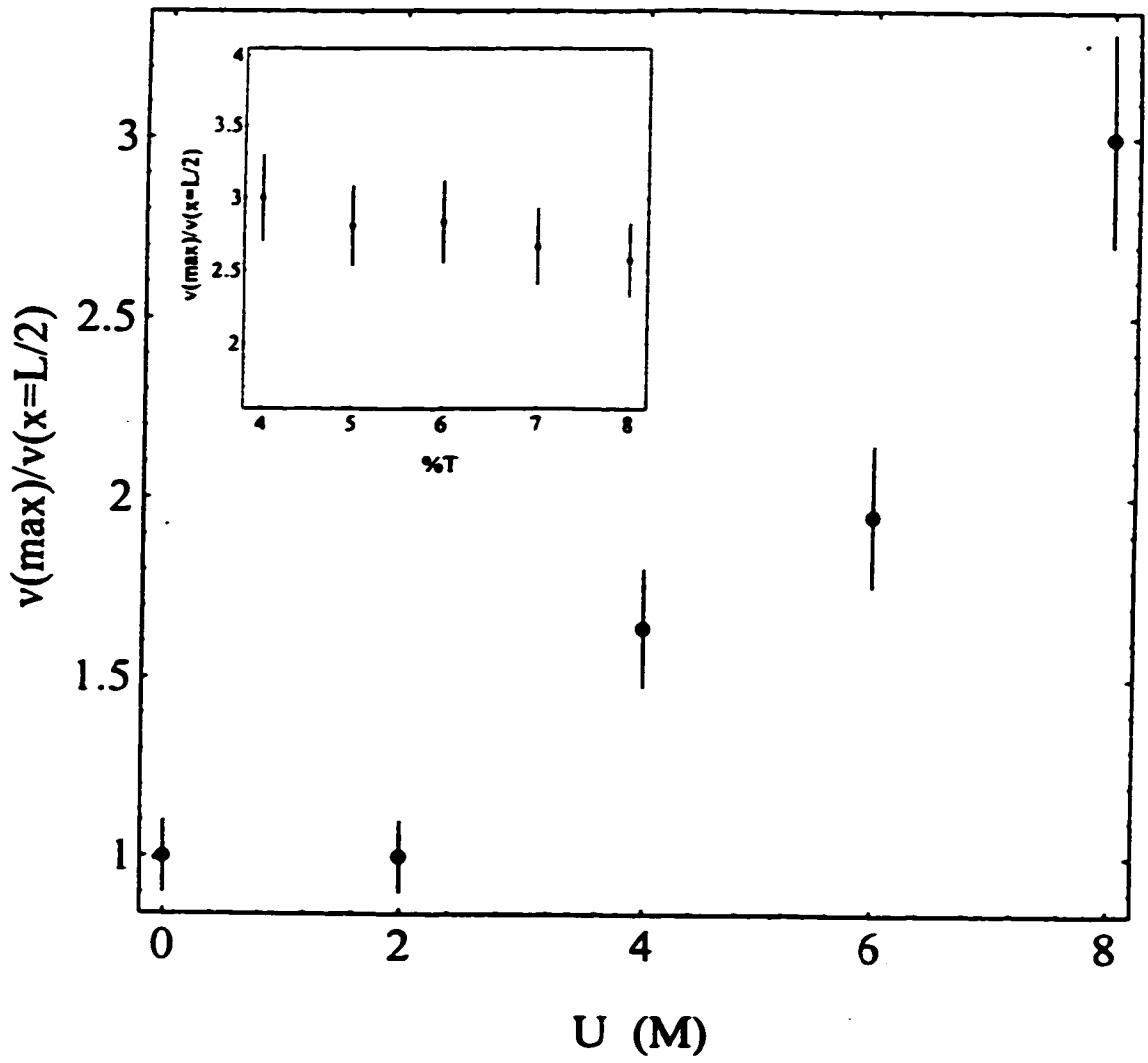


Figure 5 The ratio v(max)/v(x=L/2) vs. U, the quantity of urea in the 4%T gel, v(max) being the maximum velocity in the gel and v(x=L/2) the velocity in the middle of the gel. The inset: The same ratio v(max)/v(x=L/2) vs. %T, the concentration of polyacrylamide in the gel. Those values were obtain for a voltage of 4000 Volts.

to preserve gel integrity. The gradient observed with the 0.5×TB buffer was similar to the one performed with 0.5×TBE, while the gradient of the 1×TBE gel had lower amplitude (result not shown).

4.4 Discussion

Our results indicate that the velocity gradients near the edges of denaturing polyacrylamide gels are important at high voltages and are caused by the difference in viscosity (transference numbers) between the gel and the buffer. This difference in viscosity is due to the urea present in the gel, and not to the gel fibers (inset of Fig. 5). We have observed gradients at both ends of the gel, as opposed to Figeys et al. [2] where the small negative gradient at the bottom end of capillaries was not present. This gradient is of little importance in practice. Apparently, this gradient migrated very slowly in our system (dT/dc is small in Eq. 2) and did not have the time to exit the gel unlike in [2]. Alternatively, this weak negative gradient could be explained in part by the fact that the temperature in the buffer reservoir was 45°C, i.e. 10°C lower than in the gel. On the other hand the difference in temperature between the buffer in the top buffer reservoir and the gel does not explain the gradient observed at the top of the gel. In fact we performed an experiment where the buffer in the top reservoir was heated to the same temperature as the gel (55°C). The velocity profile obtained was almost identical to that shown in Figure 1 (results not shown).

The large velocity peak observed near the loading wells does have an impact. Its existence requires that the concepts of electrophoretic mobility and mean field intensity be redefined, and directly affects band broadening (Figure 1). In fact, the final width of a band that must cross a region

with gradients is always larger than if it had to cross a uniform region [17].

We have observed a linear dependence of the space-averaged velocity $\langle v \rangle_x$ upon voltage V. This is fully consistent with the hypothesis that the velocity gradients are mostly due to electric field gradients in the system. Temperature gradients alone cannot explain results showing a velocity that is locally three times larger than elsewhere (see Figure 2) since the viscosity of the aqueous buffer decreases typically by (only) ≈2%/°C. The ion-depletion mechanism studied by Spencer (and described in sections 4.1.1-4.1.3) is likely related to the effects observed in our investigation.

These gradients lead to three significant practical problems. First, they make it difficult to obtain useful and reproducible velocity and diffusion data for system optimization. For example, the gradients are history-, voltage- and system-dependent. Second, the presence of large velocity (and field) gradients a few cm in front of the loading wells sometimes irreversibly affects gel integrity (we have observed the formation of bubbles and the destruction of the gel in this region if very large fields are used over extended periods of time). In order to test if microscopic (i.e., invisible) gel destruction was responsible for the velocity gradients, we performed the following critical experiment. First, we prepared a gel with a regular prerun at 4000 V for 2 h. We then inverted the field direction and carried out an inverted prerun for an additional 3 h. The results (absence of velocity gradients) were identical to those obtained in absence of the initial regular prerun [1]. Therefore, the long regular prerun does not lead to detrimental microscopic gel defects. Moreover, this experiment shows that the velocity gradient are not due to urea diffusion outside the gel because a longer total prerun time leads to no gradients. Finally, note that the gradients can severely affect the efficiency and resolution of the separation because the gradients broaden the bands while the

local high field intensities may increase the negative impact of biased reptation [18-20].

In some cases, the difference between the velocity v(x=L/2) in the middle of the gel and the time-averaged velocity defined by $\langle v \rangle_t = L/\tau$ can be large enough to make it difficult to test competing electrophoresis theories. Velocities can be defined unambiguously only if multiple loadings are used, or else if the gradients are eliminated. In the next chapter, we will show that it was extremely difficult to quantify and characterize trapping gel electrophoresis [21] of DNA-streptavidin hybrid molecules because the gradients strongly interfere with the phenomenon under study.

In order to minimize the effects of the gradients, one can match the viscosity in the buffer tank with the viscosity in the gel. As mentioned above, this can be achieved by adding urea to the running buffer. This actually support Spencer idea that the transference numbers are key to understanding such gradients. However, this approach is expensive for slab gels and leads to gel loading problems, i.e., the density of normal DNA sequencing samples is not high enough to allow them to sink into wells. These two problems are less important with capillary electrophoresis. Alternatively, one could use a buffer where the retardation factor is similar for the positive and negative ions. Another solution would be to use a stronger denaturing agent, which would lead to a smaller increase in gel viscosity relative to the buffer in the tanks because it would achieve the same DNA denaturing potential at a lower concentration. If strong denaturation is not essential, one could also use a lower urea concentration in the gel in order to minimize the gradient at the loading end of the gel. However, the most practical and efficient way of minimizing the high electric field gradient at the loading end of the gel is still, in our opinion, to perform an inverted prerun because it allows the efficient elimination of this gradient without modifying the chemistry of the gel [1,3].

In this case, the system remains out-of-equilibrium during the migration but the system evolves slowly enough that DNA retains a constant electrophoretic velocity for up to 1-2h.

In conclusion, we have shown that the gel edge electric field gradients exist because of a difference in ion transference numbers caused by the difference in viscosity between the gel and tank buffers. The gel edge gradient found at the loading end of the gel leads to larger band widths and requires that the concept of electrophoretic mobility and mean field intensity be redefined.

4.5 References

[1] Mayer, P., Slater, G. W., Drouin, G., *Appl. Theor. Electroph.*, **3**, 147, 1993.

[2] Figeys, D., Renborg, A., Dovichi, N. J., *Electrophoresis*, **15**, 1512, 1994.

[3] Slater, G. W., Mayer, P., Drouin, G., *Methods in Enzymology*, **270**, 272, 1996.

[4] Swerdlow, H., Dew-Jager, K. E., Brady, K., Grey, R., Dovichi, N. J., Gesteland, R., *Electrophoresis*, **13**, 475, 1992

[5] Spencer, M., *Electrophoresis*, **4**, 36, 1983.

[6] Spencer, M., *Electrophoresis*, **4**, 41, 1983.

[7] Spencer, M., Kirk, J. M., *Electrophoresis*, **4**, 46, 1983.

[8] Houri , A., Sudaka, P., Starita-Geribaldi, M., *Appl. Theor. Electroph.*, **1**, 323, 1991.

[9] Michov, B. M., *Electrophoresis*, **9**, 105, 1988.

[10] Michov, B. M., *Electrophoresis*, **7**, 150, 1986.

[11] Kohlrausch, F., *Ann. Physik. u. Chem.*, **62**, 209, 1867.

[12] von Laue, M., *Z. anorg. u. allgem. Chem.*, **93**, 329, 1915.

[13] Slater, G. W., Noolandi, J., *Electrophoresis*, **9**, 643, 1988

[14] Slater, G. W., Mayer, P., Drouin, G., *Analusis*, **21**, M25, 1993

[15] States, J. C., Patel, L. R., Li, Q., *Biotechniques*, **10**, 46, 1991.

[16] Rousseau, J., Drouin, G., Slater, G., *Phys. Rev. Lett.*, **79**, 1945, 1997.

[17] Slater, G.W., Mayer, P., *Electrophoresis*, **16**, 771, 1995.

[18] Lumpkin, O., Déjardin, P., Zimm, B. H., *Biopolymers*, **24**, 1573, 1985.

[19] Slater, G. W. and Noolandi, J., in: Lee, L.-H. (Ed.), *New Trends in Physics and Physical Chemistry of Polymers*, Plenum Press, New York and london. pp.547, 1989.

[20] Doi, M., Kobayashi, T., Makino, Y., Ogawa, M., Slater, G. W., Noolandi, J., *Phys. Rev. Lett.*, **61**, 1893, 1988.

[21] Ulanovsky, L., Drouin, G., Gilbert, W., *Nature*, **343**, 190, 1990.

Chapter 5

On using DNA-Trapping Electrophoresis to Increase the Resolution of DNA Sequencing Gels*

As discussed in Section 1.2, DNA-trapping electrophoresis appears to be a promising method to increase the resolution of DNA sequencing ladders. In this chapter, we show that this method requires inverted gel preruns to obtain optimal resolution, and we present a method to calculate the velocity of end-labeled DNA molecules in denaturing polyacrylamide gels. We use this method to elucidate the trapping mechanism and to characterize the electric field dependence of the critical size beyond which end-labeled DNA molecules are strongly trapped. Our results show that DNA-trapping electrophoresis cannot be used to increase the resolution of DNA sequencing ladders because, in the strong trapping regime, the band widths increase more than the gain in interband distances. On the other hand, it might be possible to use DNA-trapping electrophoresis to develop isoelectric focusing-like techniques using monotonic electric field gradients.

---

*The results of this Chapter were published in: *Macromolecules*, **31**, 6499, 1998.

5.1 Introduction

DNA sequencing in polyacrylamide slab gels can typically resolve from 400 to 600 bases at single base resolution. This limit is due to the migration behavior of single-stranded DNA in polyacrylamide gels [1,2]. Optimized systems using long slab gels can resolve up to about 1000 bases [3]. However, reading accuracy decreases for molecular sizes larger than 600 bases and the thousand base limit is not always reachable [4].

Because the current DNA sequencing technologies are widely regarded as one of the bottlenecks of the human genome project, new electrophoresis-based and other sequencing technologies have been suggested [5-9]. One such promising idea was suggested by Ulanovsky, Drouin and Gilbert in a 1990 Nature paper [10]. These authors proposed to attach a large neutral label (such as the protein streptavidin, MW≈52000) to one end of the DNA molecules to be electrophoresed in a polyacrylamide gel. They noted that short streptavidin-DNA complexes (S-DNA) are only slightly retarded, while the velocity of S-DNA molecules larger than a certain critical size M^* (which depends on the applied electric field E) is dramatically reduced thus increasing the mean interband spacing [10]. Although the proposed trapping electrophoresis (TE) process attracted a lot of attention in the DNA sequencing community, no new experimental results have appeared since. Theoretical studies have been published [11-18], but no clear conclusions regarding the usefulness of TE have been reached. In this chapter, we present the first in-depth study of TE and we establish its limitations for DNA separations. In particular, we show that TE's usefulness is severely restricted by current (slab) electrophoresis instruments.

5.2 Theory

Qualitatively, TE can be explained easily. We first note that when a S-DNA migrates in a random gel, its velocity is reduced due to the extra friction generated by the streptavidin label. In a gel, one has a wide range of pore sizes and structural defects. Under certain conditions (defined below), S-DNA molecules may stop migrating, at least temporarily, because they are trapped between gel fibers. Two trapping mechanisms have been proposed: i) Ulanovsky [15] proposed the protein might get trapped in a fork-like protrusion of the gel (Figure 1A); we call this the "fork" trapping mechanism ; ii) the Slater-Drouin [13] and Viovy [16] groups, on the other hand, considered that trapping occurs when the flexible DNA chain enters a close loop that is too narrow for the label (Figure 1B); we call this the "dead end" trapping mechanism. In either case, the S-DNA molecule clearly needs to move backward over a certain minimum distance L_0 in order to get out of the trap and resume migration in the field direction. Since the backward motion is due to Brownian motion, the detrapping time will increase (exponentially) with both the electric force (field intensity times charge of DNA) and the required distance L_0. Therefore, detrapping should be substantially easier for fork trapping (where L_0 is comparable to the mean pore size of the gel) than for dead end trapping (where L_0 is a function of the contour length of the DNA molecule). One of the goals of this chapter is to establish which trapping mode dominates the TE of S-DNA.

TE can be characterized by two time scales: the detrapping time τ_d and the time τ_0 it takes to get into a new trap. If V_0 is the velocity of the S-DNA complex between traps, its net velocity V in the presence of trapping is simply given by $V=V_0\tau_0/(\tau_0+\tau_d)$. When $\tau_d \ll \tau_0$, trapping is ineffective and $V \approx V_0$: this defines the weak trapping regime. When $\tau_d \gg \tau_0$, however, trapping dominates the

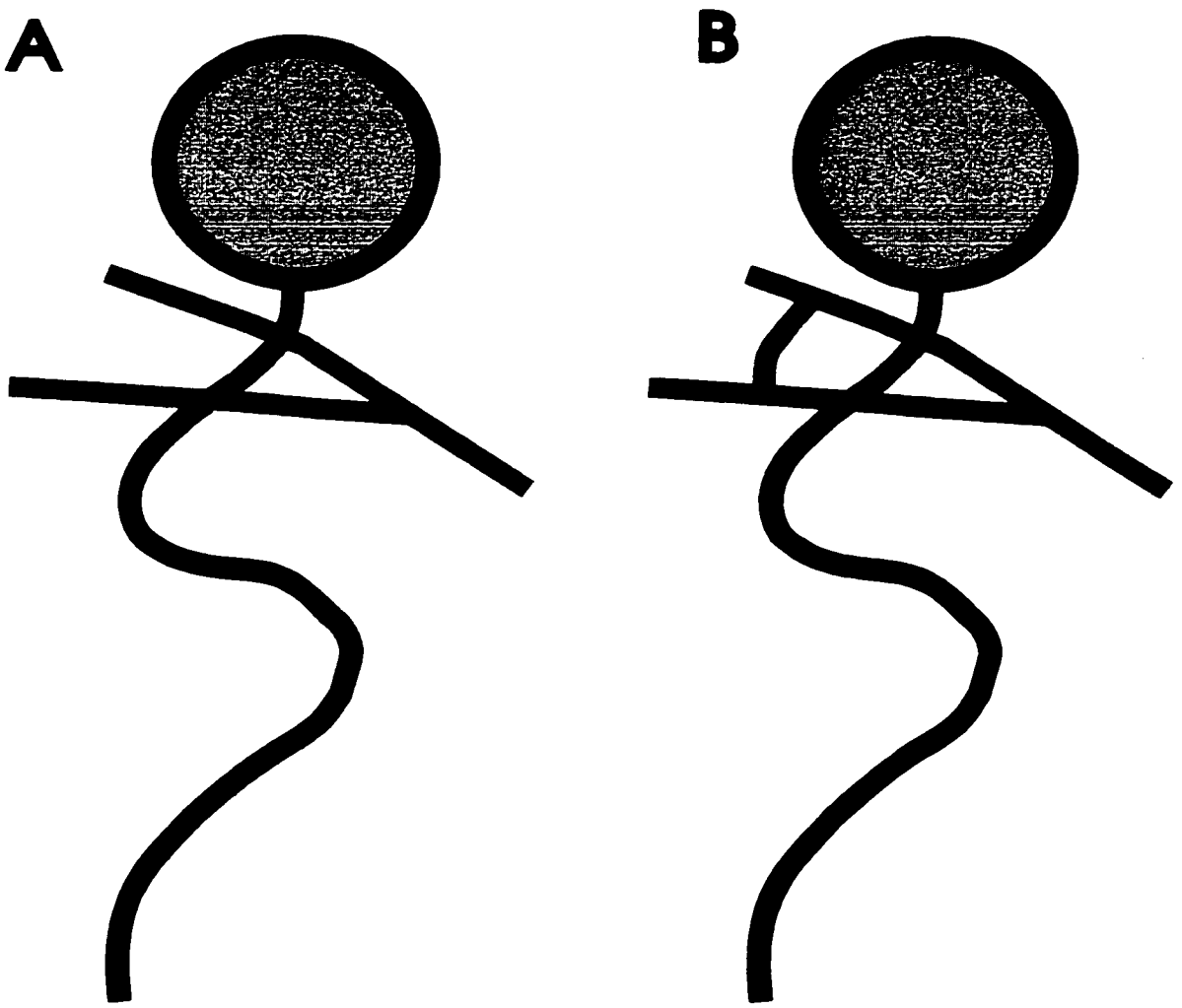


Figure 1: Schematic representation of (A) the fork trapping mechanism proposed by Ulanovsky [15] and (B) the dead end trapping mechanism proposed by the Slater-Drouin and Viovy groups.

dynamics and $V=V_0\times[\tau_0/\tau_d]\ll V_0$. The critical molecular size M^* that marks the transition from the weak to the strong trapping regime is thus the solution of the relation $\tau_0(M^*)=\tau_d(M^*)$. Another goal of this chapter is to find the field dependence of the critical size M*.

Slater and Villeneuve [11] used a modified version of the biased reptation model (BRM) [19-21] to carry out a computer simulation study of TE with dead end trapping. Their results showed a weak trapping regime where $V(M)\propto E/(M+\alpha)$, followed by a strong trapping regime for which $V(M)\propto\exp(-\beta M^3E^2)$, where β is a constant. Therefore, this numerical model predicted that $M^*\sim E^{-2/3}$. We later developed an analytical model of TE [13] based on the BRM (again considering dead end trapping) to understand the behavior of both the net velocity V and the diffusion coefficient D. The problem could be described in terms of the dimensionless parameter $\Lambda=(2M/M_a)^{3/2}\epsilon$, where M_a is the amount of DNA contained in a mean gel pore of size a, and $\epsilon=M_aqEa/(2k_BT)$ is the scaled field intensity (here, k_B is the Boltzmann constant, T is the temperature, and q is the effective charge of one nucleotide). Note that $N=M/M_a$ thus defines the number of pores occupied by the DNA molecule. Trapping was predicted to be weak for $\Lambda<1$ and strong for $\Lambda>5$. This model also gave $M^*\sim E^{-2/3}$. The predicted velocity was consistent with the numerical results of Slater and Villeneuve [11], while the model further predicted that D should increase exponentially with molecular size in the strong trapping regime (a very negative aspect of TE, of course). Using the same analytical approach, it is possible to show that fork trapping (where $L_0\approx a$) would instead lead to $M^*\propto E^{-2}$. An intuitive explanation of the field dependence of M^* for both the fork and the dead end trapping modes is given in section 5.4.5, while a more complete derivation is included in the Appendix. Défontaines and Viovy [17] also developed an analytical reptation model of dead end TE where they

allowed for the elongation of the reptation tube. They predicted numerous regimes for weak and strong tube elongation. In the limit of constant tube length (this approximation will be examined in the discussion section), they also predicted that $M^* \sim E^{-2/3}$.

5.3 Materials and Methods

5.3.1 Polyacrylamide Gels

Eight percent polyacrylamide gels with a 19:1 acrylamide (BDH) to bisacrylamide (BDH) ratio were used. They contained 5M urea (BDH) and a solution of 0.5×TBE [44.5 mM Tris (BDH), 44.5 mM boric acid (BDH), 1 mM EDTA (BDH)] was used as the running buffer. The 0.1 mm to 0.4 mm thick gels were prepared on a Pharmacia-LKB Macrophor apparatus. Polymerization was initiated with the addition of 0.1%w/v ammonium persulfate (BDH) and 0.1%v/v TEMED (BDH), and the gels were left to polymerize at room temperature for at least 12 hours.

5.3.2 DNA Samples

Radioactively labeled S-DNA 50 basepair ladders were prepared by mixing 1 μL of 50 basepair DNA ladder (1 μg/μL; Pharmacia), 3 μL of 5×sequencing buffer (200 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$, 250 mM NaCl), 4.5 μL of biotin-14-dCTP (33 pmol/μL; Life Technologies), 3 μL of [α-^{32}P] TTP (3.3 μM; Amersham), 3 μL distilled water and 0.5 μL Klenow (7.4 units/μL; Life Technologies), and this reaction was incubated for 15 minutes at room temperature. This reaction produced radioactive (α-^{32}P) double-stranded DNA fragments with a biotin molecule at both ends. We then added 5 μL 0.5M EDTA, 80 μL TE (10 mM Tris, pH 8, 1mM EDTA) and 1 μL of

0.1 mM streptavidin (diluted in TE; Boehringer Mannheim), and we incubated this reaction for 15 minutes at room temperature. This reaction allowed a streptavidin molecule to bind to each of the biotin molecules that were previously added at both ends of the DNA fragments. We finally added 100μL of stop solution (95% formamide, 1% w/v xylene cyanol, 1% w/v bromophenol blue and 10 mM EDTA) and the samples were denatured at 65°C for 5 minutes. This step lead to the separation of the two DNA strands and therefore produced single-stranded DNA fragments having a streptavidin molecule at their 5′ end. Note that these denaturing conditions (incubation at 65°C for 5 minutes in 50% v/v formamide) did not break the streptavidin-biotin bonds. These denatured samples were subsequently kept on ice until they were loaded on the gel. DNA sequencing reactions contained 1 pmole of M13mp18 single-stranded DNA, 2 pmoles of a 5'-end biotinylated primer, and were performed in a volume of 10 μL using T7 Sequencing Kits (Pharmacia). The reactions were stopped by adding 1 μL of 0.5M EDTA, and 1 μL of 0.1 mM of streptavidin (diluted in TE; Boehringer Mannheim) was added. After a 15 minute incubation period at room temperature, 12 μL of stop solution were added, and the samples were denatured as described above. The DNA fragments produced by these reactions were also radiolabeled (α-^{35}S) single-stranded DNA fragments having a streptavidin molecule at their 5′-end.

5.3.3 Electrophoresis Conditions

The running temperature (40°C) was controlled through the thermostatic plate of the Pharmacia-LKB apparatus. Experiments with voltage ranging from 500 to 4000 volts were performed on 18.5 cm-long gels using a Fisher Biotech FB600 power supply. The gels were prerun

at the selected field E (in V/cm) for at least 150/E hours (e.g., for 3 hours at 50 V/cm) as we described previously [22,23]. Unless otherwise specified, the direction of the electric field during the prerun was opposite to that of the electrophoresis. After that period of time, the electric current had fallen to about 40% of its initial value, and remained constant during the whole electrophoresis (less than 2% variation). As discussed below, this procedure is necessary to obtain optimal results.

5.4 Results

5.4.1 Electric Field Gradients

The original report on TE [10], as well as subsequent theoretical treatments [11-14,17], clearly indicated that the gel electrophoretic mobility μ of large S-DNA molecules is strongly size- and field-dependent. The former is what makes TE an attractive method to improve separations; the latter, however, means that small field gradients along the direction of migration may have devastating effects. Unfortunately, it is known that the presence of urea can lead to the formation of strong field gradients near the loading wells. This is due to the depletion of ions in this region of the gel [22-24] (see Chapter 4).

When the prerun is performed with the normal polarity (i.e., the polarity used for the separation), the field gradient forms over a period of a few hours and remains constant afterwards. In the previous chapter, we reported that the field actually reached about 300% of its mean value about 3-5 cm in front of the wells with the buffer used here [24]. It is obvious that if μ decreases exponentially with E in the strong trapping regime, most S-DNA molecules will accumulate where the local electric field E is maximum. We believe that the fact that the impact of the urea-generated

field gradient was not considered explains why no further experimental article on TE were published after the original publication of Ulanovsky et al. [10].

