To the Graduate Council:

I am submitting herewith a dissertation written by Michael Allen Stebbins entitled "Development of Strategies for Biochemical Analysis Using Capillary Electrophoresis." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Chemistry.

Michael J. Sepaniak, Major Professor

We have read this dissertation
and recommend its acceptance:

Accepted for the Council:

Associate Vice Chancellor and
Dean of The Graduate School

DEVELOPMENT OF STRATEGIES FOR BIOCHEMICAL ANALYSIS USING CAPILLARY ELECTROPHORESIS

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Michael Allen Stebbins
December 1997

DEDICATION

This dissertation is dedicated to my parents

whose love and support made this endeavor possible.

ACKNOWLEDGMENTS

There are a number of people to whom I would like to express my sincere thanks for helping me see this accomplishment through. To my family, I owe all of this to your love and support. Thanks Mom, Dad, Tammy and Kelly for always being there when I needed you. I owe special thanks to Camille Kassis whose true love and support was always there, even when it was not reciprocated. Thanks for not giving up on me Camille, you played a bigger role in this than you can possibly imagine. I can't imagine having made it without you, you mean the world to me!

I would like to thank all of the Sepaniak group members that I have worked with over the years. To the current group members (Tim, Will, Kylen, David, Jason, Jeremy, April, Vica, Susan and Shannon) thanks for the in-depth and fruitful discussions when I needed them and for the marginal ones when I needed them. April, I wish you all the best and I am sure you will do well. Good luck! To the past group members (Tracy, Brian, Craig, Christy and Beth) thanks for all the advice, I think it may have paid off. Tracy and Brian, I appreciate all the help and knowledge you have provided over the years. Your friendships are invaluable.

I would like to thank the members of Dr. Hurlburt's and Dr. Peterson's research groups for the help offered in my time spent with them. A special thanks is extended to Angelia Gibson and Christine Schar for their generous support and friendship.

I would also like to thank my committee members (Dr.'s Wehry, Alexandratos and Peterson) for their timely discussions concerning research. I owe a special debt of

gratitude to my collaborators Barry Hurlburt and Cynthia Peterson. Without your help I would not have been so successful. I feel privileged to have worked with each of you. I hope that the on-going collaborations will continue.

Finally I would like to thank my advisor, Mike Sepaniak. I appreciate you giving me the freedom to grow and mature as a scientist and the opportunity to do research with you. I would also like to thank you for your timely assistance. It has been nothing short of an incredible experience.

ABSTRACT

Capillary electrophoresis (CE) is a powerful analytical tool that has been applied to the separation of a variety of biomolecules including DNA. Recently, CE has found its way into the life sciences laboratory as a means to investigate analyses performed by the more traditional biochemical methods. The unique characteristics of CE (speed, efficiency (resolving power) and small sample requirements) are exploited for the investigation of DNA and protein molecules. Increasing pressure to extract specific information from DNA analysis in a timely fashion creates a burden on the development of effective experimental approaches. In this light, the development of computational methods for the optimization of DNA fragment separations based on content of sample and empirically derived relationships between fragment size, soluble polymer concentration and column length was undertaken. These models allowed for the optimal separation performance of all fragments in the sample. The development of a capillary-based assay for the investigation of protein-DNA interactions was also explored. Resolution of DNA from the protein-DNA complex is dependent upon buffer composition and pH. The selectively of these bio-specific interactions in demonstrated in the presence of interfering proteins. The diagnostic use of CE can further be realized in the temporal analysis of DNA restriction enzyme digests. Demonstrated is the ability to follow the progression of a digest (i.e., the creation and diminishing of fragment peaks) and to determine partial digest conditions for the generation of intermediate fragments of interest.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE

PAGE

NOMENCLATURE

a	excess resolution factor
a'	spherical ion radius
A	absorbance
b	overlap degradation factor
b'	pathlength
c	time factor
C	concentration
D	diffusion coefficient
ε	dielectric constant
ε'	molar absorptivity
E	electric field strength
F	force
F'	frictional force
H	plate height
k	instrumental fluorescence collection efficiency
K	retardation coefficient
K_D	dissociation constant
l	effective length (to detector)
L	column length

μ_e	intrinsic electrophoretic mobility
μ_{EOF}	electroosmotic flow mobility
μ_{obs}	observed mobility
$\Delta\mu$	difference in mobilities
$\overline{\mu}$	average mobility
η	viscosity
N	theoretical plate count
N'	number of base pairs
Φ	polymer concentration
$\Phi*$	entanglement threshold
Φ_f	fluorescence quantum efficiency
P	fluorescence intensity
P_o	excitation source intensity
q	charge
r	thickness of polymer strand
R_g	radius of gyration of spherical DNA molecule
R_p	radius of gyration of polymer blob
R_s	resolution
R_{so}	optimal resolution
σ	band variance
t	migration time

T	total time
T_o	optimal total time
$T_{R(n)}$	retention time of fragment n
$T_{R(n-1)}$	retention time of fragment n-1
v	velocity
v_{EOF}	velocity of electroosmotic flow
V	applied voltage
ξ	zeta potential
ξ'	average polymeric pore size
$w_{1/2}$	width at half height
W_b	base width

Abbreviations

AC	affinity chromatography
APS	ammonium persulfate
bp	base pairs
bHLH	basic-helix-loop-helix
CE	capillary electrophoresis
CEMSA	capillary electrophoretic mobility shift assay
CRF	chromatographic response function
EB	ethidium bromide

EDTA	ethylenediaminetetraacetic acid
EMSA	electrophoretic mobility shift assay
EOF	electroosmotic flow
FWHM	full width half maximum
γ-MTMS	γ-methacryloxypropyltrimethoxysilane
He-Cd	helium-cadmium laser
He-Ne	helium-neon laser
HPLC	high performance liquid chromatography
i.d.	inner diameter
IEC	ion exchange chromatography
MAT	matrix program
MALDI	matrix assisted laser desorption ionization
MC	methyl cellulose
MS	mass spectrometry
o.d.	outer diameter
PAI-1	plasminogen activator inhibitor-type 1
PCR	polymerase chain reaction
PMT	photomultiplier tube
RP-HPLC	reverse phase high performance liquid chromatography
SEC	size exclusion chromatography
SIM	simplex program
S/N	signal to noise

SSCE	size-selective capillary electrophoresis
TBE	tris-boric acid-EDTA buffer
TEMED	N,N,N', N'-tetramethylethylenediamine
trpR	tryptophan repressor protein
UV	ultraviolet
UV-VIS	ultraviolet-visible

CHAPTER 1

Introduction

BIOANALYTICAL TECHNOLOGY

The past two decades have witnessed enormous strides in bioanalytical technology. The results of these advances are recognizable in increasingly novel and efficient means for the production of therapeutic agents and biological substances. As is often the case with such notable advances, a codependency of existing technologies is typically responsible. The rapid growth and significant advances seen within the biotechnology industry would not have been realized without timely advances in separation science. Conversely, the unique demands from the biotechnology industry for analytical technology and high-resolution purification have spurred numerous innovations in the field of separations.

Virtually all methods of separation science have been applied to solving problems in biotechnology such as purity assessment, product characterization, quantitation and others. The most prominent and powerful separation methodologies are the chromatographic and electrophoretic techniques. Chromatographic methods useful in biomolecule separations include size exclusion chromatography (SEC), ion exchange chromatography (IEC), high performance liquid chromatography (HPLC) and affinity chromatography (AC). Gel-based electrophoretic methods for the analysis of proteins and DNA have also proven indispensable for biochemical research. However, a powerful

recent advancement is beginning to show significant prominence in this increasingly important area of science: Capillary Electrophoresis.

PRINCIPLES OF CAPILLARY ELECTROPHORESIS

Capillary electrophoresis (CE) is an analytical technique that is not only ideally suited for biochemical analysis, but shares with it that same sense of codependency that helped the biotechnology industry literally explode. Undoubtedly CE is a compilation of the analytical separation techniques out of which it was born: gas chromatography, liquid chromatography and conventional electrophoresis. CE owes, in large part, its columns to gas chromatography, its detection means to liquid chromatography and its separation mechanism to conventional electrophoresis. The field of capillary electrophoretic separations is rapidly approaching maturity as it begins to settle into its own niche(s) among the more traditional separation techniques. Nonetheless an overview of the principles governing CE separations is warranted.

An Overview

CE incorporates the powerful separation mechanisms of conventional electrophoresis with the instrumentation and automation capabilities of chromatography. Electrophoresis is the separation of charged molecules based on their differential migration through a fluid under the influence of an applied electric field. Capillary electrophoretic separations are typically performed in narrow-bore (25 to 75 μm i.d.) fused silica capillary columns. High voltages (10 to >30 kV) are applied across the

column, effectively producing electric fields on the order of 100 to 1000 volts/cm. The high resistivity of the column (a result of its small cross sectional area) limits the generation of high currents and thus internal heating (which can be efficiently dissipated due to the column's large surface to volume ratio), effectively lowering the risk of zone or band broadening. These attributes combined with the "plug-like" flow profile unique to capillary electrophoretic separations result in obtained efficiencies that often exceed one hundred thousand theoretical plates/meter and in special circumstances reach one million theoretical plates/meter. Detection is typically performed on-column and is often interfaced with computer aided data acquisition. Numerous attributes of CE include minimal sample volume requirements, fast analysis times and ease of automation. In addition, simple methods development, availability of a variety of modes (selectivity and wide applicability), aqueous media operation and low waste generation have contributed to make CE a standard technique in the field of separation science.

Brief Historical Perspective

Potentially the most important aspect in the development of CE as an analytical method was the transfer of electrophoretic techniques to a column format, allowing the application of higher separation fields without the deleterious consequences of heating effects and diffusion. While the traditional gel-based electrophoretic techniques require the presence of supporting media (agarose, acrylamide, etc.) to circumvent these problems, CE does not, and thus allows for separation in free solution. Free solution electrophoresis is based on the differential migration of analyte ions as they migrate from

one end of the separation column to the other. One of the earliest demonstrations of the use of 'pseudo' narrow-bore tubes for free solution electrophoretic separations was provided by Hjerten in 1967 [1]. Recognizing the heat displacement advantages inherent to capillaries, Hjerten was unfortunately limited by detection strategies useful for columns of such diminutive size and thus used 3 mm diameter tubes for separations. To counteract the thermal effects of these large tubes, they were slowly rotated about the longitudinal axis.

Virtanen compensated for the limited availability of useful detectors by describing a potentiometric detection method for free zone electrophoresis in 200 - 500 μm diameter Pyrex tubes [2]. Virtanen also described the importance of electroosmotic flow (a phenomenon unique to CE and described in the following pages) and its significant influence on the electrophoretic mobility of analyte ions.

Still limited by optical detection capabilities and thermal convection, Mikkers, Everaerts and Verheggen utilized 200 μm diameter Teflon tubes for free solution separation [3]. A variety of organic and inorganic anions were resolved in under 10 minutes using both ultraviolet (UV) and conductometric detection schemes. Due to large injection volumes necessary to counteract poor detectability, high efficiencies were not realized.

While the above accounts provided invaluable information towards the development of CE, it was the seminal works of Jorgenson and Lukacs [4,5] that led to the outbreak of interest in CE. These works demonstrated highly efficient (4×10^5 theoretical plates) electrophoretic separations of peptides, amines and amino acids in 75

μm diameter fused silica capillaries. Also demonstrated were simple experimental detection methodologies for use with these narrow-bore columns, including fluorescence detection. Jorgenson and Lukacs also expanded upon the theory of CE by discussing the importance of electroosmotic flow and molecular diffusion. These innovations are considered to be the beginning of modern day capillary electrophoresis.

Theory

Electrophoretic Migration

Electrophoretic separations are based on the differential migration or velocities of ionic components in the presence of an electric field. Upon application of an electric field (E, volts/cm) ionic analytes experience a force (F) that is proportional to the charge on that ion (q) and E. These equations are given by

$$E = \frac{V}{L} \quad (1.1)$$

and

$$F = qE \quad (1.2)$$

where V is applied voltage and L is column length. This force results in the migration of the ion towards the oppositely charged electrode. Counteracting frictional force (F') realized as the ions increase in velocity is given by

$$F' = 6\pi\eta a' v \quad (1.3)$$

where η is solution viscosity, a' is the radius of the ion (assumed spherical for this simplified case) and v is the velocity of the ion. Upon steady state conditions (reached almost immediately), the ion experiences a constant velocity given by

$$v = \mu_e E \quad (1.4)$$

where μ_e , the electrophoretic mobility, is an intrinsic property of a given ion. Rearrangement of Equation 1.4 to solve for μ_e yields a more recognizable expression for electrophoretic mobility.

Electroosmotic Flow

A particularly important phenomenon of free solution capillary electrophoretic separations is the presence of electroosmotic flow (EOF). EOF is the bulk movement of liquid through the capillary column arising from the interaction of solvated cations in the bulk of the solution with the negatively charged fused silica surface. Upon application of an electric field across the column, these solvated cations at the capillary-solution interface migrate towards the cathode. In doing so they drag along with them solution resulting in bulk flow of liquid towards the cathode. A depiction of EOF is given in Figure 1.1. The velocity of EOF (v_{EOF}) is given by

$$v_{EOF} = \frac{E\xi\varepsilon}{\eta} \quad (1.5)$$

where ξ is the zeta potential at the capillary wall, ε is the dielectric constant of the solution and η is the solution viscosity. Rearrangement of Equation 1.5 to solve for mobility of EOF (μ_{EOF}) yields

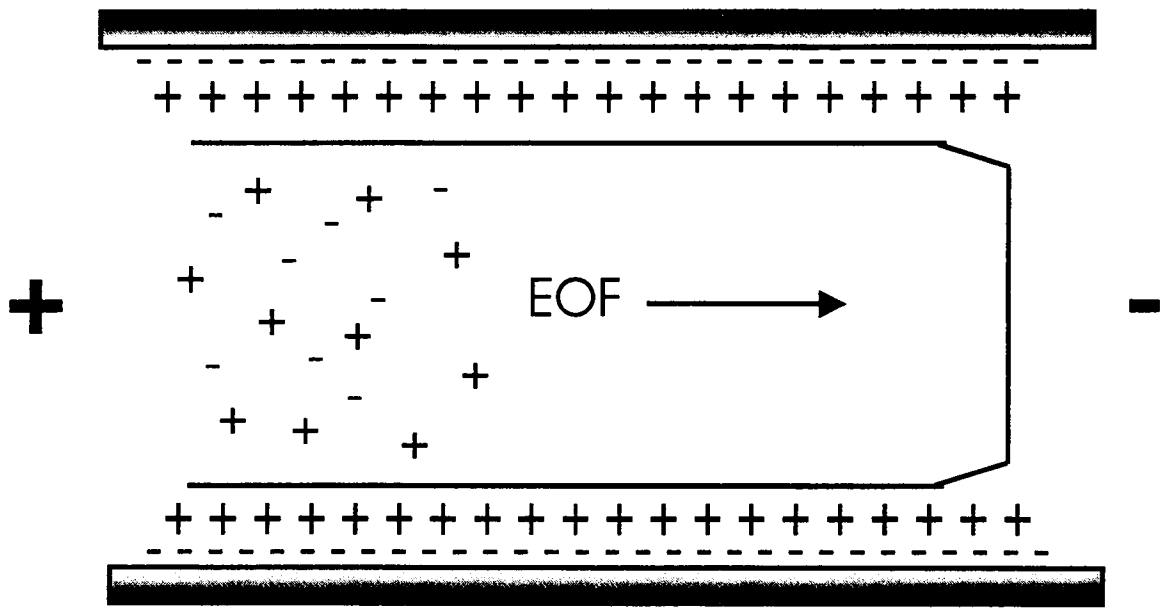


Figure 1.1. Bulk flow of liquid towards cathode resulting from electroosmotic flow (EOF).

$$\mu_{EOF} = \frac{v_{EOF}}{E} = \frac{\xi\varepsilon}{\eta} \quad (1.6)$$

μ_{EOF} is dependent upon ξ, which is determined by the surface charge on the capillary wall, for solutions of constant composition. A close inspection of Equations 1.5 and 1.6 show the independence of mobility on applied electric field as well as its lack of radial dependence. This latter unique feature provides the flat flow profile that is partially responsible for the high efficiencies attainable in CE. Because EOF is the driving force of flow and is generated at the walls of the column, a uniform distribution of flow across the column is realized, as seen in Figure 1.2 (A). In contrast, the laminar flow profiles seen in HPLC (Figure 1.2 (B)), where a pump generates the flow, create unequal flow rates for analytes at different positions within the column. This laminar flow represents a significant source of band dispersion in HPLC (not present in CE) that is generally referred to as resistance to mass transfer in the mobile phase.

Mobility and Migration

EOF serves the additional purpose of often allowing the detection of all analytes regardless of charge in a single run. This is achieved because μ_{EOF} is generally greater in magnitude than the electrophoretic mobilities of the analyte ions. The migration velocity or mobility that an individual analyte ion experiences in the column is a combination of its intrinsic mobility, μ_e, and electroosmotic mobility, μ_{EOF}. This observed mobility, μ_{obs}, is given by

$$\mu_{obs} = \mu_e + \mu_{EOF} \quad (1.7)$$

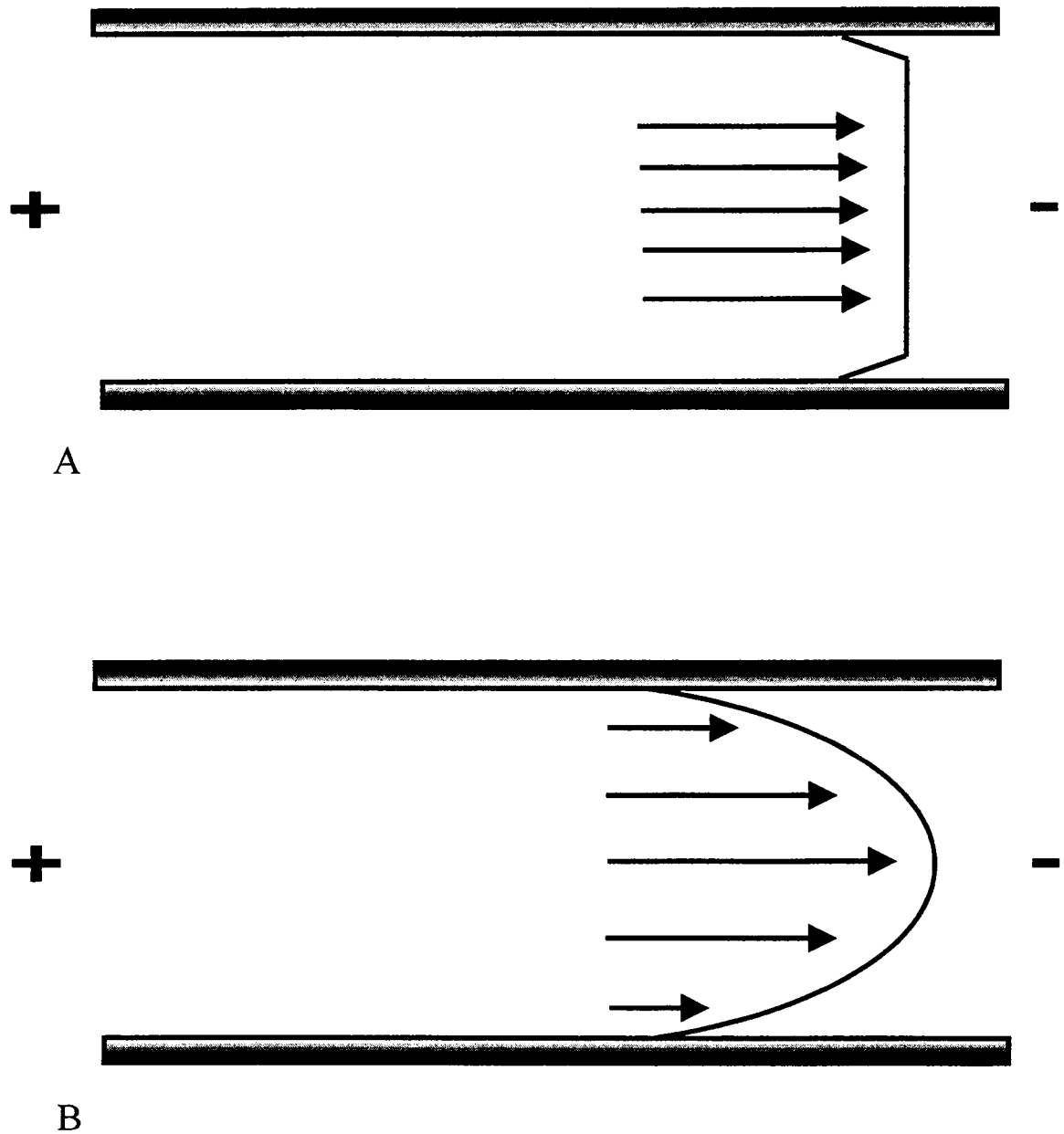


Figure 1.2. Pictorial representation of contrast between the flat flow profile of CE (A) and the laminar flow profile of HPLC (B).

Cationic analytes experience an observed mobility that is greater than that of EOF. This is due to the fact that both the cations and EOF have electrophoretic attractions toward the cathode, as seen in Figure 1.3. Neutral analytes do not possess an electrophoretic component and are thus carried by EOF. Anionic analytes experience an observed mobility that is less than that of EOF. This is due to the fact that the anions have an electrophoretic attraction to the anode ($-\mu_e$) while EOF is cathodically driven. Nevertheless, since the flow is from anode to cathode, all anions will be flushed toward the cathode by the stronger EOF allowing single run detection of all components.

The time that is required for an analyte to reach the point of detection is the migration time and is simply the quotient of migration distance and velocity. Determination of observed mobility from an electropherogram is given by

$$\mu_{obs} = \frac{l}{tE} = \frac{lL}{tV} \quad (1.8)$$

where l is the length to the detector (effective length), L is total column length, and t is migration time.

Dispersion and Efficiency

Since electrophoretic separations are based on differences in analyte mobility, the width of each analyte zone plays an important role in achieving resolution. Resolution, R_s, can be expressed as

$$R_s = \frac{1}{4}\left[\frac{\Delta\mu}{\bar{\mu}}\right]N^{1/2} \quad (1.9)$$

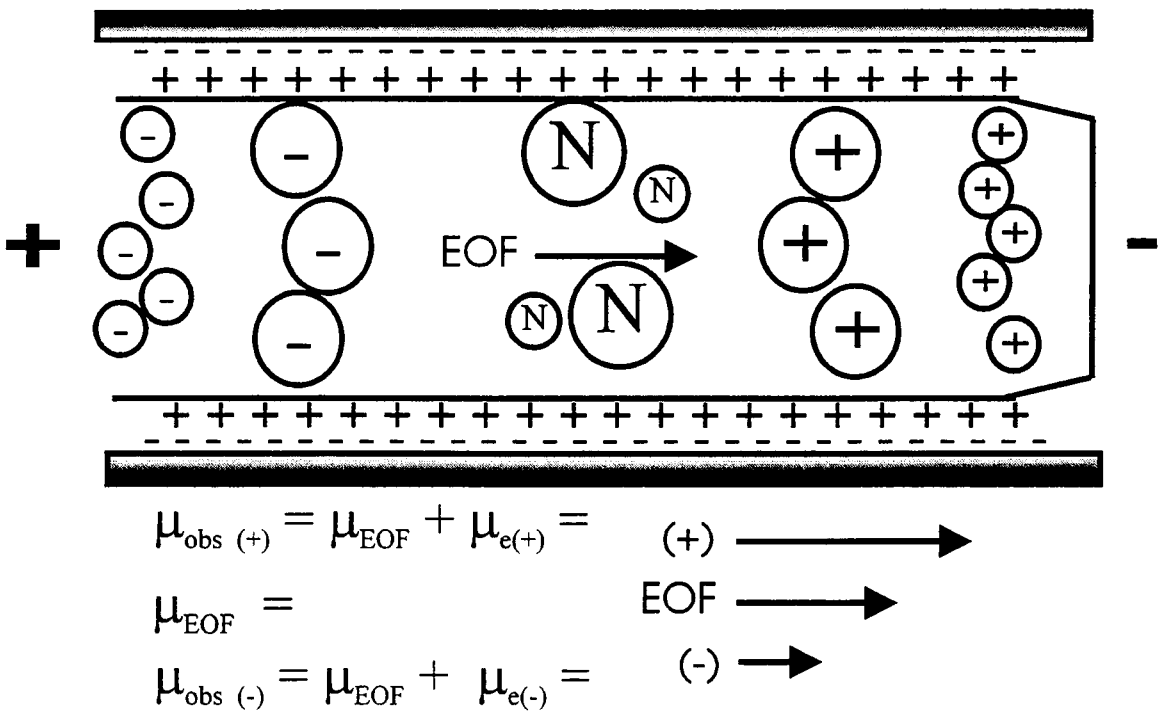


Figure 1.3. Observed mobilities for analytes in CE.

where $\Delta\mu$ is the difference in mobilities of the two species, $\overline{\mu}$ is the average mobility and N is the number of theoretical plates. There are several factors that can influence dispersion in CE and hence N. The total dispersion is given by the sum of individual dispersions that result from such factors as injection, detection, adsorption, electrodispersion, thermal effects and diffusion. Dispersion resulting from injection and detection can be partially overcome by small injection plug lengths. Dispersion resulting from adsorption to the column wall is determined by a solute's adsorption/desorption kinetics and can be approximated by the random-walk theory. Overcoming dispersion due to adsorption can be achieved by proper choice of buffer content and pH. Electrodispersion can result from changes in an analyte's μ_e or in μ_{EOF}, thus affecting μ_{obs}. Proper selection of buffer can minimize electrodispersion. The effects of thermal dispersion are negligible for narrow bore columns due to efficient heat dissipation as discussed earlier. Under ideal conditions only longitudinal diffusion is present. Therefore efficiency can be related to the molecular diffusion term in chromatography and is given by

$$\sigma^2 = 2Dt = \frac{2DlL}{\mu_e V} \quad (1.10)$$

where σ is variance and D is the diffusion coefficient of the analyte. Since variance is a measure of band diffusion and is inversely proportional to plate count, the fundamental limiting expression for N in CE is given by

$$N = \frac{\mu_e Vl}{2DL} = \frac{\mu_e El}{2D} \quad (1.11)$$

N can also be determined experimentally from an electropherogram using the following equation

$$N = 5.54\left(\frac{t}{w_{1/2}}\right)^2 \quad (1.12)$$

where $w_{1/2}$ is the width of a peak, in time units, at half height.

Instrumentation and Detection

Typical CE instrumentation is depicted in Figure 1.4. Fused silica capillaries ranging in size from 25 - 75 μm inner diameter (i.d.) and 150 - 365 μm outer diameter (o.d.) are most prevalent in CE systems. The capillary column is suspended between two buffer reservoirs whose content is identical to that which is in the column. Analyte loading is achieved by replacing the inlet of the column (anode end) with a reservoir containing the analyte(s) of interest and applying either a external pressure or a small electrical field. The former loading method, termed hydrodynamic injection, can be achieved by raising the capillary inlet, relative to the outlet, while immersed in the sample reservoir. In hydrodynamic injections, it is necessary that solution in the capillary be free flowing. The latter sample loading method, termed electrokinetic injection, is achieved by applying a small electrical field across the column resulting in the introduction of charged analytes onto the column in the absence of carrier matrix. While discrimination is realized for solutes exhibiting low mobilities, solutes having identical mobilities in free solution (i.e., DNA) show no bias. This is advantageous because electrokinetic injection

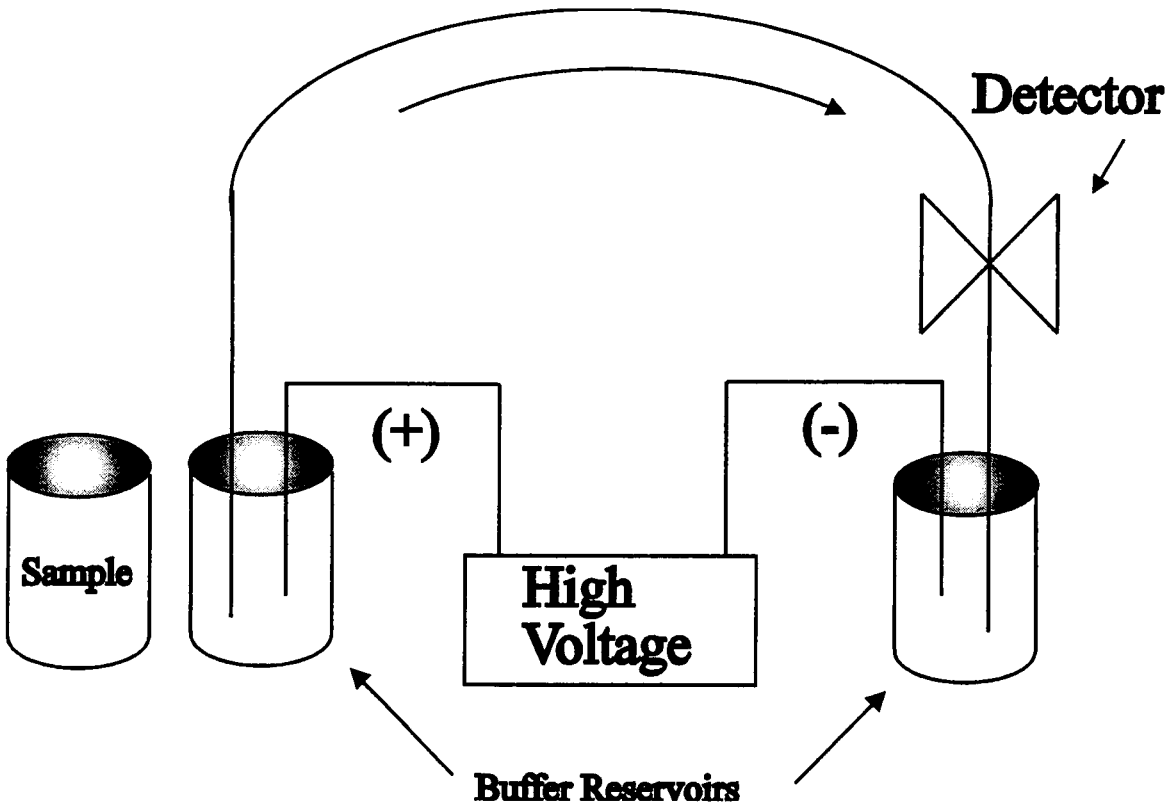


Figure 1.4. Typical CE instrumentation.

is required for the highly viscous running buffers (gels and soluble polymers) used in DNA analysis.

Following sample loading, the sample reservoir is replaced with the inlet buffer reservoir. High voltage (of the approximate range 10 - 30 kV) is then applied across the column effectively producing electric fields that can range from 100 to greater than 1000 V/cm (depending upon column length). Analyte ions migrate at different velocities from inlet (anode) to outlet (cathode) and in doing so separate into discrete bands based on their differential migration through the column. A depiction of the progression of a separation is seen in Figure 1.5. In (A) the analytes have just been injected onto the column and still exist in a single band. As the separation progresses, the analytes begin to migrate (B) into separate bands. As the analytes reach the detector at the end of the column they have separated into discrete bands that can be detected individually, (C).

As noted by the early works on column electrophoresis, detection is a significant challenge arising from the small optical path length inherent to these diminutive capillaries. The detection methods available for CE are nearly as varied as the classes of analytes that are studied: ultraviolet-visible (UV-VIS) absorbance, fluorescence, amperometry, conductivity, mass spectrometry, radioactivity, circular dichroism, Raman and others. On-column optical detection methodologies employing UV-VIS absorbance and laser induced fluorescence are the most prevalent. The absorbance of a solute, A, is dependent upon path length, b', concentration, C, and molar absorptivity, ε', and is given by

$$A = b'C\varepsilon' \quad (1.13)$$

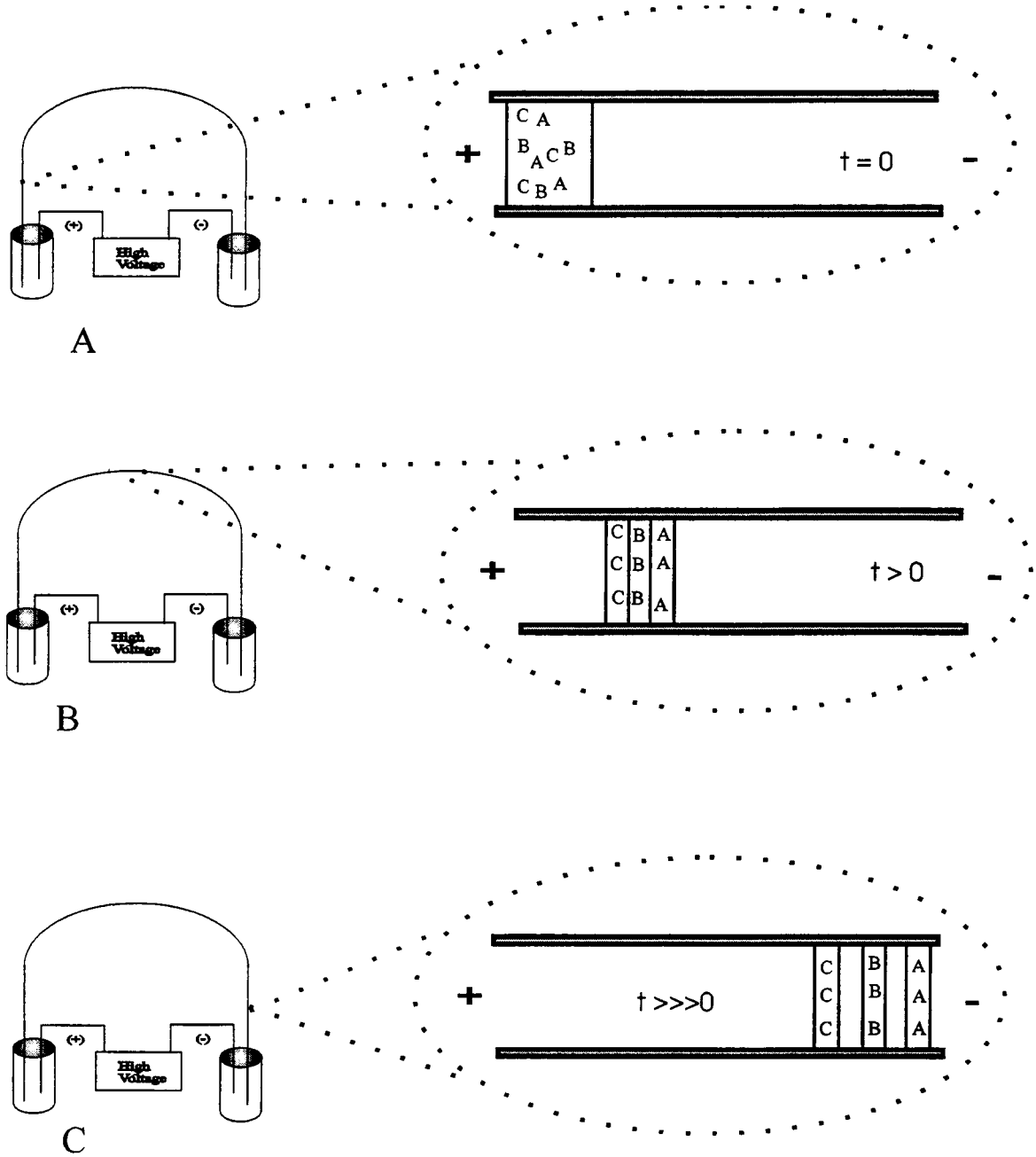


Figure 1.5. Progression of CE separation.

As is readily seen, the small pathlength (approximately the inner diameter of the capillary, 25 - 75 μm) is the major limitation to sensitivity. Although fluorometric methods are less universal, they exhibit significantly increased sensitivity. The signal intensity in fluorometry, P for small absorbances is given by

$$P = P_o 2.3\varepsilon' b' C\Phi_f k \quad (1.14)$$

where, P_o is excitation source power, Φ_f is the fluorescence quantum efficiency and k is the instrumental efficiency for collection and detection of the fluorescence emission. The linear relationship of increased fluorescence emission for increased source intensity is termed 'the fluorescence advantage.' This advantage combined with the selectivity inherent to fluorescence allows for excellent detectability. Furthermore, the unique attributes of lasers make them well suited as sources for fluorescence detection. The high photon flux inherent to lasers assists in realizing the fluorescence advantage. The high degree of collimation aids in focusing of the laser beam though capillary columns. The monochromatic nature of lasers, while to some degree limiting the range of analytes to be detected, eliminates the need to select necessary wavelengths.

