

Florida State University Libraries

Electronic Theses, Treatises and Dissertations

The Graduate School

2002

DNA Electrophoresis in Agarose Gels: A New Mobility vs. DNA Length Dependence

Afshin Beheshti



THE FLORIDA STATE UNIVERSITY
COLLEGE OF ARTS AND SCIENCES

DNA ELECTROPHORESIS IN AGAROSE GELS: A NEW MOBILITY
VS. DNA LENGTH DEPENDENCE

By

AFSHIN BEHESHTI

A Dissertation submitted to the
Department of Physics
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Degree Awarded:
Fall Semester, 2002

The members of the Committee approve the dissertation of Afshin Beheshti defended on October 23, 2002.

David H. Van Winkle
Professor Directing Dissertation

Randolph L. Rill
Outside Committee Member

Jianming Cao
Committee Member

Nicholas E. Bonesteel
Committee Member

Mark A. Riley
Committee Member

Approved:

The Office of Graduate Studies has verified and approved the above named committee members

This dissertation is dedicated to my parents for there support and sacrifice

ACKNOWLEDGEMENTS

I would like to first acknowledge my advisor, Dr. David Van Winkle, and virtually my second advisor, Dr. Randolph Rill, for their support and advice throughout my graduate years. They have taught me valuable skills which will help throughout my scientific career. Of course none of this would have been possible without my parents. After all they are the ones who brought me into this world. Since the time we both entered graduate school, Dr. Mukundan Chakrapani has provided helpful suggestions on this work. Dr. Bruce Locke (Department of Chemical Engineering, Florida State University) is acknowledged for his early contributions and helpful suggestions. I would like to also thank Dr. Gary Slater (Department of Physics, University of Ottawa) for preprints, interest, and useful discussions. The MARTECH staff has always been very helpful and provided for all my extraneous needs for this research. For example Joe Ryan has solved all my computer problems and Ken Ford made me an expert with Adobe Photoshop and Illustrator. I can't forget Dan Baxter and the physics machine shop which has accommodated for all my experimental needs. They have created some wonderful apparatuses from my drawings and diagrams. Last but not least I would like to thank my graduate committee: Dr. Mark Riley, Dr. Jianming Cao, Dr. Nick Bonesteel (for joining the committee the last minute), and Dr. Per Arne Rikvold (for being on the committee in the beginning). The research has been supported by the State of Florida - Center for Materials Research and Technology (MARTECH), NSF Research Training Grant PBI/9602233, and NSF grant BES 9521381.

TABLE OF CONTENTS

List of Tables	vii
List of Figures	viii
Abstract	xii
1. INTRODUCTION	1
1.1 DNA	3
1.2 Agarose Gels	5
1.3 The Research	8
1.4 Summary	12
2. THEORY	13
2.1 Object in an Ionic Solution	13
2.2 Polymer Chain Conformation	17
2.2.1 Freely Jointed Chain	17
2.2.2 Persistence Length	19
2.3 Free Solution Electrophoresis	22
2.4 Gel Electrophoresis	25
2.4.1 Ogston Model	25
2.4.2 Extensions of the Ogston Model	27
2.4.3 Reptation Models	31
2.4.4 Entropic Trapping	37
2.4.5 Other Models	38
2.4.6 Motivation	40
2.5 Summary	42
3. MATERIALS AND METHODS	44
3.1 Materials	44
3.1.1 DNA Samples	44
3.1.2 Agarose Gels	45
3.2 Methods	45
3.2.1 Horizontal Electrophoresis with Uniform Electric Field	46
3.2.2 Electric Field Dependence (Horizontal Geometry)	48
3.2.3 Electric Field Dependence (Vertical Geometry)	50

4. RESULTS	57
4.1 Electrophoresis data and best-fit parameters	59
4.1.1 Horizontal Electrophoresis with Uniform Electric Field	60
4.1.2 Horizontal Electrophoresis as a function of Electric Field	63
4.1.3 Vertical Electrophoresis as a function of Electric Field	68
4.2 Effective Pore Size	82
4.3 Reptation Plots and Phase Diagrams	86
4.3.1 Reptation plots accentuate deviations from fits for horizontal electrophoresis	87
4.3.2 Reptation Plots and Phase Diagrams for vertical electrophoresis	92
4.4 Residuals	99
4.5 Behavior of Derivatives	103
4.5.1 Derivatives for horizontal electrophoresis	103
4.5.2 Derivatives for vertical electrophoresis	107
4.6 Summary	111
5. DISCUSSION	113
5.1 Trends in the Electrophoretic Mobilities	118
5.2 Phase Diagrams	119
5.3 The Best-Fit Parameters and Pore Size	121
5.4 Summary	124
6. FITS TO DATA OBTAINED FROM THE LITERATURE	125
6.1 Method	125
6.2 Results	126
6.3 Summary	143
7. CONCLUSIONS	145
REFERENCES	148
BIOGRAPHICAL SKETCH	155

LIST OF TABLES

4.1 Electrophoresis Conditions and Best-Fit Parameters	64
4.2 Vertical Gel Electrophoresis Conditions and Best-Fit Parameters	70
4.3 Agarose Pore Sizes	84
4.4 Agarose Pore Size Comparison with Literature	85
6.1 Summary of the data obtained from the literature	126
6.2 The quality of the fits to data obtained from the literature	127
6.3 The parameters of the fits to data obtained from the literature	140

LIST OF FIGURES

1.1 The Watson and Crick model for DNA	4
1.2 A space filling DNA model	5
1.3 Basic repeating unit of agarose	6
1.4 Red Algae	6
1.5 A proposal for the gelation process for agarose	7
1.6 Representative dependencies of DNA mobilities on length plotted in different ways	10
2.1 Object in an ionic solution	15
2.2 The zeta potential	17
2.3 Freely Jointed Chain	18
2.4 A schematic drawing demonstrating the mechanisms occurring for different size DNA during gel electrophoresis	21
2.5 A negatively charged particle	23
2.6 The difference between free solution electrophoresis and gel elec- trophoresis	25
2.7 Ogston's model	26
2.8 Ferguson Plot	31
2.9 The Tube model	34
2.10 Reptation with stretching	37
3.1 Setup for Horizontal Gel Electrophoresis	47
3.2 Horizontal Gel Electrophoresis at $E = 2.2$ V/cm and $T = 1.0\%$	49
3.3 Vertical Gel Holder	51
3.4 Design Schematic of Vertical Gel Holder	52
3.5 The Setup for Vertical Gel Electrophoresis	53
3.6 The Gel Doc 2000 gel documentation system	54
3.7 Vertical Gel Electrophoresis at $E = 2.48$ V/cm for $T = 0.7\%$ and 1.0%	55

4.1 Vertical Gel Electrophoresis	59
4.2 Dependence of DNA mobility on length and agarose gel concentration for $E = 2.78$ V/cm	60
4.3 Fits of Literature Data	61
4.4 Dependencies of the limiting mobilities, μ_s and μ_L , and the crossover length γ on the gel concentration, T	62
4.5 Influences of electric field strength and gel concentration on fit quality .	65
4.6 Dependence of μ on L at $E = 3.51$ V/cm	66
4.7 Dependence of μ on L at $E = 1.00$ V/cm	67
4.8 Influences of electric field strength and gel concentration on fit quality for vertical electrophoresis	71
4.9 Dependence of μ on L at $E = 4.97$ V/cm	72
4.10 Dependence of μ on L at $E = 0.93$ V/cm	73
4.11 The ratios for the mobilities associated with shortest and longest DNA lengths	74
4.12 The short length mobility limit, μ_s	76
4.13 The Free Solution Mobility, μ_o	76
4.14 The Ferguson Plot For $E = 1.24$ V/cm	77
4.15 The long length mobility limit, μ_L	78
4.16 The exponential scaling factor, γ	80
4.17 The radius of gyration, R_g , as a function of γ	81
4.18 The calculated pore sizes for agarose gels	83
4.19 Representative dependencies of DNA mobilities on length plotted in different ways	86
4.20 Reptation plots for data representing good and poor fits at high and low fields	89
4.21 Reptation plots for $E = 1.31$ V/cm showing how the crossover lengths were determined	90
4.22 Phase diagram for dsDNA electrophoresed at different fields	91
4.23 Reptation plots for $E = 4.97$ V/cm	92
4.24 Reptation plots for $T = 1.0\%$	93
4.25 Reptation plots for $E = 0.93$ V/cm	94

4.26	Reptation plots for electrophoresis done in a vertical chamber at $E = 1.24$ V/cm showing how the crossover lengths were determined	95
4.27	Phase diagrams for vertical gel electrophoresis on dsDNA at different E and $\%T$	96
4.28	Phase diagram for vertical gel electrophoresis on dsDNA as a function of L and R_g	98
4.29	Residual Plots	100
4.30	The maxima and minima from the residuals	101
4.31	The 1st maxima from the residuals	102
4.32	The 2nd minima from the residuals	103
4.33	DNA length (L) dependencies of the derivatives $\Delta \log(\mu(L))/\Delta \log(L)$ on field and gel concentration	104
4.34	Dependencies of $\Delta \log(\mu(L) - \mu_L)/\Delta \log(L)$ versus L on gel concentrations at $E = 1.00$ V/cm	106
4.35	DNA length (L) dependencies of the derivatives $\Delta \log(\mu(L))/\Delta \log(L)$ on field and gel concentration	108
4.36	Dependencies of $\Delta \log(\mu(L) - \mu_L)/\Delta \log(L)$ versus L for vertical gel electrophoresis	110
5.1	Data plotted by Barkema et al.'s (1995) method	116
5.2	Fits of vertical electrophoretic data to Calladine et al.'s (1991) model	116
6.1	Influences of electric field strength and gel concentration on fit quality for Hervet (1987) and Stellwagen (1985)	130
6.2	Dependence of mobility on length at $E = 0.64$ V/cm for Stellwagen (1985)	132
6.3	Dependence of mobility on length at $E = 0.61$ V/cm for Slater, et al. (1988)	133
6.4	Dependence of mobility on length at $T = 0.6\%$ for Hervet and Bean (1987)	134
6.5	Dependence of mobility on length at $T = 0.6\%$ for Stellwagen (1985)	135
6.6	Dependence of mobility on length at $T = 0.8\%$ for Heller, et al. (1994)	136
6.7	Dependence of mobility on length at $E = 3.3$ V/cm for Stellwagen, et al. (2002)	137

6.8	Dependencies of the limiting mobilities, μ_s and μ_L , the crossover length γ , and the pore size, a , on E for Heller, et al. (1994)	138
6.9	Dependencies of the limiting mobilities, μ_s and μ_L , the crossover length γ , and the pore size, a , on $\%T$ for Slater, et al. (1988)	139

ABSTRACT

Separations were performed on double stranded DNA (dsDNA) using electrophoresis. Electrophoresis is the steady transport of particles under the influence of an external electric field. Double stranded DNA fragments ranging in length from 200 base pairs (bp) to 194,000 bp ($0.34 \text{ nm} = 1 \text{ bp}$) were electrophoresed at agarose gel concentrations $T = 0.4\% - 1.5\%$. The electric field was varied from 0.62 V/cm to 6.21 V/cm . A wide range of electric fields and gel concentrations were used to study the usefulness of a new interpolation equation, $\frac{1}{\mu(L)} = \frac{1}{\mu_L} - (\frac{1}{\mu_L} - \frac{1}{\mu_s})e^{-L/\gamma}$, where μ_L , μ_s , and γ are independent free fitting parameters. The long length mobility limit is interpreted as μ_L , the short length mobility limit is μ_s , and γ is the crossover between the long length limit and the short length limit. This exponential relation fit very well ($\chi^2 \geq 0.999$) when there are two smooth transitions observed in the “reptation plots” (plotting $3\mu L/\mu_o$ vs. L) (J. Rousseau, G. Drouin, and G. W. Slater, *Phys Rev Lett.* 1997, 79, 1945-1948). Fits deviate from the data when three different slopes were observed in the reptation plots. Reptation plots were used to determine a phase diagram for dsDNA migration regimes. The phase diagrams define different regions where mechanisms for molecular transport affect the migration of dsDNA in agarose gels during electrophoresis. The parameters from the equation have also been interpreted to provide a physical description of the structure of the agarose gel by calculating the pore sizes. The relations between the values for the pore sizes and the phase diagrams are interpreted to better understand the migration of the DNA through agarose gels.

CHAPTER 1

INTRODUCTION

A revolution in biological sciences has taken place over the last century and biophysics has played an important role. Biophysics is concerned with physical and chemical explanations of living processes. Detailed molecular descriptions are emerging for genetic difference and for the mechanisms that control their propagation and expression. Protein structure [1–6], nucleic acid structure [7–12], surface structures of biological materials [13–14], and genomics all have been subject to intense biophysical study. Scientists with strong backgrounds in mathematics, chemistry, and physics are playing more and more important roles in these developments. Mathematical models can be of great help when studying more quantitative aspects of biological systems.

In a historical perspective, physics is a general approach to understanding the material world rather than a particular set of disciplines and methods. Physics provides the description of the elementary constituents of the material world and the interactions among these constituents. It also establishes principles of empirical analysis and mathematical model building which enriches the biological sciences when applied. The physics in biology developed rapidly in the early twentieth century as a geometric model for the animal form was analyzed by D’Arcy Wentworth Thompson [15] and the basis for contemporary studies of membrane potentials and ion channels were developed by Hermann Walther Nernst [16–17]. In 1953, Watson and Crick [7] developed an understanding for the structure of deoxyribonucleic acid (DNA) by applying x-ray crystallography, a tool of physics and chemistry. In recent

years, many biophysicists have become interested in investigating the physical basis of biological function of specific molecules [18–23], the mechanisms of the motion of molecules [5, 8, 10, 24–30], the functions of cellular organelles or organs [31], and in refining existing methods so as to be able to probe these systems in greater detail [32–36]. Refining existing methods is the focus of this dissertation.

In the past 30 years, performance of molecular separations has been a critical tool in biophysics. Separations are used routinely throughout the world of biotechnology, from genomics to pharmaceuticals. A common method which is used to separate molecules is electrophoresis. Electrophoresis, or in the older literature called cataphoresis, is the steady transport of particles under the influence of an external electric field. F. Reuss [37] first described electrophoresis in a paper published in 1809, where he noticed electrophoresis of clay particles through a bed of quartz. At first, during electrophoresis it was observed that practically every surface became charged when placed in contact with an aqueous solution. Quantitative theories slowly developed to describe this change in more detail. Smoluchowski made an important contribution to the theory of electrokinetics in 1903 by deriving relations between the velocity, the electrostatic potential at the surface, and the applied field for thin double layers [38]. The first sophisticated electrophoretic apparatus was developed by Tiselius in 1937. He was awarded the 1948 Nobel prize in chemistry for his work in protein electrophoresis. He developed the “moving boundary,” which later would become known as zone electrophoresis, and used it to separate serum proteins in solution [39]. In the 1940’s and 1950’s electrophoresis was being widely used and applied to molecules ranging from proteins to amino acids. In the 1970’s interest surged in deoxyribonucleic acid (DNA) electrophoresis. In 1973, Joseph Sambrook led a team at Cold Spring Harbor Laboratory that refined DNA electrophoresis by using agarose gels [40] (see section 1.2). This has led to many new theories and discoveries regarding the mechanisms of the motion of DNA [41–42].

1.1 DNA

A critically important and fundamental constituents of the living cell are nucleic acids. In the last hundred years it has been discovered that nucleic acids are the repositories and transmitters of genetic information for every cell, tissue, and organism [31]. Nucleic acids can simply be thought of as the blueprints for an organism, which contributes to most of the organism's physical development. There are two types of nucleic acids, ribonucleic acid (RNA) and deoxyribonucleic acid (DNA). The focus of this project is on DNA.

Besides being of such fundamental biological interest, DNA has been the focus of research for physicists [7, 10, 43–44] due to its properties as a model molecule. DNA is a semi-flexible polymer chain with four monomer units connected by covalent bonds. These monomers are molecules with purine or pyrimidine bases attached to a carbon molecule. A long nucleic acid chain is formed once the monomers connect, which is referred to as a polynucleotide. The DNA can be cut into different lengths. When the molecule is placed in an aqueous solution the phosphate group will acquire a negative charge, since it is a strong acid, and the overall charge of the DNA will be negative. Thus DNA can be thought of as semi-flexible charged polymer chains of adjustable lengths.

DNA is composed of bases from two basic groups called purines and pyrimidines. DNA contains two purines, adenine (A) and guanine (G), and two pyrimidines, thymine (T) and cytosine (C). Watson and Crick first discovered in 1953, that these two basic groups pair up with each other [7]. These pairs are: adenine with thymine, and guanine with cytosine which form a double helical structure with a uniform diameter (Fig. 1.1).

To form the DNA double helix (figure 1.1), bases connect phosphate-sugar chains. For this dissertation a DNA molecule can be thought of as a long random coil [45] which is the right illustration in figure 1.1. The microscopic view on the left in

figure 1.1 represents the double helical structure of DNA discovered by Watson and Crick [7]. In the Watson and Crick model the hydrophilic phosphate and sugar backbones of the helix are on the outside and the base pairs are inside and stacked on top of each other perpendicular to the helix axis. Due to the strong van der Waals interactions between the stacking of the bases, the base pairs are rotated 36° with respect to each other with 10 base pairs (bp) in each turn of the helix. Watson and Crick saw with x-ray diffraction that the distance between each turn is 3.4 nm. The distance between base pairs is thus about 0.34 nm [7]. Although the bases are inside, they are accessible to solute molecules via two helical channels called the major and minor grooves (see Fig. 1.2).

The double stranded model that Watson and Crick conceived, allowed them to draw the conclusion that DNA has the ability to self-replicate [7]. Since A always pairs to T, and G always pairs to C, the two strands are complementary and this would permit the molecules to replicate once the two strands separate. During

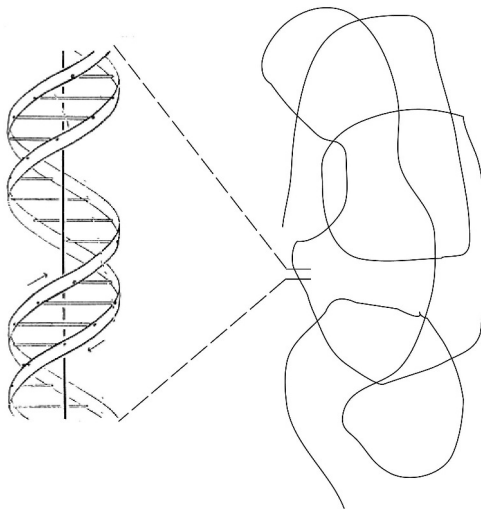


Figure 1.1. The Watson and Crick model for the structure of DNA. A DNA molecule can form into a long coil in a macroscopic view (figure on the right). Watson and Crick [7] discovered the microscopic structure of DNA is double helical (figure on the left). The two ribbons symbolize phosphate-sugar chains and the horizontal rods are the bases holding the ribbons together. The vertical line is the fiber axis.

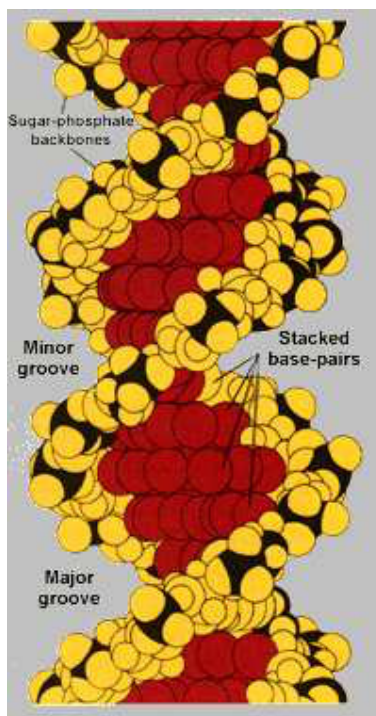


Figure 1.2. A space filling DNA model. The DNA molecule shown here displays each atom with its van der Waals radius. The black regions indicate the sugar and phosphate backbone and the yellow regions indicate the oxygen molecules. The red regions denote the base pairs in the Watson-Crick model of DNA [7]. The major and minor grooves are also designated. [31]

replication single-stranded DNA (ssDNA) occurs temporarily by using localization of torque to force the double-stranded DNA (dsDNA) to denature (to open) at a specific point. Double-stranded DNA can also denature if it is heated above its melting temperature ($\geq 90^{\circ}\text{C}$ in physiological solutions and $\geq 50^{\circ}\text{C}$ in distilled water).

1.2 Agarose Gels

Agarose forms gel media used by biologists and biochemists for electrophoretic separations of DNA. Electrophoretic separations with agarose are simple, rapid to perform, and capable of resolving mixtures of DNA fragments that cannot be separated adequately by other sizing procedures. Furthermore, the location of DNA

within the gel can be determined directly due to the uniformity of the matrix formed in an agarose gel. All the experiments presented for this work used agarose gels.

Agarose is a linear polysaccharide (average molecular mass about 120,000 Daltons) made up of the basic repeat unit agarobiose, which consists of alternating units of galactose and 3,6-anhydrogalactose (Fig. 1.3). This basic repeating unit is a purified form of agar which is a constituent of both the outermost portion of cell walls and the intercellular matrix of red algae. Figure 1.4 shows a picture of red algae which is the origin of agarose.

Some data suggest that during gelation, single strands of agarose dimerise to form a double helix. The double helix has a pitch of 1.90 nm and each strand has three-fold helical symmetry [14]. Gelation occurs when shortened agarose chains

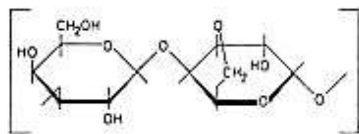


Figure 1.3. Basic repeating unit of agarose.

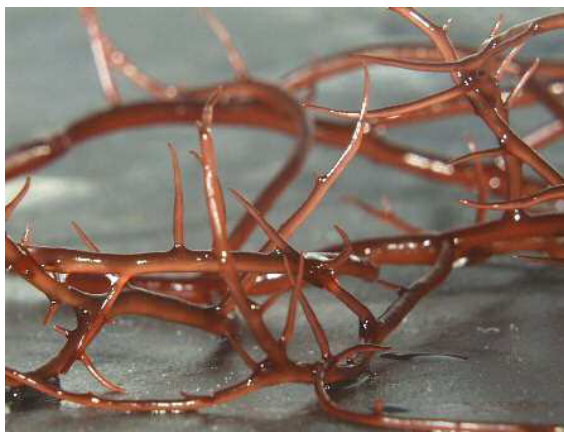


Figure 1.4. Red Algae (or *Gracilaria tikvahiae*). This figure was obtained from [//www.globaldialog.com/~jrice/algaepage/redalgae.htm](http://www.globaldialog.com/~jrice/algaepage/redalgae.htm) (maintained by John Rice, jrice@globaldialog.com).

appear to form double helices that can be joined by some galactomannans to produce a gel. The width of an agarose double helix has been determined to be no more than 1.4 nm [46]. However, gel fibers an order of magnitude wider than this have been observed by electron microscopy [47], AFM [48], light scattering [49–50], and x-ray scattering [51–52]. To explain this, Serwer [14] suggested that the agarose double helices aggregate and form fibers (suprafibers) with radii larger than the radii of the agarose double helix. It is presumably the presence of suprafibers that causes agarose gels to have strengths and pore sizes greater than those of gels that don't form suprafibers (see Fig. 1.5).

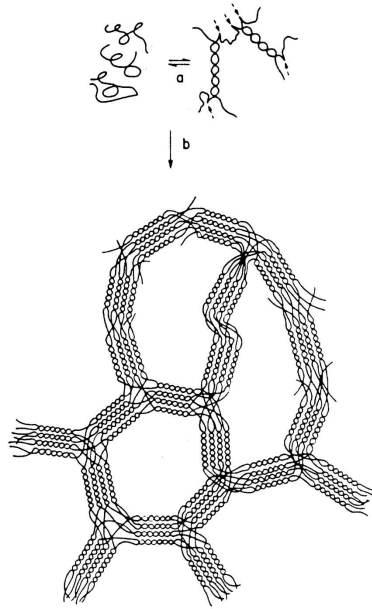


Figure 1.5. A proposal for the gelation process for agarose. During the early stages of gelation the agarose shifts from a random coil configuration to double helices. Eventually the double helices aggregate to form bundles. The figure is from Serwer (1983) [14].

1.3 The Research

The work presented in this dissertation is an in-depth study of electrophoretic separations of DNA fragments according to length. The technique plays an integral role in molecular biology and biotechnology. Understanding the molecular mechanisms of separations of different types and lengths of nucleic acid chains are of fundamental interest. This understanding should provide well-defined models for electrophoretic transport of polyelectrolytes in general. Separations of molecules are usually accomplished by electrophoresis in randomly cross-linked hydrogels, like polyacrylamide or agarose. The electrophoretic mobility, $\mu(L)$, of DNA in agarose gels decreases essentially monotonically with increasing chain contour length, but eventually reaches a limit independent of length. Noolandi et al. [53] first found a non-monotonic $\mu(L)$ dependence called “band-inversion” that occurs for long fragments separated at high gel concentrations and high electric fields. The exact transport mechanism, hence the form of the mobility dependence on length, is generally thought to be a complex function of the electric field strength and the size of the gel pores relative to the effective DNA size. Mechanisms proposed for DNA electrophoresis through gels includes simple Ogston sieving (for small DNA) [5, 24, 54–57], entropic trapping [18–20, 58–60], simple biased reptation [26, 28, 61–63], reptation with fluctuations [30, 64–66], reptation with stretching [26, 28, 30, 62, 67], the repton model [35, 68–72] and the “door-corridor” model [73–77]. The functional form of $\mu(L)$ has also been described by phenomenological models [8, 36].

Each mechanism proposed for DNA electrophoresis states that the electrophoretic mobility has a different dependence on length. For example, Lumpkin and Zimm, et al. [78] developed an adaptation of the reptation model (proposed originally by de Gennes [25]) that suggested the electrophoretic mobility to be inversely proportional to length. This is the basis of the reptation theory which is discussed in detail in section 2.4.3. Southern [10] observed that the reciprocal mobility of

DNA fragments electrophoresed in agarose gels was proportional to length (see section 2.4.5). Rodbard and Chrambach [55–56, 79] provided an extension of the Ogston model [24], which related the Ogston sieving motion of a molecule in a porous network to DNA electrophoresis. The various models each suggest a different μ versus L dependence.

Initially, experiments were performed to find over what ranges these different models apply. Horizontal gel electrophoresis was performed for dsDNA fragments spanning two decades in length (100 bp to 10,000 bp) and the data were plotted in different ways (figure 1.6). Horizontal gel electrophoresis is the standard method used for dsDNA electrophoresis in agarose gels where a uniform electric field is applied to a horizontal apparatus containing the gel to separate the DNA (see section 3.2.1). Results indicated that the existing theoretical relationships inadequately describe DNA mobility as a function of length over the entire length range of DNA fragments used. Southern’s [10] proposed empirical relationship produces a linear region between 100 bp to 4000 bp (figure 1.6d) when plotting $1/\mu$ versus L . A plot of μ versus $1/L$, which if linear can indicate reptation, shows a linear region between 3000 bp to 10,000 bp. These experiments are discussed in detail in section 4.1.1. The proposed plots to represent the existing theories are linear only for limited regions.

It shouldn’t be surprising that each of these relationships apply over a limited range of lengths. The 100 bp DNA is shorter than a persistence length, and thus is nearly rod-like. A 10,000 bp fragment acts as a random coil polymer. By judicious choice of functional form we found a fitting function that works over the entire length range for more separation conditions. As the dissertation will reveal, a simple empirical exponential equation was discovered to represent the mobility as a function of DNA length over the entire length range for many gel concentrations and electric fields. The initial results were done only for a single electric field for various agarose gel concentrations. The data from these experiments were extraordinarily well fit by

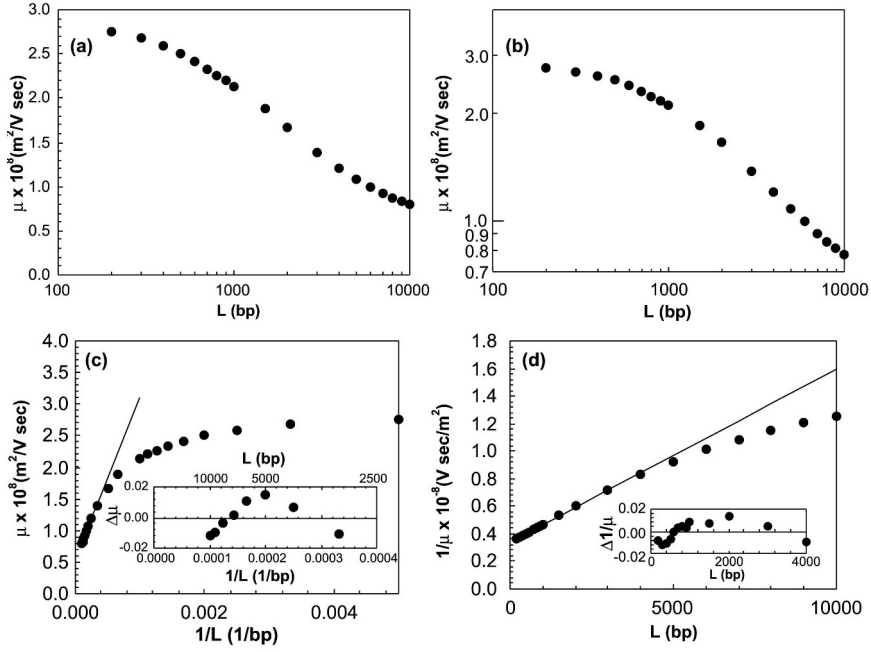


Figure 1.6. Representative dependencies of DNA mobilities on length plotted in different ways. Electrophoresis was performed on 0.8% Seakem LE agarose gel: (a) μ vs. $\log(L)$, (b) $\log(\mu)$ vs. $\log(L)$, (c) μ vs. $1/L$, (d) $1/\mu$ vs. L . A straight line was fit from 3000 bp to 10,000 bp for μ vs. $1/L$ and from 100 bp to 4000 bp for $1/\mu$ vs. L . For both graphs, the residuals (which are shown in the lower right panels) show trends indicating that deviations from the straight line were not random.

the phenomenological equation, $\frac{1}{\mu(L)} = \frac{1}{\mu_L} - (\frac{1}{\mu_L} - \frac{1}{\mu_s})e^{-L/\gamma}$, where μ_L , μ_s , and γ are independent free fitting parameters.

A more extensive study was performed to find the versatility of this equation. Horizontal electrophoresis was conducted for a wide range of electric fields, gel concentrations, and over a 485-fold dsDNA length range. The equation was able to describe a remarkably wide range of data. There were limited ranges of fields and concentrations where the equation fit the data poorly. In these regions, different plotting techniques were able to shed light on possible molecular mechanisms to explain why the data deviated from the fits. As the reader will discover in section 4.3.1, comparing the fit to the data can be used to indicate where the onset of these different mechanisms for molecular transport will occur.

A different experimental setup was developed to provide more precise data for a better understanding of the parameters for this equation. Vertical gel electrophoresis was performed over a wider range of electric fields. An uniform electric field was applied to the gel held in a vertical apparatus. There are advantages of vertical gel electrophoresis over horizontal. The vertical setup allowed for thinner gels, which produced sharper bands. Also the thinner gels dispersed heat more efficiently, hence were affected less by heating during electrophoresis [80–82]. The results for these experiments displayed the same trends as before, thus the fitting equation can be applied to different experimental setups. These experiments provided more consistent results that allowed discovery of interesting trends in the parameters from the equation. A physical picture was suggested by the parameters, which provided a better understanding for the motion of DNA during electrophoresis. Slater predicted that the parameters can be arranged in a manner which gives a physical description of the gel matrix [34]. When his proposed theory was applied to the parameters obtained from the experiments, the pore size could be estimated for the gel being used for that specific experiment.

The work has also provided other major contributions to the field, such as allowing for interpolating data. Using this equation as an interpolation standard for dsDNA electrophoresis should be a useful tool in the field. From these experiments the best fit over the longest length range for separations of dsDNA were found. From $E = 1.5 \text{ V/cm}$ to 3.5 V/cm at $T = 1.0\%$ are the best range of data for interpolating electrophoretic mobilities of dsDNA in agarose gels. By applying DNA fragments associated with the long and short length limiting mobilities and a few fragments in between, the fitting function can be used to interpolate the rest of the DNA fragments in between.

Electrophoresis was performed over much of the useful range of electric fields and agarose gel concentrations for a long DNA length range. In the past people have performed these type of experiments [10, 29, 35–36, 58], but usually only over a 1

decade range of DNA lengths and/or for limited range of fields. The work presented here provides the most extensive set of data for dsDNA electrophoresed in agarose gels. The range of data for DNA lengths, electric fields and gel concentrations can easily be applied to many aspects in research studying separations. The extensive data set provides guidance for the best set of conditions for separating DNA using electrophoresis. It is my belief that the data in this dissertation can be used as a standard for future research done with dsDNA electrophoresis in agarose gels.

1.4 Summary

The focus of this work is electrophoretic separations for dsDNA in agarose gels. Various electrophoresis experiments were performed over a wide range of fields, concentrations, and DNA lengths. The existing theories describe the data obtained from these experiments only for limited DNA length ranges. To remedy this problem, a simple empirical equation was found to fit the experimental data well in most regions. It is proposed that this equation can help determine the onset of different mechanisms governing the motion of DNA during electrophoresis and determine the structure of the media in which the separations take place.

CHAPTER 2

THEORY

As mentioned briefly in the Introduction, electrophoresis is an important technique used to separate macromolecules. For this work the macromolecule being separated is double-stranded DNA (dsDNA). Double-stranded DNA molecules can be as small as a few base pairs or $\geq 10^6$ base pairs (megabase pair, Mbp). Recall from section 1.1 that there are 0.34 nm per base pair. In other words the DNA fragments can range from 3.4 nm for one turn (10 bp) to larger than 0.34 cm for 10^6 bp. The smallest are essentially stiff rods and the longest are random coil polymers. Clearly, different mechanisms of motion during electrophoresis will apply for the vast range of sizes. Before discussing the state of knowledge of the mechanisms of electrophoretic separation of DNA, models for a particle or molecule in an ionic solution and of an ideal macromolecule will be discussed.

2.1 Object in an Ionic Solution

Forces on a polyion immersed in a fluid with mobile ions can be described to first order by approximating the polymers as smooth spheres. Fig. 2.1 shows a negatively charged spherical particle in an ionic solution. There is a distribution of ions in the surrounding interfacial region. This results in an increased concentration of counter ions (ions of opposite charge to that of the particle) close to the surface. The liquid layer surrounding the particle exists as two parts: an inner region (Stern layer) where the ions are strongly bound and an outer (diffuse) region where they are less firmly associated. The Stern layer is the region that contains the surface charge, while the

diffuse layer is composed of ions freely moving under the influence of electrical and thermal forces (as shown in Fig. 2.1). The combination of these two layers is referred to as the electric double layer. The thickness of the Stern layer is estimated to be

$$l_B = \frac{e^2}{4\pi\epsilon\epsilon_0 kT} , \quad (2.1)$$

where l_B is called the Bjerrum length, ϵ is the dielectric constant of the fluid, ϵ_0 is the permittivity of vacuum, e is the charge of an electron ($\sim 1.60 \times 10^{-19}$ C), k is the Boltzmann constant ($\sim 1.38 \times 10^{-23}$ J/K), and T is the temperature [83].

The mobile ions in the diffuse layer affect the overall charge density. From electrodynamics we know that the electrostatic potential for this system must satisfy

$$\epsilon\epsilon_0 \nabla^2 \psi = -\rho_f , \quad (2.2)$$

where ρ_f is the free charge density and ψ is the electrostatic potential. It is also known that the general Poisson-Boltzmann equation is

$$\frac{d^2\psi(x)}{dx^2} = \frac{q\rho(x)}{\epsilon_0 kT} , \quad (2.3)$$

where q is the charge and $\rho(x)$ is an arbitrary surface charge. The Poisson-Boltzmann equation describes the distribution of charges.

By using equation 2.2 and the Poisson-Boltzmann equation (eq. 2.3), the equilibrium distribution of mobile ions may be found by [84],

$$\epsilon\epsilon_0 \nabla^2 \psi = -e \sum_{i=1}^N z_i C_i \exp(-ez_i\psi/kT) , \quad (2.4)$$

where the free charge density,

$$\rho_f = \sum_{i=1}^N ez_i n_i , \quad (2.5)$$

with z_i the valence charge of ion i , C_i the bulk concentration of ions of species i , and n_i the Boltzmann distribution represented by

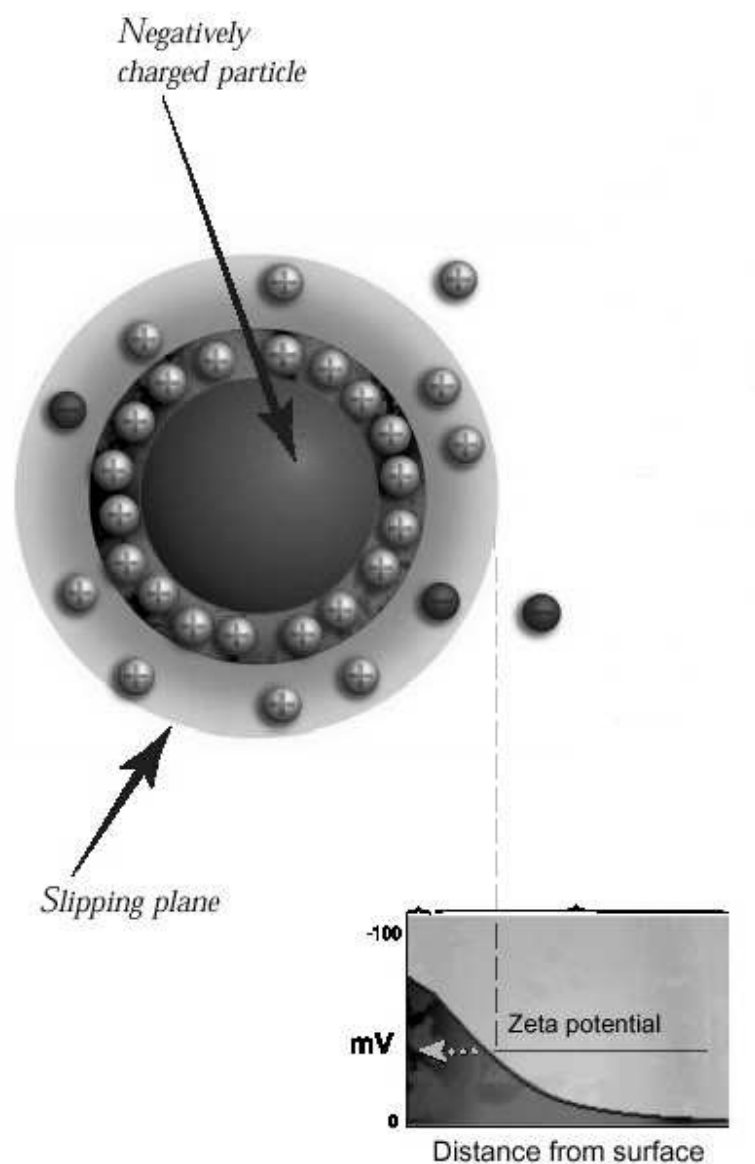


Figure 2.1. Object in an ionic solution. A negatively charged particle in a solution with an electric double layer surrounding the particle in an ionic solution. The slipping plane is also referred to as the shear plane which is an imaginary surface separating the thin layer of liquid bound to the solid surface. The electric potential at the shear or slipping plane is called the zeta potential. The picture was acquired from the following web site <http://www.silver-colloids.com/Tutorials/Intro/pcs17A.html> (maintained by Malvern Instruments Ltd.).

$$n_i = C_i \exp(-ez_i\psi/kT) . \quad (2.6)$$

Equation 2.4 is the basis of the Gouy-Chapman model of the diffuse charge cloud adjacent to a charged surface [85–86]. The Gouy-Chapman model makes the assumptions that the electrolytes are an ideal solution with uniform dielectric properties, the ions are point charges, and the potential of mean forces between the ions and the average electrostatic potential are equal. The solution to this equation yields the Debye-Hückel theory, which applies to the diffuse layer,

$$\psi(x) = \psi_s e^{-\kappa x} , \quad (2.7)$$

where ψ_s represents the potential at the surface, and κ^{-1} is the Debye thickness for a general electrolyte, which is given by

$$\kappa^{-1} = \left(\frac{\varepsilon \varepsilon_o kT}{e^2 \sum_i z_i^2 C_i} \right)^{1/2} , \quad (2.8)$$

So, according to the Debye-Hückel theory, the counterion charges are primarily in a region of thickness κ^{-1} adjacent to the interface.

Within the diffuse layer is a notional boundary known as the slipping plane within which the particle and its counterions acts as a single entity. The slipping plane is also referred to the shear plane. It can be thought of as an imaginary surface separating the thin layer of liquid bound to the solid surface from the bulk liquid. The electric potential at the shear or slipping plane is called the zeta potential, ζ . The zeta potential is a measure of the difference in charge at the surface of the membrane and the outer edge of the ionic cloud (slipping plane), which will give the overall charge a particle acquires in a specific medium. The magnitude of the ζ -potential determines the potential stability of the colloidal system. Fig. 2.2 depicts the range at which the particle will have a stable and unstable ζ -potential. The greater the zeta potential the more likely the suspension is to be stable because the charged particles repel one

another and thus overcome the natural tendency to aggregate. Therefore, the zeta potential is an important parameter characterizing colloidal dispersion.

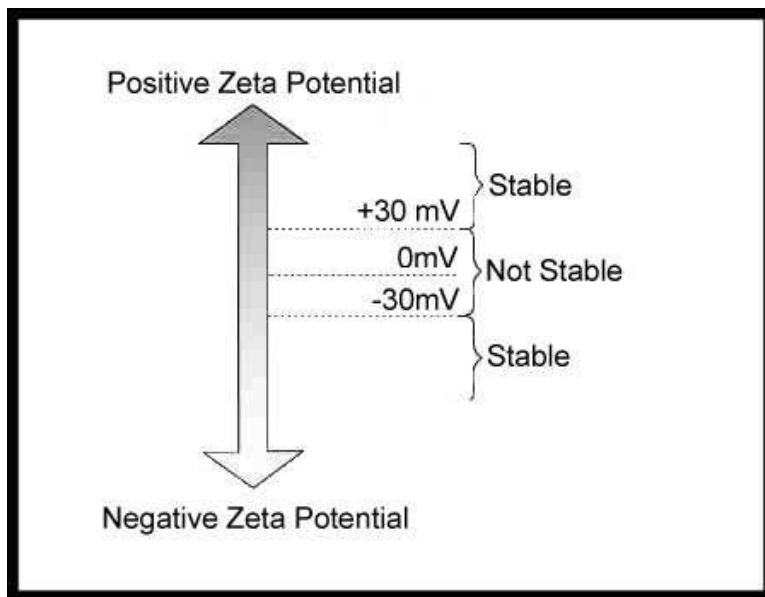


Figure 2.2. The zeta potential. The range at which particles have stable and unstable ζ -potentials. A dividing line between stable and unstable aqueous dispersions is generally taken at either +30 or -30mV.

2.2 Polymer Chain Conformation

The flexibility of the molecule also plays an integral role for separations of macromolecules during gel electrophoresis. To understand the physics behind this, the model of an ideal polymer chain is described.

2.2.1 Freely Jointed Chain

A flexible polymer can be thought of as random chains or a freely jointed chain composed of a sequence of N rigid segments, each of length l and able to point in any direction independently of each other (Fig. 2.3) [45].