After we recognized that the field gradients severely limited our empirical understanding of TE, we established an experimental method that would allow optimal results to be obtained. This method includes two main elements: a) the prerun; and b) the loading procedure. We showed previously that a simple way to eliminate the gradient near the loading well is to prerun the gel with the opposite polarity until the current is stable ([22-24] and Chapter 4). The gradient then forms at the other end of the gel. Since it takes more than 2 hours for the gradient to change position once the normal field polarity is applied (to carry out the separation), we have enough time to measure electrophoretic mobilities in spatially uniform electric fields. In order to obtain precise mobility values, we use the multiple loading method [22-24]. Briefly, we load identical samples in ($\geq$3) adjacent wells every Δt minutes. We then measure the distance x_i migrated by a given molecular size for different migration times t_i. The slope of the resulting x(t) vs t diagram gives the local velocity $V(x)=dx(t)/dt$, and indicates where the field is truly uniform. This procedure is absolutely required to obtain quantitative results for TE in denaturing polyacrylamide gels.

5.4.2 Normal vs Inverted Preruns

Figure 2A shows the autoradiogram for a 50 base S-DNA ladder electrophoresed in a (8% polyacrylamide, 5M urea and 0.5×TBE) gel when a normal prerun (field direction is the same during the prerun and the experiment) of 2h was performed at 135 V/cm prior to the experiment. Since consecutive loadings were made every Δt=3 min, the gel picture itself is a x(t) vs. t diagram. The

velocities of the M<400 base fragments are clearly higher at the beginning of the gel but reach a constant value about 5 cm beyond the loading wells (at the top of the picture). The behavior of the 400 base fragment is most interesting: its velocity is constant during the first 2 cm (first 6 minutes), then decreases for the next 2 cm (next 6 minutes), and finally increases again until it reaches a constant value in the middle of the gel. Note also that the width of this band increases very quickly in the region where the field is decreasing (from 18 to 21 minutes). We call this behaviour band anti-focusing. It is worth noting that the width of the band does not increase as quickly when the band is in the region where the field is constant (beyond 21 minutes). Larger molecules ($M \geq 500$ bases) migrate until they reach the point where the local field intensity is high enough that they get completely trapped. Interestingly, the width of the 500 base fragment band decreases (it focuses) in spite of the fact that the electric field increases in the region where it gets trapped. These results make it clear that unless one controls the field gradients, TE cannot be properly characterized, let alone exploited.

Figure 2B shows the same experiment with an inverted prerun. The situation is much simpler. For instance, the velocity of all fragments is essentially uniform throughout the gel. This allows us to observe a remarkable increase in inter-band spacing in the 400 to 800 base range. Inverted preruns generate very small positive field gradients over a few cm near the loading wells and these field gradients will only affect extremely large fragments that will get trapped in this region.

(See next page)

Figure 2: **Autoradiogram of streptavidin-DNA molecules on gels which were subjected to a normal prerun (A) or an inverted prerun (B). Gels were run at 135 V/cm and samples were loaded every 3 minutes.**

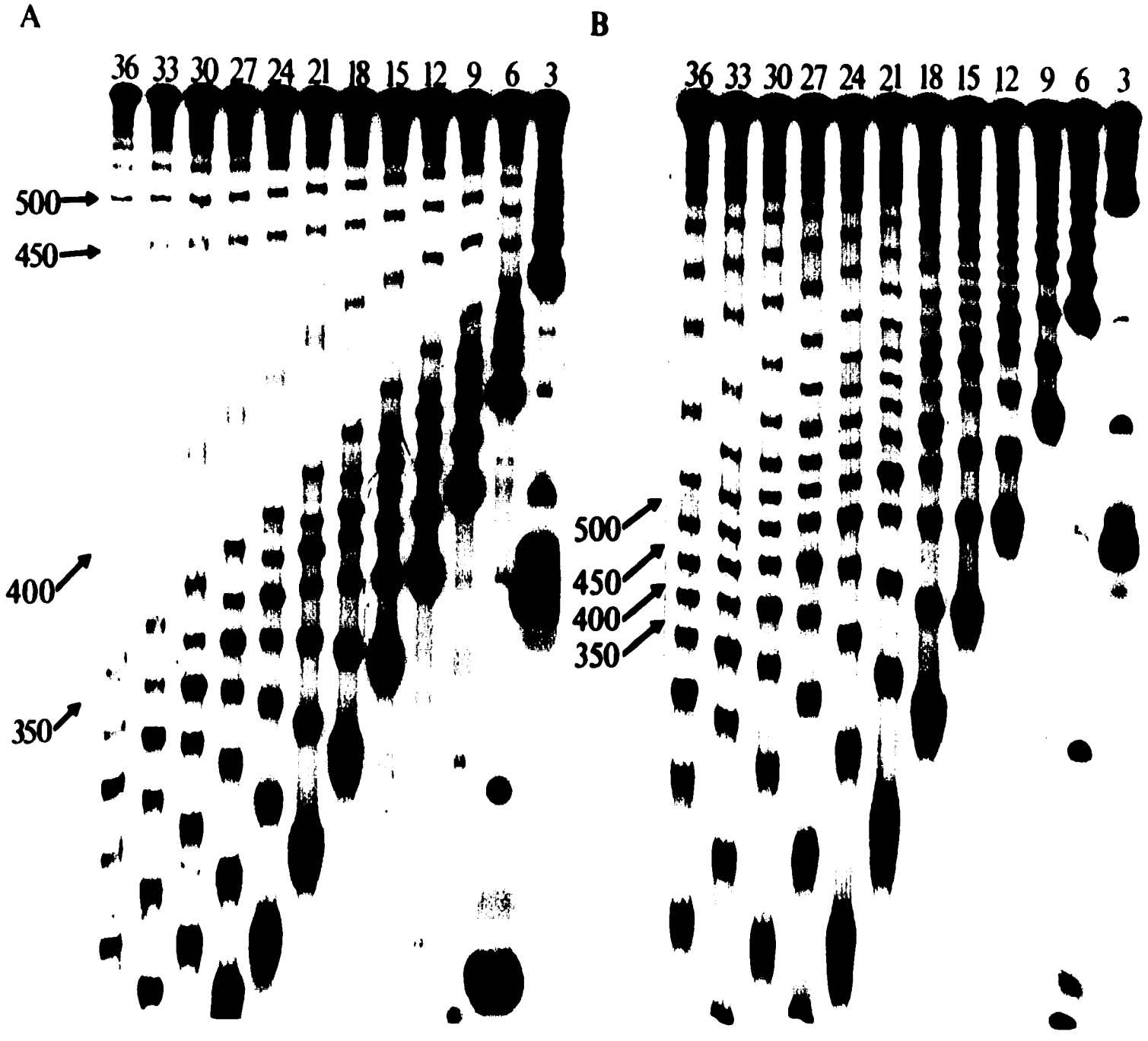

A
36 33 30 27 24 21 18 15 12 9 6 3
500
450
400
350
B
36 33 30 27 24 21 18 15 12 9 6 3
500
450
400
350

Figure 3 shows a schematic representation of the complex situation observed when a normal prerun is used. Figure 3a provides a velocity V vs. electric field E diagram for two different S-DNA sizes (400 and 500 bases). The local value E(x) of the electric field intensity ([24] and Chapter 4) along the direction of migration (x) is given in Fig. 3b. Finally, Fig. 3c presents the density of S-DNA fragments vs. position x (i.e., the band shape). The plots share common axes with one another, and thus provide a complete picture of the process.

We first note (Fig. 3a) that the velocity increases linearly with E up to a certain critical value $E^*(M)$ beyond which it quickly decreases. Note that E^* is inversely proportional to M (see Section 5.2). Figure 3b indicates that for a normal prerun, the local field reaches a maximum a few cm ahead of the loading wells and then converges towards a lower constant value in the middle of the gel. Figure 3c is essentially a convolution of Figs. 3a and 3b for a case where the 400 base band (500 base band) is in the region where the electric field is decreasing (increasing). The field intensity is seen to be stronger at the tail of the 400 base band, while the field is weaker at the front of this band. Figure 3a then tells us that for the 400 base band $V_{front}>V_{tail}$, which means that the band width will increase very quickly because of the local antifocusing field conditions. The situation is quite different for the 500 base band which is located in the region where the field is increasing. We are now in the situation where $V_{front}<V_{tail}$, and the field gradient leads to band focusing (narrowing). Moreover, the diagram makes it clear that it is impossible for a 500 base fragment to cross the region of the electric field gradient because the field exceeds the critical value $E^*(M=500)$ at one point. The latter acts like a fixed point. Finally, we note that both the focusing and the anti-focusing bands must be asymmetric.

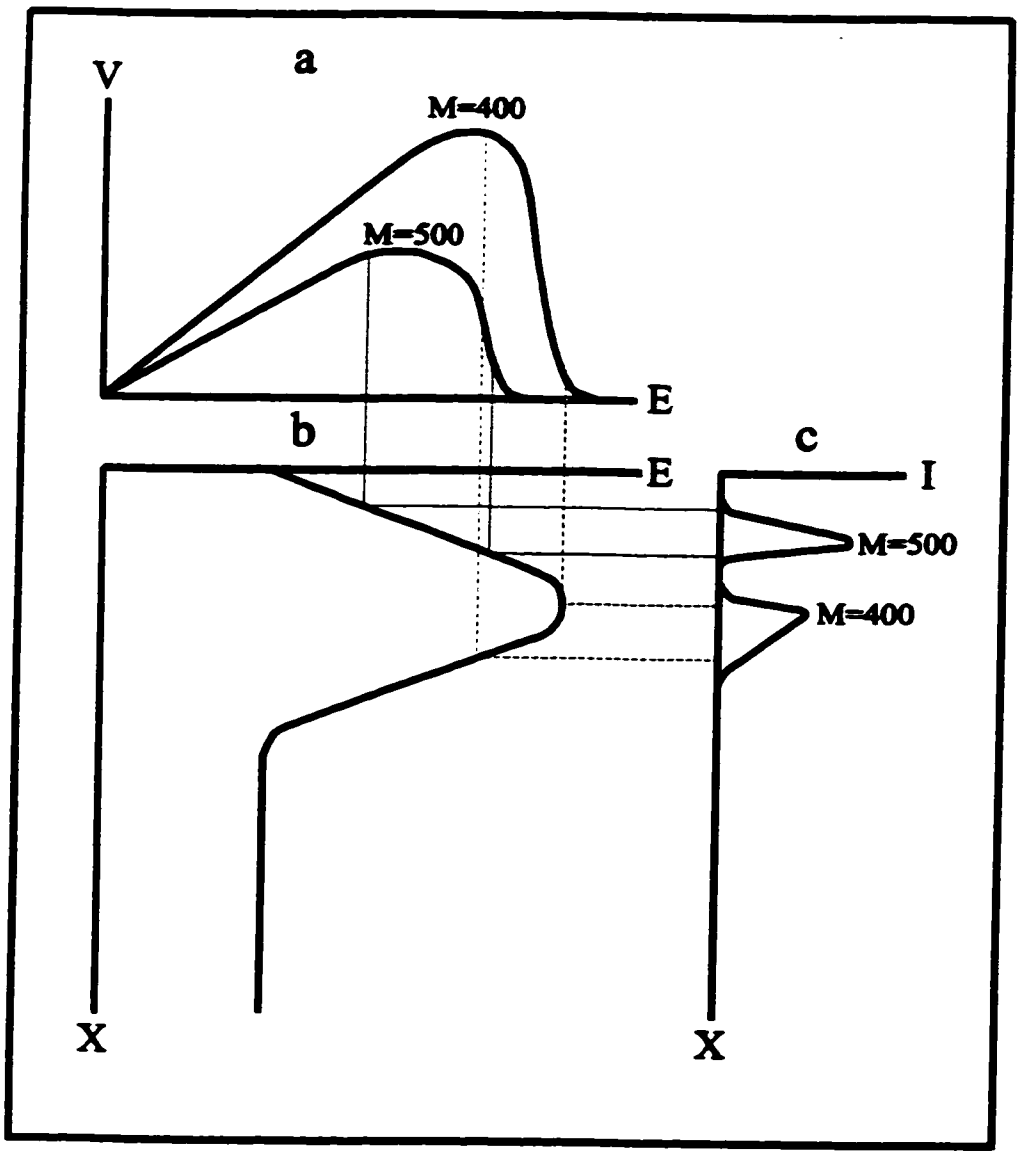


Figure 3: Schematic representation of: (a) the velocity V vs. the electric field (for fragments of 400 and 500 bases); (b) the electric field E vs. position x; and (c) of the density of molecules I vs. position x.

In practice, the normal prerun is to be avoided because the bands either never get passed the field gradient (e.g., our M≥500 base band) or become extremely wide (e.g., our 400 base band). These results do not allow us to draw quantitative conclusions about TE, and certainly do not give any useful separation. This is why inverted preruns are necessary for TE in denaturing polyacrylamide gels.

5.4.3 Determination of M^*

Using inverted preruns and multiple loadings, as described above, we can find the critical size M^* for a given applied field E. Figure 4 shows a plot of the velocity V(M) vs. M for a field of E=108 Volts/cm. Instead of seeing the usual plateau for large DNA sizes [1,2], we observe an inflection point for M≈400 bases, followed by a rapid decrease. The velocity is negligible for M>700 bases. We define M^* as the value of the molecular size at the inflection point, a simple and practical choice. Therefore, M^* is found by locating the maximum on a dV/dM vs. M plot (inset of Figure 4). Looking at Fig. 2B, we see indeed that the bands get closer and closer as the molecular size increases below M=400. Starting at $M=M^*=400$, however, the inter-band spacing starts increasing for a while because of the effect of steric trapping. Finally, note that we see bands up to the loading wells, something that we do not see with normal DNA sequencing.

5.4.4 Weak vs. Strong Trapping

Figure 5 shows a log-log plot of the electrophoretic mobility $\mu=V/E$ vs. the applied field E for DNA molecular sizes ranging from 50 (top curve) to 700 bases (bottom curve). The mobility is

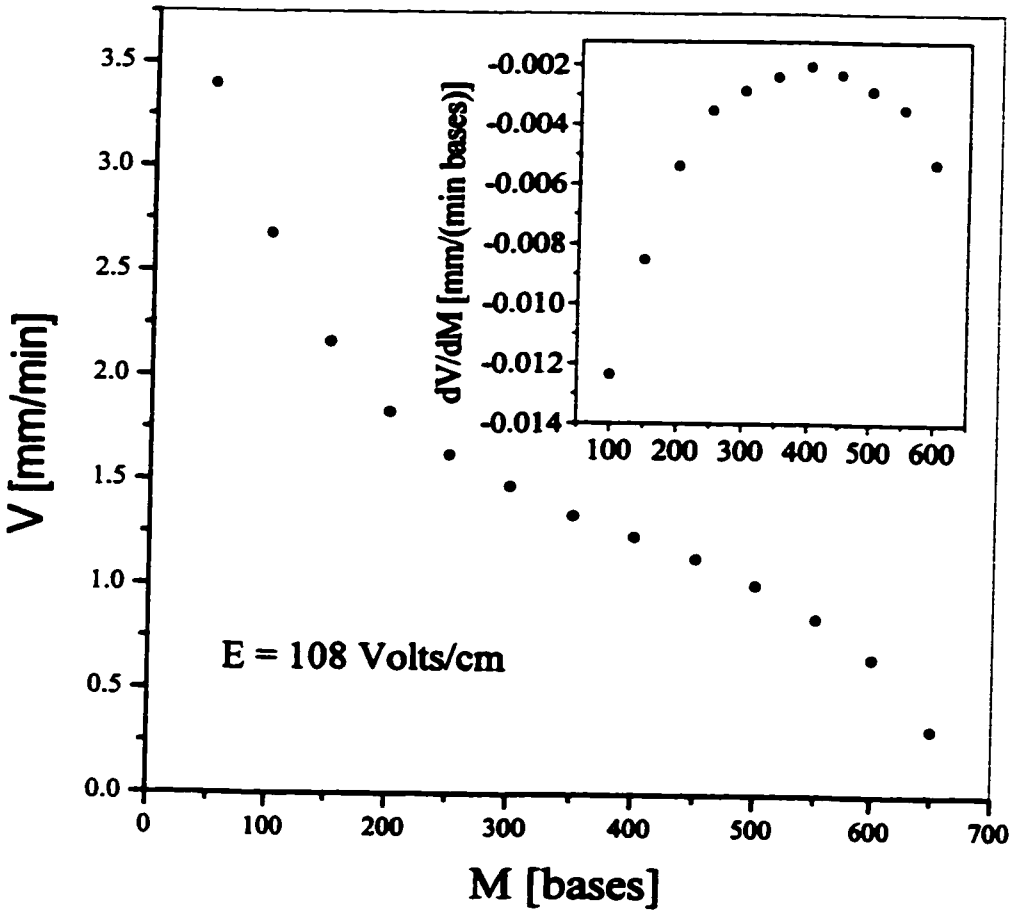


Figure 4: Velocity V vs. molecular size M for an electric field E=108 V/cm. Inset: dV/dM vs. M.

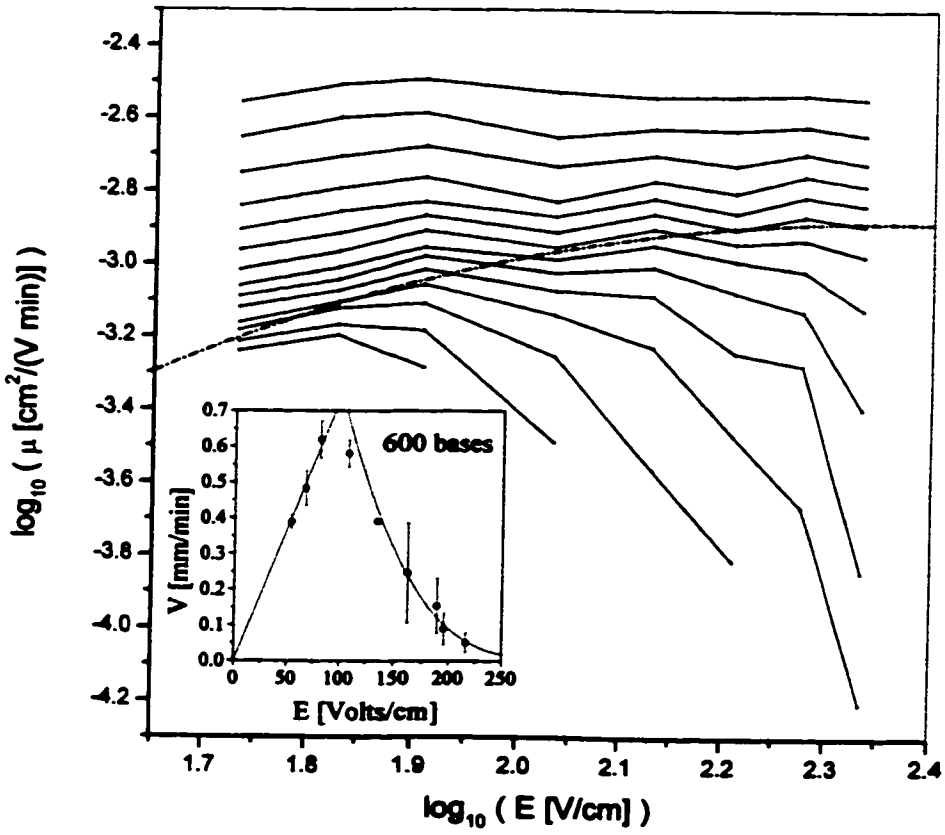


Figure 5: Log-log plot of the mobility μ vs. field intensity E (for molecules ranging from 50 (top) to 700 bases (bottom). The inset shows the velocity V vs. E for a 600 base molecule. The dashed line delimits the weak (above) and the strong (below) trapping regimes.

approximately constant (hence $V \propto E$) for $M \leq 300$ bases, indicating that these molecules remain in the weak trapping regime for these fields. The M>350 base molecules exhibit both the weak and the strong trapping regime. Indeed, their mobilities are observed to be constant up to a certain critical electric field and to decrease beyond this value. The broken curve shows the separation between the weak and strong trapping regimes as given by the $M^*(E)$ values found using the method described in Section 5.4.3. As an example, the inset of Fig. 5 shows a plot of the velocity V vs. the electric field E for a 600 base fragment. The velocity increases linearly with E for E<100 V/cm; the slope of the straight line fit then gives the trapping-free mobility $\mu = 7.39 \times 10^{-4}$ cm^2/(Volt min.). However, the velocity of the 600 base fragment then decreases very quickly for E>100V/cm. The tail of the V vs. E graph was fitted to the function $\exp(-\text{constant} \times E^2)$, the analytical form predicted by our dead end trapping model [13] and previous computer simulations [11]. The resulting fit is reasonable, but the error bars are so large that equally good agreement can be obtained using many other functions.

5.4.5 The Critical Size M^* vs. the Electric Field E

Figure 6 shows a log-log plot of the critical size M^* vs. the applied electric field E. The straight line fit has slope of -0.664, fully consistent with the relation $M^* \propto E^{-2/3}$ that we predicted for the dead end trapping model shown in Fig. 1b. In spite of the scatter of the points, our results clearly rule out fork trapping [10] which predicts a slope of -2.

The fact that $M^* \propto E^{-2}$ for fork trapping and $M^* \propto E^{-2/3}$ for dead end trapping can be understood using simple scaling arguments. For any trapping system, the time τ to get out of a trap is a function

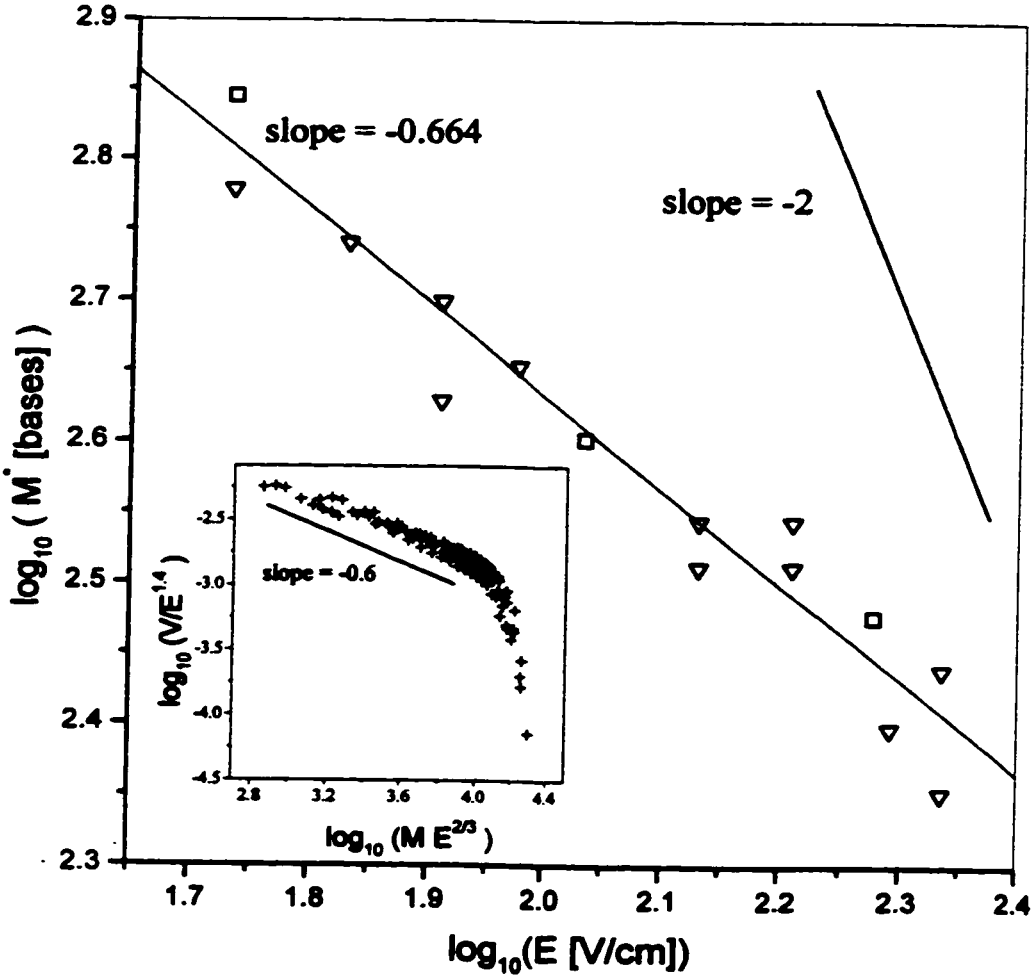


Figure 6: Log-log plot of the critical molecular size M^* vs. the field intensity E. Square points represent two identical results. The inset shows a log-log plot of $V/E^{1.4}$ vs. $M\,E^{2/3}$, where the velocity V is expressed in mm/min and the electric field E in V/cm.

of W/k_BT, where W is the work necessary to get out of the trap, k_B is the Boltzman constant and T is the temperature. This can be written as $\tau=\tau_0\times f(W/k_BT)$, where $f(W/k_BT\ll 1)\approx 1$ (weak trapping regime). The structure of the function $f(W/k_BT)$ in the limit where $W/k_BT\gg 1$ (strong trapping) is given by $\ln[f(W/k_BT\gg 1)]\propto W/k_BT$ (the Kramer approximation) [16-18]. Therefore, $W/k_BT\approx 1$ defines the critical value for the transition between weak and strong trapping. For both trapping mechanisms, the force holding the molecule in the trap is $F\propto Eh_x$, where E is the electric field and h_x is the end-to-end distance of the DNA molecules in the direction parallel to the field [11-23]. Since the DNA molecules retain their random-walk conformations (i.e., $h_x\propto M^{1/2}$) in such experiments, we obtain $W/k_BT\propto EM^{1/2}L_0$, where L_0 is the distance over which the molecule must move (backward) to get out of the trap. For fork trapping, L_0 does not depend on the molecular size and we have $W/k_BT\propto EM^{1/2}$, giving $M^*\propto E^{-2}$. On the other hand, for dead end trapping, the molecule needs to move completely out of the trap, i.e. $L_0\propto M^1$, and we obtain $W/k_BT\propto EM^{3/2}$ and $M^*\propto E^{-2/3}$.

5.4.6 A Universal Curve

The inset of Fig. 6 shows a log-log plot of $V/E^{1.4}$ vs $ME^{2/3}$ for all the data that we have obtained (i.e., for all values of E and M used in our study). A satisfactory universal curve is obtained. The -0.6 slope observed for small values of $ME^{2/3}$ corresponds to trapping-free regime where $V\propto E^1$. The abrupt decrease found for $ME^{2/3}\geq 10^{4.1}$ correspond to the strong trapping regime and indicates that $M^*E^{2/3}\approx 10^{4.1}$ or $M^*\approx 10^{4.1}E^{-2/3}$ which is consistent with the dead end trapping mechanism.