While lasers are becoming increasingly more popular as sources for CE detection schemes, they are not without drawbacks. The use of laser induced fluorescence requires that the analyte of interest exhibit fluorescent properties. While this can add a degree of selectivity, it also limits the possible number of analytes. This problem can be circumvented by derivatizing the analyte of choice either pre- or post-column. However, problems associated with fluorescent labeling have also limited this approach [6]. On-column fluorescent labeling can be achieved by incorporating an appropriate fluorophore

into the running buffer allowing the formation of an analyte-fluorophore complex. Both pre-column and on-column fluorescent labeling are utilized in this work. The large amount of light scatter characterized by on-column fluorescence detection generates a great deal of background signal. While several approaches to solve this problem include sophisticated means of spectral and spatial filtering of unwanted radiation, perhaps the greatest advances have been realized with the advent of post-column detectors. These systems, based on a sheath flow cuvette, involve the transfer of the capillary contents to a flow chamber of excellent optical quality with little to no loss of separation efficiency [7-9]. The advantage is recognized in very high detection sensitivity due to an increase in signal to noise (S/N).

DNA Analysis by Capillary Electrophoresis

DNA Basics

A DNA strand is a long polymer of nucleotide subunits, each containing a backbone in which phosphate groups alternate with deoxyribose sugar groups and are linked through phosphodiester bonds. Attached to the sugar ring of each nucleotide is one of four nitrogenous bases (adenine (A), guanine (G), thymine (T) or cytosine (C)). The structure of a single stranded DNA polymer segment is depicted in Figure 1.6. Most DNA exists in a double-helical structure where the phosphate and deoxyribose units are on the outside and the bases are oriented toward the center of the molecule, resulting in a hydrophobic interior and ionic exterior. The two chains are held together by hydrogen

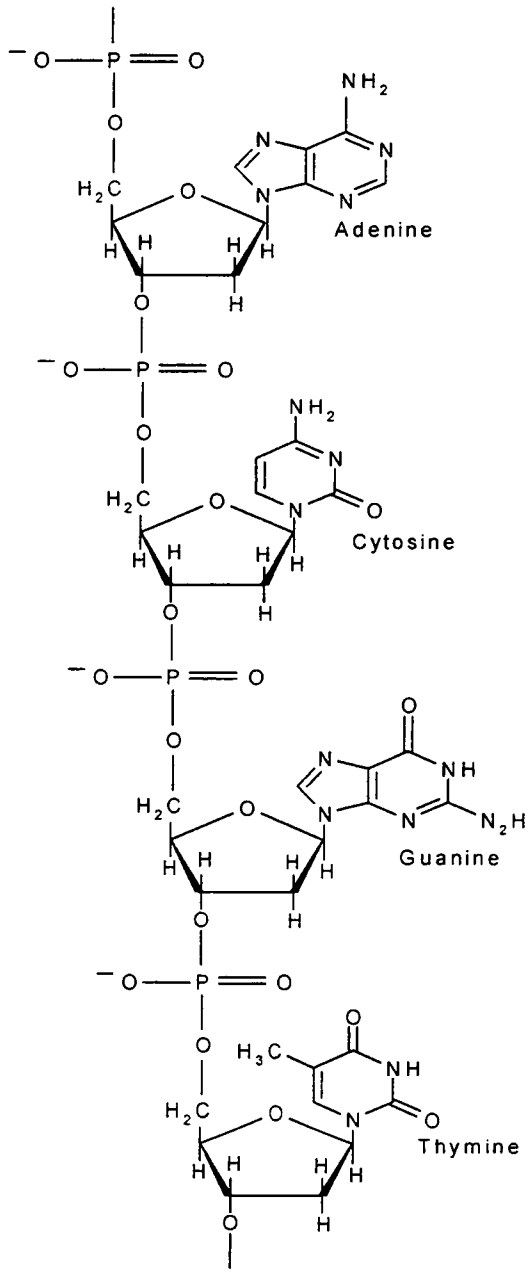


Figure 1.6. Structure of single stranded DNA segment.

bonding between the pairs of bases. The sequence of bases along the DNA strand is not limited, but it is this precise sequence that encodes the necessary genetic information.

Background

Slab-gel electrophoresis is the preeminent technique for the separation of nucleic acids. Based on sieving technology, DNA migrate though the fixed pores of a polymeric material (often acrylamide or agarose) driven by an electric field toward the appropriate electrode. Small molecules migrate easily, not experiencing significant resistance from the pores. Large molecules must travel a more tortuous path, moving laterally across the gel until encountering a pore large enough to pass through. This results in an inverse proportionality between a molecule's mobility and its size.

Hjerten, a pioneer in the field of electrophoresis, is largely responsible for the transition of slab-gels to the column format. His work describing polyacrylamide filled tubes set the framework for capillary electrophoresis of DNA [10]. Problems associated with biopolymer separations in CE include the adsorption of these molecules to the negatively charged column wall and the effect that EOF has on their migration times. In this light, capillary modification procedures which coat the inner column surface effectively reduce adsorption while significantly reducing or eliminating EOF. While the adsorption of negatively charged ions (i.e., DNA) to the negatively charged fused silica does not seem plausible, we often observe band profiles that mimic those observed when adsorption is present. In the separation of negatively charged DNA, the electrophoretic system must be operated in reversed polarity mode to allow the DNA to migrate towards

the anode. Figure 1.7 shows the difference between unmodified and deactivated (coated columns). A thorough description of the column modification procedure used in this work can be found in Chapter 2.

Theory of DNA Migration

A number of models for the migration of flexible macromolecules (e.g., DNA) through polymer matrices are currently accepted. The basis of these models is the sieving nature of a polymer matrix containing a network of randomly distributed pores of average size, ξ', and DNA molecules with a radius of gyration, R_g. The Ogston model describes a DNA molecule electrophoreseing through a polymer network as an unperturbed sphere [11, 12]. These DNA spheres must diffuse laterally until encountering a pore of appropriate size for passage. The electrophoretic mobility of the DNA, μ, is dependent upon the fractional volume of the pores available and is given by

$$\mu = \mu_0 e^{-KC(r+R_g)^2} \quad (1.15)$$

where μ_0 is free solution mobility of the DNA, K is the retardation coefficient, C is polymer concentration and r is thickness of the polymer strands.

The development of the reptation model of migration was brought about by the assumption that if a large DNA molecule behaved as an undeformable sphere, it would be trapped by gels having pore sizes too small to allow passage [13-15]. Migration in this model discounts the lateral movement of DNA strands and suggests that these coiled DNAs will reptate, head first, through the polymer network by stretching and relaxing as

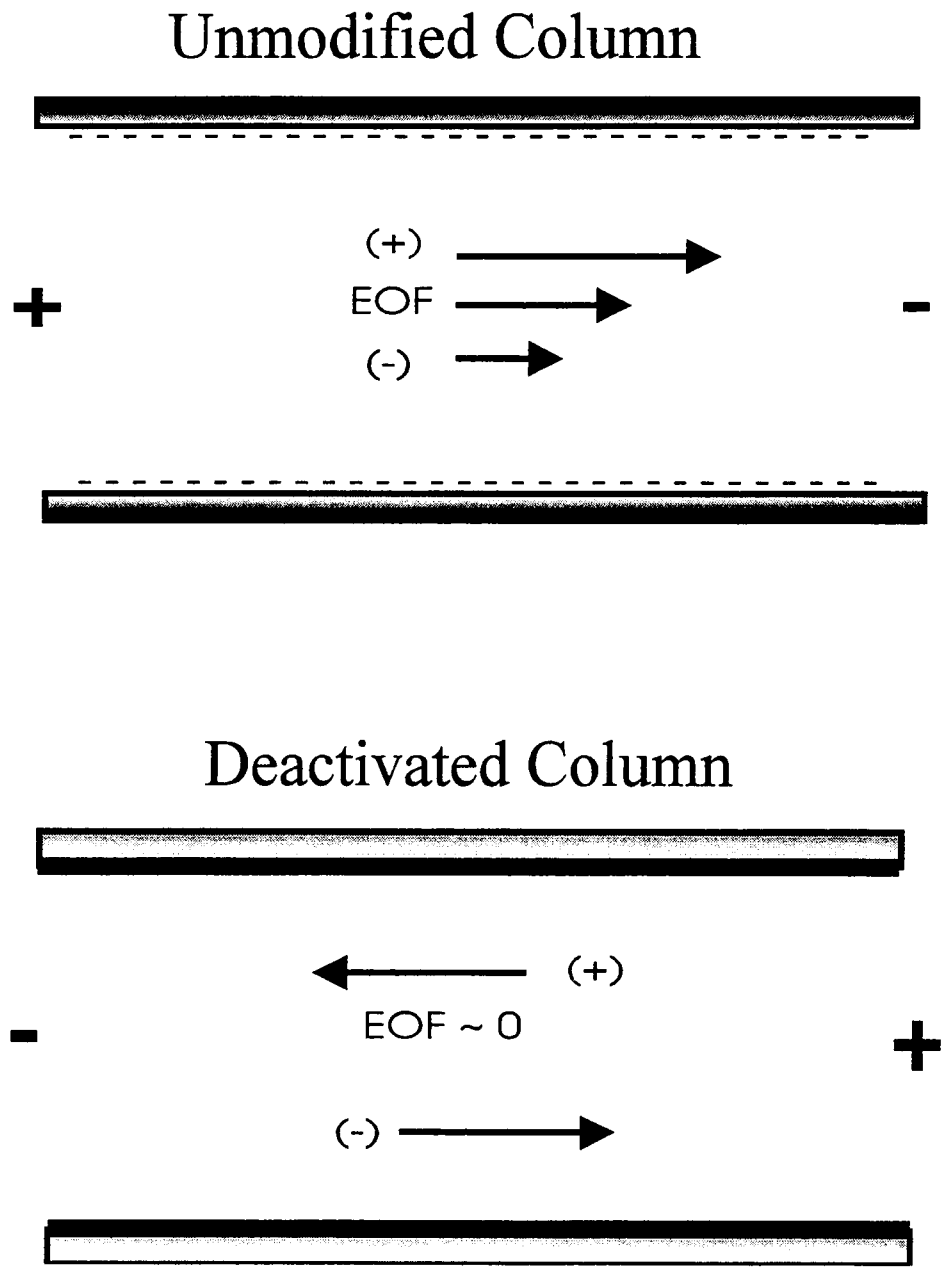


Figure 1.7. Difference between unmodified and deactivated (coated) columns.

necessary for passage. This model places greater emphasis of DNA mobility on experimental parameters (i.e., field strength, pore size). Lumpkin *et al.* described a biased reptation model that takes into account large field strengths [16]. A field induced orientation of the head of the DNA molecule is experienced resulting in greater elongation as field strength increases. Therefore as DNA size or electric field increase, the dependence of mobility on molecular size is decreased. The mobility of DNA fragments as described by the biased reptation model is given by

$$\mu = \left(\frac{1}{N'} + bE^2 \right) \quad (1.16)$$

where N' is the number of base pairs (bp) on the DNA, b is a function of polymer mesh size and DNA size (bp) and charge. It should be noted that the above account is by no means comprehensive; however, the migration mechanisms described accurately convey the complexity of DNA migration through polymeric matrices.

Sieving Matrices

The sieving matrix in slab-gel electrophoresis serves the dual propose of anticonvection and molecular sieving. The former is not necessary in CE due to the efficient heat dissipation of capillary columns. Because the net charge on a DNA molecule is directly proportional to its mass regardless of the number of bp on the DNA, free solution CE is not successful. As in traditional electrophoresis, a sieving matrix is necessary to afford the additional degree of selectivity necessary for the separation of DNA. The migration models described above also apply to DNA migrating through a

polymer network in a CE column, a depiction of which is seen in Figure 1.8. Two DNA molecules of different size exhibit identical electrophoretic mobilities; however, the selectivity imparted by the sieving matrix affords the difference in observed mobilities resulting in separation.

Early work in CE separations of DNA involved the use of cross-linked polyacrylamide sieving matrices. A major drawback to these types of gel-filled columns was that the matrix was chemically bonded to the inside of the column. Problems associated with this approach include both the reproducible manufacturing of stable gels and the inability to introduce fresh matrix when necessary. An increasingly popular approach has been the introduction of soluble polymers as sieving matrices for these types of separations. The advantage inherent to soluble polymers is that cross-linking to form the mesh network is unnecessary and thus eliminates the need to bond the matrix to the column wall. This allows the flushing and replacement of fresh polymer when necessary. A number of soluble polymers have been employed in CE DNA separations: agarose [17, 18], hydroxyethyl cellulose [19, 20], hydroxypropyl methyl cellulose [21, 22], methyl cellulose [23, 24], polyvinyl alcohol [25, 26], polyethylene oxide [27], glucomannan [28] and liquid polyacrylamide [29]. These soluble polymer incorporated separation strategies are generally referred to as size-selective CE (SSCE).

It is important to point out that soluble polymers represent dynamic systems whose pore sizes and positions are not fixed. At concentrations greater than a polymer's entanglement threshold, Φ^*, individual polymer strands begin to interact with each other,

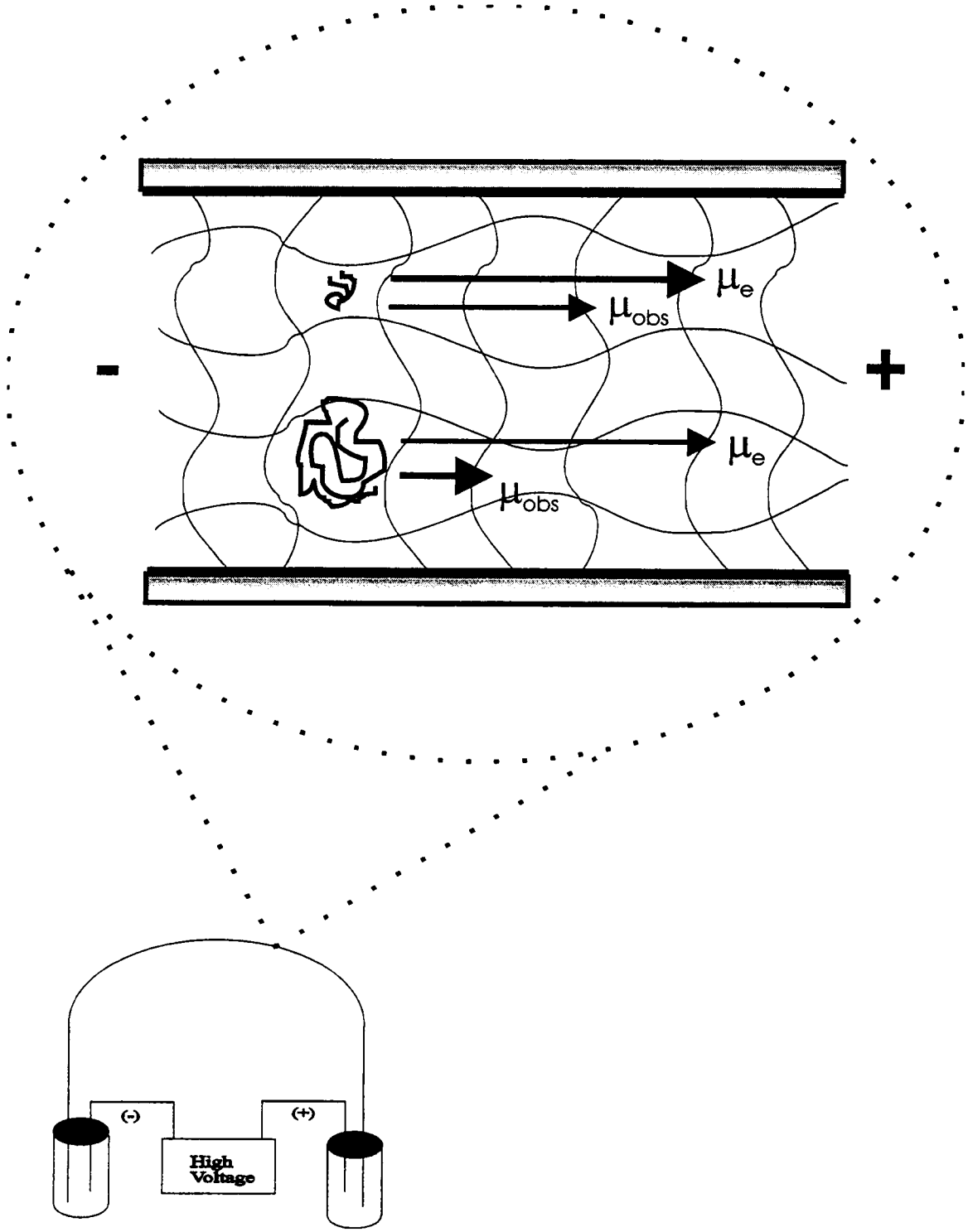


Figure 1.8. DNA migration though a polymer network in CE.

effectively producing an entangled array. An estimation of mesh (pore) size, ξ', is given by

$$\xi' = 1.43 R_p \left(\frac{\Phi}{\Phi^*} \right)^{-3/4} \quad (1.17)$$

where R_p is radius of gyration of the polymer blob and Φ is polymer concentration. The spaces between individual polymer strands act in a similar manner to the rigid, fixed pores in cross-linked gels allowing separation based on size.

STATEMENT OF PROBLEM

The unique characteristics inherent to CE present novel means to investigate analyses performed by more traditional biochemical methods. Most important among these are speed, efficiency (resolving power) and small sample requirements. As noted in the following chapters CE is already showing widespread success as a method for the analysis of a variety of biomolecules. The goal of the research presented is to exploit the unique characteristics of CE in the fundamental development of strategies for biochemical analysis of DNA and protein molecules.

ORGANIZATION

Because each topic presented represents CE approaches to different systems the following data chapters will contain their own Introduction, Experimental and Results and Discussion sections. Chapter 2 describes the optimization of separation conditions in the size-selective separation of DNA fragments. A computer program based on simplex

optimization is used to generate experimental conditions which render adequate resolution while minimizing total analysis time. The ability to adjust computational parameters to effect different separation needs is demonstrated.

Chapter 3 details the development of a capillary electrophoretic mobility shift assay for the study of protein-DNA interactions. The influences of important parameters such as running buffer constituents as well as quantitative and qualitative studies are presented. Also presented are preliminary studies aimed at determining binding sites on DNA fragments in a mixture of fragments of different sizes.

Chapter 4 describes the diagnostic capabilities of CE for the temporal analysis of DNA restriction digests. It is shown that the course of a digest can be followed, noting the creation and diminishing of fragment peaks as the digestion progresses toward completion. This temporal analysis is then used to determine when a certain sized fragment is generated during a partial digest.

Chapter 5 summarizes the findings in the previous data chapters and draws some significant conclusions about the future of CE as a biochemical analysis method.

CHAPTER 2

Optimization of Separation Conditions in the Size-Selective Capillary Electrophoretic Separation of DNA Fragments

INTRODUCTION

The use of CE as a DNA analysis method continues to grow in popularity. Investigation of DNA is traditionally a complex process that results in time-consuming steps to extract desired information. Recently, considerable effort has been employed to make CE a more robust technique for DNA analysis. CE formats for separation of DNA fragments are extremely diverse and include, among others, using entangled polymer running buffers [30-34], capillary gel electrophoresis [32-36], capillary arrays [37-39], and mixed polymer matrices [40, 41]. Typical DNA samples can range from fragments of restriction enzyme digests of double stranded DNA chains [29, 42], fragments of selected sections of DNA obtained by the polymerase chain reaction (PCR) [43-46], and fragments from Sanger reactions [27, 41]. While the above methods frequently achieve high resolution and efficiency, it is usually at the expense of analysis time, most particularly due to the fact that these techniques rely on a single set of experimental conditions. In fact, the determination and generation of such experimental conditions, for even a single analysis, can be a time consuming process.

Most separations of DNA samples are plagued by the chromatographic elution problem, wherein the complete separation of the components in a sample that are diverse with respect to their chromatographic/electrophoretic behavior is often not easily

attainable. As described in Chapter 1 models of DNA fragment migration differ according to the size of the fragment. Small DNA fragments migrate via the Ogston model, while larger fragments migrate via the biased reptation model. These models differ greatly in form and in their dependence on experimental conditions (see Equations 1.15 and 1.16) [42]. This problem is exacerbated by the need to extract specific information from DNA analyses in a timely fashion. In these 'time effective' separations, entangled polymer concentration, capillary length and applied field are perhaps the greatest concerns.

Higher concentrations of entangled polymer can lead to increased resolution for shorter fragments, but generally at the expense of analysis time. In practice, the effective mesh size of the entangled polymer in the running buffer is generally matched to the sizes of the fragments that are to be analyzed [42]. Consequently, the conditions that separate low base pair (bp) fragments with good resolution and efficiency are generally not adequate to achieve similar performance for larger bp fragments in the same sample. Shorter analysis times are effectively realized with shorter column length, but at the expense of efficiency, and hence resolution. Herein the effects of polymer concentration and capillary length are investigated to arrive at experimental conditions which render adequate resolution while minimizing total analysis time. Resolving this paradox is the focus of the work presented in this chapter.

This high degree of variability among DNA samples poses potentially the largest hurdle to wide-scale use of CE as a DNA analysis technique, especially in light of CE's increasing role in such efforts as the human genome project. However, inherent among

DNA samples is that base pair electrophoretic behavior is largely independent of its sequence. An empirical equation, based on this principle, has been derived that predicts the mobility of DNA fragments, in methyl cellulose (MC) running buffers, based on bp number and MC concentration. The unique electrophoretic behavior of DNA, along with instrumentation developed in our laboratory (described in the experimental section) and a computer optimization program based on the simplex method, facilitates the efficient optimization of experimental conditions prior to actual separation experiments.

The computer program effectively arrives at a compromise for the chromatographic elution problem by utilizing the simplex method to optimize for those conditions which are most influential in these types of separations. The ability to generate DNA fragment mobility data allows column length and MC concentration to be adjusted for optimal separation performance of samples of known content. The utility of this type of optimization can also be realized for samples whose content is not completely known. For example, it might be known that a sample contains fragments that are close in size and encompass specific size ranges. Alternately, a preliminary separation can reveal unresolved fragments of approximate sizes intermixed with easily resolved fragments. This sketchy information regarding the composition of the sample can be used with the computational methods to rapidly optimize conditions for the separation of all fragments in a sample.

The focus of the work presented in this chapter is the development of models and computational methods to optimize DNA separations based on content of sample and empirically derived relationships between fragment size, MC concentration and column

length. Illustrations of the correlation between computational and experimental work, as well as the influence of adjustable parameters are presented.

EXPERIMENTAL

Separations

Materials

Tris(hydroxymethyl)aminomethane, boric acid and ethylenediaminetetraacetic acid (EDTA) were obtained from Sigma (St. Louis, MO, USA) and used to prepare pH 8.5 tris-boric acid-EDTA buffers (TBE). Methyl cellulose (MC: 4,000 cp for 2.0 % w/v), ethidium bromide (EB), acrylamide, ammonium persulfate (APS), N,N,N′,N′-tetramethylethylenediamine (TEMED), γ-methacryloxypropyltrimethoxysilane (γ-MTMS), and samples of ΦX-174 DNA *Hae*III digest (480 μg mL^{-1}) and plasmid pBR322 DNA (0.25 μg mL^{-1}) were also obtained from Sigma. *Hinf*I enzyme (10 units μL^{-1}) with incubation buffer was obtained from Boehringer Mannheim (Indianapolis, IN, USA). Chroma Spin TE-100 columns were obtained from CloneTech (Palo Alto, CA, USA) and used to purify the *Hinf*I digest of the pBR322 plasmid DNA. Fused silica capillary (75 μm i.d. x 365 μm o.d.) was obtained from Polymicro Technologies (Phoenix, AZ, USA).

Column Preparation

As discussed in Chapter 1, DNA analysis by CE is typically performed in columns that have undergone deactivation of the capillary's inner surface. This is necessary both to eliminate electroosmotic flow and to counteract adsorption of DNA

fragments to the charged surface of the column which often results in extremely inefficient separations. Fused silica capillaries were coated with covalently bonded linear polyacrylamide, as depicted in Figure 2.1, using a modified version of Hjerten's original procedures [47, 48]. The basis of the method is the use of a bifunctional silane that reacts both with the capillary wall and with a polymerization process taking place in the column.

The capillaries, prepared in duplicate, were sequentially washed with 5 M NaOH for 1 hour, water for 10 minutes, 5 M HCl for 1 hour and water again for 10 minutes. The base treatment served to effectively remove impurities while the acid treatment served to generate the exposed silanol groups (Si-OH) necessary for reaction with the silanating agent γ- MTMS. The capillaries were then washed for 10 minutes with acetone followed by reaction with a 1:1 (v/v) solution of acetone: γ-MTMS for 24 hours. The γ-MTMS, also known as Bind-Silane, is bonded to the inside of the capillary via reaction between the methoxy groups of γ-MTMS and the silanol groups of the column surface resulting in a Si-O-Si-C bond (see Figure 2.1 (A)). This reaction has also been used in the binding of gel slabs to glass surfaces [49]. This serves to activate the surface of the capillary, preparing it for reaction with polyacrylamide. Excess γ-MTMS was removed by flushing the capillary with water for 10 minutes.

The free methacryl groups of the γ-MTMS are subsequently reacted in-situ with the acryl monomers of acrylamide to form a monomolecular layer of linear polyacrylamide (see Figure 2.1 (B)). A fresh solution of acrylamide was prepared by adding 0.4 g of acrylamide to 9 mL of 50 mM K_2HPO_4, buffered to pH 6.8. A 67 mM solution of APS was also prepared and both solutions were degassed by bubbling helium

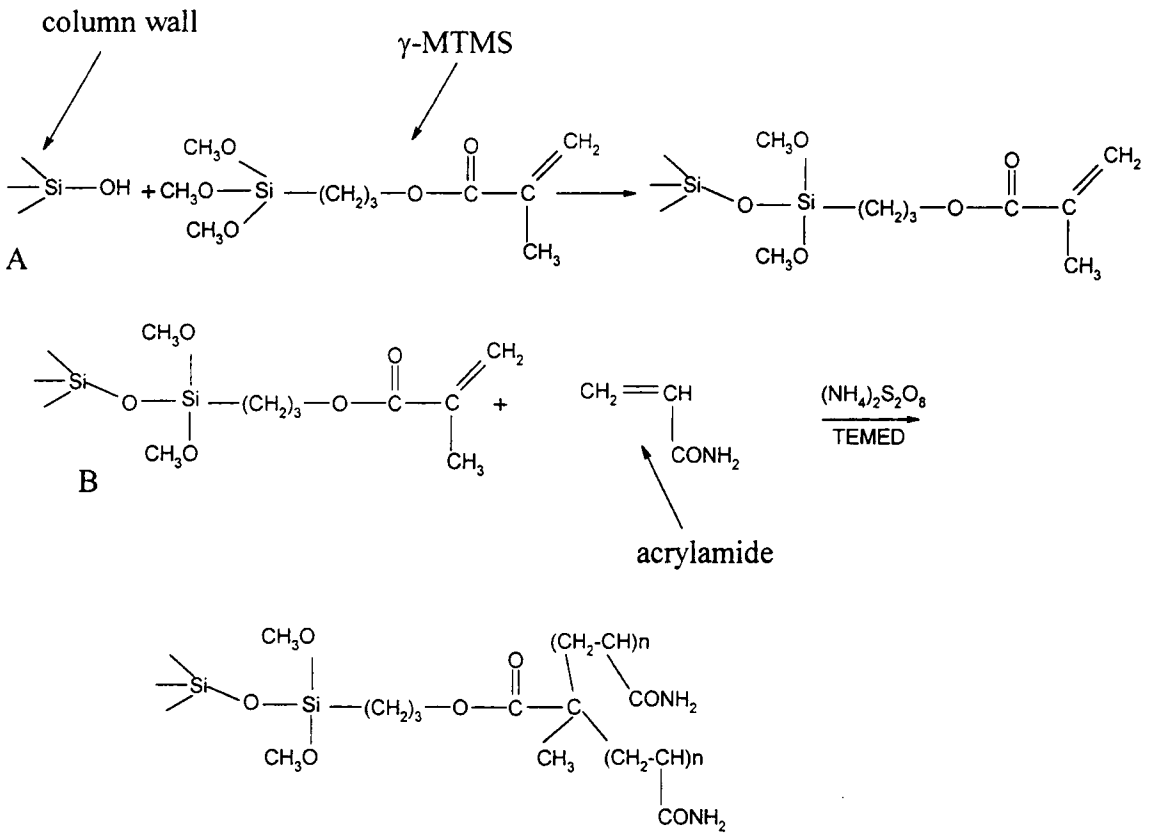


Figure 2.1. Covalently bonded linear polyacrylamide coating reaction.

through them for approximately 30 minutes. The polymerization reaction was initiated and catalyzed by adding 1 mL of the APS solution to the acrylamide solution and spiking with 7.5 μL TEMED. The resulting polyacrylamide (4 % w/v) was immediately pushed through the columns and allowed to react for 20-24 hours. Unreacted polyacrylamide was removed by flushing the capillary with 10 mM H_3PO_4.

Changes made in Hjerten's original procedures involved a more severe pretreatment of the column prior to introduction of the organosilane including a 50-fold increase in the concentration of base and acid and a 12-fold increase in reaction time for these rinses. Treatment with 5 M NaOH and 5 M HCl served to create a more reproducible coating surface by cleaning and etching the interior surface of the column, thus increasing the Si-OH concentration. Other incorporated changes involved increasing the reaction times for both the silanating and polymerization reactions up to 24 hours. While many researchers have studied, in detail, the preparation and application of linear polyacrylamide coated columns, this task was not undertaken for the studies presented in this and the following chapters. The method outlined above produced successful columns, in terms of observed efficiency and resolution on a regular basis.

Columns were stored and rinsed between runs with 10 mM H_3PO_4. Column performance was evaluated by simple visual inspection of a separation of the ΦX-174 DNA *Hae*III digest. Observed efficiencies for the 1354 bp fragment of this digest were 1-3 x 10^6 plates/meter. Severe band broadening and tailing were characteristic of poorly coated and/or degrading columns. Columns prepared in this manner were generally functional for 1-3 weeks, depending on the number of separations performed.

Polyacrylamide coated capillaries of 40 cm length were prepared for the optimization studies outlined in this chapter. The exterior polyimide coating was removed from the central 30 cm of the column to facilitate multipoint detection. This was achieved by subjecting the desired region of the column to 50 % fuming sulfuric acid. The resulting exposed region of the capillary was carefully wiped with methanol and then water.

On-Column Intercalation with Ethidium Bromide

As mentioned in Chapter 1, detection schemes in CE are nearly as varied as the analytes that are studied. When dealing with DNA separations, the two most prevalent optical detection methodologies employed are absorbance and fluorescence. While absorbance detection for DNA is considerably simpler in terms of ease of use and commercial availability of instrumentation, it is not without serious limitation. Most significant among these limitations are lack of sensitivity (in terms of achievable detection limits) and potential interference by other components that may be needed in the separation strategy. An attractive alternative involves laser induced fluorescence detection, in particular via the addition of an intercalation dye in the running buffer, resulting in on-column labeling of double stranded (ds) DNA fragments.

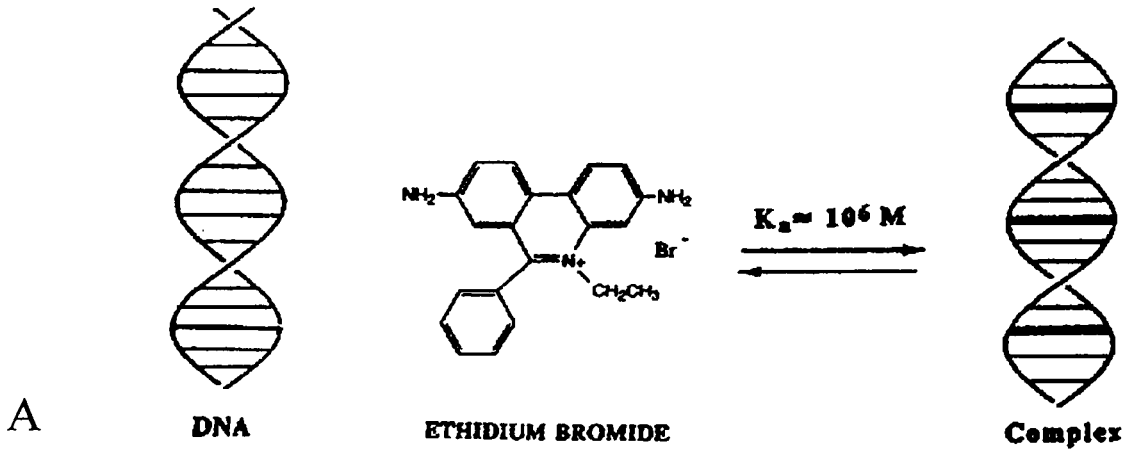
The use of ethidium bromide (EB) as a spectroscopic probe in DNA studies is well known [50]. In fact, the spectral properties of EB make it one of the most used on-column intercalation dyes for laser induced fluorescence detection of DNA in CE. Upon intercalation within the hydrophobic interior of the double helix (K_{eq} = 10^5-10^6) the

fluorescence quantum yield of EB increases approximately 20 fold [50, 51]. As depicted in Figure 2.2 (A) EB exhibits a maximum intercalation stoichiometry of one ethidium cation per four bp [50, 51]. Upon intercalation EB also exhibits a convenient spectral shift as seen in Figure 2.2 (B). Laser induced fluorescence detection of EB intercalated DNA has been achieved with helium-cadmium (He-Cd, 325 nm) and helium-neon "greenie" (He-Ne, 543.5 nm) lasers [27, 30, 52, 53]. However the ease of use, minimal cost and convenient spectral characteristics alluded to above make the He-Ne "greenie" laser an excellent choice for routine as well as in depth analyses.

EB not only affords detection of DNA fragments, but it is also known to enhance resolution in CE separations [32]. Recalling that one ethidium cation can intercalate per every four bp, a uniform reduction in the intrinsic mobility of the labeled fragments is realized [54]. EB intercalation increases the DNA chain length and induces conformational changes. This change in intrinsic mobility (μ, system retention) leads to a greater absolute difference in migration times between fragments ($\Delta\mu$, selectivity) and results in increased resolution as seen in Equation 1.9.

Apparatus

The traditional in-house fabricated laser induced fluorescence detection apparatus for CE does not facilitate multipoint detection. However, the need to easily and reproducibly adjust the point of detection was paramount to experimental verification of the developed computer optimization program. Figure 2.3 represents a schematic diagram of the CE laser fluorometry instrumentation used in this study [55]. The basic



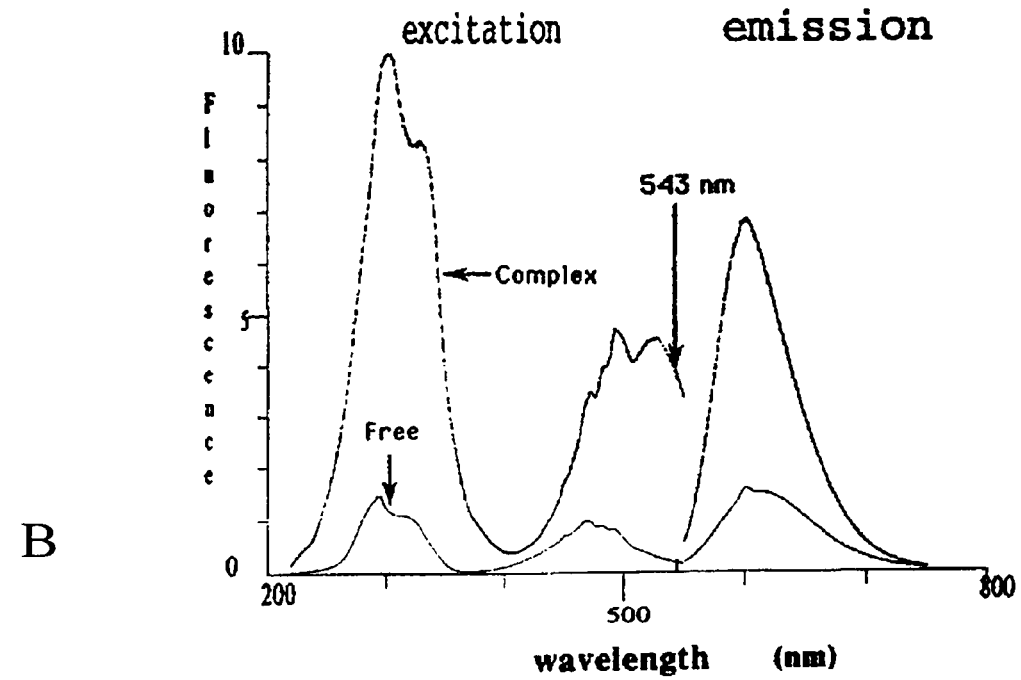


Figure 2.2. DNA-ethidium bromide (EB) intercalation event (A) and related spectroscopy (B).