The conformation of the chain can be characterized by observing the end-to-end distance, R . Since any value for the vector R can be found as frequently as the opposite value, the mean value, $\langle R \rangle = 0$. For each segment, i , we consider u_i to be a bond vector with x_i being the radius vector for the beginning of the segment and x_{i+1} being the radius for the end of the segment (Fig. 2.3). With these simple characteristics of the chain, the end-to-end vector, R , can be denoted as

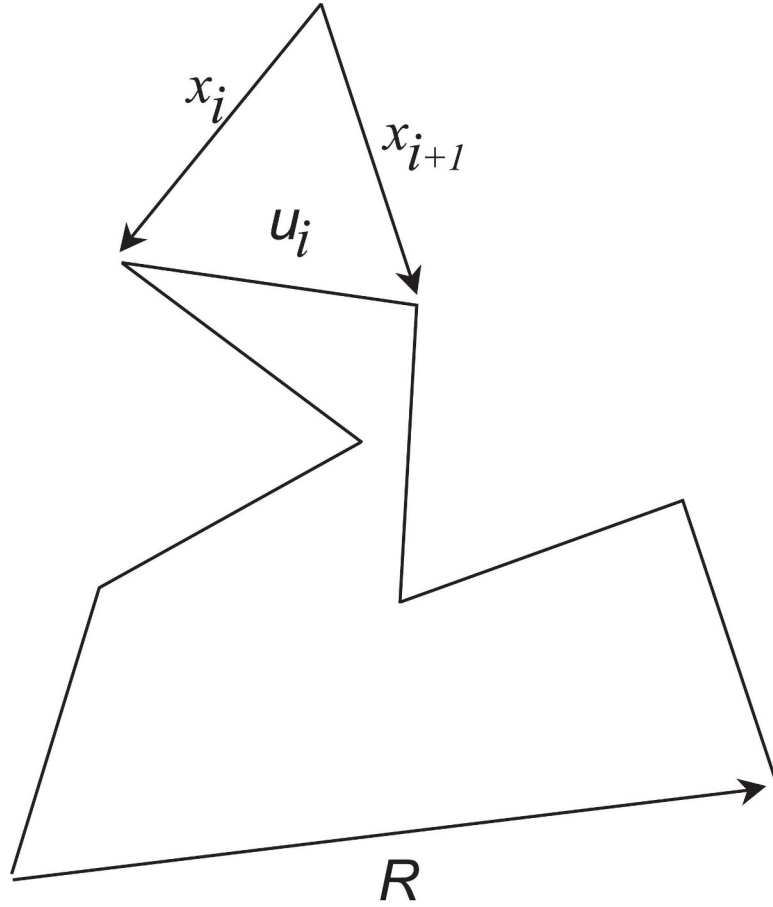


Figure 2.3. Freely Jointed Chain. The end-to-end vector is denoted by R . The radius vectors x_i and x_{i+1} , denote the beginning and the end of the i -th segment. The bond vector, u_i , is calculated by $u_i = x_{i+1} - x_i$.

$$R = \sum_{i=1}^N u_i . \quad (2.9)$$

The mean square of the end-to-end vector, $\langle R^2 \rangle$, is

$$\langle R^2 \rangle = \langle (\sum_{i=1}^N u_i) (\sum_{j=1}^N u_j) \rangle = \sum_{i=1}^N \langle u_i^2 \rangle + 2 \sum_{1 \leq i < j \leq N} \langle u_i u_j \rangle . \quad (2.10)$$

Since the chain is freely jointed, the segment directions are not correlated and the angle between the vectors u_i and u_j occurs with equal probability. Also, if $\langle u_i^2 \rangle = l^2$, the mean square of the end-to-end vector will become

$$\langle R^2 \rangle = Nl^2 . \quad (2.11)$$

So for a multilink chain with $N \gg 1$, the mean size of the molecule is represented by $R = \langle R^2 \rangle^{1/2} \sim N^{1/2}l$ and the contour length (or total length) of the molecule is Nl . Thus the mean size is much smaller than the contour length. This result implies that freely jointed chain conformations account for the process of thermal fluctuations of a molecule, while stretched or nearly straight conformations contain a minor fraction. An ideal freely jointed chain in thermodynamic equilibrium conditions corresponds to a randomly shaped coil [45].

The freely jointed chain model is only valid for chains that contain freely rotating connections between rigid segments. In theory it works nicely, but problems arise when applied to polymer chains. The flexibility of a polymer chain varies with its length. Very long polymer chains are quite flexible and an adaptation of the freely jointed model works well. Smaller length polymers can act like a semi-rigid rod with little flexibility.

2.2.2 Persistence Length

The conformation of a polymer chain is greatly affected by its length. A short length polymer can act like a rod, which has no flexibility. At this point the short

chain segments are correlated and the movement of each segment is affected by the previous segment. After a certain length (or segment) the polymer will lose recognition with the origin and will not be correlated with that portion anymore. This model of a polymer chain is called the persistent or *worm-like model* [87]. The length associated with this is referred to as a persistence length, l_p . A chain section that is smaller than the persistence length has limited flexibility and in the limit can be thought of as a rigid rod. In the opposite limit, when portions of a chain are separated by distances much longer than l_p , the flexibility leads to almost total independence of the motion of those portions. At a polymer length comparable to the persistence length the chain does not rely on the origin and a chain's orientation is destroyed. For a double-helix DNA $l_p \approx 50$ nm or ≈ 150 bp.

Long length polymers can be associated with the freely jointed chain model. It would be appropriate to correlate a persistence length of a chain to the length, l , for a freely jointed chain. As mentioned in the previous paragraph the memory of the direction of the chain prevails for only a finite distance of l_p . If the number of each of these sections is $N_p \sim L/l_p$, where L is the contour length, equation 2.11 can be related to long length polymers by:

$$\langle R^2 \rangle \sim N_p l_p^2 \sim \left(\frac{L}{l_p}\right) l_p^2 \sim L l_p . \quad (2.12)$$

Experimentally the contour length, L , and the end-to-end distance, R , can be measured, but the persistence length cannot be determined from equation 2.12 without knowledge of the coefficient. So W. Kuhn [89], and E. Guth and G. Mark [90] introduced a characteristic length in 1934 called the Kuhn segment. The effective length for a Kuhn segment is defined as

$$\langle R^2 \rangle = L l_k , \quad (2.13)$$

where l_k is the effective Kuhn length. Notice that equation 2.13 is equivalent to equation 2.11, The polymer chain behaves as a random walk of Kuhn segments. The Kuhn length and persistence length are of the same order of magnitude, so they both

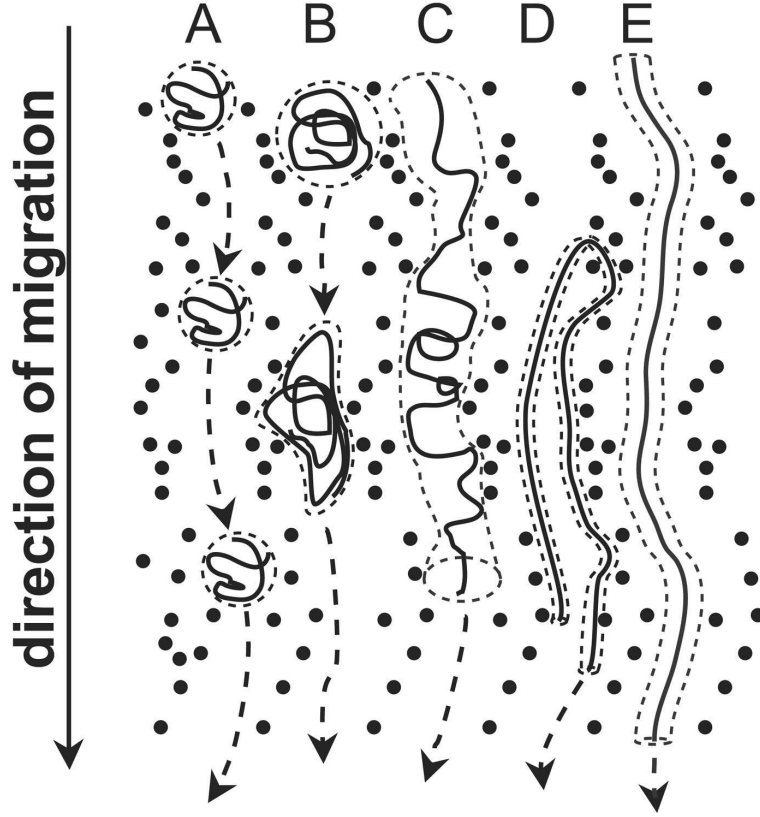


Figure 2.4. A schematic drawing demonstrating the mechanisms occurring for different size DNA during gel electrophoresis. For each different case the gel matrix is the same. (A) Ogston Sieving [24]: A small randomly coiled DNA will have a globular shape which will allow for a sieving motion as it migrates through the gel. (B) Entropic Trapping (ET) [19–20, 27, 58, 88]: The DNA molecule is selectively hopping between the largest pores. (C) Reptation (tube model) [25]: The randomly coiled DNA molecule will migrate head first through the gel and is considered trapped in a tube due to the chain environment. (D) Reptation with stretching [8, 28, 30]: At long lengths the DNA can become hooked on a pore as shown. To free the molecule from the trap the field will pull at either end until the molecule is freed. During this process the molecule becomes stretched. (E) Oriented Reptation [64, 8, 27]: The DNA will align itself with the electric field and migrate head first through the gel.

can be used to describe the flexibility of a molecule. The Kuhn length is found to be twice as long as the persistence length [45]:

$$l_k = 2l_p . \quad (2.14)$$

In order to calculate l_p a statistical definition of chain persistence is needed which is represented by the Kuhn length. The Kuhn segment can be measured experimentally, because it's a direct measure of the chain stiffness and is related to the contour length (see ref [91–93]).

The persistence length and Kuhn length are used to describe the different behaviors that occur for different DNA lengths during gel electrophoresis. Fig. 2.4 illustrates how different length DNA might be imagined to travel through an uniform porous network during gel electrophoresis. As shown in the figure, the flexibility of the molecule plays a very important role. In the following sections all these different mechanisms and theories will be described.

2.3 Free Solution Electrophoresis

Electrophoresis is the transport of charged molecules in response to an electric field. When an electric field is applied to the aqueous system with negatively charged particles described in section 2.1, the negatively charged molecules will start moving in response to the force of the field. The field also acts on the surrounding counter ions seen in Figure 1.3, which will be dragged in the opposite direction and interact hydrodynamically with the molecule (see Fig. 2.5). There will be two forces associated with this process: the drag force, F_d and the accelerating force, F_a . The drag force is the retarding force as the molecule accelerates through the gel or the force of the ions dragging in the opposite of direction. The accelerating force is due to the electric field. When F_a exceeds F_d there is an accelerated migration. In electrophoresis the molecule moves at a constant velocity which makes $F_a = F_d$.

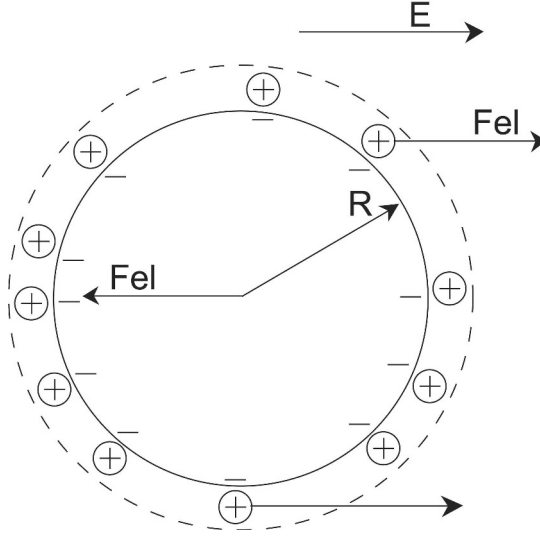


Figure 2.5. A negatively charged particle. A representation of a particle during free solution electrophoresis. A negatively charged spherical molecule with radius, R , is placed in an ionic solution with a uniform electric field, E , applied to the system. F_{el} represents the electric force acting upon the surface of the molecule and surrounding thin layer of ions. The dotted line represents a thin diffuse layer ($R > \kappa^{-1}$).

When the Debye layer is thick ($R < \kappa^{-1}$), the surrounding ions will have less effect on the overall surface charge due to its distance away. Then the following method can be used to determine the mobility of the molecule during electrophoresis. The driving force, F_a can be represented by

$$F_a = EQ \text{ with } E = V/d . \quad (2.15)$$

E is the uniform applied electric field, Q is the charge of the molecule, V is the voltage applied and d is the distance between the electrodes. The drag force is a function of size and shape of the particle and the viscosity of the fluid as expressed by the Stoke's equation,

$$F_d = fv \text{ with } f = 6\pi\eta R , \quad (2.16)$$

where v is the velocity, f is the frictional coefficient for a spherical object, η is the viscosity, and R is the effective radius of the molecule. By setting $F_a = F_d$ and rearranging the equation in terms of mobility,

$$\mu = \frac{v}{E} = \frac{Q}{6\pi\eta R} . \quad (2.17)$$

For a thin Debye layer ($R > \kappa^{-1}$) surrounding the molecule, the slipping plane (or shear plane) has a thickness of κ^{-1} (Fig. 2.1). To find a close approximation of the mobility of the particle one must solve the Navier-Stokes equation for a Newtonian fluid flowing in one dimension,

$$\eta \frac{d^2 v_x(x)}{dx^2} = E_z \rho(x) . \quad (2.18)$$

The charge density, ρ , is not defined the same way as in equation 2.2 and equation 2.4. Instead it's defined as the charge density of the fluid. Since in free solution electrophoresis the velocity becomes independent of size and shape, the mobility can be calculated from equation 2.18 [38, 84] to be:

$$\mu = \frac{\varepsilon \varepsilon_o \zeta}{\eta} . \quad (2.19)$$

One problem that arises from representing the mobility with the potential is that the diffuse layer thickness affects the potential of the inner edge layer (see Fig. 2.1). So to accurately denote the mobility, the chemistry of the interfacial region must not be ignored. A way to avoid this problem is to express the mobility as a function of the surface charge density. Using the Debye-Hückel approximation [84] the mobility will become

$$\mu = \frac{\sigma}{4\pi\eta\kappa} , \quad (2.20)$$

where σ is the surface charge density. Equation 2.20 calculates the electrophoretic mobility for a 2D molecule, while equation 2.17 treats the molecule as a point or sphere with radius, R .

2.4 Gel Electrophoresis

In gel electrophoresis separations of molecules are done in a gel or polymer matrix, which presents a series of obstacles or obstructions, providing friction that depends on molecular size. In the previous section, free solution electrophoresis was discussed, where the molecules mobility and velocity were not hindered by any obstacles. In gel electrophoresis, smaller molecules will tend to travel faster through the obstacles and this will give them a higher electrophoretic mobility (Fig. 2.6). The complexity of this subject comes into play when considering how mobility $\mu = \frac{v}{E}$ depends on E , gel structure, L , and the molecule's flexibility (characterized by l_p).

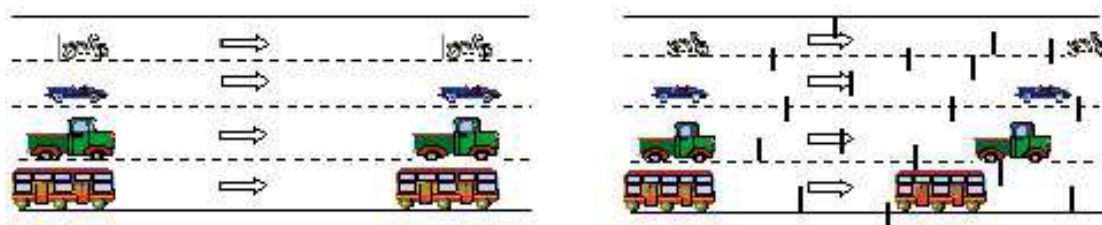


Figure 2.6. The left picture is analogous to free solution electrophoresis. In this case the molecule acts like a free draining coil and is size independent. The picture on the right is analogous to gel electrophoresis. The obstacles or obstructions in the path of the vehicles make up the gel matrix. Small molecules will travel faster through the obstacles and this will give them a higher electrophoretic mobility (This figure was obtained from <http://barronlab.chem-eng.nwu.edu/Electrophoresis.html> maintained by the Barron group at Northwestern).

2.4.1 Ogston Model

Ogston [24] developed a model of the motion of a molecule migrating through spaces in an uniform random suspension of fibers. The molecule is thought of

as a spherical object where the effective molecule radius, R , is smaller than the pore radius, a . Panel A in Figures 2.4 represents the DNA configuration and gel concentration where this model should apply.

Ogston defined the size of spaces as follows: a point of origin is taken at random in the suspension and a spherical surface is expanded from this until it reaches the nearest fiber; it may do this either (i) by making a tangential contact with some point between the ends of a fiber or (ii) by making point contact with the end of a fiber (Fig. 2.7). D is the tangential distance of fiber from the origin or the distance of one of its ends from the origin. To determine the size of the spaces, the probability distribution is calculated for a space that lies between D and $(D + dD)$, which is the combined probability that no tangential or end-contact occurs within D of the origin, and that at least one, of either sort, occurs between D and $(D + dD)$. This calculation gives the probability distribution of a point like particle being in a network of fibers,

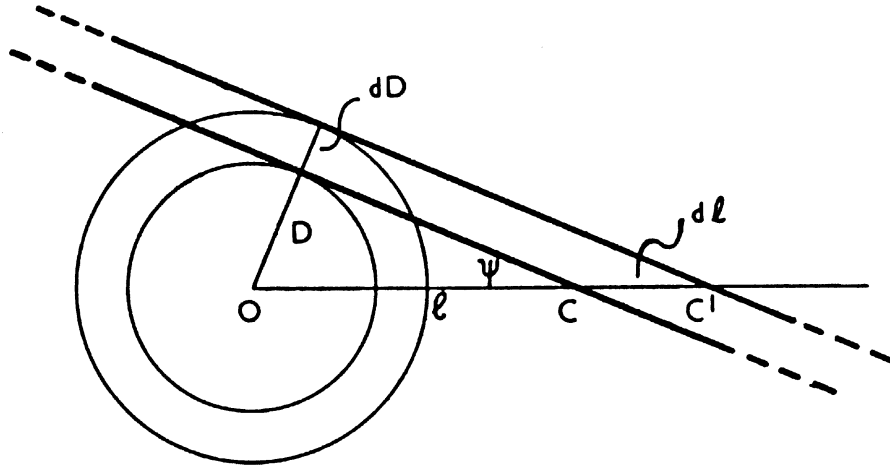


Figure 2.7. Ogston's model. The illustration is taken from the original paper Ogston published in 1958 [24]. The distance from the point of origin to the center of a fiber is l , D is the tangential distance of fiber from the origin or the distance of one of its ends from the origin, and ψ is the angle between a fiber and the line joining its center to the origin.

$$\frac{dP}{dD} = (4\pi\nu LD + 4\pi\nu D^2)e^{-(2\pi\nu LD^2 + \frac{4\pi}{3}\nu D^3)} , \quad (2.21)$$

with ν is the average number of fiber centers/cm³ and $2L$ is the length of fiber. If a spherical particle with radius r , is placed within the network its center must be at least a distance r away from the point of nearest contact with the network. So this is the probability that $D > r$ which is $P_{D>r} = \int_r^\infty \frac{dP}{dD} dD$ and the following is obtained:

$$P_{D>r} = e^{-(2\pi\nu Lr^2 + \frac{4\pi}{3}\nu r^3)} . \quad (2.22)$$

Ogston made a few assumptions for extreme cases that are essential to later developments of this theory: (i) Networks composed of very long fibers: In this case $D \ll L$ which will make the second term in equation 2.21 and equation 2.22 negligible. From equation 2.21, $\bar{D}$ is easily calculated by integrating: $\bar{D} = \int D dP = (8\nu L)^{-1/2}$ is obtained. (ii) Fibers of finite thickness: Finite fiber thickness is allowed for by adding the half-thickness to the value of D (or r). This only occurs when the fiber thickness is not sufficient in relation to the concentration for the fibers to interfere significantly with each others' distribution or orientation. (iii) Flexible Fibers: If the network is composed of interpenetrating flexible fibers, the individual segments within the volume should be considered to be randomly distributed in position and orientation. So the fibers will consist of straight segments with sharp bends between them and equation 2.21 and equation 2.22 can be used.

2.4.2 Extensions of the Ogston Model

C. J. O. R. Morris applied the Ogston model to gel electrophoresis by proposing that the electrophoretic mobility (μ) relative to free solution mobility (μ_o) is equal to f , the fractional volume available to the particle in the gel [6, 94],

$$f = \frac{\mu}{\mu_o} . \quad (2.23)$$

Rodbard and Chrambach [55–56, 79] generalized this idea with another model by Giddings et al. [54] to establish a more realistic 2 dimensional version of the Ogston Model which has been referred to as the free-volume model [8]. Three different dimensions for gels were used to describe the electrophoretic mobility. A 2-D gel was defined as a gel composed of a random meshwork of points. A random meshwork of fibers is described as a 1-D gel. They referred to a 0-D gel as a network composed of a meshwork of planes.

For 2-D gels Rodbard and Chrambach state that Giddings’ [54] result for random fiber gels may be obtained more simply and directly as an extension of Ogston’s model,

$$f = e^{-s\bar{L}}, \quad (2.24)$$

where $\bar{L}$ is the mean length of projection of the molecule along the various axes, and s is the surface per unit volume. Ogston’s assumptions are equivalent to saying that the number of contacts between molecules and gel fiber follows a Poisson distribution. For 1-D gels, the fibers are much longer than the molecule and the fractional available volume is

$$f = e^{-lS}, \quad (2.25)$$

where l is the total length of fiber per unit volume and S is the external molecular surface area. For 0-D gels the fibers are very short and

$$f = e^{-nV}, \quad (2.26)$$

where n is the number of points per unit volume and V is the volume of the molecule. By combining the results of 2-D (Eq. 2.24), 1-D (Eq. 2.25), and 0-D (Eq. 2.26) gels a general equation is obtained,

$$f = e^{-(s\bar{L}+lS+nV)} , \quad (2.27)$$

ignoring weak interaction terms.

In any real gel, either 1-D or 0-D elements predominate with only minor contributions from the 2-D element. If the 2-D element is ignored, the Ogston result will be for a spherical molecule with only the 1-D and 0-D elements being

$$\begin{aligned} S &= 4\pi(R+r)^2 \text{ and} \\ V &= 4/3\pi(R+r)^3 , \end{aligned} \quad (2.28)$$

where R is the radius of the sphere and r is the radius of the gel fiber. Including the 2-D element for a spherical particle,

$$L = 2(R+r) , \quad (2.29)$$

and combining it with Eq. 2.27, leads to a similar result for expressing the electrophoretic mobility proposed by Ferguson [5]. Rodbard and Chrambach [55] state that from equations 2.25 and 2.26 one obtains for 1-D gels:

$$a \propto \frac{1}{\sqrt{T}} , \quad (2.30)$$

and for 0-D gels

$$a \propto \frac{1}{\sqrt[3]{T}} , \quad (2.31)$$

where T is the gel concentration and a is mean pore radius. By using eqn. 2.30 and eqn. 2.31, the following is obtained directly from eqn. 2.25 and eqn. 2.26 and has previously been established theoretically by Morris [94] and empirically by Ferguson [5],

$$\log \mu = \log \mu_o - K_r T , \quad (2.32)$$

where μ_o is the free solution mobility and K_r is the retardation coefficient and is constant for any specified charged molecule at constant percentage of cross-linking of the gel. The retardation coefficients for gels composed of a random meshwork of points (or 2-D gels) is

$$K_r = c_1(R + r) , \quad (2.33)$$

for gels composed of a random meshwork of fibers (or 1-D gels)

$$K_r = c_2(R + r)^2 , \quad (2.34)$$

for gels composed of a random meshwork of planes (or 0-D gels)

$$K_r = c_3(R + r)^3 , \quad (2.35)$$

where c_1 , c_2 , and c_3 are constants for any specified electrophoretic system.

This adaption of the Ogston model has remained the most successful way of analyzing gel electrophoretic mobility of small and relatively globular objects, i.e. DNA with length ≤ 500 bp. Ferguson plots [5], represented by plotting $\ln \mu$ vs. $\%T$, (see Fig. 2.8) are commonly used to determine the molecular weight of DNA and information about the gel structure. Figure 2.8 shows that in a typical Ferguson plot the log of the mobilities for smaller size molecules as a function of $\%T$ roughly converge to a point. The mobilities for larger molecules, which are not well described by the Ogston model, do not converge to the same single point. Since the mobilities for small fragments converge at $T = 0\%$, this point is referred to as the free solution mobility, μ_o . The slopes from the Ferguson plots are values for K_r . In the past, a linear plot of K_r versus R_g gives the effective radius of the gel fibers by applying a linear least-squares fit to the data (see ref [29] for calculation). Although past

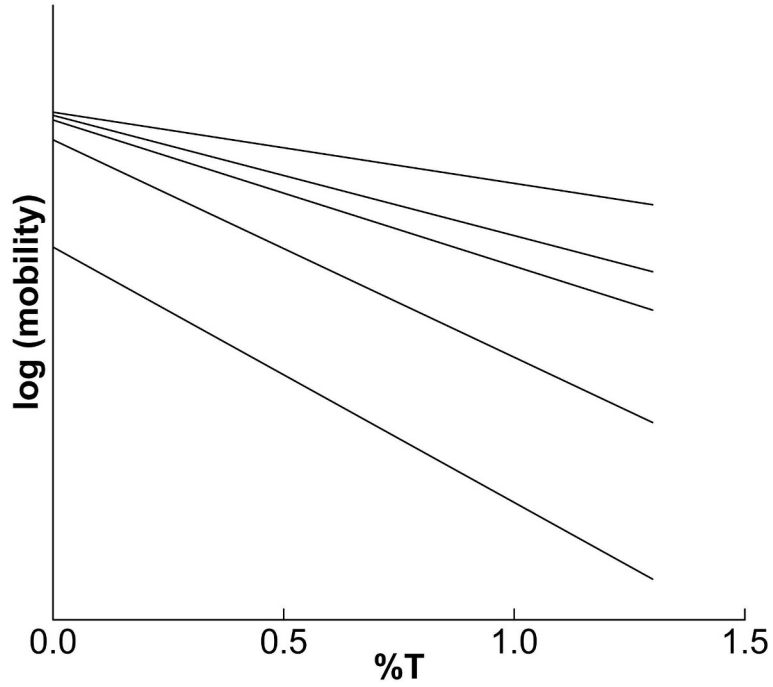


Figure 2.8. Ferguson Plot of mobility vs gel concentration ($\%T$). Each solid line represents the electrophoretic mobility for a specific dsDNA fragment of length, L_i where the fragments length increases as i increases.

results may give reasonable values [28–29], there have been experimental cases where nonlinear Ferguson plots occur [8] and raise questions about this model. Further modifications of this model and analysis of the difficulties which arise from it are addressed in references [95–98].

2.4.3 Reptation Models

For long DNA and for high gel concentrations the Ogston model and its extensions fail to describe DNA motion as it migrates through a gel during electrophoresis. Figure 2.4 sections C-E, illustrates that the DNA behaves under different mechanisms than described in sections A and B. As mentioned in the previous section, when the contour length, L , of the fragment is smaller than the persistence length, l_p , and also smaller than the mean pore size, a , the fragment can be thought of as a rigid rod and

the Ogston model will apply (Figure 2.4A) . When L is larger than l_p the fragment begins to reach a greater configurational entropy and takes the approximate shape of a globule with an effective radius of gyration, R_g , [87, 99] represented by

$$R_g^2 = \frac{l_p L}{3} \left(1 - \frac{l_p}{L} + \frac{l_p}{L} e^{-L/l_p} \right) . \quad (2.36)$$

When the fragment or chain is entangled with a gel, i.e., $R_g > a > l_k$, where l_k is the Kuhn segment of DNA, the DNA molecule can be considered as a flexible chain and the Ogston model does not apply.

P. G. de Gennes in 1971 introduced the idea of “reptation” [25] used to describe the motion of entangled polymers. He proposed a system of a single, ideal, polymeric chain, P , trapped inside a three-dimensional network, G , such as a polymeric gel. The chain P is not allowed to cross any of the obstacles; however, it may still move in a worm-like fashion between them. It was assumed that the length of the chain is very large when compared to the distance between neighboring links in the fixed network. This strongly restricts the possible chain motions. Then the only motions allowed correspond to the migration of certain “defects” along the chain (see ref [100]).

The reptation theory proposed by de Gennes did not consider any effects due to an external field, E . Lumpkin and Zimm [78] and Lerman and Frisch [101] independently theorized that the electrophoretic mobility, μ , should be inversely proportional to length, L , for polymers that behave in the reptation regime during DNA gel electrophoresis. Lumpkin and Zimm made a few key assumptions to derive the electrophoretic mobility from the reptation model. The polymer was assumed to experience a worm-like motion and also have a random polymer conformation described by Gaussian statistics. Also it was assumed that a frictional coefficient is proportional to the polymer length. The last essential assumption accounted for the counterion screening by replacing the polymer’s own charge by a overall effective charge, Q . An electric field, E , on a polymer in this system will have a total tangential electric force of

$$\frac{Q}{L} \sum \mathbf{E} \cdot \Delta \mathbf{s} = \frac{QEh_x}{L}, \quad (2.37)$$

where $\Delta \mathbf{s}$ is a length of the polymer at every point, L is the contour length, and h_x is the component of the polymer's end-to-end vector parallel to E . The velocity of the polymer in respect to it's center of mass, $\dot{X}_{cm}$, is

$$\dot{X}_{cm} = \frac{\dot{s}h_x}{L}, \quad (2.38)$$

where $\dot{s}$ is the polymer's mean velocity along a tube. The polymer's translational motion along a tube was described by relating equation 2.37 to the frictional coefficient, ξ , by

$$\xi \dot{s} = \frac{QEh_x}{L}. \quad (2.39)$$

By relating the averages of equations 2.38 and 2.39 they obtained

$$\langle \dot{X}_{cm} \rangle = \frac{\langle h_x^2 \rangle QE}{L^2 \xi}. \quad (2.40)$$

Since there initial assumptions allowed for both Q and ξ to be proportional to L , (Q/ξ) is L independent. From Gaussian statistics $\langle h_x^2 \rangle \sim L$. If the electrophoretic mobility can be represented by $\mu = v/E$, then

$$\mu = \frac{\langle \dot{X}_{cm} \rangle}{E} = \frac{\langle h_x^2 \rangle Q}{L^2 \xi} \sim \frac{1}{L}. \quad (2.41)$$

Initial experiments agreed with this finding [102–103].

A substantial improvement was made when the biased reptation model (BRM) was introduced by Slater and Noolandi in 1985 [26]. The theory suggests that the polymer chain is constrained by the gel structure to move in a “tube,” with the basic motion being a field-driven one-dimensional (1D) curvilinear diffusion (Brownian motion) along the tube axis with no lateral motions (see figures 2.4C and 2.4D).

Fig. 2.9a illustrates a polymer chain moving through a fixed network or an uniform gel matrix. The motion of the chain is shown to be confined within the obstacles which is similar to a tube-like region (Fig. 2.9b). “Defects” described by de Gennes [25] are fluctuating around the tube axis (the dotted line in fig. 2.9). A tube model is represented for this system in fig. 2.9b without the network.

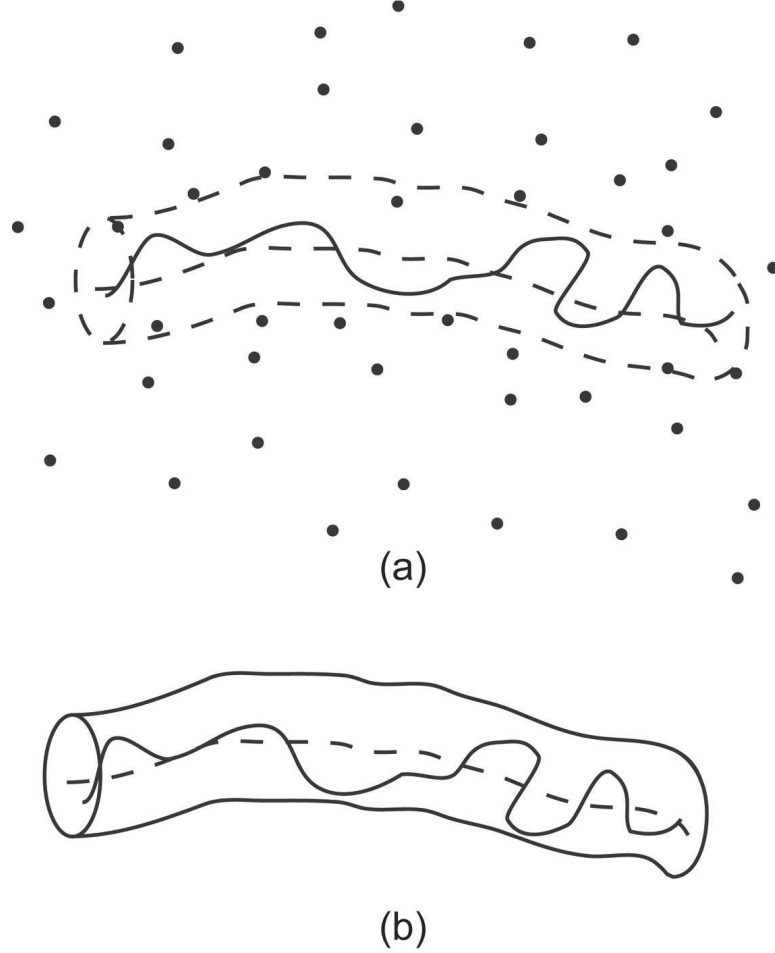


Figure 2.9. The Tube model. (a) A polymer chain in a fixed network of obstacles (denoted by dots). The dashed line represents the tube axis. (b) The Tube Model.

The BRM considers the DNA molecule as a flexible chain of blobs of size a . The tube length, L will be $L = Na$, where N is the number of blobs per chain. An

electric field, E , induces additional directional drift of the chain. From equation 2.17 the mobility is $\mu = \langle v \rangle / E$. The BRM substitutes for velocity by

$$v = \frac{v_d h_x}{Na}, \quad (2.42)$$

where $v_d = \varepsilon \frac{h_x}{N\tau_o}$, h_x is the projection of the end-to-end vector onto the field direction, τ_o is the blob time, and $\varepsilon = qEa/Tk_B$ which is the reduced field parameter with q being the effective charge per blob, T is the temperature, and k_B is the Boltzmann's constant. By substituting for v and E the mobility for the BRM becomes,

$$\mu = \frac{\langle v_d h_x \rangle}{NaE} = \mu_o \frac{h_x^2}{N^2 a^2}, \quad (2.43)$$

where $\mu_o = q/f_c$, with f_c being the effective friction coefficient per blob. For weak electric fields the chain's statistics are Gaussian and the end-to-end distance is,

$$h_x = \frac{Na^2}{3}. \quad (2.44)$$

Stronger fields will induce some orientation of the leading blob, so the tube will be elongated in the field (see Fig. 2.4E). This orientation is assumed to be governed by the Boltzmann factor, $\exp(-\varepsilon \cos \theta)$, where θ is the angle between the field direction and blob orientation. Reference [30] gives an in depth calculation for the effects of weak and strong fields on electrophoretic mobility using equations 2.44 and 2.43. The results is that

$$\frac{\mu}{\mu_o} \sim \begin{cases} N^{-1} & N < N^* \\ \varepsilon^2 & N > N^*, \varepsilon < 1 \end{cases}, \quad (2.45)$$

where $N^* \sim \varepsilon^{-2}$ is the crossover chain length from unoriented reptation to oriented reptation [8, 30]. Oriented and unoriented reptation is illustrated in figure 2.4C and 2.4E. Although the BRM worked for some initial experiments [29], it has been found that the effect of tube-length fluctuations is essential and leads to a qualitatively different answer for the mobility as a function of length [30, 65, 104].

To address this issue the model of biased reptation with fluctuations (BRF) was introduced [30, 65, 104]. The BRF model takes into account the actual motion of the chain which is not merely an electrophoretic drift along the tube, but rather is a superposition of the drift and 1D Rouse Motion [45]. In the 1D Rouse model the polymer is represented by a set of beads connected along a chain. The segments between the beads act as springs (see ref [100]). The central point of the BRF approach is that it is the terminal section (or crossover point from a Gaussian conformation to a conformation oriented with field) of blobs, m^* , rather than single end blob that must be in quasiequilibrium with the electric field. Therefore, the terminal section must be slightly stretched by the total electric force $F = qEm^*$. It was observed [30] that the tube-length fluctuations significantly enhance the orientation of the molecule under the influence of a weak field ($\varepsilon \ll 1$). For $N \leq N^*$ tube-length fluctuations and orientation effects from the field are negligible. The BRF model predicts the following mobilities,

$$\frac{\mu}{\mu_o} \sim \begin{cases} N^{-1} & N < N^* \\ \varepsilon & N > N^*, \varepsilon < 1 \end{cases} \quad , \quad (2.46)$$

where $N^* \sim \varepsilon^{-1}$. Experiments have shown that the BRF model is currently the most accurate representation for molecules in the reptation regime [64, 105]. So according to this theory, long molecules should behave like equation 2.46. A more in depth calculation and explanation of the BRM and BRF model are provided in ref [30].

For a long fragment in a large electric field, the electric forces lead to the fragment choosing a tube consisting of consecutive pores that do not follow a random walk in space. The average molecular conformation of the fragments is referred to as “stretched”, hence this regime is called the Reptation Regime with Stretching [27, 106]. The mobility of the polymer becomes independent of its length. Figure 2.10 illustrates a molecule experiencing a form of reptation with stretching. As the field becomes larger the molecule’s conformation is stretched parallel to its motion. Stretching can also occur when a long length molecule has chosen a path

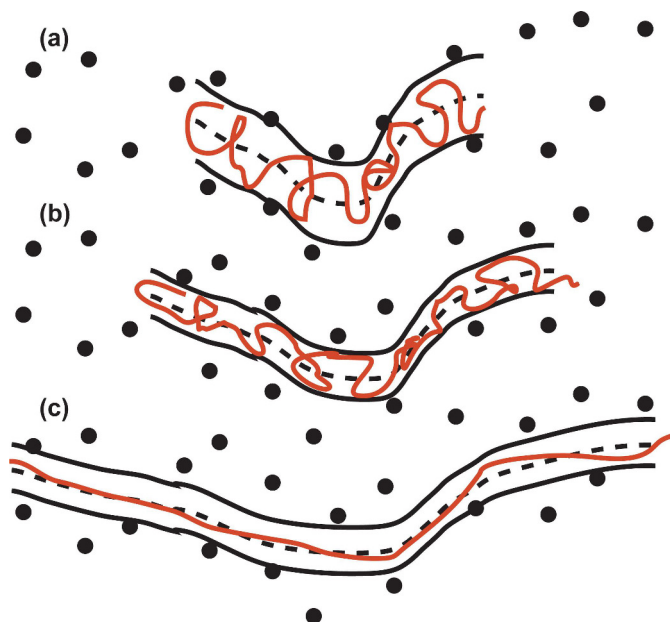


Figure 2.10. A polymer experiencing reptation with stretching. The solid black lines represent the tube model. The dotted black line is the tube axis. The thicker red line represents the polymer moving through the gel according to the tube model. The top illustration (a) shows a relaxed polymer in an electric field. The polymer starts to stretch in (b) and finally the polymer will be experiencing reptation with “stretching” in the lower panel (c).

that wraps around a fiber. This is illustrated in figure 2.4D. Due to the high fields both ends of the fiber will be experiencing the same force. The molecule will get stretched at both ends. Eventually the molecule will become free of the fiber and migrate through the gel.

2.4.4 Entropic Trapping

For a flexible polymer at low fields and concentrated gels there exists another mechanism that occurs between the Ogston regime and the reptation regime called entropic trapping (ET). This was first reported by Baumgartner and Muthukumar by computer simulations of polymer diffusion in a random network [19–20, 107]. It was experimentally confirmed later by electrophoresing single-stranded (ssDNA) in

polyacrylamide gels by Rousseau et al. [58]. ET occurs when the polymer's radius of gyration $R_g(L)$ is comparable to the gel's mean pore size, a . The molecule then selectively hops between the largest pores which thus act like “entropic traps” where the chain maximizes its conformational entropy. Figure 2.4B illustrates this behavior in a random gel network. Rousseau et al. demonstrated that at strong fields entropic-trapping (ET) was eliminated, and that denser gels favor ET, because the gels are more heterogeneous. It was shown that for the ET regime the mobility scales as [58],

$$\mu \sim 1/L^{1+\gamma} \text{ with } \gamma > 0, \quad (2.47)$$

where for large molecules the exponent γ can be a measure the entropic effect, with the reptation limit being $\gamma = 0$.

By defining L_a as the molecular size of an unperturbed DNA molecule for which $R_g(L_a) = \bar{a}$, L_a can be used as a measure of the mean pore size $\bar{a}$. The Ogston/ET transition is observed as a maximum at $L \approx L_a$. In this case, L_a is the size of the largest molecule that can migrate through a percolating pathway made of pores of sizes $a \geq R_g(L_a)$. The length associated with the Ogston/ET transition, L_{aET} , decreases for larger concentrations. This suggest that low concentration gels contains percolating pathways made of large pores. While high concentrations contains a distribution of pore sizes with some isolated pores large enough to allow ET. The critical electric field, $E^*(L)$, necessary to force the molecules to reptate appears to decrease as the molecular size, L , increases.

2.4.5 Other Models

Throughout the years others have proposed different fitting functions and models that have been valid for certain conditions. These models provided more insight into the understanding of the DNA gel electrophoresis. Calladine, et al. [36] proposed

that the migration of DNA through various types of gels during electrophoresis can be characterized by the empirical equation

$$\frac{v}{f} = \frac{1}{1 + \mathcal{L} + 0.01\mathcal{L}^3} , \quad (2.48)$$

where v is the electrophoretic velocity, f relates to half the maximum velocity, $v_{0.5}$, by $v_{0.5} = 0.5f$, and $\mathcal{L} = L/L_{0.5}$ with $L_{0.5}$ being the length associated with $v_{0.5}$. They claimed that equation 2.48 can fit over a DNA length range from ~ 10 bp to $\sim 50,000$ bp, gel concentrations ranging from 0.5% to 12% for two different type of gels, and electric fields ranging from 0.5 to 8.0 V/cm. They claimed that their function represented the data over a wide range. Their figures show that equation 2.48 seems to fit only 2/3 of their total length range.

Southern's work [10, 108] sometimes has been cited as evidence for DNA reptation; that is, Southern is said to have shown that $\mu = \alpha/L$. Southern actually showed that plots of reciprocal mobility versus length were linear over a wider range than traditional semi-log plots (μ vs. $\log(L)$). Plots of $1/\mu$ vs. L eventually deviate from linearity, but Southern also proposed an empirical relationship that extended the linear range, namely

$$L = \frac{\alpha}{\mu(L) - \mu_{\circ}} + \beta' , \quad (2.49)$$

which can be re-written

$$\mu(L) - \mu_{\circ} = \frac{\alpha}{L - \beta'} . \quad (2.50)$$

Plots of $\mu(L)$ vs. $1/L$ will not be linear unless $\beta' = 0$.

Barkema, et al. [35] have shown that some data collapse to a single curve when plotted as $T^{5/2}vL^2$ versus TLE where v is the drift velocity. They proposed a two parameter fit of the collapsed data in the form

$$\frac{T^{5/2}vL^2}{\alpha} = [(\frac{TLE}{\beta})^2 + (\frac{TLE}{\beta})^4]^{1/2}, \quad (2.51)$$

where α and β are independent free parameters. The collapsed data they analyzed fit with $\beta = 27 \pm 5$ and $\alpha/\beta = 1.0 \pm 0.1$ at 20°C.

Kozulić proposed the “door-corridor” model of gel electrophoresis [73–77]. This model relates the migration of the molecules through gels to the dislocation of the gel polymers by the migrating molecules. The model proposes that before the presence of an electric field there are no defined spaces a molecule can enter the gel. During gel electrophoresis the migrating molecule push the polymers aside to clear a space to occupy. The “door-corridor” model relates the electrophoretic mobilities of a molecule in a gel as a function of the ratio between two forces, the gel resistance (F_r) and electrokinetic force (F_e):

$$\mu = \mu_1 \exp(-\frac{F_r}{F_e}), \quad (2.52)$$

where μ_1 is the mobility of smallest segment of the migrating molecule (this is the mobility associated to a 1 bp fragment for dsDNA).

Although these models apparently worked for the specific research reported, it was found through this research that they all have limitations which were not obvious when originally reported.

2.4.6 Motivation

Despite widespread use of DNA gel electrophoresis and extensive theoretical development, important questions remain about the mechanism(s) by which the DNA migrates through a gel during electrophoresis, and how the electrophoretic mobilities of DNA can be represented over a wide range of lengths. These issues will be the focus of this dissertation. In order to reexamine these issues DNA separations were performed by gel electrophoresis on agarose gels. Experiments were performed for

an extensive range of DNA lengths, electric fields, and different gel concentrations. Several methods for fitting the resulting data were tested. This led to the conclusion that none of the standard techniques for fitting were satisfactory over the entire range conditions used. A simple phenomenological model was developed that was found to fit our data exceptionally well for a large range of gel concentrations and electric fields [32–34].