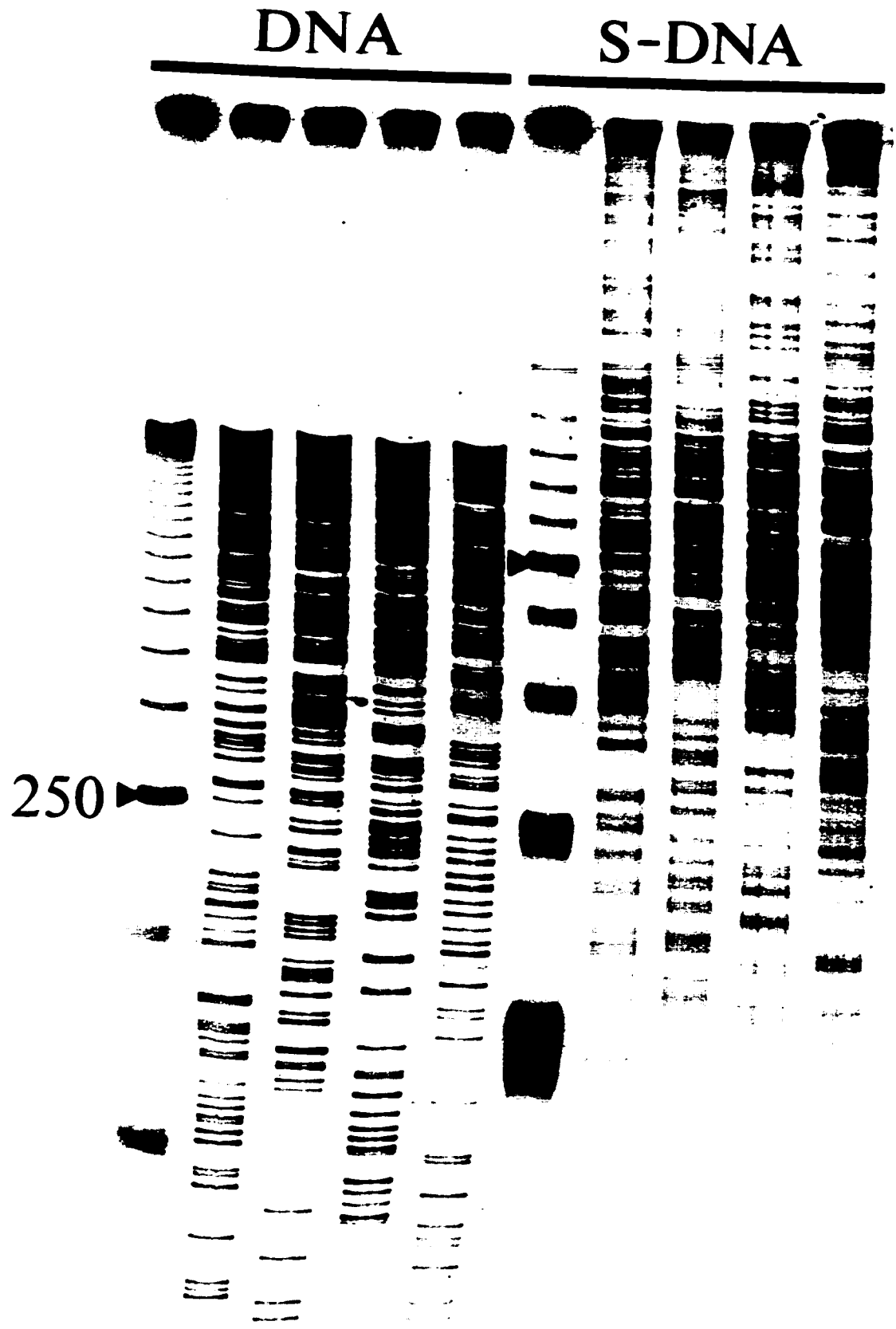
5.5 Discussion

Unfortunately, the fact that the mobility of S-DNA molecules is strongly field-dependent implies that even small field inhomogeneities can lead to detrimental anti-focusing effects. Since large field gradients do form naturally in the presence of urea [24], TE cannot be used with normal preruns for DNA sequencing. This is one of the important results of this chapter, and likely explains why new TE experimental results have not been published since the original publication of Ulanovsky et al. [10]. Here, we have demonstrated that inverted preruns are necessary to obtain optimal TE separations, and their use has allowed us to carry out the first systematic experimental study of TE. Our results agree nicely with the predicted scaling law $M^* \sim E^{-2/3}$, and we found that all our TE data could be represented by a universal curve (inset of Fig. 6).

Our experimental set-up did not allow us to measure reliable diffusion coefficients, a necessary ingredient to calculate the resolution. Therefore, in order to determine whether the resolution of DNA sequencing ladders could be improved using TE, we electrophoresed DNA and S-DNA sequencing reactions and 50-base ladders on a polyacrylamide gel with an inverted prerun (Figure 7). For unmodified DNA, the inter-band spacing decreases but the bands are narrower for large molecular size DNAs. Moreover, there are no DNA bands in the first 5-6 cm of gel because the M>1000 base fragments co-migrate at that position. For the streptavidin-DNA (S-DNA) molecules, the inter-band spacing decreases for M<400 bases, then increases for a while before it decreases again. As a result, the 400-650 base range is much better separated for S-DNA, and we find S-DNA bands over the whole gel (this is therefore a better use of the gel). However, the band widths, which decrease for S-DNAs smaller than 400 bases, increase for larger molecules, in

(See next page)

Figure 7: Autoradiogram of 50-base ladders and sequencing reactions (T, C, G and A) for DNA and S-DNA. The gel was run at 135 V/cm for 30 minutes. The two arrows mark the position of the M=250 base DNA and S-DNA molecules.


DNA
S-DNA
250

qualitative agreement with the predictions of our model [13]. This increased diffusion is due to the trapping mechanism and it greatly reduces the benefits of the increased inter-band spacing. In other words, the gain in interband separation is not enough to increase the overall resolution compared to unmodified DNA because the bands are broader.

On the other hand, the remarkable increase in inter-band spacing could possibly be used for specific applications, such as testing for point mutations or for accurate sizing. Another way to exploit TE would be to use monotonic field gradients where the electric field is increasing with the position throughout the gel. This could be achieved using gels where the gel thickness decreases from the loading end of the gel to the bottom of the gel. In such a gel, a S-DNA molecule of size M would migrate up to its own focusing point, i.e. where the local critical molecular size $M^*(E(x))=M$. This behavior was observed on Fig. 2A for the M=500 base fragment. This modified TE process would be quite analogous to protein isoelectric focusing.

From our results, we can also estimate a few microscopic parameters related to TE. The persistence length of single-stranded DNA is $p \approx 1.9$ nm in 0.5×TBE [25]. Using this value and the results of Rousseau et al.[26], we estimate that the mean pore size of our 8% polyacrylamide gels is about a=6.2 nm. The size and shape of streptavidin was studied by Hendrickson et al. [27] who report that streptavidin is effectively a sphere of radius $r_s \approx 2.5$ nm. It is possible to obtain a rough estimate of the fraction f_s of pores of size $x<r_s$ by using Lumpkin et al. [21] distribution of pore sizes $g(x)dx=2\pi\lambda x \exp(-\pi\lambda x^2)dx$, where $\lambda=1/(4a^2)$. The calculation gives

$$f_s = \int_0^{r_s} g(x)\, dx \approx \frac{1}{8.3}$$

This indicates that there is a fair amount of small pores. The fact that $p \approx a/3$ suggests that DNA is fairly stiff inside a pore; this may explain why our results agree with a theory that assumes that DNA is inextensible in its tube.

In conclusion, TE of S-DNA is characterized by a strongly field- and size- dependent electrophoretic mobility. While the latter makes it a promising method to increase separation, the former makes it extremely sensitive to field inhomogeneities. The presence of urea-related field gradients explains why it has been so difficult until now to reach definite conclusions about the usefulness and mechanisms of TE. Our study shows that the reptation model provides a satisfactory description of TE. Unfortunately, the latter predicted that increased band broadening would not make TE a good technique to improve the resolution of DNA sequencing ladders, a prediction that we have verified. Therefore, TE's usefulness would be limited to specific applications. We have suggested that with monotonic field gradients, TE could be transformed into an interesting isoelectric focusing-like technique. Alternatively, pulsed fields could be used to modulate the effect of the steric trapping [10,11,18]. However, it is extremely important to note that unless the field is uniform in space, pulsed fields are not likely to be of any use. Recently, Griess and Serwer [29] have demonstrated that ZIFE pulses [28] can be used to separate DNA and labeled DNA molecules because of a clever ratchet dynamical effect [30] (a theoretical model of the latter will be presented in Chapter 7). The latter, unfortunately, cannot be exploited for DNA sequencing purposes.

5.6 Appendix

Here, we follow the biased reptation model of Slater et al. [13]. The detrapping problem is similar to the diffusion of a particle starting on a reflecting wall at s=0 (the trap) and trying to reach an absorbing wall at $s=-L_0$ while having a velocity $(\partial s/\partial t)=\epsilon h_x/\tau_B>0$, where h_x is the end-to-end distance of the S-DNA in the field direction, $\tau_B=a^2/2D_s$ is the Brownian time, $D_s=k_BT/[\xi(N+\alpha)]$ is the curvilinear diffusion coefficient, and N is the number of gel pores occupied by the S-DNA. The particle must move backward (i.e. against the electric forces) over a distance L_0 in order to leave the steric trap. The general solution of eq 9 of Slater et al. [13] for the mean detrapping time τ_d is:

$$\frac{\tau_d}{\tau_B} = \frac{L_0^2}{a^2}\left[\frac{2}{3}\sqrt{\frac{\Omega}{\pi}}\,H(\Omega) + \frac{1}{\Omega}\,(1 - e^{\Omega}) - \sqrt{\frac{\pi}{-\Omega}}\,\mathrm{erf}(\sqrt{-\Omega})\right] \qquad \text{A1}$$

Here, the functions erf(Ω) and H(Ω) are defined by

$$\mathrm{erf}(\Omega) = \frac{2}{\sqrt{\pi}}\int_0^{\Omega} e^{-y^2}\,dy \qquad \text{A2}$$

$$H(\Omega) = \sum_{k=0}^{\infty} \frac{\Omega^k\,\Gamma(5/2)}{(k+1)\,\Gamma(k+5/2)} \qquad \text{A3}$$

where $\Gamma(x)$ is the gamma function and $\Omega \equiv 2N\epsilon^2 L_0^2/3a^2$. For small Ω's, we have

$$\frac{\tau_d}{\tau_B} \approx \frac{L_0^2}{a^2}\left[1 + \frac{2}{3\sqrt{\pi}}\sqrt{\Omega} + \frac{\Omega}{6} + \ldots\right] \qquad \text{A4}$$

Ω is thus the gauge that tells us if we are in the weak ($\Omega<1$) or strong ($\Omega>1$) trapping regimes. Dead end trapping implies that $L_0=Na$ (detrapping requires moving backward over the entire tube length) so that $\Omega=2N^3\epsilon^2/3$. For fork trapping [15], we have $L_0 \approx a$, i.e., detrapping requires moving backward over the distance of the order of a single pore and does not depend on the molecular size of the S-DNA fragments. In this case, $\Omega=2N\epsilon^2/3$. Therefore, strong trapping will occur for molecules larger than $N^*=3/(2\epsilon^2) \sim 1/\epsilon^2$ (or $M^* \sim 1/E^2$) for fork trapping or for molecules larger than $N^*=(3/(2\epsilon^2))^{1/3} \sim 1/\epsilon^{2/3}$ (or $M^* \sim 1/E^{2/3}$) for dead end trapping.

5.7 References

[1] Slater, G. W.; Drouin, G.; *Electrophoresis*, **13**, 574, 1992

[2] Grossman, P. D.; Menchen, S.; Hershey D.; *GATA*, **9(1)**, 9, 1992.

[3] Zimmermann, J.; Wiemann, S.; Voss, H.; Schwager, C.; Ansorge W.; *BioTechniques*, **17**, 302, 1994.

[4] McDonald, L. A.; Kelley, J. M.; Brandon, R. C.; Adams, M. D.; *BioTechniques*, **19**, 464, 1995.

[5] Heller, C.; Slater, G. W.; Mayer, P.; Dovichi, N.; Pinto, D.; Viovy, J.-L.; Drouin, G.; *Journal of Chromatography A*, **806**, 113, 1998

[6] Mayer, P.; Slater, G. W.; Drouin G.; *Anal. Chem.*, **66**, 1777, 1994.

[7] Davis, L. M.; Fairfield, F. R.; Harger, C. A.; Jett, J. H.; Keller, R. A.; Hahn, J. H.; Krakowski, L. A.; Marrone, B. L.; Martin, J. C.; Nutter, H. L.; Ratliff, R. L.; Shera, E. B.; Simpson, D. J.; Soper, S. A.; *GATA*, **8(1)**, 1, 1991

[8] Lindsay, S. M.; Philipp, M.; *GATA*, **8(1)**, 8, 1992.

[9] Chee, M.; Yang, R.; Hubbel, E.; Berno, A.; Huang, X. C.; Stern, D.; Winkler, J.; Lockhart, D. J.; Morris, M. S.; Fodor, S. P. A.; *Science*, **274**, 610, 1996.

[10] Ulanovsky, L.; Drouin G.; Gilbert, W.; *Nature*, **343**, 190, 1990.

[11] Slater, G. W.; Villeneuve, C.; *J. Polym. Sc. B:Polym. Phys.*, **30**, 1451, 1992.

[12] Desruisseaux, C.; Slater, G. W.; *Phys. Rev. E*, **49**, 5885, 1994.

[13] Slater, G. W.; Desruisseaux, C.; Villeneuve, C.; Guo, H. L.; Drouin G.; *Electrophoresis*, **16**,

704, 1995.

[14] Desruisseaux, C.; Slater, G. W.; *Electrophoresis*, **17**, 623, 1996.

[15] Ulanovsky, L.; *Methods in Enzymol.*, **216**, 54, 1992.

[16] Défontaines, A.-D.; Viovy, J.-L.; In *Proceedings of the First International Conference on Electrophoresis, Supercomputing and the Human Genome*; Cantor, C. R., Lim, H. A., Eds.; World Scientific: Singapore, 1991; pp. 286-312.

[17] Défontaines, A.-D.; Viovy, J.-L.; *Electrophoresis*, **14**, 8, 1993.

[18] Défontaines, A.-D.; Viovy, J.-L.; *Electrophoresis*, **15**, 111, 1994.

[19] Slater, G. W.; Noolandi, J.; *Biopolymers*, **25**, 431, 1986.

[20] Slater, G. W.; *Electrophoresis*, **14**, 1, 1993.

[21] Lumpkin, O. J.; Déjardin, P.; Zimm, B. H.; *Biopolymers*, **24**, 1573, 1985.

[22] Mayer, P.; Slater, G. W.; Drouin G.; *Appl. and Theo. Electrophoresis*, **3**, 147, 1993.

[23] Slater, G. W.; Mayer, P.; Drouin G.; *Methods in Enzymology*, **270**, 272, 1993.

[24] Desruisseaux, C.; Slater, G. W.; Drouin G.; *Electrophoresis*, **19**, 627, 1998.

[25] Tinland, B.; Pluen, A.; Sturm, J.; Weill G.; *Macromolecules*, **30**, 5763, 1997.

[26] Rousseau, J.; Drouin, G.; Slater, G. W.; *Phys. Rev. Lett.*, **79**, 1945, 1997.

[27] Hendrickson, W. A.; Pähler, A.; Smith, J. L.; Satow, Y.; Merritt, E. A.; Phizackerley R. P.; *Proc. Natl. Acad. Sci. USA*, **86**, 2190, 1989.

[28] Turmel, C.; Brassard, E.; Forsyth, R.; Hood, K.; Slater, G. W.; Noolandi, J.; (1990), In *Electrophoresis of Large DNA Molecules, Theory and Applications*; Lai, E., Birren, B. W., Eds.; Cold Spring Harbor Laboratory Press: New York, 1990; pp. 101-131.

[29] Griess, G. A.; Serwer, P.; *Biophys. J.*, **74(2)**, A71, 1998.

[30] Desruisseaux, C.; Slater, G.W.; Kist, T. B. L; *Biophysical Journal*, **75**, 1228, 1998.

Chapter 6

Trapping Electrophoresis of Labeled Single-Stranded DNA: The Effect of the Gel Concentration*

In this chapter, we present an experimental study of the dynamics of streptavidin-labeled single-stranded DNA fragments in crosslinked polyacrylamide gels of different concentrations (%T). The addition of the streptavidin label has two effects: a) it increases the friction coefficient of the molecules and, b) it causes steric trapping because some pores are smaller than the size of the label. The effect of the gel concentration on the friction coefficient and for the strength of the steric trapping were examined. Results obtained in the weak trapping regime show that the friction generated by the label is essentially independent of the gel concentration. The critical molecular size M^* that characterizes the transition between weak ($M<M^*$) and strong ($M>M^*$) trapping was measured as a function of the polyacrylamide gel concentration, and the results agree with the prediction of our biased reptation model. We show that in order to explain both the nature of the friction coefficient and the dependence of the critical molecular size M^* upon gel concentration, one must understand the relation between the mean pore size of the gel and the size of the label.

---

*The results presented in this chapter will be submitted for publication before the end of the current year.

6.1 Introduction

As discussed earlier in this thesis, Mayer, Slater and Drouin [1] proposed to use S-DNA hybrid molecules in order to achieve the separation of nucleic acids in free solution. This process, called End-Labeled Free-Solution Electrophoresis (ELFSE), actually seems to be a very promising way for DNA sequencing [2]. What makes separation possible in ELFSE is the extra friction due to the presence of the streptavidin label. Of course, the extra friction generated by the label is also present when the migration takes place in a polyacrylamide gel. Although this effect is weak compared to the effect of the steric trapping discussed in Chapter 5, it cannot be neglected for small molecular size DNAs migrating in a low electric field (the so called weak trapping regime).

A theoretical model for the friction coefficient of streptavidin during ELFSE was developed and experimental evidence on the validity of this model were described in the second chapter of this thesis. The present chapter first presents our theoretical understanding of this friction coefficient when the electrophoresis is performed in a cross-linked gel. This thus bridges the gap between trapping electrophoresis and end-labeled free-solution electrophoresis.

In the second part of this chapter, we investigate the effect of the gel concentration on steric trapping. The theory of trapping [3-9] was discussed in Section 5.2. The scaling law $M^* \sim E^{-2/3}$ was confirmed experimentally in Chapter 5. However, the effect of the gel concentration was not investigated.

It is obvious that the polyacrylamide gel concentration must have a strong effect on trapping since the latter is due to the presence of small pores. We remind the reader that in the reptation language, it is the number of gel pores occupied by the molecule, $N=M/M_a$, that matters and not the real molecular size M; note that N is itself a function of the polyacrylamide gel concentration since

M_a (the average amount of DNA in a gel pore) decreases with increasing polyacrylamide concentrations. Moreover, the critical molecular M* is the solution of the equation $\tau_d(M^*)=\tau_0(M^*)$, where $\tau_d(M)$ is the mean detrapping time (the time molecules stay inside a trap) and τ_0 is the mean trapping time (the time the molecules need to find the next trap). The latter is a function of the fraction f of pores smaller then the size of the streptavidin label, $\tau_0 \sim 1/f$; therefore it becomes obvious that the critical molecular size M* is a function of the polyacrylamide concentration through the parameter f. This has never been studied before. Experiments on the effect of the gel concentration on the critical molecular size M* were thus carried out in order to test our understanding of TE in disordered gels.

6.2 Theory of Trapping Electrophoresis

Ulanovsky, Drouin and Gilbert [10] reported that the shorter S-DNA fragments were only slightly retarded in their polyacrylamide gels while the electrophoretic velocity of the molecules larger than a certain critical size M*(E) was dramatically reduced, thus increasing the mean interband spacing. Qualitatively, the TE mechanism can be explained easily. We first note that when a S-DNA molecule migrates in a random gel, its velocity is reduced due to the extra friction generated by the streptavidin label; in other words, a S-DNA molecule having M bases endures a frictional drag larger than that of a "naked" DNA molecule of the same size. Moreover, in a gel, one has a wide range of pore sizes and structural defects [6]. If a S-DNA molecule stops migrating in the gel because it assumes a conformation wherein the protein gets sterically stuck, we say that the molecule is "trapped". In fact, trapping is believed to occur when the DNA chain enters a close loop that is too narrow for the label to follow (see Figure 1 of Chapter 5). The molecule then needs to

move backward over a distance $L_0(M)$ that is a function of the contour length L~M of the DNA fragment. In this chapter, we report a detailed study of the effects of the gel concentration in TE experiments. First, we will see that the relative friction coefficient of the label does not depend on the mean pore size of the gel, a somewhat surprising result that we explain using a generalized reptation model. Second, we will quantify the obvious fact that the polyacrylamide gel concentration (traditionally noted as %T) affects the number of potential TE traps, hence the value of the critical molecular size M*.

6.2.1 Trapping and detrapping

As discussed earlier, TE is characterized by two time scales: the trapping time τ_0 (the time it takes to find a new trap) and the detrapping time τ_d (the time it takes to get out of a trap). If V_{free} is the velocity of the S-DNA complex between traps, the net velocity $V_{S\text{-}DNA}$ in the presence of trapping is simply given by $V_{S\text{-}DNA}=V_{free}\tau_0/(\tau_0+\tau_d)$. When $\tau_d \ll \tau_0$, trapping is ineffective and $V_{S\text{-}DNA} \approx V_{free}$: this defines the weak trapping regime. When $\tau_d \gg \tau_0$, however, trapping dominates the dynamics and $V_{S\text{-}DNA}/V_{free} \approx [\tau_0/\tau_d] \ll 1$. The critical molecular size M* that marks the transition between these two regimes is thus the solution of the relation $\tau_0(M^*)=\tau_d(M^*)$.

In the BRM, the trapping time τ_0 is a function of the scaled molecular size $N=M/M_a$, the scaled electric field ϵ, and the fraction $f=f(\%T)$ of pores smaller than the size (radius) of the streptavidin. Since the average curvilinear distance traveled inside the reptation tube before encountering a trap is given by $d_0 \approx a/f$, we have

$$\tau_0(N, \epsilon, f) \approx \frac{d_0}{V_{tube}(N, \epsilon)} \approx \frac{a}{f\, V_{tube}(N, \epsilon)} \qquad (1)$$

where $V_{tube}(N,\epsilon)$ is the trap-free curvilinear velocity of the S-DNA molecule in its reptation tube. On the other hand, the mean detrapping time $\tau_d(N,\epsilon)$ does not depend upon f since it is a local property. Therefore, the critical scaled molecular size N^* is the solution of the equation:

$$f \approx \frac{a}{V_{tube}(N^*,\epsilon)\ \tau_d(N^*,\epsilon)} \qquad (2)$$

It is important to note that Eq. (2) predicts that N^* will actually be a function of the fraction f.

It is possible to obtain a rough estimate of the fraction f of pores of size $x<R_s$ using Lumpkin et al.'s pore size distribution function $g(x)dx=2\pi\lambda x \exp(-\pi\lambda x^2)dx$, where $\lambda=1/(4a^2)$ and a is the mean pore size of the gel [6]. The calculation gives

$$f = \int_0^{R_s} g(x)\ dx = 1 - e^{-\frac{\pi}{4}\left(\frac{R_s}{a}\right)^2} \qquad (3)$$

where R_s=2.5 nm is the radius of streptavidin here. According to Rousseau et al. [11], the amount of DNA per mean gel pore of size a is given by $M_a \approx 3200/(\%T)^{3/2}$ in our polyacrylamide gels. Using the value of the ssDNA persistence length p=1.9 nm estimated by Tinland [12] for our buffer, we can calculate the mean pore size a since, by definition, it must be equal to the radius of gyration $R_g(M_a)\approx(p\ L_a/3)^{1/2}$ of a molecule of size M_a (here, $L_a=M_a b$ is the contour length of this molecule, and b=0.43 nm is the linear size of a ssDNA base). This calculation gives $a \approx 29.5\text{nm}/(\%T)^{3/4}$. Eq 3 then gives $f \approx 1-\exp[-(\%T)^{3/2}/178]$. For our range of polyacrylamide concentrations (3 to 12%), this

relation predicts that the fraction of small pores is ranging from 3 to 21%! Finally, we note that the scaled critical size is given by $N^*=M^*/M_a \propto M^*(\%T)^{1.5}/3200$, while the reduced electric field scales like $\epsilon=qEa/2k_BT \sim EM_a a \sim E/(\%T)^{9/4}$.

6.2.2 The frictional contribution of the streptavidin label

In a gel, the mobility of a composite reptating molecule such as the S-DNA molecule studied here is given by:

$$\mu = \frac{1}{N} \frac{\sum_{i=1}^{N} \xi_i \, \mu_i \, \langle (\vec{u}_i \cdot \vec{x})^2 \rangle}{\sum_{i=1}^{N} \xi_i} \qquad (4)$$

where N is the number of pores occupied by the molecule, ξ_i, μ_i and $\mathbf{u}_i$ are respectively the friction coefficient, the free-solution electrophoretic mobility and the orientation (see Figure 1) of the molecular segment contained in the i^{th} pore ($\mathbf{x}$ is the orientation of the electric field as shown on the figure). This equation, suggested by Didier Long [13], essentially generalizes those of the biased reptation model for cases where the different segments are not identical.