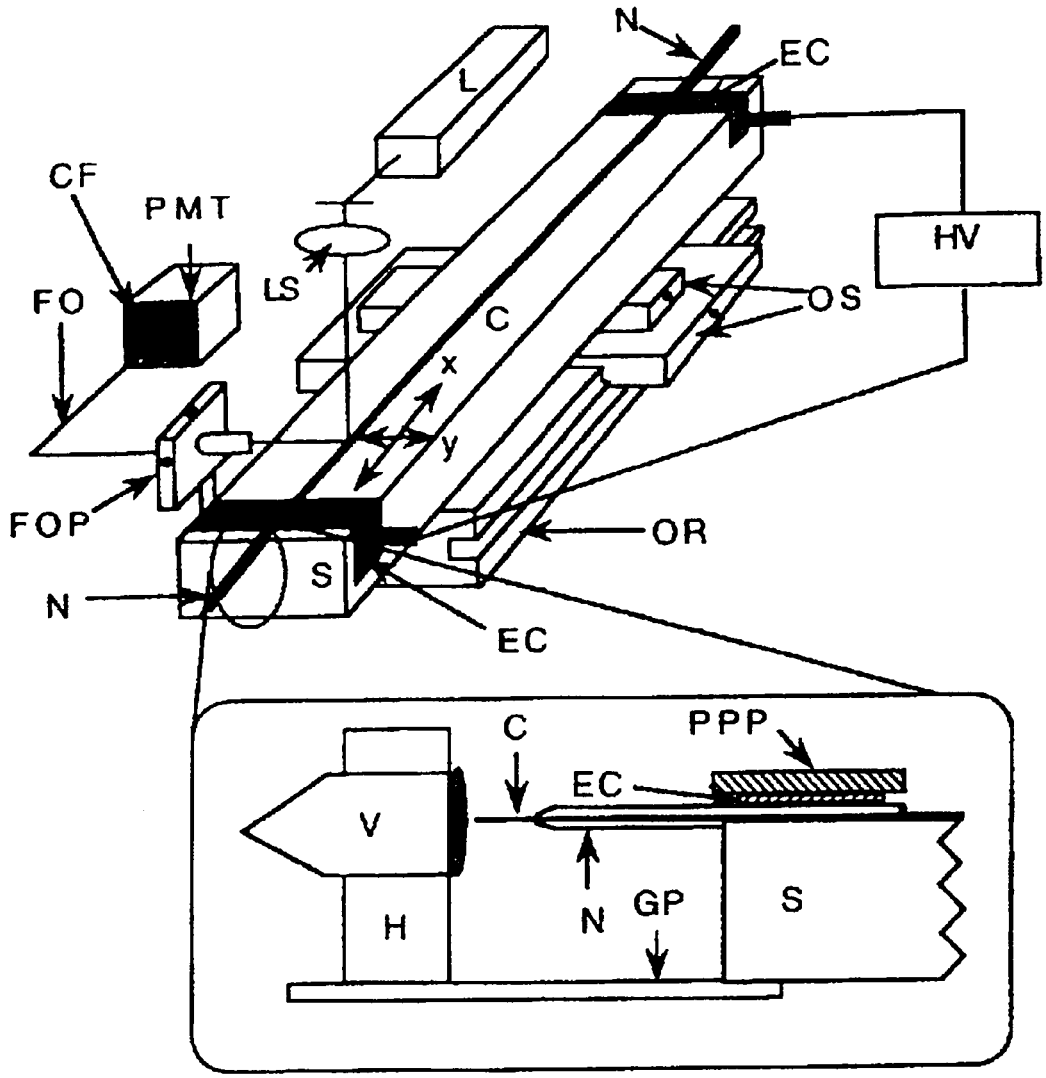


Figure 2.3. Laser fluorometry apparatus for multi-point detection. C: capillary, N: needles, S: Plexiglas substrate, OR: optical rail, OS: optical stages, EC: electrical connection, HV: high voltage, L: laser, LS: lens, CF: cut-on filter, PMT: photomultiplier tube, FO: fiber optic, FOP: fiber optic positioner, V: vial, H: vial holder, GP: guide plate, PPP: Plexiglas pressure plate. Reprinted from reference 55.

experimental apparatus and additional modifications, including data collection and processing, are described in detail in a prior report [55]. Briefly, the CE laser fluorometry instrumentation consists of a capillary column on a Plexiglas substrate attached to an optical translational device. The translational device was attached to a 20:1 speed reducer (Boston Gear, Quincy, MA, USA) via a plastic roller chain belt. A PS/2 model 50 IBM computer was used to control translation and data collection. This apparatus has a linear translation velocity range from 2.0 mm/min to approximately 600 mm/min and any portion of the column from 7-33 cm can serve as the point of detection. A PMS Electro-Optic He-Ne (Edmund Scientific, Barrington, NJ, USA) "greenie" laser (1.5 mW at 543.5 nm) in conjunction with a 25 mm *f*/1 quartz lens served as the excitation source, while a 200 μm core diameter quartz fiber optic provided collection of the EB intercalated fluorescence emission. Rejection of plasma lines emanating from the laser was accomplished by placing a 543.5 nm (10 ± 2 nm full width half maximum, FWHM) laser line filter in the path of the laser output. The emission was isolated using a 570 nm cut-on filter and monitored using a Hamamatsu (Bridgewater, NJ, USA) R282 photomultiplier tube (PMT) and a Pacific Precision Instruments (Concord, CA, USA) Model 126 photometer.

Running Buffer Preparation

Stock solutions of TBE buffer were prepared by weighing out the appropriate amount of solid and diluting in deionized H_2O to a final concentration of 0.45 M Tris, 0.45 M boric acid and 0.01 M EDTA. A 1:10 dilution of the stock solution yielded buffer

containing 45 mM Tris, 45 mM boric acid and 1 mM EDTA and referred to as 45 mM TBE. Concentrated stock solutions of MC (1.0 % w/v) were prepared by dispersing the appropriate amount of solid MC (100,000 MW) into 80-90 °C 45 mM TBE buffer. Upon dissolution of the MC in the TBE, the solution was cooled on ice with stirring to a clear viscous consistency.

Stock MC solutions were used to prepare day-to-day running buffers by serial dilution with the appropriate amount of 45 mM TBE buffer. Desired concentrations of MC running buffers ranged from 0.2 % to 0.6 %. Laser induced fluorescence detection was facilitated by adding 10 μL of concentrated EB stock (1 mg mL^{-1}) to 10 mL aliquots of MC running buffers, yielding a final [EB] of 2.5 μM. The final step in buffer preparation involved degassing the solutions under vacuum to remove any bubbles that might have been generated during the preparation steps. This step is particularly important to avoid any band dispersion and loss of resolution that may result from air bubbles present in the running buffer.

Computational Methods

Materials

Microsoft QuickBasic 4.5 (Richmond, WA, USA) was utilized for development of the matrix (MAT) and simplex (SIM) programs. Both programs utilized in this work were developed in-house. Each program incorporates an empirical mobility equation (Equation 2.1, discussed later) and the latter program incorporates a simplex algorithm patterned after that of Berridge [56]. Programming and graphics generation were

performed on a Gateway 2000 386DX25 and a Midwest Micro Soundbook II 486DX4-100, respectively. Three-dimensional (3-D) plots representing separation performance, from MAT data, and two-dimensional (2-D) simulated electropherograms, from SIM data, were generated by PSI-Plot 2.11 (Poly Software International, Ltd., Salt Lake City, UT, USA). The simulated electropherograms, based on triangulation, considered retention time and base width of each peak.

Computational Methods Development

Both the SIM and MAT programs were written specifically for the separation of DNA fragments in MC running buffers. Both utilize an empirical equation which relates percent MC and the number of bp on a DNA fragment to the mobility of that fragment. The SIM-program generates a complete set of simulation data at each [MC] and capillary length and generates optimized conditions. Mobility data for the different fragments in a sample are used to generate a simulated separation. The MAT-program increments MC percentage and capillary length over selected ranges and generates a matrix of a separation performance parameter which can then be displayed for visualization.

The SIM-program uses the simplex method for optimizing separation conditions [57]. The simplex method, briefly, treats the differing combinations of experimental parameters as a contour map. The height of a given response is related to its merit as an experimental result relative to other combinations of parameters. A simplex is a geometrical representation of one more than the number of parameters being optimized. Since this is a two-dimensional simplex (a triangle), varying capillary length and methyl

cellulose concentration ([MC]), three experiments are initially entered into the program. For each of the initial experiments, results are ranked (from best to worst) and the least optimal set of parameters is rejected and replaced with a response obtained by reflection through the midpoint of the two retained parameters. This process is repeated until an optimum is obtained [58].

A generic geometrical representation of a two dimensional simplex, superimposed on a contour map, is depicted in Figure 2.4. Data points 1, 2 and 3 are the responses that result from the input of three different sets of values for parameters 1 and 2. These points are ranked from best to worst. It is determined, in this arbitrary case, that point 1 yields the worst ranking. Reflection of point 1 through the midpoint of points 2 and 3 results in the new data point 4. A new optimization step is performed using the parameters for points 2 and 3 and now point 4. The results are again ranked and followed by reflection. This process is continued until an optimum (point 12) is reached.

The ability to optimize capillary length and [MC] is contingent upon the dependence of bp electrophoretic mobility on these parameters. The mobility of a DNA fragment is calculated from an empirically determined equation which relates mobility, μ, of that DNA fragment to [MC] (MW 100,000) in weight percent (w/v) and the number of bp on that DNA fragment. The equation used in this work is shown below:

$$\mu = 1.2\cdot 10^{-4} - \left(1.2\cdot 10^{-4}\cdot \log[MC]\right) + \left[9.2\cdot 10^{-3}\cdot (bp)^{-1}\right] - \left[2.9\cdot 10^{-3}\cdot \log[MC]\cdot (bp)^{-1}\right] \quad (2.1)$$

Equation 2.1 was developed by generating a set of plots of fragment mobility versus size at different methyl cellulose concentrations (0.2 % - 0.6 %) for the 11 fragments (bp = 72,

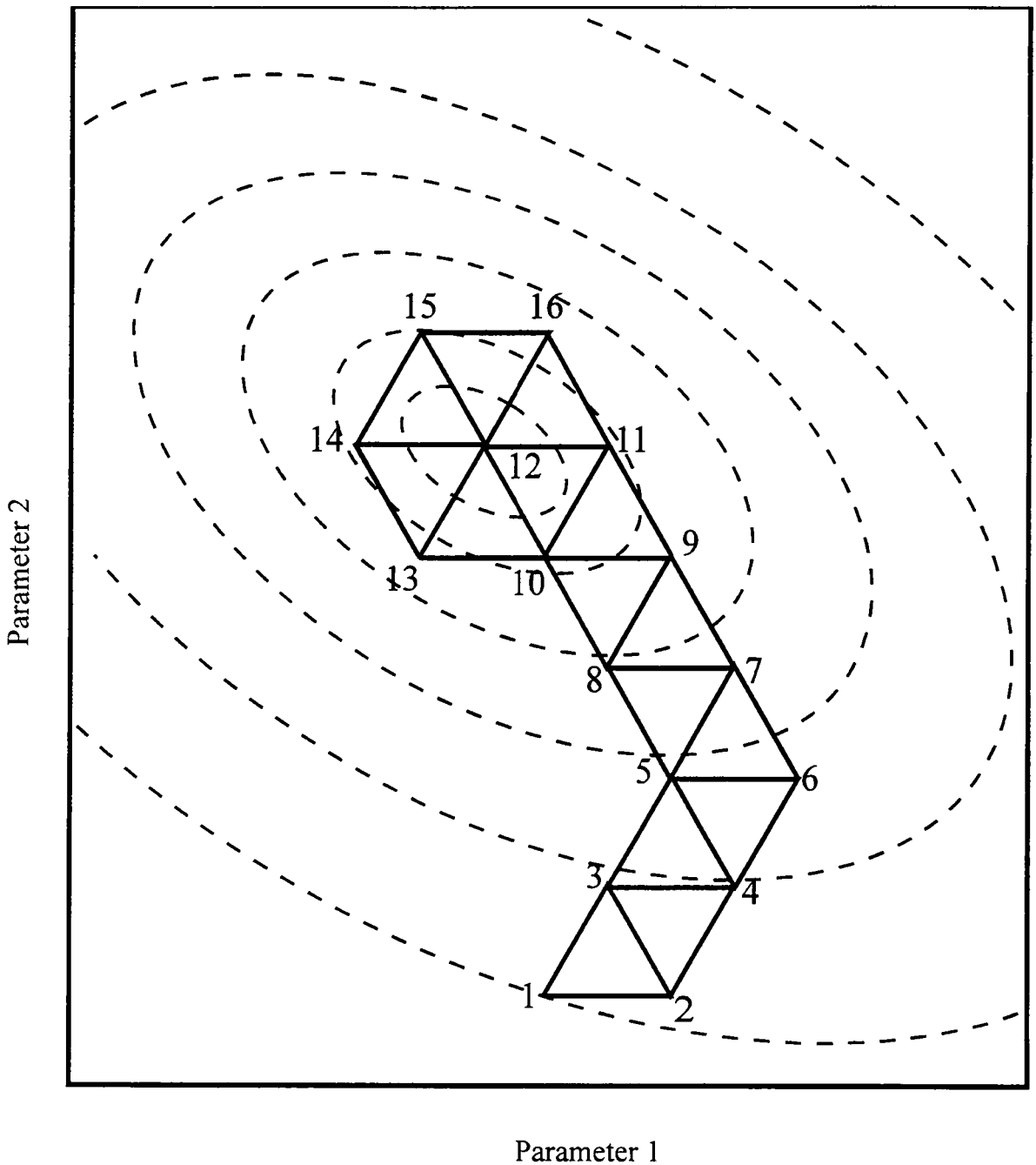


Figure 2.4. Generic representation of a 2-dimensional simplex superimposed on a contour map.

118, 194, 234, 271, 281, 310, 603, 872, 1078 and 1353) in the ΦX-174 *Hae*III digest. It was observed that each set of data had the same general form and differed only by numerical constants. A generalized equation was discovered by simply curve-fitting the data. A relationship for the numerical constants as a function of [MC] was determined by plotting the change in the constants as a function of concentration. This information was algebraically combined to generate Equation 2.1. It must be stressed that the equation is valid over [MC]'s in the 0.2 % to 0.6 % range. Equation 2.1 is not derived from theory but simply relates observed mobility at a chosen [MC] for a given bp fragment. Moreover, Equation 2.1 results in reasonable correlation between predicted and observed mobility (see Results and Discussion).

After each optimization step a chromatographic response function (*CRF*) is calculated that serves as a measure of the success of the simulated separation. The *CRF* used in this work is defined here by:

$$CRF = a\sum \ln\frac{R_{so}}{R_s} + b\sum \ln\frac{R_s}{R_{so}} + c\ln\frac{T_o}{T} \quad (2.2)$$

where a, b and c are operator selectable weighting factors. R_{so} is optimal resolution, as assigned by the experimenter and R_s is the computationally observed resolution between any two neighboring peaks. The first term relates only to peaks which are excessively separated, the second term relates only to peaks which are not completely resolved, and the third term relates to total time, T, and optimal total time, T_o, for the separation.

The experiments that are simulated by the simplex require the input of several parameters: column efficiency, applied field, optimal resolution, the operator selectable

weighting factors a, b and c, and optimal time, as well as the number of bp for each fragment in the sample. An illustration of the initial input parameters for the simplex optimization program is depicted in Table 2.1. Capillary efficiency is expressed as total plate count, N, and is given by:

$$N = \frac{l}{H} \quad (2.3)$$

where l is the effective length of the capillary and H is the observed plate height. While capillary efficiencies can vary slightly from day to day, typical plate heights are 0.3 - 0.4 μm and do not depend greatly upon [MC] or other adjustable parameters. Applied field, E, is given by:

$$E = \frac{V}{L} \quad (2.4)$$

where V is applied voltage and L is total column length. The optimal resolution parameter (R_{so}) used in this work is 1.5. The values for the operator selectable weighting factors and the optimal total time, as determined prior to the optimization, as well as the number of fragments and corresponding size in bp number are also entered.

Upon entering all necessary parameters, the program asks for the first of three initial sets of conditions ([MC] and capillary length), and then generates values for a simulated separation. For each set of bp fragments in the sample, mobility, retention time, base width and resolution are generated (see Results and Discussion). Retention time (T_R) is calculated from mobility (Equation 2.1), applied field (E) and effective length (l) and is given by:

Table 2.1. Illustration of initial input parameters for simplex optimization program.

N	column efficiency	3×10^6 N/m
E	applied field	400 V/cm
R_{so}	optimal resolution	1.5
a	excess resolution factor	5
b	overlap degradation factor	50
c	time factor	6
T_o	optimal time	200 s
bp #	number of fragments and sizes	1 (72), 2 (118), 3, (194), 4 (234), 5 (271), 6 (281), 7 (310), 8 (603), 9 (872), 10 (1078), 11 (1353)

$$T_R = \frac{l}{(\mu E)} \quad (2.5)$$

Base width of peaks in time units (W_b), resolution (R_s) and total time (T_{total}) are calculated from classical chromatographic definitions and equations:

$$W_b = 4\frac{T_R}{\sqrt{N}} = 4\frac{T_R}{\sqrt{l/H}} \quad (2.6)$$

$$R_s = \frac{\left[T_{R(n)} - T_{R(n-1)}\right]}{\overline{W_b}} \quad (2.7)$$

$$T_{total} = T_{R(last)} + 0.5W_{b(last)} \quad (2.8)$$

where $T_{R(n)}$ is the retention time of fragment n, $T_{R(n-1)}$ is the retention time of fragment $n-1$, $T_{R(last)}$ is the retention time of the last fragment and $W_{b(last)}$ is the base width of the last fragment.

These values are then used to generate a *CRF* value (using Equation 2.2) for the separation. Once this is complete, calculated data are displayed and the program will ask for the second and third sets of conditions ([MC] and capillary length) and will produce the same types of calculations respectively. At this point the program has the initial set of experiments and generates (using simplex optimization) what the next experiment will use for [MC] and capillary length. The *CRF* value, as a determination of separation quality, represents an example of the response points seen in Figure 2.4. If the predictions are acceptable, i.e., within experimental boundaries, the program generates a

new set of calculations based on these parameters and generates a new *CRF*. The program can also be instructed to run in an automatic mode where boundary conditions are pre-set and a maximum number of simplex iterations is specified. An illustration of the input/output for four iterations of the optimization program is shown in Table 2.2. This process continues until an optimum has been reached. It should be noted that the *CRF* and weighting factors (discussed below) have no effect on the simulation data, but only upon the process by which the simplex method searches for optimal conditions. An example of a complete 50 iteration simplex is shown in the Appendix.

RESULTS AND DISCUSSION

The objective of optimization in any type of chromatography is simply to find that set of conditions required to achieve the best separation for a given situation. This ultimately requires that the experimenter be familiar with the separation at hand, as well as knowledgeable of the usefulness of the results. In the size-selective capillary electrophoretic (SSCE) separation of a DNA restriction enzyme digest, for instance, the resolution of the two most difficult to resolve fragments may be the primary concern. Consequently, any attempt to implement separation conditions that provide more than the needed information may prove to be costly and time consuming. On the other hand, one may be interested in identification of all the fragments in a DNA sample. In this case, separation conditions that provide adequate resolution for identification is sufficient. It is inherent that the usefulness of the results be known prior to performing an experiment.

Table 2.2. Illustration of input/output for four iterations of the simplex optimization program.

Experiment 1
Capillary length: 40 cm
Methyl Cellulose concentration: 0.6%

peak	(BP)	(μ)	T_R (sec)	W_b	R_s
1	72	2.84×10^{-4}	352.7	1.29	0
2	118	2.30×10^{-4}	434.5	1.59	56.91
3	194	1.97×10^{-4}	506.5	1.85	41.93
4	234	1.89×10^{-4}	529.8	1.93	12.32
5	271	1.83×10^{-4}	546.5	1.99	8.48
6	281	1.82×10^{-4}	550.4	2.01	1.94
7	310	1.78×10^{-4}	560.5	2.05	4.99
8	603	1.63×10^{-4}	613.6	2.24	24.79
9	872	1.58×10^{-4}	633.3	2.31	8.61
10	1078	1.56×10^{-4}	642.0	2.34	3.77
11	1353	1.54×10^{-4}	649.8	2.37	3.29

CRF = -99.71 Total time = 650.99 s

Experiment 2
Capillary length: 35 cm
Methyl Cellulose concentration: 0.5%
CRF = -93.67 Total time = 536.96 s

Experiment 3
Capillary length: 42 cm
Methyl Cellulose concentration: 0.4%
CRF = -97.18 Total time = 601.92 s

Experiment 4
Capillary length: 37 cm
Methyl Cellulose concentration: 0.3%
CRF = -90.73 Total time = 488.92

* Calculations for μ, T_R, W_b and R_s are shown for experiment 1 only
* Experiments 1, 2 and 3 represent the initial three experiments required by the simplex
* Experiment 4 is an example of results of the first optimization step

Therefore, the ultimate goal of the program is to have it generate the ideal capillary length and [MC] for any intended separation.

The *CRF* serves both as a measure of success for simulated separations and as a means to influence the simplex for specific separation situations. *CRF* values can vary from negative to positive, with more positive values being desirable. The operator selectable weighting factors *a*, *b* and *c* all play an important role in each specific separation by placing weight on their respective terms in the *CRF* as described below. Although no limits are imposed on the values of *a*, *b* and *c* by the program, values in the range of 5 to 50 have been used in these studies.

The excess resolution factor, *a*, weights those peaks which are excessively resolved. Because only peak pairs which are excessively resolved are involved, R_{so}/R_s will always be less than one and the first term will always be negative. The effect of excess resolution is to make the first term small, thus contributing a negative value to the *CRF*. Because there are more easy-to-resolve peak pairs than difficult to resolve pairs, a relatively small value of 5 has been selected for *a* in this work.

The second term in the *CRF* relates only to peaks which are not adequately resolved. The overlap degradation factor, *b*, is the operator selectable weighting factor which can determine how much influence this term has on the value for the *CRF*. Incomplete resolution of peaks makes the second term small and contributes to a negative *CRF*. Because unresolved peaks degrade experimental results extensively, higher values for *b* are usually better. It is generally preferable to have excess resolution as opposed to inadequate resolution; however, the ability to adjust *a* and *b* allow us to determine the

weight each will have, as decided by the needs of the separation. A arbitrary value of 50 has been selected for b in this work.

The third term in the *CRF* relates to total time and optimal total time. The time factor, c, is the weight given to times shorter or longer than the optimal total time, T_o. Times shorter than T_o produce a positive contribution from the third term, while times longer that T_o produce a negative contribution to the *CRF*. A nominal value of 5 has been selected for c in this work. Optimal times should be selected realistically, with 30 seconds/fragment being a reasonable guideline.

Experimental Verification of Computational Methods

Once modeling software was developed, the degree of agreement with laboratory experiments was assessed. Figure 2.5 depicts the comparison between a previously reported separation (from our laboratory) [30] and a computer simulated separation of the ΦX-174 DNA *Hae*III digest. As seen in Figure 2.5 (A), an effective column length of 40 cm and a [MC] of 0.5 % results in the separation of all 11 fragments in under 15 minutes using a field of 300 V/cm in a 45 mM TBE running buffer. The peaks in order of elution are 72, 118, 194, 234, 271, 281, 310, 603, 872, 1078 and 1353 bp. Figure 2.5 (B) depicts the computer simulation (SIM-program) of the same separation under nearly identical conditions. The computer generated simulation shows a resolution of 1.8 for the 271/281 fragments. As seen, the degree of agreement is quite good (271/281 resolution is 2.2 in Figure 2.5 (A)) with some discrepancy in the total time for the analysis. The simulated separation time is about 7 % shorter. The analysis time difference may be attributed to

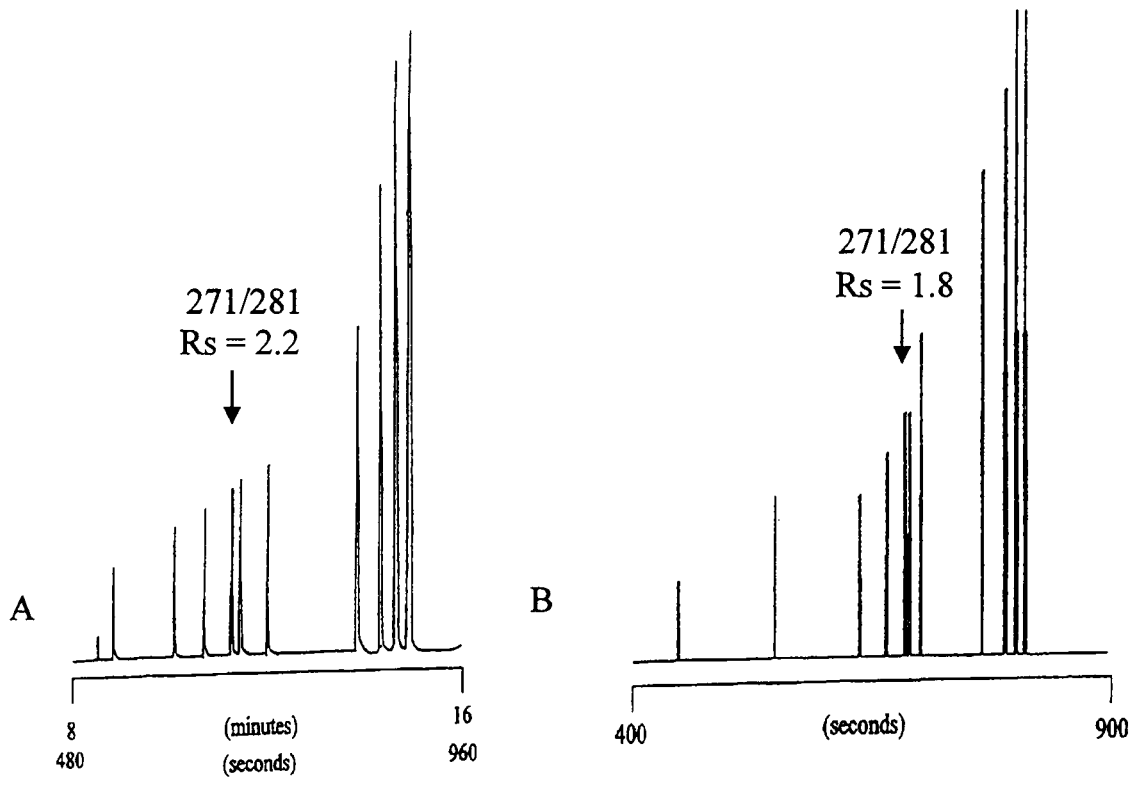


Figure 2.5. Comparison between a previously reported separation (A) and a computer simulated separation (B) of the ΦX-174 DNA *Hae*III digest.

the fact that the separation was performed at a field of 300 V/cm while the mobility equation (Equation 2.1), used for the simulated separation, was generated using data obtained at a field of 400 V/cm.

A visual representation of the *CRF*, plotted as a function of percent MC and capillary length, is depicted in Figure 2.6. This 3-D matrix provides a visual means to determine optimal conditions or optimal ranges of conditions. The selected boundaries of the matrix program are 0.05 % to 1.0 % MC and 5 to 50 cm capillary length. It can be seen, from the matrix (point labeled 2a), that at an approximate capillary length of 10 cm and [MC] of approximately 0.3 % yields the best *CRF* value.

Figure 2.7 depicts the comparison between simulated and experimental separations of the ΦX-174 DNA *Hae*III digest performed under optimized conditions (10.7 cm and 0.3 % MC), as generated by the SIM-program and as visualized in Figure 2.6. As seen in Figure 2.7 (A), the simulated optimized separation yields a separation in under 200 seconds. Resolution of the 271/281 bp fragments in the non-optimized simulated separation in Figure 2.5 (B) is excessive (R_s = 1.8), while the resolution of the same fragments in the optimized simulated separation (Figure 2.7 (A)) is 0.9 and the total separation time is four fold less. The optimized simulated separation is also consistent with an actual separation performed under similar conditions, as seen in Figure 2.7 (B), where the R_s for the 271/281 bp fragments is 1.0. Figures 2.5 - 2.7 exemplify the ability of the program to not only generate a simulated separation, but to also produce an optimized set of conditions, which accurately predict experimental results.

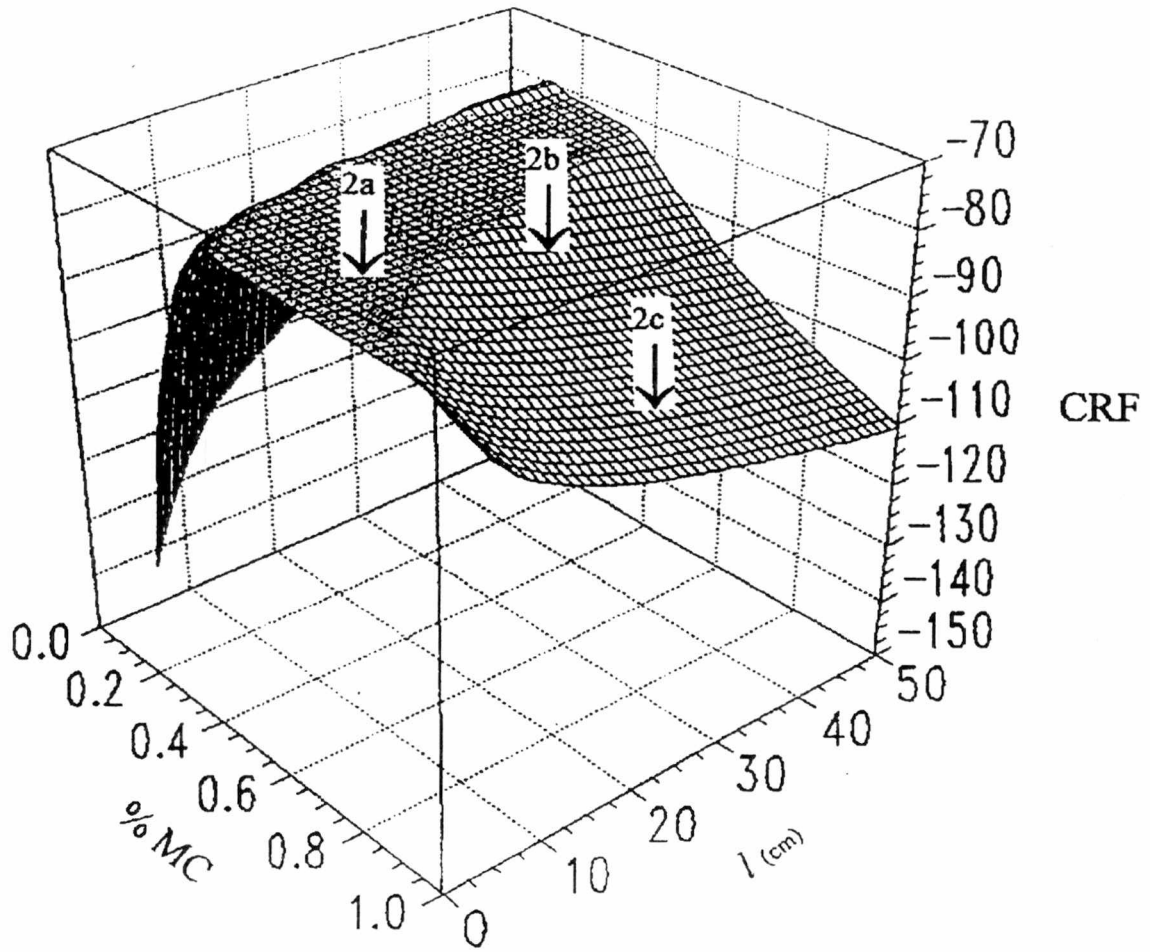


Figure 2.6. Visual representation of the *CRF*, plotted as a function of percent MC and capillary length. 2a, 2b and 2c represent best, average and poor *CRF* values, respectively, as described in the text.

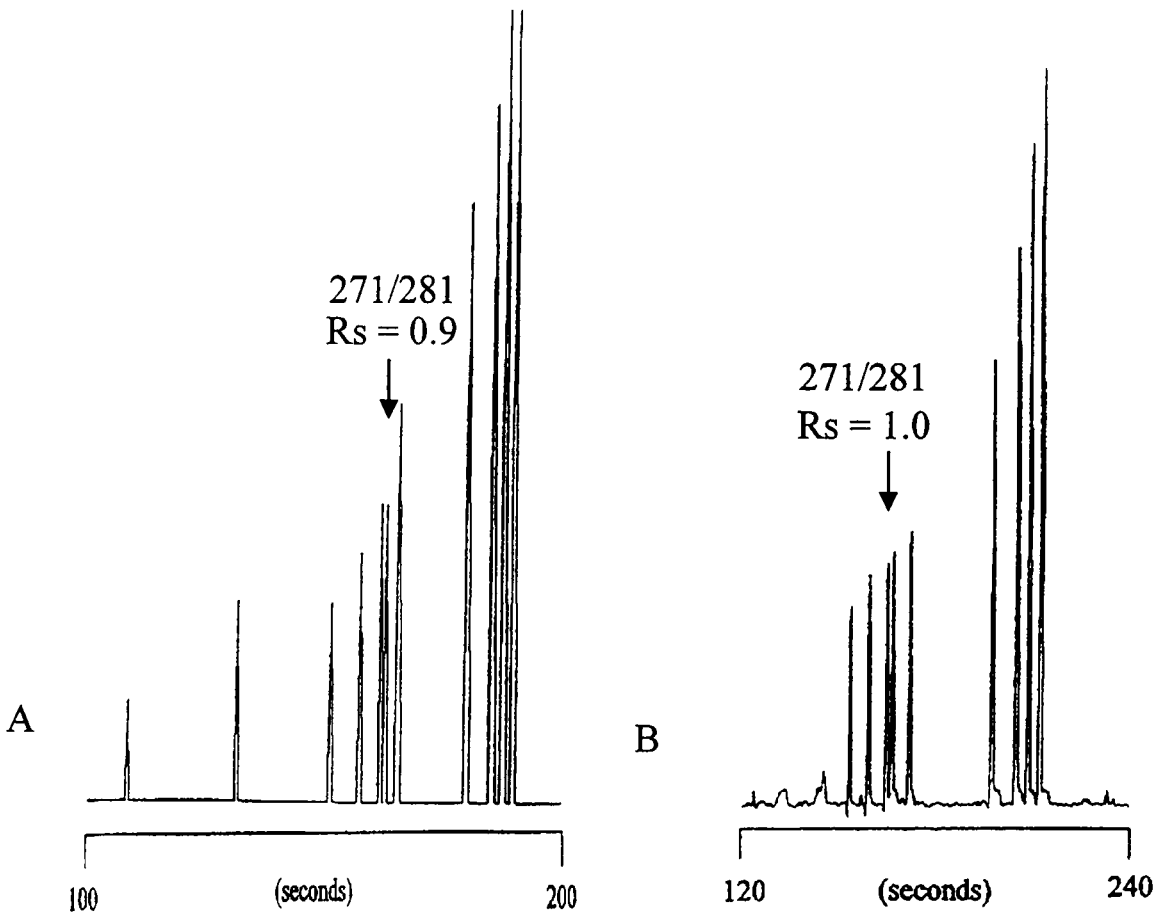


Figure 2.7. Comparison between simulated (A) and experimental (B) separations of the ΦX-174 DNA *Hae*III digest performed under optimized conditions (10.7 cm and 0.3 % MC), as generated by the SIM-program.

Figure 2.8 shows that there is agreement between points on the 3-D matrix (see Figure 2.6) and their corresponding laboratory separation. Figure 2.8 (A) represents an experimental verification of the best *CRF* value on the matrix, 0.3 % MC, l = 10 cm, as chosen by the optimization process. This value is visually represented by the point labeled 2a on Figure 2.6. The separation is performed in under 200 seconds (conserving analysis time) while providing adequate resolution of the 271/281 fragments (R_s = 1.0). Figure 2.8 (B) represents an experimental verification of an average position (*CRF* value) along the matrix, 0.3 % MC, l = 28 cm. This value is visually represented by the point labeled 2b on Figure 2.6. Total analysis time is increased approximately three fold and resolution is slightly in excess of the optimum value (R_{so} = 1.5) for the 271/281 fragments (R_s = 1.7). Figure 2.8 (C) represents the experimental verification of a poor position (*CRF* value) along the matrix, 0.6 % MC, l = 28 cm. This value is visually represented by the point labeled 2c on Figure 2.6. While resolution of the 271/281 fragments has increased to greater than 5, total analysis time is considerably longer. While the information acquired by performing optimization studies through traditional experimentation can in its own right prove beneficial, it is usually at the expense of valuable time. The use of computational methods to optimize for and then visualize possible separations prevents futile expenditure of experimental time on the generation of potentially fruitless data. Figures 2.5 - 2.8 show evidence of the utility of such a program and its ability to accurately predict future experiments.