The data were fit to a relationship that can be represented in most general form by

$$\mu(L) = \frac{1}{\beta + \alpha(1 - \exp(-L/\gamma))} , \quad (2.53)$$

where α , β , and γ are adjustable parameters.

Equation 2.53 expresses an exponential crossover from the mobilities of short DNA fragments to the mobilities of long fragments. An intuitively more appealing representation of this exponential crossover is obtained by substituting the limiting mobilities into Equation 2.53. The limiting mobility, μ_s , describes the mobility of a particle with charge q in the limit of zero length, ie.,

$$\lim_{L \rightarrow 0} \mu(L) \equiv \mu_s . \quad (2.54)$$

Note that μ_s is not equivalent to the free solution mobility of long DNA often quoted in other studies and represented by μ_o . In the limit of $L \gg \gamma$ the exponential term is negligible and the limiting mobility μ_L is

$$\lim_{L \rightarrow \infty} \mu(L) \equiv \mu_L . \quad (2.55)$$

Inverting equation 2.53 and substituting from equations 2.54 and 2.55 yields

$$\frac{1}{\mu(L)} = \frac{1}{\mu_L} - \left(\frac{1}{\mu_L} - \frac{1}{\mu_s} \right) e^{-L/\gamma} , \quad (2.56)$$

or, in particularly simple form,

$$\frac{\frac{1}{\mu_L} - \frac{1}{\mu(L)}}{\frac{1}{\mu_L} - \frac{1}{\mu_s}} = e^{-L/\gamma} . \quad (2.57)$$

The reciprocal mobility expresses the resistance that the gel matrix offers to passage of solutes; thus it is the decrease in gel resistance with decreasing DNA length from the long chain limit that obeys a simple exponential. The parameter γ , which has dimensions of length, determines the rate of decrease. In the following chapters this equation will be applied to experimental data for many different conditions.

2.5 Summary

Understanding the basis behind the migration of DNA molecules through a network plays an integral role in the world of biophysics and biochemistry. Gel electrophoresis is the most widely used technique to separate DNA molecules. Several theories have been used to describe migration of DNA through a gel network during electrophoresis. Ogston [24] described the motion of a molecule, acting like a rigid rod, migrating through spaces in an uniform random suspension of fibers. The Ogston model was later applied to gel electrophoresis by Rodbard and Chrambach [55]. Molecules behaving as a long random coil were described by a different mechanism referred to as “reptation” by de Gennes [25]. The Biased Reptation Model (BRM) was introduced by Slater and Noolandi [26] to apply the reptation model to gel electrophoresis. This model described the entire DNA chain to migrate within a tube defined by the path through the gel of the first chain segment by an electrophoretic drift. The chain will migrate along the tube’s axis biased by the field direction, but the random distribution of chain configurations is considered to be unperturbed during reptation. The biased reptation with fluctuations (BRF) model [30] modified the BRM model to account for the effect of tube-length fluctuations which leads to a qualitatively different answer for the mobility as a function of length. Entropic trapping [107] accounted for a regime observed between the Ogston model and

reptation. When the size of the DNA fragment is equivalent to the largest pore size of the network, the molecule will selectively hop between the largest pores. This will act as an entropic trap, because the molecule maximizes its conformational entropy in large pores. Other theories and models have been proposed [8] which describe the motion of DNA during electrophoresis for limited conditions.

A new simple phenomenological equation used to describe the migration of the DNA during gel electrophoresis [32] has been discovered. This equation describes the behavior of the mobility of DNA over more extensive ranges of DNA sizes, gel concentrations, and electric fields than other descriptions. In the following chapters data will be presented to show the usefulness of this equation.

CHAPTER 3

MATERIALS AND METHODS

Electrophoresis experiments were used to separate dsDNA in agarose gels. DNA samples were obtained from commercial suppliers and were synthesized to contain a distribution of known fragment sizes. This type of sample is referred to as a DNA ladder. Ladders have only recently become available and the measurements reported here are some of the most complete performed to date. These standards allowed for easy analysis of electrophoretic mobilities of DNA fragments at different lengths. The same type of agarose was used throughout this work to allow for comparisons for each experiment. Conventional electrophoresis was performed on the DNA samples in agarose gels. Two different apparatuses were used. The first apparatus had a horizontal geometry and the second apparatus was vertical.

3.1 Materials

3.1.1 DNA Samples

Different set of mixtures were used for each different set of experiments. For the first set of experiments the DNA samples were a mixture of $N \times 100$ bp and $N \times 1000$ bp “ladder” standards ($N = 1 \dots 10$) from Promega Corp (Madison, WI, USA).

The second set of electrophoresis experiments used DNA mixtures of $N \times 100$ bp (100 bp PCR Molecular Ruler), $N \times 1000$ bp (1 kb Molecular Ruler), $N \times 2500$ bp (2.5 kb Molecular Ruler) “ladder” standards from Bio-Rad Laboratories (Hercules, CA, USA) and a lambda phage DNA from Sigma (St. Louis, MO, USA). The 2.5 kb

Molecular Ruler was found to extend in practice to 47.5 kb, well beyond the 35 kb limit reported by the supplier.

The third set of electrophoresis experiments used a similar DNA mixture of $N \times 100$ bp (100 bp PCR Molecular Ruler), $N \times 1000$ bp (1 kb Molecular Ruler), and $N \times 2500$ bp (2.5 kb Molecular Ruler) “ladder” standards from Bio-Rad Laboratories (Hercules, CA, USA) along with a lambda phage DNA from Promega (Madison, WI, USA). A Pulse Field Marker from Sigma was also used (St. Louis, MO, USA) which contained fragments (in kbp): 194.0, 145.5, 97.0, 48.5, 23.1, 9.4, 6.6, 4.4, 2.3, 2.0, 0.6, and 0.1. Typically the only fragments that can be resolved are 194000 bp and 23100 bp. Occasionally with some electric fields and gel concentrations 9400 bp, 6600 bp, and 4400 bp also can be resolved. Although it is assumed that the last band is 194000 bp, the sample also contains 145000 bp, 97000 bp, and 48502 bp. For these experiments the separations between these molecules can not be distinguished. For this reason it is assumed that the band will be called 194000 bp.

3.1.2 Agarose Gels

Gels of Seakem LE agarose (Whittaker BMA Bioproducts, Rockland, ME, USA) were cast in $1 \times$ TBE buffer (obtained by dilution of a standard stock solution of $10 \times$ TBE which is 0.89 M TRIS, 0.89 M boric acid, 20 mM Na_2EDTA , 1mM cacodylic acid, pH ~ 8.2 (components from Sigma-Aldrich, St. Louis, MO, USA)). All solutions were prepared using water purified by reverse osmosis and further deionized by passage through a Milli-Q Plus Purification Pak (Millipore Corp., Bedford, MA, USA).

3.2 Methods

Three different set of experiments were performed for this project. This section will be divided into three subsections, one for each experiment. The first set of

experiments involved performing horizontal gel electrophoresis at a single field with varying concentration. Horizontal gel electrophoresis was also performed for the second set of experiments with varying electric field and gel concentration. The last experiments were performed on a vertical electrophoresis apparatus with varying field and gel concentration.

3.2.1 Horizontal Electrophoresis with Uniform Electric Field

The electrophoretic mobilities, $\mu(L)$, of fragments of “100 bp ladder” and “1000 bp ladder” standards were determined for gels of eleven agarose concentrations of $T = 0.5, 0.6, 0.7, \dots, 1.5\%$. The DNA sample used was a mixture with $5\mu\text{L}$ of 100 bp ladder, $2\mu\text{L}$ of 1000 bp ladder, $2\mu\text{L}$ ddH₂O, and $2\mu\text{L}$ of blue/orange 6 $\times$ loading dye solution (provided by Promega (Madison, WI, USA)) for tracking the progress of the DNA. The sample was placed in a well in the gel for electrophoresis.

Gels were cast in a horizontal gel holder (see Fig. 3.1) after the agarose was dissolved in $1\times$ TBE buffer by boiling in a microwave oven. Constant volume aliquots from a single stock solution of $10\times$ TBE buffer prepared in deionized water were used for all experiments. Gel concentrations are expressed in $\%T = (\text{grams agarose}/100 \text{ mL buffer})\times 100$. On average approximately 4 mL of water evaporated per 170 mL of gel solution. This was not replaced, nor measured individually. Thus the actual gel concentrations were higher by $\leq 3\%$ than the intended concentrations reported. The gels were cast and allowed to equilibrate at ambient temperature ($\sim 22^\circ\text{C}$) for one hour before use as suggested by the supplier. The gel holder had two evenly spaced dividers so samples could be run simultaneously on gels of three different concentrations. Gels were 25 cm long with an effective thickness of ~ 4 mm. A comb was placed at the top of the gel to allow the formation of $0.5 \times 0.1 \times 0.3$ cm wells.

Electrophoresis experiments were performed in horizontal gel geometry at ambient temperature ($\sim 22^\circ\text{C}$). Experiments were performed using a chamber purchased from Bethesda Research Laboratories, Inc. (Rockville, MD, USA) with a digital

power supply (Bio-Rad Model 100/500 Constant Voltage Power Supply, Hercules, CA, USA) (see Fig. 3.1). Approximately 3 liters of $1\times$ TBE buffer filled the electrode trays to a level approximately 1 mm above the gel. During electrophoresis the applied voltage was controlled to produce a constant field of 2.8 V/cm to reduce heating and distortions of DNA conformation induced by strong fields [29, 43]. All experiments were performed with the same equipment, using constant volume aliquots from a single stock solution of TBE buffer for all gels and tray buffers. Gels were post-stained with ethidium bromide (0.5 mg/L) and then photographed using Polaroid Type 667 film. Distances from the well bottom to band midpoints, measured directly from photographs, were used to calculate mobilities. The following equation was used to calculate the mobilities,

$$\mu = \frac{v}{E} = \frac{d}{t} \frac{l}{V}, \quad (3.1)$$

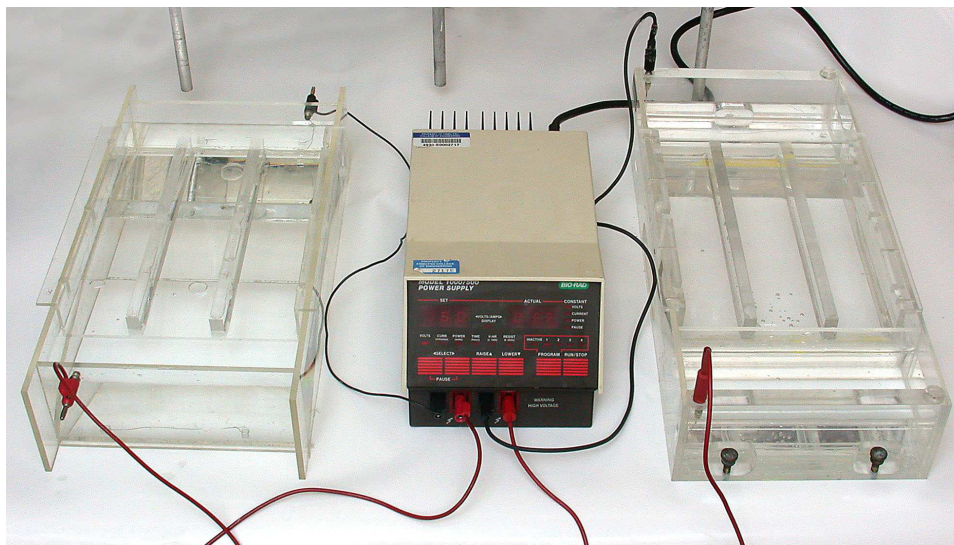


Figure 3.1. Setup for horizontal gel electrophoresis. The same digital power supply (Bio-Rad Model 100/500 Constant Voltage Power Supply, Hercules, CA, USA) was used for both electrophoresis chambers. The Electrophoresis chamber on the right was purchased from Bethesda Research Laboratories, Inc. (Rockville, MD, USA) and the chamber on the left was locally constructed.

where v is the velocity of the DNA fragments, E is the electric field applied to the system, d is the measured distance from the band to the well, t is the total time of electrophoresis, l is the distance between the electrodes, and V is the voltage applied to the system. Precision in mobilities varied from $\pm 0.6\%$ to 1.6% (shortest to longest fragments), on 0.5% gels; and from $\pm 0.8\%$ to 6% on 1.5% gels.

3.2.2 Electric Field Dependence (Horizontal Geometry)

DNA ladders were mixed into two different samples for good visualization of each fragment. The first DNA sample was a mixture with $5\mu\text{L}$ of 2.5 kb Molecular Ruler, $3\mu\text{L}$ of 100 bp PCR Molecular Ruler, $2\mu\text{L}$ ddH₂O, and $2\mu\text{L}$ of blue/orange $6\times$ loading dye solution (provided by supplier) for tracking the progress of the DNA. The second DNA sample was a mixture with $4\mu\text{L}$ of 1 kb Molecular Ruler, $4\mu\text{L}$ of lambda phage DNA, $2\mu\text{L}$ ddH₂O, and $2\mu\text{L}$ of blue/orange $6\times$ loading dye solution. The two samples were placed in adjacent wells in each gel for electrophoresis. Gels were cast by the same method as in section 3.2.1.

Electrophoresis experiments were performed in horizontal gel geometry at ambient temperature ($\sim 22^\circ\text{C}$). Experiments were performed using two similar chambers operated simultaneously from the same power supply (Bio-Rad Model 100/500 Constant Voltage Power Supply, Hercules, CA, USA). One chamber was locally constructed (Fig. 3.1) and the other was purchased from Bethesda Research Laboratories, Inc. (Rockville, MD, USA) (see Fig. 3.1). Approximately 3 liters of $1\times$ TBE buffer filled the electrode trays to a level approximately 1mm above the gel. Electrophoresis was performed at constant voltage over a range of fields from 0.71 to 5.00 V/cm. Although the dimensions of the two chambers were similar, the electrodes had different separations, producing somewhat different fields. Gels were post-stained with ethidium bromide (0.5 mg/L) and then placed on a UV plate. The ethidium bromide intercalates with the DNA. The gels are placed on a UV plate and the DNA bands fluoresce due to ethidium binding. The gels were photographed using

Polaroid Type 667 film. Several photographs were taken of each gel to optimally record different regions. Figure 3.2 shows an example of a photographed gel.

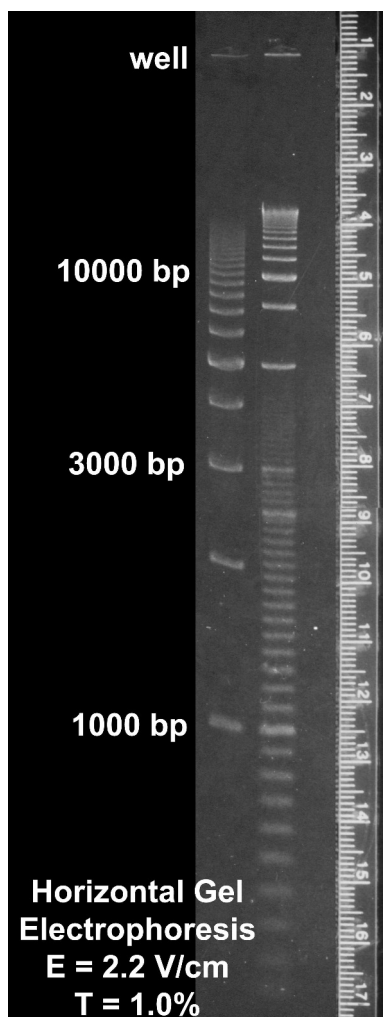


Figure 3.2. Horizontal Gel Electrophoresis at $E = 2.2$ V/cm and $T = 1.0\%$. Intermittent DNA fragments are labelled.

The photographs were digitized using a conventional scanner and displayed on a computer screen as plots of optical density versus distance using GIMP (GNU Image Manipulation Program, MindVision Software, Lincoln, NE, USA). Distances from the well bottom to band intensity maxima were used to calculate mobilities. Equation 3.1 was used to calculate the mobilities. Precision in mobilities varied from

$\pm 0.6\%$ to 1.6% (shortest to longest fragments), on 0.5% gels, and from $\pm 0.8\%$ to 6% on 1.5% gels. Lines shown in plots are non-linear least squares fits to Eq. 2.56 generated by MathCad (MathSoft, Inc., Cambridge, MA, USA). Gel temperatures were measured before and during electrophoresis using an electronic thermistor thermometer. Electrophoretic mobilities were calculated based on the run time and voltage gradient. Mobilities were corrected for the temperature dependence of the viscosity of water by multiplying the observed mobility by the ratio of the viscosity of water at the temperature measured at the end of the run to the viscosity at 20°C . The final temperature of the gel usually was higher than the initial temperature, primarily due to resistive heating. The increase was relatively small ($\leq 2^\circ\text{C}$) for all fields ≤ 3.1 V/cm. At $E = 3.5$ V/cm the temperature increased by 3°C and at $E = 5.0$ V/cm the temperature increased by 7°C . The viscosity corrections did not seem to introduce a bias. If these corrections adversely impacted the ability of Eq. 2.56 to fit the mobility data, the largest deviations from fits to Eq. 2.56 would be observed for data taken at the higher voltages, not the lower voltages, contrary to the observed results.

3.2.3 Electric Field Dependence (Vertical Geometry)

DNA ladders were mixed into three different samples similar to section 3.2.2. The first DNA sample was a mixture with $5\mu\text{L}$ of 2.5 kb Molecular Ruler, $3\mu\text{L}$ of 100 bp PCR Molecular Ruler, $2\mu\text{L}$ ddH₂O, and $2\mu\text{L}$ of blue/orange $6\times$ loading dye solution (provided by supplier) for tracking the progress of the DNA. The second DNA sample was a mixture with $4\mu\text{L}$ of 1 kb Molecular Ruler, $4\mu\text{L}$ of lambda DNA, $2\mu\text{L}$ ddH₂O, and $2\mu\text{L}$ of blue/orange $6\times$ loading dye solution. The third DNA sample was a Pulse Marker from Sigma. The sample is embedded in 1% low melting point agarose. For convenience it is supplied in a graduated syringe. The agarose is extruded from the syringe and with a clean razor a 1 mm slice is taken from the end. Since the agarose is low melting the slice is put in a capillary tube and placed in a water bath at $\sim 60^\circ$.

Once the agarose has melted, 10 μ L ddH₂O and 2 μ L of blue/orange 6 $\times$ loading dye solution is added to the sample for tracking purposes. The three samples were placed in adjacent wells in each gel for electrophoresis.

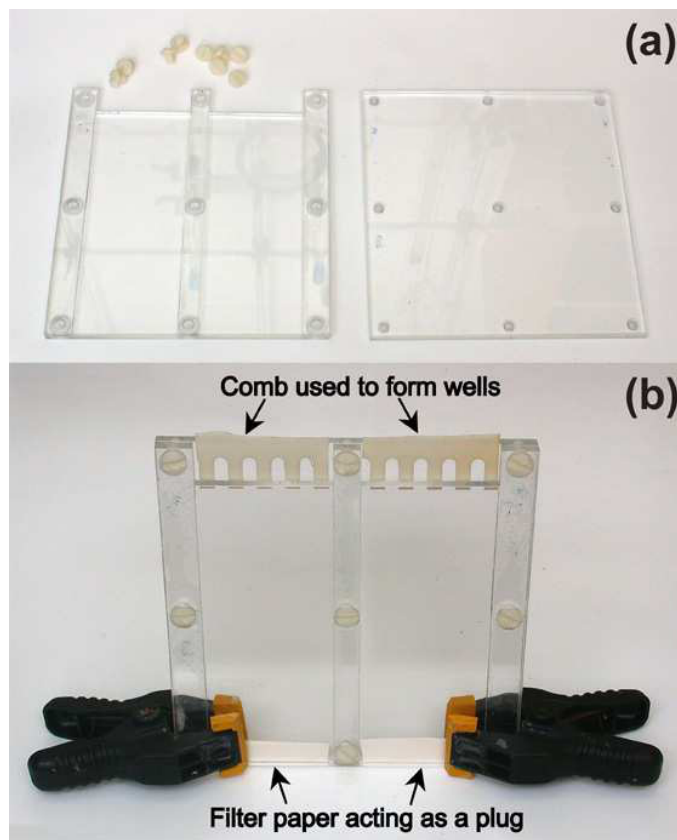


Figure 3.3. Vertical gel holder for vertical electrophoresis experiments. (a) The disassembled gel holder. This holder consists of two parts and nine screws. The plates are screwed together so they are leak proof. (b) The assembled gel holder. Filter paper is put on the bottom of the holder to prevent the gel from leaking out during gelation. The comb is placed on the top to form wells in the gels.

Gels were cast in a vertical gel holder (see figures 3.3 and 3.4) after the agarose was dissolved in 1 $\times$ TBE buffer by boiling in a microwave oven. New vertical gel holders were designed for these experiments made from plexiglas. Two similar holders were made with three spacers $\sim$ 1 mm thick which are permanently attached. The design schematic (figure 3.4) shows that a spacer in the middle evenly divides the holder

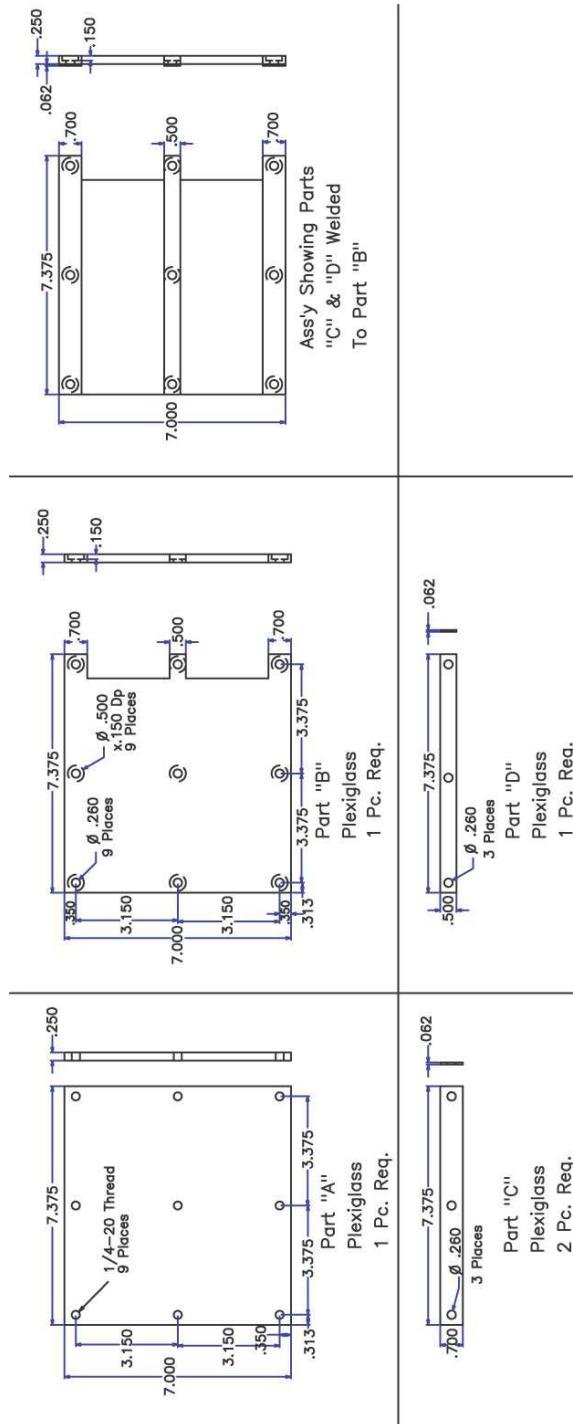


Figure 3.4. The Design Schematic of the vertical gel holder in figure 3.3.

into two sections to run gels simultaneously with two different concentrations. Before casting the gels, filter paper is put at the bottom of the holder to act like a plug. The holders are screwed together and tape is put at the bottom of each holder to prevent the gel from leaking during gelation. Constant volume aliquots of $10\times$ TBE buffer prepared in deionized water were used for all experiments. Gel concentrations are expressed in $\%T = (\text{grams agarose}/100 \text{ mL buffer}) \times 100$. The agarose was dissolved in $1\times$ TBE buffer and allowed to cool to $\sim 60^\circ\text{C}$. The evaporated water was replaced and the gels were cast and allowed to equilibrate at ambient temperature ($\sim 22^\circ\text{C}$) for one hour before use as suggested by the supplier. Gels were 18 cm long with an effective thickness of $\sim 1 \text{ mm}$. A comb was placed at the top of the gel to allow the formation of $0.7 \times 0.1 \times 0.3 \text{ cm}$ wells.

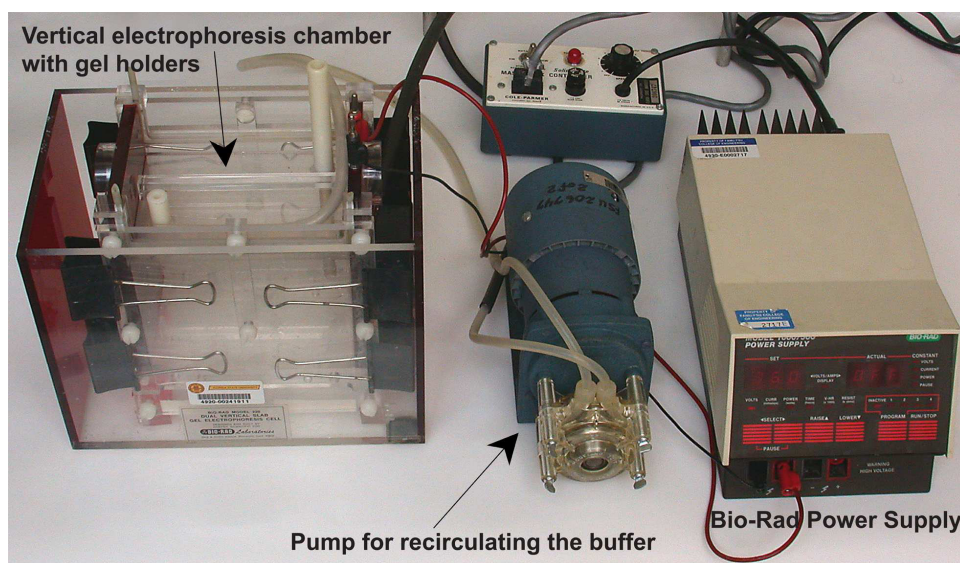


Figure 3.5. The setup for vertical electrophoresis experiments. The electrophoresis chamber contains the gel holder with approximately 1.5 liters of $1\times$ TBE buffer filling the bottom of the chamber and 0.5 liters of $1\times$ TBE buffer filling the top. The pump recirculated the buffer from the bottom of the chamber to the top. The power supply is a Bio-Rad Model 100/500 Constant Voltage Power Supply (Hercules, CA, USA).

Electrophoresis experiments were performed in vertical gel geometry at ambient temperature ($\sim 22^\circ\text{C}$). Experiments were performed using a chamber made by

Bio-Rad (Hercules, CA, USA), operated by a digital power supply (Bio-Rad Model 100/500 Constant Voltage Power Supply, Hercules, CA, USA) with two holders put on each side of the vertical electrophoresis chamber so four gel concentrations can be run simultaneously (see Fig. 3.5). Approximately 1.5 liters of $1\times$ TBE buffer filled the bottom of the vertical electrophoresis chamber tray and 0.5 liters of $1\times$ TBE buffer filled the top. A pump was used to recirculate the buffer during the electrophoresis experiment to keep steady temperature and a uniform pH. Electrophoresis was performed at constant voltage over a range of fields from 0.62 to 6.21 V/cm. Gels were post-stained with ethidium bromide (0.5 mg/L) for one hour and then photographed using the Bio-Rad Gel Doc 2000 gel documentation system (Hercules, CA, USA). The Gel Doc 2000 is a high performance gel imaging system which uses a CCD camera to capture images in real time, with a motorized zoom lens for convenient zoom, focus, and iris adjustments (see Fig. 3.6). Using software provided from Bio-Rad, several images were acquired of each gel to optimally record different regions. Figure 3.7



Figure 3.6. Bio Rad's Gel Doc 2000 gel documentation system. A CCD camera is attached to the top of the apparatus. The whole apparatus is attached to a computer with software provided by Bio-Rad to acquire an image of the gel. Picture acquired from Bio-Rad web site (www.biorad.com).

shows gels run at $E = 2.48$ V/cm for two different concentrations run at the same time.

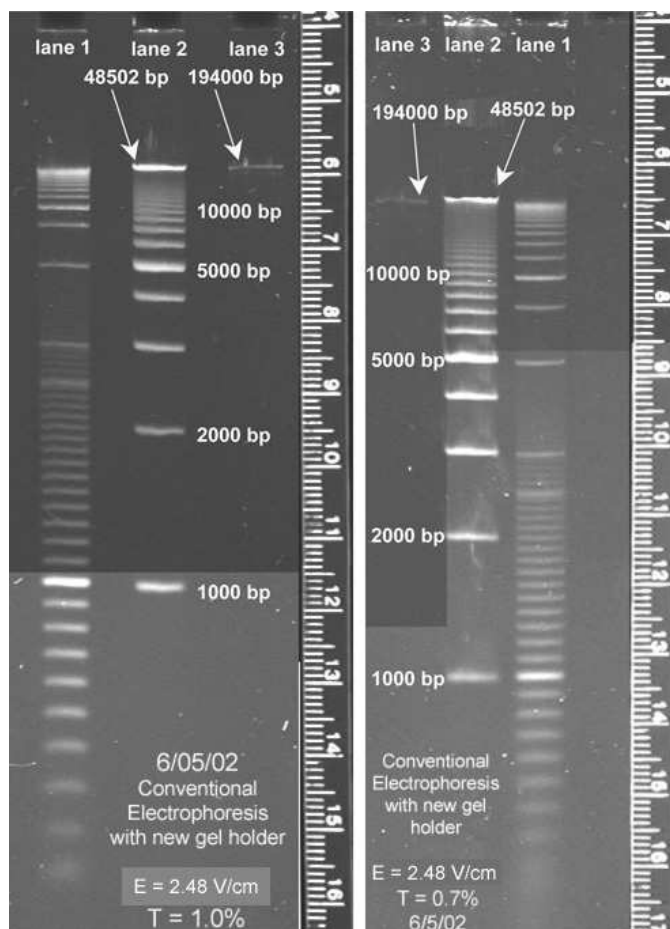


Figure 3.7. Vertical Gel Electrophoresis at $E = 2.48$ V/cm for $T = 0.7\%$ and 1.0% . The image was acquired using the Gel Doc 2000.

The mobilities of the DNA fragments were gathered and calculated using the same methods described in section 3.2.2. Gel temperatures were measured before and during electrophoresis using an electronic thermistor thermometer. Electrophoretic mobilities were calculated based on the run time and voltage gradient. Mobilities were corrected for the temperature dependence of the viscosity of water by multiplying the observed mobility by the ratio of the viscosity of water at the temperature measured

at the end of the run to the viscosity at 22°C. The temperature of the gel after electrophoresis usually either stayed the same or increased by $\leq 1.5^{\circ}\text{C}$ for all fields. The mobilities measured were reproducible to within approximately 6%.

CHAPTER 4

RESULTS

The primary goal of this project is to reexamine the mechanisms by which the DNA migrates through a gel during electrophoresis. As discussed in chapter 1, gene mapping, sequencing and other aspects of genomics depend on accurate determinations of double stranded DNA fragment lengths. DNA separations according to length are commonly accomplished by electrophoresis in agarose gels. Length determination depends on observations, first made over 30 years ago [10–11, 109–110], that the electrophoretic mobility of DNA in agarose gels decreases to some non-zero limit with increasing chain contour length, regardless of base sequence. Despite widespread use and extensive theoretical development, important questions remain (e.g. see the review by Viovy [8] and references therein).

In order to reexamine these issues a series of electrophoretic separations were first performed at varying agarose concentrations at a uniform electric field ($E = 2.78$ V/cm) and several methods for fitting the resulting data were tested. This led to the conclusion that none of the standard techniques for fitting were satisfactory for DNA molecules ranging from 100 to 10000 base pairs. A simple phenomenological model was developed that was found to fit our data exceptionally well over the entire range of DNA sizes. Section 3.2.1 discusses the first set of materials used and experiments performed.

The simplicity of this function suggests that some commonality dominates DNA electrophoresis over a larger size range than expected, at least under certain conditions. New data were gathered to describe the length dependence of mobility

over a range of electric fields and gel concentrations. Additional electrophoresis data were obtained at gel concentrations from 0.5% to 1.3%, and several fields from 0.71 to 5.00 V/cm. The range of DNA lengths examined was also extended to approach the regime of length-independent mobilities by using a 2.5 kbp DNA ladder (maximum length 47.5 kbp) and a phage lambda DNA (48.5 kbp).

These additional measurements showed that equation 2.56 fits the data well for high fields and low gel concentrations within the ranges examined. For example, data taken at $T = 0.5\%$ yielded excellent fits ($\chi^2 = 0.9993$) at all fields (0.71 to 5.00 V/cm), for DNA sizes through the maximum enumerable length ($= 32.5$ kbp). Data taken at high field (5.00 V/cm) also yielded excellent fits ($\chi^2 = 0.9995$) at concentrations of 0.5, 1.0 and 1.3% for DNA sizes through the maximum enumerable length (22.5 kbp).

There are two different geometries for electrophoresis experiments that were performed, horizontal and vertical. As described in section 3.2.2, the gels that are cast for horizontal electrophoresis are ~ 4 mm thick. The thinner the gel the less effect temperature will have during electrophoresis, because on both sides of the gel are walls which act as a heat sink [80–82]. The heat will dissipate faster than the horizontal geometry where one side of the gel is exposed to the air. This should reduce systematic errors between experiments. The last set of experiments were a different type of conventional electrophoresis, vertical gel electrophoresis. Figure 4.1 depicts the simple setup for vertical gel electrophoresis. As the figure shows, the DNA is put in the wells on top of a gel that is between two plates (see section 3.2.3). Once the field is turned on, the DNA feel a force down towards the positive electrode.

Success in fitting different data sets to equation 2.56 was correlated with the shapes of “reptation plots” described by Rousseau et al. [58–59], and of plots of derivatives of $\log(\mu)$ versus $\log(L)$. Plot analysis indicates that equation 2.56 provides an excellent description of DNA electrophoresis in agarose when there is a smooth transition from the Ogston sieving to reptation regimes with increasing

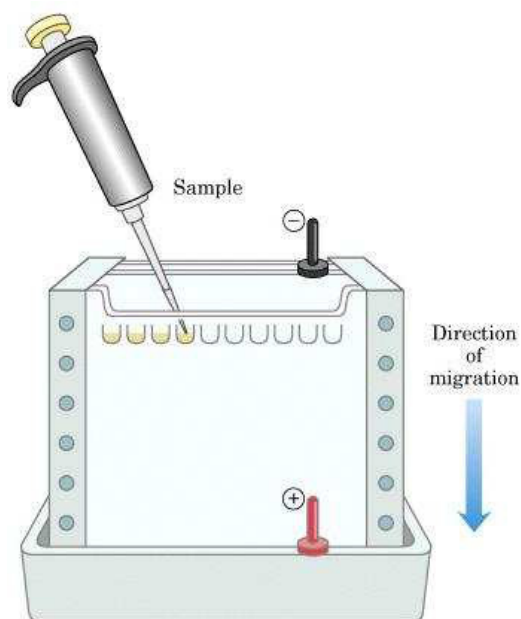


Figure 4.1. An illustration of vertical gel electrophoresis.

length. Unique features characterized both reptation plots and derivative plots of data that could not be fit well. These features were consistent with the occurrence of entropic trapping. As suggested for single stranded DNA, entropic trapping of double stranded DNA is favored by increasing gel concentration at low fields. It was also found that the parameters from equation 2.56 can be analyzed to understand aspects of the structure of agarose gels. In the following sections the results from these electrophoresis experiments will be discussed in detail.

4.1 Electrophoresis data and best-fit parameters

All experiments for this dissertation displayed similar behaviors when fitting the electrophoretic mobilities of dsDNA to equation 2.56. The trends are the same for both the horizontal and vertical setup, but the data quality is generally better for the vertical setup. For this reason, and in order to preserve the historical integrity of the research, the results are separated into sections for each set of experiments.

These sections will discuss the quality of the fit and the range in gel concentration and E field where it applies for each setup. There is more emphasis on discussing the parameters from the vertical electrophoresis experiments. The results are more consistent and contain less scatter than the data from the horizontal electrophoresis experiments.

4.1.1 Horizontal Electrophoresis with Uniform Electric Field

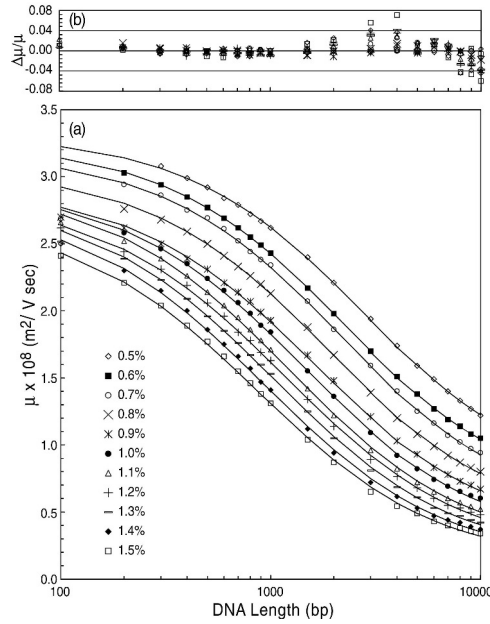


Figure 4.2. (a) Dependence of DNA mobility on length and agarose gel concentration for $E = 2.78$ V/cm. DNA samples were a mixture of $N \times 100$ bp and $N \times 1000$ bp “ladder” standards ($N = 1 \dots 10$) from Promega Corp. Lines are non-linear least squares fits to Eq. (1) generated by a MathCad (MathSoft, Inc., Cambridge, MA) program. Residuals of the fits are shown in frame (b), and were consistent in magnitude with experimental precision.

The electrophoretic mobilities, $\mu(L)$ of fragments of “100 bp ladder” and “1000 bp ladder” standards were determined for gels of eleven agarose concentrations of $T = 0.5, 0.6, 0.7, \dots, 1.5\%$ (Fig. 4.2a) as described in section 3.2.1. Plots of μ versus $\log(L)$ plot (figure 4.2) shows the complexity of the mobility for dsDNA

during electrophoresis. As the DNA length increases the mobility decreases. As the concentration increases the mobility of a DNA fragment will also decrease.

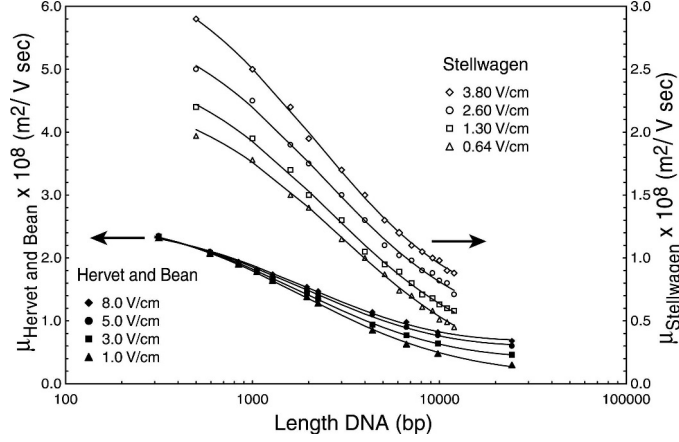


Figure 4.3. Fits of literature data to Eq. 2.53 taken from Hervet and Bean [111] for $T = 1.0\%$ (left scale), and from Stellwagen [29] for $T = 0.6\%$ (right scale) at indicated electric fields

The data were fit to a relationship that can be represented in most general form by equation 2.53. Excellent fits were obtained for all gel concentrations using an algorithm generated in MathCad (MathSoft, Inc., Cambridge, MA) (Fig. 4.2a). Residuals in $\mu(L)$ ($\frac{\mu(L) - \mu(L)_f}{\mu(L)}$, where $\mu(L)_f$ is the fit mobilities to equation 2.56) were small ($< 2\%$) and generally randomly distributed for $T \leq 1.0\%$ (Fig. 4.2b). Some non-random trends were noticeable for the highest gel concentrations, but residuals remained less than 2% for $L \leq 2000$ bp, and less than 4% for longer fragments in all gels except for $T = 1.5\%$. The residuals in the fits were also within the precision of the data for all gel concentrations. Chi-squared (χ^2) for gels with $T \leq 1.1\%$ were ≥ 0.9996 ; values for higher T drifted downward, reflecting the higher residuals for long DNAs, but χ^2 remained ≥ 0.999 for all fits. The values for χ^2 is calculated by

$$\chi^2 = 1 - \frac{\sum_i^n (\mu_i(L) - \mu_i(L)_f)^2}{(\sum_i^n \mu_i(L)^2) - \frac{(\sum_i^n \mu_i(L))^2}{n}}, \quad (4.1)$$

where i is a DNA fragment, $\mu(L)_f$ is the mobility fit to equation 2.56 and n is the number of DNA fragments. Data from several other sources were also fit well by Eq. 2.53, as is illustrated in Fig. 4.3 for some sample cases. A more comprehensive analysis of literature data is provided in chapter 6.

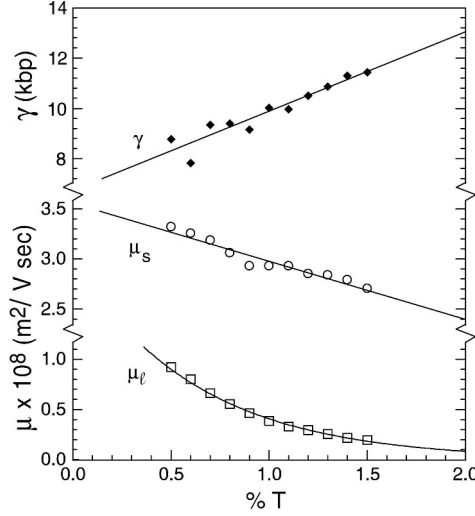


Figure 4.4. Dependencies of the limiting mobilities, μ_s and μ_L , and the crossover length γ on the gel concentration, T (refer to Eq. 2.56). The following equations, represented by solid lines, describe the dependencies: $\mu_s = -0.58T + 3.56$, $\mu_L = 1.99e^{-1.59T}$, and $\gamma = 2.78T + 7.49$, determined using least-squares fitting methods.

The parameters resulting from fits of our data to equation 2.53 showed clear trends with gel concentration (Fig. 4.4). The mobility limit for short DNA, μ_s , decreased linearly by about 24% with increasing concentration from $T = 0.5\%$ to 1.5% . The linear decrease of μ_s with gel concentration observed is consistent with the expectation that the mobility of a point particle through a collection of obstacles will be proportional to the number density of obstacles (at densities well below the percolation limit). Likewise the values of γ changed modestly and linearly with gel concentration, increasing from 8 kbp at $T = 0.5\%$ to 11.8 kbp at $T = 1.5\%$. By contrast, the mobility limit for long DNA, μ_L , decreased with increasing gel concentration, and the ratio of μ_s/μ_L rose from 3 for $T = 0.5\%$, to 12 for gels of

$T = 1.5\%$. In the figure captions, equations are shown representing best fits to allow the reader to extract values for the three parameters.

4.1.2 Horizontal Electrophoresis as a function of Electric Field

For this section, electrophoretic mobilities were determined over a range of fields from 0.71 to 5.00 V/cm and gel concentrations from 0.5% to 1.3%. Section 3.2.2 describes the use of three standard ladders and phage lambda DNA that provided data which extended, in principle, over a 485-fold length range. In practice the readable length range was often shorter because of poor resolution of the largest fragments in the most concentrated gels, and of the shortest fragments in the most dilute gels. The experimental conditions examined and the enumerable ranges of fragment lengths are listed in Table 4.1. In many cases the longest DNAs were unresolved. The sharp upper boundary of the slowest (often broad) band was presumed in these cases to correspond to the longest DNA in that lane; thus the mobility of the longest fragments could be estimated. Data obtained previously (section 4.1.1 and reference [32]) for 100 to 10,000 bp fragments electrophoresed at a field of 2.78 V/cm are included in Table 4.1 and Figure 4.5.

Data were fit to equation 2.56 using a non-linear least squares procedure programmed with MathCad (Mathsoft, Inc.). The best-fit parameters and corresponding chi-squared (χ^2) values are also summarized in Table 4.1. Data described in section 4.1.1 for a field of 2.78 V/cm were fit well to equation 2.56 ($\chi^2 > 0.999$) for all gel concentrations, T , with a trend of decreasing χ^2 with increasing T . The present studies showed that lower fields further diminished the ability to fit data at high gel concentration (Table 4.1, see also Fig. 4.5). Residuals for data that yielded the worst fits were largest for the longest DNA. Truncation of these data sets allowed fits to be obtained with $\chi^2 > 0.999$ for all conditions examined, in most cases without reducing the maximum length included in the analysis, L'_{max} , to less than 10 kbp.