Again, we define M_a as the number of DNA bases in a pore of mean size a (see Figure 2). When we add a streptavidin label at one end of our DNA fragment, $N>M/M_a$ since the bulky label also occupies some tube space. Let M' be the number of DNA bases required to fill the pore (or tube segment) occupied by the streptavidin (note that the later cannot occupy more than one pore

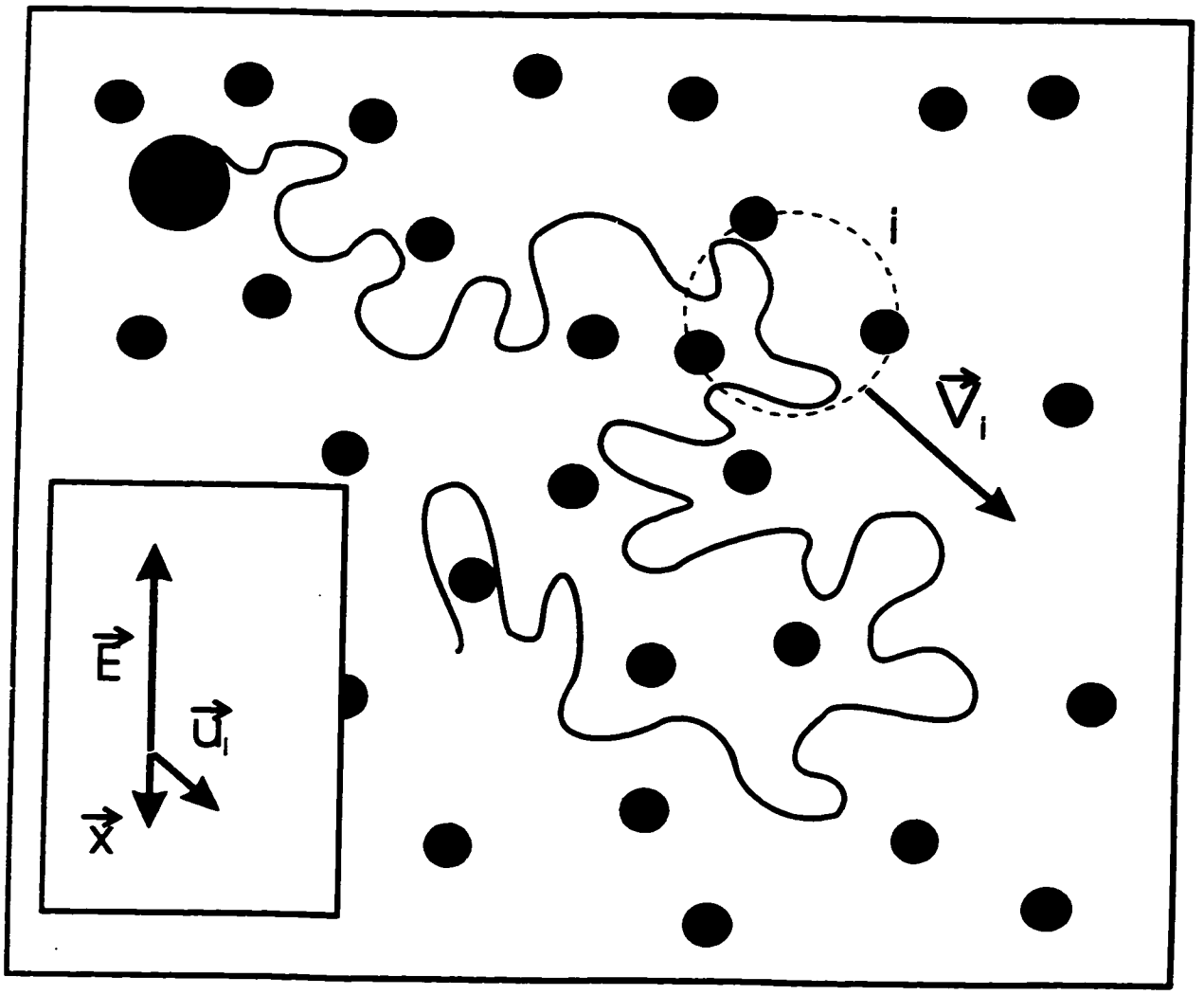


Figure 1: Schematic representation of a S-DNA molecule in a gel. **E** is the electric field, **x** is the direction of migration of the molecule, $\mathbf{u}_i$ is the orientation of the DNA contained in the i_{th} pore.

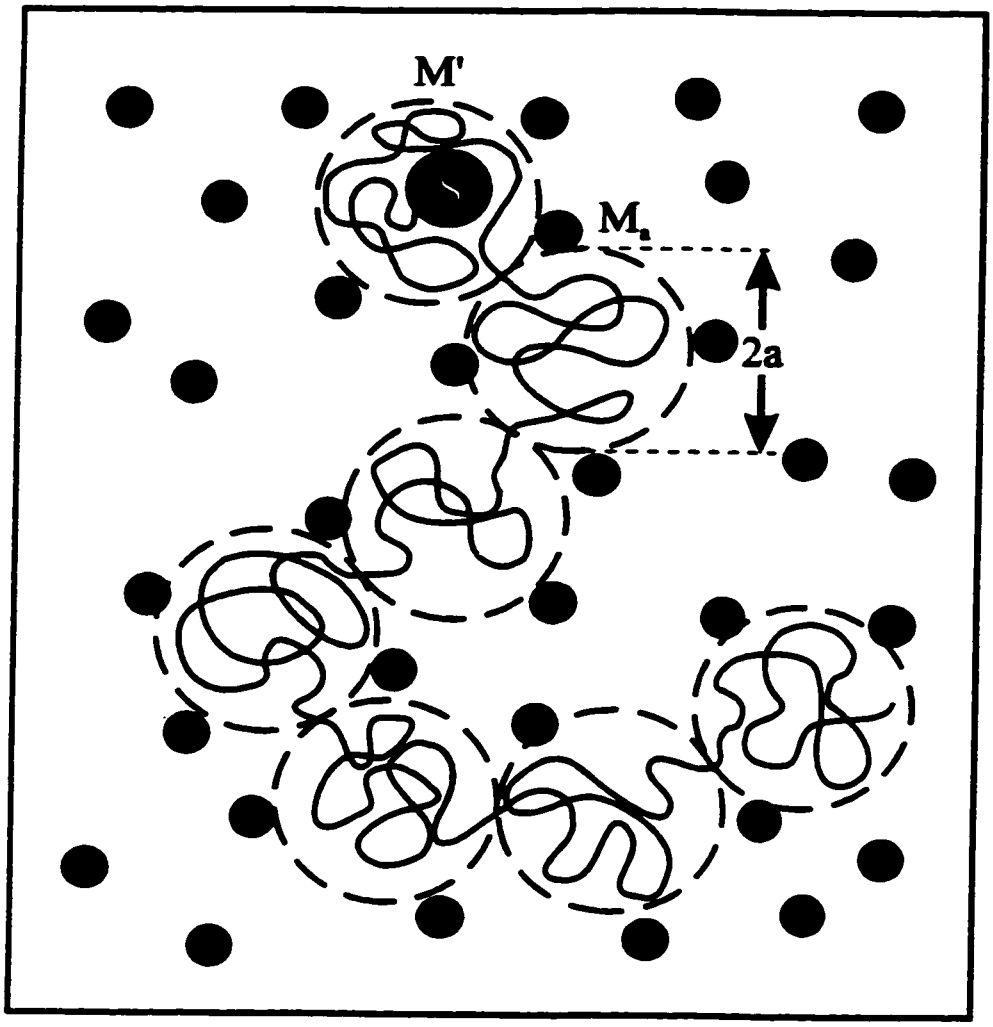


Figure 2: Schematic representation of the S-DNA molecule in a gel. There is on average M_a bases in each pore of size a. The pore containing the label contains only M′ DNA bases because of the space occupied by the label.

since this would not allow any migration in the gel). We then have

$$M = M' + (N - 1)\, M_a \qquad (5)$$

Our molecule is thus composed of $N-1=(M-M')/M_a$ identical pores containing M_a ssDNA bases and of one pore containing M' (where $0 \le M' < M_a$) ssDNA bases and the label.

The mobility of a reptation segment (or blob) containing only DNA is μ_0 (the size-independent free-flow mobility of DNA), while its friction coefficient is ξ_0. Although the reptation segment containing the label has the same friction coefficient ξ_0 (because this property is governed by the hydrodynamic size of the blob, which is itself given by the mean pore size a imposed by the gel), its mobility is $\mu_S=\mu_0 \times M'/(M'+\alpha)$, the free-flow mobility of a streptavidin-DNA molecule of M' bases (see eq 10 of Chapter 2). One can also say that $\mu_S=(M'/M_a)\times\mu_0$ since the charge is reduced by a factor M'/M_a for this special segment. Using the fact that in the low field limit (we are studying only the case where the reptation tube retains its random-walk conformation) we have $\langle(\mathbf{u}_i \cdot \mathbf{x})^2\rangle = 1/3$, Eq. (4) can be rewritten as:

$$\mu_{S\text{-}DNA} = \frac{\mu_0 (N - 1) + \mu_S}{3 N^2} = \frac{\mu_0}{3} \frac{M M_a}{[\, M + (M_a - M')\,]^2} \qquad (6)$$

Since the gel electrophoretic mobility of a naked DNA molecule can be expressed as $\mu_{DNA}=\mu_0 M_a/3M$ [4-6], Eq. (6) is equivalent to:

$$\mu_{S\text{-}DNA}(M) = \mu_{DNA}(M) \frac{M^2}{(M + \alpha)^2} \qquad (7)$$

where $\alpha=(M_a-M')$. In fact Eq. 7 can be written as:

$$\mu_{S\text{-}DNA}(M) = \mu_{DNA}(M + \alpha) \times \frac{M}{M + \alpha} \qquad (8)$$

This relation reflects the fact that the addition of the α-label affects the velocity both by adding extra friction (the $M/(M+\alpha)$ term) and by increasing the length of the reptation tube (the fact that we have $\mu_{DNA}(M+\alpha)$ and not $\mu_{DNA}(M)$ in the equation). We see that this equation is exactly the same as the one obtained for S-DNA in free solution (eq 10, Chapter 2: note that: $\mu_{DNA}(M)=\mu_{DNA}(M+\alpha)=\mu_0$ in this case) and in dilute polymer solutions (eq 7, Chapter 3).

6.3 Materials and Methods

6.3.1 Polyacrylamide Gels

Polyacrylamide gels with a 19:1 acrylamide (BDH) to bisacrylamide (BDH) ratio were used. They contained 5M urea (BDH), and a solution of 0.5×TBE [44.5 mM Tris (BDH), 44.5 mM boric acid (BDH), 1 mM EDTA (BDH)] was used as the running buffer. The 0.2 mm to 0.4 mm thick gels were prepared on a Pharmacia-LKB Macrophor apparatus. Polymerization was initiated with the addition of 0.1%w/v ammonium persulfate (BDH) and 0.1%v/v TEMED (BDH), and the gels were left to polymerize at room temperature for at least 12 hours.

6.3.2 DNA Samples

Radioactively labeled S-DNA 50 basepair ladders were prepared by mixing 1 μL of 50

basepair DNA ladder (1 μg/μL; Pharmacia), 3 μL of 5×sequencing buffer (200 mM Tris-HCl, pH 7.5, 100 mM $MgCl_2$, 250 mM NaCl), 4.5 μL of biotin-14-dCTP (33 pmol/μL; Life Technologies), 3 μL of [α-^{32}P] TTP (3.3 μM; Amersham), 3 μL distilled water and 0.5 μL Klenow (7.4 units/μL; Life Technologies), and this reaction was incubated for 15 minutes at room temperature. This produced radioactive (α-^{32}P) double-stranded DNA fragments with a biotin molecule at both ends. We then added 5 μL 0.5M EDTA, 80 μL TE (10 mM Tris, pH 8, 1mM EDTA) and 1 μL of 0.1 mM streptavidin (diluted in TE; Boehringer Mannheim), and we incubated this reaction for 15 minutes at room temperature. This reaction allowed a streptavidin molecule to bind to each of the biotin molecules that were previously added at both ends of the DNA fragments. We finally added 100μL of stop solution (95% formamide, 1% w/v xylene cyanol, 1% w/v bromophenol blue and 10 mM EDTA) and the samples were denatured at 65°C for 5 minutes. This step lead to the separation of the two DNA strands and therefore produced single-stranded DNA fragments having a streptavidin molecule at their 5′ end. Note that these denaturing conditions (incubation at 65°C for 5 minutes in 50% v/v formamide) did not break the streptavidin-biotin bonds. These denatured samples were subsequently kept on ice until they were loaded on the gel.

6.3.3 Electrophoresis Conditions

The running temperature (40°C) was controlled through the thermostatic plate of the Pharmacia-LKB apparatus. Experiments were performed on 18.5 cm-long gels using a Fisher Biotech FB600 power supply. The gels were prerun at the selected field E (in V/cm) for at least 150/E hours (e.g., for 3 hours at 50 V/cm) as we described previously (Chapter 4 and 5). The

direction of the electric field during the prerun was opposite to that of the electrophoresis. After that period of time, the electric current had fallen to about 40% of its initial value, and remained constant during the whole electrophoresis (less than 2% variation). As discussed in Desruisseaux et al. [14] (Chapter 5), this procedure is necessary to obtain reliable velocity data.

6.4 Results

6.4.1 The Friction Coefficient

We predicted that the velocity of S-DNA fragments should be related to the velocity of naked DNA fragments through the relation $\mu_{S\text{-}DNA}=\mu_{DNA}\times M^2/(M+\alpha)^2$ when the streptavidin-DNA complex is in the weak trapping regime. Therefore, if we plot $(\mu_{DNA}/\mu_{S\text{-}DNA})^{1/2}$ as a function of 1/M we should obtain a straight line of slope α crossing the origin at $\mu_{DNA}/\mu_{S\text{-}DNA}=1$. This is to be expected, of course, only if the DNA and S-DNA molecules are both in the reptation regime. In other words, we should not observe this linear behavior for molecules that are in other separation regimes (i.e., Ogston, entropic trapping, steric trapping, oriented reptation, etc.).

Figure 3 shows $(\mu_{DNA}/\mu_{S\text{-}DNA})^{1/2}$ as a function of 1/M for a 3% polyacrylamide gel run at the very low field of E=250V/18.5cm=13.5V/cm. We observe the predicted linear behavior for molecular sizes M larger than ≈350 bases, indicating that these molecules are not in the strong trapping regime. The value of the streptavidin friction coefficient obtained from the slope of the graph is α=47.2. We plotted this graph for different polyacrylamide concentrations %T and were able to obtain a value for α(%T) for %T ranging from 3% to 9%. Unfortunately, it was impossible to increase %T further since the large fraction (f) of small pores one has for %T>10% induces

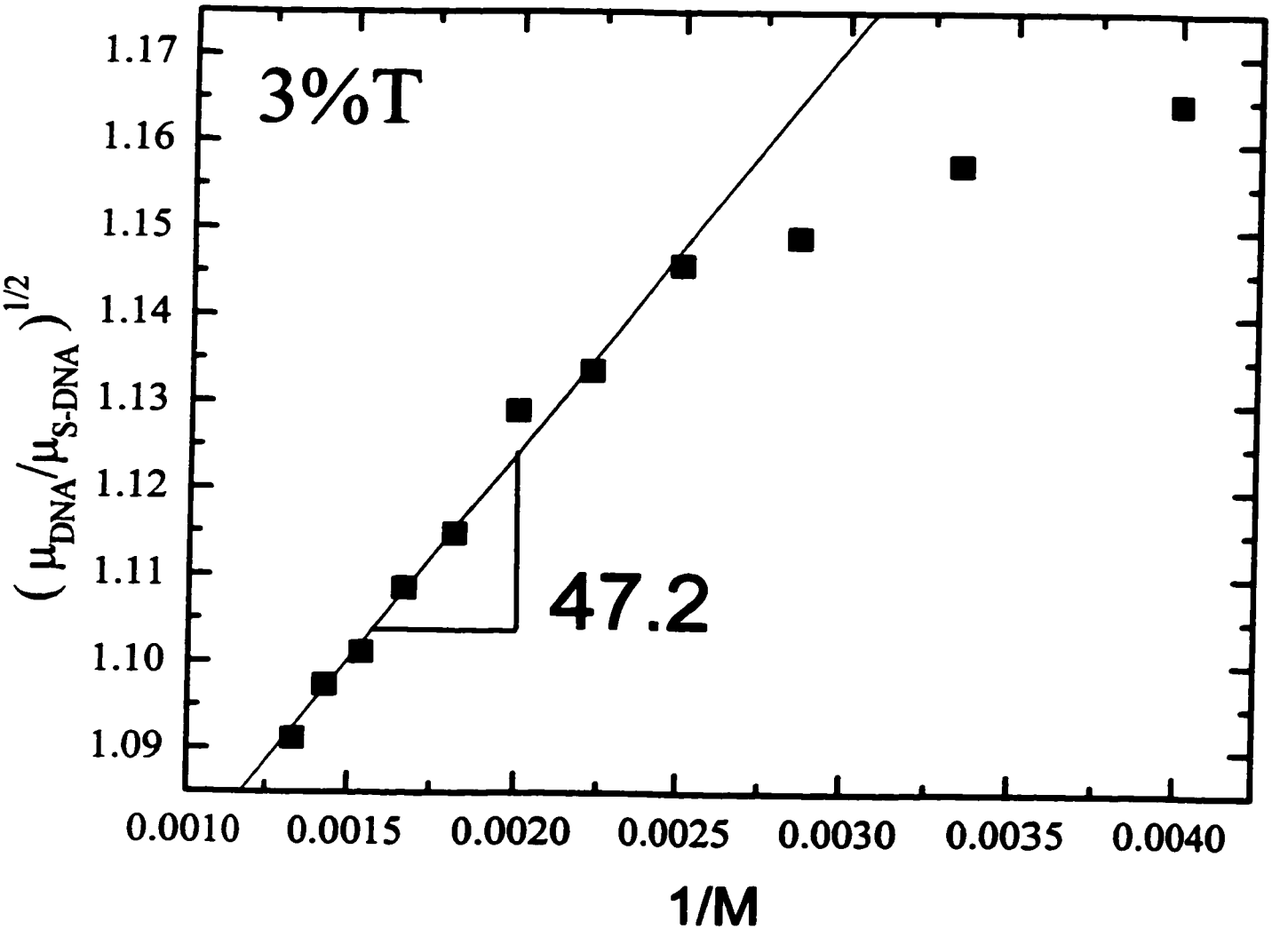


Figure 3: $(\mu_{DNA}/\mu_{S\text{-}DNA})^{1/2}$ as a function of 1/M for a 3% polyacrylamide gel run at 250V/18.5cm=13.5V/cm. The straight line represents the best fit y=1.03+47.2 x, indicating that α=47.2.

strong trapping even at low electric field intensities.

We also observe on Figure 3 that the small molecules do not satisfy eq 7. The curvature on Figure 3, for molecules smaller than M<400 bases (or 1/M>0.0025), is explained by the fact that small molecules are not in the reptation regime. Those DNA molecules migrate from one pore to another much like would do a spherical particle (the so-called Ogston regime). This is in agreement with the experimental observations of Rousseau et al. [11].

Different concentrations of polyacrylamide were thus studied and the friction coefficient α of streptavidin was estimated using the method shown in Figure 3. The experiments were performed at electric field intensities that were low enough to insure that we were in the weak trapping regime. We used electric fields of E=13.5 and E=27 V/cm; the values of α were similar in both cases, which indicates that we were indeed in the weak trapping regime and that our estimates were appropriate. Figure 4 shows a plot of α as a function of the polyacrylamide concentration %T. We see that α is essentially constant when we increase the concentration of polyacrylamide. The horizontal line corresponds to the mean of our 28 different values of α. The mean value of α=32±3 is in excellent agreement with the value α=32.4 observed in free solution in TAPS buffers with the same ionic strength (I≈0.05 mol/L) (Chapter 2, Figure 5). We also observe from Figure 4 that the fluctuations are quite large. Those fluctuations can be due to the fact that the fitting procedure oversimplifies the problem by assuming that the trapping is so weak that it can be neglected. Moreover, it is well-known that the gel structure (e.g., as characterized by the mean pore size a and the fraction f) fluctuates substantially from gel to gel because of the poorly controlled polymerization procedure.

It is also possible to estimate α by looking at the mobility of these small molecules that are

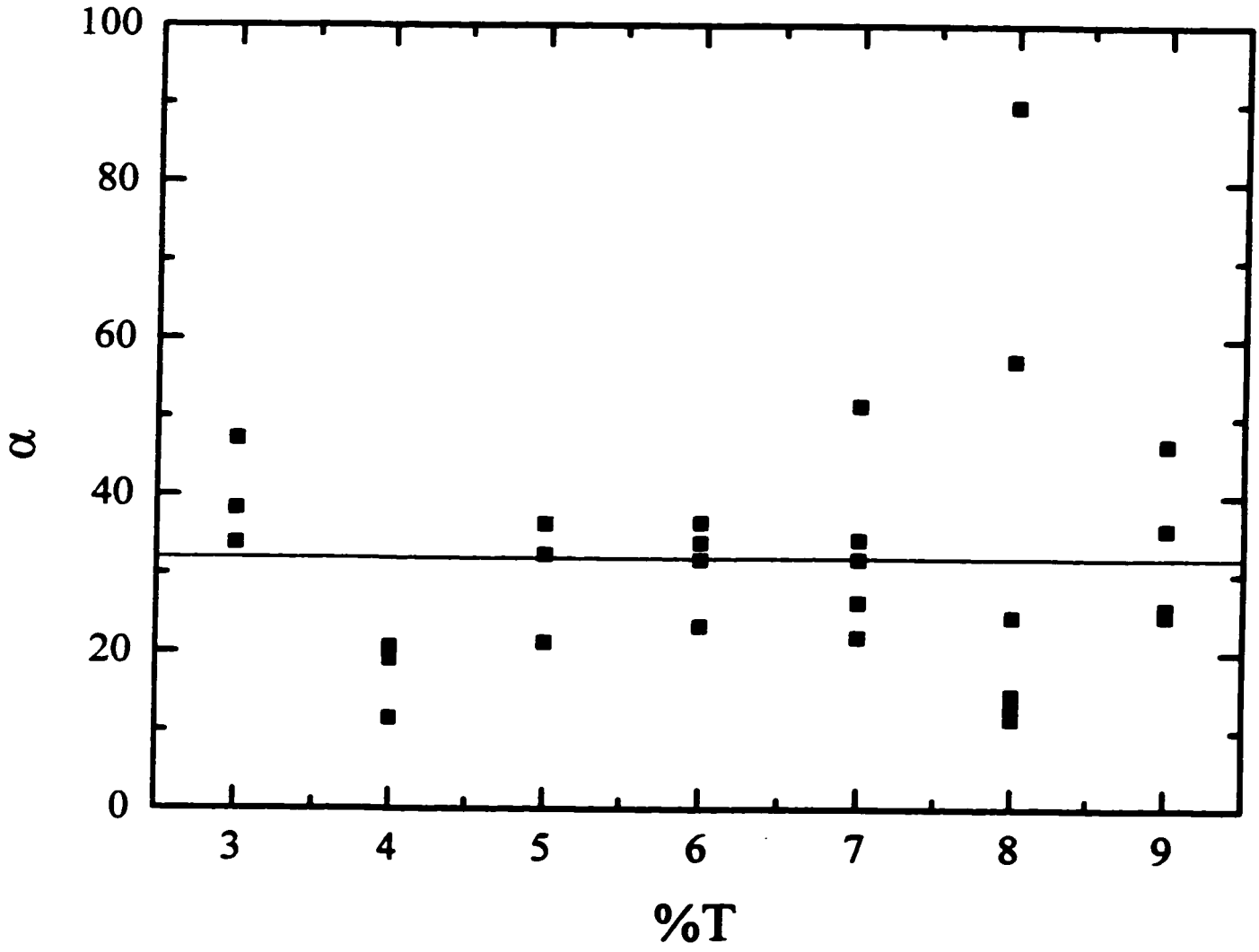


Figure 4: Effective friction coefficient α as a function of the polyacrylamide gel concentration %T. The straight line represents the mean value α=32±3.

not yet large enough to be in the reptation regime. We know that in free solution, the relation between the mobility of DNA and S-DNA fragments is simply given by $\mu_{S\text{-}DNA}=\mu_{DNA}\times M/(M+\alpha)$ (eq 8, Chapter 2). This should also be true here in the limit where the molecules are small and where the pores are large compared to the size of the label (very low concentrations of polyacrylamide). This means that for small molecules $\lim_{\%T\to 0}[\mu_{DNA}(M)/\mu_{S\text{-}DNA}(M+\alpha)]\times[M/M+\alpha]=1$. This ratio is plotted in Figure 5 for a molecular size M=50 and using the mean value of α obtained from Figure 4. We can see that these short S-DNA molecules are migrating, for concentrations %T<4%, as if they were in free solution. Also, the curve converges towards the right limit for large concentration %T; in that limit, the molecules are reptating (eq 7) and we should get $[\mu_{DNA}(M=50)/\mu_{S\text{-}DNA}(M=50)]\times[50/50+\alpha]=(M+\alpha)/M=82/50=1.64$, as shown on Figure 5. Figure 5 thus shows a smooth transition from free solution to gel electrophoresis.

6.4.2 M* vs %T

Further experiments were performed at high electric field (E= 4000V/18.5cm=216V/cm) using inverted preruns and multiple loadings, as described in Chapter 5 [14]. The goal here is to study how M*, the critical size at which trapping starts to occur, is affected by the concentration (%T) of polyacrylamide. As discussed in Chapter 5, M* can easily be determined by plotting dV/dM vs. M, since one then observes a maximum at M=M*.