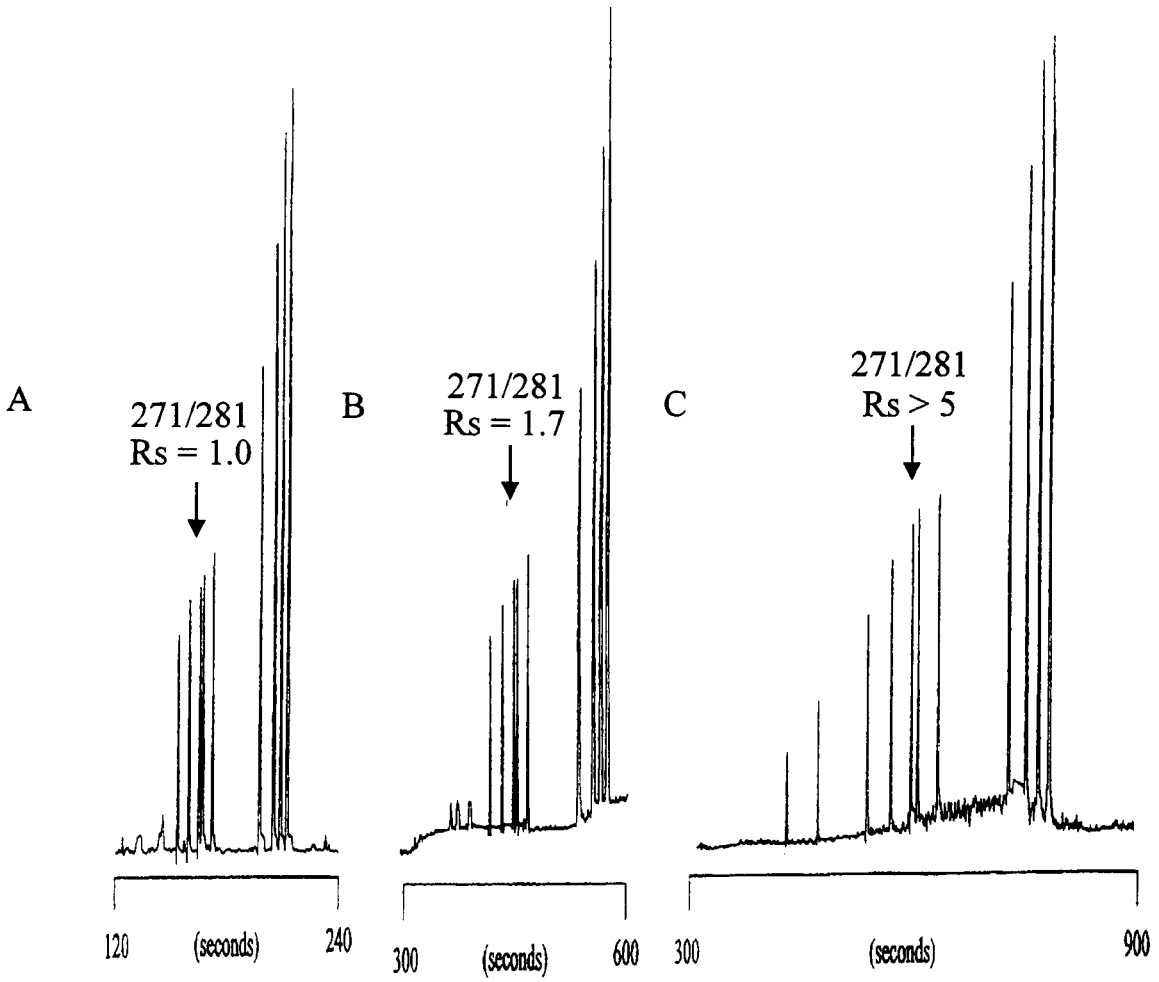


Figure 2.8. Agreement with points on the 3-D matrix. (A) represents verification of the best *CRF* value (0.3 % MC, l = 10 cm), (B) an average *CRF* value (0.3 % MC, l = 28 cm) and (C) a poor *CRF* value (0.6 % MC, l = 28 cm).

Utility of Computational Methods to Predict Separation Behavior

These methods can also be utilized to simulate and optimize separations of samples that were not considered in the initial data set (i.e. fragments other than those from the ΦX-174 DNA *Hae*III digest). This can be seen with the SSCE separation of the pBR322 *Hinf*I digest shown in Figure 2.9. Figure 2.9 (A) is a previously reported separation of pBR322 *Hinf*I digest fragments from our laboratory [42]. The fragments in order of elution are 154, (220, 221), 298, 344, 396, (506, 517) and 1631 bp. As can be seen 7 'sets' of fragments can be resolved in under 320 seconds utilizing a field of 500 V/cm, with the 220/221 and 506/517 fragments co-eluting and unresolved. An effective column length of 40 cm and 0.5 % MC was used in this experiment. Expansion of the computer generated simulation for this separation, Figure 2.9 (B), also shows that the 220/221 and 506/517 fragment pairs are not resolved (see insert, Figure 2.9 (B)). The computer generated simulation of this separation shows a slightly longer analysis time. Again, this can be attributed to the fact that a field of 400 V/cm was employed for the mobility equation training set used to generate the simulated separation, as opposed to the 500 V/cm used in this run.

It is important to point out that not only is good agreement of total analysis time achievable, but the ability to predict separation performance for DNA samples that were not initially considered is also possible. This is due to the "base pair is a base pair" behavior alluded to earlier. The utility of this program to determine the conditions that will result in resolution of the 220/221 and 506/517 fragment pairs is illustrated in the following figures. Figure 2.10 (A) depicts the 3-D matrix for optimization of the

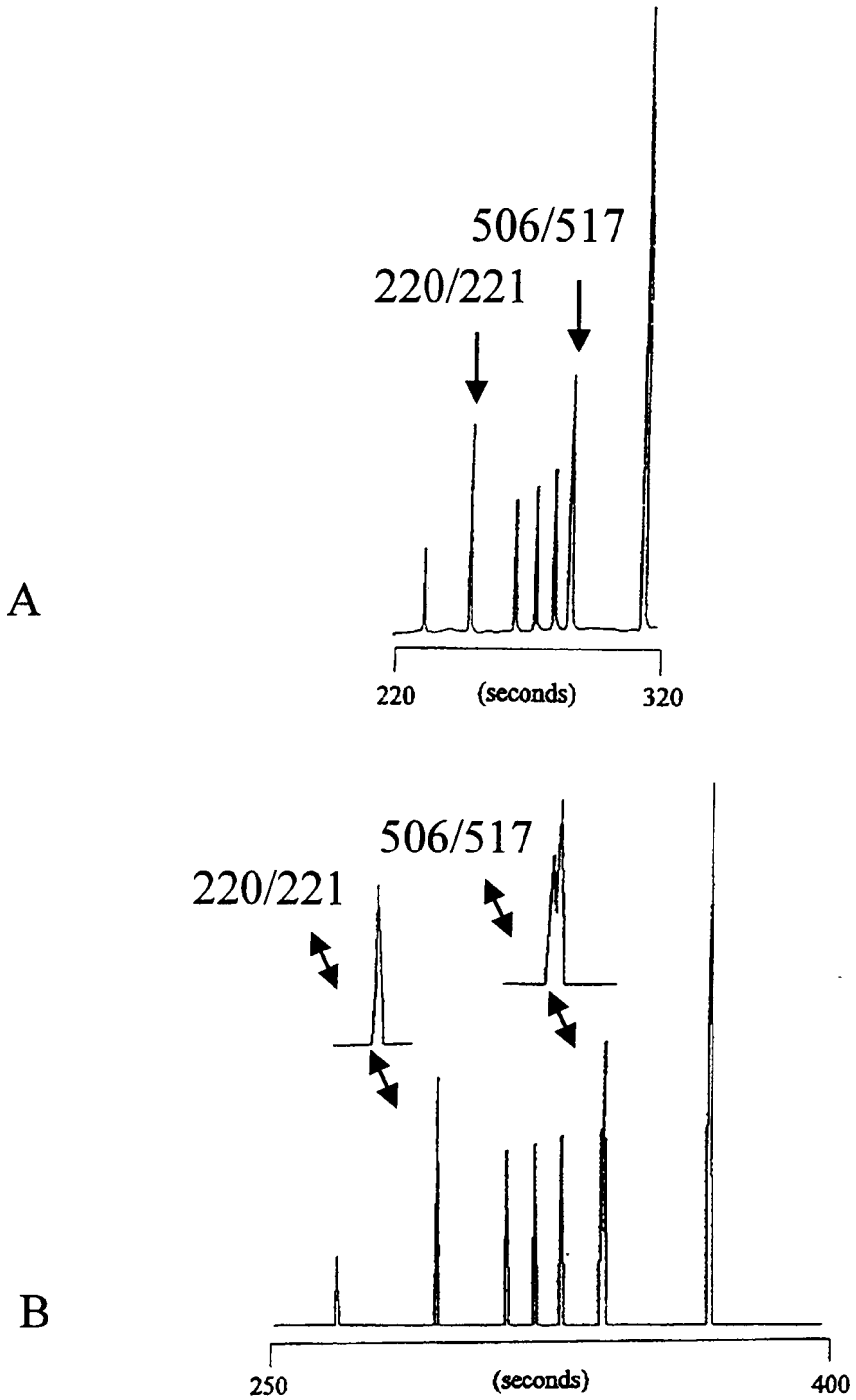


Figure 2.9. Comparison between a previously reported separation (A) and a computer simulated separation (B) of the pBR322 DNA *Hinf*I digest.

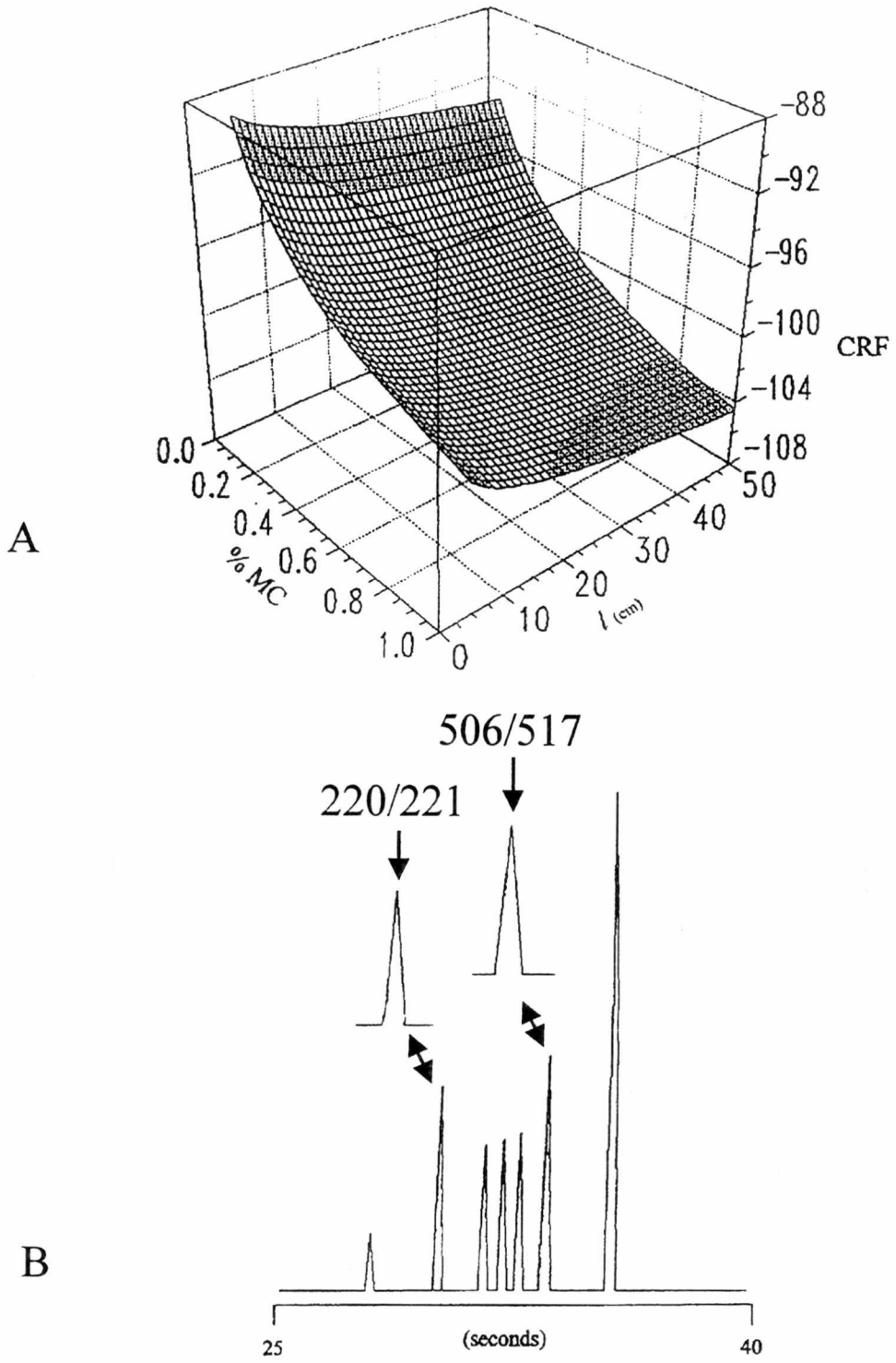


Figure 2.10. Adjustment of computational parameters to influence separation quality (1). (A) depicts the 3-D matrix for optimization using a *b* value of 15. (B) depicts a computer generated simulated separation using optimum conditions, 0.05 % MC and l = 5 cm, from (A).

separation using an overlap degradation factor, *b*, of 15. Optimum conditions are shown for short capillaries and low [MC]. Because the excess resolution term, *a*, is dominant and conditions that yield fast separation times are emphasized, a relatively low value of *b* does not afford sufficient emphasis for resolving the 220/221 and 506/517 fragment pairs. Figure 2.10 (B) depicts the simulated separation using optimum conditions from the 3-D matrix in Figure 2.10 (A), albeit conditions that are outside of our experimental boundaries (0.05 % MC, l = 5 cm). Because the optimum conditions emphasize fast separation times, the resolution of the 506/517 fragments is worse than in Figure 2.9 (B).

Because the 220/221 and 506/517 fragment pairs are extremely difficult to resolve, a *b* value of 15 is too small to place any emphasis on their resolution. Increasing the influence *b* has on the *CRF* is seen in Figure 2.11. When *b* is increased from 15 to 50 it becomes evident that the *CRF* is becoming more positive at higher [MC] and longer capillary lengths, drastically changing the response surface, as seen in Figure 2.11 (A). However, even at a [MC] of 1.0 % and a capillary length of 50 cm (optimum conditions from Figure 2.11 (A)), the 220/221 and 506/517 fragment pairs are still not adequately resolved, as seen in Figure 2.11 (B). Experimental verification of computationally predicted optimum conditions with *b* values of 15 and 50 are beyond presently attainable experimental limits. Current instrumental constraints permit detection only within the 7-33 cm window of the column. Additionally [MC] percentages of 0.05 % and 1.0 % fall outside of those concentrations considered in the mobility equation training set. A

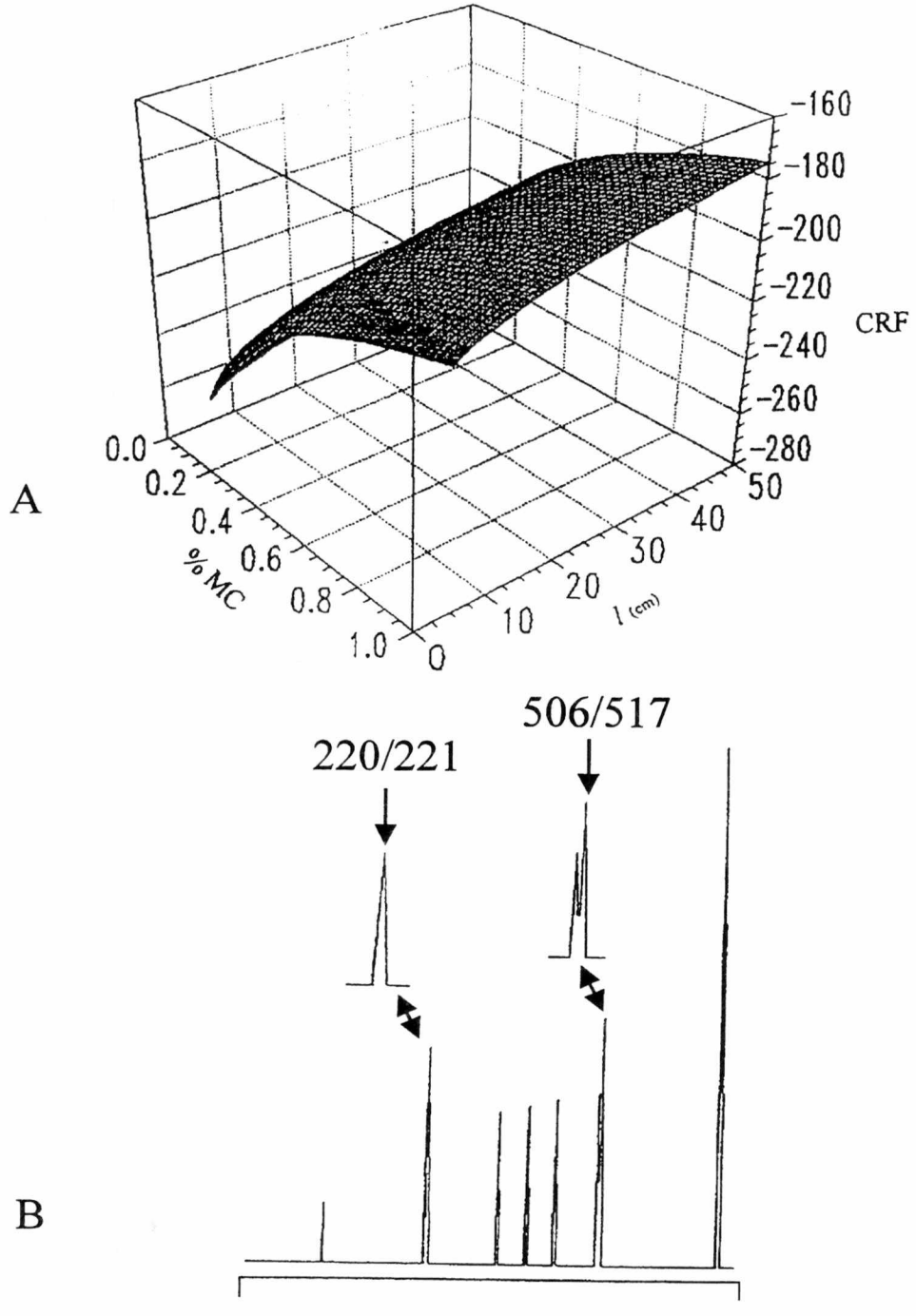


Figure 2.11. Adjustment of computational parameters to influence separation quality (2). (A) 3-D matrix for optimization using a b value of 50 and (B) computer generated simulated separation using optimum conditions, 1.0 % MC and l = 50 cm, from (A).

simulation (not shown) calculates that a capillary length of 100 cm and a [MC] of 2.0 % will produce resolution of 1.49 for the 506/517 fragment pairs, but only 0.6 for the 220/221 fragment pairs. Obviously, there is no way to know if Equation 2.1 is valid at such large [MC]. Nevertheless the function of the SIM- and MAT-programs in predicting conditions that solve challenging separation problems is clearly illustrated.

Depending on the mix of DNA fragments in a sample, the *a*, *b* and *c* terms may be adjusted in optimizing conditions. The SIM-program contains complete simulation data for each set of experimental conditions, providing a guide for adjustment of weighting terms for additional optimizations. For example, if resolution is inadequate, the *b* term may be increased and another optimization performed, as shown above. If resolution is adequate but time is excessive the *c* term may be increased, or optimal total time decreased. Reduction of the excess resolution factor, *a*, increases the contribution of unresolved peaks and forces the *CRF* toward higher [MC]. Reducing the overlap degradation factor, *b*, allows the excess resolution and time factors to produce the best *CRF* values at shorter capillary lengths and lower [MC]. On the other hand, increasing *b* tends to increase [MC]. Reducing the time factor, *c*, has the effect of producing higher *CRF* values at higher [MC] but with little change in the overall shape of the *CRF* response surface. Of the effects described, those of lowering the excess resolution factor and increasing the overlap degradation factor are the most dramatic.

Although the effects of changing *a*, *b* and *c* are readily apparent on the value of *CRF* from MAT plots and provide a guide to selecting optimal conditions, the actual effects on separations can be visualized only by the SIM-program simulations in which

resolution of the bp of interest may be examined in detail. Some differences in times between simulated and experimental chromatograms can perhaps be reduced by further refinement of the mobility equation.

While mobilities are largely independent of base pair sequence, observed differences in retention times suggest the need to accommodate the different migration mechanisms in a more generalized mobility expression. Expansion of the initial data set to generate mobility data that encompasses a greater range of bp fragment sizes will ultimately make the program more universal. Nevertheless, the simplex optimization described herein provides a means for the selection of specific conditions for a desired separation in which optimization may be based on any desired combination of resolution behavior and separation time.

CHAPTER 3

Design and Optimization of a Capillary Electrophoretic Mobility Shift Assay for the Study of Protein-DNA Interactions

INTRODUCTION

Analytical methodologies involving complex biomolecules are continuing to show widespread success throughout both the physical and life sciences. The analysis of these complex substances pervades such disciplines as chemistry, biochemistry, molecular biology, genetics and others. The emergence of new methodologies almost always arises from the need to improve one or more variables in traditional techniques. CE has had its share of success as a viable technique for the analysis of DNA and other large biopolymers (e.g., proteins). In fact, CE is quickly finding its way into many of the life science laboratories as both a routine analytical tool and as an attractive instrumental replacement for more traditional, yet cumbersome, techniques. CE is poised to replace slab-gel electrophoresis for automated DNA sequencing [59-68].

One of the most promising approaches that has arisen from this work involves the analysis of antisense DNAs which could result in a more thorough understanding of the suppression of virus replication and oncogene expression [69-71]. The screening of genetic diseases and forensic analysis have benefited tremendously from CE separations of PCR amplified products [72-77]. CE has also played a role in replacing denaturing gradient gel electrophoresis for the detection of point mutations [46, 78]. Analytical

methodologies involving other biomolecules are just as prevalent. The use of CE for on-line protein digestion, peptide derivatization and separation has been explored for protein characterization [41, 79]. CE has also been employed as a powerful complement to reverse phase HPLC (RP-HPLC) for peptide mapping [80-83]. However, possibly the most exciting area of CE separations is its interface with mass spectrometry (MS). The information available from the combination of the high resolving power of CE and the mass information of MS stands to make this an extremely powerful analytical tool. Matrix-assisted laser desorption ionization (MALDI) circumvents some of the incompatibility problems plaguing CE interfaces with MS and has recently been applied to the analysis of proteins [84-86]. With the success that CE has enjoyed in the analysis of DNA and proteins it is evident that the technology can be implemented for the study of protein-DNA interactions.

Studies of protein-DNA interactions permeate most areas of biochemical research. The model of genetic regulation, proposed by Jacob and Monod, for protein synthesis by DNA-binding repressor proteins is now over thirty years old [87]. Modern research has revealed that the majority of genetic regulation results from the actions of DNA-binding repressors and activators. The study of the interaction between nucleic acids and regulatory proteins is essential to the understanding of genetic regulation at the molecular level. The binding of a particular protein to a specific region of a DNA molecule results in the activation or repression of a variety of genetic signals. Currently the established method for these studies is the electrophoretic mobility shift assay (EMSA), also known as gel retardation assays [88, 89]. This widely used technique can be employed to study

complex formation and stability [88, 90, 91], define binding sites on nucleic acids [90, 92, 93], analyze DNA conformations in complexes [90, 94, 95] and study dynamic processes involving transcription and splicing [90, 96].

The principles of the mobility shift assay are illustrated in Figure 3.1. Upon loading of a solution containing a mixture of protein-DNA complexes and unbound molecules (Figure 3.1 (A)), free DNA will enter the gel as a physically discrete band when subjected to electrophoresis (Figure 3.1 (B)). This is based on the principle that the mobility of specific protein-DNA complexes can differ substantially from that of unbound DNA [88, 89, 97]. DNA-bound proteins are pulled into the gel by the negatively charged DNA, albeit at a mobility that is retarded relative to that of the unbound DNA. Since the free DNA band is not affected by any re-equilibration, i.e. dissociation/re-association of the protein-DNA complexes, the assay provides an accurate value for the quantity of uncomplexed DNA. This is based on the fact that no appreciable dissociation/re-association takes place prior to the electrophoresis event. Knowledge of the concentration of input DNA and the concentration of uncomplexed DNA yields the amount of DNA that exists in the complexed form.

However, this method suffers from some significant limitations, most importantly, long assay time, difficulty with accurate quantitation and difficulty in determining dissociation constants [97]. Capillary electrophoresis is well-suited as a replacement for such a traditional method. The current research, termed capillary electrophoretic mobility shift assay (CEMSA), will incorporate the advantages of both EMSA and CE, while

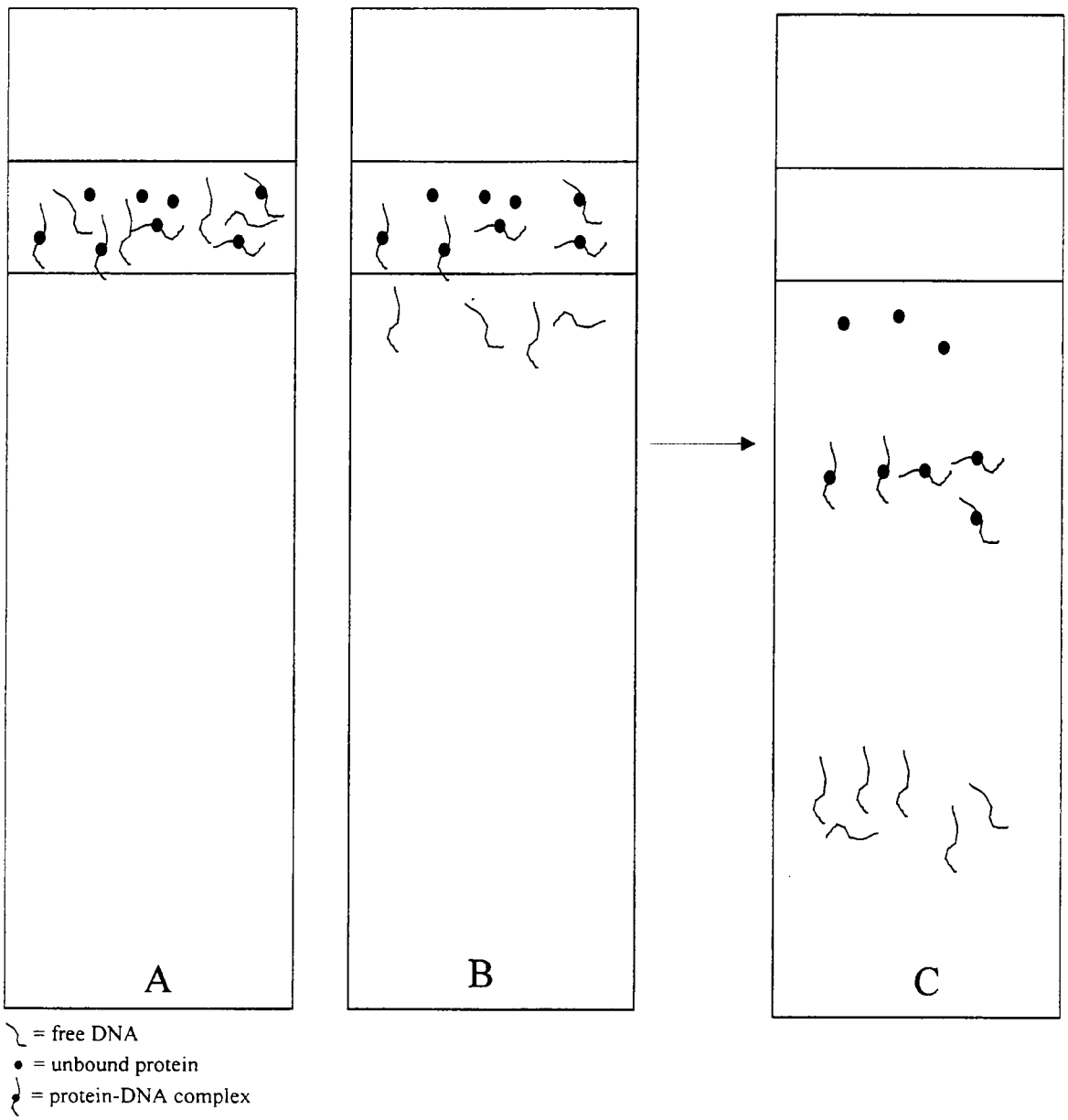


Figure 3.1. Gel-based mobility shift assay.

overcoming some of the major limitations of the former. This look at CEMSA provides both a quantitative and qualitative basis for the development of the technique.

Inherent to the development of a novel instrumental based assay is the need to characterize it with a well known system. The interaction between the *trp* repressor (protein) of *Escherichia coli* and the *trp* operator (DNA) serves as a well-characterized system for the development of CEMSA. The *trp* repressor protein controls transciption initiation in the *trp*, *arch* and *trpR* operons [98]. Transcription is the process by which a class of enzymes constructs an RNA copy of one strand of a DNA sequence. This RNA then directs the synthesis of cellular proteins (translation). An operon is a cluster of genes, an operator and its structural genes, whose transcription is blocked by repressor proteins and activated by stimulatory proteins. Simply stated, the binding of repressor to operator prevents transcription of the structural genes that are controlled by the operator.

Repression is the result of competition between the repressor and RNA polymerase for overlapping binding sites in the promoter/operator region of the regulated operons. DNA binding in the *trp* repressor/*trp* operator system is regulated by allosteric control and is described below. The *trp* repressor is responsible for the regulation of L-tryptophan biosynthesis in *E. coli*. This is accomplished via a simple negative feedback loop [99]. When the concentration of L-tryptophan is low, *trp* repressor is inactive (aporepressor) and the *trp* operon is expressed, resulting in the biosynthesis of L-tryptophan. As depicted in Figure 3.2 (A), the inactive repressor is not confomationally competent for high affinity binding with the operon. A relatively high concentration of L-tryptophan leads to formation of the active repressor, which in binding the operator,

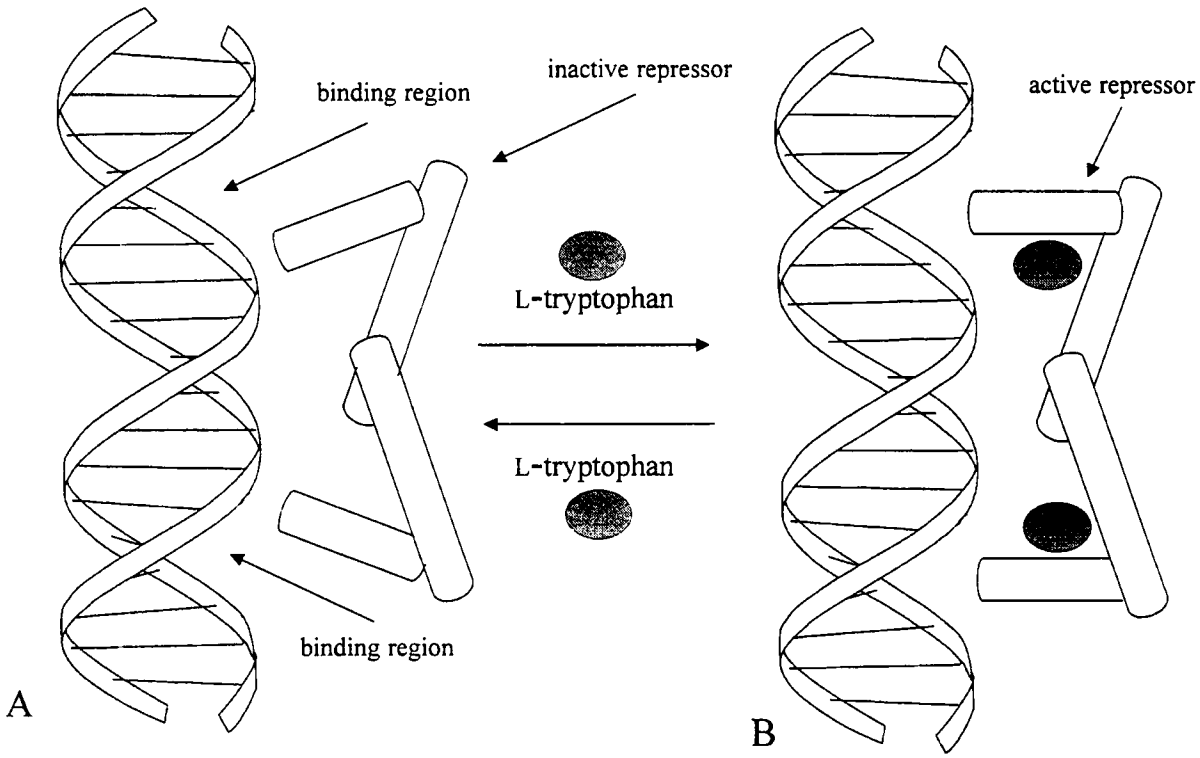


Figure 3.2. Interaction between the *trp* repressor and the *trp* operator. (A) the inactive repressor is not conformationally competent for high affinity binding with the operator, (B) the binding of L-tryptophan induces a conformational change that alters the orientation of the recognition helices so as to fit into the two successive major groves of the operator DNA.

inhibits transcription initiation in the *trp* operon and thus represses the production of L-tryptophan. The binding of L-tryptophan induces a conformational change in the *trp* repressor that alters the orientation of the recognition helices so as to fit into the two successive major grooves of the operator DNA, as seen in Figure 3.2 (B) [99]. The binding of *trpR* to DNA, once thought to be the result of direct hydrogen bonds between the protein's side chains and the bases in the major grove of the DNA, involves water molecules mediating hydrogen bonding contacts between the protein's recognition helices and the bases that contribute most to the identity of the operator [91].

In this dissertation, an investigation of protein-DNA interactions by CE with laser fluorometric detection is performed and combines the rapid and minimal sample consumption methods of CE with the selective separation influence of mobility shift assays. The interaction between the *trp* repressor and the *trp* operator is used to explore the capabilities of CEMSA. The use of fluorescently tagged operator not only lends itself to laser induced fluorescence detection but also avoids the use of radiolabeled detection. In CEMSA, as in traditional EMSAs, the mobility of a specific protein-DNA complex differs from that of free DNA. Evaluation of running buffer conditions (composition and pH) allows optimization of complex formation and resolution of free DNA from complex. Quantitative studies show reaction of repressor with operator results in the diminishing of free operator signal and the simultaneous creation of the repressor-operator peak that is well resolved from the free operator. Qualitative studies involving nonspecific interactions between operator and a complex protein sample are

demonstrated. It is shown that the specificity of operator for repressor can be used to selectively separate the repressor from a complex sample that includes non-specific protein. These studies demonstrate the technique's ability to selectivity differentiate between bioaffinity specific and nonspecific interactions.

EXPERIMENTAL

Materials

In addition to the materials mentioned in Chapter 2, myoglobin, lysozyme and L-tryptophan were also obtained from Sigma. Fused silica capillaries (50 μm i.d. x 365 μm o.d.) were coated with linear polyacrylamide using the method outlined in Chapter 2.