Table 4.1. Electrophoresis Conditions and Best-Fit Parameters.

E (V/cm)	%T	L_{max}^a (bp)	L_{min}^a (bp)	μ_s × 10^{8b} (m² / V sec)	μ_L × 10^{8b} (m² / V sec)	γ^b (kbp)	χ^{2c}	L_{max}^d (bp)	χ^{2, d}
0.71	0.5	35000	1000	2.61	0.42	29.70	0.9999		
	1.0	20000	300	2.54	1.64 × 10 ⁻⁷	2.81 × 10 ⁷	0.9981	5000	0.9992
	1.3	20000	200	2.49	2.60 × 10 ⁻⁹	1.06 × 10 ⁹	0.9647	2600	0.9990
1.00	1.0	48502	400	2.37	0.23	29.99	0.9997		
	1.1	30000	200	2.21	0.05	74.08	0.9987	25000	0.9991
	1.2	27500	200	2.14	0.01	144.40	0.9964	1300	0.9990
1.23	0.7	47500	1000	2.60	0.66	19.86	0.9997		
	0.8	35000	200	2.55	0.23	23.84	0.9988	27500	0.9990
	0.9	25000	200	2.44	0.08	49.58	0.9971	3000	0.9992
1.31	0.5	35000	400	2.54	0.39	19.69	0.9999		
	1.0	25000	200	2.52	0.15	25.75	0.9986	17500	0.9992
	1.3	48502	200	2.48	0.05	46.75	0.9967	14000	0.9991
1.51	1.0	35000	200	2.24	0.29	17.86	0.9997		
	1.1	32500	300	2.26	0.14	23.78	0.9980	20000	0.9989
	1.2	48502	200	2.17	0.04	64.31	0.9974	15000	0.9991
1.80	0.5	47500	300	2.48	0.51	17.23	0.9998		
	1.3	30000	200	2.39	0.22	11.93	0.9980	13000	0.9991
2.00	1.0	48502	200	2.22	0.23	13.96	0.9982	20000	0.9992
	1.1	48502	200	2.24	0.16	13.06	0.9971	10000	0.9991
	1.2	17500	200	2.15	0.07	30.99	0.9979	10000	0.9990
2.22	0.5	35000	300	2.68	0.65	12.90	0.9995		
	1.0	35000	200	2.64	0.43	9.56	0.9994		
	1.3	32500	200	2.56	0.29	8.42	0.9981	17500	0.9991
2.47	0.7	32500	200	2.59	0.51	9.50	0.9997		
	0.8	35000	500	2.64	0.47	8.77	0.9998		
	0.9	48502	200	2.57	0.39	8.98	0.9992		
2.51	0.7	35000	200	2.44	0.45	14.19	0.9999		
	0.8	35000	200	2.44	0.39	12.41	0.9996		
	0.9	48502	200	2.43	0.27	11.03	0.9986	15000	0.9991
2.78	0.5	10000	200	3.32	0.92	9.30	0.9999		
	0.6	10000	200	3.25	0.80	8.47	1.0000		
	0.7	10000	200	3.18	0.66	9.80	0.9998		
	0.8	10000	200	3.06	0.55	9.84	0.9997		
	0.9	10000	200	2.93	0.46	9.64	0.9997		
	1.0	10000	200	2.93	0.39	10.40	0.9997		
	1.1	10000	200	2.93	0.33	10.35	0.9993		
	1.2	10000	200	2.85	0.30	10.82	0.9995		
	1.3	10000	200	2.84	0.25	11.14	0.9996		
	1.4	10000	200	2.79	0.22	11.52	0.9994		
3.10	1.0	32500	200	2.65	0.58	8.61	0.9999		
	1.1	30000	200	2.62	0.55	7.99	0.9998		
	1.2	27500	200	2.56	0.33	6.06	0.9980	5000	0.9994
3.51	0.5	35000	900	2.58	0.75	13.34	0.9993		
	1.0	48502	200	2.58	0.52	6.82	0.9997		
	1.3	25000	200	2.51	0.38	6.05	0.9989	8000	0.9990
5.00	0.5	22500	1000	2.53	0.75	18.15	0.9996		
	1.0	20000	400	2.55	0.67	5.78	0.9995		
	1.3	17500	200	2.51	0.50	4.22	0.9995		

^a The maximum and minimum lengths of DNA fragments that could be enumerated.

^b Best fit parameters obtained from equation 2.56 using full range of observed lengths.

^c The χ² values for fits to equation 2.56 using the parameters listed.

^d The maximum length DNA fragment which could be included in the fitting procedure to yield the new χ², values (all ≥ 0.999).

The maximum lengths, L'_{max} , that could be included to obtain fits with $\chi'^2 \geq 0.999$ are listed with the corresponding χ'^2 values in the last columns of Table 4.1.

The combinations of field and gel concentration yielding good fits to equation 2.56 are conveyed well in a three-dimensional plot of χ^2 (for untruncated data sets) as a function of both variables (Fig. 4.5). This plot is seen as a relatively flat, L-shaped surface where $\chi^2 > 0.999$, and a decreasing bowl-shaped surface for low fields and high gel concentrations. In other words, equation 2.56 describes the mobilities of double stranded DNA very well for moderate to high fields and moderate to low gel concentrations within the ranges examined. For example, data taken with $E = 3.51$ V/cm produced excellent fits over the full concentration range tested of $T = 0.5\%$ to 1.3% (Fig. 4.6). The data for $T = 1.3\%$ were fitted over the readable length range of 0.20 to 25.0 kbp with $\chi^2 = 0.9989$. Although better fits were obtained for

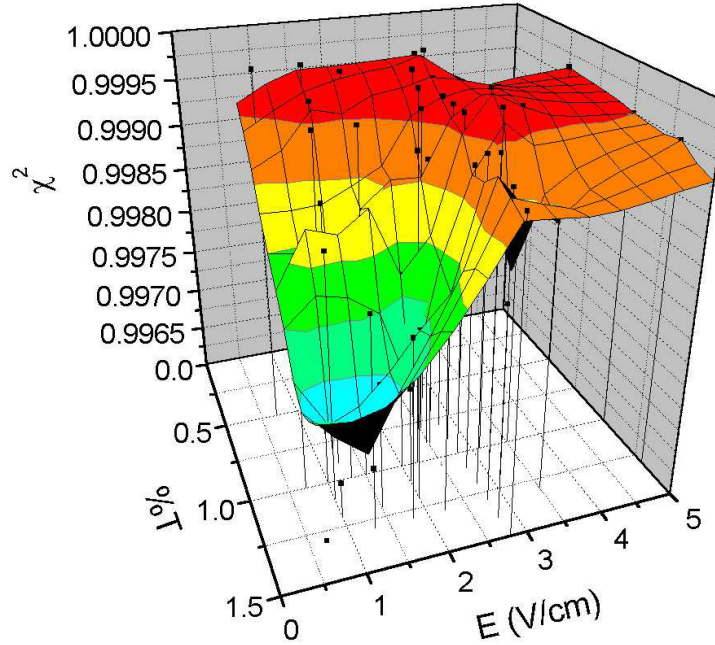


Figure 4.5. The influences of electric field strength and gel concentration on fit quality. Chi-squared (χ^2) values calculated using fitting parameters listed in Table 4.1 are plotted against E and T for ranges of $E = 0.71$ V/cm to 5.00 V/cm and $T = 0.5\%$ to 1.5% .

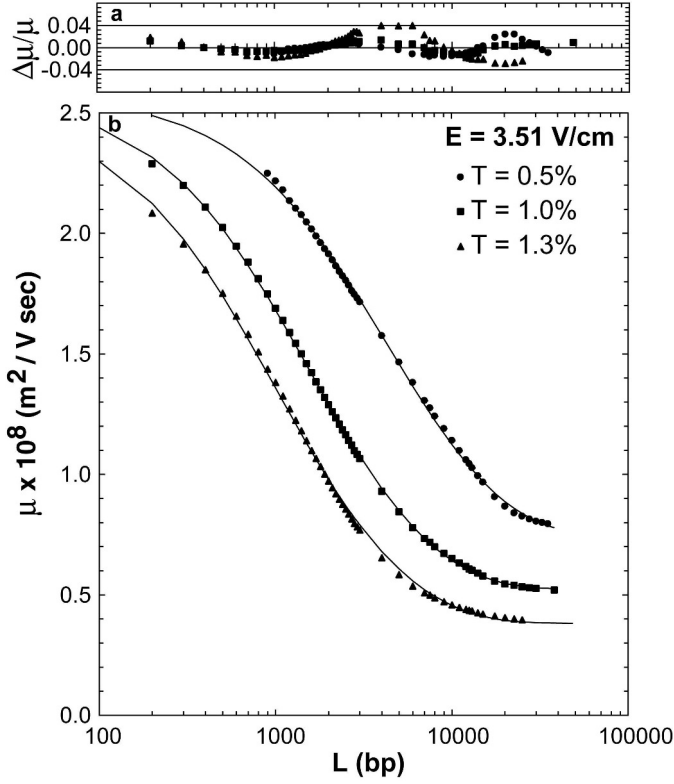


Figure 4.6. Dependence of mobility on length at moderately high field, $E = 3.51$ V/cm, for $T = 0.5\%$, 1.0% and 1.3% . Solid lines were calculated from the best fits to equation 2.56 using parameters in Table 4.1. The residuals expressed as percent deviations between calculated and measured values are shown in the upper panel.

data taken with lower gel concentrations, the largest individual residuals $\Delta\mu/\mu$ for data taken with $T = 1.3\%$ did not exceed 4% (Fig. 4.6). Data for $E = 1.00$ V/cm illustrate the deterioration of fits with increasing gel concentrations at the lower fields. Data for $E = 1.00$ V/cm, $T = 1.0\%$ were fitted extremely well ($\chi^2 = 0.9997$) over a broad readable range from 0.4 to 48.5 kbp (Fig. 4.7). Fits became notably poorer at higher gel concentrations. Deviations from the best-fit curves were most evident at long lengths, and residuals exceeded 10% for $L > 3.0$ kbp when $E = 1.00$ V/cm, $T = 1.2\%$ (Fig. 4.6, see also Fig. 4.20). The poorest fits occurred when the observed mobilities approached the length-independent limit at shorter lengths than the best-fit function.

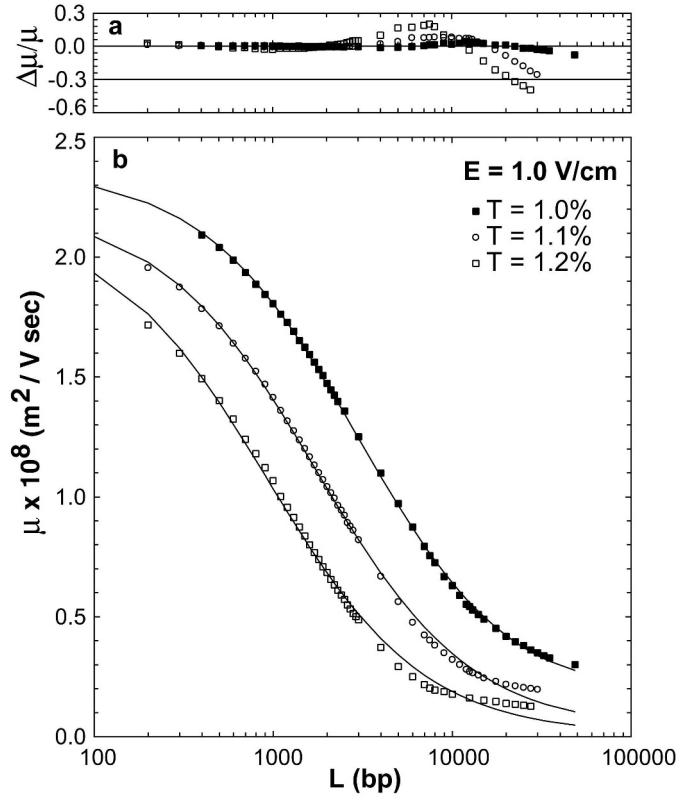


Figure 4.7. Dependence of mobility on length at moderately low field $E = 1.00$ V/cm, for $T = 1.0\%$, 1.1% and 1.2% . Solid lines were calculated from the best fits to equation 2.56 using parameters in Table 4.1. The residuals expressed as percent deviations between calculated and measured values are shown in the upper panel.

Examination of table 4.1 shows that there were modest variations in best-fit parameter values with field and gel concentration for data that conformed well to equation 2.56. The two limiting reciprocal mobilities change to varying degrees with field. The limiting mobility for short DNA, μ_s , was independent of field within the data precision. A slight, linear decrease in μ_s of about 2% per 0.1% change in T with increasing gel concentration was noted previously for the $E = 2.78$ V/cm. By contrast, very little decrease in μ_s with T was observed at higher fields in the present studies, and the mean value measured for $E = 3.51$ and 5.00 V/cm at all gel concentrations was $\mu_s = 2.54(\pm 0.03) \times 10^8$ m²/V sec.

The limiting mobility for long DNA, μ_L , increased with increasing field, about 80% from $E = 0.71$ V/cm to 5.00 V/cm. A quadratic decrease in μ_L with T was noted at 2.78 V/cm. The present, less comprehensive data at other fields is roughly consistent with this dependence. It is interesting to note that, at high fields, the limiting mobilities of long DNA in 0.5% and 1.0% gels are 30% and 26% of the mobility of the short limit DNA. Since μ_s is field independent, but μ_L decreases with decreasing field, there is a greater difference between the mobilities of long and short DNA at low fields. The long DNA limit is about 16% of the short limit for $E = 0.71$ V/cm, $T = 0.5\%$, and 10% of the short limit for $E = 1.0$ V/cm, $T = 1.0\%$.

The behavior of the exponential scaling factor γ is more complex. For low gel concentration, $T = 0.5\%$, there was an initial decrease in γ as field increased, but γ increased again at high fields. By contrast, there was a steady decrease in γ with field for $T = 1.0\%$. The behavior of γ at higher T values cannot be stated because good fits were obtained only at the higher fields in these cases. The dependence of γ on T at constant field was also complex. The newly acquired data indicate that γ decreases with increasing T . This result is at variance with the modest increase in γ with increasing T noted from the extensive data taken at $E = 2.78$ V/cm. Careful examination of Table 4.1 shows other inconsistencies between the parameters taken at this field and the new data. These small discrepancies appear to be the result of the longer length range over which the function was fit with the new data.

4.1.3 Vertical Electrophoresis as a function of Electric Field

Vertical gel electrophoresis was performed for a range of E from 0.62 V/cm to 6.21 V/cm at $T = 0.4\%, 0.7\%, 1.0\%$ and 1.3% . The previous horizontal setup had some drawbacks. The buffer was not recirculated which can cause more heating in the gel. The gel thickness could vary from run to run. Typically the thickness should be around 4 mm which at higher fields will provide for more heating. The vertical setup is more desirable (see section 3.2.3), because these problems were addressed

and more precise data was achieved. The range of lengths could be extended, because of higher resolution. Section 3.2.3 describes the materials and experimental details. The DNA lengths ranged from 200 bp to 194 kbp. Although it is assumed that the last band is 194,000 bp it also contains 145,000 bp, 97,000 bp, and 48,502 bp. Since the bands are so close together it is assumed that the band will be called 194,000 bp. Similar to the results from section 4.1.2, The length range would often vary, because of poor resolution of the largest fragments in the most concentrated gels, and of the shortest fragments in the most dilute gels. The experimental conditions and enumerated lengths used are expressed in Table 4.2.

Data were fit to equation 2.56 using a non-linear least squares procedure programmed with MathCad (Mathsoft, Inc.). The best-fit parameters and corresponding chi-squared (χ^2) values are summarized in Table 4.2. Similar to the results from section 4.1.2, lower fields further diminished the ability to fit data at high gel concentrations (Table 4.2, see also Fig. 4.8). The longest DNA yielded the worst fits. Truncation of the data sets allowed fits to be obtained with $\chi^2 > 0.999$ for all the cases examined. Some cases required lowering the maximum length, L'_{max} , to less than 4 kbp to obtain the desired χ^2 value. The maximum lengths, L'_{max} , that could be included to obtain fits with $\chi'^2 \geq 0.999$ are listed with the corresponding χ'^2 values in the last columns of Table 4.2.

The untruncated chi-squared (χ^2) values are plotted against E and T for ranges of $E = 0.62$ V/cm to 6.21 V/cm and $T = 0.4\%$ to 1.3% on figure 4.8. Notice that the χ^2 has a saddle shape. For all concentrations χ^2 roughly increases with field for $0.62 \text{ V/cm} \leq E \leq 2.17 \text{ V/cm}$ and starts to decrease for $E \geq 2.48 \text{ V/cm}$. For $T = 0.4\%$ and $E \leq 1.86 \text{ V/cm}$ the $\chi^2 \geq 0.999$. For all other concentrations the quality of the fit is poorest at the lowest fields. For example, when data taken for $E = 4.97 \text{ V/cm}$ (fig. 4.9) is compared to $E = 0.93 \text{ V/cm}$ (fig. 4.10), it is observed that for $T \geq 0.7\%$ the fits are worse for the lower field. This can easily be observed when analyzing the residuals (fig. 4.9b and 4.10b). The residuals are expressed as

Table 4.2. Vertical Gel Electrophoresis Conditions and Best-Fit Parameters.

E (V/cm)	%T	L_{max}^a (bp)	L_{min}^a (bp)	μ_S × 10⁸^b (m²/V sec)	μ_L × 10⁸^b (m²/V sec)	γ^b (kbp)	χ^{2c}	L_{max}^d (bp)	χ^{2, d}
0.62	0.4	48502	1000	2.67	0.21	52.799	0.9993		
0.93	0.4	194000	1000	2.54	0.34	25.304	0.9992		
1.24	0.4	48502	200	3.67	0.81	14.922	0.9997		
1.55	0.4	194000	200	3.18	0.77	14.488	0.9992		
1.86	0.4	194000	200	3.12	0.78	12.603	0.9990		
2.17	0.4	194000	200	3.07	0.89	11.911	0.9971	17500	0.9992
2.48	0.4	194000	200	3.13	0.90	12.648	0.9979	15000	0.9991
3.10	0.4	194000	200	3.14	0.98	9.423	0.9964	12000	0.9991
4.97	0.4	194000	200	3.34	1.22	9.630	0.9983	8000	0.9990
6.21	0.4	194000	200	3.11	1.20	9.214	0.9959	4400	0.9993
0.62	0.7	48502	200	2.86	8.36 × 10 ⁻⁶	7.32 × 10 ⁵	0.9973	2700	0.9990
0.93	0.7	194000	200	2.38	0.09	54.264	0.9968	25000	0.9990
1.24	0.7	48502	200	3.71	0.49	15.255	0.9993		
1.55	0.7	194000	200	3.61	0.53	13.125	0.9994		
1.86	0.7	47500	300	3.61	0.62	9.834	0.9997		
2.17	0.7	194000	200	3.07	0.55	10.151	0.9998		
2.48	0.7	194000	200	3.42	0.72	9.036	0.9997		
3.10	0.7	194000	200	3.14	0.69	7.713	0.9996		
4.97	0.7	194000	200	3.08	0.93	6.188	0.9984	15000	0.9990
6.21	0.7	194000	200	3.24	1.03	6.371	0.9987	10000	0.9991
0.62	1.0	48502	200	2.48	1.82 × 10 ⁻¹⁴	1.83 × 10 ¹⁴	0.9938	2900	0.9990
0.93	1.0	194000	300	2.41	1.37 × 10 ⁻¹³	2.33 × 10 ¹³	0.9947	2900	0.9990
1.24	1.0	48502	200	3.34	0.17	25.855	0.9974	14000	0.9990
1.55	1.0	47500	200	3.25	0.33	12.773	0.9977	14000	0.9990
1.86	1.0	47500	400	3.69	0.47	8.502	0.9984	10000	0.9991
2.17	1.0	194000	200	2.87	0.38	9.372	0.9989	15000	0.9990
2.48	1.0	194000	200	3.39	0.52	7.209	0.9990		
3.10	1.0	19400	300	3.36	0.61	5.852	0.9994		
4.97	1.0	48502	200	3.04	0.77	4.749	0.9989	48502	0.9991
6.21	1.0	194000	200	3.13	0.88	4.417	0.9975	12500	0.9991
0.62	1.3	48502	200	2.12	1.87 × 10 ⁻¹⁰	1.11 × 10 ¹⁰	0.9904	1800	0.9991
0.93	1.3	194000	300	1.96	3.80 × 10 ⁻⁶	6.10 × 10 ⁵	0.9928	2600	0.9992
1.24	1.3	48502	300	3.32	0.02	179.189	0.9967	2600	0.9991
1.55	1.3	194000	200	3.47	0.27	11.920	0.9958	9000	0.9990
1.86	1.3	47500	400	3.97	0.36	8.594	0.9970	7000	0.9991
2.17	1.3	194000	200	2.76	0.21	12.412	0.9966	10000	0.9990
2.48	1.3	194000	200	3.26	0.45	7.464	0.9985	11000	0.9990
3.10	1.3	47500	200	3.37	0.50	5.159	0.9984	6000	0.9992
4.97	1.3	48502	200	2.68	0.49	4.048	0.9983	4000	0.9993
6.21	1.3	194000	200	2.70	0.59	3.555	0.9991		

^a The maximum and minimum lengths of DNA fragments that could be enumerated.^b Best fit parameters obtained from equation 2.56 using full range of observed lengths.^c The χ² values for fits to equation 2.56 using the parameters listed.^d The maximum length DNA fragment which could be included in the fitting procedure to yield the new χ², values (all ≥ 0.999).

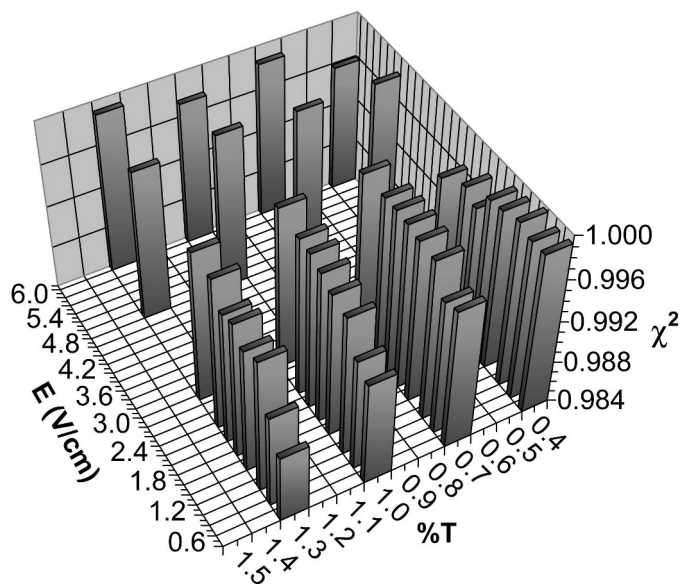


Figure 4.8. The influences of electric field strength and gel concentration on fit quality for vertical electrophoresis. Chi-squared (χ^2) values calculated using fitting parameters listed in Table 4.2 are plotted against E and T for ranges of $E = 0.62$ V/cm to 6.21 V/cm and $T = 0.4\%$ to 1.3% for vertical gel electrophoresis.

percent deviations between calculated and measured values, $\Delta\mu/\mu$, where $\Delta\mu$ is the fit subtracted from the data. The residuals for $E = 4.97$ V/cm are generally $< 4\%$ at any length or gel concentration do not exceed 8% while for $E = 0.93$ V/cm residuals exceed 50% for the largest DNA fragments and high gel concentrations. Figures 4.9 and 4.10 also demonstrate that as the gel concentration increases the value of the reduced χ^2 decreases. Although the residuals for $E = 4.97$ V/cm are small, they do show that the biggest deviations are for the higher gel concentrations. When $E = 0.93$ V/cm, the fits show more noticeable deviations from the data as $\%T$ increases (fig. 4.10b). In general the ability of the best-fit curve to represent data worsens as it approached the long-length mobility limit. These observed trends are similar to the fits to experiments reported in section 4.1.2.

Calculating the ratios of the mobilities associated with the shortest and longest DNA lengths provides new insight. The ratios of these mobilities indicate how far the

shortest lengths were separated from the longest lengths. Figure 4.11 plots the ratios as a function of E and $\%T$. The obstacles in the gel will have less effect on the shortest length DNA compared to the longest DNA fragments, so the mobility associated with shortest DNA length, μ_{200} , is always going to be faster than the mobility for the longest DNA length, μ_b . The ratios decrease with increasing E for $T = 0.7\% - 1.3\%$ (figure 4.11a). For $T = 0.4\%$ μ_{200}/μ_b decrease by 57% with increasing E . As E increases the distances between the ratios for different gel concentrations becomes closer by approximately 40%. As the gel concentration decreases the distance between the ratios decrease linearly as a function of E (figure 4.11b). The two lowest fields,

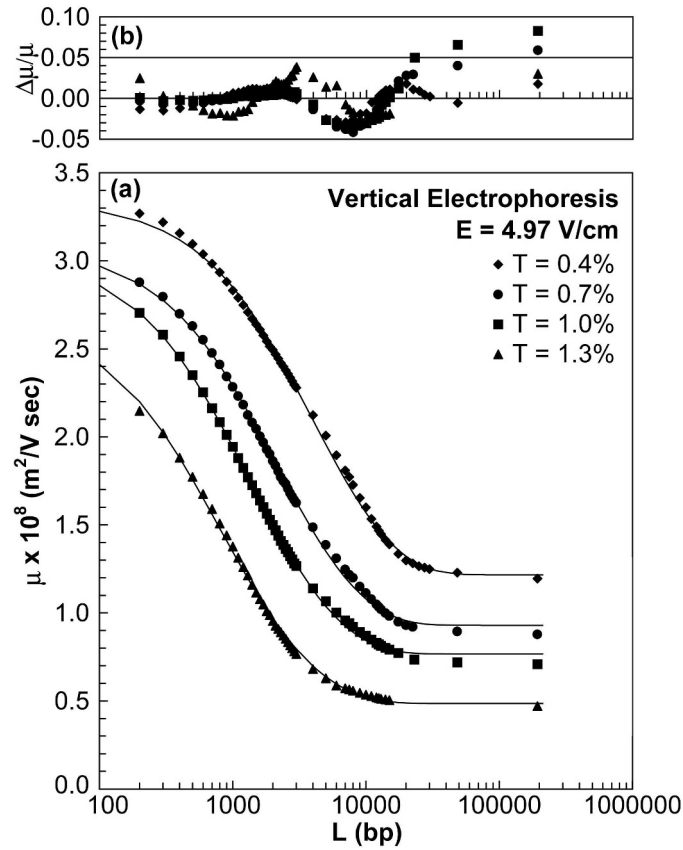


Figure 4.9. Dependence of mobility on length at a high field, $E = 4.97$ V/cm, for $T = 0.4\%, 0.7\%, 1.0\%$ and 1.3% . Solid lines were calculated from the best fits to equation 2.56 using parameters in Table 4.2. The residuals expressed as percent deviations between calculated and measured values are shown in the upper panel.

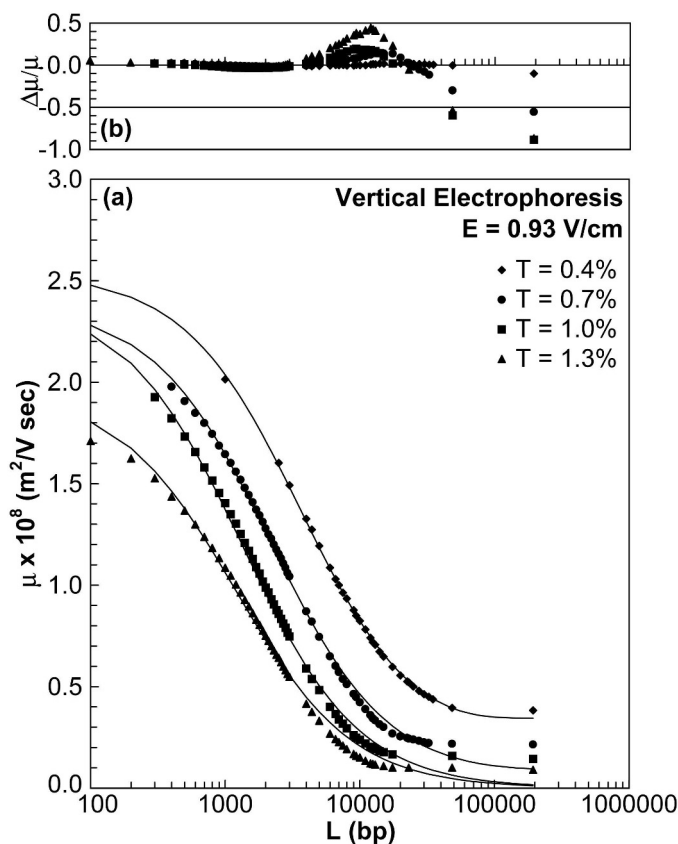


Figure 4.10. Dependence of mobility on length at a low field, $E = 0.93 \text{ V/cm}$, for $T = 0.4\%, 0.7\%, 1.0\%$ and 1.3% . Solid lines were calculated from the best fits to equation 2.56 using parameters in Table 4.2. The residuals expressed as percent deviations between calculated and measured values are shown in the upper panel.

$E = 0.62$ and 0.93 V/cm , have the greatest differences between μ_{200} and μ_b . From $E = 0.93 \text{ V/cm}$ to 1.24 V/cm slopes for μ_{200}/μ_b versus $\%T$ have a distinct change. A least-squares linear fit was performed on figure 4.11b (not shown) and plotted as a function E (figure 4.11c). There is only a 3% change in the slopes for $E \geq 1.24 \text{ V/cm}$. While a 11% decrease from $E = 0.62 \text{ V/cm}$ to 6.21 V/cm . A clear jump exists between $E = 0.93 \text{ V/cm}$ to 1.24 V/cm . A characteristic field may exist that can be associated with this change that affects the mechanisms which govern the motion of the molecule.

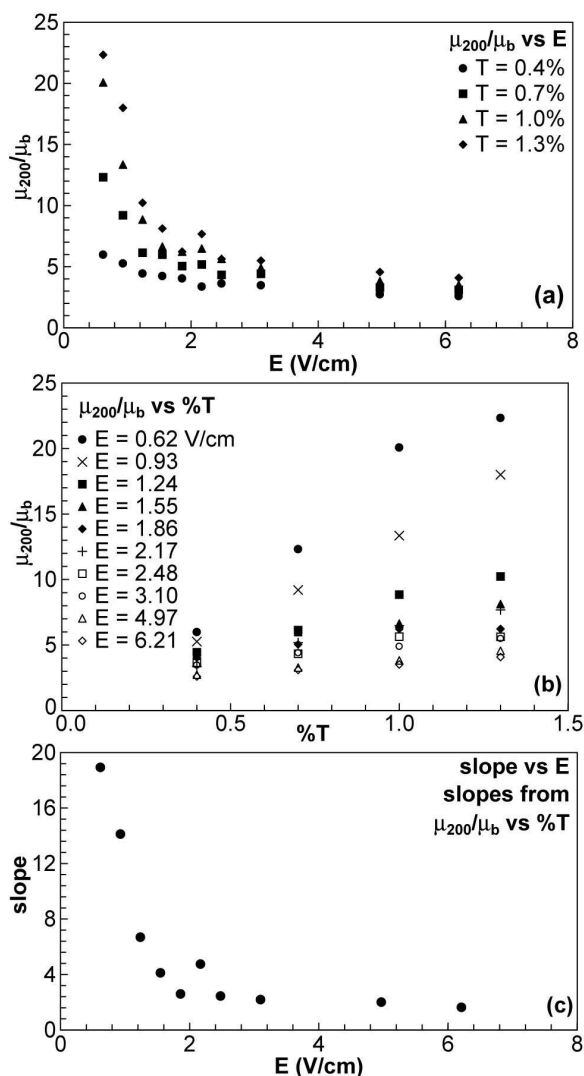


Figure 4.11. The ratios for the mobilities associated with shortest, μ_{200} , and longest, μ_b , DNA lengths as a function of E (a) and $\%T$ (b) for vertical gel electrophoresis. Slopes were extracted from least-squares linear fits to μ_{200}/μ_b versus $\%T$ (not shown). Figure 4.11c plots these slopes as a function of E .

Some interesting trends in the values for the three parameters from equation 2.56 are observed. The short length mobility limit, μ_s , varies slightly with both field and concentration (fig. 4.12). For $E \geq 1.24$ V/cm, on average there seems to be a 4.0% linear decrease in μ_s per 0.5 V/cm (see fig. 4.12a). The μ_s values for the two lowest fields are approximately 60% lower than the rest of the fields. The short length

mobility limit increases slightly with concentration for the larger fields and decreases slightly for the lower fields (see fig. 4.12b). The DNA fragments associated with μ_s will act as rigid rods and Ogston sieve through the gel [24]. Since the pores are much larger than these fragments, the DNA fragment will travel through the pores without deforming. The obstacles in the gel will cause only a slight difference in mobility for different gel concentrations. For example a 200 bp DNA fragment will be about 68 nm in length. The persistence length for a dsDNA is 50 nm, so this fragment is a little larger than one persistence length. The molecule acts like a semi-rigid rod. Serwer and Hayes [112] approximated the pore size for $T = 0.4\%, 0.7\%, 1.0\%$, and 1.3% to be 232 nm, 154 nm, 118nm and 97 nm . Even for the largest concentration gel the pore size will be approximately double the size of the smallest DNA characterized. The obstructions in the gel will not be too much of a hindrance for the migrating DNA. The short length mobility doesn't vary much with E above $E = 1.24$ V/cm. Since the DNA will not deform at this length and will sieve through the gel matrix at most field intensities, the mobilities of short length DNA are expected be approximately constant with E .

As mentioned in section 2.4.6, μ_s should not be confused with the free solution mobility, μ_o . Figure 4.13 shows that μ_o has no dependence on E with the mobilities approximately converge to a single point at $T = 0\%$. This value is associated with the free solution mobility and for each electric field it will be the same. When a least-squares linear fit is applied to figure 4.13 (not shown) the free solution mobility at $E = 0$ V/cm is calculated to be 2.9×10^{-8} m² / V sec. The value for μ_o also is within the range of values reported elsewhere [8, 28–29, 103, 113–114].

The free solution mobility is calculated for each field from the Ferguson plots. Section 2.4.2 and figure 2.8 describe how Ferguson plots are used to extract the free solution mobility. Figure 4.14 shows a specific Ferguson plot, $\ln \mu$ versus $\%T$, for $E = 1.24$ V/cm. Only selected DNA lengths are shown, so the plot will not get cluttered. A linear least-squares fit (the dotted lines) is applied is each length. As

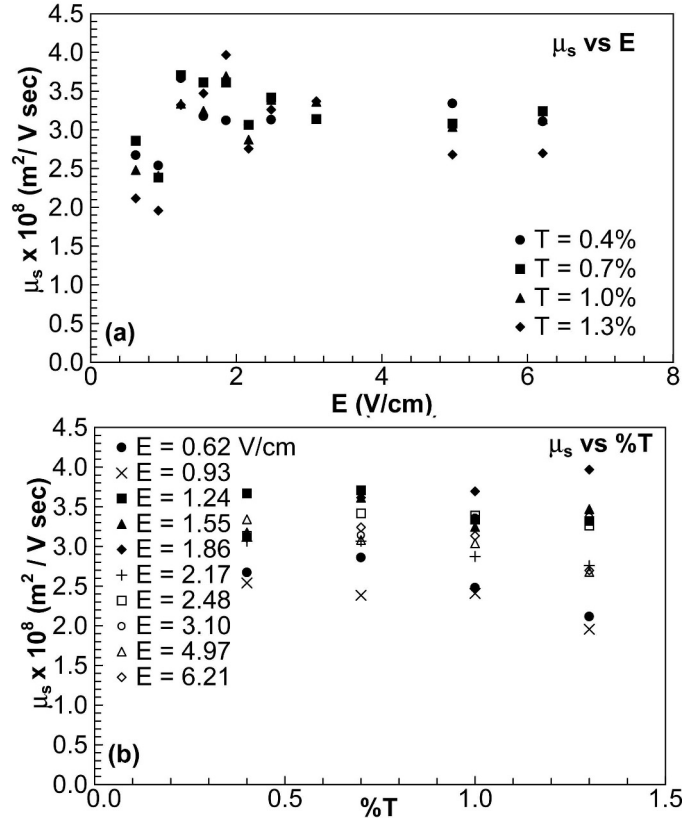


Figure 4.12. The short length mobility limit, μ_s , as a function of electric field, E , (the upper panel) and gel concentration, $\%T$ (the lower panel). The values were obtained from best-fits using equation 2.56.

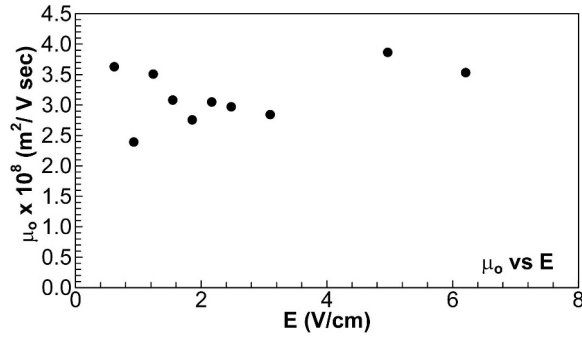


Figure 4.13. The free solution mobilities, μ_o , as a function of electric field, E . The free solution mobilities were extrapolated from the Ferguson plots, $\ln \mu(L)$ vs. $\%T$ for each field intensity (see section 2.4.2 and reference [5]).

the plot shows, for $L \leq 3000$ bp the data approximately converges to a single point at $T = 0\%$. This value is the free solution mobility, μ_o . For electrophoresis the Ferguson plot assume that the molecule is small enough so Ogston sieving [55] will apply. At $L > 3000$ bp the DNA is 20 times the persistence length, so it will behave as a random coil. Ogston sieving will not apply to DNA this large. The Ferguson plot reflects this by showing the larger lengths not converging to a point (see figure 4.14).

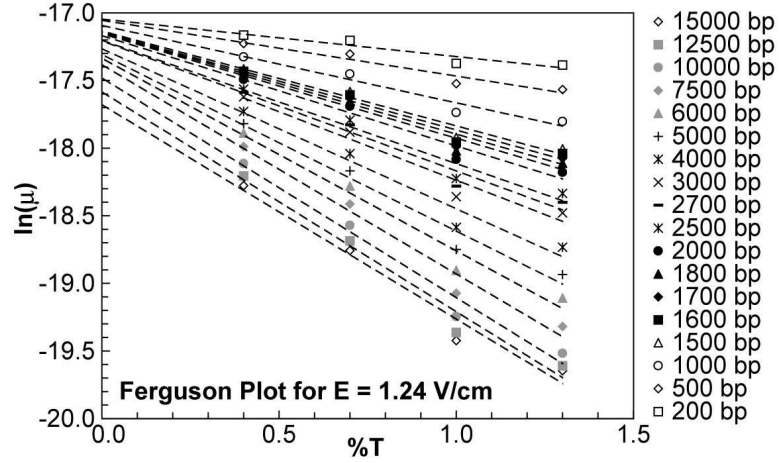


Figure 4.14. The Ferguson Plot [5] at $E = 1.24$ V/cm for selected DNA lengths. The dotted lines represent linear least-squares fit to each DNA length. See section 2.4.2 for a detailed explanation.

The limiting mobility for long DNA, μ_L , increases with increasing field by about 83% from $E = 0.62$ V/cm to 6.21 V/cm at the lowest gel concentration ($T = 0.4\%$) (fig. 4.15a). For $T \geq 0.7\%$ μ_L becomes extremely small for the lower fields. These values are excluded from figure 4.15. As seen on table 4.2, the excluded values of μ_L also associate with poor quality fits with $\chi^2 \leq 0.999$. Figure 4.15b depicts μ_L decreasing with increasing $\%T$. For each E , with increasing $\%T$, μ_L decreases on average 51% from $T = 0.4\%$ to 1.3% .

The biased reptation model predicts relationships for μ_L with gel concentration and electric field for low fields [8, 26, 30]. The reptation model for large fields is not expected to be valid due to many different reasons including the formation of

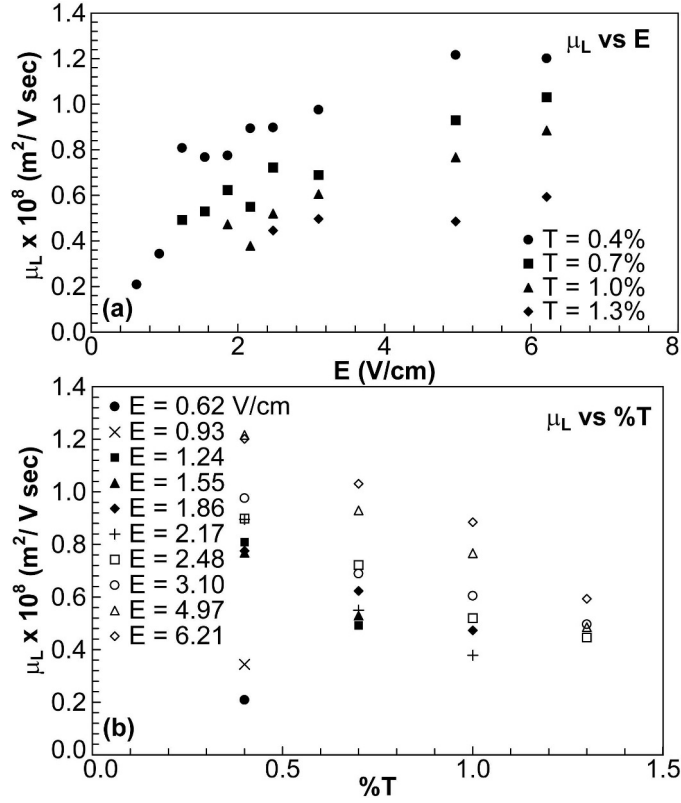


Figure 4.15. The long length mobility limit, μ_L , as a function of gel concentration, E , (the upper panel) and electric field, $\%T$ (the lower panel). The values were obtained from best-fits using equation 2.56. Long length mobilities data were excluded for $E < 1.24$ V/cm at $T = 0.7\%$, $E < 1.86$ V/cm at $T = 1.0\%$, and $E < 2.48$ V/cm at $T = 1.3\%$, because the best fit values of μ_L were unphysically small. All μ_L values for $T = 0.4\%$ were included.

lateral hernias, the stretching of the molecule, and the instability of the reptation tube [8, 59]. At the low field limit the reptation model predicts a linear increase of the long length mobility with field, $\mu_L \propto E$. Also if the persistence length, l_p , is smaller the gel pore size, the reptation model will predict that $\mu_L \propto a^2 E$. In our case dsDNA has a persistence length of 50 nm, which is smaller than the agarose gel pore size, a . Since the mean pore size is related to the concentration by $a(T) \propto 1/T^x$ with $1/2 \leq x \leq 3/4$ for gels and polymer solutions [34], the reptation model predicts $\mu_L \propto E/T^{2x}$ for low fields. By excluding $E = 6.21$ V/cm and 4.97

V/cm, the rest of the data points shown should be within the predictions of the low field reptation model. By performing a least-squares fit to the points in that region for μ_L as a function T (not shown), the exponent is found to be $0.26 \leq x \leq 0.50$. Only the 0.5 agrees with the previous predictions. There is a linear increase in μ_L with field intensity as predicted by the reptation model. A linear least-squares fit obtains χ^2 values ranging from 0.7249 to 0.9491 for $T = 0.4\%$ to 1.3% . A numerical analysis reveals that for $T = 0.4\%$ the data seems to be represented better by a log dependence which has a $\chi^2 = 0.9079$ when the data is fit to a log equation of the form $\mu_L = a \ln(\%T) + b$ with a and b being free independent parameters. If the two lowest fields are removed, $E = 0.62$ V/cm and 0.93 V/cm, a least-squares linear fit will represent the data well with a χ^2 value of 0.9249.

The crossover length or exponential scaling factor, γ , decreases with increasing field (fig. 4.16a). As the field decreases and concentration increases, γ eventually becomes extremely large. Though clearly unphysical, best fit crossover lengths become of the order of 10^7 bp. While the best fit γ parameter grows, μ_L becomes extremely small, not matching the actual mobilities of the longest length DNA. In these cases the data flatten to a limiting mobility, faster than the fit can. Similar to μ_L , these values are excluded from figure 4.16.