Figure 6 shows a log-log plot of the critical size M* as a function of the concentration of polyacrylamide %T. We see that M* decreases with %T approximately as $M^*=1527/(\%T)^{0.70}$. This relation is not trivial. We know that when we increase the concentration of polyacrylamide in our

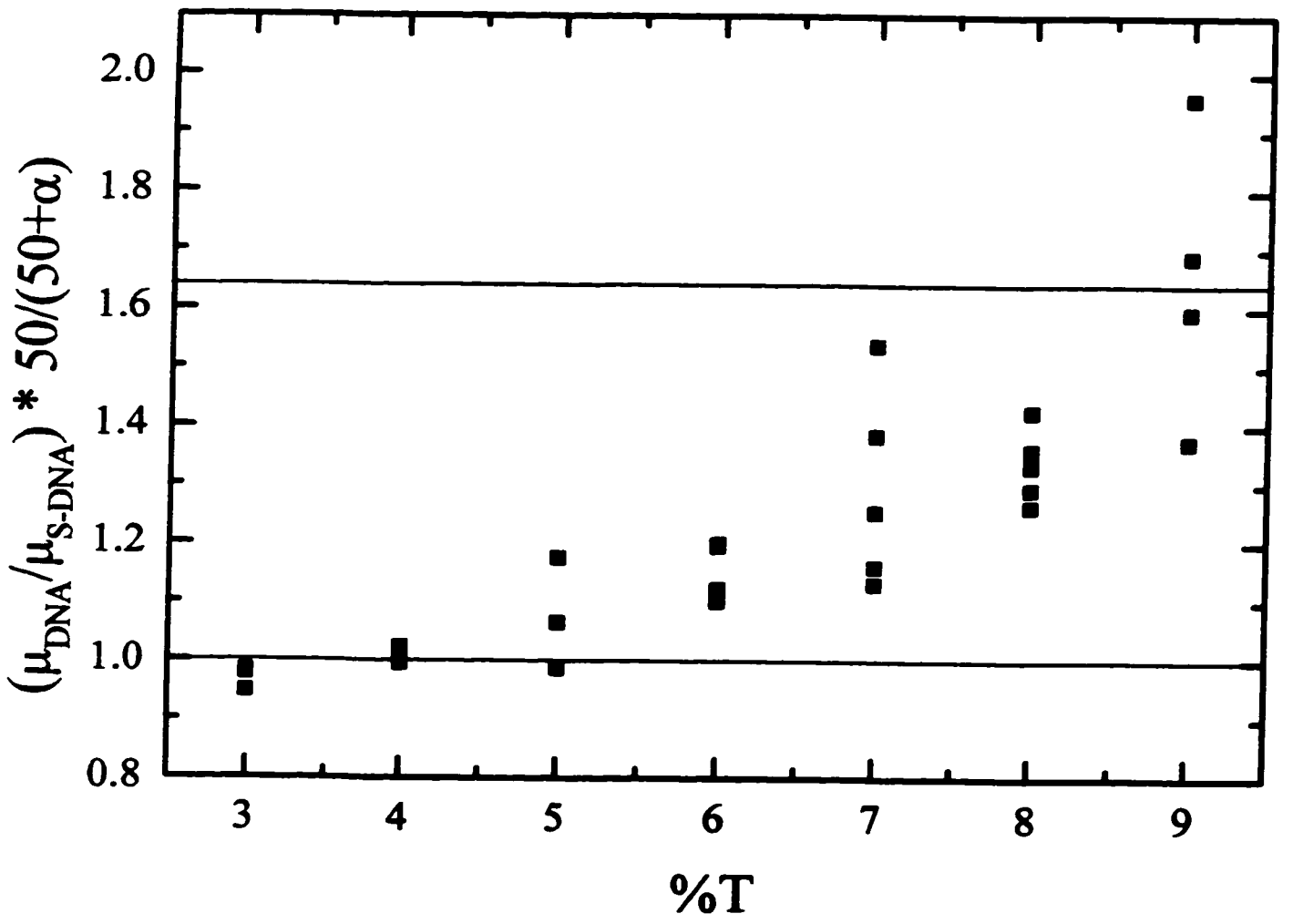


Figure 5: Plot of $\mu_{DNA}/\mu_{S\text{-}DNA} \times 50/(50+\alpha)$ as a function of %T. The value of α used in order to obtain this graph was $\alpha=32$, the mean value of α for concentrations ranging from 3%T to 9%T. We see that an ELFSE-like separation is observed for small concentrations (i.e., $\mu_{DNA}/\mu_{S\text{-}DNA} \times 50/(50+\alpha) \sim 1$), while the molecules are in the reptation regime in the opposite limit (i.e. $\mu_{DNA}/\mu_{S\text{-}DNA} \times 50/(50+\alpha) \sim 1.64$).

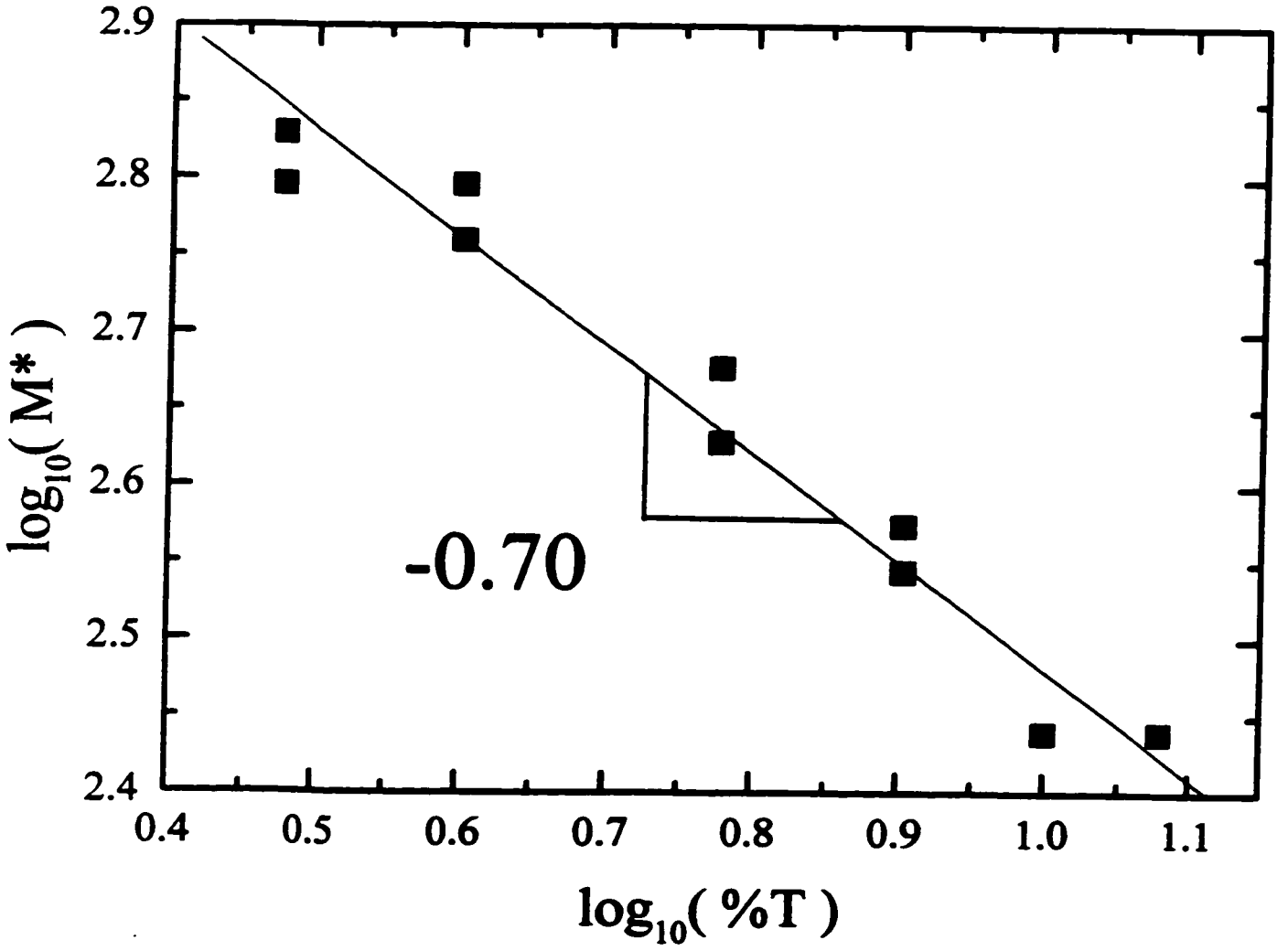


Figure 6: Log-log plot of the critical size M* at which trapping starts to occur as a function of the concentration of polyacrylamide %T. The straight line represents the best linear fit which gives $M^*=1527/(\%T)^{0.70}$ bases.

gels, the mean pore size a decreases while the fraction f of pores smaller than the size of the streptavidin increases. The smaller pore size increases the number of pores occupied by the DNA molecule (N) but it also decreases the value for the reduced electric field (ϵ). In order to test our reptation model, our data should be studied in terms of the reduced variables (N and ϵ) instead of M and E.

Figure 7 thus shows a log-log plot of N^* as a function of ϵ. We transformed the molecular size M^* (in bases) into the more convenient N^* (the number of pores) using the relation $N^*=M^*/M_a$, where $M_a \approx 3200/(\%T)^{1.5}$ [11]. We remind the reader that the reduced electric field $\epsilon=qEa/2k_BT=(M_ae)Ea/2k_BT$, where e is the charge of an electron and k_B is the Boltzmann constant. Since $M_a \approx 3200/(\%T)^{1.5}$ and $a \approx 29.52nm/(\%T)^{0.75}$, the reduced electric field scales like $\epsilon \propto (\%T)^{-2.25}$. In Chapter 5, we saw that the theoretical model and our experimental data (for different electric fields in 8%T polyacrylamide gels) both gave the scaling law $N^* \propto \epsilon^{-2/3}$. We can clearly see in Figure 7 that this relationship does not hold when we change the reduced electric field ϵ by changing the concentration of polyacrylamide; we obtain $N^* \sim \epsilon^{-0.35}$ instead. This is due to the fact that when %T is changed, the number of pores smaller than the size of streptavidin is also changed. In fact, the data (from left to right) correspond to 12%T, 10%T, 8%T, 6%T, 4%T and 3%T gels, so that the fraction of small pores (f) is decreasing from left to right. As we show below, incorporating the concept of f in our reptation theory can explain the effective slope of -0.35 observed in Figure 7.

As discussed previously, the critical molecular size N^* is the solution of the equation $\tau_0(N^*)=\tau_d(N^*)$; when we vary ϵ without changing the gel concentration %T or the fraction of small pores f, this gives $N^* \sim \epsilon^{-2/3}$ as observed in Chapter 5. In the current set of experiments, however, the

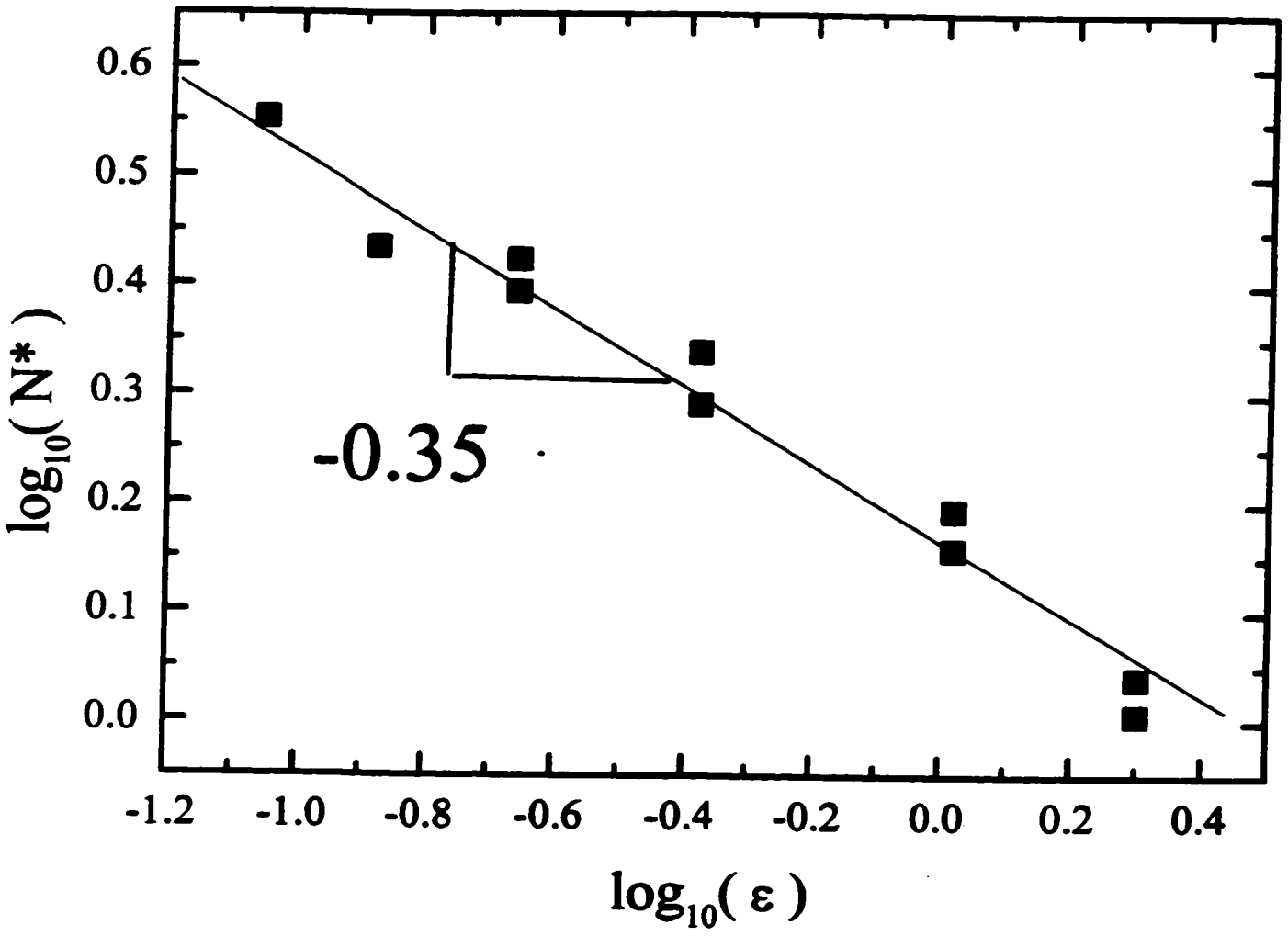


Figure 7: Log-log plot of N* as a function of ϵ.

fraction f of small pores increases with the concentration of polyacrylamide. It is possible to estimate the fraction f of small pores via two independent methods. First, if we use the distribution of pore sizes estimated by Zimm's group (eq 3) [6] in conjunction with the mean pore size calculated by Rousseau et al. [11], we obtain $f_{Zimm} \approx 1-\exp[-(\%T)^{1.5}/178]$. We can also estimate f_{BRM} using the measured critical molecular sizes N*(%T) and our reptation model (see the Appendix at the end of this chapter for the details). Figure 8 shows that we actually have $f_{BRM} \approx 5.4 \times f_{Zimm}$, indicating that the two numbers are proportional to one another. The factor of 5.4 represents the unknown constant in the equation for τ_0. The fact that the fraction of small pores calculated using both methods are consistent with one another seems to indicate that the biased reptation model is appropriate to explain our results. In other words, the slope of -0.35 observed in Fig. 7 is due to the gel concentration dependence of the key parameter f(%T).

6.5 Discussion

In this chapter, we presented a new theoretical model that explains the effect of adding a streptavidin label at one end of DNA molecules reptating in polyacrylamide gels. The effect of steric trapping was discussed in Chapter 5 and a free solution study of the label friction coefficient α was presented in Chapter 1. The current study focused on the effect of the gel concentration on both trapping and label friction. The concept of the relative streptavidin friction coefficient α was generalized to the case where the electrophoresis is performed in a crosslinked gel. The concept of the friction parameter α is rather subtle and requires a new scaling approach to electrophoresis theory. The new model takes into account not only the fact that there is more friction because of the

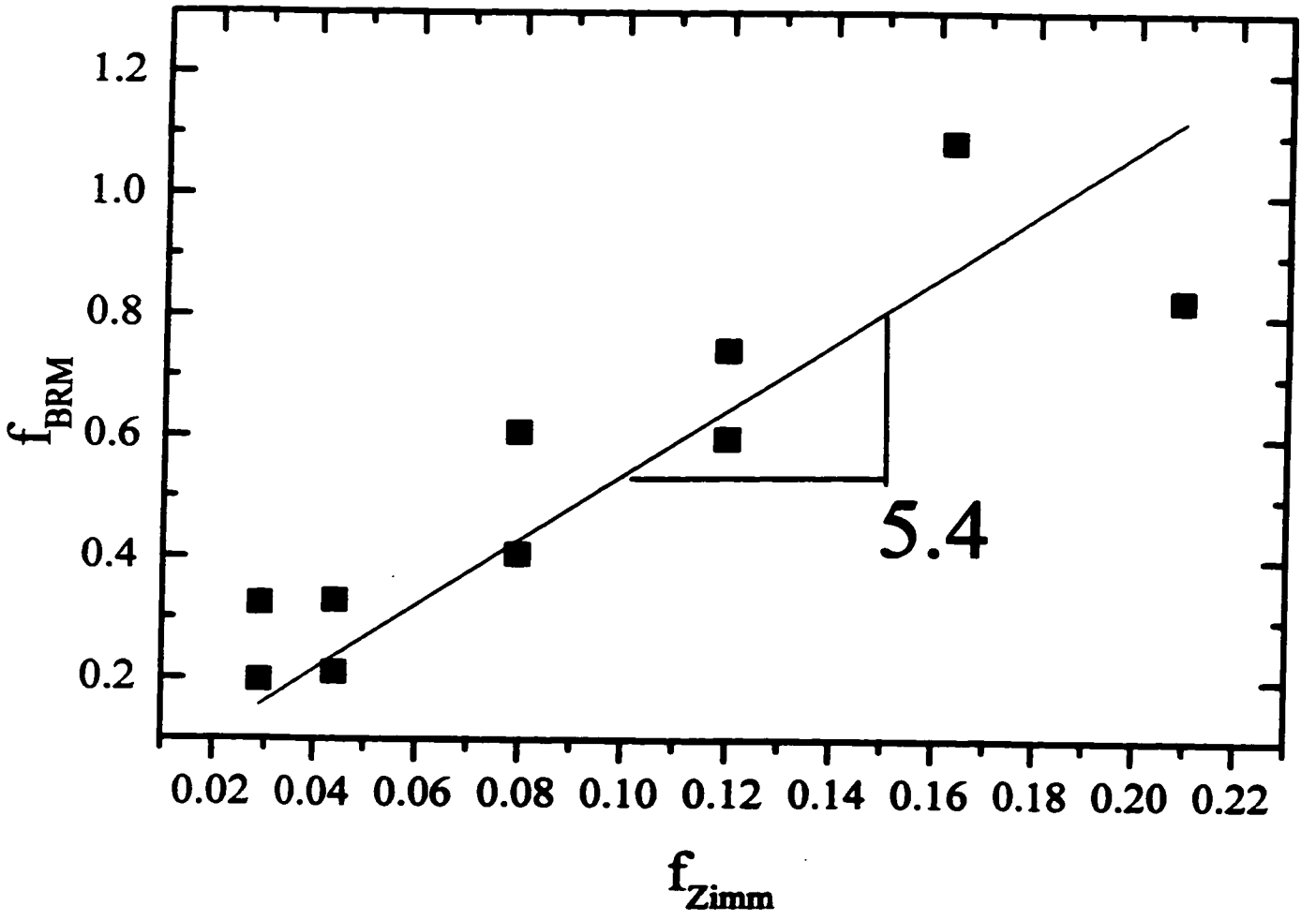


Figure 8 : Plot of f_{BRM} vs. f_{Zimm}. A straight fit y = 5.4 x is also shown.

presence of the label (like we observe in ELFSE experiments), but also the fact that the reptation tube is longer due to the presence of the label. In the reptation limit, the relation between the velocity of naked DNA and S-DNA fragments of the same size is given by $V_{S\text{-}DNA}(M)=V_{DNA}(M)\times M^2/(M+\alpha)^2$, where one $M/(M+\alpha)$ factor comes from the fact that there is some added friction due to the presence of the label, while the second $M/(M+\alpha)$ factor is due to the fact that a S-DNA molecule of size M occupies approximately the same space (i.e., the same reptation tube) as a DNA molecule having $M+\alpha$ bases.

In fact, the results of Chapter 2, 3 and 6 show that one single equation can be used (if there is no steric trapping) in free-solution, dilute solutions and gels:

$$V_{S\text{-}DNA}(M) = V_{DNA}(M + \alpha) \times \frac{M}{M + \alpha} \qquad (9)$$

It is quite remarkable that one single equation can summarize the situation in those three cases although the dynamics of the analyte is so different. It is not clear yet if trapping occurs in concentrated polymer solutions since the traps disappear with time due to the temporary nature of the polymer entanglements.

Experimental verification of the new model was presented in this chapter (e.g., see Figure 3) and it was indeed possible to observe the transition from gel electrophoresis to free solution electrophoresis for small S-DNA molecules in low concentration polyacrylamide gels (Figure 5). Remarkably, the value of α observed for small molecules was in fair agreement with the value of α obtained for the largest molecules that were in the reptation regime. This shows that the relation

$\mu_{S\text{-}DNA}(M)=\mu_{DNA}(M+\alpha)\times M/(M+\alpha)$ is also valid for molecules that are in the so-called Ogston sieving regime.

The relative friction coefficient (α) of the streptavidin label does not seem to vary significantly with polyacrylamide concentration. This suggests that our definition of the reptation blob is valid: the DNA blob containing the streptavidin and some DNA has the same friction coefficient as the blobs containing only DNA. It is obvious that the size (and hence friction coefficient) of these blobs increases when we increase the polyacrylamide concentration; however, since the friction coefficients of all the blobs are equal this does not change the relative friction coefficient of the streptavidin label. The value of $\alpha \approx 32\pm3$ that we obtained for electrophoresis in gels of ionic strength I=0.045 mol/L is in very good agreement with the value of α=32.4 obtained in ELFSE with TAPS buffers of the same ionic strength (see Chapter 2). This is a strong prediction of these theoretical models since they share the same definition for α.

A study of the trapping mechanism was also carried out as a function of the concentration of polyacrylamide. The study was performed at high electric field (E=216 V/cm) in order to observe the transition from weak to strong trapping. The critical molecular size M* at which trapping starts to occur is decreasing with the concentration of polyacrylamide roughly like $M^*=1527/(\%T)^{0.70}$. In order to understand the physical mechanism responsible for this apparent scaling law, we plotted the scaled critical molecular size N* as a function of the reduced electric field ϵ. We observed that the critical molecular size decreases like $N^*\sim\epsilon^{-0.35}$ when the reduced field ϵ is varied by changing the concentration of polyacrylamide %T (as a comparison, we saw in Chapter 5 that when the reduced field ϵ is reduced by changing the electric field E, this relation reads $N^*\sim\epsilon^{-2/3}$ instead). This

difference is not surprising since the small values of ϵ actually correspond to large gel concentrations %T here, implying that the fraction f of small pores also decreases when ϵ increases.

We also estimated the fraction f of small pores using two independent approaches. The first consisted in using the theoretical distribution of pore sizes proposed by Lumpkin et al. [6], together with the experimental results of Rousseau et al. [11]. The second method used the present experimental data and the biased reptation model of TE. The fact that the two estimates of the parameter f were linearly correlated demonstrates that we have a consistent description of the problem. Apparently, the S-DNA molecule needs to migrate over a curvilinear distance 5.4 times larger than the mean distance a/f between consecutive steric trap; this might be due to the fact that it is unfavorable for DNA to be in small pores because of a reduction in its entropy [11] so that the molecule naturally avoids these pores.

In conclusion, we now have a reliable reptation model of TE and a fundamental understanding of the empirical friction parameter α. In the first theoretical models of TE [3, 7-9], the friction coefficient of streptavidin was not fully understood and an empirical parameter was required in order to take it into account. Moreover, the effect of the fraction f of small pores was not considered. We now know that the effect of f is very important in TE and that we cannot have a good understanding of the problem without taking it into account (at least empirically).

6.6 Appendix

The trapping time τ_0 is a function of the molecular size M, the electric field E and the fraction f(%T) of pores smaller than the size (radius) of the streptavidin. The average curvilinear distance traveled inside the reptation tube before encountering a trap is given by $d_0 \propto a/f$. In the BRM, the curvilinear velocity of a S-DNA molecule is given by $V_{tube}=\epsilon h_x/\tau_B(N)$, where h_x is the end-to-end distance of the tube in the field direction and $\tau_B(N)\sim N\sim M$ is the Brownian time required to diffuse over a curvilinear distance of one gel pore (a) when E=0. In absence of tube elongation or orientation, one has the random-walk value $\langle|h_x|\rangle \approx a[2N/3\pi)]^{1/2}$. The trapping time $\tau_0 \propto d_0/V_{tube}$ is then given by

$$\tau_0(N,\ \epsilon,\ f) \propto \frac{\tau_B(N)}{f\ \epsilon\ \sqrt{2N/(3\pi)}} \qquad \text{(A1)}$$

On the other hand, the mean detrapping time only depends on the molecular size M and on the electric field E (see Appendix of Chapter 5).

The transition from weak ($N<N^*$) to strong trapping ($N>N^*$) happens when $\tau_0(N^*,\epsilon,f)=\tau_d(N^*,\epsilon)$. Therefore, from the equation for $\tau_0(N,\epsilon,f)$ and $\tau_d(N,\epsilon)$, one gets:

$$f \propto \frac{\tau_B(N^*)}{\epsilon\ \sqrt{2N^*/(3\pi)}\ \tau_d(N^*,\epsilon)} \qquad \text{(A2)}$$

for the fraction of pores smaller than the size of the label. This equation actually overestimate f

because of entropic reasons. Indeed, it is more difficult for DNA to have access to smaller pores (there is an reduction in its entropy). Figure 8 suggests that the rhs of eq A2 overestimates f by a factor of ≈5.4. In other words, only ≈1/5.4=19% of the small pores of radius $a \leq R_S$ are actually visited by the DNA molecules. This entropic effect was discussed by Rousseau et al [11].

6.7 References

[1] Mayer, P.; Slater, G. W.; Drouin, G.; *Anal. Chem.*, **66**, 1777, 1994.

[2] Ren, H.; Karger, A. E.; Oaks, F.; Menchen, S.; Slater, G. W.; Drouin, G.; *Electrophoresis*, **20**, 2501, 1999.

[3] Slater, G. W.; Villeneuve, C. *J. Polym. Sc. B:Polym. Phys.*, **30**, 1451, 1992.

[4] Slater, G. W.; Noolandi, J. *Biopolymers*, **25**, 431, 1986.

[5] Slater, G. W. *Electrophoresis*, **14**, 1, 1993.

[6] Lumpkin, O. J.; Déjardin, P.; Zimm, B. H. *Biopolymers*, **24**, 1573, 1985.

[7] Slater, G. W.; Desruisseaux, C.; Villeneuve, C.; Guo, H. L.; Drouin, G.; *Electrophoresis*, **16**, 704, 1995.

[8] Défontaines, A.-D.; Viovy, J. L.; *Electrophoresis*, **14**, 8, 1993.

[9] Défontaines, A.-D.; Viovy, J. L.; *Electrophoresis*, **15**, 111, 1994.

[10] Ulanovsky, L.; Drouin, G.; Gilbert, W.; *Nature*, **343**, 190, 1990.

[11] Rousseau, J.; Drouin, G.; Slater, G. W. *Phys. Rev. Lett.*, **79**, 1945, 1997.

[12] Tinland, B.; Pluen, A.; Sturm, J.; Weill, G.; *Macromolecules*, **30**, 5763, 1997.

[13] Long, D.; personal communication.

[14] Desruisseaux, C.; Slater, G. W.; Drouin, G.; *Macromolecules*, **19**, 6499, 1998.

Chapter 7

Trapping Electrophoresis and Ratchets: a Theoretical Study for DNA-Protein Complexes*

Recently, Griess and Serwer [1] showed that it was possible to use trapping electrophoresis and unbiased but asymmetric electric field pulses to build a type of correlation ratchet which would allow the efficient separation of naked DNAs from identical DNAs that form a complex with a bulky object such as a protein. Here, we present a theoretical investigation of this novel macromolecular separation process. We start by looking at the general features of this electrophoretic ratchet mechanism in the zero-frequency limit. We then examine the effects of finite frequencies on velocity and diffusion. Finally, we use the biased reptation model and computer simulations to understand the band broadening processes. Our study establishes the main experimental regimes that can provide good resolution for specific applications.

---

*The results of this chapter were published in *Biophysical Journal*, **75**, 1228, 1998.