Running Buffer and Sample Preparation

Running buffers containing methyl cellulose (MC) were prepared in the manner described in Chapter 2. Purified *trp* repressor was prepared as described in Czernik *et al.* [93] and stored at -20°C in small aliquots at concentrations greater than 1 mg mL^{-1} in 10 mM sodium phosphate pH 7.6, 150 mM NaCl and 0.1 mM EDTA. Purified E12 protein was prepared as described in Makeli *et al.* [100]. The concentrations of *trp* repressor (60 μM) and E12 (12.3 μM) were determined at 280 nm using absorptivities of 13,940 cm^{-1} M^{-1} and 17,900 cm^{-1} M^{-1}, respectively. The symmetrical 26 nucleotide DNA was synthesized incorporating a fluorescein conjugated base at the 3′ end using a Perseptive Biosystems (Framingham, MA, USA) Expedite synthesizer by the Molecular Biology Core facility at the University of Arkansas for Medical Sciences (U.A.M.S.). The

oligonucleotide was annealed by heating to 100 °C for 5 minutes and slowly cooling to room temperature over a period of 1 hour. After self-annealing, the duplex sequence is as follows:

5′ - GAGCGTACTAGTTAACTAGTACGCTC - Fl
Fl - CTCGCATGATCAATTGATCATGCGAG - 5′

where Fl designates a fluorescein molecule. The underlined portions represent the base pairs (bp) that have been shown to be most important for repressor binding in genetic and biochemical studies. DNA concentrations were based on spectral measurements using a Hewlett-Packard (Palo Alto, CA, USA) Model 8452A Diode Array Spectrophotometer and the conversion of 1.0 absorbance units (abs) at 260 nm = 50 $\mu g\ mL^{-1}$. Working solutions of DNA were diluted to the desired concentrations with 45 mM TBE, 0.5 mM L-tryptophan. Purified 3205 bp plasmid DNA and *Hinf*I restriction enzyme samples were provide by Dr. Barry Hurlburt at U.A.M.S.

Preparation of *trpR*-DNA complexes involved the addition of the appropriate amounts of each analyte (depending on desired ratio) and diluting to the desired volume with running buffer. A 1:1 *trpR*-DNA complex represents an approximate concentration of $2.6x10^{-6}$ M for each analyte. All other ratios are based on this concentration. Myoglobin and lysozyme stock solutions were prepared by weighing out the appropriate amount of protein and diluting to the desired concentration with running buffer. Protein-DNA samples involving myoglobin, lysozyme and E12 were prepared in the same fashion as the *trpR*-DNA complexes.

Apparatus and Electrophoresis

The electrophoresis and fluorescence apparatus was assembled in-house. A Spectra-Physics (Mountain View, CA, USA) Model 162A air cooled argon ion (Ar^+) laser, operating at 488 nm, was employed as the excitation source. Rejection of plasma lines emanating from the laser was accomplished by placing a 488 nm (10 ± 2 nm FWHM) laser line filter in the path of the laser output. A 25 mm diameter, *f*/1, quartz lens was used to focus the laser radiation through the capillary. The collection optics consisted of a 600 μm (core) diameter fiber optic which was positioned within 200 μm of the capillary at the point through which the beam was focused. The other end of the fiber was carefully placed against a narrow bandpass filter (20 nm bandpass centered at 520 nm) to isolate the fluorescence emission. The filter was placed against a shutter assembly on the face of the photodetector, a Hamamatsu (Bridgewater, NJ, USA) model HC120-01 PMT assembly. The PMT power supply was assembled in-house, as were a noise filter and offset circuit for the signal output [101]. Electropherograms were recorded using a strip chart recorder. Figure 3.3 represents a schematic diagram of the CE laser fluorometry apparatus used in these studies.

A Hipotronics (Brewster, NY, USA) Model 840A high voltage power supply was used to apply electrophoretic fields. Injections were performed electrokinetically (for running buffers incorporating MC) by placing the inlet of the capillary into the sample vial and applying -2 kV for 20 seconds. Following injection, the capillary inlet was replaced into the running buffer vial and a running potential of -20 kV was applied across the 50 cm (40 cm to the detector) capillary. Hydrostatic injections (for non-MC

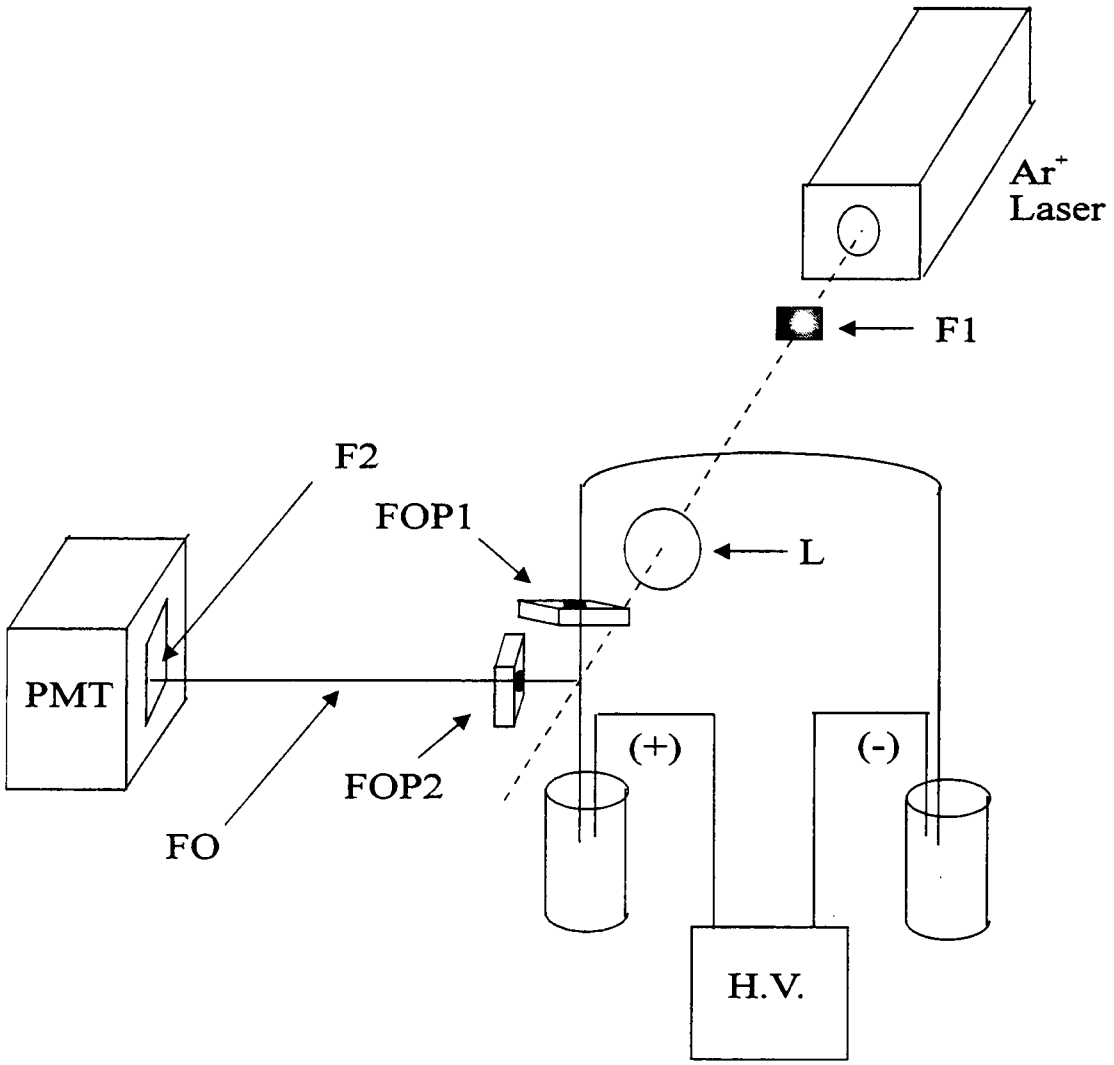


Figure 3.3. Laser fluorometry apparatus for capillary electrophoretic mobility shift assay. F1: 488 nm laser line filter; L1: focusing lens; FOP1 & 2: fiber optic positioners; FO: collection fiber optic; F2: 520 nm narrow bandpass filter; PMT: photomultiplier tube.

containing running buffers) were performed by placing the inlet of the capillary in the sample vial and raising it 10 cm for 10 seconds.

RESULTS AND DISCUSSION

Factors Influencing Complex Formation

In an attempt to develop a capillary-based assay, evidence of complex formation was of primary concern. Two requirements present themselves when dealing with the creation of protein-DNA complexes. The first requirement is the resolution of free DNA and the protein-DNA complex. The second requirement is that the protein-DNA complex remains intact throughout the course of the separation, at least to some degree. As mentioned previously, in the gel-based assay, the amount of free DNA is quantified resulting in the calculation of the amount of DNA present in the complexed form. In both cases, however, it is paramount that the free DNA signal is reduced in response to protein addition to be sure that a complex is being formed.

Figure 3.4 depicts experimental evidence of the formation of a complex that remains intact, at least somewhat, during the separation. An electropherogram depicting free DNA, at an approximate concentration of 2.6×10^{-6} M, is shown in Figure 3.4 (A). The more intense peak, at approximately 6 minutes, represents the "active" portion of the DNA sample that contains the bp region known for high affinity binding of the *trp* repressor. The additional peaks, surrounding the peak of interest, are similarly sized fragments that do not contain full length binding regions and are a result of using non-gel purified DNA. Figure 3.4 (B) depicts evidence of a complex peak upon addition of an

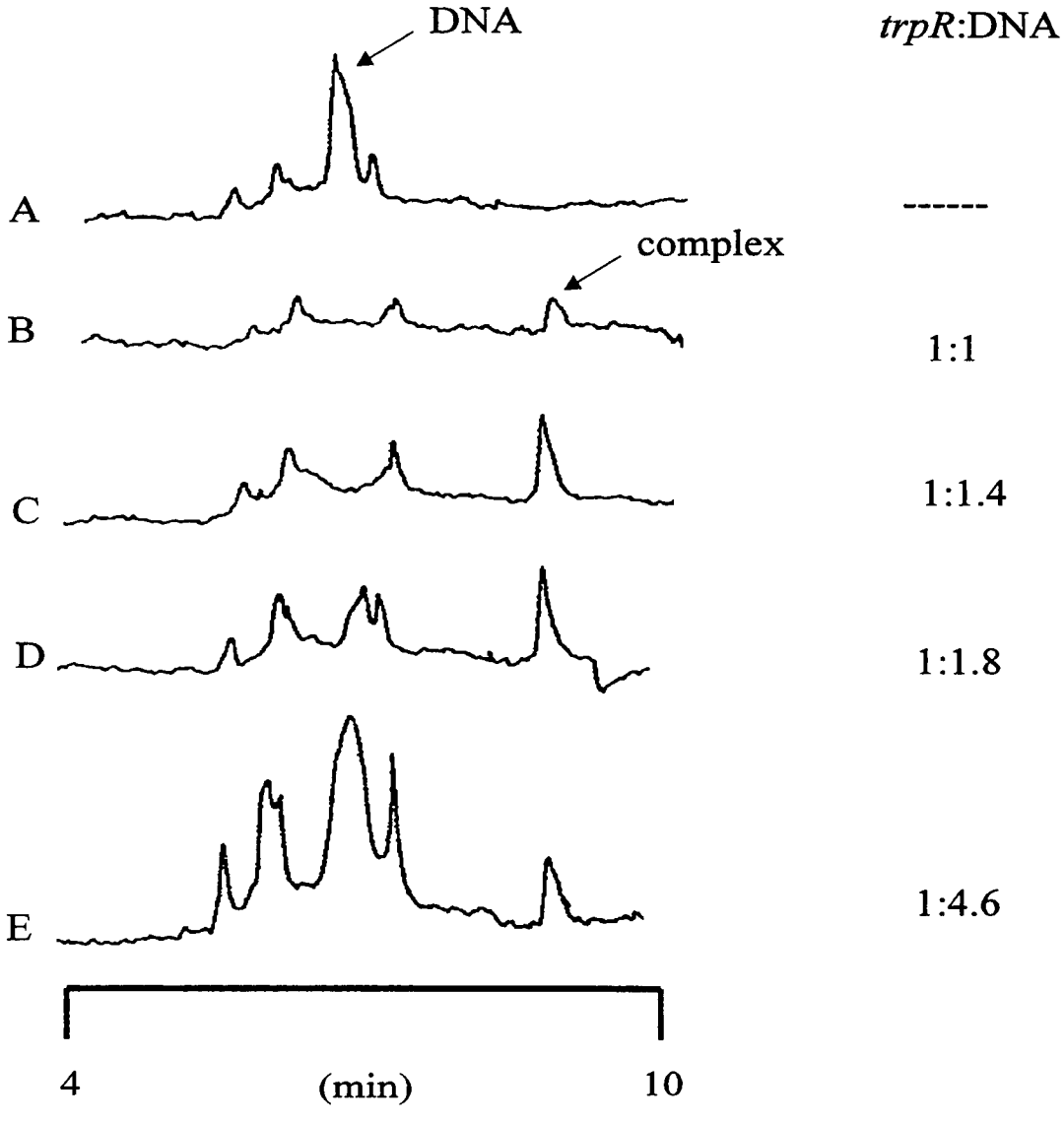


Figure 3.4. Experimental evidence of *trp*R:DNA complex formation. (A) free DNA, (B) 1:1 *trp*R:DNA complex, (C-E) addition of DNA to complex beyond 1:1.

equimolar amount of *trp* repressor (1:1, *trpR*:DNA). Binding of the *trp* repressor to the DNA is evidenced by the loss of the "active" DNA peak. Subsequently, the created complex remains intact, at least somewhat, as evidenced by the creation of a complex peak that is well resolved from the free DNA signal. Both Figures 3.4 (A) and (B) show cursory evidence of the two requirements described earlier: 1) resolution of free DNA from the protein-DNA complex and 2) complex remaining intact throughout the separation.

It is shown that as more DNA is added to a 1:1 (*trpR*:DNA) complex (Figure 3.4 (C-E)), the complex peak grows until the *trp* repressor exists only in the complexed form. At this point the addition of more DNA results in an increase in the free DNA signal only. At repressor concentrations greater than 100-fold above the dissociation constant (K_D = 0.2 - 0.5 nM), the binding is 1:1 [102, 103]. The fact that the complex peak grows upon addition of DNA above a 1:1 amount may suggest that the complex dissociates, somewhat, under electrophoretic conditions, or simply that the concentration of active, full length strands of DNA is relatively low in the non-gel purified DNA sample. Poor labeling efficiency between fluorophore and DNA can play a significant role in the concentration of these active DNA strands. Since detection is afforded by excitation of the fluorescently labeled DNA, those DNA strands that are not labeled and that bind protein would not be detected, i.e., both labeled and unlabeled DNA are competing for protein. A more efficient labeling protocol would facilitate working at lower DNA and repressor concentrations thus realizing less of the influences of competing DNA. It can be assumed that wall adsorption is negligible due to coating of the capillary walls with

linear polyacrylamide. Figure 3.4 further shows that resolution of complex and free DNA exists. At this point attention was turned to developing running buffer conditions that maximized efficiency and resolution of free DNA from complex.

For site-specific binding of *trp* repressor to the operator DNA, a saturating amount of L-tryptophan is required [91]. As recalled from Figure 3.2 and the previous discussion, the *trp* repressor binds two molecules of L-tryptophan to become competent for high affinity operator binding [104, 105]. The binding of L-tryptophan induces a conformational change allowing the *trp* repressor's recognition helices to fit properly into the two successive major grooves of the DNA that represent the bp recognition regions.

Figure 3.5 shows the dependence of complex formation on L-tryptophan. An electropherogram of free DNA in the absence of repressor, and without L-tryptophan in the running buffer (45 mM TBE, pH 8) is depicted in Figure 3.5 (A). Upon incubation of DNA with a stoichiometic amount of repressor (Figure 3.5 (B)), and no L-tryptophan in the running buffer, a complex is apparently created (evidenced by diminishing of the free DNA peak), but is not maintained; i.e., it appears to dissociate since no complex peak is observed. The reduction in the free DNA signal can be attributed to non-specific binding by the low affinity (not conformationally competent) repressor to the DNA. The complex that is briefly formed appears quickly to dissociate and becomes part of the tailing edge of the free DNA peak; thus no complex peak is observed. An electropherogram of free DNA in the absence of repressor, with L-tryptophan in the running buffer (45 mM TBE, 0.5 mM L-tryptophan, pH 8) is depicted in Figure 3.5 (C). Upon incubation of DNA with a stoichiometic amount of repressor (Figure 3.5 (D)), the presence of L-tryptophan

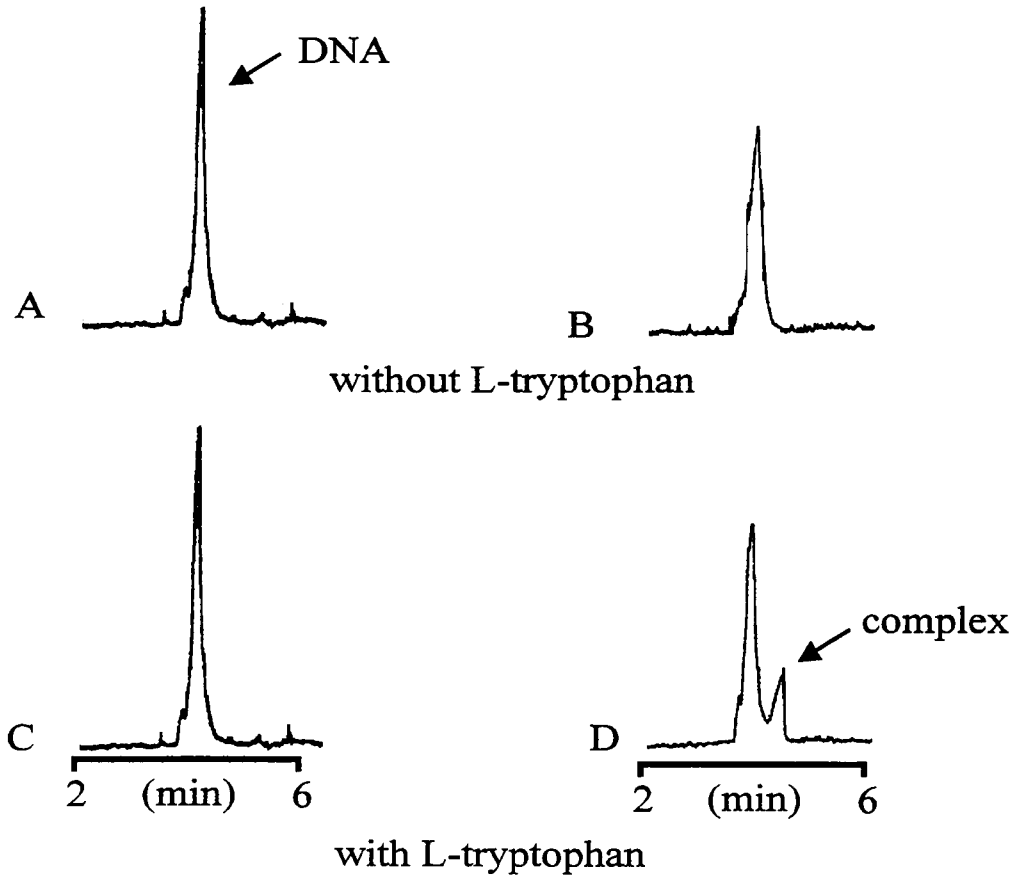


Figure 3.5. Dependence of complex formation on L-tryptophan. (A) and (B) free DNA and a 1:1 *trp*R:DNA complex, respectively without L-tryptophan in the running buffer. (C) and (D) free DNA and a 1:1 *trp*R:DNA complex, respectively with L-tryptophan (0.5 mM) in the running buffer.

induces the conformational change that is required for high affinity binding of repressor to the DNA. A complex is created (evidenced by diminishing of the free DNA peak) and maintained (evidenced by the presence of a complex peak).

Once complex formation and resolution were observed, a cursory study of the effects of running buffer constituents on the performance of the capillary-based assay was conducted. TBE has been used as the buffer in traditional gel-based assays of *trp* repressor binding to DNA and was thus chosen as a starting point for this capillary-based work [91]. It has been shown that the pH of a gel-based assay for *trp* repressor plays an important role in the ability of *trp* repressor to retard the electrophoretic mobility of the DNA fragment. As the pH of the buffer approaches the isoelectric point (pI) of *trp* repressor (5.9), retardation is enhanced [91]. In the capillary-based assay, this was not observed. A comparison of pH 5, 6, 8, and 10 buffers (45 mM TBE, 0.5 mM L-tryptophan) resulted in no discernible difference in mobilities, i.e., resolution of the free DNA relative to the complex; however, an increase in signal was observed as pH increased. This can be attributed in part to the pH-sensitive fluorescence of fluorescein-based dyes. The intensity of fluorescein emission increases upon increasing pH, while exhibiting a significant reduction in emission below pH 7 [106]. Since pH 8 buffer resulted in adequate signal for free DNA and optimal signal and reproducibility for the complex, it was chosen as the pH for use in further studies.

It has been demonstrated that running buffer additives play an important role in the optimization of protein separations. Various protein separations and enzyme micro assays, studied by CE, have employed such running buffer additives as polyethylene

glycol, dextran, cross-linked and linear polyacrylamide gels [66, 107-110]. In each case, the additives served as polymer networks imparting sieving. The addition of soluble polymers to running buffers for the analysis of DNA restriction fragments serves the same size selective sieving purpose [42].

Moreover, in gel-based assays the cross-linked acrylamide matrix imparts a caging effect [90]. This serves to effectively raise the local concentration of protein and DNA. Therefore, when a protein-DNA complex dissociates under electrophoretic conditions, they may re-associate rapidly to reform the complex. In this capillary-based work, the linear polyacrylamide is chemically bonded to the column wall and thus serves only to inhibit wall adsorption of analytes. Therefore a similar caging phenomenon may be desired for the capillary-based assay. It should be noted that because of the speed of these CEMSA separations, relative to gel-based assays, this may not be needed.

Since the presence of MC imparts size selectivity in DNA restriction digest assays, the effect of MC in the running buffer was investigated. The influence of MC on the protein-DNA interaction also served an additional purpose. Future studies involving the analysis of protein-DNA interactions in the presence of other DNA fragments that do not contain a specific binding region would require the separation of the DNA fragments prior to interaction with the binding protein. Preliminary studies aimed at determining protein binding sites on relatively large DNA fragments is presented later in this chapter.

The molecular weight and percent of MC used in these studies was not expected to yield a mesh that would be small enough to impart size selection between the free DNA and the complex. However, the use of MC conditions has been shown to influence

efficiency [42]. The addition of MC not only improved resolution, but it also led to an increase in efficiency for both the free DNA and the complex as seen in Figure 3.6. What appear to be multiple peaks from the free DNA response can again be attributed to the fact that non-gel purified DNA was used. Figure 3.6 (A) and (B) represent free DNA and an approximate 1:1 (*trpR*:DNA) complex, respectively, in the absence of MC as a sieving agent. Running buffer conditions are 45 mM TBE, 0.5 mM L-tryptophan, pH 8. Upon addition of 0.75 % MC to the running buffer, Figure 3.6 (C) and (D), both an increase in efficiency and resolution are observed for the free DNA and the complex. The increase in efficiency afforded to the complex peak may be attributed to the same phenomenon seen in gel-based assays. The MC is a non-rigid soluble polymer network that may impart the caging effect alluded to above. Not only does it serve to keep any dissociated protein and DNA in proximity, it may also slow down the diffusion of any wall adsorbing species. In effect, it acts not only as a sieving medium, resulting in increased resolution, but it also may help prevent wall adsorption. The final result is increased resolution and efficiency.

The observation that MC improved efficiency led to the optimization of MC conditions. Three different methyl cellulose percentages (0.25, 0.5 and 0.75 %) were compared as to their influence on signal, resolution and efficiency. All running buffers consisted of 45 mM TBE, 0.5 mM L-tryptophan, pH 8 and the soluble polymer. Running buffers containing 0.25 % MC yielded broad, highly inefficient responses for free DNA, as well as decreased resolution and efficiency of the complex peak as compared to 0.75 % MC. This can be attributed, in part, to the fact that a 0.25 % MC solution is below the

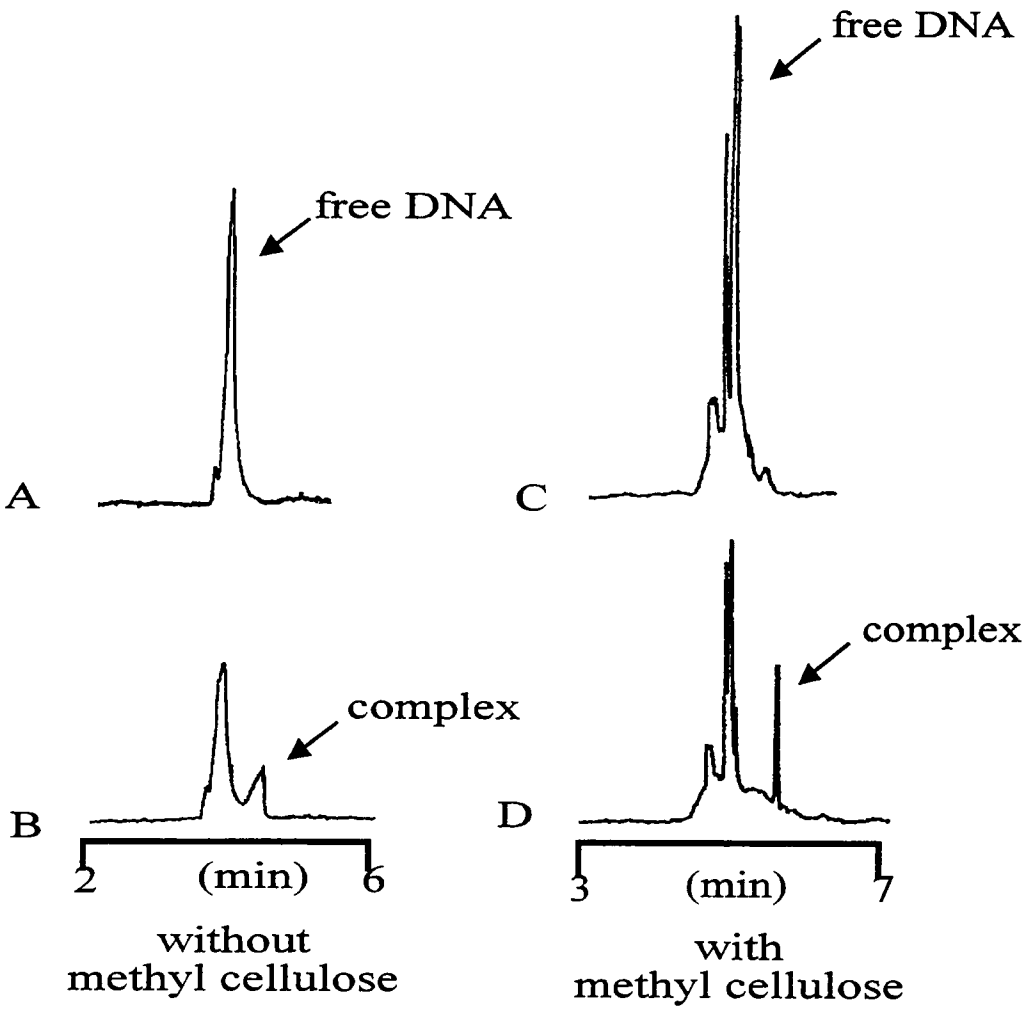


Figure 3.6. Effect of methyl cellulose on efficiency and resolution of free DNA and complex. (A) and (B) represent free DNA and a 1:1 *trp*R:DNA complex, respectively in the absence of methyl cellulose. (C) and (D) represent free DNA and a 1:1 *trp*R:DNA complex, respectively with 0.75 % methyl cellulose in the running buffer.

entanglement threshold (3.5 %) for MC and therefore does not contain the entangled polymer network responsible for effective sieving. In such a situation the caging effect necessary to keep dissociated protein and DNA in proximity is unlikely. A comparison between 0.5 % MC and 0.75 % MC resulted in slightly greater efficiency for free DNA from the 0.5 % MC buffer, but significantly greater efficiency for the complex peak from the 0.75 % MC buffer. While the effects of soluble polymer percentages and size (MW) have been investigated for DNA fragments [42], there are no reports that they have been studied for DNA-protein separations.

Quantitative and Qualitative Studies

Upon optimization of running buffer conditions, simple quantitative studies were performed to show the changes in peak height as a function of increasing repressor concentration. In this work, high concentrations of operator DNA were required to offset poor fluorescent labeling efficiency and hence poor detectability. At concentrations greater than 100 fold above the K_D, binding of repressor to operator is stoichiometric [102, 103]. This allows easy manipulation of the amounts of either analyte in solution. Figure 3.7 shows a plot depicting change in peak height as a function of increasing [*trpR*]. Each data point represents a ratio of *trpR*:DNA, beginning with a ratio of 0:1 and increasing in increments of 0.25 units. An approximate ratio (1:1, where [DNA] $\cong$ 3.2×10^{-6} M) represents an equal molar amount of each analyte, since the stoichiometry is 1:1. In the plot, both free DNA peak height and complex peak height are monitored as a function of increasing [*trpR*]. As the concentration of *trp* repressor approaches that of the

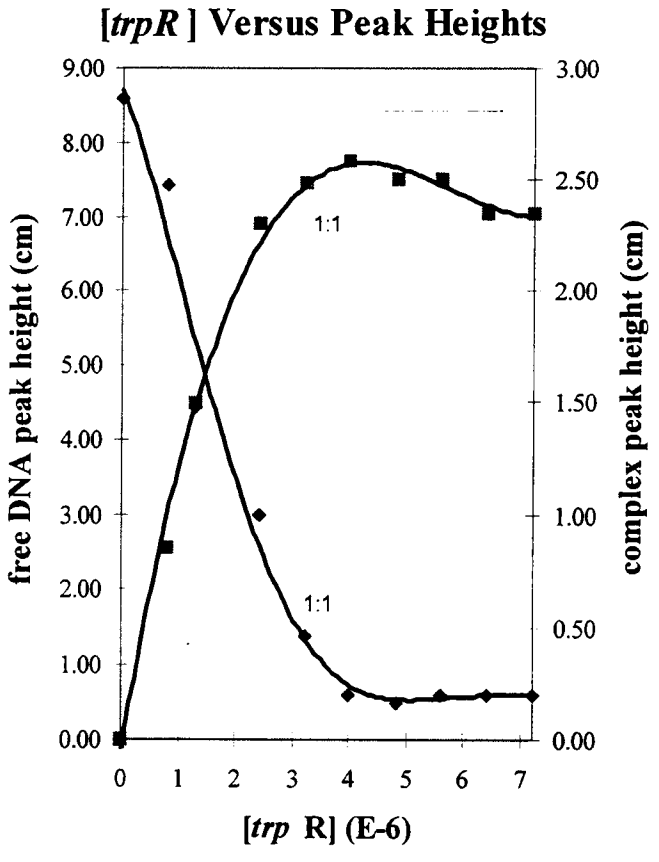


Figure 3.7. Change in free DNA and complex peak height as a function of [*trpR*].

DNA in solution, a subsequent decrease in the free DNA peak height is observed. This expected behavior can be attributed to the consumption of free DNA by the increasing amounts of *trp* repressor. Up to, and slightly beyond, the stoichiometric ratio of 1:1, this effect is relatively linear. This slight deviation in linearity can be attributed to the low concentration of active DNA alluded to earlier. Once the ratio approaches and exceeds 1:1, as [*trpR*] > [DNA], a leveling effect is seen in the peak height of the free DNA. This can be attributed to the fact that all of the DNA present in the reaction mixture is complexed with the repressor. Once equal molar amounts react, no more complex can ideally be formed, partly due to the large association constant and the high concentrations employed, yielding no subsequent decrease in the free DNA signal as more repressor is added. This same effect is also seen in the plot of complex peak height as a function of increasing [*trpR*]. As the concentration of *trp* repressor approaches that of the DNA, a subsequent increase in the complex peak height is observed. Again, up to and slightly beyond the stoichiometric ratio of 1:1, this effect is relatively linear. As the ratio approaches and exceeds 1:1, a leveling effect is seen in the peak height of the complex.

The *trp* repressor/*trp* operator interaction is one of high specificity. Not only must the correct bp sequence be present on the DNA, but L-tryptophan must also be present to allow the repressor to properly bind. Simple qualitative studies, involving both non-DNA-binding and DNA-binding proteins, can demonstrate the high selectivity of the operator for the repressor. An example of *trp* repressor/*trp* operator selectivity involving non-DNA binding proteins is depicted in Figure 3.8. Figure 3.8 (A) represents an electropherogram of free DNA in 45 mM TBE, 0.5 mM L-tryptophan, 0.75 % MC, pH 8.

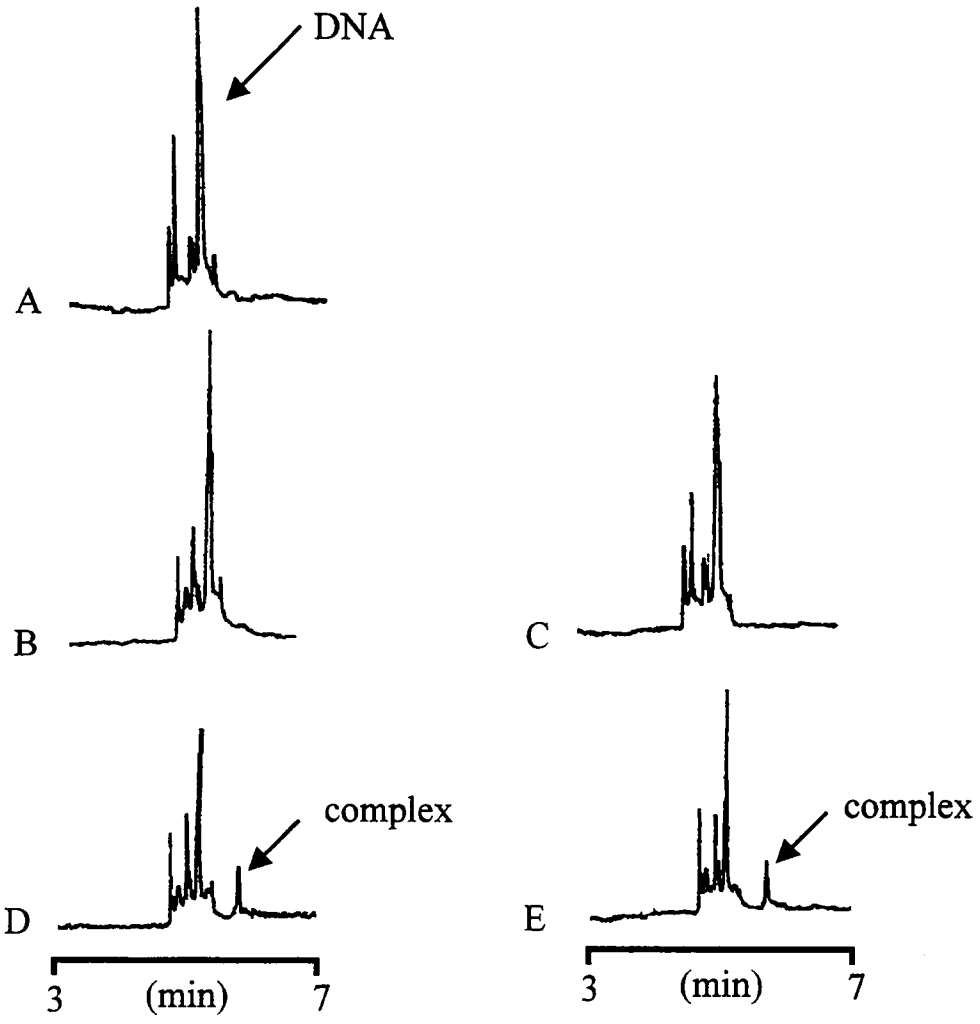


Figure 3.8. Effect of non-DNA binding proteins on the selectivity of the *trp* repressor/*trp* operator interaction. (A) free DNA, (B) and (C) 1:1 ratio of DNA with myoglobin and lysozyme, respectively. (D) and (E) equal molar amounts of *trp* repressor added to both (B) and (C).

Upon reaction of the DNA with an equal molar amount of either myoglobin or lysozyme, no complex peak is observed, as seen in Figure 3.8 (B) and (C) respectively. This is anticipated due to the fact that neither protein recognizes the specific nucleotide sequences on the operator as binding sites. Myoglobin and lysozyme, while representing interfering species, cause relatively no change in the free DNA signal. However, upon addition of *trp* repressor, in equal molar amounts with the operator, complex creation is observed, even in the presence of the non-specific proteins, as seen in Figure 3.8 (D) and (E). This demonstrates the selectivity of the DNA for the repressor. A greater degree of selectivity can be seen when the interfering analyte is also a DNA-binding protein.