Another trend is that γ decreases with increasing field about 83% from $E = 0.62$ V/cm to 6.21 V/cm. The long length mobility increased with the same value with respect to field. The values of γ and μ_L obtained from fits suggest that these two parameters are coupled. A least-squares power law fit was done to the data (not shown), which produced a dependence on field of $\gamma = 26.57E^{-1.33}$ with $\chi^2 = 0.8904$. The solid line represents a least-squares fit to $\gamma = cE^{-1}$ with c being a free parameter calculated to be 26.57.

Slater [34] proposed that γ is related to the crossover length for a molecule from unoriented reptation to oriented reptation, $N^*(\varepsilon)$ [30]. The ratio γ/L_a is the number of gel pores occupied by the molecules of size γ with L_a being the size of a molecule

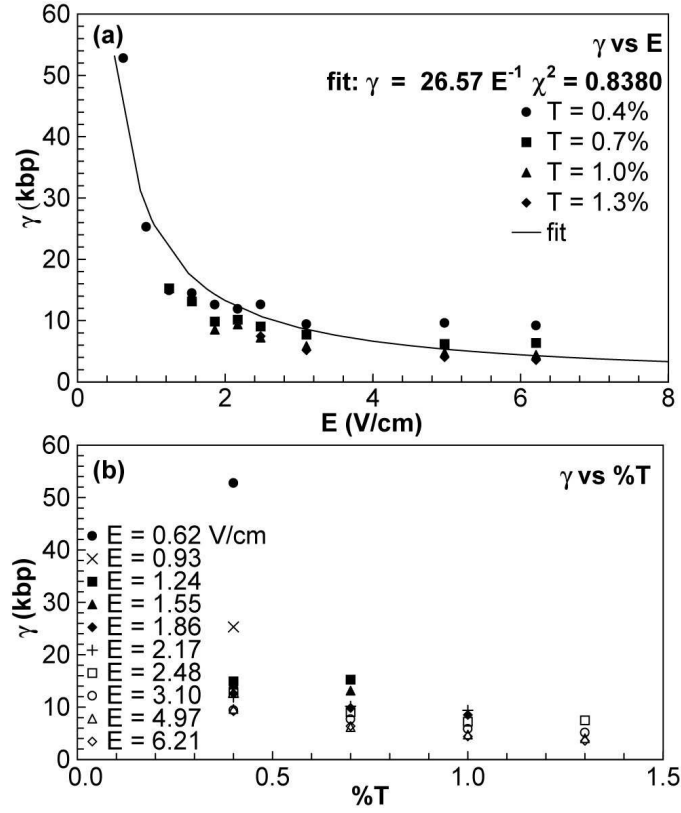


Figure 4.16. Dependency of the crossover length, γ , as a function of electric field, E , (the upper panel) and gel concentration, $\%T$ (the lower panel). The values for were obtained from best-fits using equation 2.56. The solid line represents a least-squares fit to $\gamma = cE^{-1}$ with c being a free parameter. Data were excluded for $E < 1.24$ V/cm at $T = 0.7\%$, $E < 1.86$ V/cm at $T = 1.0\%$, and $E < 2.48$ V/cm at $T = 1.3\%$, because the best fit values of γ were unphysically large. All γ values for $T = 0.4\%$ were included.

which can occupy a pore. In section 4.2, a detailed analysis of L_a will be discussed. If γ is the largest molecule that can have be individually separated, then the ratio $\gamma/L_a = N^*(\varepsilon)$. Since $N^*(\varepsilon) \sim 1/E$, then $\gamma \sim 1/E$. Figure 4.16a demonstrates that a power law fit $\gamma = cE^{-x}$ with $x = -1$ provides a decent fit.

Figure 4.16b shows that γ decreases $\sim 54\%$ with increasing gel concentrations from $T = 0.4\%$ to 1.3% . This value is quite close to the increase observed for μ_L with respect to $\%T$. As the gel concentration increases the pores become smaller

and the long length mobility limit will be reached sooner. The crossover will occur at shorter DNA lengths.

To find a more physical understanding of γ it is useful to calculate the radius of gyration, R_g , associated with each value of best fit γ . In section 2.4.3 the radius of gyration was discussed for an arbitrary length, L . To calculate R_g , γ replaces L and now has a form of,

$$R_g^2 = \frac{l_p \gamma}{3} \left(1 - \frac{l_p}{\gamma} + \frac{l_p}{\gamma} e^{-\gamma/l_p} \right), \quad (4.2)$$

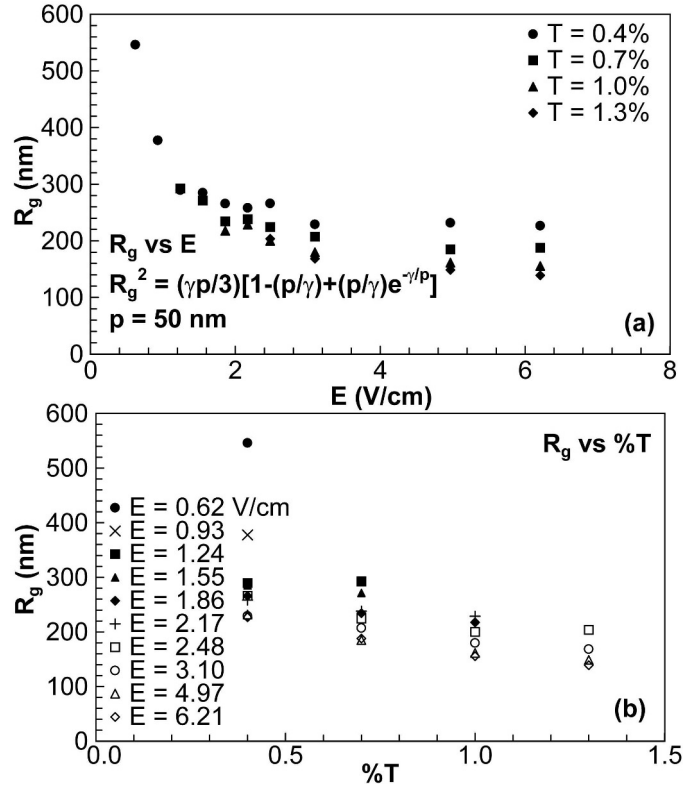


Figure 4.17. The radius of gyration, R_g , calculated from the crossover length, γ , as a function of electric field, E , (the upper panel) and gel concentration, $\%T$ (the lower panel). Data were excluded for $E < 1.24$ V/cm at $T = 0.7\%$, $E < 1.86$ V/cm at $T = 1.0\%$, and $E < 2.48$ V/cm at $T = 1.3\%$. All R_g values for $T = 0.4\%$ were included.

where l_p is the persistence length for a dsDNA ($l_p = 50$ nm). The same trends occur for R_g that were observed for γ (fig. 4.17). The crossover radius of gyration decreases as a function of field in a similar fashion as γ and falls between the values of ~ 546 nm to ~ 139 nm. For each concentration, R_g as a function of E is fit to an equation of the form $R_g = cE^{-x}$, where c and x are free parameters (not shown). The parameter c ranges from 271 to 362 nm and x ranges from 0.28 to 0.38.

At higher fields the molecule has a smaller radius gyration. This can be attributed to the deformation of the molecules due to the field. According to the reptation model at high field intensities the molecule can reptate with stretching. Also as the concentration increases R_g decreases. This also can be due to the molecule deforming, because it's trying to squeeze through the pores. As the network gets tighter the molecule will have to either find new pathways, or stretch out so it can migrate through the gel.

4.2 Effective Pore Size

The characteristics of the gel plays a very important role in understanding the separations of macromolecules. As discussed in section 1.2, the internal structure of agarose is still not fully understood. In the past, Ferguson plots [5] have been used to try to characterize the pore size, a , and fiber radius of agarose gels. This method provides a rough estimate which does not always correspond with the results from microscopy experiments (see table 4.4). Slater [34] proposed a method to calculate the effective pore size by the parameters from equation 2.56. The pore size values were obtained by the following combination of the best-fit parameters,

$$a = \left(\frac{L_a l_p b}{3} \right)^{1/2}, \quad (4.3)$$

where b is a conversion from bp to nm ($= 0.34 \text{ nm/bp}$), l_p is the persistent length, and L_a is the length of the molecules which occupies the mean pore size represented by,

$$L_a = \frac{3\gamma\mu_L}{\mu_o} . \quad (4.4)$$

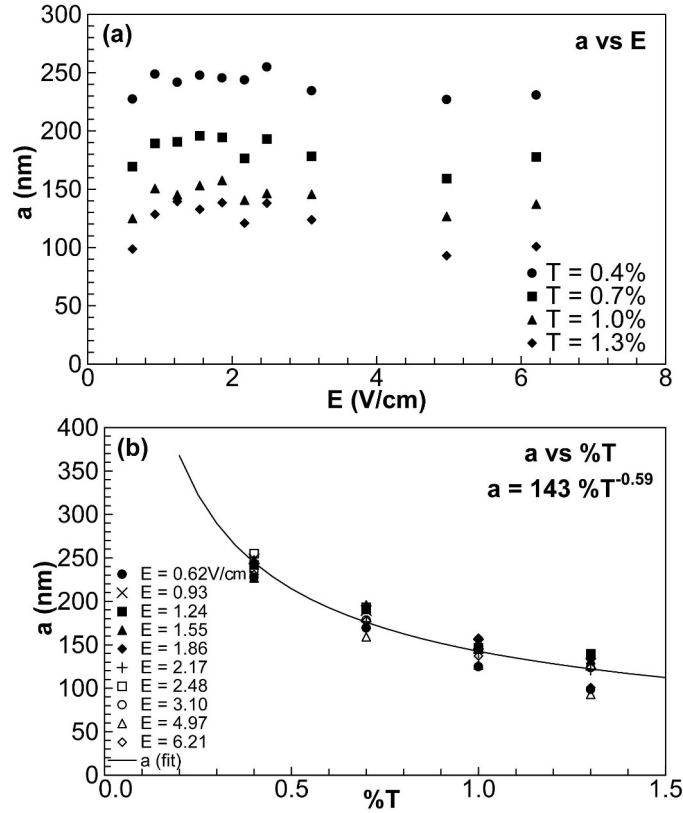


Figure 4.18. The calculated pore sizes for agarose gels from vertical gel electrophoresis experiments. The pore size values were obtained by using the method Slater [34] proposed, represented by equation 4.3. The solid line represents a least-squares power law fit to all the pore size values, with a value of $a = 143T^{-0.59}$.

There are very noticeable trends when the calculated pore size, a , is plotted as a function of E and $\%T$. Only data from the vertical gel electrophoresis experiments were used, since it provided less scatter in the parameters. The pore size stays fairly

Table 4.3. Agarose Pore Sizes Calculated from Vertical Electrophoresis Experiments

T	a (nm) ¹	a (nm) ²
0.4%	240	246
0.7%	182	176
1.0%	143	143
1.3%	121	122

¹ The pore size is the calculated average from Fig. 4.18a.

² The calculated values from eqn. 4.3 for data from vertical gel electrophoresis.

constant with respect to E (fig. 4.18a). The pore should be fairly constant with respect to E . The electric field should only affect the shape of the molecule.

The pore size has a power law dependence with respect to % T of $a = 143T^{-0.59}$ (represented by the solid line in fig. 4.18b). Table 4.3 compares the average values of a calculated from figure 4.18a with the values calculated by the power law dependence. These two values are quite similar (for $T = 1.0\%$ it is the same), which indicates the least-squares power law fit is a reasonable equation to represent the pore sizes for the data presented.

The result obtained from equation 4.3 also seems to be in agreement with Ogston's [24] and de Gennes's [25] predictions. Ogston predicted for a random array of straight chains $a \propto T^{-0.50}$ and de Gennes predicted for a network of flexible chains: $a \propto T^{-0.75}$. The pore size dependence obtained from the parameters fall in between these two predictions.

Table 4.4 compares the pore size equations obtained using Slater's method to the standard method (see below) and measured values of the pore size using electron microscopy and AFM. The standard method for calculating the pore size from electrophoretic data is to use Ferguson plots (see section 2.4.2 and reference [5]). Recall from section 2.4.2, that the extended Ogston model has the form $\mu = \mu_o e^{-K_r T}$, where μ_o is the free solution mobility, K_r is the retardation coefficient, and T is the gel concentration. For this type of gel network $K_r = c_2(R + r)^2$ (see section 2.4.2 and equation 2.34). So, when the apparent mobility of a macromolecule is equal to

Table 4.4. Agarose Pore Size Comparison with Literature.

Author	Type of agarose	Buffer	Pore size Equation	Calculated gel pore size			
				0.4% (nm)	0.7% (nm)	1.0% (nm)	1.3% (nm)
My data ^a	LE	TBE	$a = 143T^{-0.59}$	246	176	143	122
Stellwagen [115] ^b	LE	TBE	$a_a = 99T^{-1.00}$	248	141	99	76
			$a_r = 137T^{-0.50}$	217	164	137	120
			$a = 25 + 70T^{-1.0}$	200	125	95	79
Stellwagen [29]	LE	TAE	$a = 25 + 70T^{-1.0}$	200	125	95	79
Slater [28]	EG ^c	2×TBE	$a = 89T^{-0.65}$	161	112	89	75
Serwer [112]	LE	phos ^e	$a = 118T^{-0.74}$	232	154	118	97
Serwer [116]	SeaPlaque	2×phos	$a = 117T^{-1.71}$	546	215	117	75
Holmes [117]	LE	TBE	$a = 66T^{-0.95}$	158	93	66	51
Griess [47] ^d	LE	phos	$a = 162T^{-1.10 \pm 0.1}$	444	240	162	121
Griess [118]	LE	phos	$a = 148T^{-0.87}$	328	202	148	118
Griess [118]	SeaPlaque	phos	$a = 114T^{-1.12}$	318	170	114	85
Righetti [119]	Isogel	-	$a = 140T^{-0.70}$	266	180	140	117
Tinland [120]	EG	TBE	$a = 488T^{-0.70}$	927	626	488	406
Maaloum [48] ^f	EG	TBE	$a \propto T^{-0.60}$	-	-	254 ± 45	-

^a The pore size values were obtained by using the method Slater [34] proposed from equation 2.56.

^b Stellwagen calculated the pore size using the absolute mobilities (a_a) and also the relative mobilities (a_r). The absolute mobilities account for electroendosmosis.

^c EG denotes “electrophoretic grade” agarose.

^d They measured the values of pore size and fiber width (15-30 nm) by looking at images obtained from electron microscopy on the agarose gels.

^e Phos is abbreviated for 0.025 M phosphate.

^f They observed the pore sizes in the agarose gel by conducting AFM imaging on different concentrations of agarose gels.

1/2 of its mobility at zero gel concentration, the median pore size of the gel should be equal to the average diameter of the macromolecules [6, 29, 55–56, 121]. Using the values of the radius of gyration as a measure of molecular size, the average pore radius of agarose gels are found (since $a \propto T^{-0.50}$). Values for the pore sizes obtained by the standard method are equivalent to the values from equation 4.3.

4.3 Reptation Plots and Phase Diagrams

Rousseau et al. [58–59] used plots of $3\mu L/\mu_o$ versus L , termed “reptation plots,” to assess the relative contributions of various mechanisms of retardation of single stranded DNA (ssDNA) in polyacrylamide gels. This plotting method is based on the prediction of the biased reptation model that

$$\frac{\mu(L)}{\mu_o} = \frac{L_{a,RP}}{3L} + f(E) , \quad (4.5)$$

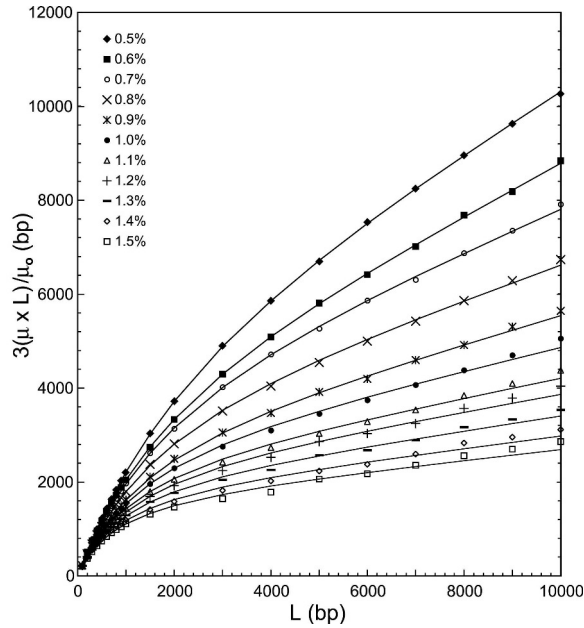


Figure 4.19. A reptation plot ($3\mu L/\mu_o$ versus L) for agarose gel concentrations $T = 0.5\%$ to 1.5% for a constant field $E = 2.78$ V/cm.

where μ_o is the free solution mobility, L is molecular size, $L_{a,RP}$ is the molecular size for which the radius of gyration $R_g(L_{a,RP}) = \langle a \rangle$, the mean size of gel pores visited by the reptating chain, and $f(E)$ is a field-dependent term due to DNA orientation. With increasing length, “reptation plots,” are expected to increase monotonically with decreasing slope in the Ogston sieving regime [58] then become linear with slope $3f(E) > 0$ and intercept $L_{a,RP}$ if there is a smooth transition to the biased reptation

regime. Analogous plots using the data from this project are consistent with the high field, relatively low gel concentration regime (Fig. 4.19). Slight deviations of the fits from the data are most pronounced for the higher concentrations.

Rousseau et al. discuss what the “reptation plots” reveal when there are deviations in the smooth transition from sieving to reptation. In the low field limit entropic trapping (ET) is predicted to cause the mobility to scale as $1/L^{1+\gamma}$, where $\gamma > 0$ [58]. The mobility will decrease in this region, because the DNA is in a trap for a period of time (see section 2.4.4 for a detailed explanation). The reptation plot in the ET regime should then be represented by a decreasing dependence on L . A dip in the plot is expected to distinguish the transition from the Ogston to entropic trapping regimes. The maximum in the curve identifies the size of the largest molecules that percolate through a connected pathway of pores of sizes $a > R_g(L_{a,ET})$. Two distinct slopes were noted in data for ssDNA data at higher fields for Rousseau et al. These slope changes were interpreted in terms of models for biased reptation with and without fluctuations [59].

4.3.1 Reptation plots accentuate deviations from fits for horizontal electrophoresis

Reptation plots for double stranded DNA electrophoresis in agarose reveal subtleties not noticed in the respective mobility versus length plots. Plot shapes were strongly correlated with the qualities of the fits of equation 2.56 to the data. Plots increased monotonically for data that could be fit exceptionally well, and the slopes decreased monotonically to a constant limit for the longest fragments. Examples of this type of reptation plot are shown in figure 4.20 for the two data sets represented in figure 4.5, plus a third data set at a lower electric field (0.71 V/cm). Solid lines were calculated using the best-fit parameters of equation 2.56 appropriate to each data set. Data for separations at $E = 3.51$ V/cm and in gels ranging from $T = 0.5\%$ to 1.3% all fit very well through the entire length range (Figure 4.20a). Data for

$E = 0.71$ V/cm at the low concentration of $T = 0.5\%$ (Fig. 4.20c) also were fit exceptionally well ($\chi^2 = 0.9999$) through the observation limit of 35 kbp. The smooth transition exhibited by the fit and the data in these cases for double stranded DNA look strikingly similar to plots of data for ssDNA at high fields that were interpreted as a smooth transition from sieving to reptation [58].

The reptation plots are particularly effective at accentuating deviations of the fitted equation from the data for the longer DNA. This feature is illustrated well in the plot for $E = 1.00$ V/cm, $T = 1.0\%$ (Fig. 4.20b), which was fit to equation 2.56 with $\chi^2 = 0.9997$. Substantial deviations in the reptation-plot appear for DNA > 25 kbp, but these were perceived as only minor deviations in the long length limit in plots of mobility versus length (Fig. 4.7).

Reptation plots for double stranded DNA have a different shape when electrophoretic data is fitted poorly to equation 2.56. Deviations from the curves calculated from equation 2.56 were seen as dips of the data below the fit in the intermediate length range of about 3 kbp to 12 kbp, followed by a cross-over for longer DNA to a linear increase with a larger slope than the fit. Such deviations are apparent in reptation plots of the data for $E = 1.0$ V/cm, $T = 1.1\%$ and 1.2% (Fig. 4b) and for $E = 0.71$ V/cm, $T = 1.0\%$ and 1.3% (Fig. 4.20c). The data for $E = 0.71$ V/cm, $T = 1.3\%$ show a clear peak in the reptation plot at $L \sim 4$ kbp. In general, the most distinctive slope variations (e.g. dips) were observed for the more concentrated gels at low fields. Decreasing the gel concentration at low field produced better conformity with equation 2.56 and increased the slope of the curve in the intermediate length range (e.g. compare the curves for $T = 1.0\%$ to 1.2% at $E = 1.00$ V/cm in Fig. 4.20b).

In 1997, Rousseau, et al. [58] proposed a phase diagram describing three different regimes that occur in electrophoresis of single stranded DNA in polyacrylamide gels. They distinguished between the Ogston, entropic trapping (ET), and reptation regimes in a schematic phase diagram extrapolated from the reptation plots

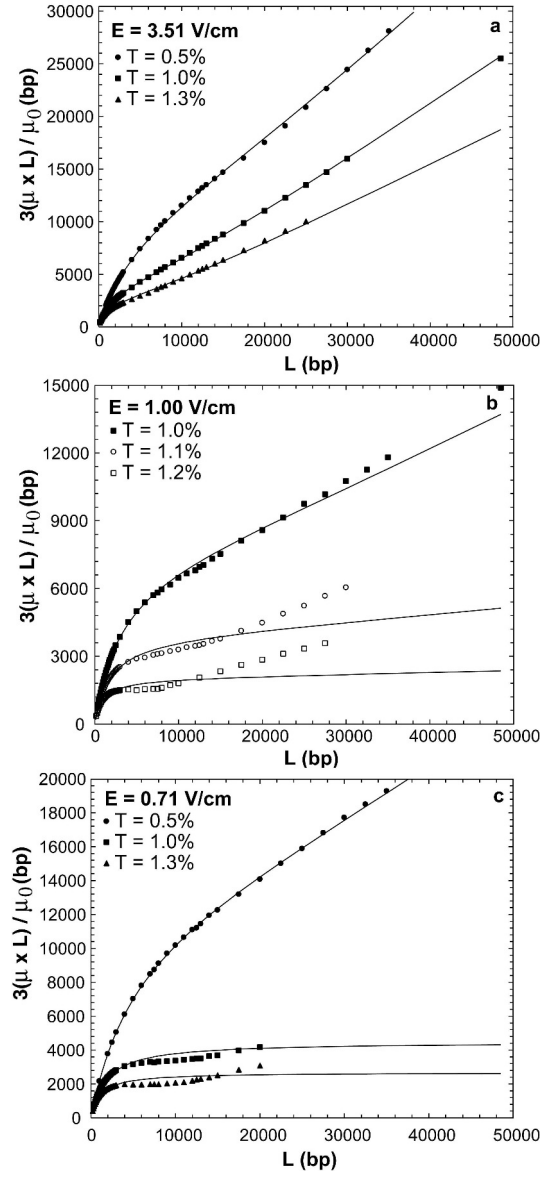


Figure 4.20. Reptation plots for data representing good and poor fits at high and low fields for horizontal gel electrophoresis. The function $(3\mu L/\mu_0)$ is plotted against L , where μ_0 is the length-independent free solution mobility obtained from Ferguson plots in which mobilities are extrapolated to zero gel concentration [5]. Solid lines were calculated from best fits to equation 2.56. (a) $E = 3.51$ V/cm and $T = 0.5\%$, 1.0% , and 1.3% , (b) $E = 1.00$ V/cm and $T = 1.0\%$, 1.1% , and 1.2% , and (c) $E = 0.71$ V/cm and $T = 0.5\%$, 1.0% , and 1.3% .

$(3\mu L/\mu_o)$. These diagrams were constructed using $L_{a,RP}$ and $L_{a,ET}$ values determined from their reptation plots. In 2000, Rousseau et al. [59] examined in more detail separations on ssDNA in the reptation limit and compared predictions of the biased reptation models with and without fluctuations. They showed and analyzed reptation plots that exhibited two apparently linear ranges with different positive slopes in the long length range, as distinct from the negative slope characteristic of ET.

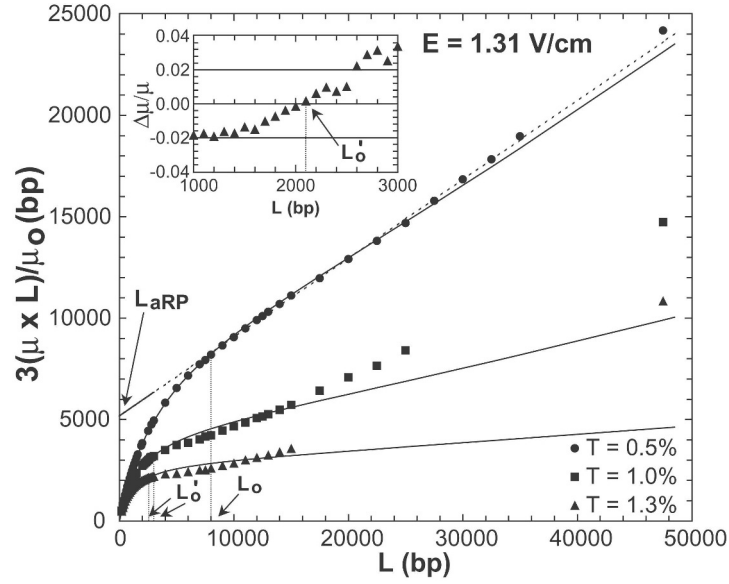


Figure 4.21. Reptation plots for $E = 1.31$ V/cm and $T = 0.5\%$, 1.0% , 1.3% showing how the crossover lengths, L_o and L'_o were determined. Solid lines were calculated to equation 2.56. The dotted line represents a linear least-squares fit to the reptation regime for $T = 0.5\%$. L_o is the crossover length from Ogston sieving to reptation when there is a smooth transition between the regimes. The insert shows residuals from $L = 1000$ bp to 3000 bp for $T = 1.3\%$. L'_o is the length for which the fit systematically lies above than the data.

Similar phase diagram were constructed (Figure 4.22) for electrophoresis of double stranded DNA in agarose based on lengths associated with characteristic transitions in reptation plots (Figure 4.21). The reptation plots were consistent with a smooth crossover from sieving to reptation when excellent fits were obtained to equation 2.56. Data for the longer DNA fragments were well fit by a straight line in these cases.

A crossover length L_o was associated with the shortest length remaining on a linear-least-squares fit of the long length data of a reptation plot (see illustration in Figure 4.21). This length is determined in the same manner as βM_{aRP} in ref [58]. The crossover lengths determined in this fashion when excellent fits were obtained for equation 2.56 are plotted as open symbols in the phase diagram (Figure 4.22).

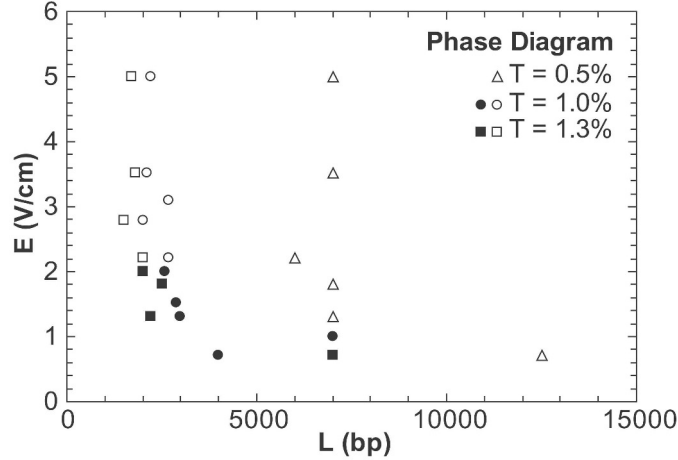


Figure 4.22. Phase diagram for dsDNA electrophoresed at different fields in agarose gels with $T = 0.5\%$, 1.0% , and 1.3% . The open and solid symbols represent the values for L_o and L'_o extrapolated from reptation plots as shown in Figure 4.21.

Best fits deviated from the experimental data in systematic ways when the quality of the fit diminished. DNA lengths associated with systematic deviations of the data from the fits were used to define boundaries between regions of the phase diagram dominated by different mechanisms of separation. Figure 4.21 shows how characteristic DNA sizes for the phase diagram were extracted from reptation plots in these cases. L'_o is the length for which the fit systematically lies above than the data. L'_o is analogous to $M_{a,ET}$ in ref [58]. These lengths are shown in Figure 4.22 as closed symbols.

4.3.2 Reptation Plots and Phase Diagrams for vertical electrophoresis

Section 4.3.1 discussed the significance of using the reptation plot, $(3\mu L/\mu_o)$ vs. L , to view the electrophoretic data. Similar trends are observed when comparing vertical gel electrophoresis data to horizontal gel electrophoresis data. When equation 2.56 provided best-fit curves with $\chi^2 \geq 0.999$, the reptation plots would increase monotonically and the slopes decrease monotonically to a constant limit for the longest fragments. An example of this would be for $E = 4.97$ V/cm (fig. 4.23), which shows two smooth transitions for all concentrations. As mentioned in section 4.1.3, the higher fields produce best-fit curves with χ^2 a bit smaller than 0.999. The reptation plots show that the larger DNA lengths deviate from the fit, reducing the quality of the fit (see figures 4.23 and 4.24). Although the χ^2 are less than 0.999, the fits are still better than for low fields and high gel concentrations.

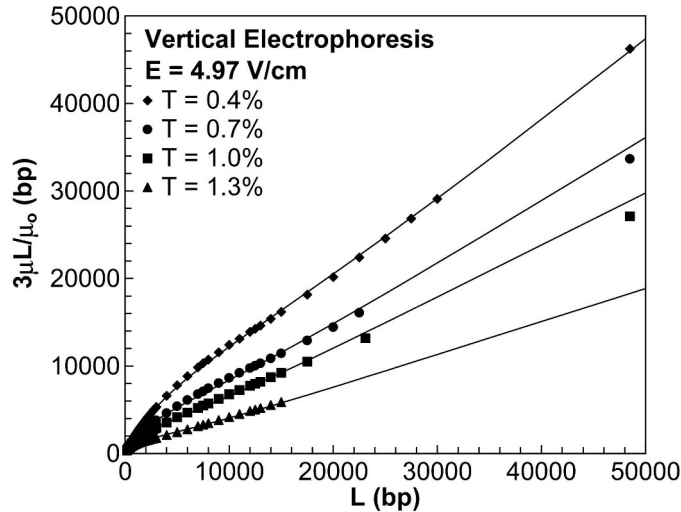


Figure 4.23. Reptation plots for electrophoresis done in a vertical chamber at $E = 4.97$ V/cm and $T = 0.4\%, 0.7\%, 1.0\%$, and 1.3% . The function $(3\mu L/\mu_o)$ is plotted against L , where μ_o is the length-independent free solution mobility obtained from Ferguson plots in which mobilities are extrapolated to zero gel concentration [5]. Solid lines are calculated from equation 2.56.

As the field increases the reptation plots change with an interesting trend. Rousseau, et al. [58] observed for low fields and high concentrations a region with a negative slope between the two smooth regions (the Ogston sieving and reptation regimes). Described in section 4.3.1, as the concentration decreases this negative slope disappears (fig. 4.25). A region of negative slope was never observed in the reptation plots for the electrophoretic data taken with the normal apparatus. A negative slope was observed a region of in data obtained at the lowest fields using vertical gel electrophoresis (Fig. 4.25). With increasing field the negative slope becomes flat in the reptation plots. Eventually for higher fields, the reptation plots show only two smooth regions and this third region disappears. Figure 4.24 shows this trend for $T = 1.0\%$. For $E \leq 0.93$ V/cm there is a region of negative slope. For 1.24 V/cm $\leq E \leq 2.17$ V/cm the slope changes from a positive to flat and back to

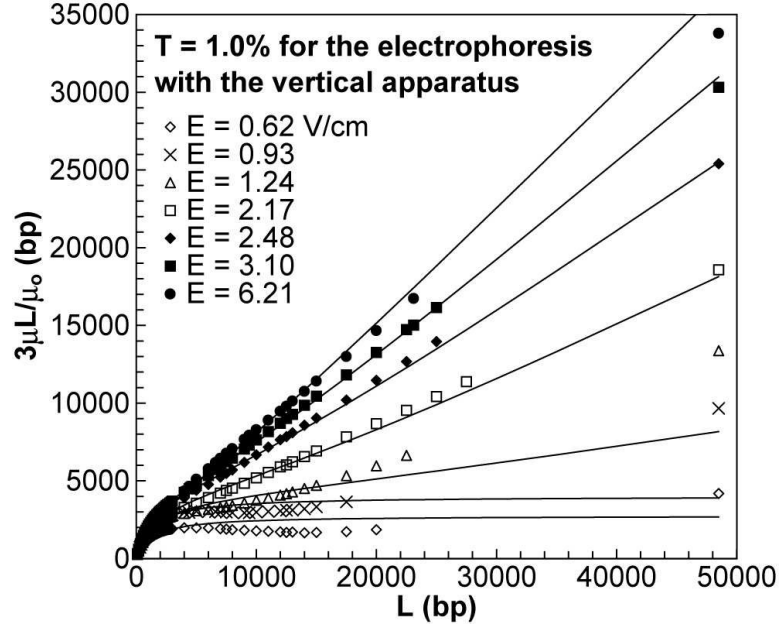


Figure 4.24. Reptation plots for electrophoresis done in a vertical chamber at $T = 1.0\%$ V/cm and $E = 0.62, 0.93, 1.24, 2.48, 3.10$, and 6.21 V/cm. The function $(3\mu L/\mu_0)$ is plotted against L , where μ_0 is the length-independent free solution mobility obtained from Ferguson plots in which mobilities are extrapolated to zero gel concentration [5]. Solid lines are calculated from equation 2.56.

positive with increasing L . When $E \geq 2.48$ V/cm there are only two smooth regions with a single crossover length.

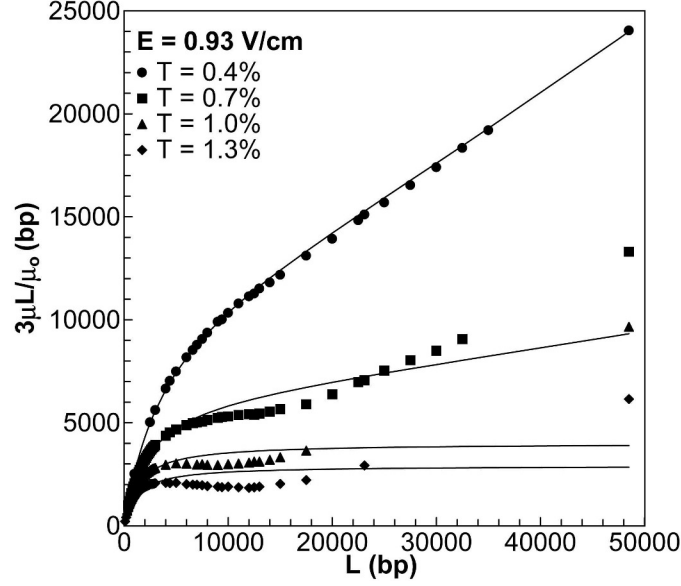


Figure 4.25. Reptation plots for electrophoresis done in a vertical chamber at $E = 0.93$ V/cm and $T = 0.4\%$, 0.7% , 1.0% , and 1.3% . The function $(3\mu L/\mu_o)$ is plotted against L , where μ_o is the length-independent free solution mobility obtained from Ferguson plots in which mobilities are extrapolated to zero gel concentration [5]. Solid lines are calculated from equation 2.56.

In 1997 Rousseau, et al. [58] proposed a phase diagram describing three different regimes that occur in electrophoresis of single stranded DNA in polyacrylamide gels. They distinguished between the Ogston, entropic trapping (ET), and reptation regimes in a schematic phase diagram extrapolated from the reptation plots $(3\mu L/\mu_o)$. A detailed description is available in section 4.3.1. Figure 4.26 demonstrates how the lengths associated with characteristic transitions in reptation plots were determined. The dotted line represents a linear least-squares fit to the reptation regimes, which is always the transport mechanism for the long length limit. In figure 4.26 $T = 0.4\%$ and 0.7% show only two smooth regions, while for the other concentrations there are three regions. A crossover length βL_{aRP} was associated with

the shortest length remaining on a linear-least-squares fit of the long length data of a reptation plot. This length is determined in the same manner as βM_{aRP} in ref [58].

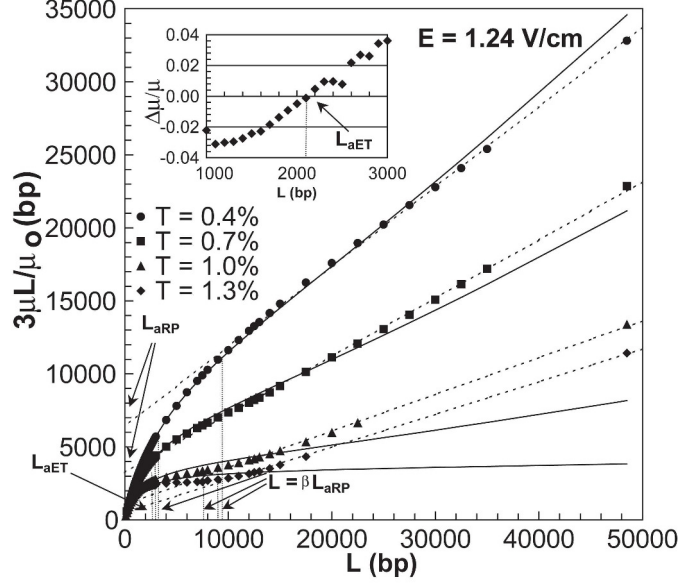


Figure 4.26. Reptation plots for electrophoresis done in a vertical chamber at $E = 1.24$ V/cm and $T = 0.4\%$, 0.7% , 1.0% , and 1.3% , showing how the transition points L_{aRP} and L_{aET} are determined. Solid lines are calculated from equation 2.56. The dotted line represents a linear least-squares fit to the reptation regimes. L_{aRP} is the crossover length from the Ogston sieving to reptation when there is a smooth transition between the regimes. L_{aET} is the length for which the fit systematically lies above the data and represents the transition from Ogston sieving to entropic trapping. The insert shows residuals from $L = 1000$ bp to 3000 bp for $T = 1.3\%$. The residual is $\Delta\mu/\mu$, where $\Delta\mu$ is the fit subtracted from data.

Experimental data deviated from the best fit in a systematic way when the quality of the fit diminished. DNA lengths associated with systematic deviations of the fits from the data were used to define additional boundaries between regions of the phase diagram. L_{aET} is the length for which the fit systematically lies above the data and represents the transition from Ogston sieving to entropic trapping (see illustration in Figure 4.26). L_{aET} is analogous to $M_{a,ET}$ in ref [58].

Phase diagrams for vertical gel electrophoresis on dsDNA in agarose gels were constructed using the characteristic lengths determined from the reptation plots

(Fig. 4.27). Three regimes were determined from reptation plots by the method described in Figure 4.26. The circles represent the transition from Ogston sieving to reptation (RP), the squares represent the transition from Ogston sieving to entropic trapping (ET), and the triangles represent the transition from ET to reptation. The transition from Ogston to entropic trapping increases in length with decreasing field. The entropic trapping transition occurs for all concentrations except for $T = 0.4\%$. The critical length characterizing the transition from ET to reptation increases with decreasing E . The results from the phase diagram are consistent with the theory for entropic trapping [107]. At low fields the molecule will tend to have a random coil or a globular configuration with the radius of gyration, R_g , being its characteristic size. When the field increases there is an increase of force applied to the molecule. The stronger the force the more pressure is applied to the globular configuration.

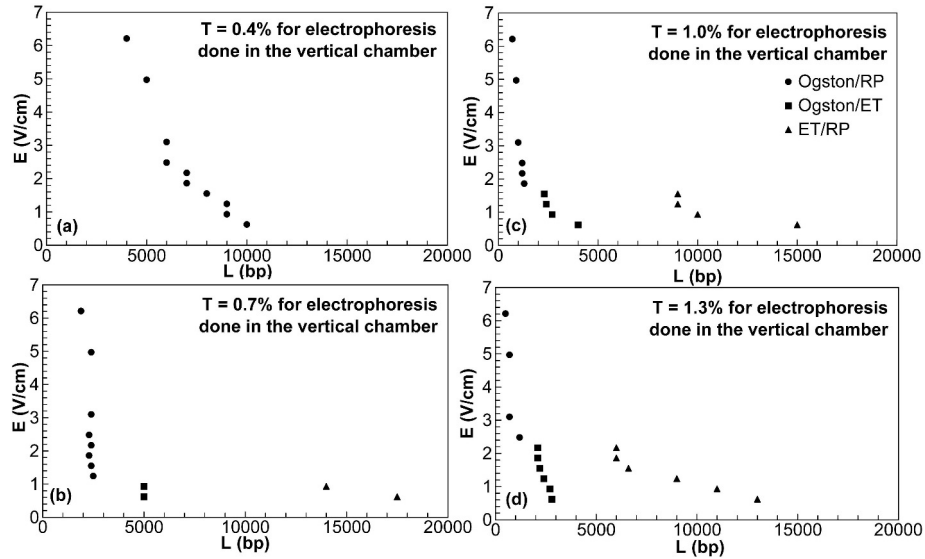


Figure 4.27. Phase diagram for dsDNA electrophoresed at different fields in agarose gels with $T = 0.4\%$ (a), 0.7% (b), 1.0% (c), and 1.3% (d). The solid circle (Ogston to reptation transition, RP) and triangle (entropic trapping, ET, to reptation transition) symbols represent the values for L_{aRP} extrapolated from reptation plots as shown in Figure 4.26. The solid square symbol (Ogston to entropic trapping) represent the values for L_{aET} extrapolated from reptation plots as shown in Figure 4.26.

Eventually there will be enough force applied so the molecule will stretch out and will reptate (migrate head first) or sieve through a gel matrix. Entropic trapping occurs for molecule's radius of gyration, R_g , approximately equivalent to the gel's mean pore size, a , which forces the molecules to selectively hop between the largest pores. Molecules with R_g much larger than a , are forced to squeeze through the pores and the motion through the gel becomes more like to reptation. At this point the molecule is no longer affected by entropic trapping.

The gel concentration plays an important role in the DNA migration mechanism. Figure 4.28 demonstrates the effect of gel concentration on the phase diagrams. The transition between Ogston sieving and reptation drops from 5000 ± 1000 bp for 0.4% gels to 500 bp for 1.3% gels. These drops in critical length consistent with the Ogston sieving and reptation theories [6, 24–25, 30, 78]. As the gel concentration increases the gel matrix will become tighter, with smaller pore sizes, so only smaller molecules will sieve. Ogston sieving to entropic trapping also decreases with length by 46% from $T = 0.7\%$ to 1.3%. The decrease can also be attributed to the tighter gel network. The molecule will need to have a smaller radius of gyration to experience entropic trapping. The transition from entropic trapping to reptation decreases with length by 21% from $T = 0.7\%$ to 1.3%. At low electric fields the molecule will deform less and would favor selectively hopping between the isolated large pores. As the gel concentration increases a larger field is required to deform the molecule so it can overcome entropic trapping in a tighter matrix. These observed trends with entropic trapping are similar to what others have observed for ssDNA in polyacrylamide gels [58–59] and are consistent with the proposed entropic trapping theories [8, 19–20].