7.1 Introduction

When a particle moves in an asymmetric but periodic potential under the action of non-equilibrium fluctuations, a net drift can be observed even though the net applied force is zero [2]. Similar effects can be observed for asymmetric fluctuations and symmetric potentials [3], or for various schemes where the potential itself is fluctuating [4]. These systems are often referred to as correlation ratchets (CR). Such thermodynamic systems have attracted a lot of attention recently as a model for molecular motors [5,6]. Applications to the field of separation science have also been proposed. For example, Rousselet et al. [7] have reported the successful (albeit inefficient) separation of particles using a simple electrostatic potential with no net force. Slater et al. [8] proposed a system where asymmetric steric interactions may be used to separate macromolecules based solely on their internal entropy. Chacron and Slater [9] suggested an electrophoresis system where a correlation ratchet uses a strong field gradient to force the migration of the molecules towards unique fixed (spatial) points where the resulting bands self-focus (much like for isoelectrofocusing of proteins).

Gel electrophoresis is the main separation tool of modern molecular biology laboratories [10]. For example, current DNA sequencing and mapping methods are entirely based on the ability of gel electrophoresis to separate DNA fragments that differ by less than 1% in size. Proteins are also routinely separated by gel electrophoresis [11,12]. One early example of a simple CR-like electrophoretic process was ZIFE (for Zero-Integrated Field Electrophoresis, see Figure 1) [13]. In this process, an unbiased (average field is zero) pulsed field is applied, with short high field pulses of intensity ε_H and duration T alternating with longer low field pulses of intensity $-\varepsilon_L$

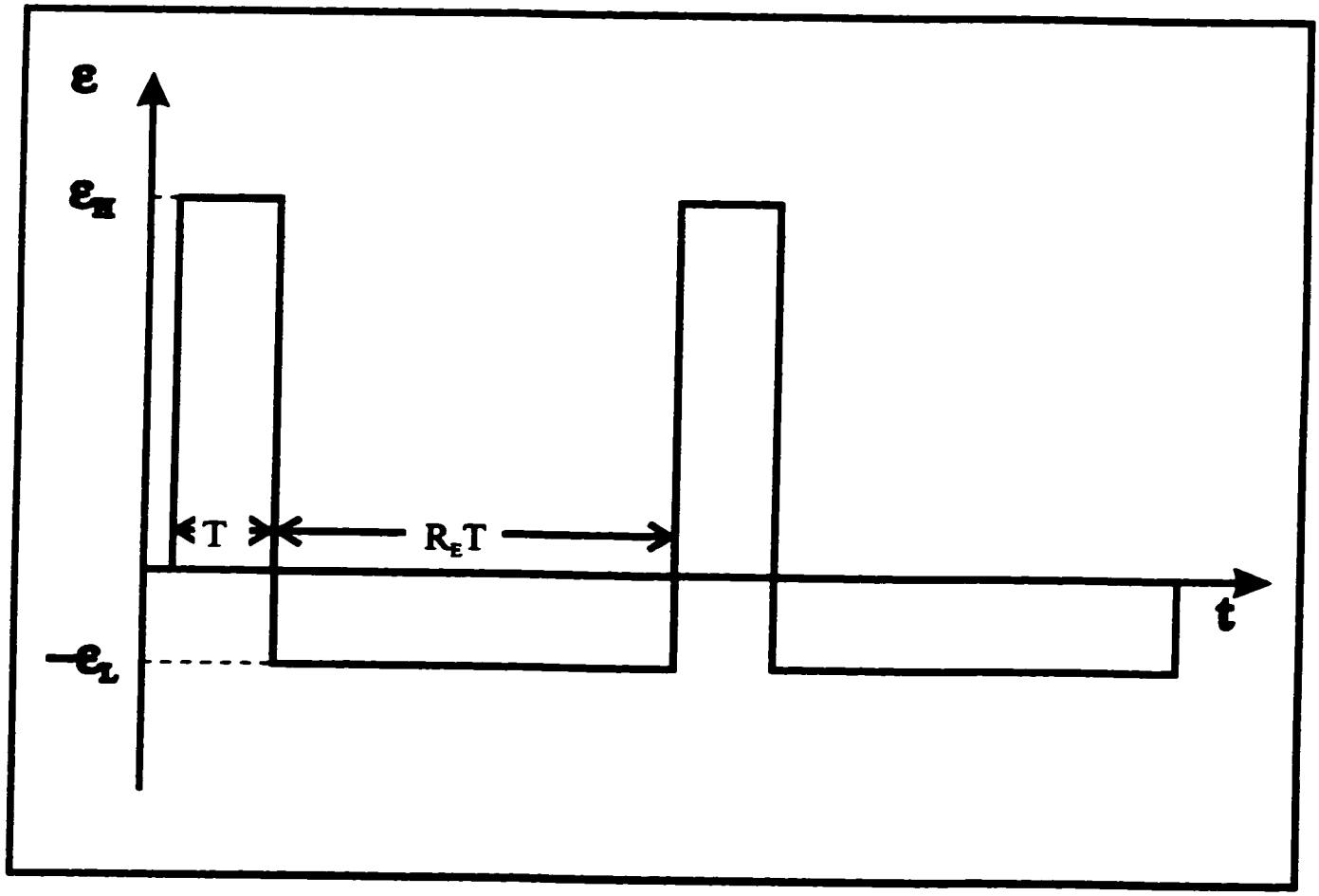


Figure 1: The time profile of the applied electric field for Zero-Integrated Field Electrophoresis (ZIFE). Here ε_H and ε_L denote the high and the low field intensities, respectively, and the ratio R_E is given by $R_E=\varepsilon_H/\varepsilon_L$. The duration of one complete period is $(1+R_E)T$ and the mean field is zero.

and duration $(\varepsilon_L/\varepsilon_H)\times T$ (see Figure 1). Due to the non-linearities (the electrophoretic mobility μ of DNA increases with field intensity), a large DNA molecule acquires a finite velocity in the direction of the high field pulses in spite of the absence of a net (integrated) external applied field.

A simple ratchet process can also be constructed for large spherical particles being electrophoresed in tight gels. Indeed, it is well-known that a large particle quickly gets trapped (stop moving) in a tight gel if the field ε is too high (i.e., if it exceeds some trapping critical value ε_T) [14,15]. This is apparently due to the fact that the field drives such a particles into dead-ends where it stays trapped for extended periods of time [14]. If the field is low enough ($\varepsilon<\varepsilon_T$), the Brownian motion eventually frees the particle from the trap and the migration resumes. ZIFE-like pulses, where we alternate between a high ($\varepsilon_H>\varepsilon_T$) and a reverse low ($\varepsilon_L<\varepsilon_T$) field, will clearly lead to a net motion in the direction of the low field intensity for this system. Note that ZIFE pulses actually make particles and DNA molecules move in opposite directions.

Ulanovsky et al. [16] suggested to attach a bulky object (the protein streptavidin in the original paper) to one end of the DNA molecules prior to electrophoresis in order to increase the separation between different DNA sizes. The idea is simple: while the protein's size is responsible for the trapping in dead-ends, it is the electric force on the DNA molecule (hence the DNA charge) that restricts the escape from the traps. Therefore, larger DNA molecules are trapped for longer periods of time and their mobility is more severely reduced. This process was called Trapping Electrophoresis (TE). The original experimental results [16] were promising, but theoretical investigations [17-23] and a more recent experimental study [24] (see Chapter 5) have demonstrated that TE suffers from an explosive increase of the diffusion coefficient. It is still unclear whether one

can use pulsed fields (a reverse pulse is clearly an efficient way to detrap molecules) to improve the situation [16, 21, 23].

Serwer's group studied various problems related to the gel electrophoretic migration of DNA molecules that carry a bulky object [15, 25]. In recent articles, Griess and Serwer [1] presented a simple ratchet idea for the separation of a naked DNA molecule from an identical DNA molecule that carries a bulky object (in the rest of this chapter, this will be called a S-DNA complex, and the object will be called the label). These authors subjected a mixture of DNA and S-DNA to a ZIFE-like pulse sequence. They observed that the DNA was moving in the direction of the high field pulses, while the S-DNA was moving in the opposite direction. As mentioned above, these two effects are due qualitatively to non-linearities and trapping, respectively.

In this Chapter, we present the first theoretical analysis of this original macromolecular separation technology for a case where the label is attached at one end of the DNA (a similar theory can be derived for other cases). We first examine the general features (i.e., the possible operating regimes) of the process for zero-frequency and low frequency pulses. This analysis is model-independent and uses only the well-known electrophoretic properties of DNA and the known TE results. We then use computer simulations to obtain quantitative results within the framework of the Biased Reptation Model (BRM). Finally, we draw conclusions about the possible usefulness of this novel idea.

7.2 General Theoretical Principles

7.2.1 Electrophoresis of a DNA molecule in a DC electric field

It is well-known that when the electric field is low enough, long reptating DNA molecules retain their random walk conformations (see Figure 2b) during the electrophoretic migration. This is the linear regime characterized by a field-independent mobility μ. Heller et al. [26], for example, have clearly established this behaviour for double-stranded DNA electrophoresed in agarose gels. When the field ε exceeds a critical reptation value ε_R, however, the molecule becomes oriented in the field direction (Figure 2d) (see also eq 1, Chapter 1) [27]. This orientation reduces the retarding effect of the gel. The mobility then becomes field-dependent [$\mu(\varepsilon)$ is a monotonically increasing function of ε; see Figure 3a].

7.2.2 Electrophoresis of a S-DNA molecule in a DC electric field

If it were not of steric trapping, S-DNA complexes would have the same general electrophoretic properties as naked DNAs (except for an extra friction coefficient which we will note α; in other words, the streptavidin has a friction coefficient equivalent to that of a DNA fragment containing α monomers (see Chapter 6)). However, DNA reptation leads to situations such as the one shown in Figure 2c. In this case, the bulky object cannot follow the reptating molecule because the latter previously chose a tube section that was too narrow for the S-label. The oriented (Figure 2c) S-DNA molecule is thus strongly trapped since it is normally the charged DNA head that drives the migration. If the field is low enough, i.e. lower than a trapping critical value ε_T, trapping is of little importance because the Brownian forces dominate the electric

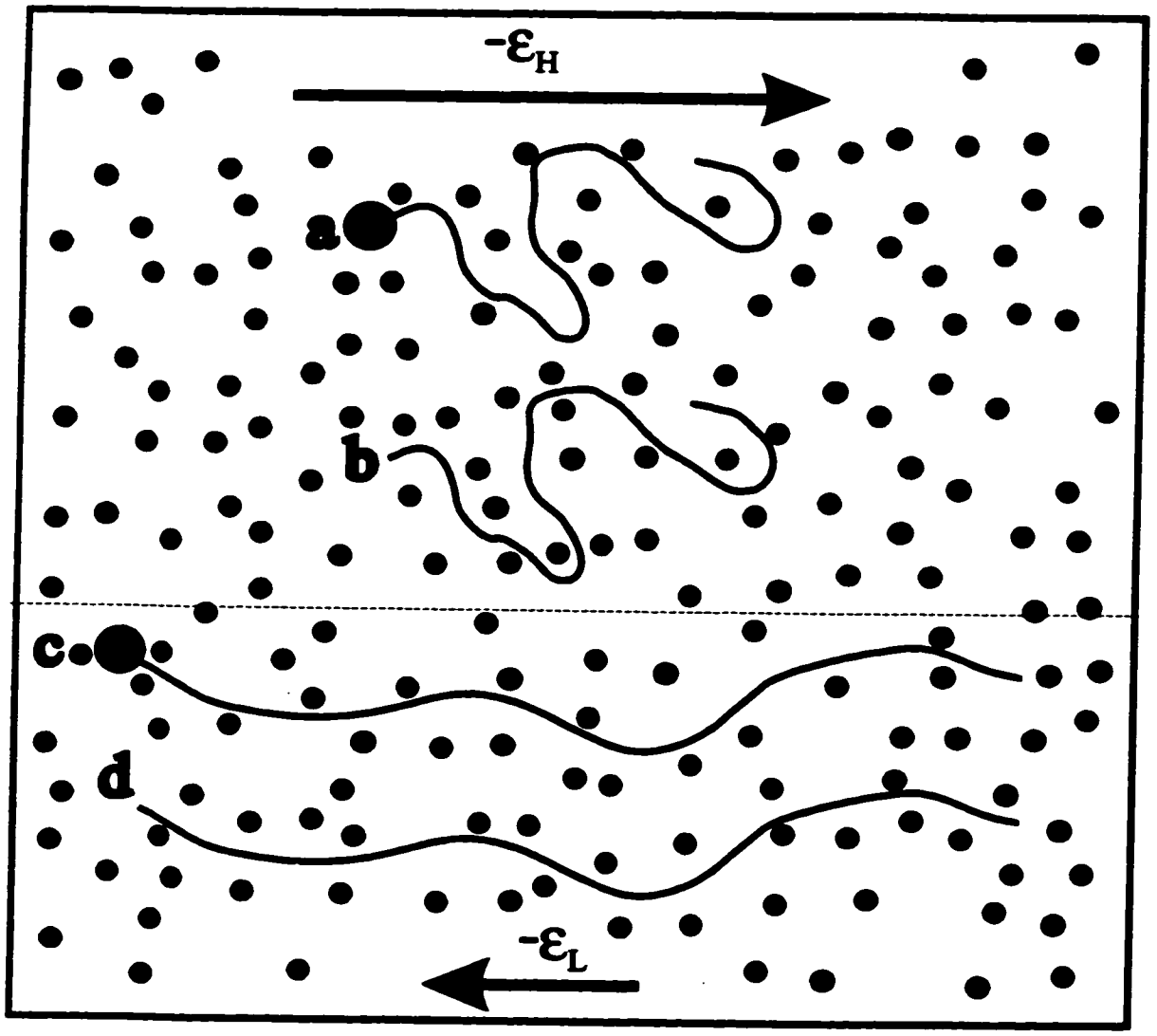


Figure 2: Schematic picture of the gel network (black dots); a and c show a protein-DNA complex (or S-DNA); b and d show a free DNA chain. In a and b the chains are in a low electric field, while the field is high for c and d. Note that the S-DNA molecule in c is trapped.

forces and S-DNA molecules detrap very easily (Figure 2a). Since previous experimental investigations have demonstrated that $\varepsilon_T<\varepsilon_R$ in practice [16, 24], the mobility is field-independent in this weak-trapping regime (Figure 3a). When $\varepsilon>\varepsilon_T$, however, trapping becomes a serious problem and the mobility quickly decreases with electric field because the electric forces actually hinder detrapping by Brownian motion. Finally, when the field exceeds the stopping critical value ε_S, the velocity is negligible because the detrapping time is larger than the duration of the experiment (Figure 3a). In practice, we previously found that the critical fields ε_T and ε_S are very close ([24] and Chapter 5).

7.2.3 ZIFE-like pulses: the zero-frequency limit

Given the DC behaviour of DNA and S-DNA described in Figure 3a, it is relatively easy to understand the dynamics of these molecules in the zero-frequency limit since the transient behavior upon field reversal is then irrelevant. In other words, we can consider that the molecules instantaneously attain their steady-state mobility and diffusion coefficient after the field direction is changed.

Let us now discuss the effect of the type of pulsed field (called ZIFE) shown in Figure 1. The high-field pulse of intensity ε_H is of duration T, while the reverse pulse of intensity $-\varepsilon_L$ is of duration $R_E T$, where $R_E= \varepsilon_H/\varepsilon_L$ is the field ratio ($R_E >1$). Note that this automatically implies that the mean field intensity $\langle E\rangle=0$. The net velocity of a molecule (either DNA or S-DNA) is then given by:

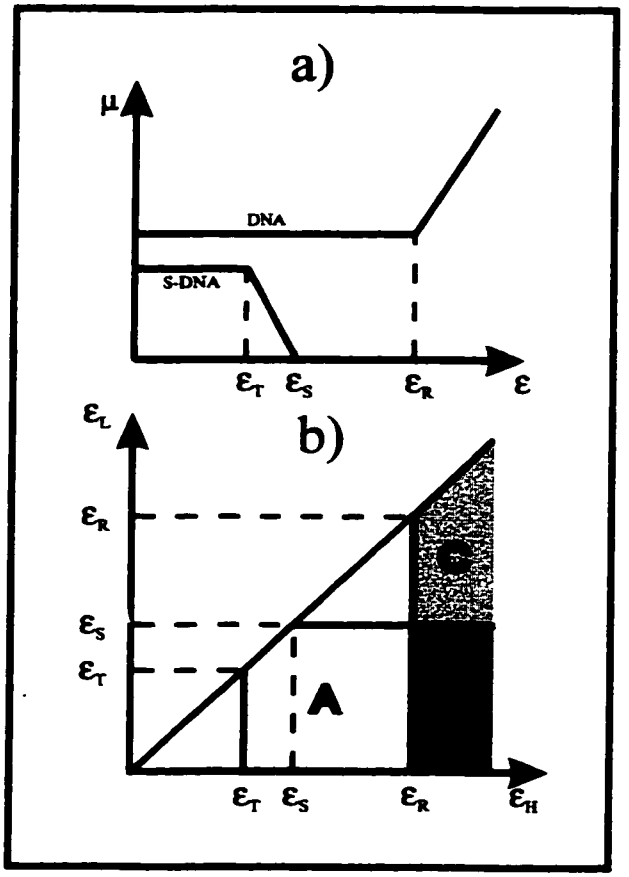


Figure 3 a) A schematic drawing of the electrophoretic mobility μ for DNA and S-DNA molecules as a function of field strength ε. For DNA, the mobility is practically constant for $\varepsilon<\varepsilon_R$ and increases linearly with field beyond this value. For S-DNA, the mobility is constant up to $\varepsilon=\varepsilon_T$, then decreases quickly and becomes negligible for $\varepsilon>\varepsilon_S$. b) Phase diagram for the zero-frequency ZIFE ratchet process. The net velocity of S-DNA is negative in regions A and B (and zero elsewhere), while the velocity of DNA is positive in regions B and C (and zero elsewhere).

$$V = \frac{\varepsilon_H \mu(\varepsilon_H) - R_E \varepsilon_L \mu(\varepsilon_L)}{1+R_E} = \frac{\varepsilon_H}{1+R_E} \times [\mu(\varepsilon_H)-\mu(\varepsilon_L)] \qquad (1)$$

The choice of the field intensities ε_H and ε_L is crucial for the two opposing ZIFE ratchets because of the presence of the differential mobility $\Delta\mu=\mu(\varepsilon_H)-\mu(\varepsilon_L)$ in eq (1).

It is clear that normal DNA will indeed have a net (positive) velocity unless both field intensities are lower than ε_R, in which case $\Delta\mu=0$. Therefore, a sufficient condition for DNA to have a net velocity is $\varepsilon_H>\varepsilon_R$. This process was shown to be useful for the separation of chromosomal DNA [13].

It is also obvious that the net velocity of a S-DNA molecule will be zero if both fields are either lower than ε_T [in which case $\mu(\varepsilon_H)=\mu(\varepsilon_L)$] or higher than ε_S (mobility is zero in both directions). Otherwise, the net velocity will be negative because $\Delta\mu\leq 0$ in the presence of trapping. The most efficient situation is found when $\varepsilon_H \gg \varepsilon_S$ and $\varepsilon_L \leq \varepsilon_T$.

Figure 3b shows a "phase diagram" describing the net zero-frequency ZIFE velocities for DNA and S-DNA molecules having the same number of nucleotides. Both velocities are zero in the non-shaded areas (only the regions below the $\varepsilon_H=\varepsilon_L$ line are considered because $\varepsilon_L \leq \varepsilon_H$). In region A, the velocity of the S-DNA molecule is negative while that of DNA is zero. In region B, the velocity for the S-DNA molecule is negative and the DNA velocity is positive. Finally, the velocity of the S-DNA molecule is zero and the DNA velocity is positive in region C. Obviously, one would have more regimes if the DNA critical field ε_R were smaller than ε_T (see Discussion), but it would still be easy to draw the corresponding phase diagram.

7.2.4 Finite but Low Frequencies

7.2.4.1 Transients upon field reversal: DNA

We must first examine what happens to a DNA molecule immediately after the field direction (and intensity) changes. Immediately after the field is reversed from ε_H to $-\varepsilon_L$, with $\varepsilon_H > \varepsilon_L$, the DNA molecule has an orientation corresponding to the field ε_H (Figure 2d) but must migrate now in a field of intensity ε_L. Consequently, the molecule's velocity is higher at the beginning of the pulse $[V=-\mu(\varepsilon_H)\varepsilon_L]$ and decreases as the molecule loses some of its orientation. Finally, the molecule reaches its steady-state velocity $V=-\mu(\varepsilon_L)\varepsilon_L$. Such field reversals are often characterized by velocity oscillations and overshoots [28], but these effects can be neglected in our study since we assume that the pulse duration is long compared to the time scales over which these effects take place.

If we now reverse the field back from $-\varepsilon_L$ to $+\varepsilon_H$, the molecule first has the (low) orientation corresponding to the field $-\varepsilon_L$ (Figure 2b) and a velocity $V=+\mu(\varepsilon_L)\varepsilon_H$, but eventually orients in the (high) field direction and reaches its steady-state velocity $V=+\mu(\varepsilon_H)\varepsilon_H$. Here again, we will neglect the velocity oscillations.

The dashed lines of Figure 4 show a schematic plot of mean position $|\langle x\rangle|$ *vs* εt for a DNA molecule (note that since $\alpha=0$ for naked DNA, the ordinate-axis is simply $|\langle x\rangle|$). The time-axis has been rescaled with the field intensity in order to be able to directly compare the two parts of the ZIFE pulses (indeed, we note that we have $\varepsilon t=\varepsilon_H T$ at the end of the high-field pulse, while $\varepsilon t=\varepsilon_L R_E T=\varepsilon_H T$ at the end of the low-field pulse). Moreover, the slopes give directly the mobilities on this type of diagram. In this case, we chose $\varepsilon_H > \varepsilon_R > \varepsilon_L$. Curve **a** is for the high field

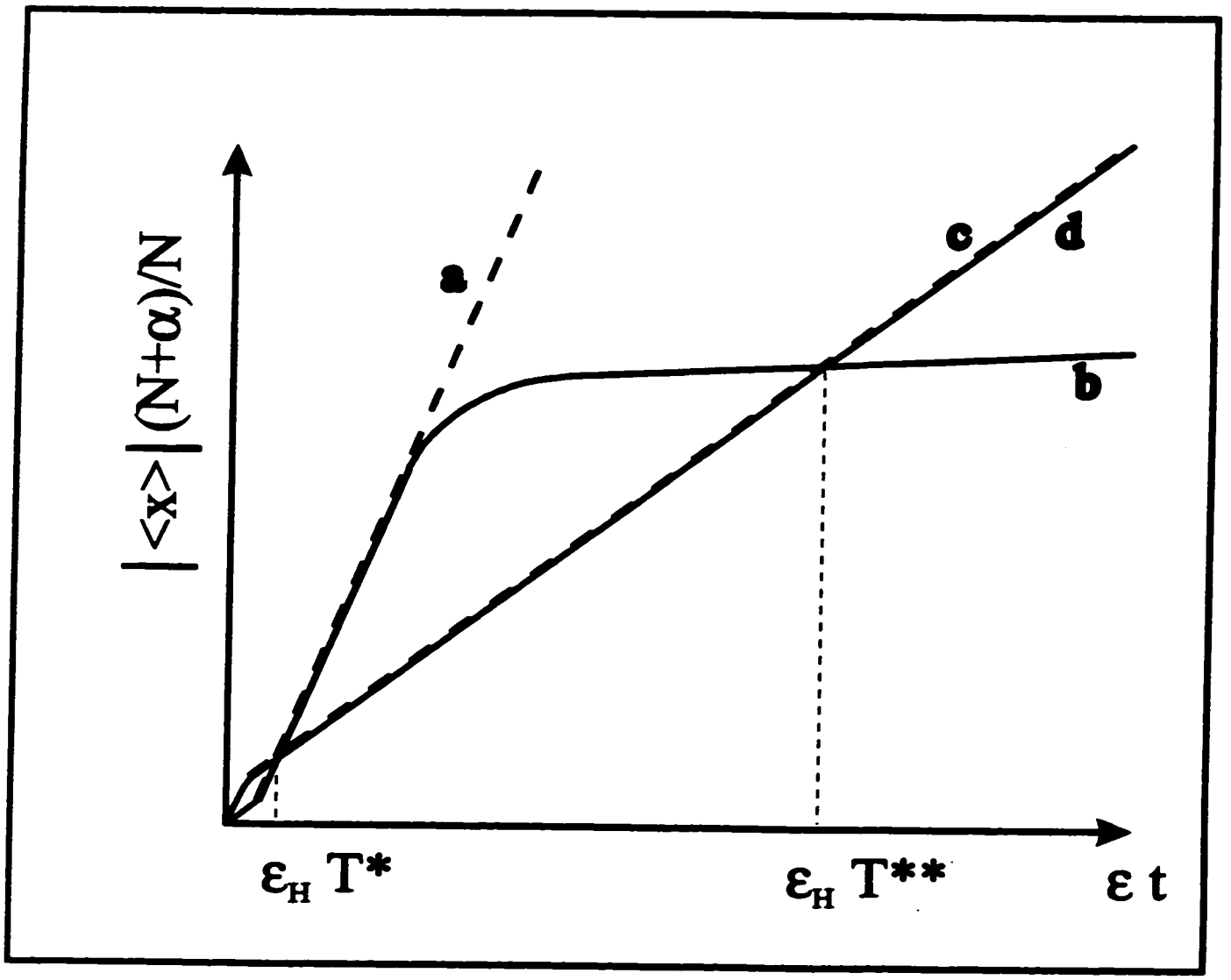


Figure 4: Schematic plot of the (modified) mean position $|\langle x\rangle|(N+\alpha)/N$ *vs* time εt for DNA (dashed line; note that $\alpha=0$ in this case) and S-DNA (solid line), and for low (c and d) and high (a and b) field strengths.

pulse, while curve c is for the low field one. When the pulse duration $T<T^*$, the net velocity is predicted to be negative (i.e., in the direction of ε_L); this prediction has yet to be tested experimentally. In practice, however, times $T>T^*$ are more likely, and the net velocity of the DNA molecule will be positive (i.e., in the direction of the high field pulses).

7.2.4.2 Transients upon field reversal: S-DNA

The behavior of the S-DNA molecule is expected to be similar to that of the DNA molecule for very short times (Figure 4, solid curves). Here we have chosen $\varepsilon_S>\varepsilon_H>\varepsilon_T>\varepsilon_L$ so that the molecule is trapped only in the high field direction. Because the hydrodynamic friction coefficient of the uncharged label slows down the S-DNA molecule (see Chapter 6), we have rescaled the ordinate-axis by the factor $(N+\alpha)/N$ in order to be able to superpose the DNA and S-DNA curves. Note that if the molecules do not orient (low fields), there is no difference between DNA and S-DNA (besides the trivial rescaling of the position axis).