E12 is a basic-helix-loop-helix (bHLH) protein known to bind the enhancers of immunoglobulin genes [111]. Together with other bHLH proteins, E12 is involved in mammalian muscle development [111-113]. While E12 does not represent a protein species that specifically binds to the region on the *trp* operator, it is known to bind DNA. Figure 3.9 depicts the effect of DNA-binding protein on the selectivity of the *trp* repressor/*trp* operator interaction. Upon reaction of the DNA with an equal molar amount of E12, an almost complete diminishing of the free DNA signal is observed, Figure 3.9 (A); however, no complex is observed. Non-specific interactions are probably responsible for this apparent binding to DNA. The binding constant for E12 is large enough to observe non-specific interactions but is not large enough to observe a complex peak. The high concentration of both DNA and E12 may also play a role in the presence of non-specific interactions. More efficiently labeled DNA would allow the use of lower concentrations thus lessening the effects of any non-specific binding. The addition of an

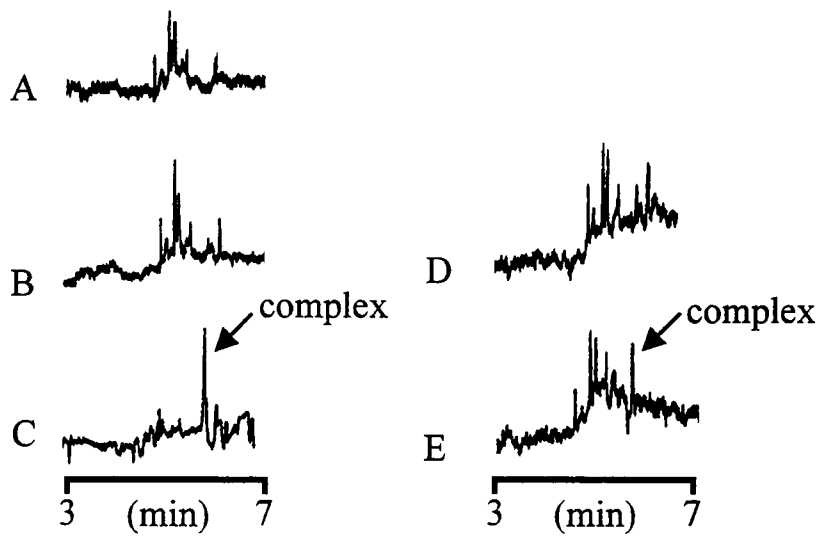


Figure 3.9. Effect of the DNA binding protein E12 on the selectivity of the *trp* repressor/*trp* operator interaction. (A) 1:1 E12:DNA complex. (B) and (C) addition of equimolar amount of *trp* repressor to (A) immediately and after 24 hours, respectively. (D) and (E) addition of equimolar amount of *trp* repressor to DNA, E12, lysozyme and myoglobin complex immediately and after 24 hours, respectively.

equal molar amount of *trp* repressor yields no complex peak as seen in Figure 3.9 (B). However, over time (24 hours), the dominant and highly specific interaction of the *trp* repressor for its operator is evident as depicted in Figure 3.9 (C). This can be attributed to the fact that the large binding constant for *trp* repressor to DNA eventually overcomes the smaller binding constant of E12. Figure 3.9 (D) and (E) also depicts a similar situation when all three interfering species (myoglobin, lysozyme and E12), are added to the DNA, followed by addition of repressor. The same highly specific behavior is apparent. Running buffer conditions for Figure 3.9 included 45 mM TBE, 0.5 mM L-tryptophan, 0.75 % MC, pH 8

Further improvements to the assay could result in the determination of a variety of parameters inherent to protein-DNA interactions. A more efficient fluorescent labeling protocol would result in a greater number of DNA molecules that could be detected thus allowing work at lower concentrations of analyte. This combined with better DNA preparation methods would insure that each DNA molecule is of the proper length and fluorescently labeled. Increased detectability can also be realized through the refinement of the current fluorescent detection apparatus. Refinements such as lens-based collection of fluorescence, spatial rejection of unwanted specular scatter, use of extended light path capillaries and use of sheath flow cell devices can aid in increased detectability by improving the signal to noise ratio (S/N). An additional alternative involves the use of velocity programming for increased detection zone residence. This entails the significant decrease of the separation field as an analyte band approaches the detection zone resulting

in an increase in the measurement time of the analyte. The advantage afforded by this approach is an increase in S/N.

The ability to work at lower analyte concentrations, realized by increased detectability, would facilitate the quantitative investigation of the nature of the these interactions such as equilibrium association and dissociation rate constants. The speed of capillary-based assays should also allow the determination of kinetic information, generally not feasible due to the "deadtime" of gel-based assays. As refinements are employed CEMSA should continue to make strides towards being a true competitor for the traditional gel-based assay. The applicability of CEMSA to the determination of protein binding sites on DNA is currently being investigated and is briefly outlined in the following section. This work demonstrates methodologies for the study of protein-DNA interactions by CE. Through the combination of the selective separation influence of conventional mobility shift assays and the rapid and minimal sample consumption attributes of CE, CEMSA exploits the advantages of both. The use of polyacrylamide coated columns suppressed DNA and protein adsorption to the column, while sufficiently eliminating electroosmotic flow for resolution of DNA from the complex. Assay times are generally under 6 minutes and, in common with most CE systems, the ability to rapidly adjust system parameters (running buffer constituents, pH, applied voltage, etc.) is present. With capillary columns in hand, an entire series of experiments can literally be run without interruption. This is highly favorable when compared to the long turn-around times of traditional gel-based assays. The *trp* repressor/*trp* operator interaction, well

characterized by other assay procedures, served as an appropriate system for the development of CEMSA.

Future Studies

The determination of protein binding sites on DNAs of interest is a time consuming and labor intensive process for the biochemical researcher. Present and future research in our laboratory is centered around evaluation of CEMSA as an analytical tool for determination of these binding sites. The initial strategy involves subjecting a DNA plasmid containing the region known to bind *trpR* to digestion with a restriction enzyme. Knowledge of the size and sequence of the DNA as well as the cut sites of the restriction enzyme on that DNA will yield a number of fragments whose bp size is known. A thorough description of restriction enzymes in presented in Chapter 4. The resulting fragments are fluorescently labeled and detected as described in the Experimental section. *trpR* is added to the digested sample and CE is run again. Fragments that contain the binding site will either diminish in peak size or exhibit a shift.

Figure 3.10 is an electropherogram representing the CEMSA analysis of a DNA digest that contains a fragment housing the *trpR* binding site. A 3205 bp plasmid was digested with the restriction enzyme *Hinf*I and yielded fragments of 22, 65, 75, 264, 396, 517, 536 and 1330 bp in length. Figure 3.10 (A) depicts the laser fluorometric detection of the digested fragments. Upon addition of *trpR* to the digest a significant diminishing of the 264 bp fragment is observed in the presence of other fragments, as seen in Figure 3.10 (B). This is due to the fact that only the 264 bp fragment contains the binding region

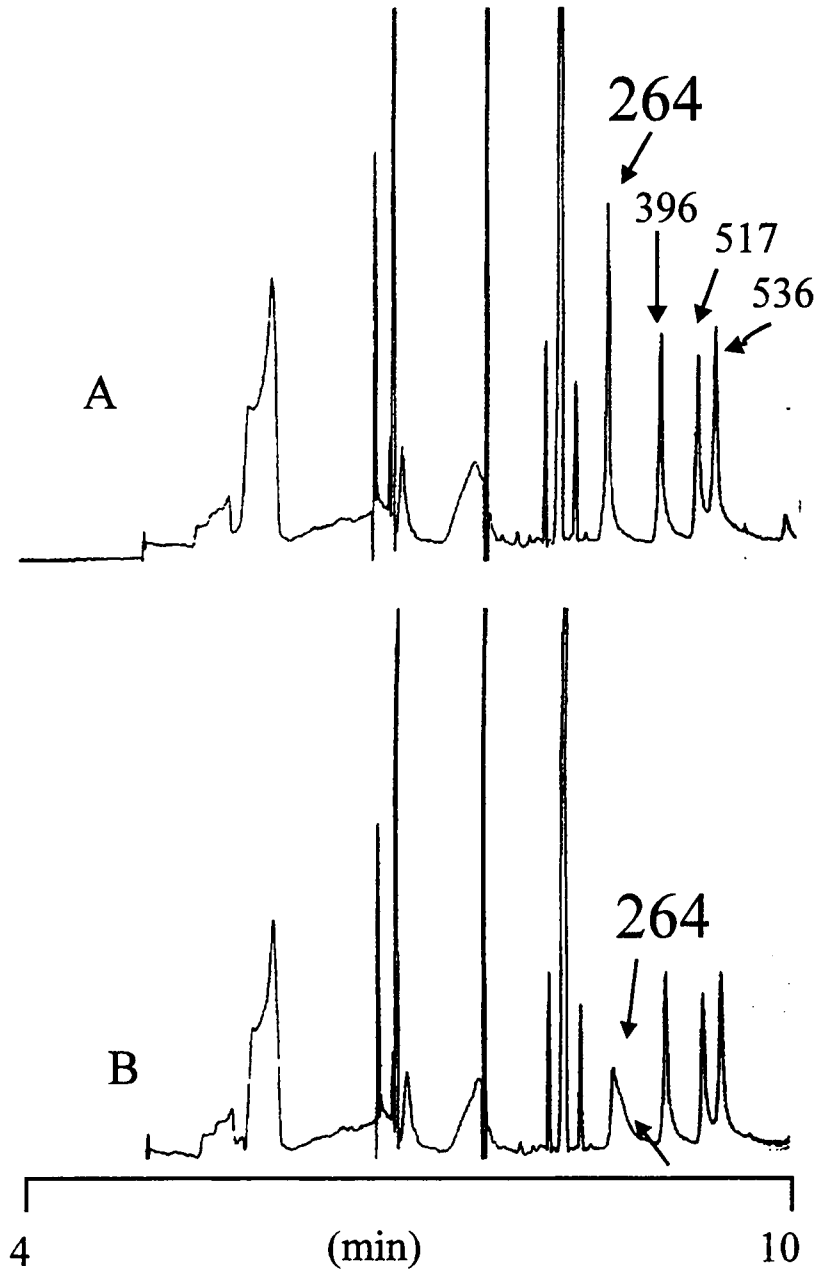


Figure 3.10. CEMSA analysis of *HinfI* digest containing fragment housing the *trpR* binding site. (A) laser fluorometric detection of the digested fragments and (B) diminishing of the 264 bp fragment upon addition of *trpR*.

for *trpR*. The absence of a discernible shift and complex peak presumably arises from the large size of the DNA fragment relative to the protein. In the case described earlier a 26 bp fragment bound by *trpR* resulted in a change in mobility of approximately 10 % for the protein bound DNA (complex) relative to the free DNA. The 264 bp fragment is 10-fold greater in size and would result in an approximately 1 % change in mobility for the complex. This small mobility difference between the free DNA and the complex would put the complex peak in the tailing baseline of the 264 bp fragment peak (depicted by the arrow in Figure 3.10 (B)). Nonetheless evidence of the protein-DNA interaction is seen in the diminishing of the fragment of interest.

Future work will entail working with other digests containing known fragments with the binding site. Other samples (DNA digests with binding sites and appropriate proteins) where the fragment sizes and binding sites are unknown to the experimenter will serve as blind tests. Once the methodology is appropriately tested, the digestion/analysis process will be repeated with enzymes that contain overlapping cut regions to quickly narrow binding regions on DNA.

CHAPTER 4

Temporal Analysis of DNA Restriction Digests by Capillary Electrophoresis

INTRODUCTION

As described in Chapters 2 and 3, CE is already showing widespread success as a method for the analysis of a variety of biomolecules. The ability to exploit the unique characteristics of CE will continue to make it a more robust tool for biochemical analysis. Studies of DNA restriction digests by CE have typically centered around the use of these samples to adjust and optimize separation conditions that will lead to increased resolution and reproducibility, as well as extremely fast analysis times. A great deal of this work is intended to make CE a more rigorous tool for genome sequencing, where fast analysis times and high reproducibility are extremely important. Researchers are continuously demonstrating how new sieving media (e.g., new polymers or new combinations of polymers) produce greater degrees of resolution. Also new capillary coating procedures, and even the absence of column coating combined with certain polymeric matrices, produce excellent run-to-run reproducibility. In another illustration of the utility of this technique, we have chosen to investigate CE as a diagnostic tool for the biochemical researcher by exploiting its advantages for the temporal analysis of the DNA restriction digestion process.

DNA restriction enzyme digests generate sets of fragments resulting from the enzyme cutting the double stranded (ds) DNA at specific sequence sites. During the

digestion, the enzyme cleaves particular phosphodiester linkages within the DNA as it comes in contact with the specific sequences of bp sites that it recognizes. As the digestion progresses towards completion, large fragments are further recognized and digested into smaller fragments until all restriction sites are cleaved. The size of the resulting fragments, in bp number, is usually known. However, a certain fragment that may be required for further research may not be generated from a complete digestion. Having access to these incompletely digested fragments and being able to determine the time during the digestion when these fragments are generated provides the researcher a greater range of DNA fragments to utilize for biochemical research.

Biochemical researchers traditionally approach this problem by performing partial DNA restriction digestions. These partial digestions serve the purpose of generating specific DNA fragments that are not available from digesting the DNA to completion with commercially available restriction enzymes. Once the time point is known during the digestion when a certain sized fragment is generated, the fragment can then be isolated for cloning and/or other purposes. At certain times during the digestion, the reaction is terminated resulting in a partially digested mixture that contains fragments that are different from those that result from complete digestion. This termination is typically achieved via heat inactivation or phenol addition/extraction (enzyme extracted with phenol and DNA recovered). The analysis of partial digests is traditionally performed by slab gel electrophoresis. The sizes of the fragments are determined by comparison of the migration distances to a standard containing DNA fragments of known size. However, in many cases, gel conditions cannot be easily optimized to produce the resolution that is

necessary to determine the presence of a partial digest fragment. A problem with analyzing partial digests by slab gel electrophoresis is the lack of resolution inherent to agarose gels. Also, because it is not known prior to the analysis when a certain size fragment is present, the procedure must be performed numerous times, varying the digestion times and amounts of restriction enzyme in the reaction mixture. The entire process can be extremely costly and time consuming, requiring multiple analyses performed to arrive at the desired information empirically.

An alternative approach involves using the high resolving power, speed and small sample requirements of CE as a diagnostic tool to monitor the time course of the digestion. As illustrated herein, a restriction digest can be performed under any desirable conditions and injected directly onto the capillary column, eliminating both the need to terminate the digestion for the purpose of sampling and any clean-up step that may also be needed. The ability to flush the column between runs allows the introduction of fresh sieving matrix, effectively eliminating carryover effects from the previous injection (most importantly the enzyme used in the digestion reaction).

A temporal analysis of the ΦX-174 DNA *Hae*III digestion is presented in this chapter. ΦX-174 DNA was digested with *Hae*III restriction enzyme under conditions that allowed the monitoring of digestion as it proceeded toward completion. Separation by a polymer solution of methyl cellulose (MC) in a polyacrylamide coated capillary allowed high resolution and a high degree of reproducibility between sequential runs. At pre-selected time intervals an injection of the digest, directly from the reaction mixture, was made. Sensitive detection was achieved by using ethidium bromide (EB) as an

intercalation dye and allowing intercalation to occur on-column. It is demonstrated that the course of the digestion (i.e., the creation and diminishing of fragment peaks) can be followed using this methodology.

Also demonstrated is the use of a temporal analysis by CE to determine when a cloning and expression vector containing a gene of interest (i.e. pET3a-PAI-1) has been linearized by the *Bst*YI restriction enzyme, an enzyme known to have 10 cut sites on this vector, as described below. A partial digest of the pET3a-PAI-1 DNA with the restriction enzyme *Bst*YI is desired for subsequent cloning of the DNA fragment which contains the cDNA (complementary DNA) encoding PAI-1 into another plasmid vector. Plasmids are circular double stranded (ds) DNA molecules. The digestion strategy required a partial digestion with *Bst*YI to avoid fragmentation of the coding sequence at an internal *Bst*YI restriction site.

As a first step in the cloning, a limited *Bst*YI digest is desired to produce cut plasmid with an average of one cut per molecule at any of the potential 10 cleavage sites. Subsequent digestion with a second enzyme, *Nde*I, cleaves at a unique site within the pET3a-PAI-1 plasmid at a position which lies immediately upstream of the coding sequence for PAI-1. One of the several *Nde*I/*Bst*YI fragments produced will be approximately 1350 bp in length and will contain the intact cDNA for PAI-1. Ultimately, this 1350 bp fragment from the *Nde*I/*Bst*YI double digestion will be isolated from the digestion mixture and sub-cloned into another plasmid vector. The immediate goal of this work was to use CE rapidly and conveniently to evaluate the temporal framework of the initial *Bst*YI digestion step to determine conditions which produce predominantly

single-cut plasmid. CE has a distinct advantage over agarose electrophoresis in its ability to detect low molecular weight fragments that will not be observed in the slab gel.

EXPERIMENTAL

Materials

In addition to those materials listed in Chapter 2, samples of ΦX-174 DNA *Hae*III digest (390 μg mL^{-1}), ΦX-174 phage DNA (10 units mL^{-1}), *Hae*III restriction enzyme (14 units μL^{-1}) with incubation buffer and a 50 base pair step ladder (0.25 μg μL^{-1}) were obtained from Sigma. pET3a-PAI-1 DNA (0.7 μg mL^{-1}) was the gift of Dr. David Ginsburg, Howard Hughes Medical Institute, University of Michigan Medical Center (Ann Arbor, MI, USA). *Bst*YI restriction enzyme (10,000 units mL^{-1}) with incubation buffer was obtained from New England Biolabs (Beverly, MA, USA). Fused silica capillary (50 and 75 μm i.d. x 365 μm o.d.) was coated with linear polyacrylamide by the method outlined in Chapter 2.

Running Buffer and Sample Preparation

All running buffers used in this work contained 45 mM TBE with the appropriate percentage of MC. These MC containing buffers were prepared in a manner previously described in Chapter 2. Digestion conditions for the *Hae*III digestion of ΦX-174 DNA and the *Nde*I digestion of pET3a-PAI-1 DNA consisted of 1 μL of DNA, 2 μL of the appropriate incubation buffer, 17 μL H_2O and 0.5 μL or 1 μL restriction enzyme, respectively. Digestion conditions for the *Bst*YI digestion of pET3a-PAI-1 consisted of 1

μL of DNA, 2 μL incubation buffer, 17 μL H_2O and the appropriate amount of diluted restriction enzyme in a total reaction volume of 20 μL.

Apparatus and Electrophoresis

The electrophoresis and fluorescence apparatus was assembled in-house. With the exception of laser, collection optics and appropriate filters, the apparatus and associated components is similar to the one described in Chapter 3 and depicted in Figure 3.3. A He-Ne "greenie" laser provided excitation of the ethidium bromide (EB) intercalated DNA and a 25 mm *f*/1 lens provided collection of the resulting fluorescence. The emission was isolated using a 570 nm cut-on filter. Rejection of plasma lines emanating from the laser was accomplished by placing a 543.5 nm (10 nm ± 2 nm FWHM) laser line filter in the path of the laser output. Electropherograms were recorded using a data collection routine programmed in-house and a strip chart recorder. A Hipotronics (Brewster, NY, USA) Model 840A high-voltage power supply was used to apply electrophoretic fields. Injections were performed electrokinetically by placing the inlet of the capillary into the sample vial and applying -2 kV for 20 seconds and -4 kV for 40 seconds for the ΦX-174 DNA and the pET3a-PAI-1 DNA, respectively. Following injection, the capillary inlet was replaced into the inlet running buffer vial and a running potential of -24 kV (400 V/cm) was applied across the 60 cm (47 cm to detector) capillary.

RESULTS AND DISCUSSION

Temporal Analysis of ΦX-174 DNA

To obtain the best separation conditions for DNA fragments in CE, it is important to be able to optimize for those conditions which are the most important for a given separation situation. When considering DNA restriction fragment digests, important considerations may include resolution of fragment peaks, efficiency, reproducibility and analysis time. The degree to which these considerations affect the separation is ultimately dependent upon what information is desired from the separation. In a large number of restriction digest separations, resolution of the two or three most difficult to resolve fragments may be the primary goal, while in other cases the identification of each fragment is all that is necessary, leaving resolution to be sacrificed to some degree.

When attempting to follow the time course of a restriction digest to completion, other considerations become more important. Since this technique deals with the injection of the digest directly from the reaction mixture, analysis time is extremely important. The greater the number of injections that can be made during the course of the digestion the greater the sampling rate, allowing a more accurate determination of when certain fragments begin to appear and disappear. If CE conditions are such that an analysis takes 30 minutes to complete, generally only a couple of time points can be analyzed before the digestion goes to completion, yielding a potentially poor overview of the digestion course.

Using the computer program outlined in Chapter 2 we are able to optimize separation conditions for the fast separation of restriction digests. Taking into

consideration both excess and inadequate resolution of fragments and analysis time, conditions were selected, prior to actual experimentation, that resulted in the best separation in the shortest amount of time. This advantage allowed experiments to be performed sequentially every 10-12 minutes, permitting considerably more time points to be experimentally determined during the course of the digestion. This procedure circumvented the many time consuming steps required for slab gel analysis. The speed of CE runs allows time between runs to rinse the column (with dilute acid) and fill it with fresh sieving matrix. This step helps to overcome any carryover effects (i.e., memory effects) from the previous runs. Injections can be made directly from the digestion reaction mixture without being concerned with the deleterious effects of the restriction enzyme resulting in near real-time analysis of digestions.

Figure 4.1 shows an electropherogram depicting an effectively complete digestion of the ΦX-174 plasmid DNA digested with *Hae*III restriction enzyme under conditions that allow the digestion to progress at a slower rate than under normal circumstances. Using a 0.5 % MC running buffer and detecting at 47 cm from the outlet of the 60 cm column, all 11 fragments are separated using a field of 400 V/cm. The electropherogram was generated by injecting directly from the digestion reaction mixture two hours after the initiation of the digestion. The efficiency for the 1353 bp fragment is approximately 2 x 10^6 plates/meter providing evidence that the presence of enzyme and other digestion materials does not influence the ability to generate extremely efficient separations. Digestion conditions were based upon manufacturer's suggestions with the exception of incubation temperature and concentration of restriction enzyme. The digestion was

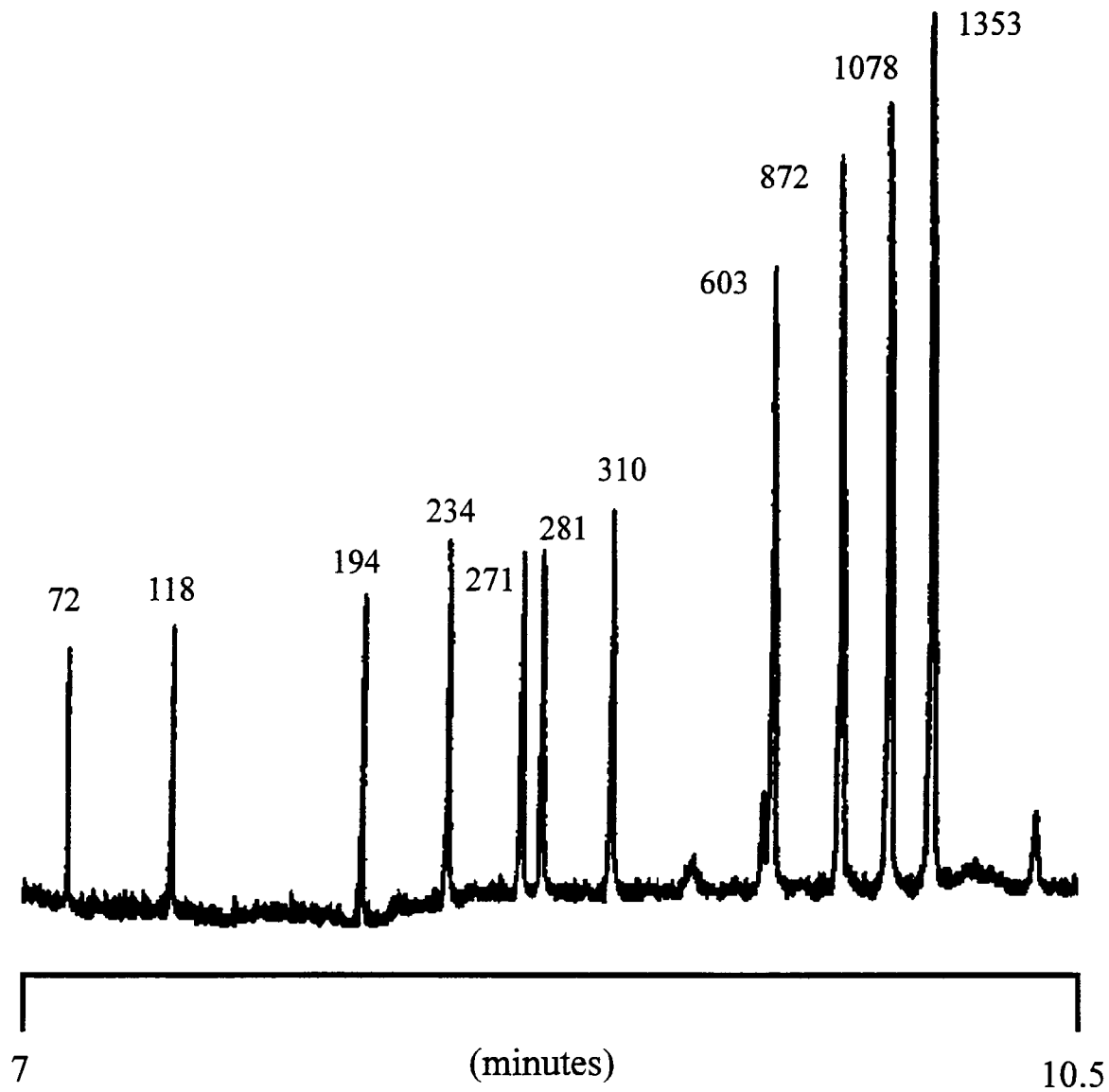


Figure 4.1. Effectively complete digestion of ΦX-174 plasmid DNA with *Hae*III restriction enzyme under conditions that allowed digestion to progress at a slower rate than under normal circumstances.

carried out on ice as opposed to 37 °C and the normal amount of enzyme used was halved, thereby slowing the reaction rate to facilitate monitoring of the digestion.

It can be seen in Figure 4.2 that the progression of a restriction digest can be followed by CE. At approximately 7 minutes into the digestion reaction (see t = 7 min separation in Figure 4.2) a single prominent fragment peak is evident relatively late in the electropherogram. Comparison of the migration time of this peak (11.3 minute) to that of the last eluting peak of the 50 bp step ladder (3147 bp, 11 minute, Figure 4.2) suggests that this peak corresponds to a single cut resulting in the linearization of the ΦX-174 plasmid DNA (5,386 bp). The 50 bp step ladder was used as a known size standard for migration time comparison only. A closer look at the 7 minute run shows evidence of additional fragmentation prior to the 11 minute fragment. Following the digestion through the seven samplings presents some interesting findings.

After approximately 26 minutes into the reaction a significant amount of digestion has occurred. Comparison of migration times to that of the almost completed digest (see t = 118 minute separation in Figure 4.2), where the elution order is 72, 118, 194, 234, 271, 281, 310, 603, 872, 1078 and 1353 bp, it is evident that the 72 and 118 bp fragments are beginning to appear. Also apparent is the disappearance of the 11.3 minute fragment. The last four migrating peaks of the complete digest (603, 872, 1078, and 1353 bp) are also becoming evident late in the elution window, and are labeled accordingly. However, it appears as though each of these late eluting peaks in the 26 minute sampling has a significant shoulder, potentially indicating the presence of partial fragments that are not present upon complete digestion. A closer look at the 26 minute sampling also indicates

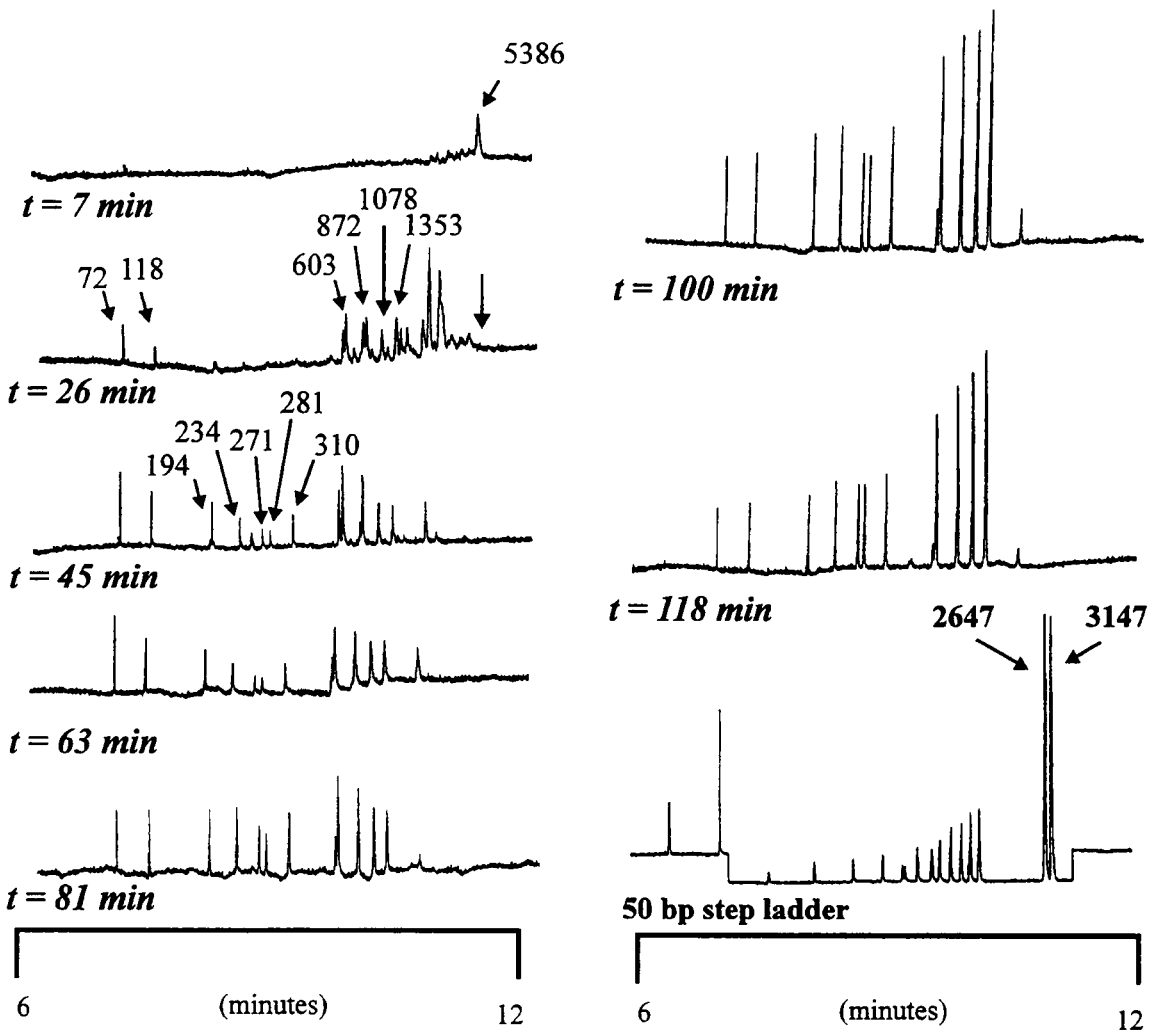


Figure 4.2. Progression of the ΦX-174 DNA *Hae*III digest. t = 7 min shows evidence of the 5386 bp single cut species. t = 26 min shows evidence of the 72, 118, 603, 872, 1078 and 1353 bp fragments. t = 45 shows evidence of the 194, 234, 271, 281 and 310 bp fragments. t = 118 shows an effectively complete digest.

two intense peaks eluting beyond the 1353 bp fragment, an additional indication of partial fragments. Sampling the digest at approximately 45 minutes shows that the digestion is beginning to appear more like the completely digested sample. The 72 and 118 fragments are becoming more intense and there is evidence of the mid-eluting fragments (194, 234, 271, 281 and 310 bp). Another indication that the digestion is progressing toward completion is the diminishing of the fragments eluting beyond the 1353 bp fragment, as initially seen it the 26 minute sampling. As the digestion progress from the first hour through the second hour it begins to take on a more complete appearance, until at approximately 118 minutes the digestion has effectively run its course, resulting in the familiar 11 fragment pattern inherent to this DNA and restriction enzyme.

Recall from the introduction that DNA restriction enzyme digests generate sets of fragments resulting from the enzyme cutting the ds DNA as it comes in contact with the specific sequences of bp sites that it recognizes (i.e. restriction sites). As the digestion progresses towards completion, the restriction sites on the larger fragments that were initially generated are further recognized and digested into smaller fragments. This continues until all restriction sites are cleaved. While the restriction sites at which an enzyme cleaves are pre-determined, the process by which larger fragments are digested to smaller fragments is random.

While it is interesting to follow the progression of a restriction enzyme digest, it can also serve the purpose of determining when bp fragments of certain size are present and what size bp fragments result from their digestion. Knowledge of the cut sites for a particular enzyme on a given DNA can provide valuable information about the generation

of these partial digestion fragments. Figure 4.3 depicts a restriction map for the ΦX-174 DNA *Hae*III digest. This plasmid DNA has 11 restriction sites that generate the familiar 11 fragments seen in Figures 4.1 and 4.2. The 11 unique restriction sites are depicted by the bp number where the enzyme cleaves and are shown around the outside of the restriction map (434, 668, 978, 1172, 1775, 3128, 4206, 4487, 4758, 4876 and 4948 bp). The 11 fragments that result from cleavage at each of these sites is shown along the inside of the restriction map, between the two cut sites that generate that fragment.

Figure 4.4 shows an example of the creation and diminishing of fragment peaks. As mentioned above each of the four latest eluting bp fragments in the complete digest (603, 872, 1078 and 1353 bp) indicate the presence of a significant shoulder. If the time course of the digestion is followed from the 26 minute sampling through completion it is observed that as the shoulder appearing on the 872 bp fragment decreases over time, the intensity of the 603 and 194 fragment peaks increases. It is likely that this shoulder is actually a 797 bp fragment and its digestion produces the 603 and 194 bp fragments. Referring to the restriction map in Figure 4.3 it is possible that a 797 bp fragment can actually be generated during the course of the digestion. Cleavage of the DNA at restriction sites 1775 and 978 would result in a 797 bp fragment. As this partial fragment is subsequently cleaved the 602 and 194 bp fragments result. The generation of the 797 bp fragment, however, could also be the result of any number of larger partial fragments that encompass the 1775 and 978 restriction sites.

Another example of the creation and diminishing of fragment peaks for this digest is shown in Figure 4.5. If the time course of the digestion is followed from the 26 minute

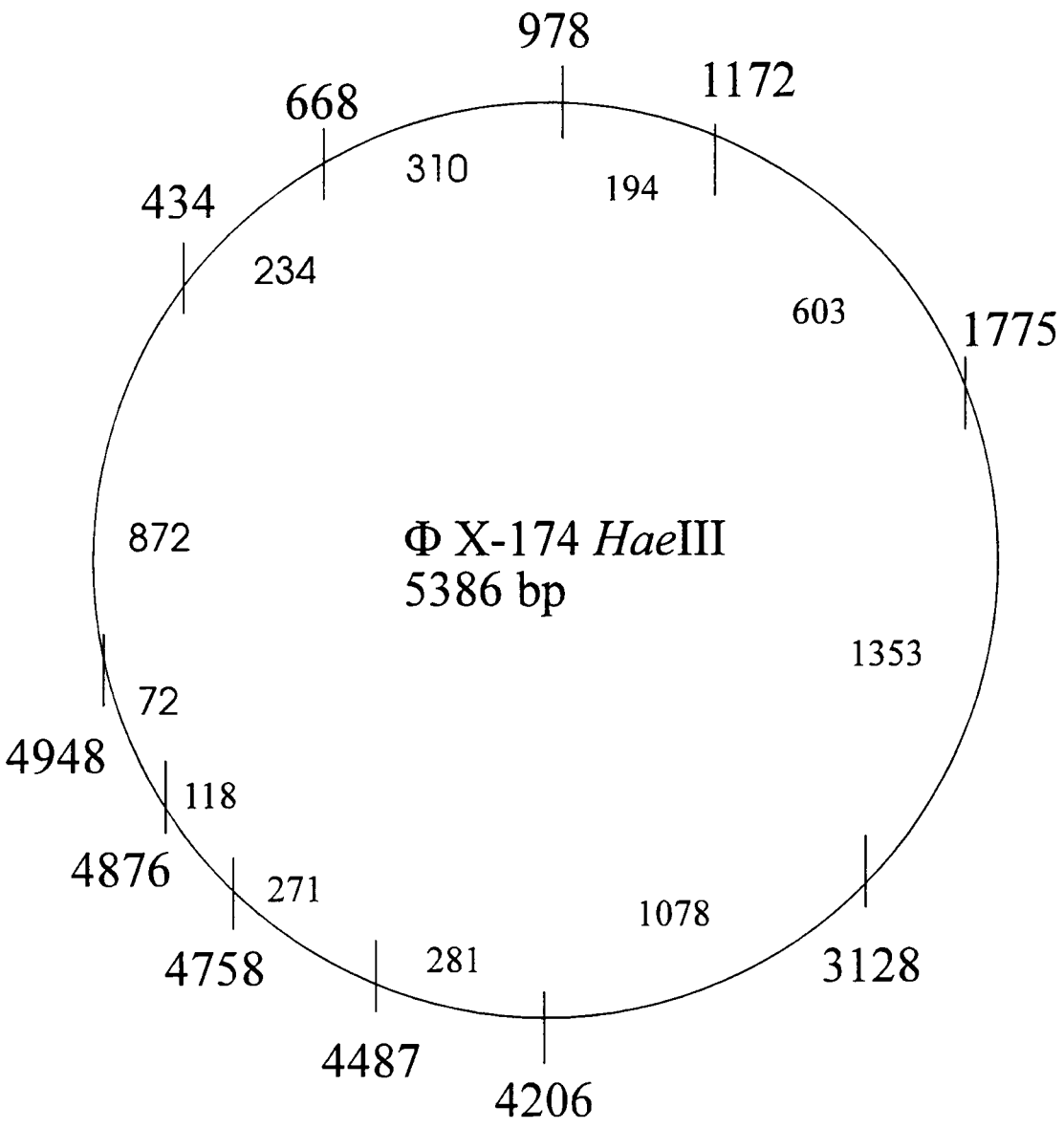


Figure 4.3. Restriction map for the ΦX-174 DNA *Hae*III digest depicting enzyme cut sites (outside of circle) and resulting fragments (inside of circle).