Figure 4.28b represents the phase diagram as a function of the radius of gyration, which allows easy comparisons with the pore size and other physical characteristics of a molecule. Pore size estimates for these specific concentrations are listed on table 4.3 and shown on figure 4.18. For $T = 0.4\%, 0.7\%, 1.0\%$, and 1.3% the calculated pore sizes from equation 4.3 are 246 nm, 176 nm, 143 nm, and 122 nm. The transition

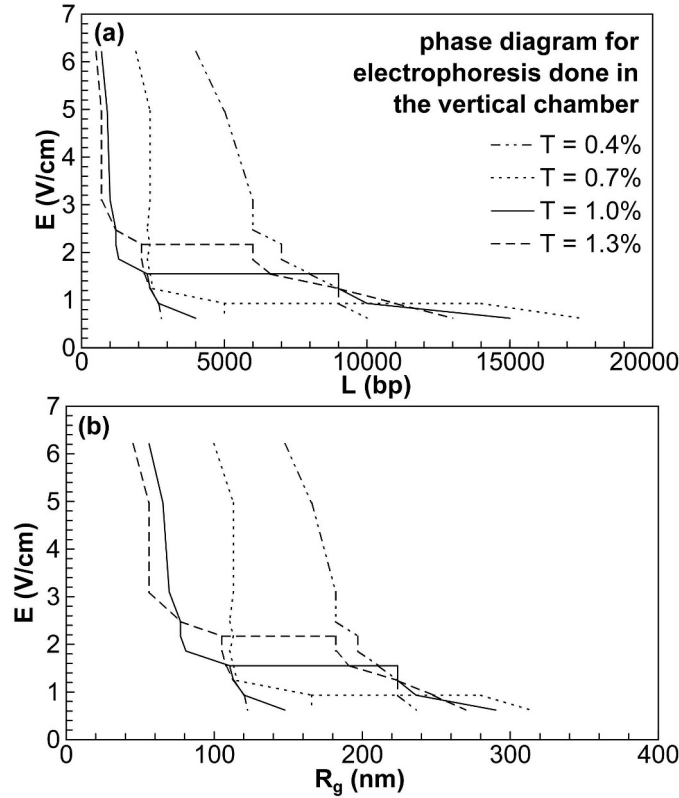


Figure 4.28. Phase diagram for vertical gel electrophoresis on dsDNA at different fields in agarose gels. The lines represent the data points for $T = 0.4\%$, 0.7% , 1.0% , and 1.3% obtained from figures 4.27 and 4.26. The upper panel (fig. 4.28a) has the phase diagrams as a function of L . The phase diagram in the lower panel (fig. 4.28b) is a function of the radius of gyration, R_g , calculated using equation 2.36.

lines in the phase diagrams between Ogston sieving and reptation are found to range from 150 to 240 nm for $T = 0.4\%$, 100 to 115 nm for $T = 0.7\%$, 60 to 80 nm for $T = 1.0\%$ and 50 to 77 nm for $T = 1.3\%$. According to the Ogston sieving model [24] when the molecule is smaller than the pore size it sieves through the gel matrix. Once the molecule is larger than the pore size it will start to deform and reptate through the gel according to the reptation theory [8, 25, 28, 65, 104, 122–123]. Thus these results provide another measure of pore size, yielding slightly smaller values than the calculation based on Slater’s pore size equation (equation 4.3). Since the calculated pore size is actually the center of a broad distribution in sizes the smaller

values attained here may provide a measure of the limiting size. The transition lines between Ogston sieving and entropic trapping are found to be approximately 170 to 180 nm for $T = 0.7\%$, 148 nm for $T = 1.0\%$ and 110 to 120 nm for $T = 1.3\%$. These values are approximately the same as the calculated size of the center of the pore distribution. The transition lines between entropic trapping and reptation are found to be approximately 280 nm for $T = 0.7\%$, 220 to 250 nm for $T = 1.0\%$ and 180 to 250 nm for $T = 1.3\%$. These values are much larger than the calculated pore sizes. As stated in section 2.4.4, entropic trapping occurs when the molecule's radius of gyration, R_g , is about equal to the gel's mean pore size, a , which forces the molecule to selectively hop between the largest pores. When R_g becomes much larger than a , the molecule is forced to squeeze through the pores and has the possibility to reptate through the gel. At this point the molecule will no longer be affected by entropic trapping. The results from the phase diagram are convincingly consistent with this interpretation.

4.4 Residuals

A closer examination of the residuals reveals some systematic deviations of the data from the fits. The residual is $\Delta\mu/\mu$, where $\Delta\mu$ is the fit value subtracted from datum for that length. Figure 4.29 displays the residuals for three different fields, $E = 0.62, 2.17$, and 6.21 V/cm for all the concentrations used the vertical electrophoresis experiments. For each concentration the residuals exhibit systematic deviations. For all E and $\%T$ the residuals have a maximum and one or two minima. The maxima and minima are indicated in figure 4.29. Notice that in figure 4.29a there is not always a first minimum of zero slope. For consistency the minimum after the 1st maximum is always labelled as the 2nd minimum.

The residuals at $T = 0.4\%$ do not exceed 5% (fig. 4.29a). Recall from sections 4.1.3 and 4.3.2, that the best-fit curves provides an exceptional quality fit over

the entire length range and the reptation plots contained two smooth regions for all fields. Even though the residuals are small, systematic deviations still exist. For $T = 0.4\%$ there are no first minima, and the maxima and the minima tend to decrease with increasing field. For the rest of the concentrations at $E = 0.62$ V/cm, the fits are not very good for DNA lengths around 10 kbp. When looking at the reptation plots it is evident that entropic trapping is occurring. Equation 2.56 only fits well when Ogston sieving and reptation occur. The fits are poor when entropic trapping is present. Although for the data where no entropic trapping occurs the residuals are small, systematic trends were similar to these for $T = 0.4\%$.

Trends are observed with the field in the maxima and minima (see fig. 4.30). The 1st minima for all the concentrations except for $T = 0.4\%$ are weakly dependent on concentration and field intensity. The value for the 1st minimum corresponds to R_g of

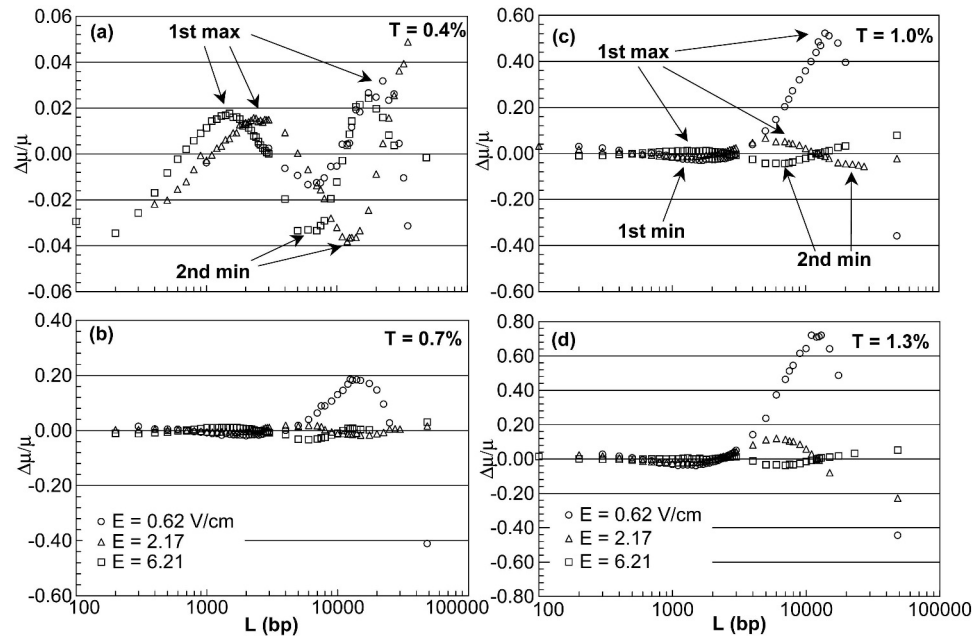


Figure 4.29. The residual plots for $T = 0.4\%$, 0.7% , 1.0% , and 1.3% and $E = 0.62$ V/cm, 2.17 V/cm, and 6.21 V/cm from fitting to equation 2.56 to the data. The residual is $\Delta\mu/\mu$, where $\Delta\mu$ is the fit subtracted from data. The maximum and minima are labelled.

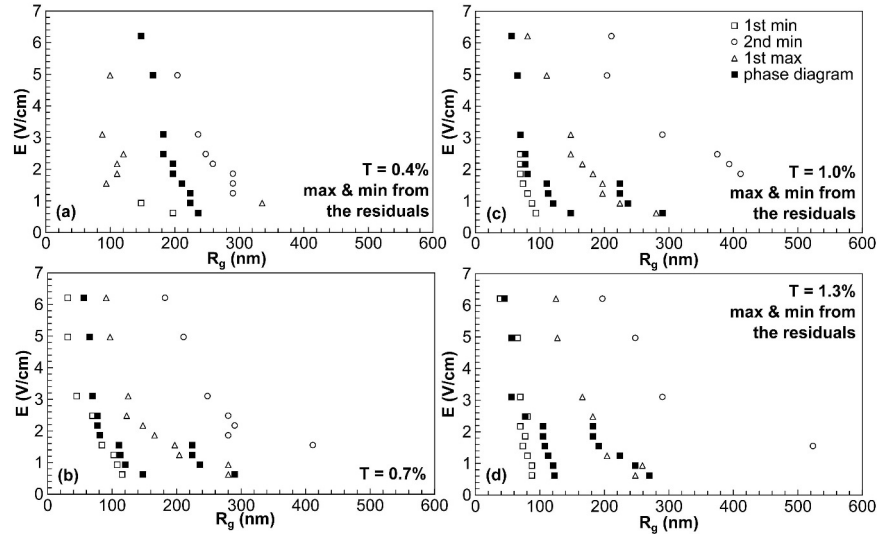


Figure 4.30. The maximum and minima from the residuals for $T = 0.4\%$, 0.7% , 1.0% , and 1.3% plotted on with the phase diagram. The filled in squares are the phase diagrams for each concentration (Figures 4.27 and 4.28). The open symbols are the minima and maxima taken from the residuals as indicated in figure 4.29.

approximately between 50 to 100 nm. For $T = 0.4\%$ the 1st minimum in the residuals only occurs for the two lowest fields. The positions of both the 1st maximum and 2nd minimum increase with decreasing field intensity. The 2nd minimum decreases with increasing concentration. Since these residual are so systematic this might be another indication for when different mechanisms of molecular transport have a transition from one regime to another.

The 1st maxima increase with decreasing field intensity (fig. 4.31). At $E = 0.93$ V/cm the 1st maxima have the greatest increase. This result is similar to that seen for the phase diagrams. For $T = 1.0\%$ and 1.3% the phase diagrams still indicate that entropic trapping should be occurring between a $R_g \approx 120$ nm to 220 nm. The 1st maxima for $T = 1.0\%$ and 1.3% change from 200 nm to 250 nm for $E \leq 1.55$ V/cm for plots of E versus R_g (figure 4.31). This can be an indication of a transition from entropic trapping to the reptation regime. At some point the molecule will be

fully reptating, free of entropic trapping. At low fields ($E \leq 0.93$ V/cm) there is a slight increase with decreasing $\%T$ and at higher fields ($E \geq 1.86$ V/cm) there is a slight decrease with decreasing $\%T$. Even with these slight changes the 1st maxima seem to be fairly independent of gel concentration.

For high fields ($E \geq 3.10$ V/cm) the 2nd minima increase slightly with decreasing field (fig. 4.32a). At $E = 3.10$ V/cm the 2nd minima start to increase drastically with decreasing field. The concentration plot (fig. 4.32b) show that the 2nd minima do not exist for $E \leq 1.24$ V/cm. The 2nd minima increase with increasing $\%T$ for $E \geq 2.48$ V/cm. The higher fields seem to be fairly concentration independent.

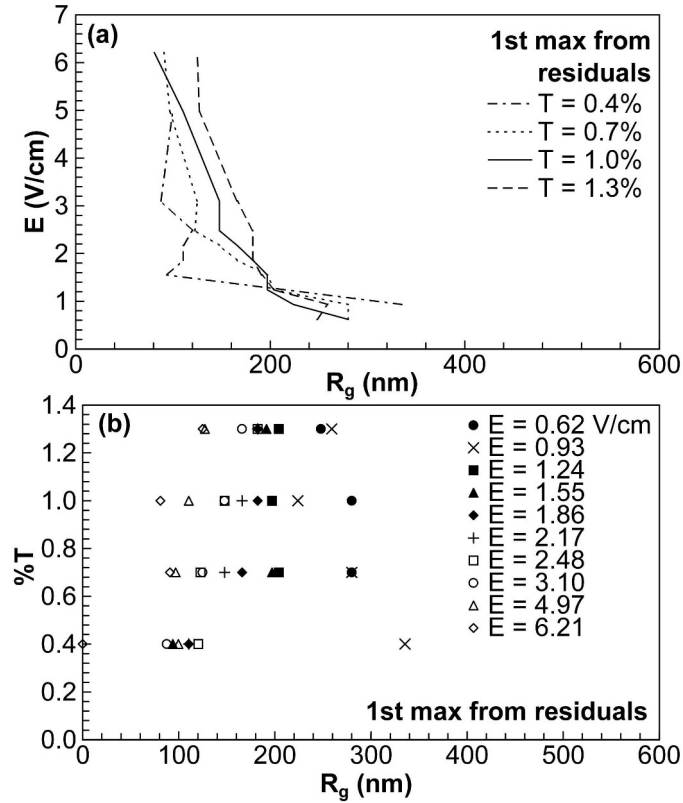


Figure 4.31. The 1st maxima from the residuals plotted as a function of the radius of gyration for both field intensity (a) and gel concentration (b). The data was acquired as described in figure 4.29.

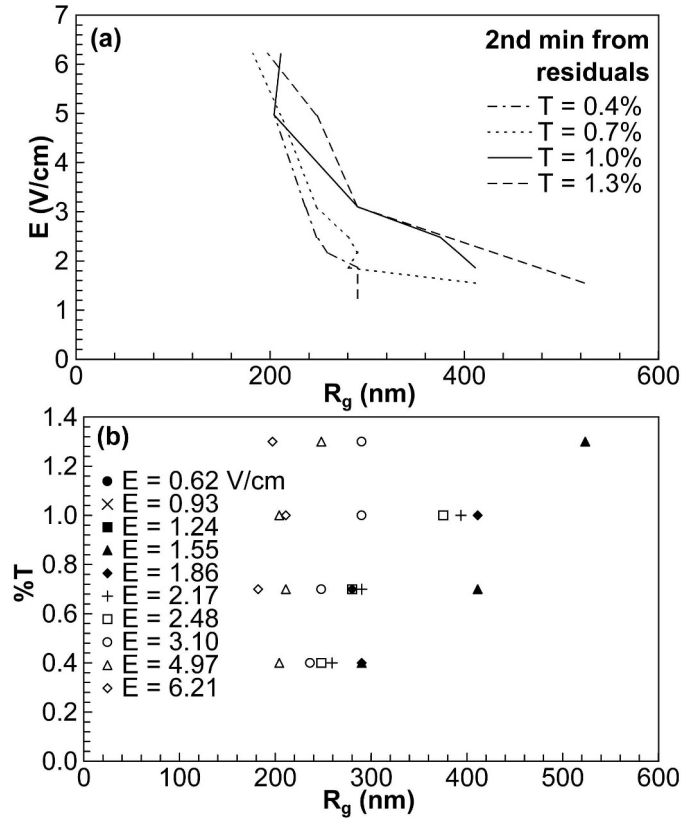


Figure 4.32. The 2nd minima from the residuals plotted as a function of the radius of gyration for both field intensity (a) and gel concentration (b). The data was acquired as described in figure 4.29.

4.5 Behavior of Derivatives

Examination of the derivative of $\log(\mu(L))$ with respect to $\log(L)$ as a function of DNA length is expected to assist in distinguishing between different mechanisms of DNA electrophoresis, since slopes of -1 or $-(1 + \gamma)$, $\gamma > 0$ are predicted for classic biased reptation or entropic trapping, respectively [58].

4.5.1 Derivatives for horizontal electrophoresis

The dependencies of derivatives of $\log(\mu(L))$ with respect to $\log(L)$ on DNA length are shown in figure 4.33 for data at five electric fields and three gel concen-

trations from horizontal electrophoresis experiments. The derivatives in all cases decrease gradually from a maximum of about -0.1 to -0.15, for the shortest DNA, to a minimum of -0.4 to -1.1 dependent on field and gel concentration, then increase sharply for lengths exceeding about 10 kbp.

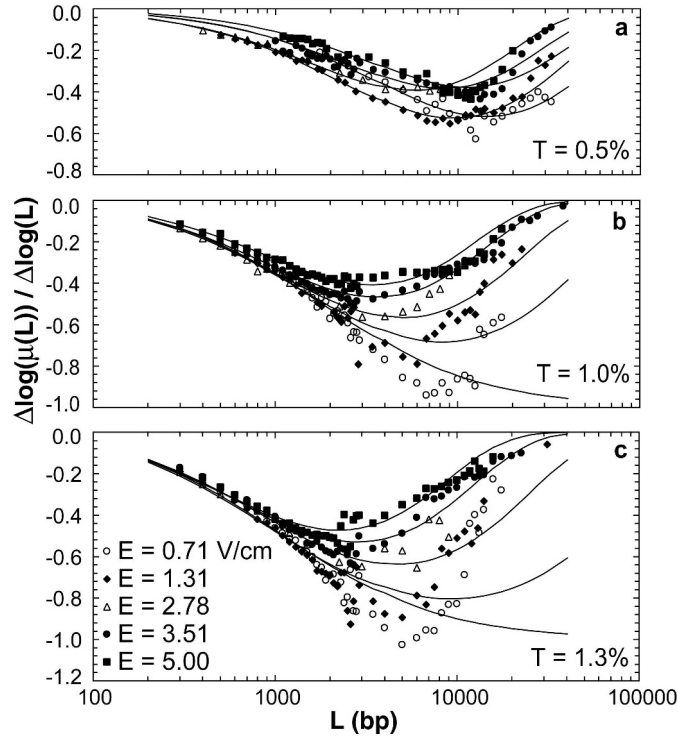


Figure 4.33. DNA length (L) dependencies of the derivatives $\Delta \log(\mu(L))/\Delta \log(L)$ on field and gel concentration. Symbols are three-point sliding average slopes of the data. Solid lines were calculated from the best fits to equation 2.56. Gel concentrations were $T = 0.5\%$, 1.0% , and 1.3% (a-c top to bottom), with fields as indicated

Both the magnitudes of the minimum derivatives, and the lengths corresponding to the minima, depend strongly on the gel concentration. The largest (negative) derivatives observed at low gel concentration ($T = 0.5\%$) fall in a narrow range of -0.4 to -0.55, and decrease in magnitude with increasing field. It should be recalled that all data obtained with $T = 0.5\%$ were fit well by equation 2.56, a result reflected by the close correspondence between the derivatives calculated using the best-fit parameters

to equation 2.56 (solid lines) and the derivatives of the actual data (points). The lengths corresponding to minima in the derivatives fall between 8.0 and 12 kbp with one exception. A consistent trend is not apparent. At higher gel concentrations, there was a clear shift of the minimum to smaller lengths with increasing field.

Derivative plots were more varied at higher gel concentrations, as shown in Fig. 4.33b, c for $T = 1.0\%$ and $T = 1.3\%$. Minimum derivatives fell in the range of -0.4 to -0.6 at the highest fields ($E = 2.78$ to 5.0 V/cm), where data were well fit by equation 2.56. By contrast, derivatives for data taken at low fields and high gel concentrations approached the predicted reptation limit of -1.0 ($E = 0.71$ V/cm, $T = 1.0, 1.3\%$ and $E = 1.31$ V/cm, $T = 1.3\%$ in Fig. 4.33b, c). The lesser ability of equation 2.56 to fit data for these conditions resulted in severe deviations of the calculated from the observed derivatives for lengths above 3.0 kbp. The poorer of the best-fit functions were not able to reproduce either the steep drop of the derivative as it approached the minimum near -1.0, or the sharp rise at greater lengths.

Reptation plots prepared from our data are reminiscent of those reported by Rousseau et al. [59] for electrophoresis of single stranded DNA in polyacrylamide. Examination of the derivatives of $\log \mu(L)$ with respect to $\log(L)$, however, suggests that separation mechanisms for double stranded DNA do not simply parallel the behavior of single stranded DNA. Data for double stranded DNA that were well fit by equation 2.56 yielded reptation plots similar to plots for single stranded DNA that were attributed to a smooth transition from the Ogston sieving to the reptation regime. Despite this similarity, derivative plots for double stranded DNA yielded maximum negative slopes of -0.4 to -0.6, far from the reptation limit. Slopes approaching the reptation limit of -1.0 were observed for double stranded DNA only if the data were not described well by equation 2.56.

Theories that describe reptation and entropic trapping predict linear relationships between mobility and L^{-a} with $a = 1$ and $(1+\gamma, \gamma > 1)$, respectively. For ssDNA [58] slopes of $\log \mu(L)$ versus $\log(L)$ plots were ≤ -1.0 , and calculated values of $1 + \gamma$,

were from 1.1 to nearly 1.5 for cases attributed to entropic trapping. The failure to observe such large negative slopes for double stranded DNA can be rationalized in terms of the contribution of a length-independent term to DNA mobility, expressed as $f(E)$ in equation 4.5. This equation can be rewritten as

$$\mu(L) - f(E)\mu_o = \frac{L_a\mu_o}{3L}, \quad (4.6)$$

where L_a is the length corresponding to molecular size M_a . The limiting mobility for large L is then $f(E)\mu_o$. Slopes of $\log(\mu(L))$ versus $\log(L)$ plots will approach -1 or $-(1 + \gamma, \gamma > 1)$ only if the limiting mobility is zero or very small.

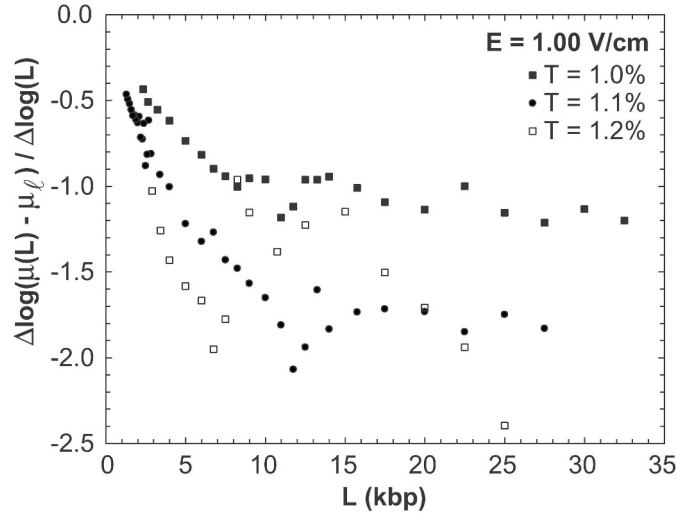


Figure 4.34. Dependencies of $\Delta \log(\mu(L) - \mu_L) / \Delta \log(L)$ versus L on gel concentrations at $E = 1.00$ V/cm. Symbols are three-point sliding average slopes of the data. The long DNA mobility limit, μ_L , for $T = 1.0\%$ was predicted by the fit to equation 2.56. For $T = 1.1\%$ and 1.2% , μ_L values were set equal to 90% of the mobility of the longest fragment listed in Table 4.1.

In the limit $L \rightarrow \infty$, the length-independent term $f(E)\mu_o$ should be identical to the μ_L parameter of equation 2.56 under conditions where the best-fit equation provides a good description of the mobilities of long DNA. Figure 4.34 shows that the slope of $\log(\mu(L) - \mu_L)$ versus $\log(L)$ does, in fact, reach a length-independent

limit, for lengths ≥ 10 kbp, at the predicted reptation limit of -1.0 for data taken at $E = 1.00$ V/cm, $T = 1.0\%$. These data were well fit by equation 2.56, and examination of Figure 4.7 shows that the calculated best-fit equation closely follows the mobilities at long lengths for this case.

Data taken at $E = 1.00$ V/cm, at the higher gel concentrations ($T = 1.1\%, 1.2\%$), were poorly fit. For these data, the calculated best fits yielded values of μ_L much smaller than the actual mobilities. In these cases replacement of the best-fit values of μ_L by actual mobilities measured for the longest DNA provided more accurate estimates of the limiting mobilities. Setting values of the limiting mobilities equal to 90% of the actual mobilities of the longest lengths reported in these cases again caused the slope of $\log(\mu(L) - \mu_L)$ versus $\log(L)$ to approach a length-independent limit for lengths greater than about 10 kbp (Fig. 4.34). The limits in these cases were about -1.8.

4.5.2 Derivatives for vertical electrophoresis

In section 4.5.1, an examination of the derivative of $\log(\mu(L))$ with respect to $\log(L)$ as a function of DNA length was performed on horizontal gel electrophoresis experiment. A similar analysis is done in this section for vertical gel electrophoresis. The dependencies of derivatives of $\log(\mu(L))$ with respect to $\log(L)$ on DNA length are shown in figure 4.35 for vertical gel electrophoresis data. The derivatives in all cases decrease gradually from a maximum of about -0.1 to -0.15, for the shortest DNA, to a minimum dependent on field and gel concentration, then increase sharply for lengths exceeding about 10 kbp. This confirms the horizontal gel results.

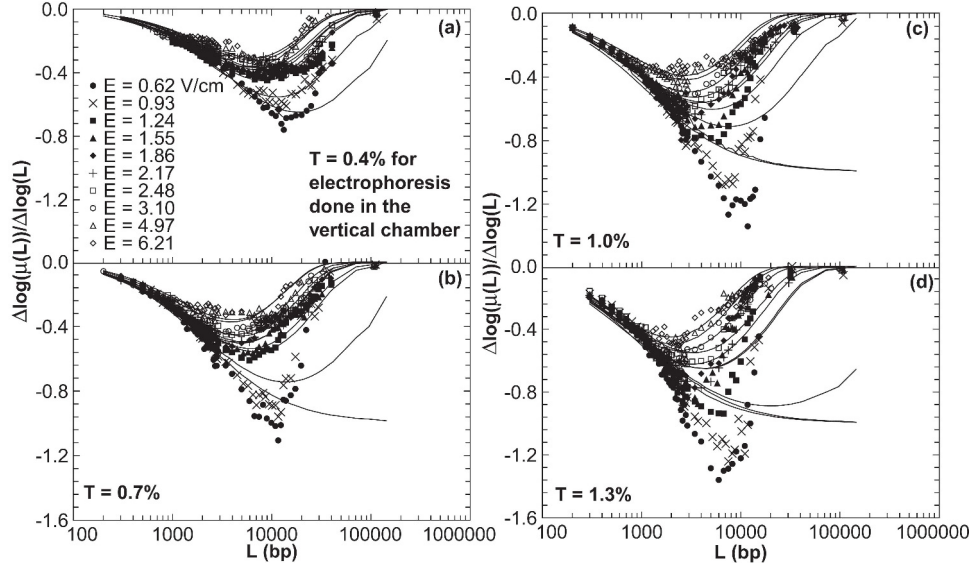


Figure 4.35. DNA length (L) dependencies of the derivatives $\Delta \log(\mu(L))/\Delta \log(L)$ on field and gel concentration. Symbols are three-point sliding average slopes of the data. Solid lines were calculated from the best fits to equation 2.56. Gel concentrations were $T = 0.4\%$, 0.7% , 1.0% , and 1.3% (a-d), with fields as indicated.

At $T = 0.4\%$ (Fig. 4.35a) the limit is greater than -0.7 for all field intensities. The reptation plots and phase diagrams indicate that entropic trapping does not occur for any E or $\%T$. The reptation plots show a transition from Ogston sieving to reptation between 5000 bp to 9000 bp. At $T = 0.7\%$ (Fig. 4.35b) the limit at low fields ($E = 0.62$ V/cm and 0.93 V/cm) is between -1.0 around $10,000$ bp. This would indicate that at this point the molecule should be reptating and that it will never reach the entropic trapping region. Reptation plots show (figures 4.24, 4.25, 4.26, and 4.23) that these conditions have an entropic trapping region for $L = 5000$ bp. The rest of the fields at $T = 0.7\%$ are ≥ -1.0 which would indicate the molecules might exhibit the Ogston sieving or reptation regime. At $T = 1.0\%$ and 1.3% (Fig. 4.35c and d), the limit is about -1.0 and -1.4 for $E = 0.62$ V/cm and 0.93 V/cm. Slopes of $-(1+\gamma)$, $\gamma > 0$ are predicted for entropic trapping [58], so this indicates that entropic trapping may occur which is in agreement with the reptation plots for these fields.

The rest of the fields have a limit ≥ -1.0 . At $E = 1.55$ V/cm for both $T = 1.0\%$ and 1.3% , the reptation plots and phase diagrams show an entropic trapping region occurring while in the derivative plots they reach a limit of approximately 0.9 for $T = 1.3\%$ and approximately 0.8 for $T = 1.0\%$, which is consistent with Ogston sieving or reptation.

In section 4.5.1, plotting $\Delta \log(\mu(L) - \mu_L) / \Delta \log(L)$ versus L was introduced. This allowed for a better representation for distinguishing between different mechanisms of DNA electrophoresis. Figure 4.36 shows these plots for vertical gel electrophoresis. The long DNA mobility limit, μ_L , for all fields at $T = 0.4\%$ was predicted by the fit to equation 2.56. For $T = 0.7\%, 1.0\%$ and 1.3% , μ_L values were set equal to the mobility of the longest fragment listed in table 4.2. For data sets which did not contain $L = 194,000$ bp, μ_L values were set equal to 90% of the mobility of the longest fragment.

At $T = 0.4\%$, figure 4.36a reaches a limit of -1.0 for $L \leq 20,000$ bp. Recall from section 4.5.1, when entropic trapping does not occur $\gamma \leq 1$. From the reptation plots (fig. 4.26) and table 4.2 it is observed that entropic trapping does not occur for these regions which is in agreement with the results. For $L > 20,000$ bp a limit of ~ -2.5 for $E = 1.24$ and 1.55 V/cm.

For $T = 0.7\%$ (figure 4.36b), a limit of -1.0 is reached for both $E = 1.24$ and 1.55 V/cm. The reptation plots show no entropic trapping region for these fields. Only $E = 0.62$ V/cm reaches a limit of -3.0 and becomes less than -1.0 at $L = 5000$ bp. The phase diagrams (figures 4.27 and 4.28) confirms this result by showing the Ogston sieving to entropic trapping transition occurring at $L = 5000$ bp.

The phase diagrams (figures 4.27 and 4.28) show that there is an entropic trapping (ET) transition for all the fields listed in figures 4.36c ($T = 1.0\%$) and 4.36d ($T = 1.3\%$). For all fields at $T = 1.0\%$ (figure 4.36c) the plot increases less than -1.0 at $L \sim 2500$ bp which corresponds with the Ogston sieving to ET transition in the phase diagrams. The limit for $E = 0.62$ V/cm is -3.0 at $L \sim 15,000$ bp. The

phase diagrams show an ET to reptation transition for $T = 1.0\%$ at $L = 15,000$ bp. Similar trends are observed for the rest of the fields at $T = 1.0\%$ and 1.3% .

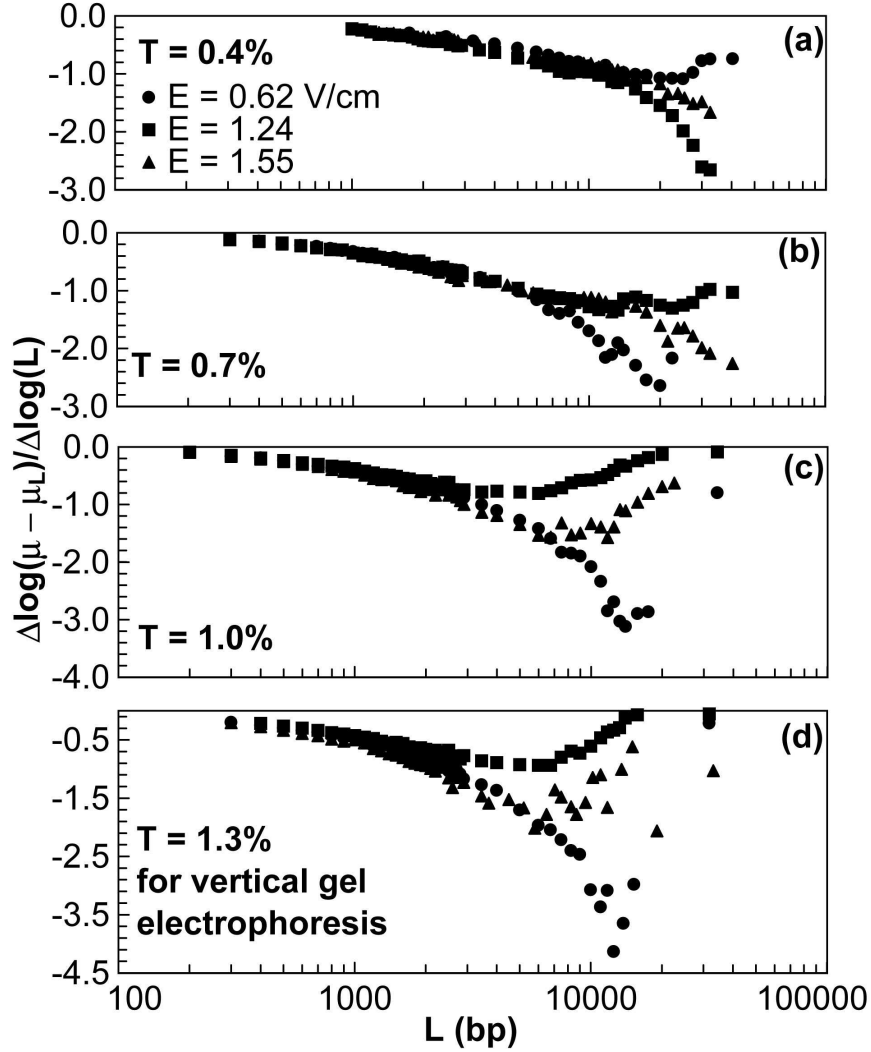


Figure 4.36. Dependencies of $\Delta \log(\mu(L) - \mu_L) / \Delta \log(L)$ versus L on gel concentrations at $E = 0.62, 1.24$, and 1.55 V/cm. Gel concentrations are $T = 0.4\%, 0.7\%, 1.0\%$, and 1.3% (a-d). Symbols are three-point sliding average slopes of the data. The long DNA mobility limit, μ_L , for all fields at $T = 0.4\%$ was predicted by the fit to equation 2.56. For $T = 0.7\%, 1.0\%$ and 1.3% , μ_L values were set equal to the mobility of the longest fragment listed in table 4.2. For data sets which did not contain $L = 194,000$ bp, μ_L values were set equal to 90% of the mobility of the longest fragment.

4.6 Summary

The results presented here comprise the most complete set of experiments performed to study DNA separation by electrophoresis in agarose gels. In addition, analyzing the data in distinct ways has tested several models. None of these models seemed to adequately fit the data presented over the entire range of DNA lengths, gel concentrations, and/or electric fields. The new methods used to analyze the electrophoretic mobilities provided further insight into the mechanisms involved for a molecule moving through a media. It has been shown that a simple exponential function (eq. 2.56) describes the electrophoretic mobility for DNA lengths over three orders of magnitude. The valid length range corresponds to the range of most interest for gene cloning and mapping, i.e., a few hundred to $\sim 50,000$ bp. For horizontal electrophoresis, good conformity with eq. 2.56 was observed over the full range of gel concentrations examined (0.5% to 1.3%) for electric fields of 2.5 to 5.0 V/cm. Data for low fields, down to 0.62 V/cm, were also fit well by eq. 2.56 when the gel concentration was low (0.4–0.7%). Data described best by eq. 2.56 over a large DNA length range were those obtained in moderate to low gel concentrations ($\leq 1\%$) at moderately high fields (2.5 - 5 V/cm).

The use of vertical gel electrophoresis helped provide a more uniform data set than horizontal gel electrophoresis. The quality of the fit, χ^2 , shows clear trends (fig. 4.8 and table 4.2) which has lead to an understanding for what DNA lengths, gel concentration, electric fields the different mechanisms involved occur during the migration of DNA molecules. Equation 2.56 describes the data best at moderate fields ($1.24 \text{ V/cm} \leq E \leq 3.10 \text{ V/cm}$) and for $T = 0.7\%$ and 1.0% . Data at low fields was also fit well when $T = 0.4\%$.

An analysis of reptation plots provides a context for understanding when DNA electrophoresis conforms to eq. 2.56 and when it does not. This context derives from considerations by Rousseau et al. [58–59] and others (see [8] and references therein) of

the relative contributions of Ogston sieving, entropic trapping, and biased reptation with and without fluctuations to electrophoretic transport of polyelectrolytes in gels. By this view the simple exponential decrease in drag with decreasing DNA length (expressed best in eq. 2.57) is associated with a smooth transition from the Ogston sieving to biased reptation regimes. Failure of this relation is associated with more complex behavior of the DNA chain. Entropic trapping is a reasonable possibility. Chain trapping is expected to occur when pores are significantly smaller than the average molecular cross-section of the mobile polymer, and is minimized by low gel concentrations and high fields.

This chapter provides great insight into the ability of equation 2.56 to describe the mobility dependence on length for varying fields and concentrations over a vast DNA length range. The next chapter will further discuss the meaning of these results and point out similarities and differences between the two different electrophoresis setups.

CHAPTER 5

DISCUSSION

The critical importance of the gel matrix for DNA separations by electrophoresis is well established. Stellwagen and co-workers have shown that the mobility of DNA in free solution is independent of length for $L > 400$ bp [124], and decreases with decreasing length for $L < 400$ bp. This behavior contrasts sharply to the general decrease in mobility with increasing length when electrophoresis occurs in gels. Mobilities of DNA in agarose gels also become independent of length, but this limit occurs only for quite long $L > 20$ kbp or longer, depending on gel concentration [8, 28, 125–126]. The challenge has been to describe how DNA molecules (and by inference other semi-rigid chains) moving in a field interact with the gel matrix, and how these interactions change as the chain length increases over several orders of magnitude.

The usual practice in plotting calibration curves for determining DNA lengths is to graph electrophoretic mobility vs. logarithm of length for a series of standards (e.g. see ref. [40] and [110]). This practice has little theoretical basis, and plots of μ vs. $\log_{10}(L)$ are seldom linear for more than a decade in L (Fig. 1.6a). Plots of $\log_{10}(\mu)$ versus $\log_{10}(L)$ are of more fundamental interest and are approximately sigmoidal (Fig. 1.6b), showing that mobility does not scale as a simple power of length ($\mu \neq \alpha L^\beta$) over any significant range of L . Migration by biased reptation is proposed to dominate when the effective DNA radius significantly exceeds the mean pore size (see section 2.4.3). As originally described by Lumpkin and Zimm [61, 78], and Slater and Noolandi [122], the entire DNA chain is envisioned to migrate within a

tube defined by the path through the gel of the first chain segment. The migration of the chain center of mass is biased in the field direction, but the random distribution of chain configurations is considered to be unperturbed during reptation. For this idealized model the ratio of the gel mobility of the reptating chain, μ , to the free solution mobility, μ_o , is predicted to depend inversely on length, i.e.,

$$\frac{\mu}{\mu_o} = \frac{\alpha}{L} . \quad (5.1)$$

It is well known that very long DNAs migrate through gels by an inchworm-like motion in which the DNA coil is alternatively stretched and relaxed. As discussed in section 2.4.3 at sufficiently long lengths, stretching during reptation is predicted to decrease or cause total loss of the dependence of mobility on length [8, 125].

Southern's work [10], which is discussed in section 2.4.5, showed that plots of reciprocal mobility versus length were linear over a wider range than traditional semi-log plots (μ vs. $\log(L)$). In fact, plots of $\mu(L)$ versus $1/L$ for the data collected in our study were seldom linear for a full decade (e.g. see the quasi-linear region for the $L = 10,000$ to 3000 bp range in length in Fig. 1.6c). By contrast, plots of $1/\mu(L)$ vs. L were relatively linear over the $L = 100$ to 4000 bp range (Fig. 1.6d). In both cases, the non-random trends make the fits unacceptable, but Southern's approach yields quasi-straight lines for a 40-fold range in length compared to a 3-fold range for $\mu(L)$ vs. $1/L$.

Southern's approach is a limiting form of the relations presented here, as is seen by expanding the $[1 - e^{-L/\gamma}]$ term in Eq. 2.53:

$$\frac{1}{\mu(L)} = \beta + \alpha \left(\frac{L}{\alpha} - \frac{L^2}{2!\gamma^2} + \frac{L^3}{3!\gamma^3} - \dots \right) . \quad (5.2)$$

Thus, Southern's results support the exponential dependence of mobility on length suggested here more strongly than they support the $\mu(L) \propto 1/L$ dependence predicted by the classic reptation model (see section 2.4.3).

Rousseau et al. has shown that plots of $3\mu L/\mu_o$ versus L are illuminating for electrophoresis of single stranded DNA in polyacrylamide gels [58]. These plots are referred to as “reptation plots.” For data taken with relatively high fields and low gel concentrations these plots increased monotonically, with slopes that decrease smoothly. By contrast, plots for data taken with low fields and high gel concentrations exhibited dips, and the slopes became non-monotonic. This slope inversion is considered a signature of the entropic trapping regime. Entropic trapping (see section 2.4.4) is envisioned to occur when a random coil polymer expands entropically into a gel vacancy, then must assume a smaller, low probability configuration to continue its forced migration through the gel.

Other models, which were described in section 2.4.5, fail to produce reasonable fits. Barkema et al. [35] claimed that their proposed model (see section 2.4.5) showed the data collapse to a single curve when plotted as $T^{5/2}vL^2$ versus TLE . From this plotting method they proposed an empirical equation (eq. 2.51). Figure 5.1 shows all the data for this dissertation analyzed this way and reasonable fits were not obtained. The solid black line in both figures is Barkema’s proposed model, which clearly deviates from both data sets.

The empirical equation proposed by Calladine et al. [36] also failed to fit the data for the entire range studied (see section 2.4.5). Figure 5.2a shows the vertical electrophoresis data plotted by Calladine’s method [36], $\mu/\mu_{1/2}$ versus $L/L_{1/2}$ where $\mu_{1/2}$ is 1/2 of the maximum mobility and $L_{1/2}$ is the length associated with that mobility. Their model, $\mu = \frac{2\mu_{1/2}}{(1+(L/L_{1/2})+0.01(L/L_{1/2})^3)E}$, seems to fit the mobilities for small DNA lengths, but at $L/L_{1/2} \sim 1$ the data starts to deviate from their fit.

The plotting method Calladine et al. use to display the quality of their fit can be very misleading. When trying to fit their equation to the data they always show plots of the electrophoretic velocities (or mobilities) normalized to the maximum velocity (they normalize to a variable f which equals $2v_{1/2}$). This velocity is related to the shortest length DNA used in the experiment. Also the DNA lengths were

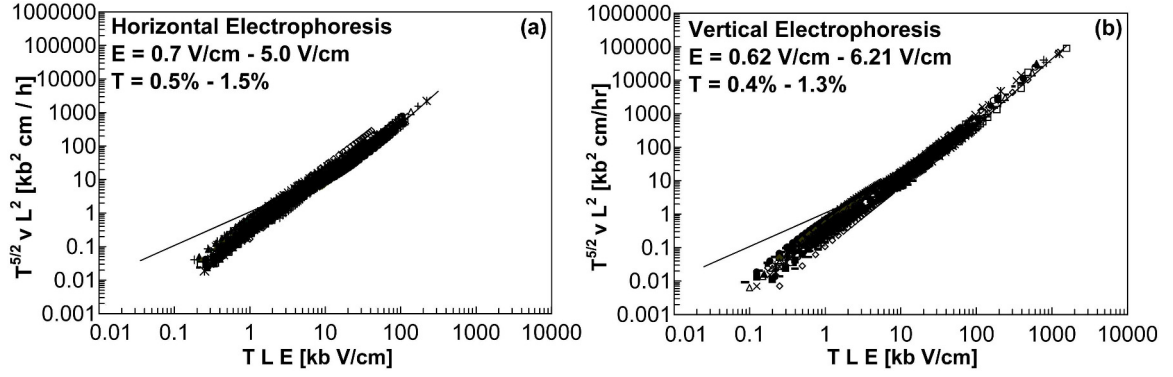


Figure 5.1. All the data for both horizontal (a) and vertical (b) electrophoresis using Barkema et al.'s plotting method [35], $T^{5/2}vL^2$ versus TLE . The solid black line in both plots represent a fit to the data using Barkema et al.'s equation: $\frac{T^{5/2}vL^2}{\alpha} = [(\frac{TLE}{\beta})^2 + (\frac{TLE}{\beta})^4]^{1/2}$, with α and β being free parameters.

normalized to the length associated with this velocity. Figure 5.2b shows the data from this dissertation plotted and fit by Calladine's technique. Similar results are obtained when comparing Calladine's data. Figure 5.2b plots the electrophoretic

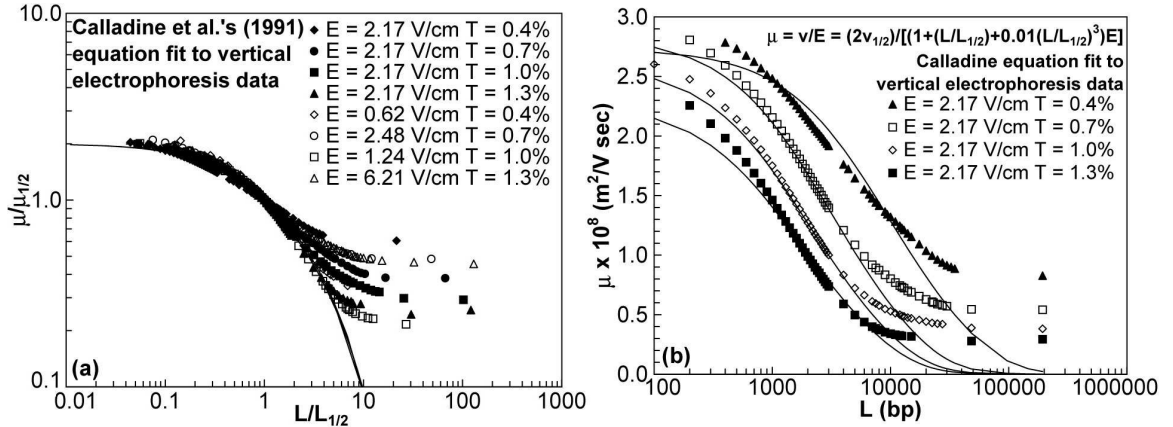


Figure 5.2. Fits of vertical electrophoretic data to Calladine et al.'s model [36], $\mu = \frac{2\mu_{1/2}}{(1+(L/L_{1/2})+0.01(L/L_{1/2})^3)E}$, where $\mu_{1/2}$ is 1/2 the maximum mobility and $L_{1/2}$ is the length associated with that mobility. Figure 5.2a plot the data using Calladine's method, $\mu/\mu_{1/2}$ versus $L/L_{1/2}$. Figure 5.2b is the typical method to plot electrophoretic mobilities. The solid black line represents a fit to the data using Calladine et al.'s equation.

mobilities for $E = 2.17$ V/cm and $T = 0.4\%, 0.7\%, 1.0\%$, and 1.3% as μ versus $\log L$ and the solid black lines represent Calladine's fit. The conventional way of plotting shows that Calladine's model is in fact a very poor representation of the electrophoretic mobilities (figure 5.2b). The maximum range of DNA lengths it covers for any gel concentration is from $L = 1000$ bp to 5000 bp. Calladine's plotting method (figure 5.2a) compresses all the mobilities related to short lengths, so their model seems to fit over a larger range of DNA lengths.