The dynamics of S-DNA in high fields $\varepsilon_S>\varepsilon_H>\varepsilon_T$ is quite different since trapping then dominates. Not long after it has re-oriented in the high field direction, the S-DNA molecule gets trapped and stops moving for a long (but finite) period of time. The most remarkable thing is that lines b (high field) and d (low field) cross twice: first for very short pulses (like naked DNA) because of the re-orientation process described before for DNA, and then for much longer pulse durations $T=T^{**}$. Therefore, we predict that the net ZIFE velocity of the S-DNA molecules will change sign twice as the pulse duration is increased. The situation for $T\ll T^*$ (high frequency pulses) will not be discussed since it is both experimentally irrelevant and theoretically model-dependent.

7.3 Simulation Results

7.3.1 The Biased Reptation Model (BRM)

Although the BRM's [29-31] weaknesses have been well documented [32], this model still represents a useful tool to understand the physics of gel electrophoresis processes, at least qualitatively. As we will see, the model-independent analysis presented in the previous sections agrees with the results of the BRM. We have thus used the simulation method developed previously by Slater and Villeneuve [17] to study TE. Briefly, this modified BRM algorithm is as follows (see [17, 31] for more details).

The polyelectrolytes move by reptating between the gel obstacles. The electric forces bias both the motion inside the reptation tube as well as the mean orientation of the tube itself. Each curvilinear displacement of length ±a (a is the mean pore size) is of duration [33]

$$\tau(h_x) = \tau_B \times \frac{\tanh(\delta(h_x))}{\delta(h_x)}. \qquad (2)$$

where h_x is the end-to-end distance of the DNA molecule in the field direction (x), $\tau_B=a^2/2D_c$ is the Brownian time for the unbiased (field E=0) case where $D_c=k_BT/\xi(N+\alpha)$ is the curvilinear diffusion coefficient of the polymer in its reptation tube, $\delta = \varepsilon h_x/a$ is the bias factor, $\varepsilon = qEa/2k_BT$ is the scaled (dimensionless) electric field intensity, and q and ξ are respectively the charge and the friction coefficient of a primitive reptation segment of length a. These jumps occur with probabilities [33]

$$p_{\pm} = \frac{1}{1 + \exp[\mp 2\delta(h_x)]} , \qquad (3)$$

where the ± refers to the (arbitrarily chosen) direction of the motion inside the tube. A new tube section of length a is created after each "jump" of duration τ. If a tube section is created by a charged end-segment of the polymer chain, its orientation is biased by the field and follows a Boltzmann distribution function $\exp(-\varepsilon \cos(\theta))$, where θ is the angle between this new tube section and the field axis [29, 30]. Tube sections created by the uncharged S-end are randomly oriented.

The BRM has to be modified to take into account the S-DNA steric trapping that occurs when the label faces a small opening (Figure 2). A fraction f«1 of the pore-to-pore passages are thus marked "too-narrow", and any move that tries to make the label move through such passages is rejected (however, the time τ is added to the current time). Trapping occurs when the label is pinned by a narrow passage, and detrapping requires the molecule to move backward over a curvilinear distance L=Na (the contour length of the reptation tube), where N is the number of reptation segments forming (or the number of gel pores occupied by) the molecule. This detrapping process is the only one allowed by reptation [17-23]. Since the bias factor $\delta = \varepsilon h_x/a$ depends on both the field ε and the end-to-end distance h_x, detrapping is very unlikely for high field intensities and long, oriented molecules.

The simulations were carried out on Unix workstations using a Fortran code. The following conditions were used for the simulations: molecular size of the DNA molecules N=20; fraction of small pores, f=0.001; high field intensity ε_H=1.0; low field intensity ε_L=0.04, so that R_E=25. Note

that for this molecule, $\varepsilon_R \approx 0.84$ and $\varepsilon_T \approx \varepsilon_S \approx 0.05$ [17]. For simplicity, we chose $\alpha=0$ so that it would easier to compare the results for DNA and S-DNA (for the streptavidin-DNA complex, α was found to be smaller than unity because the label occupies less than one pore; see Chapter 6). Therefore, the distances are measured in units of the mean pore size a, while the times are measured in units of τ_B. Note that $\tau_B \sim 1/D_c \sim N$ with these units. The high-field pulse ε_H defines the positive direction of migration.

7.3.2 The effect of frequency

Figure 5 shows how the mean (steady-state) velocity of DNA (squares) and S-DNA (circles) varies as a function of the pulse duration T, while Figure 6 shows a similar plot for the net diffusion coefficient. Remarkably, the S-DNA diffusion coefficient has a large maximum around $\log_{10}T \approx 2.8$ but is actually lower than that of DNA for $\log_{10}T \leq 1$ and $\log_{10}T \geq 4$.

The dotted line in Figure 5 represents the DNA velocity calculated using the constant field approximation and Figure 4. To obtain this curve, two simulations in constant field were performed, one at $-\varepsilon_L$ and one at ε_H. In both cases, we first oriented the molecules with respect to the other field to reproduce the conditions observed after a field reversal. The distance migrated during a high field pulse of duration T is $x_H(T)>0$, while the distance migrated during the next low field pulse is $x_L(R_E T)<0$. The total duration of a complete cycle is given by $(1+R_E)T$. The steady state velocity is thus given by:

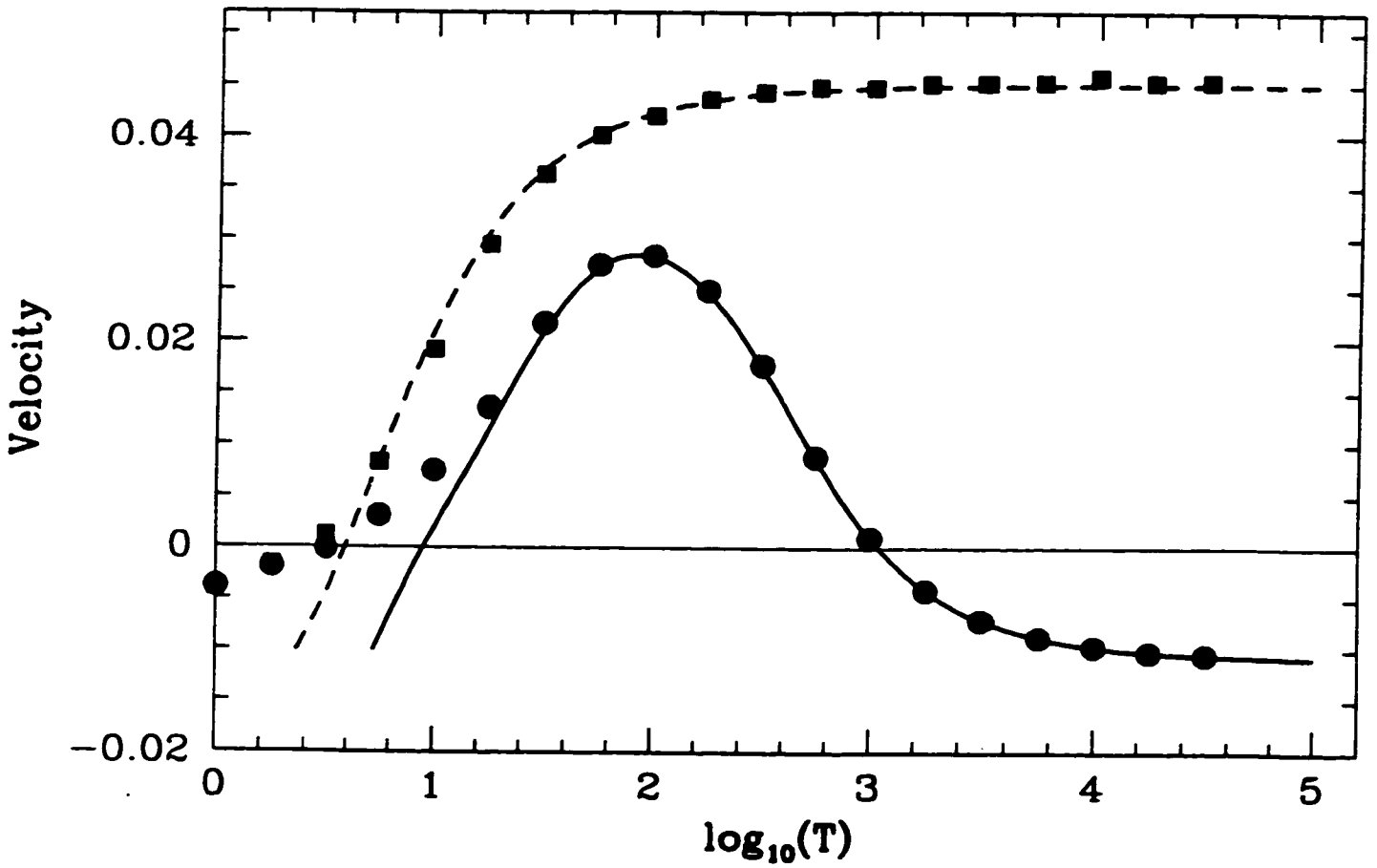


Figure 5: Mean electrophoretic velocity for DNA (■ and dashed line) and S-DNA (● and solid line) as a function of pulse duration T. The simulation parameters used for Figs. 5-10 are described in the text. The data points come from the simulations, while the lines come from a simple theory which is also described in the text.

$$V_{ss}(T) = \frac{x_H(T) + x_L(R_E\ T)}{(1 + R_E)\ T}. \qquad (4)$$

The solid line plotted on Figure 5 represents the same calculation for S-DNA molecules. We note that the constant field approximation method is valid for $\log_{10}T \geq 0.8$ for DNA and for $\log_{10}T \geq 1.5$ for S-DNA. This simply reflects the fact that for pulses that are too short, the molecules do not have time to completely loose the orientation acquired during the previous pulse. Note that the label makes the S-DNA reorientation time longer since this molecule must do a complete flip-flop after every field reversal, while a DNA molecule has perfect head/tail symmetry.

The dotted line in Figure 6 represents the diffusion coefficient estimated using the same approach. The increase of the variance during the high field pulse is $\Delta x^2_H(T)$, while it is equal to $\Delta x^2_H[(1+R_E)T]$ during the low field pulse that follows. The resulting diffusion coefficient is then simply equal to:

$$D_{ss}(T) = \frac{\Delta x_H^2(T) + \Delta x_L^2(R_E\ T)}{2\ (1 + R_E)\ T}. \qquad (5)$$

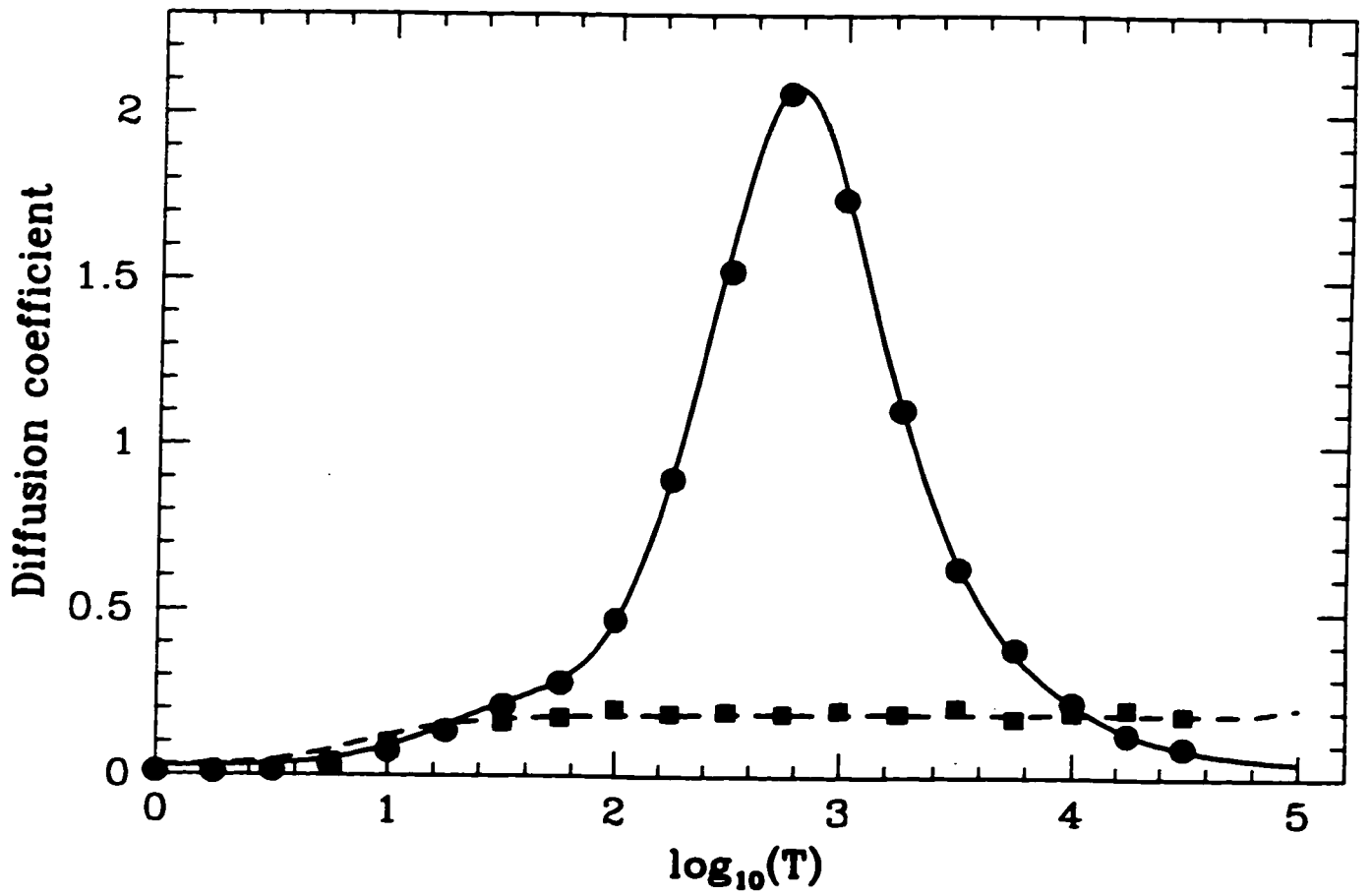


Figure 6: Diffusion coefficient for DNA (■ and dashed line) and S-DNA (● and solid line) as a function of pulse duration T. The data points come from the simulations, while the lines come from the simple theory described in the text.

The solid line represents the same calculation for S-DNA molecules. The two lines provide excellent approximations for $\log_{10}T \geq 1.5$. It should be mentioned that the constant field approximation can be used for any system where it is possible to estimate the initial state (at the beginning of every pulse) of the system.

Figure 7 shows a log-log plot of mean position $|\langle x \rangle|$ *vs* modified time εt during a high field (thick lines) and low field (thin lines) pulse for DNA (dotted lines) and S-DNA (solid lines). As discussed in sections 7.3.1 and 7.3.2, the net velocity of DNA and S-DNA is null when their two corresponding curves are crossing. Here we see that the curves for DNA are crossing once for $\log_{10}T \approx 0.6$ while the curves for S-DNA are crossing twice for $\log_{10}T \approx 1$ and $\log_{10}T \approx 3$. This is consistent with the schematic representation of Figure 4 (which was not log-log).

Figure 8 presents a log-log plot of the variance $|\Delta x^2|$ *vs* time εt during a high-field (thick lines) and low-field (thin lines) pulse for DNA (dotted lines) and S-DNA (solid lines). These curves are very useful to understand the peak observed around $\log_{10}T=2.8$ in Figure 6. During a high field pulse, the variance of the S-DNA band first increases very quickly for times $\log_{10}T<2.8$, and plateaus for $\log_{10}T>3$. The fact that the diffusion coefficient increases for $\log_{10}T<2.8$ is due to the fact that $\Delta x^2 \propto t^{\gamma}$ with $\gamma>1$, in this regime; see Desruisseaux and Slater [18] for a study of the anomalous diffusion properties of S-DNA. The diffusion coefficient then decreases with T for $\log_{10}T>3$ since $\Delta x^2 \propto t^0$ (no increase) when there is complete trapping. The DNA diffusion coefficient is constant for $\log_{10}T>2$ because we then have $\Delta x^2 \propto t$ (normal diffusion) during both the high and low field pulses.

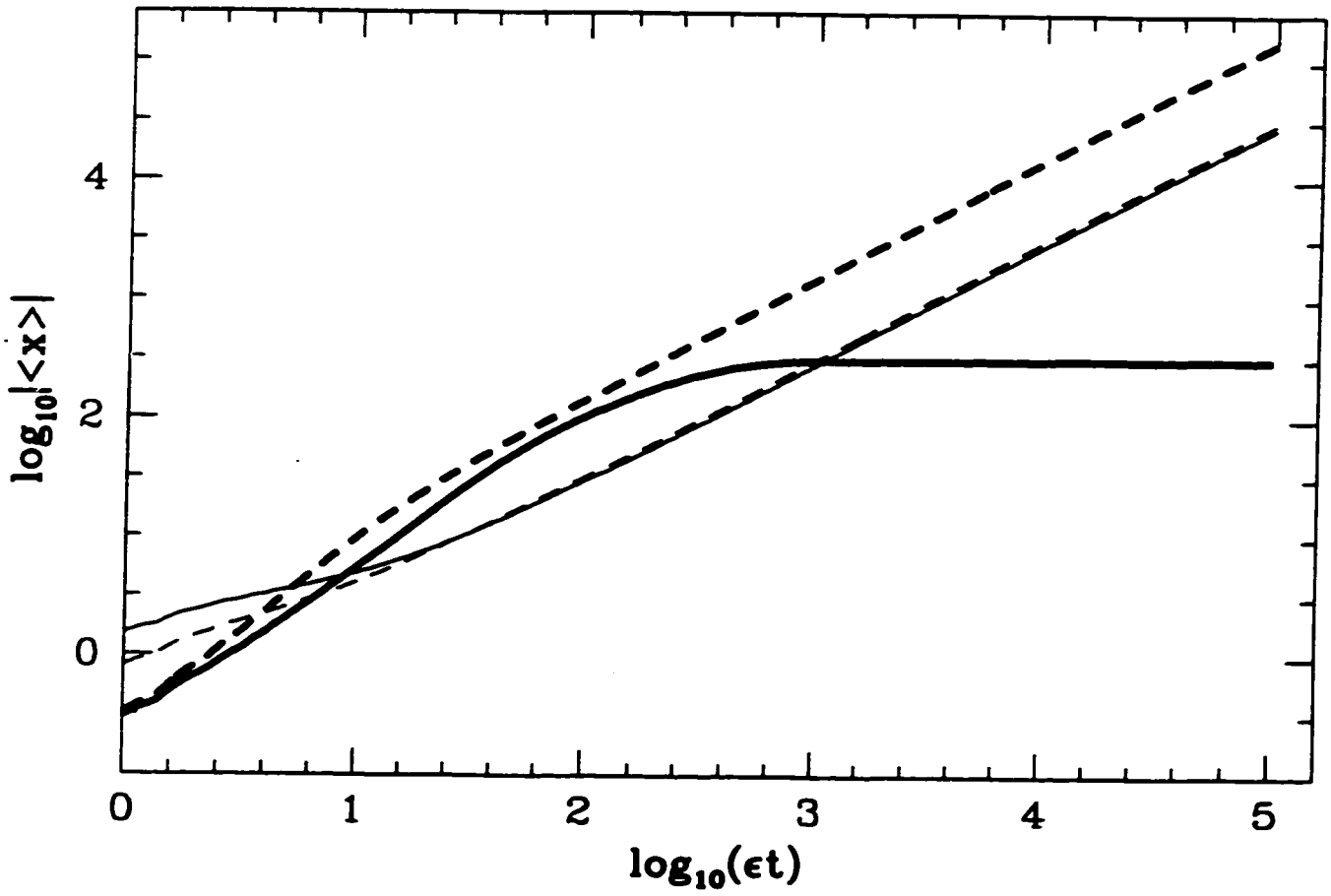


Figure 7:Log-log plot of the mean position $|\langle x\rangle|$ *vs* ϵt for DNA (dashed lines) and S-DNA (solid lines) during the high (ε_H; thick lines) and low (ε_L; thin lines) field pulses.

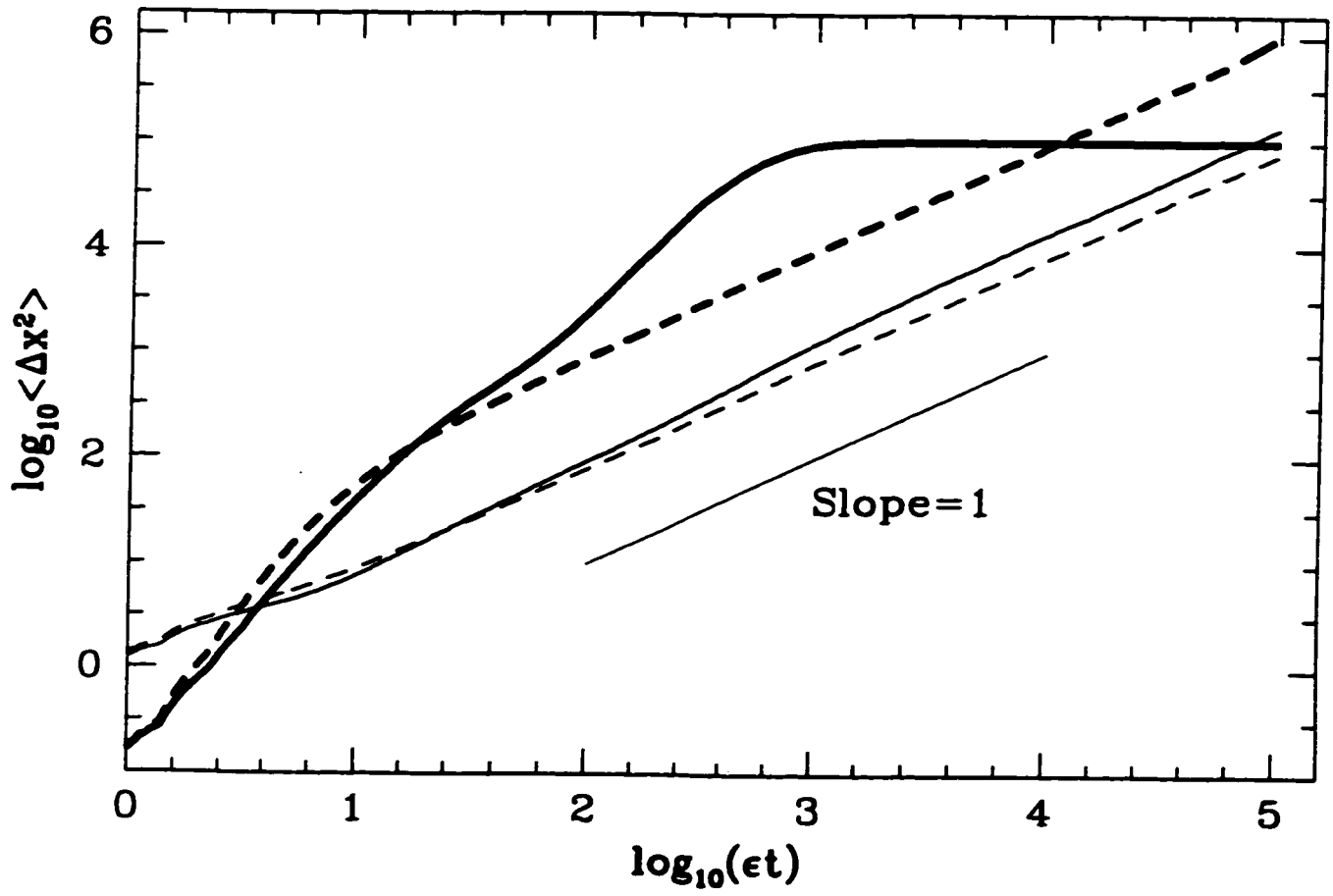


Figure 8: Log-log plot of band variance Δx^2 *vs* εt for DNA (dashed lines) and S-DNA (solid lines) during the high (ε_H; thick lines) and low (ε_L; thin lines) field pulses.

7.3.3 Asymmetry of the DNA and S-DNA bands

The very peculiar asymmetric dynamics of the S-DNA molecules in the system under investigation naturally leads to asymmetric bands. The central-limit theorem, however, implies that the band shape should become Gaussian, hence symmetric, for long enough times. In order to investigate the band asymmetry and its temporal evolution, we also computed the band skew (Sk):

$$Sk = \frac{\langle (x-\langle x\rangle)^3\rangle}{(\langle x^2\rangle - \langle x\rangle^2)^{3/2}}. \qquad (6)$$

One thus has Sk=0 for a symmetric band, Sk<0 for a band with a back-tail and Sk>0 for one with a front tail. Figure 9 shows Sk *vs* time for both DNA (dashed lines) and S-DNA (solid lines), and for pulse durations $\log_{10}T$=1, 2, 3, and 4. Note that the simulation conditions are identical to those used for Figures 5 and 6, and that the initial band was a delta-peak. The direction of the high field (ε_H) defines the positive (+x) direction of migration, and the pulse sequence starts with a high-field pulse.