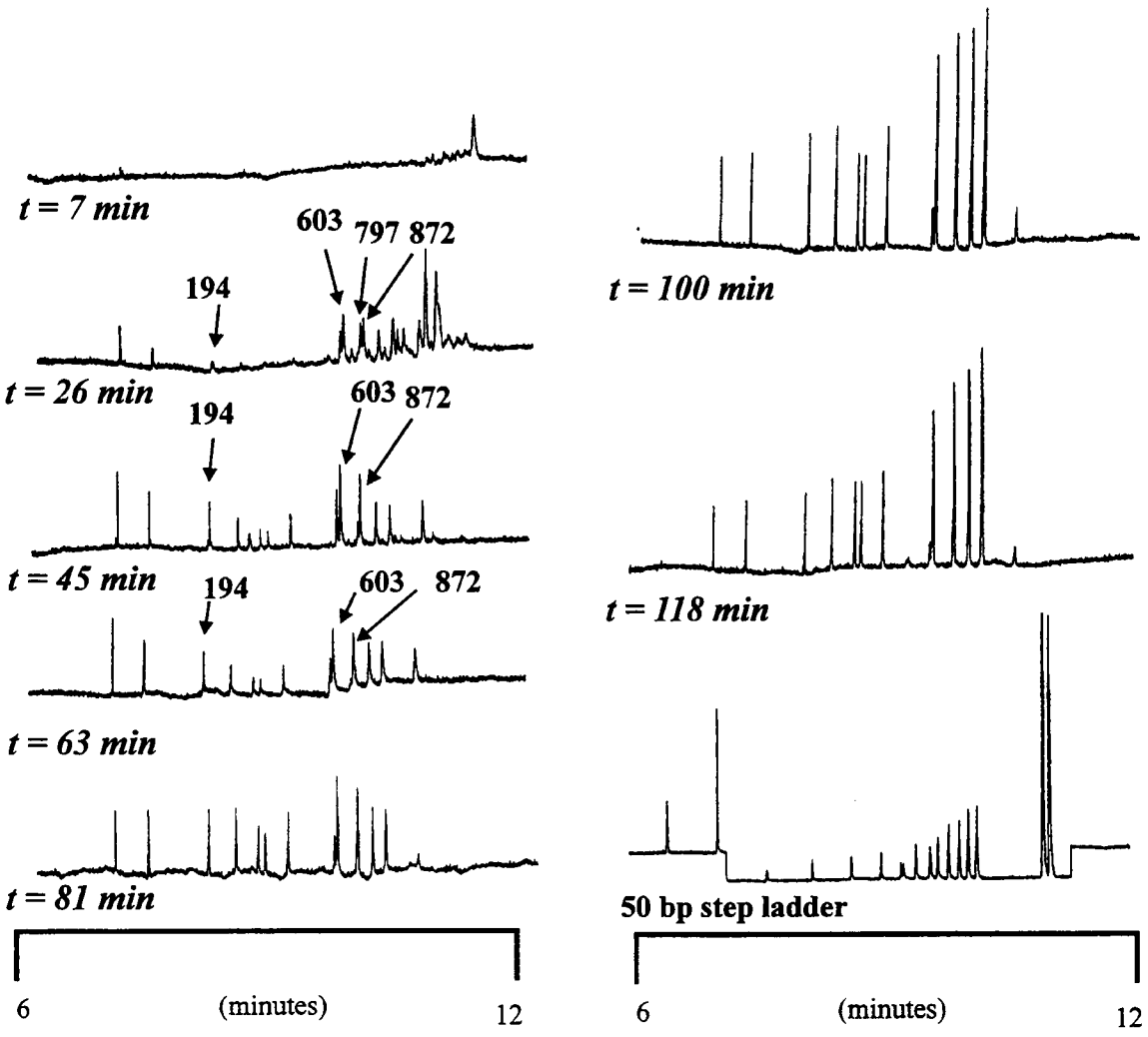


Figure 4.4. Temporal analysis of ΦX-174 DNA *Hae*III: creation of the 194 and 603 bp fragments from the 797 bp partial fragment.

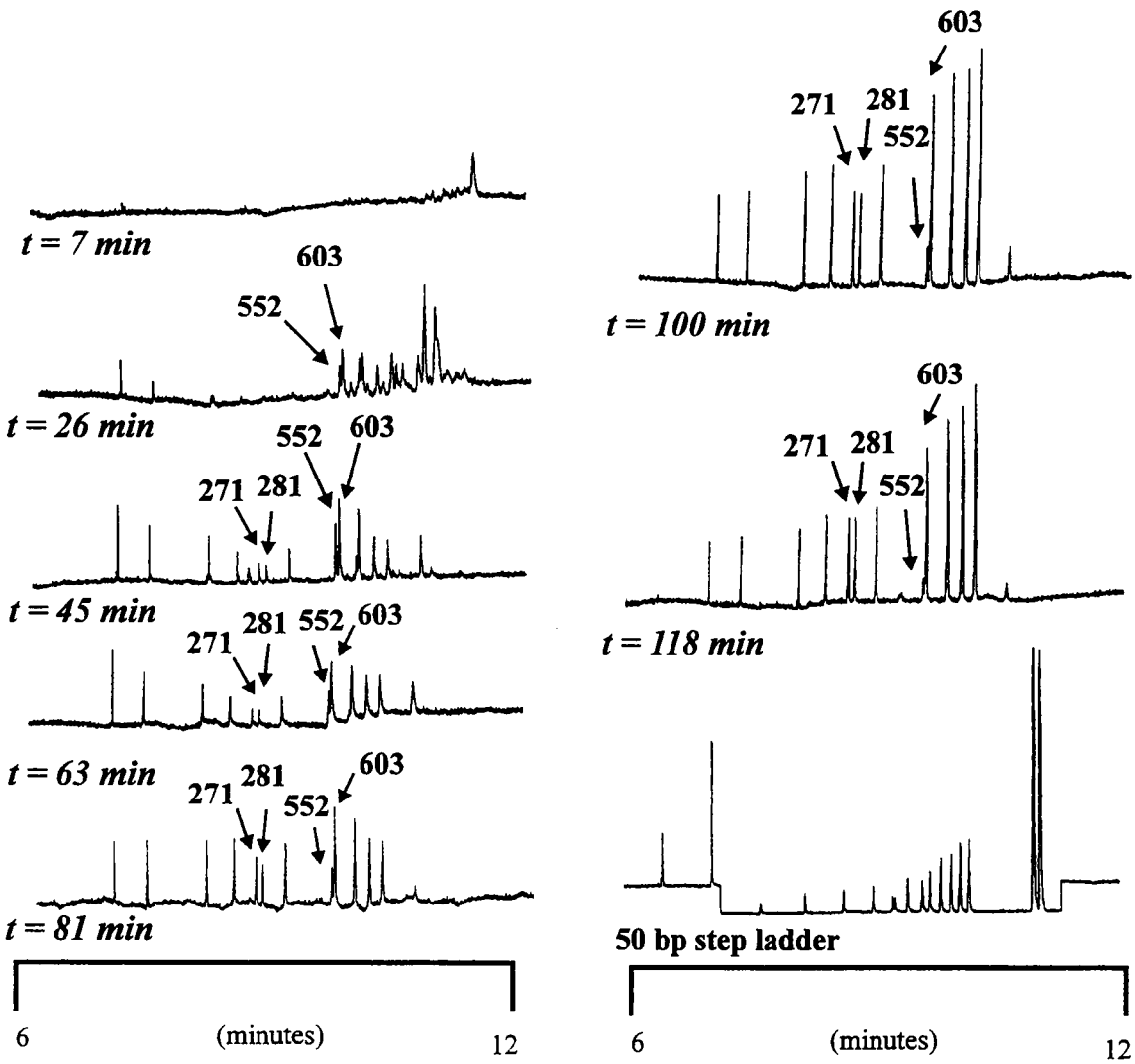


Figure 4.5. Temporal analysis of ΦX-174 DNA *Hae*III: creation of the 271 and 281 bp fragments from the 552 bp partial fragment.

sampling through completion it is observed that, as the shoulder appearing on the 603 bp fragment decreases over time, the intensity of the 271 and 281 bp fragment peaks increases. It is likely that this shoulder is actually a 552 bp fragment and its digestion produces the 271 and 281 bp fragments. This trend can be followed visually as the digestion progress through the 45, 63, 81, 100 and 118 minute samplings. Referring to the restriction map in Figure 4.3 it is possible that a 552 bp fragment can actually be generated during the course of the digestion. The 552 bp fragment could only result from the DNA being cleaved at restriction sites 4758 and 4206. As this partial fragment is subsequently cleaved the 271 and 281 bp fragments result. The generation of the 552 bp fragment, however, could also result from any number of larger partial fragments that encompass the 4758 and 4206 restriction sites being further cleaved.

Yet another example is illustrated in Figure 4.6. The fragment peak labeled 2431 is seen in each of the samplings at a migration time of approximately 10.8 minutes. When this peak is followed through the progression of the digestion it appears to be diminishing in intensity as the 1078 and 1353 fragment peaks increase in intensity. Comparison of the migration time of this peak (10.8 minutes) to that of the second to last eluting peak of the 50 bp step ladder (2647, 10.9 minutes) suggests that this peak corresponds to a 2431 bp fragment, the partial fragment leading to the 1078 and 1353 bp fragments. Again, referring to the restriction map in Figure 4.3 it is possible that this 2431 bp fragment can actually be generated during the course of the digestion. The 2431 bp fragment could only result from the DNA being cleaved at restriction sites 4206 and 1775. As this partial fragment is subsequently cleaved the 1078 and 1353 bp fragments

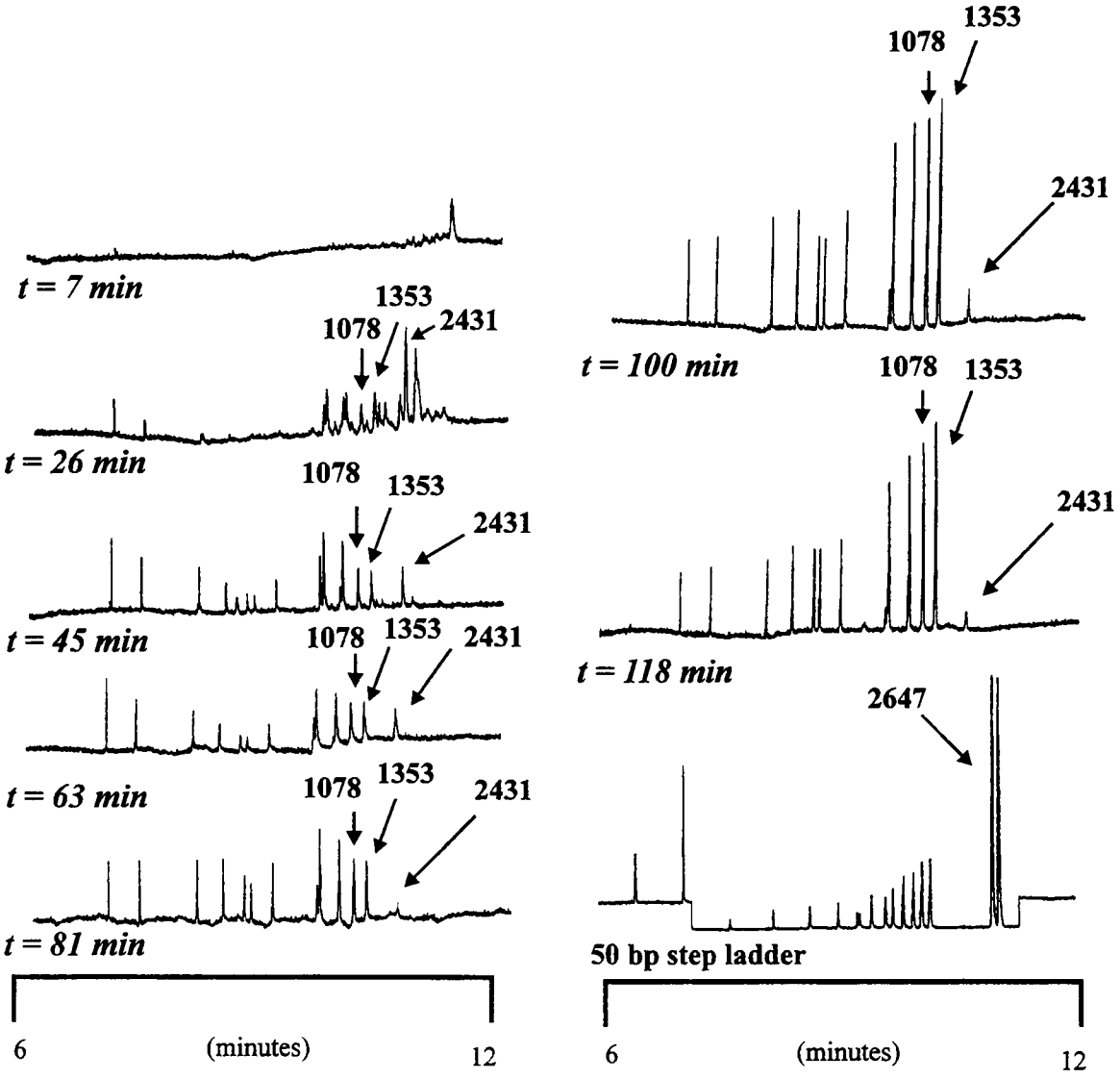


Figure 4.6. Temporal analysis of ΦX-174 DNA *Hae*III: creation of the 1078 and 1353 bp fragments from the 2431 bp partial fragment.

result. The generation of the 2431 bp fragment, however, could be the result of any number of larger partial fragments that encompass the 4206 and 1775 restriction sites.

While it may prove difficult to repeat exact sampling times from digestion to digestion, three separate digests of the ΦX-174 plasmid DNA with the restriction enzyme *Hae*III yielded highly reproducible results in terms of the digestion trends discussed above. Intra-run pooled standard deviations of migration times (from clearly identifiable peaks) averaged 2.7 s (n = 5-6) while the inter-run pooled standard deviation was 4.2 s (n = 3). While run to run reproducibility is important, an advantage inherent to restriction digest analysis is that there are generally recognizable fragment patterns. This relaxes the need for extreme precision (i.e., the standard deviations obtained herein are more than adequate). While the digestion of ΦX-174 plasmid DNA may not prove to be a challenging separation, it did serve as method development for temporal digests of other potentially more important DNA digestions.

Temporal Analysis of pET3a-PAI-1 DNA

Partial digestions in the biochemical laboratory are often required in recombinant DNA work to avoid fragmentation of a particular DNA sequence of interest. An example of such a partial digestion strategy is given with the limited *Bst*YI digest of pET3a-PAI-1 which avoids cleavage of the cDNA encoding PAI-1 at an internal *Bst*YI site. pET3a-PAI-1 DNA is a 5954 bp cloning vector that contains the 1,206 bp encoding gene plasminogen activator inhibitor-1 (PAI-1). Figure 4.7 depicts the restriction map for the *Bst*YI and *Nde*I restriction enzymes on pET3a-PAI-1 DNA. This plasmid DNA has 10

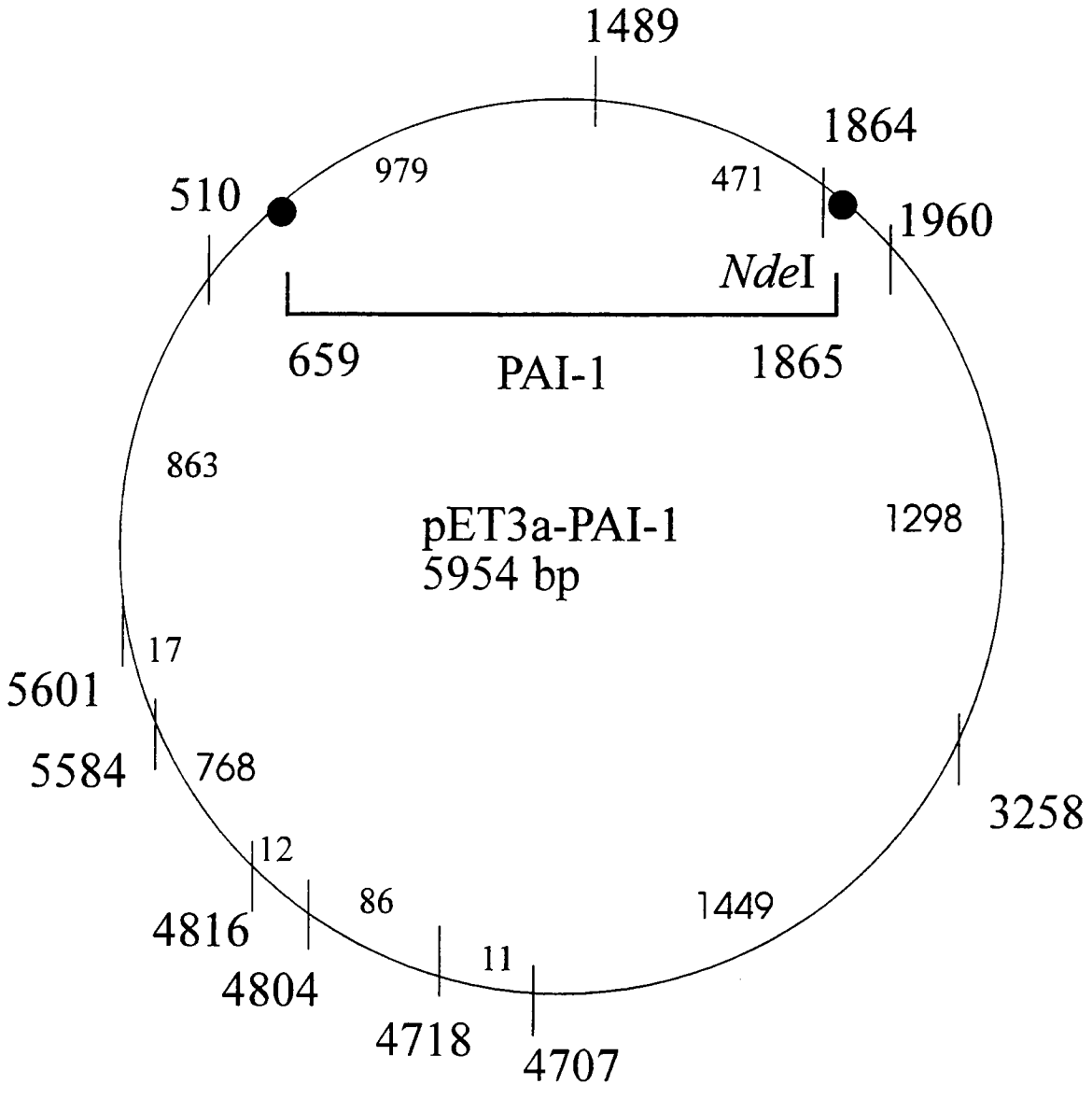


Figure 4.7. Restriction map for the *Bst*YI and *Nde*I restriction enzymes on pET3a-PAI-1 DNA depicting enzyme cut sites (outside of circle) and resulting fragments (inside of circle). The bracketed region between the 659 and 1865 bp fragments is the PAI-1 portion.

*Bst*YI restriction sites that generate 10 fragments when the digestion is taken to completion. The 10 unique restriction sites are depicted by the bp number where the enzyme cleaves and are shown around the outside of the restriction map (510, 1489, 1960, 3258, 4707, 4718, 4804, 4816, 5584, 5601). The 10 fragments that result from cleavage at each of these sites is shown along the inside of the restriction map, between the two cut sites that generate that fragment. The single cut site for *Nde*I at 1864 bp is also shown in Figure 4.7. The region bracketed between the black circles represents the PAI-1 portion of the plasmid. It is this region that is desired for subsequent cloning into another DNA of interest. However, as mentioned, this PAI-1 region contains an internal *Bst*YI restriction site at 1489 bp. Complete digestion with *Bst*YI will result in cleavage of the region of interest.

An experimental approach which yields plasmid cut at a single site is required as an initial step to isolate the intact region. A limited digest with *Bst*YI will be performed to produce an average of one cut per DNA molecule at any of the potential cut sites. Monitoring of the limited digest will be aided by comparison with the digestion of the same DNA with the *Nde*I enzyme that has a single unique restriction site. This site is at 1864 bp and lies immediately upstream of the PAI-1 region of interest.

Figure 4.8 shows the ability of temporal analysis by CE to determine ideal conditions for producing a single cut of the cloning vector pET3a-PAI-1. The first electropherogram in Fig. 4.8 shows the digestion of pET3a-pAI1 with the single cut enzyme *Nde*I. The resulting single fragment, 5954 bp, has a migration time of

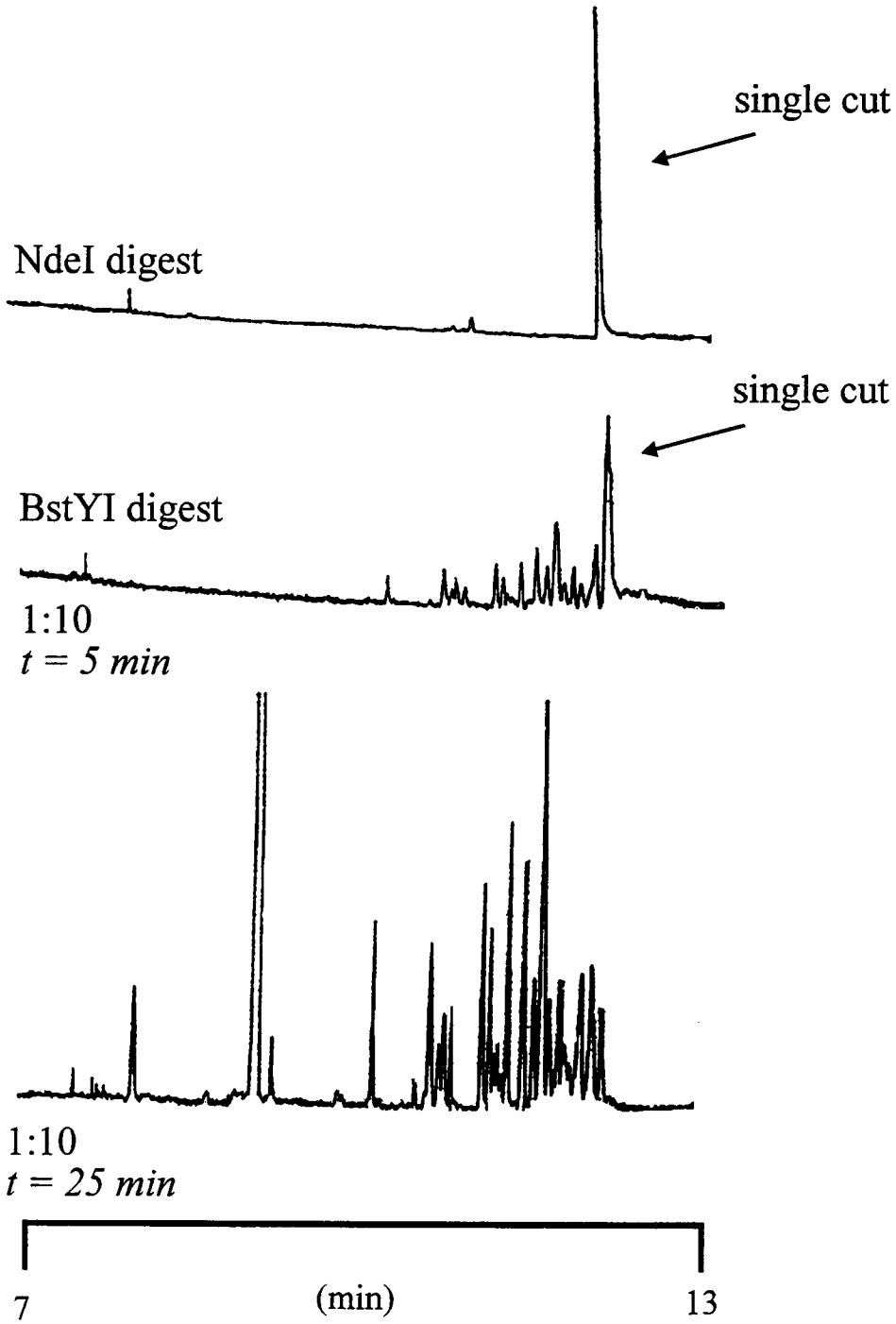


Figure 4.8. Temporal analysis of pET3a-PAI-1 DNA with *Bst*YI at a 1:10 dilution.

approximately 12 minutes. This digestion was used simply as a reference point for the migration time of a single cut species. The goal of this series of experiments was to determine what conditions produced a single cut with the restriction enzyme *Bst*YI. Digestion conditions were based upon manufacturer's suggestions with the exception of concentration of restriction enzyme. The digestion was carried out at 60 °C and the amount of enzyme used was adjusted appropriately. The second electropherogram depicts the initial investigation for the limited digest, a 1:10 dilution of enzyme. After sampling the digestion at five minutes into the reaction it is seen that a considerable amount of digestion is already occurring (see Figure 4.8, t = 5 minutes). It is evident, however, that the most intense peak at this time is still the peak that corresponds to a single cut appearing at 12 minutes. Further sampling of the 1:10 digest, at 25 minutes, shows an envelope of fragments, signifying that the digestion is still incomplete, but is well beyond conditions that would produce predominantly single cut fragments. It is evident that a 1:10 dilution of *Bst*YI is too concentrated to produce a single cut under these experimental conditions.

Figure 4.9 represents the temporal analysis investigation of a 1:100 dilution of enzyme. The top electropherogram again represents a true single cut species from the *Nde*I enzyme as a migration time marker. At five minutes into the digestion, a faint peak is seen at approximately 12 min. This signifies that this concentration of *Bst*YI may be the appropriate amount to produce a single cut. Twenty five minutes into this digestion it becomes more evident that a single cut is achieved. However, close inspection of the run

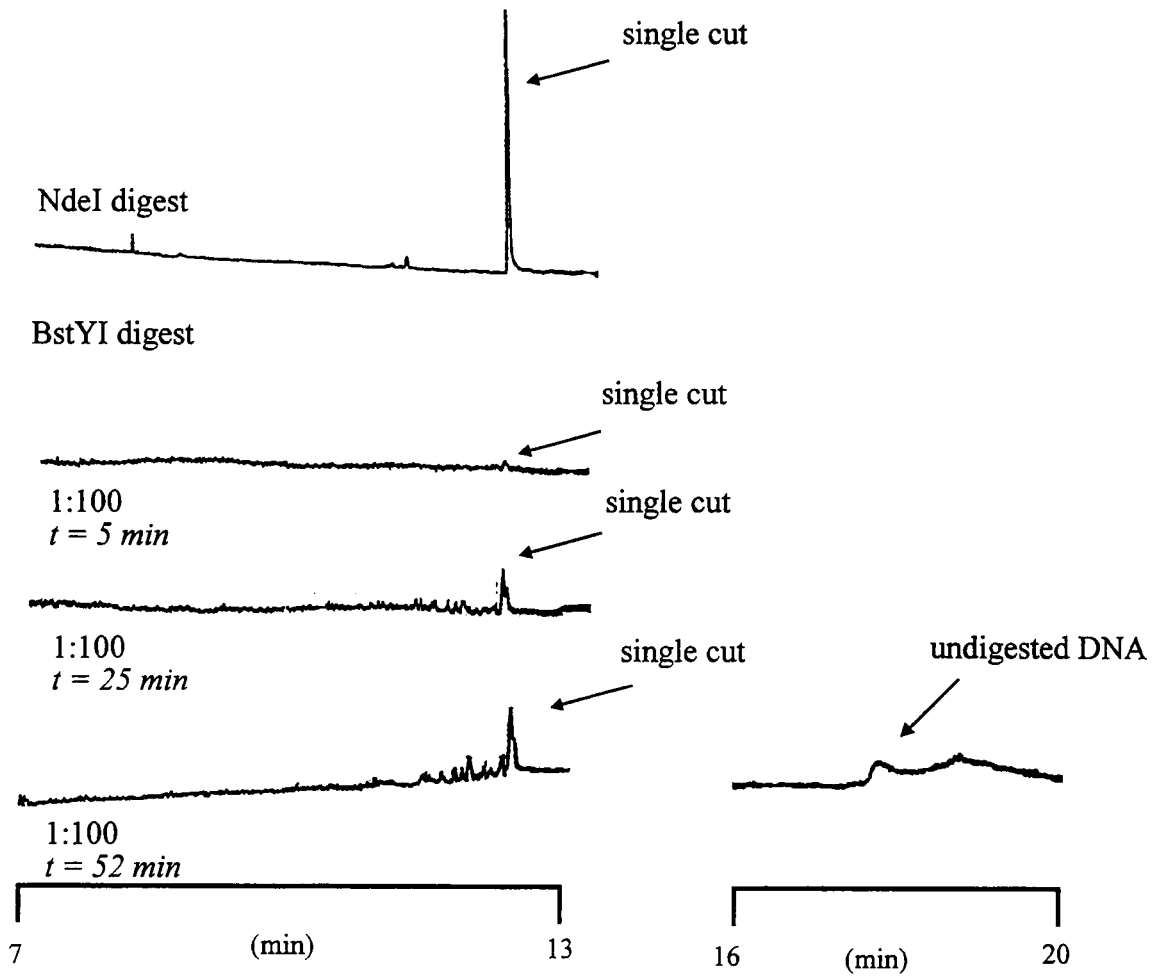


Figure 4.9. Temporal analysis of pET3a-PAI-1 DNA with *Bst*YI at a 1:100 dilution.

shows a small envelope of fragments appearing directly prior to the single cut fragment. Following the digestion out to 52 minute begins to show a more extensive envelope of fragmentation prior to the single cut species. A continuation of the 52 minute electropherogram shows undigested DNA appearing at approximately 18 min (as compared to an injection of uncut DNA not shown). Evidence of undigested DNA suggests that, while this enzyme concentration is to some degree producing predominately single cut fragments, it is doing so at a rate which is not completely consuming the plasmid DNA. Nonetheless, this series of experiments utilizing temporal analysis CE has shown that a 1:100 dilution of *Bst*YI will produce a single cut when digested in under 20 min. This entire series of experiments was conducted in under 3 hours, including analysis time. The more traditional approach to evaluate the partial digestion conditions using agarose gel electrophoresis gave similar results, indicating that a dilution of enzyme from 1:100 to 1:500 was necessary to achieve a single-cut plasmid. In contrast to the rapid analysis time afforded by CE, the use of slab gels to evaluate the digestion conditions for this system required several days and many more test reactions.

This work demonstrates methodology for the temporal analysis of DNA restriction enzyme digests by CE. Beyond providing interesting bioanalytical information about the progression of these digests, these studies prove beneficial in the facile determination of partial digest conditions that may be warranted for future studies. Continuation of the work presented in this chapter is focused on using temporal analysis to further optimize single cut conditions of the pET3a-PAI-1 DNA with *Bst*YI while minimizing the amount of uncut plasmid. Subsequent digestion with two single-cut enzymes (*Nde*I and *Nru*I)

will produce a variety of fragments containing the 1350 bp fragment alluded to earlier. This fragment will be isolated and collected using larger diameter columns for PCR amplification, further demonstrating the utility of CE for biochemical analysis.

CHAPTER 5

Concluding Remarks

In general, the transfer of traditional biochemical techniques to CE formats proved to be a challenging and rewarding endeavor. The insight gained into the advantages and limitations of both was instrumental in the success described in the preceding pages. While conclusions were drawn in each of the data chapters, an overall summary of these findings as well as insight into the future of CE as a biochemical analysis method is warranted and presented in this chapter.

The high degree of variability among DNA samples (seen as electrophoretic diversity) makes the determination of experimental conditions, for even a single analysis, a time consuming process. Computational models based on the content of sample and empirically derived relationships between fragment size, soluble polymer concentration and column length allow for the optimization of separation performance for all fragments in a sample. Operator selectable factors provide a means to influence each specific separation by placing weight on their respective terms in a chromatographic response factor. The subsequent generation and visualization (simulated electropherograms) of simulation data for each set of experimental conditions aids in the adjustment of weighting factors. The 11 fragments of the ΦX-174 DNA *Hae*III digest served as an appropriate representation of fragments spanning a wide range of base pair size. Excellent correlation between computational and experimental work is demonstrated as is

the ability to predict separation needs for the pBR322 DNA *Hinf*I digest not considered in the development of the program.

While the optimization program accurately predicted experimental performance, refinements in the mobility expression to accommodate the different migration mechanisms are needed in order to decrease discrepancies in observed migration times. Expansion of the initial data set to encompass a greater range of bp fragment sizes and methyl cellulose concentrations will also make the program more universal. Instrumental improvements including expansion of detection zone and increased column length would also prove valuable. Nonetheless the ability of the computational methods to produce optimized experimental conditions is clearly illustrated.

The development of a capillary-based assay for the study of protein-DNA interactions that combines the rapid and minimal sample consumption methods of CE with the selective separation influence of mobility shift assays was undertaken. The well characterized interaction between the *trp* repressor of *E. coli* and the *trp* operator served as the basis of the assay. The use of fluorescently tagged DNA not only lends itself to laser induced fluorescence detection but also avoids the use of radiolabeled detection. It is demonstrated that the composition and pH of the running buffer are critical for maximized efficiency and resolution of DNA from the *trp*R-DNA complex. Quantitative studies showed reaction of *trp*R with DNA resulted in the diminishing of free DNA signal and the simultaneous creation of the well resolved *trp*R-DNA complex. It is also shown that the specificity of *trp*R for DNA can be used to selectively separate *trp*R from a sample containing non-specific proteins.

While the demonstration of mobility shift assays in CE was successful, refinements could extend the utility of the technique to the determination of a variety of parameters inherent to protein-DNA interactions. Improvements such as more efficient or alternative fluorescent labeling methods and increased detectability will continue to make this assay a true competitor to the traditional gel-based assay. Preliminary studies aimed at protein binding site determination on DNA fragments proved promising.

The diagnostic capabilities of CE are demonstrated with the temporal analysis of DNA restriction enzyme digests. ΦX-174 DNA was digested with *Hae*III under conditions that allowed the monitoring of the digestion as it progressed towards completion. Separation by methyl cellulose allowed high resolution and a high degree of reproducibility between sequential runs. At pre-selected time intervals an injection of the digest directly from the reaction mixture was made, allowing the creation and diminishing of fragment peaks to be followed. Also demonstrated is the ability to use temporal analysis to determine ideal conditions for producing a single cut within a DNA of interest (i.e. pET3a-PAI-1) which contains 10 potential cleavage sites. This initial attempt to follow a restriction digest through completion not only provides meaningful information for the biochemical researcher, but also furthers the use of CE as a diagnostic tool for the biochemical laboratory. Current work entailing temporal analysis is directed at the determination of conditions that will produce a fragment of interest for subsequent collection and PCR amplification.

In conclusion, it has been shown that CE is a viable method for biochemical analysis. The preceding chapters, each detailing an individual aspect, set the stage for the

continuation and evaluation of CE's ever increasing role in biotechnology. While the mentioned methodologies are not without their drawbacks, simple refinements will produce more robust and mature techniques.