Relations commonly proposed to describe the dependence of DNA electrophoretic mobility on length are valid only over limited size ranges, generally a factor of ten or less. These size ranges can be related to models for transport of a particular molecule type through a gel network. Ogston sieving and reptation are two of the best known models and are used to describe electrophoresis of rigid molecules and of unperturbed, worm-like chains, respectively. The models describing the length dependence of mobility for these cases are quite distinct and essentially unrelated. Sieving occurs for relatively short lengths, low gel concentrations, and high fields. Reptation seems to apply for long lengths, low gel concentrations, and high fields.

It was found that a remarkably simple relationship with only a single, exponential term in L describes the electrophoretic mobilities of DNA molecules as small as 100 bp and as large as $194,000$ bp. The 100 bp molecules are substantially shorter than the DNA persistence length and behave approximately as rods with a modest axial ratio (≈ 14), while at 10 kbp the molecules are about 70 persistence lengths and are well described as random coils. The fitting function works best for relatively low gel concentrations and relatively high electric fields, where sieving and reptation apparently work. The exponential relationship described here should be a useful interpolation function for determining DNA lengths over a wide range. The simplicity of this relationship may be of more fundamental significance and suggests that some common feature dominates the electrophoresis of double stranded DNA fragments in agarose gels, regardless of length.

5.1 Trends in the Electrophoretic Mobilities

It was pleasing to find that the trends in the quality of the fits to electrophoretic data for the different experiments were consistent. In section 4.1.2 the χ^2 values for the horizontal electrophoresis were discussed. A three dimensional plot of χ^2 as a function of E and $\%T$ (figure 4.5) showed that the quality of the fit degrades as the field decreases with increasing gel concentration. The same trends were observed for χ^2 plots for vertical electrophoresis experiments (fig. 4.8). The horizontal electrophoresis experiments fit extremely well ($\chi^2 \geq 0.999$) to data with $E = 2.5$ V/cm to 5 V/cm for all $\%T$ and at low concentrations for all E . The vertical electrophoresis experiments fit extremely well at $T = 0.4\%$ for $E \leq 2.48$ V/cm. At $T = 0.7\%$ equation 2.56 fit the data with $\chi^2 \geq 0.999$ for $1.24 \text{ V/cm} \leq E \leq 3.10$ V/cm. At $T = 1.0\%$ equation 2.56 fit the data with $\chi^2 \geq 0.998$ for $1.86 \text{ V/cm} \leq E \leq 4.97$ V/cm. Finally at $T = 1.3\%$ the fits have $\chi^2 \geq 0.998$ for $E \geq 2.48$ V/cm. In other words, equation 2.56 describes the mobilities of double stranded DNA (and potentially ssDNA) very well for moderate to high fields and moderate to low gel concentrations within the ranges examined.

Deviations of the fits from data, where they occur were typically most pronounced for the longest DNA. The deviations and behaviors of fits were explained when the data were represented as “reptation plots” ($3\mu L/\mu_o$ vs. L). Recall that Rousseau, et al. [58] observed three regions with different slopes when looking at the reptation plot for ssDNA electrophoresed in polyacrylamide gels at low fields. For higher fields the middle region disappeared. A similar behavior was observed when the reptation plot method was applied to the present experiments. Three regions were observed for reptation plots in the conditions where the fits broke down (low fields and moderate to high gel concentrations). Equation 2.56 fit extremely well when two characteristic regions were observed in the reptation plots. The different regions are thought to correspond to different mechanisms dominating the interactions of the

molecules within the gel. The first region (short length, steep slope) corresponds to Ogston sieving and the last region (long length, lower positive slope) corresponds to reptation. Our new exponential fitting relation describes the data extraordinarily well when there is a smooth transition from the Ogston sieving regime to the reptation regime.

Rousseau, et al. [58], proposed that the third, middle region observed at low fields is associated with “entropic trapping.” They observed a region with a negative slope between the Ogston sieving and reptation regions for ssDNA electrophoresed in polyacrylamide gels. A similar region was seen in reptation plots for dsDNA electrophoresed in agarose gels. When equation 2.56 is applied to data containing entropic trapping the fit produces poor χ^2 values. Therefore, it is concluded that equation 2.56 does not account for entropic trapping.

5.2 Phase Diagrams

Phase diagrams were extracted from the reptation plots by identifying the different regimes and assigning boundary points (described in sections 4.3.1 and 4.3.2). The phase diagrams for both the horizontal (fig. 4.22) and vertical (figures 4.27 and 4.28) electrophoresis experiments displayed similar trends. The phase diagrams increase to the left with length as the gel concentrations increases and shift up with electric field. Rousseau, et al. [58] had also proposed a phase diagram for ssDNA electrophoresed in polyacrylamide and a similar trend was observed. As the gel concentration is increased the gel matrix is tighter and there will be smaller pore sizes. Smaller pores mean the transition from sieving to other transport mechanisms will occur at smaller DNA lengths.

The phase diagram of Rousseau, et al. [58] indicates that the length associated with the transition between Ogston sieving and reptation of single stranded DNA in polyacrylamide is independent of field and increases with decreasing %*T*. Similar trends were observed for L_o and βL_{aRP} for double stranded DNA in agarose

(Figures 4.22 and 4.27). This is understandable since the limiting length for Ogston sieving should depend primarily on the pore size distribution in the gel. Increasing electric field can elongate the random coil configuration of long DNA, but this should be significant only for lengths longer than those that can sieve. Rousseau et al. [58] further suggest that the transition from entropic trapping to reptation for single stranded DNA in polyacrylamide occurs at lower electric fields with decreasing % T , and is length independent. For both horizontal and vertical electrophoresis on double stranded DNA the field associated with the transition from excellent to poorer fits also occurs at lower fields as the gel concentration decreases.

The two different electrophoresis methods produced similar phase diagrams. There were only slight variations between the transitions. Since the transition lines were similar for both experiments and the vertical electrophoresis experiments produced cleaner results, the only results for vertical electrophoresis will be discussed. The transition line between Ogston sieving and reptation for $T = 1.0\%$ is at $L \approx 1500$ bp. At $T = 1.3\%$ the transition point for Ogston sieving to entropic trapping (ET) is at $E = 2.17$ V/cm for vertical electrophoresis (see fig. 4.28). It occurs at $L = 2300$ bp and $E = 1.55$ V/cm for the vertical setup. These values give a clear idea of where the transitions occur and can be used to determine the prime conditions for electrophoresis.

Plotting the phase diagram as a function of the radius of gyration, R_g , of the molecule draws a physical picture for the transitions (Figure 4.28b). Comparisons of these phase diagrams with the pore sizes for a specific gel concentration give a clearer picture for how the molecule migrates during electrophoresis. The transition lines describing the onset for different mechanisms occurring during molecular transport for Ogston sieving to reptation are all found to be smaller than the pore sizes. According to the Ogston sieving model [24] molecules smaller than the pore size should sieve through the gel matrix. Molecules exceeding the pore size will start to deform and reptate through the gel according to the reptation

theory [8, 25, 28, 65, 104, 122–123]. The transition lines for Ogston sieving to entropic trapping are approximately equivalent to the pore sizes. The values for the transition lines for entropic trapping to reptation are larger than the pore sizes. As stated in section 2.4.4, entropic trapping occurs when the molecule’s radius of gyration, R_g , is equivalent to the gel’s mean pore size, a , which forces the molecule to selectively hop between the largest pores. When R_g becomes much larger than a , the molecule is forced to squeeze through the pores and has the possibility to reptate through the gel. At this point the molecule will no longer be affected by entropic trapping [1, 6, 8, 37, 39].

5.3 The Best-Fit Parameters and Pore Size

The three free independent parameters in equation 2.56 provided for a better understanding of the mobility of the DNA during gel electrophoresis. The parameters also added information about the agarose gel structure. Polymer conformations were discussed in section 2.2 and a diagram (fig. 2.4) was presented illustrating the mechanisms occurring for different size DNA during gel electrophoresis. A physical picture starts to reveal itself by relating the parameters to figure 2.4, Slater [34] recently proposed methods to estimate the gel pore size from the parameters, and from the phase diagrams obtained from the reptation plots.

The short length mobility limit, μ_s was usually only slightly affected by the $\%T$ and E . For horizontal electrophoresis very little decrease in μ_s with T was observed at higher fields and no dependence was observed for μ_s on E , which indicates that its field independent. Vertical electrophoresis was in agreement with these trends and the values were fairly close. Some of the results for μ_s provided different results. As discussed in section 4.1.3 the short length mobility limit, μ_s , is related to a DNA fragment which acts like a rigid rod. The smallest fragment electrophoresed was $L = 200$ bp, which is approximately 68 nm. Figure 2.4A best illustrates how a small

length molecule will migrate through the gel by Ogston sieving. Tables 4.3 and 4.4 lists the pore sizes for agarose gels at different % T . For all the gel concentrations used, the pore size estimates are much larger than the DNA length associated with μ_s . The short length mobility limit seems to have little dependence on % T (fig. 4.12b).

The long length mobility limit, μ_L , have the same trends for all concentrations. An increase with field and a decrease with % T . For both horizontal and vertical electrophoresis μ_L increased $\sim 80\%$ from the lowest field to the highest field. Recall that at large enough lengths the mobility should be length independent and will only depend on the field. So, the prediction $\mu_L \propto E$ is not seen for the data collected in this dissertation (fig. 4.15). However it is observed that μ_L increases with increasing E .

The DNA length associated with the long length mobility, μ_L , is best illustrated using figure 2.4(C, D, and E). The two largest observable fragments were $L = 48,502$ bp and 194,000 bp. When entropic trapping did not occur, the fits using equation 2.56 reached a limiting mobility very close to the fit value μ_L (see fig. 4.9). Since equation 2.56 does not account for entropic trapping, the values for μ_L will deviate greatly from the data in such cases (see fig 4.10). Assuming that the persistence length, l_p , for dsDNA is ~ 50 nm, $L = 48,502$ bp and 194,000 bp are approximately 330 and 1319 times l_p . These lengths are best described using the biased reptation model as illustrated in figure 2.4(C, D, and E). The phase diagrams show that these lengths are well within the reptation regime (see figure 4.27). As described in section 4.1.3, the dependencies observed for μ_L as a function of E seems to agree with the reptation model ($\mu_L \propto E$) [8, 30]. The biased reptation model also predicted $\mu_L \propto 1/T^{2x}$. The results for μ_L as a function of % T does not agree with this (see fig. 4.15). Only for the higher fields does the concentration dependence predicted by the biased reptation model occur. Recall from section 4.1.3, that power law fits to plots of μ_L versus % T for $E \geq 4.97$ V/cm produced results that agreed with the reptation model.

The crossover length, γ , decreases as a function of E . Slater [34] proposed that γ is proportional to $N^*(\varepsilon)$ [30], which is defined as the crossover length of the molecule from unoriented reptation to oriented reptation. Since $N^*(\varepsilon)$ is inversely proportional to E , γ can assume an equation of the form $\gamma = cE^{-1}$, with c being a free parameter. This provides a reasonable representation for both horizontal and vertical electrophoresis. The values obtained for the crossover length, γ , fall between 3.6 kbp to 52.8 kbp (see fig. 4.16). Even the smallest size γ is approximately 25 persistence lengths. A molecule of that size can be considered a random coil. Recall, that a random coil will be in a globular state without any external forces acting on it. So there will be a radius of gyration associated with it. The calculated radius of gyrations for γ are shown in figure 4.17. The length that is associated with $N^*(\varepsilon)$ is also considered the length of the largest molecule that can be distinguished from other fragments or largest resolved fragment. The largest resolved fragment for most cases was $L \simeq 35$ kbp. For the lowest gel concentrations, the DNA fragments could be resolved to $L = 48.5$ kbp. Only a few values of γ fall with the definition for $N^*(\varepsilon)$. Even though γ can be reasonably fit to $\gamma = cE^{-1}$ (fig. 4.16), the relation to $N^*(\varepsilon)$ is not clear.

The Radius of gyration corresponding to γ , $R_g(\gamma)$, can easily be compared to the pore size calculated using Slater's method [34]. At $T = 1.0\%$ and 1.3% the pore sizes are calculated to be 146 nm and 122 nm (table 4.3). All values for $R_g(\gamma)$ are greater than these pore sizes. Figure 2.4(C, D, and E) would best illustrate the mechanisms involved during electrophoresis for the molecules associated with this size. At $T = 0.4\%$ and 0.7% all $R_g(\gamma)$ for $E \leq 3.1$ V/cm are larger than the gel's pore size, thus the mechanisms portrayed in figure 2.4C, D, and E would dominate. For $E = 4.97$ V/cm and 6.21 V/cm the $R_g(\gamma)$ is either equivalent to the pore size or smaller. Since the molecules are under the influence of higher fields the same mechanisms, such as reptation, would apply. Entropic trapping should occur if the value for $R_g(\gamma)$ was equivalent or smaller than the pore for low fields. For $T = 0.4\%$

the values for $R_g(\gamma)$ decrease with increasing E from 546 nm to 227 nm. The average pore size calculated for $T = 0.4\%$ is 240 nm. At higher fields the molecules associated with $R_g(\gamma)$ are equal to the pore size and will experience entropic trapping. The reptation plots do not show an entropic trapping region, but for $E \geq 2.17$ V/cm at $T = 0.4\%$ the quality of the fit becomes worse as the field increases. There might be a region where entropic trapping does occur, but can not be detected by the reptation plots. The rest of $R_g(\gamma)$ values still decrease with increasing E , but is always greater than the average pore size for a given gel concentration.

5.4 Summary

This chapter provides great insight into the ability of equation 2.56 to describe the mobility dependence on length for varying fields and concentrations over a large DNA length range. The reciprocal of the electrophoretic mobility of a particle is a measure of the effective drag offered by the medium. Equation 2.56 states, in essence, that the decrease in drag from some long-length limit follows an exponential in length. The exponential relationship described here should also be a useful interpolation function for determining DNA lengths over a wide range.

One question that still needs attention is how equation 2.56 relates with electrophoresis experiments done with different type gels, with ssDNA, and other types of experiments. The next chapter provides a preliminary investigation using data gathered from the literature to discover the versatility of this equation for many types of separation experiments reported.

CHAPTER 6

FITS TO DATA OBTAINED FROM THE LITERATURE

The focus of this dissertation has so far been on analysis of dsDNA electrophoresis in agarose gels at Florida State University. Now that a thorough investigation has been done on using a simple exponential equation (eq. 2.56) to fit our data, the question can data from different media, molecules, and types of electrophoresis experiments fit as well? This question has been studied by gathering data from the literature. The results show that equation 2.56 fits separations data for different types of experiments.

6.1 Method

Data were obtained from several publications with varied experimental conditions. Table 6.1 lists the sources of data that were analyzed with the specific type of materials used and the type of electrophoresis experiments performed in each case. The mobilities were calculated from mobility versus length plots in the publications using TechDig (obtained from an independent source on the web, Ron Jones, ronjones@xnet.com, Mundelein, IL, USA). The figures from the literature were first digitized using a conventional scanner. Images were then displayed in TechDig where the data points on the plot were individually selected and digitized. The obtained x, y coordinates were then converted into μ versus L data. As before, fits to equation 2.56 from the mobilities were generated by MathCad (MathSoft, Inc., Cambridge, MA, USA).

Table 6.1. Summary of the data obtained from the literature

Source	Gel ^a	Experiment ^b	Molecule Separated
Slater (1988) [106]	Agarose	Conv	dsDNA
Stellwagen (1985) [29]	Agarose	Conv	dsDNA
Hervet (1987) [111]	Agarose	Conv	dsDNA
Calladine (1991) [36]	Agarose	Conv	dsDNA
Heller (1994) [64]	Agarose	CHEF	dsDNA
Park (2001) [127]	Agarose	EC	dsDNA
Chiari (1995) [128]	Agarose	Conv	dsDNA
	Hybrid		
	pAAm		
Brahmasandra (2001) [129]	pAAm	Conv	ssDNA
Rousseau (1997) [58]	pAAm	Conv	ssDNA
Stellwagen (2002) [130]	pAAm	CE	dsDNA

^a pAAm = polyacrylamide gels, Hybrid = a gel with a mixture of agarose and polyacrylamide

^b Conv. = conventional gel electrophoresis, CHEF = Contour-clamped homogeneous electric field, CE = capillary electrophoresis, and EC = electrochromatography

6.2 Results

Conventional electrophoresis in agarose gels is a common technique used for separating macromolecules. Table 6.1 shows five sources with the same experimental conditions as the data reported in this dissertation, which is dsDNA separated by conventional electrophoresis in agarose gels. The experiments performed by the sources with the same experimental conditions [106, 111, 29, 36] were done in a horizontal setup similar to section 3.2.1 and 3.2.2. The rest of the sources either used different media, such as polyacrylamide (pAAm) (see the following references for more information on pAAm [13, 131]), or different type electrophoresis experiments. Table 6.2 lists the maximum and minimum lengths gathered from the figures using TechDig. The free solution mobilities were taken from the values quoted in the sources.

Table 6.2. The quality of the fits to data obtained from the literature

Source ^a	E (V/cm)	% T	% C	L_{max} ^b (bp)	L_{min} ^b (bp)	$\mu_o \times 10^8$ ^c (m ² /Vsec)	χ^2 ^d
Brahmasandra (2001)	40.0	6.0	5.0	735	26	2.80	0.9923
	40.0	9.0	5.0	717	28	2.80	0.9968
	40.0	12	5.0	707	29	2.80	0.9962
Park (2001)	0.50	2.0	-	29418	310	4.00	0.9940
	1.50	2.0	-	29418	310	4.00	0.9926
	3.00	2.0	-	29418	310	4.00	0.9947
	5.00	2.0	-	29418	310	4.00	0.9950
	8.00	2.0	-	29418	310	4.00	0.9946
	10.0	2.0	-	29418	310	4.00	0.9956
Chiari (1995)	10.0	10	4.7	4412	65	2.80	0.9932
	8.00	10	4.7	4447	66	2.80	0.9961
	6.00	10	4.7	4410	65	2.80	0.9977
	4.00	10	4.7	4408	65	2.80	0.9986
	3.00	10	4.7	4389	65	2.80	0.9962
	10.0	2.0	10	4274	63	2.80	0.9987
	10.0	4.0	3.3	4289	63	3.00	0.9982
	10.0	1.2	-	4275	122	3.00	0.9977
Stellwagen (2002)	3.30	3.5	3.0	1375	127	4.50	0.9995
	3.30	5.7	3.0	1645	125	4.50	0.9961
	3.30	3.5	3.0	1391	176	4.50	0.9949
	3.30	5.7	3.0	1385	125	4.50	0.9901
Stellwagen (1985)	3.80	0.2	-	12216	506	2.95	0.9953
	3.80	0.4	-	12216	506	2.95	0.9985
	3.80	0.5	-	12216	506	2.95	0.9992
	3.80	0.6	-	12216	506	2.95	0.9989
	3.80	0.7	-	12216	506	2.95	0.9992
	3.80	0.8	-	12216	506	2.95	0.9993
	3.80	0.9	-	12216	506	2.95	0.9995
	3.80	1.0	-	12216	506	2.95	0.9999
	3.80	1.25	-	12216	506	2.95	0.9998
	3.80	1.5	-	12216	506	2.95	0.9996
continued on next page							

Table 6.2 – continued from previous page

Source ^a	E (V/cm)	% T	% C	L_{max} ^b (bp)	L_{min} ^b (bp)	$\mu_o \times 10^8$ ^c (m ² /Vsec)	χ^2 ^d
	2.60	0.6	-	12216	506	2.91	0.9997
	2.60	1.5	-	12216	506	2.91	0.9984
	1.30	0.6	-	12216	506	2.81	0.9993
	1.30	1.5	-	12216	506	2.81	0.9861
	0.64	0.2	-	12216	506	2.52	0.9983
	0.64	0.3	-	12216	506	2.52	0.9993
	0.64	0.4	-	12216	506	2.52	0.9990
	0.64	0.5	-	12216	506	2.52	0.9994
	0.64	0.6	-	12216	506	2.52	0.9995
	0.64	0.7	-	12216	506	2.52	0.9995
	0.64	0.8	-	12216	506	2.52	0.9985
	0.64	0.9	-	12216	506	2.52	0.9983
	0.64	1.0	-	12216	506	2.52	0.9940
	0.64	1.25	-	12216	506	2.52	0.9935
	0.64	1.5	-	12216	506	2.52	0.9945
Rousseau (1997)	18.30	4.0	5.0	1406	37	3.80	0.9296
	18.3	7.0	5.0	498	47	3.80	0.9693
	18.3	10	5.0	451	47	3.80	0.9929
	77.2	4.0	5.0	1102	41	3.80	0.9946
	9.62	4.0	5.0	1200	52	3.80	0.9900
Slater (1988)	0.61	0.4	-	48572	1999	2.50	0.9946
	0.61	0.5	-	48572	1999	2.50	0.9978
	0.61	0.6	-	48500	1999	2.50	0.9963
	0.61	0.7	-	48572	1999	2.50	0.9978
	0.61	0.8	-	48572	1999	2.50	0.9965
Hervet (1987)	8.00	0.6	-	24176	327	3.34	0.9993
	8.00	1.0	-	23056	319	3.34	0.9993
	8.00	2.0	-	23695	319	3.34	0.9991
	5.00	0.6	-	24374	328	3.30	0.9999
	5.00	1.0	-	23388	320	3.30	0.9995
	5.00	2.0	-	24144	317	3.30	0.9992
	3.00	0.6	-	23874	326	3.10	0.9998
	3.00	1.0	-	23056	319	3.10	0.9999
	3.00	2.0	-	23739	313	3.10	0.9972
	1.00	0.6	-	23666	319	3.07	0.9997
	1.00	1.0	-	23222	320	3.07	0.9997
	1.00	2.0	-	23853	311	3.07	0.9981

continued on next page

Table 6.2 – continued from previous page

Source ^a	E (V/cm)	% T	% C	L_{max} ^b (bp)	L_{min} ^b (bp)	$\mu_o \times 10^8$ ^c (m ² /Vsec)	χ^2 ^d
Heller (1994)	10.00	1.0	-	48915	2006	8.47	0.9981
	10.0	0.8	-	48888	2021	8.47	0.9972
	5.00	1.0	-	48915	1987	1.92	0.9950
	5.00	0.8	-	48859	2020	1.92	0.9993
	3.00	1.0	-	48915	1987	1.64	0.9999
	3.00	0.8	-	48820	2020	1.63	0.9983
	2.00	1.0	-	48451	1987	1.63	0.9988
	2.00	0.8	-	48334	2019	1.63	0.9988
	1.50	1.0	-	48451	1987	1.22	0.9965
	1.50	0.8	-	48326	2019	1.22	0.9976
	1.00	1.0	-	48451	1987	1.35	0.9922
	1.00	0.8	-	48746	2019	1.35	0.9970
	0.75	1.0	-	48451	1987	8.30	0.9770
	0.75	0.8	-	48257	2000	8.30	0.9956
	0.50	1.0	-	48451	1987	1.21	0.9451
	0.50	0.8	-	48689	2000	1.21	0.9971
	0.35	1.0	-	48451	1987	1.30	0.9955
	0.35	0.8	-	48675	2000	1.30	0.9971
	0.25	1.0	-	48451	1987	1.13	0.9963
	0.25	0.8	-	48631	2019	1.13	0.9967
	0.20	1.0	-	47991	1987	1.29	0.9959
	0.20	0.8	-	48163	2019	1.29	0.9955
	0.13	1.0	-	47991	2006	8.96	0.9972
	0.13	0.8	-	48126	2019	8.96	0.9946
Calladine (1991)	1.00	0.5	-	48187	9	2.41	0.9967
	1.00	1.0	-	47931	9	2.41	0.9957
	1.00	1.5	-	47127	9	2.41	0.9885
	1.00	2.0	-	47695	9	2.41	0.9914
	2.50	0.5	-	48211	9	2.96	0.9986
	2.50	1.0	-	48227	9	2.96	0.9971
	2.50	1.5	-	48239	9	2.96	0.9960
	2.50	2.0	-	47697	9	2.96	0.9928

^a The sources are the same references as in table 6.1^b The maximum and minimum lengths extrapolated from figures in the specified literature^c The free solution mobility stated in the specified literature^d The χ^2 values for fits to equation 2.56 for data taken from the specified source

Table 6.2 lists the quality of the fits from the specified sources evaluated by the χ^2 . The same noticeable trends are observed from the table for each data set. Figure 6.1 plots χ^2 against E and $\%T$ for Stellwagen [29] and Hervet and Bean [111]. For these two sources the same trends occur. At low fields with increasing concentration the quality of the fit degrades ($\chi^2 < 0.999$). For all fields at low concentrations ($T \leq 0.6\%$) the $\chi^2 \geq 0.999$. For all concentrations at large fields $\chi^2 \geq 0.999$. For Stellwagen's data (fig. 6.1a) the quality of fit is ≥ 0.999 at $E \geq 2.5$ V/cm, while for Hervet and Bean (fig. 6.1b) it occurs at $E \geq 5$ V/cm. A similar behavior occurs for the other data sets.

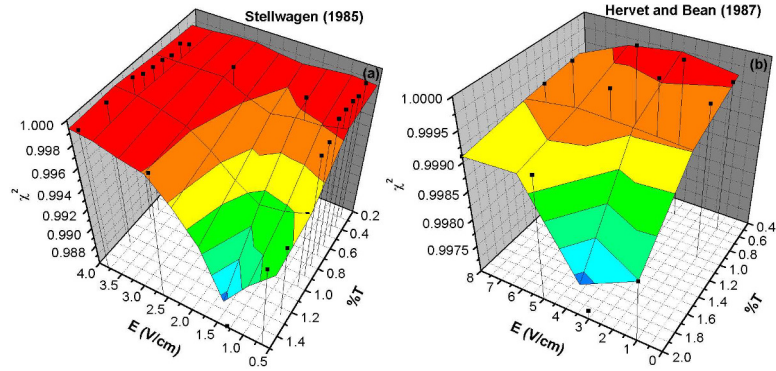


Figure 6.1. The influences of electric field strength and gel concentration on fit quality for Stellwagen (1985) [29] (a) and Hervet and Bean (1987) [111] (b). Chi-squared (χ^2) values calculated using fitting parameters listed in table 6.2 are plotted against E and $\%T$.

Although the trends are the same, χ^2 values do not correspond from experiment to experiment, even for sources with the same experimental conditions. This is quite evident in figure 6.1. Fits from equation 2.56 to both data sets show qualitatively the same trends in χ^2 , but the absolute values of χ^2 vary significantly among the data analyzed. For example the residuals from Stellwagen's mobility versus length plot (see figure 6.2(a and b)) becomes $\geq 10\%$ at $T \geq 0.8\%$, while for Slater, et al. [106] see (figure 6.3(a and b)) the residuals $\geq 10\%$ at $T \geq 0.5\%$. The difference in fields between Stellwagen's data (dsDNA electrophoresed at $E = 0.64$ V/cm) and

Slater's data ($E = 0.61$ V/cm) are small enough that the mobilities should be very similar. Figures 6.4 and 6.5 show that even the mobility versus DNA length plot can have a different shape for electrophoresis experiments with the same conditions. The mobilities for Hervet and Bean (Fig. 6.4) converge while for Stellwagen (Fig. 6.5) they diverge. The shape of Stellwagen's data is similar to the data presented in previous chapters.

The reptation plots, $3\mu L/\mu_o$ versus L , for each source have the same trends (figures 6.2c, 6.3c, 6.4c, and 6.5c). As the gel concentration increases for low fields the reptation plots seem to develop an "entropic trapping" region. A negative slope or a slope equal to zero appears between two positive slopes. For example in figure 6.3 for $T = 0.4\%$ there only exists two smooth regions. At $T = 0.5\%$ there appears a fairly flat slope from $L \sim 10,000$ bp to $\sim 25,000$ bp. As the gel concentration increases this flat region becomes more prominent.

The focus so far has been on conventional electrophoresis of dsDNA in agarose gels. Table 6.1 includes data from several other sources that used other media, molecules, or methods of separations. Equation 2.56 fits these data as well (table 6.2) and shows the same trends in χ^2 with changing E or $\%T$ as noted previously. Figures 6.6 and 6.7 are examples of these fits. Figure 6.6 was taken from a contour-clamped homogeneous electric field [132–133], or CHEF, experiment. The experiments were done with dsDNA in agarose gels with fields ranging from $E = 0.13$ V/cm to 10 V/cm. The reptation plot (fig. 6.6c) and residuals (fig. 6.6b) show that as the field decreases the fit becomes worse. This also corresponds with the "entropic trapping" region which appears at low fields and is not accounted for in equation 2.56.

Equation 2.56 works well with both a different experimental setup and separation media. Figure 6.7 shows electrophoretic data from a capillary electrophoresis (CE) [1, 134–139] experiment done on a dsDNA in a polyacrylamide gel. Capillary electrophoresis is a relatively new technique that allows for high separation efficiency requiring only small amounts of material. Stellwagen, et al. used two different type

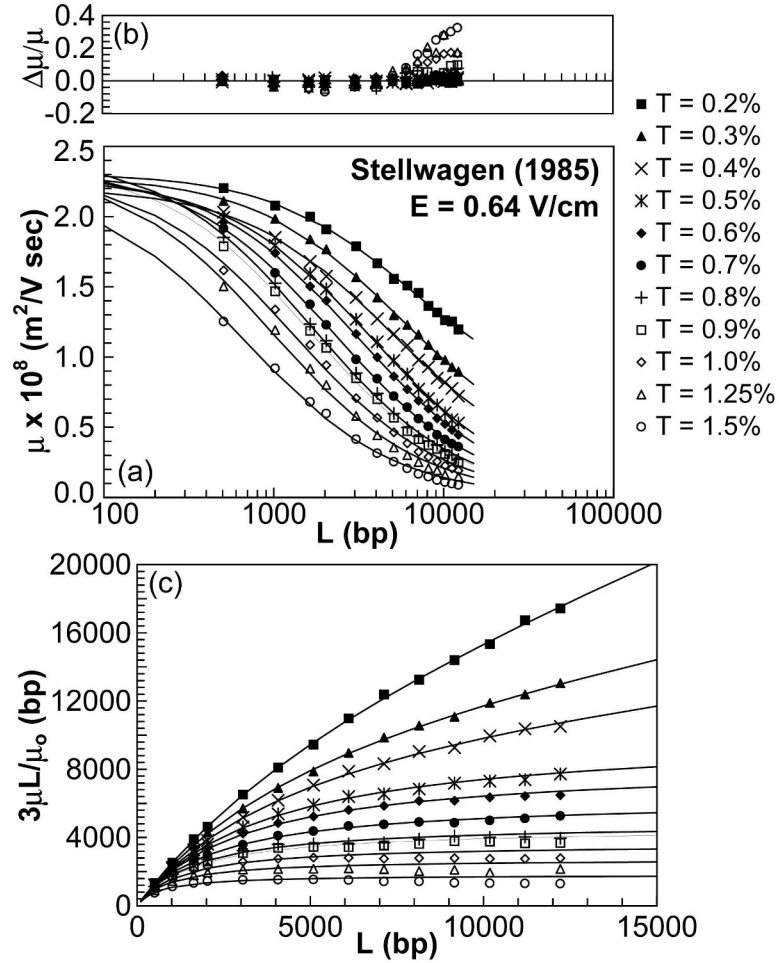


Figure 6.2. Dependence of mobility on length at $E = 0.64$ V/cm, for $T = 0.2\% - 1.5\%$ for Stellwagen [29]. Solid lines were calculated from the best fits to equation 2.56 using parameters in Table 4.2. The residuals expressed as percent deviations between calculated and measured values are shown in the upper panel (b). The lower panel (c) shows the reptation plots for the same E and $\%T$. The function $(3\mu L/\mu_o)$ is plotted against L , where μ_o is the length-independent free solution mobility obtained from the literature.

of dsDNA, 123 bp ladder and a 12A multimer (for more information see ref. [130]). A multimer is a subclass of the molecule to represent aggregates of one or more of the original monomers. The plots show that for the 123 bp ladder equation 2.56 represents the data very well with a $\chi^2 = 0.9995$ for $T = 3.5\%$. The quality of the fit decreases as the concentration increases. When looking at the reptation plot

(fig. 6.7c), it seems as if an “entropic trapping” region is appearing at the higher concentration, although the sparseness and uncertainty of the data may overwhelm our ability to fit.

Parameters from fits to the published data provide some interesting and familiar results. Table 6.3 lists all the best fit parameters to equation 2.56 for all the same

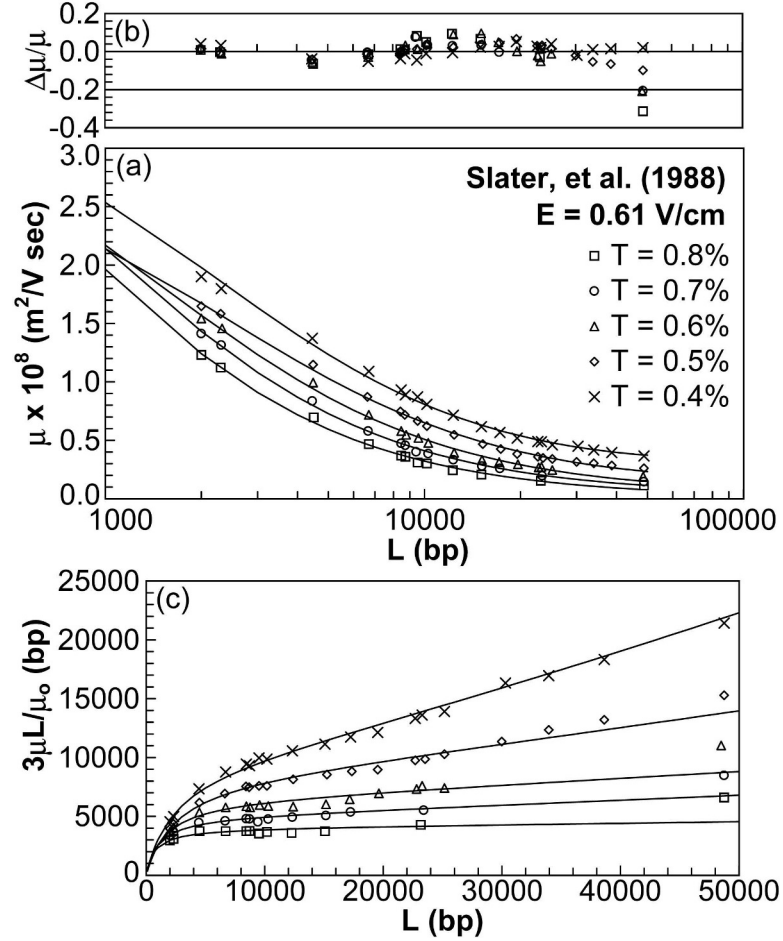


Figure 6.3. Dependence of mobility on length at $E = 0.61$ V/cm, for $T = 0.4\% - 0.8\%$ for Slater, et al. [106]. Solid lines were calculated from the best fits to equation 2.56 using parameters in Table 4.2. The residuals expressed as percent deviations between calculated and measured values are shown in the upper panel (b). The lower panel (c) shows the reptation plots for the same E and $\%T$. The function $(3\mu L/\mu_0)$ is plotted against L , where μ_0 is the length-independent free solution mobility obtained from the literature.

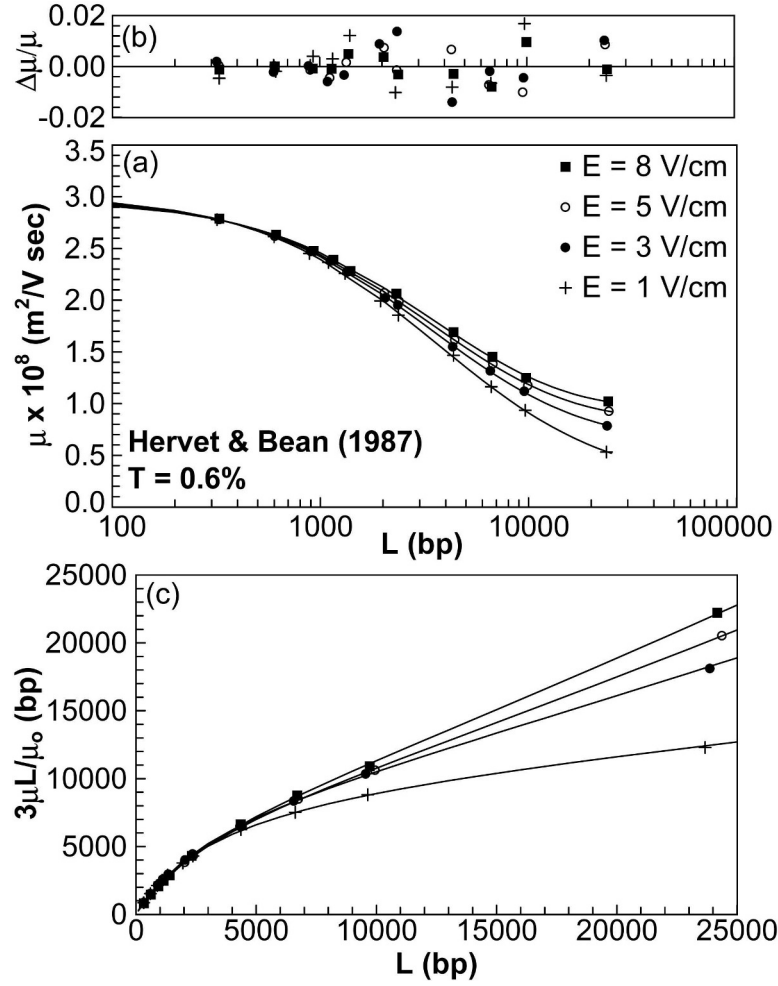


Figure 6.4. Dependence of mobility on length at $T = 0.6\%$, for $E = 1 - 8 \text{ V/cm}$ for Hervet and Bean [111]. Solid lines were calculated from the best fits to equation 2.56 using parameters in Table 4.2. The residuals expressed as percent deviations between calculated and measured values are shown in the upper panel (b). The lower panel (c) shows the reptation plots for the same E and $\%T$. The function $(3\mu L/\mu_0)$ is plotted against L , where μ_0 is the length-independent free solution mobility obtained from the literature.

sources listed in table 6.1. Notice that μ_L becomes extremely small and γ increases to a very large value for low fields and higher concentrations. This can be attributed to the regions where entropic trapping occurs. For example for Heller, et al. [64] at $E \leq 0.35 \text{ V/cm}$ the parameters do not seem reasonable when compared with the values of parameters when fit to our data. In figure 6.6c, even though the

data is sparse, there appears to be a region which can be associated with entropic trapping. In this region the fit is expected to become poor, with μ_L becoming small and γ becoming large. For all data taken from different sources the trends in the electrophoretic mobilities were very similar to our data. Figure 6.8 show the parameters for Heller's data as a function of E . Figure 6.9 show the parameters

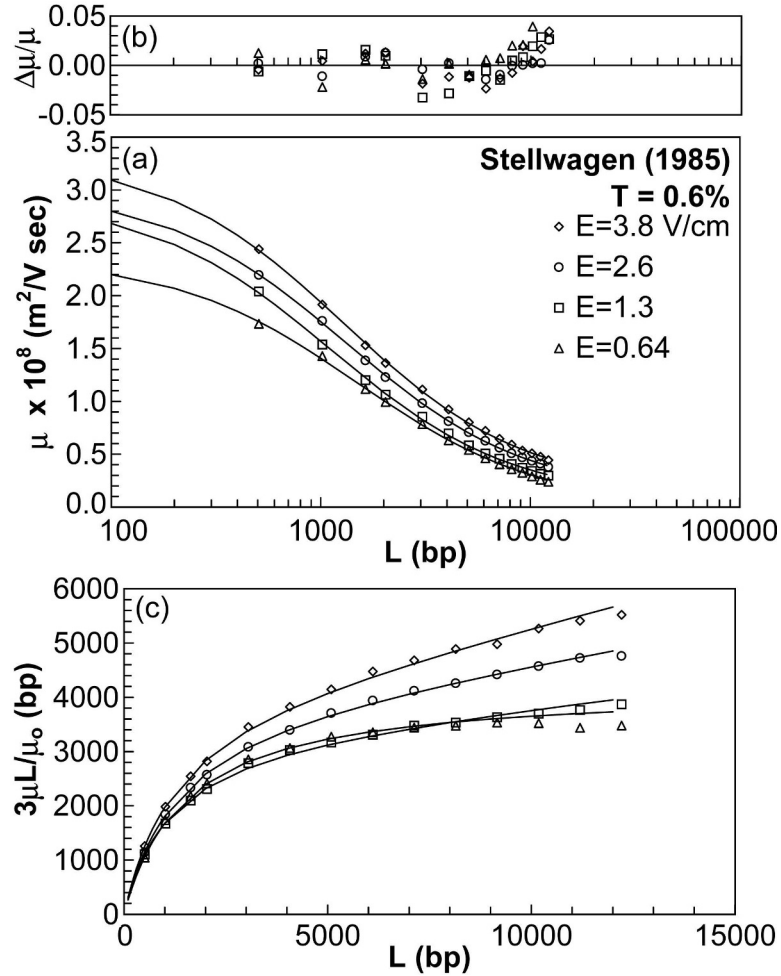


Figure 6.5. Dependence of mobility on length at $T = 0.6\%$, for $E = 0.64 - 3.80 \text{ V/cm}$ for Stellwagen [29]. Solid lines were calculated from the best fits to equation 2.56 using parameters in Table 4.2. The residuals expressed as percent deviations between calculated and measured values are shown in the upper panel (b). The lower panel (c) shows the reptation plots for the same E and $\%T$. The function $(3\mu L/\mu_o)$ is plotted against L , where μ_o is the length-independent free solution mobility obtained from the literature.