The short time evolution of the S-DNA band skew is a function of the initial conditions (i.e., whether we start with a high-field or low-field pulse, and whether the molecules are initially relaxed or oriented). However, many characteristic results are independent of the initial conditions. For very long pulses (e.g., $\log_{10}T$=4), the skew first goes up for t<3 (a few molecules start moving), then down (the rest of the molecules complete their very first move; note that the skew is then negative!), and then up again (some molecules are left behind in very deep traps), before it saturates (all

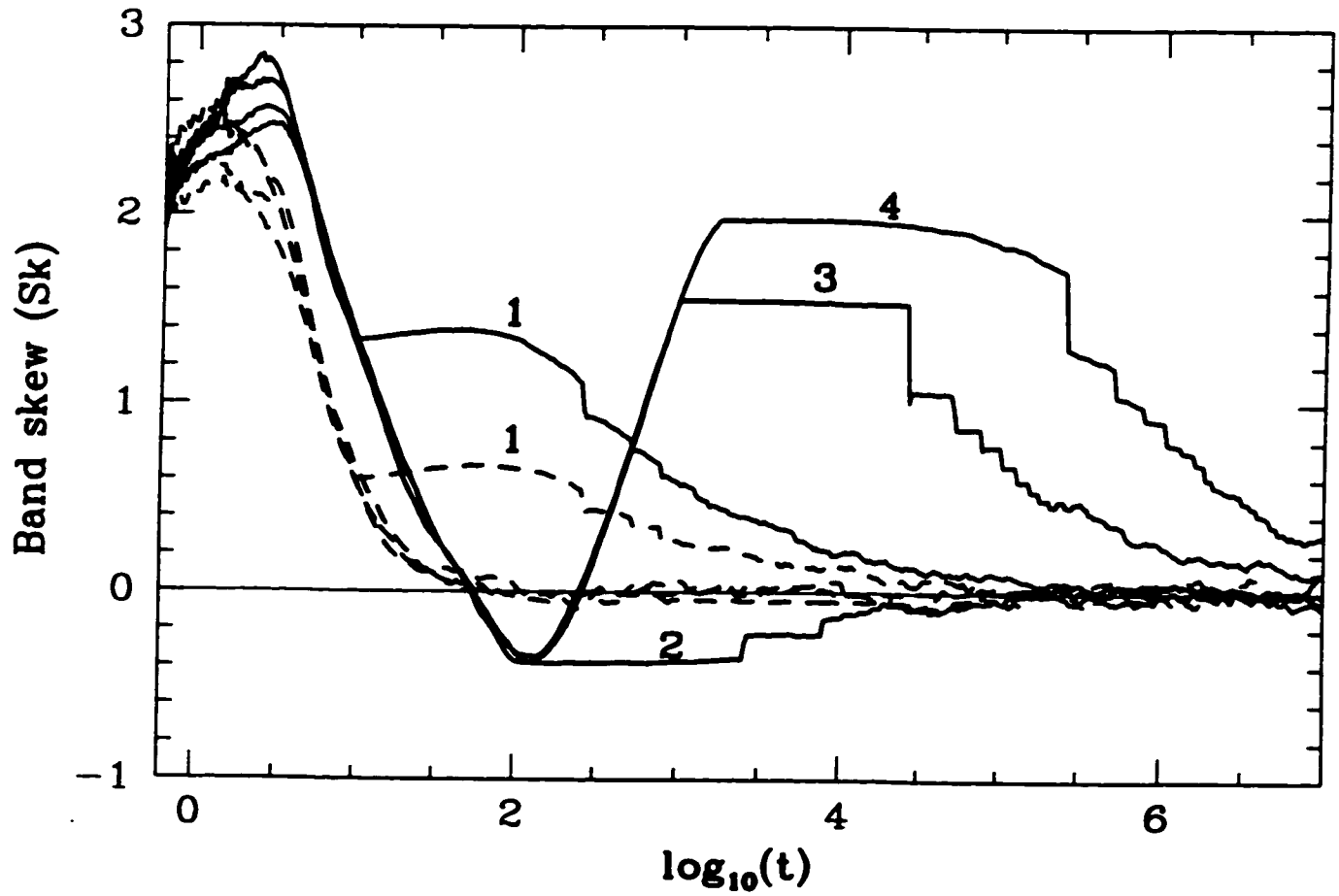


Figure 9: Band skew (Sk) for DNA (dashed lines) and S-DNA (solid lines) as a function of time t for pulse durations $\log_{10}T$=1, 2, 3, and 4, as indicated. The dashed curves (DNA) are almost identical for $\log_{10}T$=2, 3, and 4.

molecules have fallen in deep traps). The field reversal at $\log_{10}t=4$ frees the molecules and the skew slowly decreases as $t^{-1/2}$ for very long times (more than 100 pulses), as expected for a simple directed walk problem (see Appendix A). The situation is quite similar for $\log_{10}T=3$. In comparison, the skew goes down to zero very quickly for DNA (the initial peak is due to the fact that we have a wide distribution of tube renewal times in the presence of a field, a distribution that is related to the distribution of end-to-end distance h_x). For $\log_{10}T=1$, on the other hand, the evolution of the skew is almost identical for S-DNA and DNA because the pulse duration is shorter than both the mean time between traps and the mean tube renewal time.

The situation for $\log_{10}T=2$ is qualitatively different. Because the pulse duration T is now comparable to the mean time between traps, the skew remains negative (but small) after the initial decay. This is an interesting phenomenon, and it provides a simple experimental method to estimate the mean trapping time.

7.3.4 The resolution between DNA and S-DNA

There are many ways to mathematically express the resolution between two bands. One popular definition of resolution (R) is given by [34]:

$$R = \frac{\left| \langle x_1 \rangle - \langle x_2 \rangle \right|}{2\left(\sigma_1 + \sigma_2 \right)}, \qquad (7)$$

where $\langle x_1 \rangle$ and $\langle x_2 \rangle$ are the centre of mass positions of bands 1 and 2, respectively, and σ_1 and σ_2 are

the standard deviations of these bands. The resolution between two bands normally grows like $R \propto t^{1/2}$. In such cases, $R/t^{1/2}$ is a fundamental parameter which tells us how fast two bands are being resolved from each other (large values of $R/t^{1/2}$ means that less time will be required to resolve the bands). Figure 10 shows the resolution ratio $R/t^{1/2}$ between a naked DNA molecule and the corresponding S-DNA molecule *vs* pulse duration T. The data points come from our Monte Carlo simulations while the solid line was obtained using the constant field approximation discussed previously. Clearly, long pulse durations are preferable. Again, the critical value $\log_{10}T \approx 2$ shows up as a minimum.

7.4 Discussion

This Chapter establishes the general theoretical framework that is necessary to understand and explain the results of the Griess and Serwer ratchet electrophoretic separation process for S-DNA and DNA molecules. For the sake of simplicity, we have focused our study on the following conditions: a) the label (S) is attached at the end of the DNA molecule; b) the reptation orientation field ε_R is higher than the trapping field ε_T ($\approx \varepsilon_S$); c) the pulse frequency is small compared to the frequency of the intra-molecular DNA stretching and relaxation modes. Clearly, a model that would cover all possible cases would lead to an essentially infinite number of separation regimes, and hence would be useless at this stage. The theoretical approach described in this article will easily be adapted to other situations once more detailed experimental data becomes available.

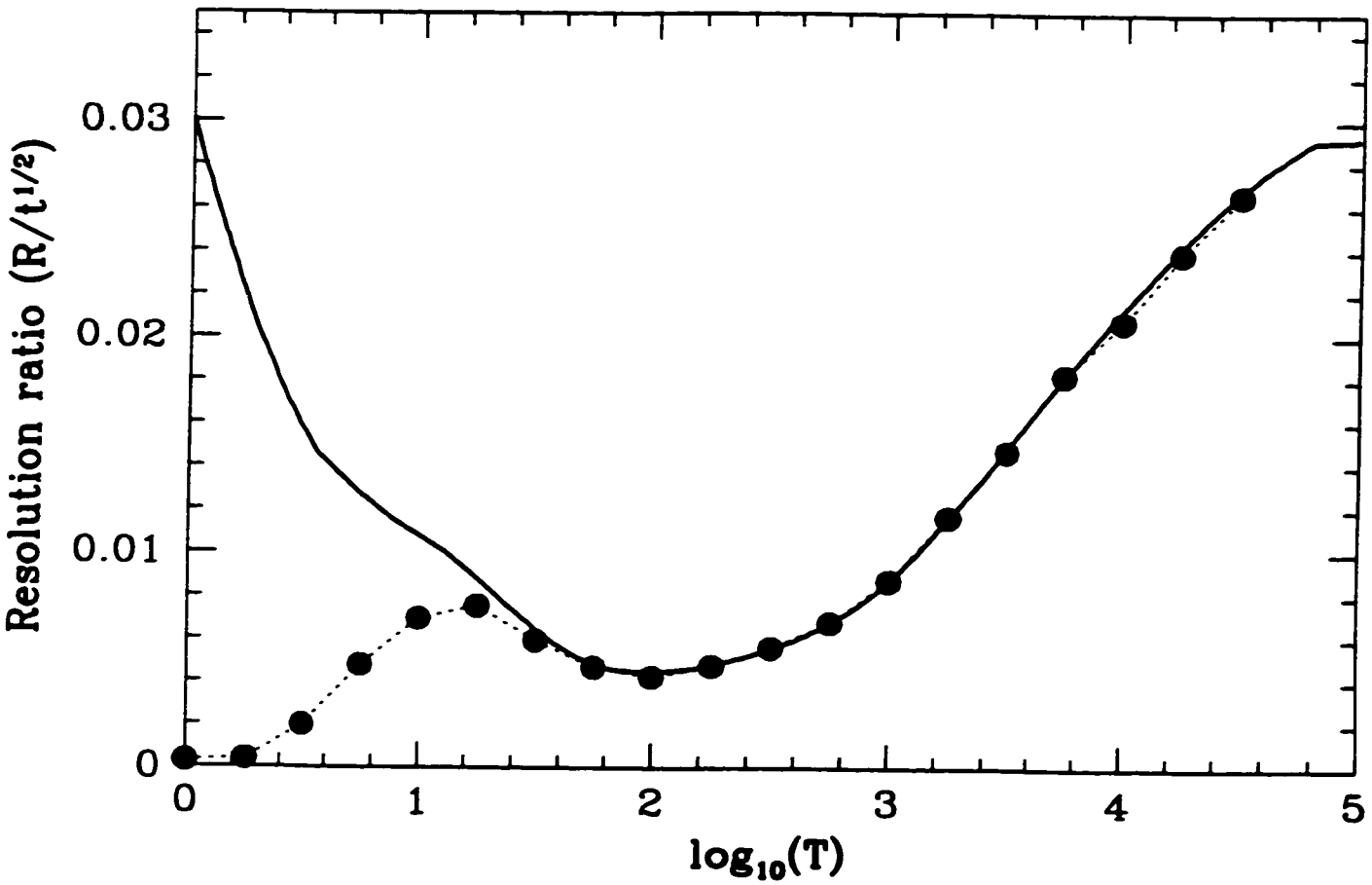


Figure10: Resolution ratio $R/t^{1/2}$ *vs* pulse duration T for the ratchet separation between a DNA band and the corresponding S-DNA band. The solid line comes from the simple theory described in the text. The dotted line connects the data points obtained from the Monte Carlo simulations.

We have shown that such ratchets should be using very low frequencies in order to optimize resolution. However, the pulse duration can also be used as a spectroscopic tool to estimate the microscopic times such as the mean trapping and detrapping times. For instance, the mean S-DNA velocity is predicted to change sign (and its diffusion coefficient to reach a maximum value) for intermediate pulse durations that correspond closely to these characteristic times. We have also demonstrated that this trapping mechanism should lead to asymmetric bands, and that the latter should slowly become symmetric over hundreds of pulses.

The field-dependence of the gel electrophoretic mobility of charged spherical particles is qualitatively different from that of DNA because particles do not "orient" in the field (i.e., their mobility is essentially constant in absence of trapping). In the presence of trapping, however, their behavior is somewhat similar to that of S-DNA, i.e. their mobility vanishes beyond a certain critical field intensity [14, 15, 35]. With zero-frequency ZIFE-type pulses, particles can have either a zero net velocity (if they are not trapped in either direction), or a net negative velocity (if they are trapped in the positive - high field - direction). At high frequency, the net velocity of trapped particles should actually be zero when the pulse duration T gets smaller than the time it takes to fall into a dead-end (trap). The only way to make particles move in opposite directions would thus be to slightly bias the pulsed field, as described previously for a different system [8].

The first part of the chapter described general model-independent principles that apply to S-DNA ratchets. The second part used a specific electrophoresis model (the BRM) for Monte Carlo computer simulation purposes. The BRM has largely been replaced by the BRF (biased reptation model with fluctuations) over the last few years. The main difference between the two is that the

latter correctly predicts that the mobility of DNA should increase linearly with field intensity when $\varepsilon > \varepsilon_R$ in most experimental situations, while the first predicted a quadratic increase [32, 36]. The qualitative predictions of the simulation results are thus not affected by this difference. Neither model would correctly describe the dynamics at very high frequencies.

Perhaps the most important problem not treated in our study is that of DNA/S-DNA mixtures. For example, can this ratchet process separate many DNA and S-DNA molecules of different lengths simultaneously? Can this ratchet separate S-DNA molecules of identical lengths but different anchoring points for the bulky S-label? Can it separate S-DNA molecules with more than one label?

The Griess and Serwer ratchet process is very flexible and can be modified in very many ways. For example, the pulse shape (which does not have to be square), frequency and amplitude (both of which can be changed during the separation) offer tunable parameters. Moreover, one can add a small DC component to the ZIFE pulses in order to bias the separation process towards a given direction (this can easily increase the efficiency). More interesting, however, would be two-dimensional schemes where to different separation conditions are used in two orthogonal directions.

7.5 Appendix

The motion of our polylectrolytes during one complete pulse cycle is similar to that of a particle that makes a jump over a distance $x \geq 0$ (the displacement during the cycle) every time unit (the cycle duration). Let $P(x)dx$ be the probability distribution function for these jumps. The mean velocity of the particle is then simply given by $V=\langle x \rangle$. It is easier to calculate the moments of the distribution of particles after t jumps if we use the relative position $s= x-\langle x \rangle$, such that the mean position of the particles is always $s=0$. The problem is then reduced to evaluating the variance $\langle (x-\langle x \rangle)^2 \rangle = \langle s^2 \rangle$ and third moment $\langle (x-\langle x \rangle)^3 \rangle = \langle s^3 \rangle$. Indeed, after t independent (i.e., uncorrelated) jumps, the mean position is given by:

$$\langle S(t) \rangle = \left\langle \sum_{i=1}^{t} s_i \right\rangle = t \langle s \rangle = 0, \qquad \text{(A1)}$$

where the moments of s are defined by

$$\langle s^j \rangle = \int_{-\langle x \rangle}^{\infty} P(s)\, s^j\, ds. \qquad \text{(A2)}$$

The second moment of the distribution of final positions S(t) is

$$\langle S^2(t) \rangle = \left\langle \left(\sum_{i=1}^{t} s_i \right)^2 \right\rangle = \left\langle \sum_{i=1}^{t} s_i^2 + \sum_{i}^{t} \sum_{j \neq i}^{t} s_i\, s_j \right\rangle. \qquad \text{(A3)}$$

Since the jumps are uncorrelated, the last term is zero and we obtain

$$\langle S^2(t) \rangle = t \langle s^2 \rangle \sim t. \qquad \text{(A4)}$$

Similarly,

$$\langle S^3(t) \rangle = \left\langle \left(\sum_{i=1}^{t} s_i \right)^3 \right\rangle = \left\langle \sum_{i=1}^{t} s_i^3 + 3\sum_{i \neq j} s_i^2 s_j + \sum_{i \neq j \neq k} s_i s_j s_k \right\rangle = t \langle s^3 \rangle \sim t. \qquad \text{(A5)}$$

Therefore, the skew $\langle (x-\langle x \rangle)^3 \rangle / \langle (x-\langle x \rangle)^2 \rangle^{3/2} = \langle S^3 \rangle / \langle S^2 \rangle^{3/2} = \Gamma / t^{1/2}$, where $\Gamma = \langle s^3 \rangle / \langle s^2 \rangle^{3/2}$ is a time-independent property of the probability distribution function P(x)dx. Note that the skew is zero for a symmetric distribution function, as it should. For an asymmetric distribution function, the skew will thus decrease as $1/t^{1/2}$ and the band will become symmetric for long enough times.

7.6 References

[1] Griess, G. A.; Serwer, P.; *Biophysical J.*, **74(2)**, A71, 1998.

[2] Magnasco, M. O.; *Phys. Rev. Lett.*, **71**, 1477, 1993.

[3] Chialvo, D. R.; Millonas, M. M.; *Lett., A* **209**, 26, 1995.

[4] Astumian, R. D.; *Science*, **276**, 917, 1997.

[5] Duke, T. A. J.; Holy, T. E.; Leibler, S.; *Phys. Rev. Lett.*, **74**, 330, 1995.

[6] Jülicher, F.; Ajdari, A.; Prost, J.; *Rev. Mod. Physics.*, **69**, 1269, 1997.

[7] Rousselet, J.; Salomé, L.; Ajdari, A.; Prost, J.; *Nature*, **370**, 446, 1994.

[8] Slater, G. W., Guo, H. L.; Nixon, G. I.; *Phys. Rev. Lett.*, **78**, 1170, 1997.

[9] Chacron, M. J.;Slater, G. W.; *Phys. Rev. E*, **56**, 3446, 1997.

[10] Andrews, A. T.; Electrophoresis: Theory, Techniques and Biochemical and Clinical Applications. Clarenton, Oxford, 1986.

[11] Guttman, A.; *Electrophoresis*, **17**, 1333, 1996.

[12] Dunn, M. J.; Corbett, J. M.; *Methods Enzymol.*, **271**, 177, 1996.

[13] Turmel, C.; Brassard, E.; Forsyth, R.; Hood, K.; Slater, G. W.; Noolandi, J.; 1990. *In* Electrophoresis of large DNA Molecules - Theory and Applications. E. Lai and B. W. Birren, editors. Cold Spring Harbor Laboratory Press, Plainview, New York, 101-131, 1990.

[14] Griess, A. G.; Serwer, P.; *Biopolymers*, **29**, 1863, 1990.

[15] Serwer, P., Griess, A. G.; *Analusis Magazine*, **21**, M16, 1993.

[16] Ulanovsky, L.; Drouin, G.; Gilbert, W.; *Nature*, **343**, 190, 1990.

[17] Slater, G. W.; Villeneuve, C.; *J. Polymer Sci. B*, **30**, 1451, 1992.

[18] Desruisseaux, C.; Slater, G. W.; *Phys. Rev. E*, **49**, 5885, 1994.

[19] Défontaines, A. D.; Viovy, J. L.; Theoretical Model of Trapping Electrophoresis. *In*

Proceedings of the First International Conference on Electrophoresis, Supercomputing and the Human Genome. World Scientific, Singapore. 286-313, 1991.

[20] Défontaines, A. D.; Viovy, J. L.; *Electrophoresis*, **14**, 8, 1993.

[21] Défontaines, A. D.; Viovy, J. L.;. *Electrophoresis*, **15**, 111, 1994.

[22] Slater, G. W.; Desruisseaux, C.; Villeneuve, C.; Guo, H. L.; Drouin, G.; *Electrophoresis*, **16**, 704, 1995.

[23] Desruisseaux, C.; Slater, G. W.; *Electrophoresis*, **17**, 623, 1996.

[24] Desruisseaux, C.; Slater, G. W.; Drouin, G.; *Macromolecules*, **31**, 6499, 1998.

[25] Serwer, P., Hayes, S. J.; Moreno, E. T.; Park, C. Y.; *Biochemistry*, **31**, 8397, 1992.

[26] Heller, C.; Duke, T. A. J.; Viovy, J. L.; *Biopolymers*, **34**, 249, 1994.

[27] Holzwarth, G.; Mckee, C. B.; Steiger, S.; Crater, G.; *Nucleic Acids Research*, **15**, 10031, 1987

[28] Sabanayagam, C. R.; Holzwarth, G.; *Electrophoresis*, **17**, 1052, 1996.

[29] Lumpkin, O. J.; Déjardin, P., Zimm, B. H.; *Biopolymers*, **24**, 1573, 1985.

[30] Slater, G. W.; Noolandi, J.; *Biopolymers*, **25**, 431, 1986.

[31] Slater, G. W.; *Electrophoresis*, **14**, 1, 1993.

[32] Duke, T. A. J.; Viovy, J. L.; Semenov, A. N.; *Biopolymers*, **34**, 239, 1994.

[33] Slater, G. W.; Rousseau, J.; Noolandi, J.; *Biopolymers*, **26**, 863, 1987.

[34] Giddings, J. C., Unified Separation Science. John Wiley & Sons, New York (pages 90 and 102), 1991.

[35] To, K.-Y; Boyde, T. R. C.; *Electrophoresis*, **14**, 597, 1993.

[36] Semenov, A. N.; Duke, T. A. J.; Viovy, J. L.; *Phys. Rev. E*, **51**, 1520, 1995.

Chapter 8

Conclusion

Many aspects of the molecular dynamics of the streptavidin-DNA complex were discussed in this thesis. The dynamical properties of this molecule were studied not only in gels (like in my earlier theoretical Master thesis), but also in free-solutions and in dilute polymer solutions. We were able to develop a theoretical model that explained the situation in all of the above mentioned media and we also obtained experimental confirmation of these new theoretical ideas. The new theoretical approach is based on the theories of Long et al. for the hydrodynamics of polyelectrolytes and polyampholytes during electrophoresis. This thesis can be seen, at least partly, as a test of the fundamental theories of this group.

The free-solution properties of naked ssDNA were examined in Chapter 2. When the persistence length is larger than the Debye length ($p>\lambda_D$), we observed that the free mobility of DNA decreases with increasing ionic strength like $\mu_0 \sim \ln(I^{-1/2})$. We were also able to see the transition from this regime to the regime where the free mobility is $\mu_0 \sim \lambda_D/p$. In order to interpret our experimental results, we used the blob picture proposed by Long et al.. Since the latter does

not require any assumption about the "effective charge" of DNA, we conclude that the concept of effective charge is not relevant in electrophoresis (as predicted by this group).

The electrophoretic properties of S-DNA were also studied in free-solution. This new separation process, commonly referred to as ELFSE, was suggested in 1994 by Mayer et al. and the theoretical approach adopted in the first paper was insufficient to provide a good picture of the situation. In order to understand ELFSE, one cannot neglect the hydrodynamics of the S-DNA complex. The new theory of ELFSE predicts that the conformation adopted by a S-DNA molecule is the same as if there were no label (the DNA adopts a random coil conformation). The implication of this is that there should be an ionic strength dependence of the relative friction coefficient of the label. This dependence is simply due to the fact that the flexibility of DNA is a function of its ionic environment. Our experimental results confirmed this prediction; furthermore, we were able to obtain a persistence length vs. ionic strength curve that was in good agreement with the existing theory. For all practical purposes, we used the label as if it were a "persistence length probe"!

A theoretical model for the migration of end-labeled DNA molecules in dilute polymer solutions was developed in Chapter 3. This model is simply a hybrid between our model for the migration of S-DNA in free-solution and the one developed by Hubert et al. for the migration of naked DNA in dilute polymer solutions. Our new theoretical model leads to a surprisingly simple relation between the velocities of the DNA and S-DNA molecules:

$$V_{S\text{-}DNA}(M) \approx V_{DNA}(M + \alpha) \times \frac{M}{M + \alpha}$$

By performing experiments at different polymer concentrations, we were able to observe the transition from free-solution to "gel" electrophoresis and confirm the validity of our "universal equation". For instance, we saw that the large molecules moved faster than the smaller ones at low polymer concentrations, while we observed the opposite situation at high polymer concentrations. At intermediate polymer concentrations, the small molecules are in an ELFSE-like separation regime, while the larger molecules are in a gel electrophoresis-like regime. This fascinating situation (which leads to band inversion) was correctly predicted by our theoretical model.

Chapter 4 discussed the electric field gradient that appears near the edges of denaturing polyacrylamide gels. We showed that these gradients are caused by the difference in viscosity between the inside and the outside of the gel (due to the presence of urea). It was important to study the electric field gradients since these gradients make it difficult to obtain useful and reproducible velocity and diffusion data for system optimization. The gradients can also lead to polyacrylamide gel deterioration (since the electric field can be very high in the gradient region). We proposed many ways to reduce the electric field gradients in order to test existing electrophoresis theories. The most practical way of minimizing the high electric field gradient is, in our opinion, to perform an inverted prerun because it allows for efficient elimination of the gradient without changing the chemistry of the gel. It is very important to understand the gradients in order to test trapping electrophoresis (Chapters 5 and 6) since the electric field gradient strongly interferes with the phenomenon under study.

A theoretical study on the dynamical properties of S-DNA fragments during gel electrophoresis was carried out during my M. Sc. program. One of the strongest predictions of my M. Sc. thesis was that the transition between weak and strong trapping should occur for molecules larger than M^* where $M^* \sim E^{-2/3}$. This prediction was verified experimentally in Chapter 5. Like stated previously, it is not easy to obtain reproducible results on TE without a good understanding of the electric field gradients. However, the existence of the electric field gradient allowed us to observe some band focussing (and antifocussing) on our autoradiogram, in agreement with our understanding of these processes. We also predicted that TE could be transformed into an interesting-isoelectric focussing-like technique.

The modeling of the friction coefficient of the streptavidin label, first done for free-solution electrophoresis (Chapter 2) and dilute polymer solutions (Chapter 3), was generalized to gel electrophoresis in Chapter 6. A blob picture of the streptavidin-DNA molecules, where each blob has the size of the pore, was used to understand the properties of our hybrid molecules in gels. Interestingly enough, the definition of α remains the same. Moreover, the relation between the DNA and S-DNA velocities is exactly the same in the three models even though the separation mechanisms are very different! The observed value of $\alpha \approx 32 \pm 3$ is in excellent agreement with the value obtained in TAPS buffer having the same ionic strength (i.e., $I \approx 0.045$ mol/L). This is a very satisfying result indeed.

A study of the gel trapping mechanism with varying concentrations of polyacrylamide was also performed in order to test our understanding of the nature of the trapping process. We observed that the scaled molecular size decreases like $N^* \sim \epsilon^{-0.35}$ when the scaled field ϵ is reduced

by changing the concentration of polyacrylamide instead of varying the electric field E (in which case $N^* \sim \epsilon^{-2/3}$). This difference can be explained by the fact that the fraction f (of pores having a size smaller than that of the label) increases when we increase the concentration of polyacrylamide.

A theoretical study of the DNA/S-DNA ratchet system, first studied experimentally by Griess and Serwer, was presented in Chapter 7. We have shown the existence of several separation regimes depending on the frequency and the intensity of the electric field pulses. We have further shown that the best separation is obtained at very low frequencies. The pulse duration can also be used as a spectroscopic tool to estimate the microscopic times (mean trapping and detrapping times). We also predicted that transient band asymmetry should slowly disappear over several hundred pulses. Most interestingly, our theory is very powerful in the sense that it is model-independent: the general principles of this study should be universal.

Now that the concept of the friction coefficient for trapping electrophoresis and for ELFSE separations has been clarified, our group concentrates its efforts in order to optimize the separation techniques. This is especially true for ELFSE since the experimental results obtained by Dr. Hongji Ren are quite promising. We are not very far from the point where DNA sequencing using ELFSE will become a reality. In order to be useful, users generally want a new sequencing technique to be able to sequence at least 500 bases per run in less than 2 hours. The success of ELFSE depends on the capacity to engineer larger labels. The theories developed in this thesis are crucial since they will guide the design of the ideal labels. Our group has already

started an experimental study of the diffusion coefficient of end-labeled DNA during electrophoresis and some "applied" theoretical calculations are being developed.

"So at first I did not know what it wanted. But in the end I understood this language. I understood it, I understand it, all wrong perhaps. That is not what matters. It told me to write the report. Does this mean I am freer than I was? I do not know. I shall learn."

Samuel Beckett

Molloy