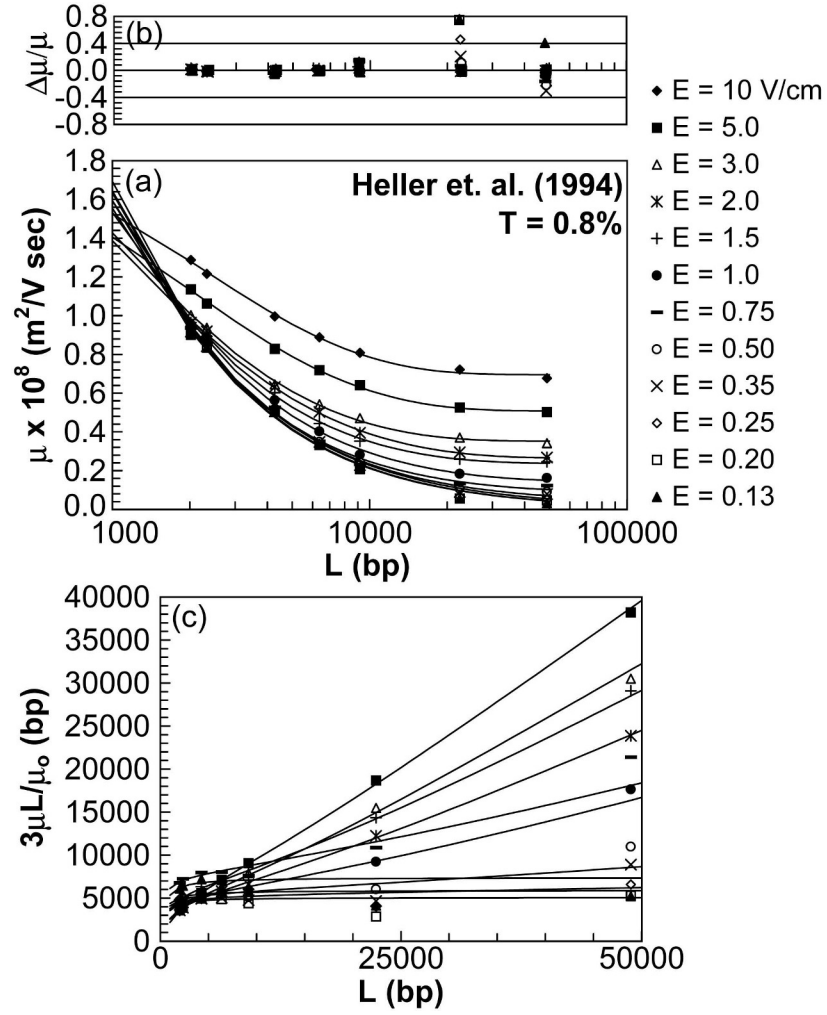


Figure 6.6. Dependence of mobility on length at $T = 0.8\%$, for $E = 0.13 - 10.00$ V/cm for Heller, et al. [64]. Solid lines were calculated from the best fits to equation 2.56 using parameters in Table 4.2. The residuals expressed as percent deviations between calculated and measured values are shown in the upper panel (b). The lower panel (c) shows the reptation plots for the same E and $\%T$. The function $(3\mu L/\mu_0)$ is plotted against L , where μ_0 is the length-independent free solution mobility obtained from the literature.

for Slater's data with as a function of $\%T$. The long length mobility in all cases increased with field (fig. 6.8) and decreased with concentration (fig. 6.9). Figure 6.8c demonstrates that for a different type of experiment μ_L will follow the same trend. The cross-over length, γ , in all cases decreases with E and increases with $\%T$. When

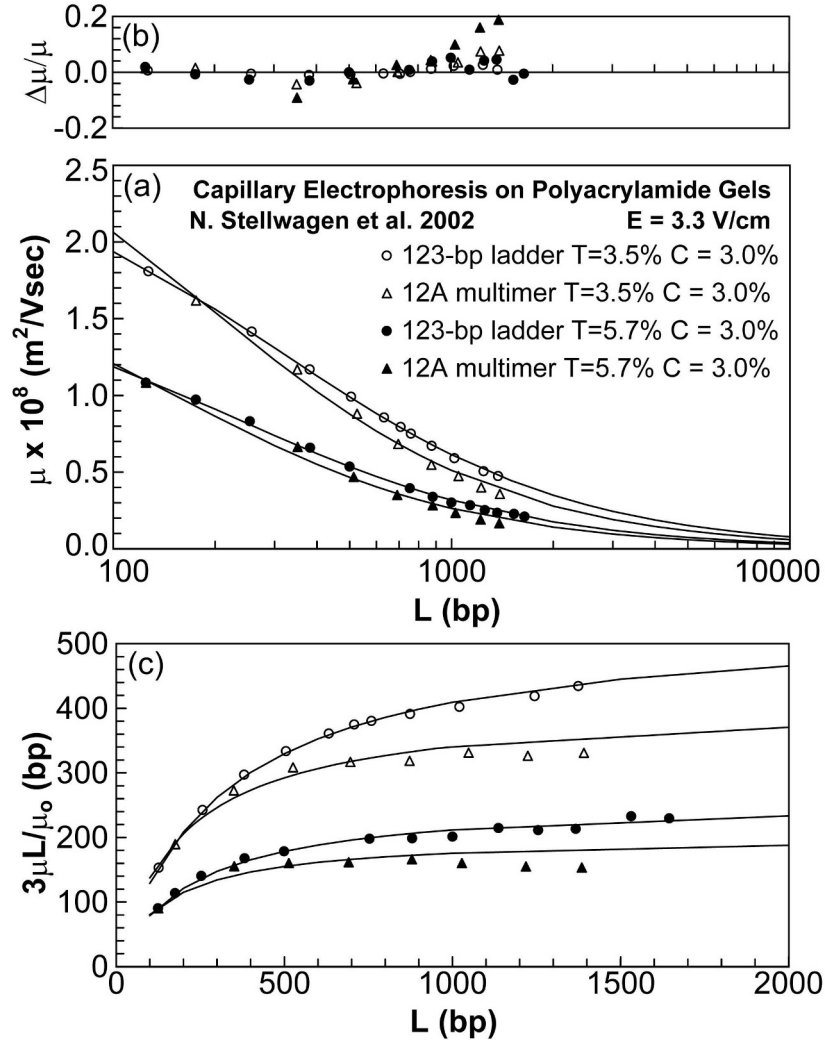


Figure 6.7. Dependence of mobility on length at $E = 3.3 \text{ V/cm}$, for $T = 3.5\%$ and 5.7% and $C = 3.0\%$ for Stellwagen, et al. [130]. $\%C$ is the cross-link concentration. Solid lines were calculated from the best fits to equation 2.56 using parameters in Table 4.2. The residuals expressed as percent deviations between calculated and measured values are shown in the upper panel (b). The lower panel (c) shows the reptation plots for the same E and $\%T$. The function $(3\mu L/\mu_0)$ is plotted against L , where μ_0 is the length-independent free solution mobility obtained from the literature.

a fit to an equation of the form $\gamma = cE^{-1}$, is applied to figure 6.8b (not shown) the result is $\gamma = 33E^{-1}$, with $\chi^2 = 0.6921$. This was done since for our data it was shown that γ can be best-represented by that dependence. The short length mobility

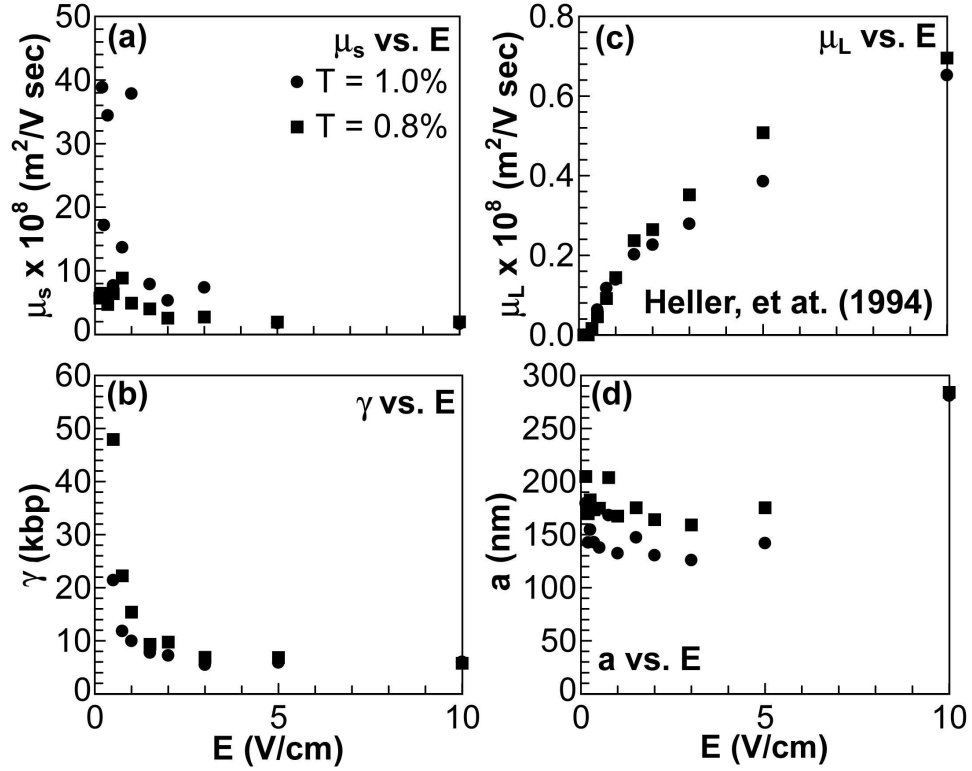


Figure 6.8. Dependencies of the limiting mobilities, μ_s and μ_L , the crossover length γ , and the pore size, a , on E for $T = 0.8\%$ and 1.0% from Heller, et al. [64](refer to Eq. 2.56).

limit, μ_s varied from data set to data set. In one series of experiments done with conventional electrophoresis μ_s varied from 2.5×10^{-8} to 5.0×10^{-8} (see figure 6.9a). Experiments with different conditions yield a different behavior for μ_s at lower fields. For example in the CHEF experiment performed by Heller, et al. [64], figure 6.8a shows at low fields the μ_s is extremely large compared to other experiments and then decreases to a fairly constant value for $E \geq 1.0$ V/cm. For $E \geq 1.0$ V/cm μ_s decreases only slightly.

The pore size, a , is also listed for each data set on table 6.3. Pore sizes were calculated from the best fit parameters using equation 4.3 described in section 4.2. The pore sizes calculated for agarose gels are in agreement with the previous chapters. For example in table 6.3 the pore size from the best-fit parameters for Slater, et al.

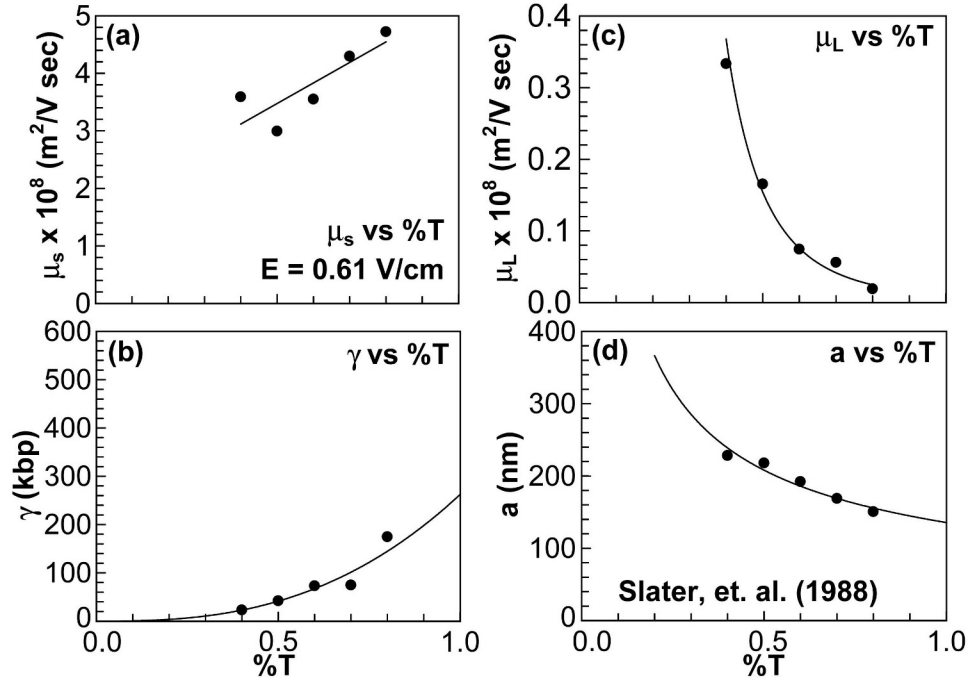


Figure 6.9. Dependencies of the limiting mobilities, μ_s and μ_L , the crossover length γ , and the pore size, a , on %T for $T = 0.8\%$ and 1.0% from Slater, et al. [106](refer to Eq. 2.56). The following equations, represented by solid lines, describe the dependencies: $\mu_s = 3.57T + 1.69$, $\mu_L = 0.01T^{-3.90}$, $\gamma = 262.18T^{2.67}$ and $a = 135.62T^{-0.62}$, determined using least-squares fitting methods.

(1988) at $T = 0.4\%$ was calculated to be 228 nm. In section 4.2 the pore size from the data taken for this dissertation at $T = 0.4\%$ was 242 nm. Table 4.4 lists agarose pore sizes for $T = 0.4\%$ ranging from 158 nm to 248 nm.

The experiments performing electrophoresis with polyacrylamide gels calculated pore sizes to range from ~ 4 nm for $T = 12\%$ and $C = 5\%$ to ~ 8 nm for $T = 4\%$ and $C = 5\%$. These values are in close agreement with what has been reported in the literature (see ref. [13, 131, 140–141]).

Table 6.3. The parameters of the fits to data obtained from the literature

Source ^a	E (V/cm)	$\%T$	$\mu_s \times 10^8$ ^b (m ² /Vsec)	$\mu_L \times 10^8$ ^b (m ² /Vsec)	γ ^b (kbp)	a ^c (nm)
Brahmasandra (2001)	40.0	6.0	1.97	6.50×10^{-6}	26286	7
	40.0	9.0	1.85	1.40×10^{-3}	63	5
	40.0	12	1.04	0.01	4	4
Park (2001)	0.50	2.0	2.49	8.80×10^{-7}	2.96×10^6	105
	1.50	2.0	2.48	0.02	187	109
	3.00	2.0	2.58	0.16	16	103
	5.00	2.0	2.77	0.27	9	101
	8.00	2.0	2.77	0.33	7	102
	10.0	2.0	2.86	0.39	6	102
Chiari (1995)	10.0	10	0.91	4.64×10^{-7}	7.26×10^5	45
	8.00	10	1.20	2.36×10^{-11}	1.74×10^{10}	50
	6.00	10	1.37	1.12×10^{-3}	450	55
	4.00	10	1.85	0.12	5	62
	3.00	10	2.13	0.20	4	71
	10.0	2.0	2.73	0.24	5	84
	10.0	4.0	1.54	0.06	7	47
	10.0	1.2	1.95	0.08	36	130
Stellwagen (2002)	3.30	3.5	2.55	9.66×10^{-7}	8.37×10^5	55
	3.30	5.7	1.71	5.30×10^{-10}	7.36×10^8	38
	3.30	3.5	3.12	2.33×10^{-6}	2.62×10^5	48
	3.30	5.7	2.01	6.17×10^{-7}	4.91×10^5	34
Stellwagen (1985)	3.80	0.2	3.02	0.94	22	347
	3.80	0.4	3.04	0.55	29	306
	3.80	0.5	3.13	0.59	18	245
	3.80	0.6	3.13	0.59	14	215
	3.80	0.7	3.13	0.60	10	185
	3.80	0.8	3.22	0.54	8	162
	3.80	0.9	3.14	0.50	9	157
	3.80	1.0	3.15	0.46	8	147
	3.80	1.25	3.07	0.41	7	133
	3.80	1.5	2.83	0.33	7	113
continued on next page						

Table 6.3 – continued from previous page

Source ^a	E (V/cm)	% T	$\mu_s \times 10^8$ ^b (m ² /Vsec)	$\mu_L \times 10^8$ ^b (m ² /Vsec)	γ ^b (kbp)	a ^c (nm)
	2.60	0.6	3.00	0.19	21	150
	2.60	1.5	2.26	0.15	16	120
	1.30	0.6	2.92	0.12	26	139
	1.30	1.5	2.33	3.27×10^{-4}	5967	109
	0.64	0.2	2.31	0.83	17	307
	0.64	0.3	2.29	0.42	28	283
	0.64	0.4	2.21	0.26	39	262
	0.64	0.5	2.28	5.96×10^{-9}	1.43×10^9	240
	0.64	0.6	2.31	6.97×10^{-10}	1.01×10^{10}	218
	0.64	0.7	2.35	9.31×10^{-8}	5.65×10^7	188
	0.64	0.8	2.44	1.99×10^{-7}	2.04×10^7	166
	0.64	0.9	2.38	3.23×10^{11}	1.19×10^{11}	161
	0.64	1.0	2.31	3.25×10^{-5}	93186	143
	0.64	1.25	2.35	5.19×10^{-6}	441570	124
	0.64	1.5	2.22	3.70×10^{-8}	4.09×10^7	101
Rousseau (1997)	18.30	4.0	2.79	9.20×10^{-6}	32058	8
	18.3	7.0	3.21	1.26×10^{-6}	108158	6
	18.3	10	13.2	3.01×10^{-12}	1.15×10^{10}	3
	77.2	4.0	3.00	8.50×10^{-8}	3.03×10^6	8
	9.62	4.0	4.33	5.50×10^{-9}	5.19×10^7	8
Slater (1988)	0.61	0.4	3.59	0.33	23	228
	0.61	0.5	3.00	0.17	42	218
	0.61	0.6	3.55	0.07	73	192
	0.61	0.7	4.30	0.06	75	169
	0.61	0.8	4.72	0.02	175	151
Hervet (1987)	8.00	0.6	2.97	0.97	9	215
	8.00	1.0	2.84	0.77	6	155
	8.00	2.0	2.70	0.55	4	101
	5.00	0.6	2.99	0.86	10	215
	5.00	1.0	2.82	0.66	7	155
	5.00	2.0	2.74	0.43	5	100
	3.00	0.6	3.00	0.69	13	225
	3.00	1.0	2.87	0.44	11	162
	3.00	2.0	2.77	0.24	9	106
	1.00	0.6	3.02	0.28	37	240
	1.00	1.0	2.87	0.03	215	173
	1.00	2.0	2.77	2.00×10^{-9}	1.12×10^9	111

continued on next page

Table 6.3 – continued from previous page

Source ^a	E (V/cm)	% T	$\mu_s \times 10^8$ ^b (m ² /Vsec)	$\mu_L \times 10^8$ ^b (m ² /Vsec)	γ ^b (kbp)	a ^c (nm)
Heller (1994)	10.0	1.0	1.68	0.65	6	89
	10.0	0.8	1.97	0.70	6	90
	5.00	1.0	1.76	0.39	6	142
	5.00	0.8	1.95	0.51	7	175
	3.00	1.0	7.36	0.28	5	126
	3.00	0.8	2.73	0.35	7	159
	2.00	1.0	5.33	0.23	7	131
	2.00	0.8	2.56	0.26	10	164
	1.50	1.0	7.91	0.20	8	147
	1.50	0.8	4.02	0.24	9	175
	1.00	1.0	37.85	0.14	10	132
	1.00	0.8	4.91	0.14	15	167
	0.75	1.0	13.69	0.12	12	53
	0.75	0.8	8.87	0.09	22	64
	0.50	1.0	7.68	0.06	21	138
	0.50	0.8	6.41	0.05	48	175
	0.35	1.0	34.43	0.01	136	143
	0.35	0.8	4.71	0.02	145	173
	0.25	1.0	17.18	3.03×10^{-6}	522553	155
	0.25	0.8	6.32	2.05×10^{-7}	1.08×10^7	183
	0.20	1.0	38.86	7.61×10^{-10}	2.03×10^9	143
	0.20	0.8	6.48	1.45×10^{-7}	1.51×10^7	170
	0.13	1.0	6.08	1.30×10^{-8}	1.31×10^8	57
	0.13	0.8	5.73	1.00×10^{-6}	2200100	65
Calladine (1991)	1.00	0.5	2.57	2.76×10^{-4}	26425	227
	1.00	1.0	2.35	1.02×10^{-8}	2.86×10^8	143
	1.00	1.5	2.24	1.08×10^{-14}	1.49×10^{14}	107
	1.00	2.0	2.10	2.71×10^{-10}	4.07×10^9	88
	2.50	0.5	2.85	0.56	12	193
	2.50	1.0	2.78	0.26	11	129
	2.50	1.5	2.65	2.68×10^{-4}	7492	107
	2.50	2.0	2.45	1.59×10^{-6}	774998	84

^a Sources are the same references as in table 6.1^b Best fit parameters obtained from equation 2.56 using the full range of lengths extrapolated from figures in the specified source^c Pore sizes calculated using equation 4.3 from the parameters

6.3 Summary

Equation 2.56 have been shown to have a surprising ability to fit mobilities in a broad range of separation experiments. Publications were searched to find a diverse set of experiments, including other separations in agarose gels, separations of ssDNA in polyacrylamide gels, capillary electrophoresis, contour-clamped homogenous electric field electrophoresis, and electrochromatography. Different types of experiments provide different types of results. For example capillary electrophoresis provides extremely high resolution and increases the separation efficiency of the system for electrophoretic separations (see ref. [1]) with only small amounts of material. Field strengths of about 300 V/cm are used for DNA separations. The experiments are performed in a capillary typically 25 – 100 μm in diameter. Since the field is much higher than for conventional electrophoresis we would certainly expect the details of the separation mechanism to be very different. This is reflected in the differences in the absolute values of the parameters of equation 2.56 that yield a good fit. Perhaps the most notable observation is that good fits are still possible.

Polyacrylamide gels have a much tighter network of pores than agarose gels [8]. A smaller network allows the DNA molecule to get trapped easier during electrophoresis. At lower fields the molecule is more prone to become entropically trapped. At higher fields a polyacrylamide gels will provide more obstacles when separating large molecules [8]. Although these differences exist in electrophoresis for agarose gels and polyacrylamide gels equation 2.56 still provided reasonable fits in the applicable areas. When entropic trapping occurs the fit will degrade , but still can provide for nice indication where entropic trapping will occur. Chiari et al. [128] electrophoresed dsDNA using polyacrylamide gels and agarose gels. Since agarose gels have larger pore sizes, larger molecules will be electrophoresed easier. When looking at fits to Chiari et al.'s data tables 6.2 and 6.3 show that better fits were provided for agarose gels.

It's not surprising that single-stranded DNA (ssDNA) provide very similar results to dsDNA when fitting equation 2.56 to electrophoretic data. The persistence length for ssDNA is $l_p \approx 5$ nm [58]. Chances are ssDNA will often be in a globular state smaller than the pore size in agarose gels. In that case Ogston sieving model will apply. In polyacrylamide gels, as discussed above, the pores are much smaller and the molecule will behave similarly to a dsDNA being electrophoresed in an agarose gel.

In summary, this simple exponential equation (equation 2.56) seems to apply to many different systems. This chapter showed the same behaviors exist for the parameters when applied to different experiments. The equation can also give information about the pore sizes of gel's other than an agarose gel. This chapter only provided a preliminary investigation on different systems.

CHAPTER 7

CONCLUSIONS

A simple exponential equation fit electrophoretic mobility of dsDNA in agarose gels depends on DNA length. This equation is empirical and has the form $\frac{1}{\mu(L)} = \frac{1}{\mu_L} - (\frac{1}{\mu_L} - \frac{1}{\mu_s})e^{-L/\gamma}$, where μ_L , μ_s , and γ are independent, free fitting parameters. Extensive electrophoresis experiments in agarose gels were performed on dsDNA fragments ranging in length from 100 bp to 194,000 bp to find the versatility of this equation. The electric field was varied from $E = 0.71$ V/cm to 5 V/cm on a horizontal electrophoresis apparatus and 0.62 V/cm to 6.21 V/cm on a vertical electrophoresis apparatus. For each set of experiments equation 2.56 fits extremely well ($\chi^2 \geq 0.999$) for all gel concentrations with E between 1.5 V/cm and 3.5 V/cm. Equation 2.56 was applied to separations data taken from the literature. The data from the literature contained a diverse set of experimental conditions, such as different separations media, separations of ssDNA, and different experimental conditions. The various types of separations experiments could all be fit by equation 2.56. This is not too surprising for some of the more limited data sets from the literature. On the other hand, data ranging from 10 bp to 50 kbp were fit better by equation 2.56 than by the phenomenological equation used by other authors. Figure 6.1 and table 6.2 show that at lower fields the quality of the fit decreases for all gel concentrations. Also the quality of the fit usually decreases at higher fields, but the fits are still better than for fields below 1.5 V/cm. All separations data gathered from the literature decreases in χ^2 with decreasing E and increasing %T (see table 6.2 and figure 6.1), in agreement with the experiments performed for this dissertation. In other words, equation 2.56

describes the mobilities of double stranded DNA, and potentially ssDNA, very well for moderate to high fields and moderate to low gel concentrations.

“Reptation plots” (plotting $3\mu L/\mu_o$ vs. L) helped give new perspective on understanding the trends with the fit. Rousseau, et al. [58] observed three regions with different slopes when looking at the reptation plot for ssDNA electrophoresed in polyacrylamide gels at low fields. For higher fields the middle region disappeared and only two smooth regions exist, described by Ogston sieving and reptation. At low fields and high concentration the “entropic trapping” regime occurs and accounts for the decreasing quality of the fit.

Slater [34] found an interesting relation between the parameters, relating them to the pore size (equation 4.3 discussed in section 4.2). By comparing tables 4.4 and 6.3 the pore sizes were calculated from Slater’s suggestion and were found to be in close agreement with what has been reported in the past. Table 4.4 also lists pore sizes determined by other sources using other methods. Values obtained by imaging methods, such as AFM or electron microscopy, produce values that are close to this new method. Even for polyacrylamide gels this method provides reasonable values for the pore size [58, 129–130].

In summary, this simple exponential equation provides an amazingly good representation of electrophoretic data. The versatility of equation 2.56 clearly covers an extraordinary range of DNA length, gel concentration, and electric field. Not only does this equation model dsDNA electrophoresis in agarose gels, but it also provides an excellent representation for other gels, other molecules (for example ssDNA), and other type of experiments.

In addition to exploring fit quality, this dissertation has also provided for the best settings for interpolating electrophoresis data. It was shown that the best fits were done at $T = 1.0\%$ for E between 1.5 V/cm to 3.5 V/cm. This is evident when looking at the quality of the fits. With this knowledge, an experiment can be run

with a few DNA markers to set the parameters for the fit within this region and lengths for unknown DNA fragments can be determined precisely by interpolation.

REFERENCES

- [1] J. W. Jorgenson and K. D. Lukacs. *Science*, 222:266–272, 1983.
- [2] G. A. Griess and P. Serwer. *Biophys. J.*, 74:A71–A71, 1998.
- [3] S. J. Huang and P. Serwer. *Biophys. J.*, 82:217, 2002.
- [4] S. Dorrell. *Drug Discov. Today*, 7:539–540, 2002.
- [5] K. A. Ferguson. *Metabolism*, 13:985–1002, 1964.
- [6] J. S. Fawcett and C. J. O. R. Morris. *Separation Science*, 1:9–26, 1966.
- [7] J. D. Watson and F. H. C. Crick. *Nature*, 171:737–738, 1953.
- [8] J. L. Viovy. *Rev. Mod. Phys.*, 72:813–872, 2000.
- [9] E. M. Southern. *J. Mol. Biol.*, 98:503–517, 1975.
- [10] E. M. Southern. *Anal. Biochem.*, 100:301–318, 1979.
- [11] N. C. Stellwagen. *Biochemistry*, 22:6180–6185, 1983.
- [12] P. Serwer and J. L. Allen. *Biochemistry*, 23:922–927, 1984.
- [13] M. Chakrapani D. H. Van Winkle B. C. Patterson and R. L. Rill. *Langmuir*, 18:6449–6452, 2002.
- [14] P. Serwer. *Electrophoresis*, 4:375–382, 1983.
- [15] D. W. Thompson. *On Growth and Form*. Cambridge University Press, 1942.
- [16] W. H. Nernst. *Z. Phys. Chem.*, 3:613–637, 1888.
- [17] W. H. Nernst. *Z. Phys. Chem.*, 4:129–81, 1889.
- [18] A. Baumgartner and M. Muthukumar. *J. of Chem. Phys.*, 84:440–443, 1986.
- [19] A. Baumgartner and M. Muthukumar. *J. of Chem. Phys.*, 87:3082–3088, 1987.
- [20] M. Muthukumar and A. Baumgartner. *Macromolecules*, 22:1937–1941, 1989.

- [21] M. Djabourov A. H. Clark D. W. Rowlands and S. B. Rossmurphy. *Macromolecules*, 22:180–188, 1989.
- [22] T. B. Liu Z. K. Zhou C. H. Wu V. M. Nace and B. Chu. *Macromolecules*, 30:7624–7626, 1997.
- [23] R. L. Rill Y. J. Liu D. H. Van Winkle and B. R. Locke. *J. Chromatogr. A*, 817:287–295, 1998.
- [24] A. G. Ogston. *Trans. Faraday Soc.*, 54:1754–1757, 1958.
- [25] P. G. de Gennes. *J. Chem. Phys.*, 55:572–579, 1971.
- [26] G. W. Slater and J. Noolandi. *Phys. Rev. Lett.*, 55:1579–1582, 1985.
- [27] G. W. Slater P. Mayer and G. Drouin. *High Resolution Separation and Analysis of Biological Macromolecules Pt A*, 270:272–295, 1996.
- [28] G. W. Slater J. Rousseau J. Noolandi C. Turmel and M. Lalande. *Biopolymers*, 27:509–524, 1988.
- [29] N. C. Stellwagen. *Biopolymers*, 24:2243–2255, 1985.
- [30] A. N. Semenov T. A. J. Duke and J. L. Viovy. *Phys. Rev. E*, 51:1520–1537, 1995.
- [31] C. K. Mathews and K. E. van Holde. *Biochemistry*. The Benjamin\Cummings Publishing Company, Inc., Menlo Park, CA, 1996.
- [32] D. H. Van Winkle A. Beheshti and R. L. Rill. *Electrophoresis*, 23:15–19, 2002.
- [33] R. L. Rill A. Beheshti and D. H. Van Winkle. *Electrophoresis*, 23:2710–2719, 2002.
- [34] G. W. Slater. *Electrophoresis*, 23:1410–1416, 2002.
- [35] G. T. Barkema C. Caron and J. F. Marko. *Biopolymers*, 38:665–667, 1996.
- [36] C. R. Calladine C. M. Collis H. R. Drew and M. R. Mott. *J. Mol. Biol.*, 221:981–1005, 1991.
- [37] F. F. Reuss. *Memoires de la Société Imperiales de Naturalistes de Moskou*, 2:327–336, 1809.
- [38] M. v. Smoluchowski. *Bull. Int. de l’Aced. des Sciences de Crac.*, 8:182–200, 1903.
- [39] A. Tiselius. *Trans. Faraday Soc.*, 33:524, 1937.

- [40] J. Sambrook E. F. Fritsch and T. Maniatis. *Molecular Cloning. A Laboratory Manual Vol. 2*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.
- [41] P. Dejardin O. J. Lumpkin and B. H. Zimm. *J. Polym. Sci. Pol. Sym.*, 73:67–69, 1985.
- [42] C. Heller and J. L. Viovy. *Biopolymers*, 35:485–492, 1995.
- [43] G. W. Slater C. Turmel M. Lalande and J. Noolandi. *Biopolymers*, 28:1793–1799, 1989.
- [44] J. L. Viovy T. Duke and F. Caron. *Contemp. Phys.*, 33:25–40, 1992.
- [45] A. Y. Grosberg and A. R. Khokhlov. *Statistical Physics of Macromolecules*. American Institute of Physics, New York City, NY, 1994.
- [46] S. Arnott A. Fulmer W. E. Scott I. C. M. Dea R. R. Moorhouse and D.A. Rees. *J. Mol. Biol.*, 90:269–284, 1974.
- [47] G. A. Griess K. B. Guiseley and P. Serwer. *Biophys. J.*, 65:138–148, 1993.
- [48] M. Maaloum N. Pernodet and B. Tinland. *Electrophoresis*, 19:1606–1610, 1998.
- [49] P. Y. Key and D. B. Sellen. *J. Poly. Sci.*, 20:659–679, 1982.
- [50] C. Rochas A. M. Hecht and E. Geissler. *Macromol. Symp.*, 138:157–163, 1999.
- [51] T. K. Attwood and D. B. Sellen. *Biopolymers*, 29:1325–1328, 1990.
- [52] S. A. Foord and E. D. T. Atkins. *Biopolymers*, 28:1345–1365, 1989.
- [53] J. Noolandi J. Rousseau G. W. Slater C. Turmel and M. Lalande. *Phys. Rev. Lett.*, 58:2428–2431, 1987.
- [54] J. C. Giddings E. Kucera C. P. Russel and M. N. Myers. *J. Phys. Chem.*, 13:4397–4408, 1968.
- [55] D. Rodbard and A. Chrambach. *Proc. Natl. Acad. Sci. USA*, 65:970–977, 1970.
- [56] D. Rodbard and A. Chrambach. *Anal. Biochem.*, 40:95–134, 1971.
- [57] N. C. Stellwagen. *Electrophoresis*, 19:1542–1547, 1998.
- [58] J. Rousseau G. Drouin and G. W. Slater. *Phys. Rev. Lett.*, 79:1945–1948, 1997.
- [59] J. Rousseau G. Drouin and G. W. Slater. *Electrophoresis*, 21:1464–1470, 2000.

- [60] S. W. P. Turner M. Cabodi and H. G. Craighead. *Phys. Lett. Rev.*, 88:128103–1–128103–4, 2002.
- [61] O. J. Lumpkin P. Dejardin and B. H. Zimm. *Biopolymers*, 24:1573–1593, 1985.
- [62] G. W. Slater and J. Noolandi. *Europhys. Lett.*, 1:347–353, 1986.
- [63] G. W. Slater and J. Noolandi. *Biopolymers*, 28:1781–1791, 1989.
- [64] C. Heller T. Duke and J. L. Viovy. *Biopolymers*, 34:249–259, 1994.
- [65] T. A. J. Duke A. N. Semenov and J. L. Viovy. *Phys. Rev. Lett.*, 69:3260–3263, 1992.
- [66] T. Duke and J. L. Viovy. *Phys. Rev. E*, 49:2408–2416, 1994.
- [67] G. W. Slater J. Rousseau and J. Noolandi. *Biopolymers*, 26:863–872, 1987.
- [68] G. T. Barkema J. F. Marko and B. Widom. *Phys. Rev. E*, 49:5303–5309, 1994.
- [69] M. E. J. Newman and G. T. Barkema. *Phys. Rev. E*, 56:3468–3473, 1997.
- [70] M. Widom and I. Al-Lehyani. *Physica A*, 244:510–521, 1997.
- [71] M. J. Chacron and G. W. Slater. *Phys. Rev. E*, 56:3446–3450, 1997.
- [72] G. T. Barkema and M. E. J. Newman. *Physica A*, 244:25–39, 1997.
- [73] B. Kozulić. *Appl. Theor. Electrophor.*, 4:117–123, 1994.
- [74] B. Kozulić. *Appl. Theor. Electrophor.*, 4:125–136, 1994.
- [75] B. Kozulić. *Appl. Theor. Electrophor.*, 4:137–148, 1994.
- [76] B. Kozulić. *Appl. Theor. Electrophor.*, 4:149–159, 1994.
- [77] B. Kozulić. *Anal. Biochem.*, 231:1–12, 1995.
- [78] O. J. Lumpkin and B. H. Zimm. *Biopolymers*, 21:2315–2316, 1982.
- [79] A. Chrambach and D. Rodbard. *Science*, 172:440–451, 1971.
- [80] A. Guttman T. Lengyel M. Szoke and M. Sasvari-Szekely. *J. Chromatogr. A*, 871:289–298, 2000.
- [81] M. Szoke M. Sasvari-Szekely and A. Guttman. *J. Chromatogr. A*, 830:465–471, 1999.

- [82] S. F. Zakharov H. T. Chang and A. Chrambach. *Electrophoresis*, 17:84–90, 1996.
- [83] F. Oosawa. *Polyelectrolytes*. Marcel Dekker, New York, 1971.
- [84] W. B. Russel D. A. Saville and W. R. Shopwalker. *Colloidal Dispersions*. Cambridge University Press, Cambridge, England, 1989.
- [85] G. Gouy. *J. Phys. Radium*, 9:457–468, 1910.
- [86] D. L. Chapman. *Phil. Mag.*, 25:475–481, 1913.
- [87] H. Benoit and P. Doity. *J. Phys. Chem.*, 57:658–963, 1953.
- [88] M. Muthukumar and A. Baumgartner. *Macromolecules*, 22:1941–1946, 1989.
- [89] W. Kuhn. *Kolloid Z.*, 68:2–15, 1934.
- [90] E. Guth and H. Mark. *Monatsh. Chem.*, 65:93, 1934.
- [91] S. B. Smith J. Liphardt C. Bustamante and D. Smith. *Current Opinion in Structural Biology*, 10:279–285, 2000.
- [92] M. T. Zamora P. G. Rouxhet B. C. van der Aa R. M. Michel M. Asther and Y. F. Dufrêne. *Langmuir*, 17:3116–3119, 2001.
- [93] G. Beaucage S. Rane S. Sukumaran M. M. Satkowski L. A. Schechtman and Y. Doi. *Macromolecules*, 30:4158–4162, 1997.
- [94] C. J. O. R. Morris. *Protides of the Biological Fluids, 14th Colloquium*, pages 543–551. Elsevier Publishing Co, 1967.
- [95] E. Arvanitidou D. Hoagland and D. Smisek. *Biopolymers*, 31:435–447, 1991.
- [96] G. W. Slater and H. L. Gou. *Electrophoresis*, 17:977–988, 1996.
- [97] G. W. Slater and H. L. Gou. *Electrophoresis*, 17:1407–1415, 1996.
- [98] G. W. Slater and J. R. Reunient. *J. Chromatogr. A*, 772:39–48, 1997.
- [99] K. Wuthrich M. Billeter and W. Braun. *J. Mol. Biol.*, 180:715–740, 1984.
- [100] M. Doi and S. F. Edwards. *The Theory of Polymer Dynamics*. Oxford University Press, New York City, NY, 1986.
- [101] L. S. Lerman and H. L. Frisch. *Biopolymers*, 21:995–997, 1982.
- [102] N. C. Stellwagen. *Biochemistry*, 22:6186–6193, 1983.

- [103] S. P. Edmondson and D. M. Gray. *Biopolymers*, 23:2725–2742, 1984.
- [104] T. Duke J. L. Viovy and A. N. Semenov. *Biopolymers*, 34:239–247, 1994.
- [105] A. Pluen B. Tinland J. Sturm and G. Weill. *Electrophoresis*, 19:1548–1559, 1998.
- [106] G. W. Slater and J. Noolandi. *Electrophoresis*, 9:643–646, 1988.
- [107] A. Baumgartner and M. Muthukumar. *Adv. Chem. Phys.*, 94:625–708, 1996.
- [108] J. K. Elder and E. M. Southern. *Anal. Biochem.*, 128:227–231, 1983.
- [109] M. Takahashi T. Ogino and K. Baba. *Biochem. Biophys. Acta*, 174:183–187, 1969.
- [110] T. Maniatis A. Jeffrey H. van deSande. *Biochemistry*, 14:3787–3794, 1975.
- [111] H. Hervet and C. P. Bean. *Biopolymers*, 26:727–742, 1987.
- [112] P. Serwer and S. J. Hayes. *Anal. Biochem.*, 158:72–78, 1986.
- [113] N. C. Stellwagen S. Magnúsdóttir C. Gelfi and P. G. Righetti. *Biopolymers*, 58:390–397, 2001.
- [114] N. C. Stellwagen and D. L. Holmes. *Electrophoresis*, 11:649–652, 1990.
- [115] N. C. Stellwagen. *Electrophoresis*, 13:601–603, 1992.
- [116] P. Serwer J. L. Allen and S. J. Hayes. *Electrophoresis*, 4:273–276, 1983.
- [117] D. L. Holmes and N. C. Stellwagen. *Electrophoresis*, 11:5–15, 1990.
- [118] G. A. Griess E. T. Moreno R. A. Easom and P. Serwer. *Biopolymers*, 28:1475–1484, 1989.
- [119] P. G. Righetti B. C. W. Brost and R. S. Snyder. *J. Biochem. Bioph. Meth.*, 4:347–363, 1981.
- [120] B. Tinland N. Pernodet and A. Pluen. *Biopolymers*, 46:201–214, 1998.
- [121] T. C. Laurent and J. Killander. *J. Chromatogr.*, 14:317–330, 1964.
- [122] G. W. Slater and J. Noolandi. *Biopolymers*, 25:431–454, 1986.
- [123] O. Lumpkin S. D. Levene and B. H. Zimm. *Phys. Rev. A*, 39:6557–6566, 1989.
- [124] N. C. Stellwagen C. Gelfi and P. G. Righetti. *Biopolymers*, 42:687–703, 1997.

- [125] B. H. Zimm and S. D. Levene. *Q. Rev. Biophys.*, 25:171–204, 1992.
- [126] J. L. Viovy. *Electrophoresis*, 14:1088–1088, 1993.
- [127] Y. G. Park. *Biochem. Engineering J.*, 7:213–221, 2001.
- [128] M. Chiari L. Dalesio R. Consonni and P. G. Righetti. *Electrophoresis*, 16:1337–1344, 1995.
- [129] S. N. Brahmasandra D. T. Burke C. H. Mastrangelo and M. A. Burns. *Electrophoresis*, 22:1046–1062, 2001.
- [130] N. Stellwagen C. Gelfi and P. G. Righetti. *Electrophoresis*, 23:167–175, 2002.
- [131] R. L. Rill B. R. Locke Y. J. Liu J. Dharia and D. VanWinkle. *Electrophoresis*, 17:1304–1312, 1996.
- [132] D. Vollrath and R. W. Davis. *Nucleic Acids Res.*, 15:7865–7876, 1987.
- [133] D. Vollrath. *Methods in Molecular Biology Vol. 12:Pulsed-Field Gel Electrophoresis*, volume 12, pages 19–30. The Humana Press Inc., Totowa, NJ, 1992.
- [134] R. Virtanen. *Acta. Polytech. Scand.*, 169:1–67, 1974.
- [135] J. W. Jorgenson and K. D. Lukacs. *J. Chromatogr.*, 218:209–216, 1981.
- [136] J. W. Jorgenson and K. D. Lukacs. *Clin. Chem.*, 27:1551–1553, 1981.
- [137] J. W. Jorgenson and K. D. Lukacs. *Anal. Chem.*, 53:1298–1302, 1981.
- [138] R. L. Stalkier. *Anal. Chem.*, 68:569R–586R, 1996.
- [139] A. S. Cohen and B. L. Karger. *J. Chromatogr.*, 397:409–417, 1987.
- [140] H. J. Naghash and O. Okay. *J. Appl. Pol. Sci.*, 60:971–979, 1996.
- [141] A. Deriu F. Cavatorta D. Asnaghi A. Bossi and P. G. Righetti. *Physica B*, 234-236:271–272, 1997.

BIOGRAPHICAL SKETCH

Afshin Beheshti

Now the part of the dissertation which everyone enjoys and will pay attention to the most. Afshin Beheshti was born on January 25th, 1975 in Tehran, Iran. At the age of 4 in 1979 his family moved from Tehran, Iran to Minneapolis, Minnesota. There, he grew up in a suburb called Apple Valley where he started to gain interest in the sciences. During High School he participated in the math teams and science teams which allowed his interest in science to blossom. To alleviate the side of his brain focused on science he took piano and saxophone lessons since the 4th grade, which he still avidly plays.

In 1993 he started his undergraduate education at the University of Minnesota (Twin Cities Campus). During his undergraduate career he researched in various High Energy Physics laboratories. This gave him the opportunity to work at Brookhaven National Laboratory for a summer. Towards the end of his undergraduate studies his interest started to swing towards biophysics. It took him 3 years and two quarters to obtain a bachelors of science in physics. For six months after his completion he worked at 3M. Although he enjoyed the job and worked hard, he learned that the corporate life isn't for him yet and he still needs to learn more. Lucky for him, he was already accepted to Florida State University for graduate school in physics.

In the fall of 1997 he started his graduate work in physics. He spent five years and some change discovering, exploring, and reinventing the world of biophysics. To keep his sanity throughout graduate school he played music and learned many new

instruments. He was able to record and release several albums of original music under the moniker The Afrosheens which has done well in the college music scene. Now the end of this chapter of his life is complete and it is time to move on to the next exciting challenge.