REFERENCES

REFERENCES

1. Hjerten, S., *Chromatogr. Rev.*, **1967**, *9*, 122.
2. Virtanen, R., *Acta Polytech. Scand. Appl. Phys. Ser*, **1974**, *123*.
3. Mikkers, F.E.P., Everaerts, F.M. and Verheggen, P.E.M., *J. Chromatogr.*, **1979**, *169*, 11.
4. Jorgenson, J. and Lukacs, K.D., *Anal. Chem.*, **1981**, *53*, 1298.
5. Jorgenson, J. and Lukacs, K.D., *Science*, **1983**, *222*, 266.
6. Sepaniak, M.J., Swaile, D.F., Powell, A.C. and Cole, R.O. In Capillary Electrophoresis, Theory and Practice, Grossman, P.D. and Colburn, J.C. (eds), Academic Press, Inc., San Diego, 1992, 159.
7. Cheng, Y.F. and Dovichi, N.J., *Science*, **1989**, *480*, 141.
8. Cheng, Y.F., Wu, S., Chen, D.Y. and Dovichi, N.J., *Anal. Chem.*, **1990**, *62*, 496.
9. Zhang, J.Z., Chen, D.Y., Wu, S., Harke, H.R. and Dovichi, N.J., *Clin. Chem.*, **1991**, *37*, 1492.
10. Hjerten, S., *J. Chromatogr.*, **1983**, *270*, 1.
11. Ogston, A.G., *Trans. Faraday Soc.*, **1958**, *54*, 1754.
12. Viovy, J.L. and Duke, T. *Electrophoresis*, **1993**, *14*, 322.
13. de Gennes, P.G., *J. Chem. Phys.* **1971**, *55*, 572
14. Doi, M. and Edwards, S.F., *JCS Faraday Trans. II*, **1978**, *79*, 1789.
15. Lerman, L.S. and Frisch, H.L., *Biopolymers*, **1982**, *21*, 995.
16. Lumpkin, O.J., Dejardin, P. and Zimm, B.H., *Biopolymers*, **1985**, *24*, 1573.

17. Kleparnik, K., Fanali, S. and Bocek, P., *J. Chromatogr. A.*, **1993**, *638*, 283.

18. Hjerten, S., Srichaivo, T. and Palm, A., *Biomed. Chromatogr.*, **1994**, *8*, 73.

19. Barron, A.E., Blanch, H.W. and Soane, D.S., *Electrophoresis*, **1994**, *15*, 597.

20. Grossman, P.D. and Soane, D.S., *J. Chromatogr. A.*, **1991**, *559*, 257.

21. Hjerten, S., Valtcheva, L., Elenbring, K. and Eaker, D.J., *J. Liq. Chromatogr.*, **1989**, *12*, 2471.

22. Zhu, M., Hansen, D.L., Burd, S. and Gannon, F., *J. Chromatogr. A.*, **1989**, *480*, 311.

23. McGregor, D.A. and Yeung, E.S., *J. Chromatogr. A.*, **1993**, *652*, 67.

24. MacCrehan, W.A., Rasmussen, H.T. and Northrop, D.M., *J. Liq. Chromatogr.*, **1992**, *15*, 1063.

25. Gilges, M., Husmann, H., Kleemiβ, M.H., Motsch, S.R. and Schomburg, G., *J. High Resolut. Chromatogr.*, **1992**, *15*, 452.

26. Kleemiβ, M.H., Gilges, M. and Schomburg, G., *Electrophoresis*, **1993**, *14*, 515.

27. Fung, E.N. and Yeung, E.S., *Anal. Chem.*, **1995**, *67*, 1913.

28. Izumi, T., Yamaguchi, M., Yoneda, K., Isobe, T., Okuyama, T. and Shinoda, T., *J. Chromatogr. A.*, **1993**, *652*, 41.

29. Heiger, D.N., Cohen, A.S. and Karger, B.L., *J. Chromatogr. A.*, **1990**, *516*, 33.

30. Clark, B.K. and Sepaniak, M.J., *J. Microcol. Sep.*, **1993**, *5*, 275.

31. Nathakarnkitkool, S., Oefner, P.J., Bartsch, G., Chin, M.J. and Bonn, B.K., *Electrophoresis*, **1992**, *13*, 18.

32. Guttman, A. and Cooke, N., *J. Chromatogr. A .*, **1991**, *559*, 285

33. Harke, H.R., Bay , S., Zhang, J. Z., Rocheleau ,M.J. and Dovichi, N.J., *J. Chromatogr. A.*, **1992**, *608*, 143.

34. Luckey, J.A. and Smith, L.M., *Anal. Chem.* **1993**, ***65***, 2841.

35. Navin, M.J., Rapp, T.L. and Morris, M.D., *Anal. Chem.* **1994**, *66*, 1179.

36. Swerdlow, H., Zhang, J.Z., Chen, D.Y., Harke, H.R., Grey, R., Wu, S., Dovichi, N.J. and Fuller, C., *Anal. Chem.* **1991**, *63*, 2835.

37. Huang, X.C., Quesada, M.A. and Mathies, R.A., *Anal. Chem.* **1992**, *64*, 2149.

38. Tayor, J.A. and Yeung, E.S., *Anal. Chem.* **1993**, *65*, 956.

39. Uenos, K. and Yeung, E.S., *Anal. Chem.* **1994**, *66*, 1424.

40. Ng, C.L., Lee, H.K. and Li, S.F.Y., *J. Chromatogr. A.*, **1994**, *659*, 427.

41. Chang, H.T. and Yeung, E.S., *J. Chromatogr.B.*, **1995**, *669*, 113.

42. Clark, B.K., Nickles, C.L., Morton, K.C and Sepaniak, M.J., *J. Microcol. Sep.* **1994**, *6*, 503.

43. Stalbom, B., Torven, A. and Lundberg, L.G., *Anal. Biochem.* **1988**, *8*, 2896.

44. Bianchi, N., Mischiati, C., Feriotto, G. and Gambri, R., *Nucleic Acids Res.* **1993**, *21*, 3596.

45. Kuypers, A.W.H.M., Willems, P.M.W., van der Schans, M.J., Linssen, P.C.M., Wessels, H.M.C., de Bruijn, C.H.M.M., Everaerts, F.M. and Mensinkk, E.J.B.M., *J. Chromatogr. A .*, **1993**, *621*, 149.

46. Khrapks, K., Hanekamp, J.S., Thinly, W.G., Belenkii, A., Foret, F. and Karger, B.L., *Nucleic Acids Res.* **1994**, *22*, 364.

47. Hjerten, S., *J. Chromtogr. A.*, **1985**, *347*, 191.

48. Hjerten, S., *Electrophoresis*, **1993**, *14*, 390.

49. Bianchi-Bosisio, A., Loehrlein, C., Synder, R.S. and Righetti, P.G., *J. Chromtogr. A.*, **1980**, *189*, 317.

50. Davidson, J., *Spectrosc.*, **1992**, *7*, 31.

51. LePecq, J.B. and Paoletti, C., *J. Liq. Chromatogr.*, **1991**, *14*, 869.

52. Berka, J., Pariat, Y.F., Muller, O., Hebenbrock, K., Heiger, D.N., Foret, F. and Karger, B.L., *Electrophoresis*, **1995**, *16*, 377.

53. Demana, T., Lanan, M. and Morris, M.D., *Anal. Chem.*, **1991**, *63*, 2795.

54. Guttman, A. and Cooke, N., *Anal. Chem.*, **1991**, *63*, 2038.

55. Clark, B.K., Vo-Dinh, T. and Sepaniak, M.J., *Anal. Chem.* **1995**, *67*, 680.

56. Berridge, J.C., *Techniques for the Optimization of HPLC Separations*, John Wiley and Sons, **1985**.

57. Deming, S.N. and Morgan, S.L., *Anal. Chem.* **1973**, *45*, 278A.

58. Powell, A.C. and Sepaniak, M.J., *Anal. Inst.* **1993**, *21*, 25.

59. Drossman, H., Luckey, J.A., Kostichka, J.A., D'Cunha, J. and Smith, L.M., *Anal. Chem.*, **1990**, *62*, 900.

60. Baba, Y., Matsuura, T., Wakamoto, K., Morita, Y., Nishitsu, Y. and Tsuhako, M., *Anal. Chem.*, **1992**, *64*, 1221.

61. Carson, S., Cohen, A.S., Belenkii, A., Ruiz-Martinez, M.C., Berka, J. and Karger, B.L., *Anal. Chem.*, **1993**, *65*, 3219.

62. Chen, D.Y., Swerdlow, H.P., Harke, H.R., Zhang, J.Z. and Dovichi, N.J., *J. Chromatogr. A.*, **1994**, *680*, 479.

63. Lu, H., Arriaga, E., Chen, D.Y. and Dovichi, N.J., *J. Chromatogr. A.*, **1994**, *680*, 497.

64. Slater, G.W. and Drouin, G., *Electrophoresis*, **1992**, *13*, 574.

65. Slater, G.W., Mayer, P. and Drouin, G., *Electrophoresis*, **1992**, *13*, 961.

66. Ruiz-Martinez, M.C., Berka, J., Belenkii, A., Foret, F., Miller, A.W. and Karger, B.L., *Anal. Chem.*, **1993**, *65*, 2851.

67. Zagursky, R.J. and McCormick, R.M., *BioTechniques*, **1990**, *9*, 74.

68. Manabe, T., Chen, N., Terabe, S., Yohda, M. and Endo, I, *Anal. Chem.*, **1994**, *66*, 4243.

69. DeDionisio, L., *J. Chromatogr. A.*, **1993**, *652*, 101.

70. Bourque, A.J. and Cohen, A.S., *J. Chromatogr. B.*, **1994**, *662*, 343.

71. Vilenchik, M., Belenky, A. and Cohen, A.S., *J. Chromatogr. A.*, **1994**, *663*, 105.

72. McCord, B.R., McClure, D.L. and Jung, J.M., *J. Chromatogr. A.*, **1993**, *652*, 75.

73. McCord, B.R., Jung, J.M. and Holleran, E.A., *J. Liquid Chromatogr.*, **1993**, *16*, 1963.

74. Schwartz, H.E., Ulfelder, K., Sunzeri, F.J., Busch, M.P. and Brownlee, R.G., *J. Chromatogr. A.*, **1991**, *559*, 267.

75. Arakawa, H., Uetanaka, K., Maeda, M. and Tsuji, A., *J. Chromatogr. A.*, **1994**, *664*, 89.

76. Butler, J.M., McCord, B.R., Jung, J.M. and Allen, R.O., *J. Chromatogr. B.*, **1994**, *658*, 271.

77. Gelfi, C., Righetti, P.G., Magnani, C., Cremonesi, L. and Ferrari, M., *Clin. Chim. Acta*, **1994**, *229*, 181.

78. Gelfi, C., Righetti, P.G., Cremonesi, L. and Ferrari, M., *Electrophoresis*, **1994**, *15*, 1506.

79. Nashabeh, W. and El Rassi, Z., *J. Chromatogr. A.*, **1992**, *596*, 251.

80. Grossman, P.D. *Capillary Electrophoresis: Theory and Practice*, Grossman, P.D. and Colburn, J.C. (eds.). Academic Press, San Diego, **1993**, 256.

81. Weinberger, R., Practical Capillary Electrophoresis. Academic Press, Boston, Massasuchuets, **1993**, 68.

82. Guzman, N.A. (ed.), *Capillary Electrophoresis Technology*. Chromatographic Science Series, Marcel Dekker, New York, **1993**, 605.

83. Rush, R.S. In: *Applications of Enzyme Biotechnology*. Kelly, J.W. and Baldwin, T.O. (eds.), Plenum Press, New York, **1991**, 233.

84. Keough, T., Takigiku, R, Lacey, M.P. and Purdon, M., *Anal. Chem.*, **1992**, *64*, 1594.

85. Castoro, J.A., Chiu, R.W., Monnig, C.A. and Wilkins, C.L., *J.Am. Chem. Soc.*, **1992**, *114*, 7571.

86. Yu, W., Vath, J.E., Huberty, M.C. and Martin, S.A., *Anal. Chem.*, **1993**, *65*, 3015.

87. Jacob, F. and Monod, J., *J. Mol. Biol.*, **1961**, *3*, 318.

88. Fried, M.G.and Crothers, D.M., *Nucl. Acids Res.*, **1981**, *9*, 6505.

89. Garner, M.M. and Revzin, A., *Nucl. Acids Res.*, **1981**, *9*, 3047.

90. Lane, D., Prentkl, P. and Chandler, M., *Microbiol. Revs.*, **1992**, *56*, 509.

91. Carey, J., *Proc. Natl. Acad. Sci. USA*, **1988**, *85*, 975.

92. Carey, J., Lewis, D.E.A., Lavoie, T.A. and Yang, J., *J. Biol. Chem.*, **1991**, *266*, 24509.

93. Czernik, P.J., Shin, D.S. and Hurlburt, B.K., *J. Biol. Chem.*, **1994**, 269, 27869.

94. Kolb, A,. Spassky, A., Chapon, C., Blazy, B. and Buc, H., *Nucl. Acids Res.*, **1983**, *11*, 7833.

95. Wu, H.M. and Crothers, D.M., *Nature*, **1984**, *308*, 509.

96. Konarska, M.M. and Sharp, P.A., *Cell*, **1986**, *46*, 845.

97. Garner, M.M. and Revzin, A., In: *Gel Electrophoresis of Nucleic Acids-A Practical Approach*, D. Rickwood and B.D. Hames (eds), Oxford University Press, New York, **1990**, 201.

98. Yanofsky, C. and Crawford, I.R., In: *Eshherichia coli and Salmonella typhimurium - Cellular and Molecular Biology*, Neidhardt, F. (ed.), American Society for Microbiology, Washington, **1991**, 1453.

99. Sigler,P.B., In: *Transcriptional Regulation,* McKnight, S.L. and Yamamoto, K.R. (eds.), Cold Spring Harbor Press, New York, **1992**, 475.

100. Makeli, S.J., Czernik, P.J. and Hurlburt, B.K., **in preparation**.

101. Staller, T.D., Ph.D. Thesis, University of Tennessee, 1995.

102. Hurlburt, B.K. and Yanofsky, C. *J. Biol. Chem.*, **1990**, *265*, 7853.

103. Riggs, A.D., Suzuki, H. and Bourgeois, *J. Mol. Biol.*, **1970**, *48*, 67.

104. Marmorstein, R.Q. and Sigler, P.B., *J. Biol. Chem.*, **1989**, *264*, 9149.

105. Hurlburt, B.K. and Yanofsky, C. *J. Biol. Chem.*, **1992**, *267*, 16783.

106. Haugland, R.P., *Molecular Probes: Handbook of Fluorescent Probes and Research Chemicals*, **1992**, Molecular Probes, Inc. USA.

107. Ganzler, K., Greve, K.S., Cohen, A.S. and Karger, B.L., *Anal. Chem.*, **1992**, *64*, 2665.

108. Wu, D. and Regnier, F.E., *Anal. Chem.*, **1993**, *65*, 2029.

109. Thromann, W., Molteni, S., Caslavska, J. and Schmutz, J., *Electrophoresis*, **1994**, *15*, 3.

110. Guttman, A., Cohen, A.S., Heiger, D.N. and Karger, B.L., *Anal. Chem.*, **1990**, *62*, 137.

111. Murre, C., McCaw, P.S. and Baltimore, D., *Cell*, **1989**, *56*, 777.

112. Lasser, A.B., Davis, R.L., Wright, W.E., Kadesch, T, Murre, C., Voronova, A., Baltimore, D. and Weintraub, H., *Cell*, **1991**, *66*, 305.

113. Olson, E.N., *Genes and Dev.*, **1990**, *4*, 1035.

APPENDIX

APPENDIX

Complete output for a 50 iteration simplex optimization.

Column Efficiency = 3000000 Plates/Meter, Applied Field = 400 Volts/Cm
Optimal Resolution = 1.5 , Excess Rs Factor = 5 , Overlap Factor = 50
Optimal Time = 200 Seconds, Time Factor = 6

Base Pair No. 1 = 72
Base Pair No. 2 = 118
Base Pair No. 3 = 194
Base Pair No. 4 = 234
Base Pair No. 5 = 271
Base Pair No. 6 = 281
Base Pair No. 7 = 310
Base Pair No. 8 = 603
Base Pair No. 9 = 872
Base Pair No. 10 = 1078
Base Pair No. 11 = 1353

Experiment Number 1
Capillary Length = 40 Centimeters
Methyl Cellulose Concentration = .6 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	2.835259E-04	352.7015	1.287884	0
2	118	2.30151E-04	434.4974	1.58656	56.91252
3	194	1.974227E-04	506.5275	1.849577	41.92502
4	234	1.887363E-04	529.8397	1.934701	12.32055
5	271	1.829844E-04	546.4946	1.995516	8.475311
6	281	1.816899E-04	550.3884	2.009734	1.944347
7	310	1.78408E-04	560.5129	2.046704	4.991854
8	603	1.629561E-04	613.6622	2.240778	24.79278
9	872	1.579128E-04	633.2608	2.312341	8.608855
10	1078	1.557525E-04	642.0445	2.344415	3.77245
11	353	1.538938E-04	649.799	2.37273	3.287775

CRF = -99.70695 SQI =-92.62594 Total Time = 650.9853

Experiment Number 2
Capillary Length = 35 Centimeters
Methyl Cellulose Concentration = .5 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	2.959007E-04	295.7073	1.154323	0
2	118	2.412925E-04	362.6304	1.415564	52.08252
3	194	2.07808E-04	421.0618	1.643657	38.2002
4	234	1.989209E-04	439.8733	1.71709	11.19483
5	271	1.930361E-04	453.283	1.769436	7.692317
6	281	1.917117E-04	456.4146	1.78166	1.763698
7	310	1.88354E-04	464.5509	1.813421	4.526381
8	603	1.72545E-04	507.1141	1.979571	22.44309
9	872	1.673852E-04	522.7464	2.040593	7.776928
10	1078	1.651749E-04	529.7415	2.067899	3.405199
11	1353	1.632733E-04	535.9114	2.091984	2.966363

CRF = -93.66967 SQI =-87.74406 Total Time = 536.9574

Experiment Number 3
Capillary Length = 37 Centimeters
Methyl Cellulose Concentration = .3 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.305725E-04	279.8176	1.062365	0
2	118	2.725087E-04	339.4387	1.288724	50.71782
3	194	2.369054E-04	390.4513	1.4824	36.81727
4	234	2.27456E-04	406.6721	1.543984	10.71963
5	271	2.211988E-04	418.1759	1.58766	7.346782
6	281	2.197905E-04	420.8553	1.597833	1.68224
7	310	2.162203E-04	427.8044	1.624216	4.313465
8	603	1.99411E-04	463.866	1.761129	21.30458
9	872	1.939247E-04	476.9892	1.810953	7.347649
10	1078	1.915746E-04	482.8407	1.833168	3.211442
11	1353	1.895526E-04	487.9912	1.852723	2.794734

CRF = -90.73106 SQI =-85.36781 Total Time = 488.9176

Experiment Number 4
Capillary Length = 32 Centimeters
Methyl Cellulose Concentration = .2 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.580929E-04	223.4057	.9120499	0
2	118	2.972865E-04	269.1007	1.098599	45.45301
3	194	2.600013E-04	307.6908	1.256142	32.7765
4	234	2.501055E-04	319.865	1.305843	9.503717
5	271	2.435528E-04	328.4709	1.340977	6.502841
6	281	2.42078E-04	330.472	1.349146	1.487737

7	310	2.383391E-04	335.6562	1.370311	3.812631
8	603	2.207358E-04	362.4242	1.479591	18.78526
9	872	2.149903E-04	372.1098	1.519132	6.459782
10	1078	2.125292E-04	376.4189	1.536724	2.820253
11	1353	2.104117E-04	380.207	1.552189	2.452718

CRF = -83.53774 SQI =-79.67112 Total Time = 380.9831

Experiment Number 5
Capillary Length = 32 Centimeters
Methyl Cellulose Concentration = .2 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.580929E-04	223.4057	.9120499	0
2	118	2.972865E-04	269.1007	1.098599	45.45301
3	194	2.600013E-04	307.6908	1.256142	32.7765
4	234	2.501055E-04	319.865	1.305843	9.503717
5	271	2.435528E-04	328.4709	1.340977	6.502841
6	281	2.42078E-04	330.472	1.349146	1.487737
7	310	2.383391E-04	335.6562	1.370311	3.812631
8	603	2.207358E-04	362.4242	1.479591	18.78526
9	872	2.149903E-04	372.1098	1.519132	6.459782
10	1078	2.125292E-04	376.4189	1.536724	2.820253
11	1353	2.104117E-04	380.207	1.552189	2.452718

CRF = -83.53774 SQI =-79.67112 Total Time = 380.9831

*** Exp. Values 28 Cm, -2.980232E-08 Percent Rejected

*** Exp. Values 34 Cm, 0 Percent Rejected

Experiment Number 6
Capillary Length = 34.75 Centimeters
Methyl Cellulose Concentration = .375 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.154268E-04	275.4204	1.078992	0
2	118	2.588726E-04	335.5898	1.314712	50.273
3	194	2.241948E-04	387.4979	1.518068	36.64817
4	234	2.14991E-04	404.0866	1.583057	10.69853
5	271	2.088965E-04	415.8758	1.629242	7.339992
6	281	2.075249E-04	418.6245	1.640011	1.681595
7	310	2.040475E-04	425.7587	1.66796	4.313332
8	603	1.876751E-04	462.9009	1.813469	21.33734
9	872	1.823315E-04	476.4674	1.866617	7.372909
10	1078	1.800424E-04	482.5251	1.890349	3.224788

11	1353	1.78073E-04	487.8617	1.911255	2.80754

CRF = -90.71465 SQI =-85.35262 Total Time = 488.8173

Experiment Number 7
Capillary Length = 29.75 Centimeters
Methyl Cellulose Concentration = .275 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.364783E-04	221.0395	.9358922	0
2	118	2.77826E-04	267.7036	1.13347	45.0999
3	194	2.418617E-04	307.5105	1.302014	32.68914
4	234	2.323165E-04	320.1452	1.35551	9.508603
5	271	2.259959E-04	329.0989	1.393421	6.514339
6	281	2.245733E-04	331.1836	1.402247	1.491347
7	310	2.20967E-04	336.5888	1.425133	3.823477
8	603	2.039872E-04	364.6061	1.54376	18.87392
9	872	1.984453E-04	374.7884	1.586872	6.504916
10	1078	1.960714E-04	379.3262	1.606086	2.842387
11	1353	1.940289E-04	383.3192	1.622992	2.473178

CRF = -83.57819 SQI =-79.6622 Total Time = 384.1307

*** Exp. Values 27 Cm, 9.999996E-02 Percent Rejected

Experiment Number 8
Capillary Length = 32.8125 Centimeters
Methyl Cellulose Concentration = .30625 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.29173E-04	249.2041	1.004696	0
2	118	2.712487E-04	302.4208	1.219246	47.85797
3	194	2.357309E-04	347.9869	1.402951	34.75414
4	234	2.263042E-04	362.4823	1.461391	10.12127
5	271	2.20062E-04	372.7642	1.502844	6.937338
6	281	2.186571E-04	375.1593	1.5125	1.588575
7	310	2.150955E-04	381.3712	1.537544	4.073367
8	603	1.983266E-04	413.617	1.667547	20.12158
9	872	1.928535E-04	425.3553	1.714872	6.940798
10	1078	1.90509E-04	430.59	1.735976	3.033817
11	1353	1.884918E-04	435.1979	1.754553	2.640236

CRF = -87.18024 SQI =-82.50325 Total Time = 436.0752

Experiment Number 9
Capillary Length = 28.9375 Centimeters
Methyl Cellulose Concentration = .16875 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.696247E-04	195.7222	.8402508	0
2	118	3.076689E-04	235.1351	1.009453	42.61534
3	194	2.69679E-04	268.2588	1.151656	30.65434
4	234	2.595962E-04	278.678	1.196387	8.874839
5	271	2.529196E-04	286.0346	1.227969	6.068885
6	281	2.514169E-04	287.7441	1.235308	1.38804
7	310	2.476074E-04	292.1712	1.254314	3.556389
8	603	2.296714E-04	314.9881	1.352269	17.50718
9	872	2.238173E-04	323.2268	1.387638	6.013838
10	1078	2.213096E-04	326.8893	1.403362	2.624512
11	1353	2.191521E-04	330.1075	1.417177	2.281954

CRF = -83.0358 SQI =-80.01633 Total Time = 330.8161

*** Exp. Values 27 Cm, 9.999999E-02 Percent Rejected

*** Exp. Values 31.1875 Cm, .09375 Percent Rejected

Experiment Number 10
Capillary Length = 30.10938 Centimeters
Methyl Cellulose Concentration = .2296875 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.48699E-04	215.8694	.9085305	0
2	118	2.888288E-04	260.6161	1.096856	44.62657
3	194	2.521177E-04	298.5647	1.256571	32.24963
4	234	2.423743E-04	310.567	1.307085	9.363394
5	271	2.359224E-04	319.0601	1.34283	6.410141
6	281	2.344703E-04	321.0361	1.351146	1.466941
7	310	2.307891E-04	326.1569	1.372698	3.759972
8	603	2.134567E-04	352.6402	1.484159	18.54021
9	872	2.077997E-04	362.2403	1.524562	6.381473
10	1078	2.053765E-04	366.5144	1.542551	2.787056
11	1353	2.032916E-04	370.2733	1.558371	2.424363

CRF = -83.45068 SQI =-79.74252 Total Time = 371.0524

Experiment Number 11
Capillary Length = 27.04688 Centimeters
Methyl Cellulose Concentration = .1984375 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.586253E-04	188.5455	.8372536	0
2	118	2.977658E-04	227.0818	1.008378	41.75948

3	194	2.60448E-04	259.6187	1.152861	30.10954
4	234	2.505436E-04	269.8819	1.198435	8.729758
5	271	2.439852E-04	277.1364	1.23065	5.973095
6	281	2.425091E-04	278.8233	1.23814	1.366547
7	310	2.38767E-04	283.1932	1.257545	3.501942
8	603	2.211483E-04	305.755	1.357733	17.25384
9	872	2.153978E-04	313.9177	1.39398	5.932827
10	1078	2.129345E-04	317.5492	1.410106	2.590148
11	1353	2.108152E-04	320.7416	1.424282	2.252608

CRF = -82.94231 SQI =-80.09509 Total Time = 321.4537

Experiment Number 12
Capillary Length = 24.57031 Centimeters
Methyl Cellulose Concentration = .1976562 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.588931E-04	171.1534	.7974063	0
2	118	2.980068E-04	206.1221	.9603256	39.78835
3	194	2.606727E-04	235.6433	1.097865	28.68659
4	234	2.50764E-04	244.9546	1.141247	8.316925
5	271	2.442027E-04	251.5361	1.17191	5.690502
6	281	2.427259E-04	253.0664	1.17904	1.301899
7	310	2.389822E-04	257.0308	1.19751	3.33624
8	603	2.213558E-04	277.498	1.292867	16.43705
9	872	2.156028E-04	284.9026	1.327365	5.651821
10	1078	2.131384E-04	288.1967	1.342713	2.467463
11	1353	2.110181E-04	291.0925	1.356204	2.145863

CRF = -82.60296 SQI =-80.33706 Total Time = 291.7706

Experiment Number 13
Capillary Length = 23.39844 Centimeters
Methyl Cellulose Concentration = .1367187 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.839116E-04	152.3686	.7274473	0
2	118	3.20532E-04	182.4969	.8712875	37.69016
3	194	2.816689E-04	207.6768	.9915025	27.03455
4	234	2.713544E-04	215.5708	1.029191	7.813218
5	271	2.645244E-04	221.1369	1.055765	5.339242
6	281	2.629872E-04	222.4294	1.061936	1.220746
7	310	2.590901E-04	225.7751	1.077909	3.126996
8	603	2.407419E-04	242.9826	1.160062	15.37782
9	872	2.347533E-04	249.1812	1.189655	5.275981
10	1078	2.32188E-04	251.9342	1.202799	2.301433

11	1353	2.299809E-04	254.352	1.214342	2.000546

CRF = -82.11358 SQI =-80.65686 Total Time = 254.9592

*** Exp. Values 20.04297 Cm, 9.023434E-02 Percent Rejected

Experiment Number 14
Capillary Length = 19.03125 Centimeters
Methyl Cellulose Concentration = .165625 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.708934E-04	128.2798	.6790847	0
2	118	3.088112E-04	154.0687	.8156053	34.50733
3	194	2.707437E-04	175.7312	.9302821	24.81554
4	234	2.606403E-04	182.5432	.9663432	7.183265
5	271	2.539501E-04	187.3522	.9918011	4.911815
6	281	2.524444E-04	188.4697	.9977168	1.12338
7	310	2.486271E-04	191.3634	1.013035	2.878188
8	603	2.306544E-04	206.2745	1.091971	14.16727
9	872	2.247884E-04	211.6573	1.120467	4.865996
10	1078	2.222756E-04	214.0501	1.133134	2.123501
11	1353	2.201137E-04	216.1525	1.144263	1.84628

CRF = -81.5553 SQI =-81.07344 Total Time = 216.7246

Experiment Number 15
Capillary Length = 14.07813 Centimeters
Methyl Cellulose Concentration = .1640625 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.715367E-04	94.72903	.5830557	0
2	118	3.093904E-04	113.757	.7001724	29.65636
3	194	2.712836E-04	129.7362	.7985243	21.32418
4	234	2.611698E-04	134.7603	.8294473	6.172157
5	271	2.544727E-04	138.3068	.8512765	4.220307
6	281	2.529654E-04	139.1309	.8563488	.9651962
7	310	2.491442E-04	141.2648	.8694828	2.47289
8	603	2.31153E-04	152.2598	.9371569	12.17176
9	872	2.252809E-04	156.2286	.9615844	4.180401
10	1078	2.227655E-04	157.9926	.9724423	1.824244
11	1353	2.206014E-04	159.5426	.9819822	1.586085

CRF = -80.49503 SQI =-81.83263 Total Time = 160.0336

Experiment Number 16
Capillary Length = 12.90625 Centimeters
Methyl Cellulose Concentration = .103125 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	4.03051E-04	80.05346	.5146114	0
2	118	3.377638E-04	95.52718	.6140819	27.41882
3	194	2.977312E-04	108.3717	.6966507	19.59895
4	234	2.871063E-04	112.3822	.7224316	5.652253
5	271	2.800706E-04	115.2053	.7405797	3.859349
6	281	2.784872E-04	115.8604	.7447906	.8820052
7	310	2.744728E-04	117.5549	.7556837	2.258665
8	603	2.555723E-04	126.2485	.8115692	11.09407
9	872	2.494035E-04	129.3712	.8316428	3.800694
10	1078	2.46761E-04	130.7566	.8405487	1.656995
11	1353	2.444875E-04	131.9725	.8483651	1.439893

CRF = -82.05541 SQI =-84.5305 Total Time = 132.3967

*** Exp. Values 3.585938 Cm, .1304687 Percent Rejected

Experiment Number 17
Capillary Length = 18.44531 Centimeters□
Methyl Cellulose Concentration = .1351562 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.846918E-04	119.8707	.6445691	0
2	118	3.212344E-04	143.5503	.7718989	33.43465
3	194	2.823237E-04	163.3348	.8782842	23.97857
4	234	2.719965E-04	169.5363	.9116309	6.929387
5	271	2.651581E-04	173.9086	.9351419	4.735103
6	281	2.63619E-04	174.924	.9406015	1.082595
7	310	2.597172E-04	177.5519	.9547325	2.773068
8	603	2.413464E-04	191.0668	1.027405	13.63667
9	872	2.353505E-04	195.9345	1.05358	4.678317
10	1078	2.32782E-04	198.0964	1.065204	2.040666
11	1353	2.305722E-04	199.995	1.075413	1.773857

CRF = -81.27332 SQI =-81.25736 Total Time = 200.5327

Experiment Number 18
Capillary Length = 19.61719 Centimeters
Methyl Cellulose Concentration = .1960937 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.594318E-04	136.4458	.711445	0
2	118	2.984918E-04	164.3026	.8566931	35.52838
3	194	2.611248E-04	187.8143	.9792862	25.6122
4	234	2.512074E-04	195.229	1.017948	7.425019

5	271	2.446402E-04	200.4698	1.045273	5.080152
6	281	2.431622E-04	201.6883	1.051627	1.162212
7	310	2.394152E-04	204.8449	1.068086	2.978315
8	603	2.217732E-04	221.1402	1.153051	14.67297
9	872	2.160151E-04	227.0349	1.183787	5.045007
10	1078	2.135486E-04	229.6572	1.19746	2.202484
11	1353	2.114264E-04	231.9624	1.209479	1.915403

CRF = -81.81062 SQI =-80.90545 Total Time = 232.5671

Experiment Number 19
Capillary Length = 17.93945 Centimeters
Methyl Cellulose Concentration = .1728515 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.679947E-04	121.873	.6645114	0
2	118	3.062014E-04	146.4678	.7986138	33.61944
3	194	2.683111E-04	167.1516	.9113923	24.19155
4	234	2.582547E-04	173.6605	.9468818	7.005264
5	271	2.515956E-04	178.2568	.9719433	4.79079
6	281	2.500969E-04	179.325	.9777676	1.095759
7	310	2.462974E-04	182.0914	.9928513	2.807618
8	603	2.284083E-04	196.3529	1.070612	13.82285
9	872	2.225696E-04	201.5038	1.098697	4.748936
10	1078	2.200685E-04	203.7939	1.111184	2.0726
11	1353	2.179167E-04	205.8063	1.122157	1.802148

CRF = -81.38714 SQI =-81.1991 Total Time = 206.3674

Experiment Number 20
Capillary Length = 14.58398 Centimeters
Methyl Cellulose Concentration = .1263672 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.892556E-04	93.66586	.5664252	0
2	118	3.253433E-04	112.0661	.677697	29.57949
3	194	2.861538E-04	127.4139	.7705096	21.19556
4	234	2.757526E-04	132.2198	.7995726	6.121923
5	271	2.688651E-04	135.6069	.820055	4.182472
6	281	2.67315E-04	136.3932	.8248103	.9561411
7	310	2.633852E-04	138.4283	.8371168	2.449005
8	603	2.448827E-04	148.8874	.9003664	12.03943
9	872	2.388438E-04	152.6519	.9231312	4.12884
10	1078	2.36257E-04	154.3233	.9332389	1.800753
11	1353	2.340313E-04	155.7909	.9421141	1.565168

CRF = -80.39771 SQI =-81.8784 Total Time = 156.262

Experiment Number 21
Capillary Length = 12.90625 Centimeters
Methyl Cellulose Concentration = .103125 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	4.03051E-04	80.05346	.5146114	0
2	118	3.377638E-04	95.52718	.6140819	27.41882
3	194	2.977312E-04	108.3717	.6966507	19.59895
4	234	2.871063E-04	112.3822	.7224316	5.652253
5	271	2.800706E-04	115.2053	.7405797	3.859349
6	281	2.784872E-04	115.8604	.7447906	.8820052
7	310	2.744728E-04	117.5549	.7556837	2.258665
8	603	2.555723E-04	126.2485	.8115692	11.09407
9	872	2.494035E-04	129.3712	.8316428	3.800694
10	1078	2.46761E-04	130.7566	.8405487	1.656995
11	1353	2.444875E-04	131.9725	.8483651	1.439893

CRF = -82.05541 SQI =-84.5305 Total Time = 132.3967

Experiment Number 22
Capillary Length = 10.2168 Centimeters
Methyl Cellulose Concentration = .1552734 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.752739E-04	68.06227	.4917546	0
2	118	3.127551E-04	81.66772	.5900549	25.15313
3	194	2.744199E-04	93.0763	.6724827	18.07247
4	234	2.642455E-04	96.66007	.6983757	5.228506
5	271	2.575082E-04	99.18903	.7166476	3.574434
6	281	2.559919E-04	99.77656	.7208925	.8174131
7	310	2.521478E-04	101.2977	.7318829	2.094128
8	603	2.340487E-04	109.1311	.7884796	10.30461
9	872	2.281415E-04	111.9568	.8088956	3.537956
10	1078	2.25611E-04	113.2125	.8179682	1.543724
11	1353	2.234339E-04	114.3157	.8259385	1.342088

CRF = -85.4465 SQI =-88.78099 Total Time = 114.7286

Experiment Number 23
Capillary Length = 16.38818 Centimeters
Methyl Cellulose Concentration = .1401855 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.82212E-04	107.193	.6115057	0
2	118	3.190017E-04	128.4333	.7326757	31.60336

3	194	2.802426E-04	146.1964	.8340089	22.67599
4	234	2.699556E-04	151.7674	.8657897	6.554858
5	271	2.631439E-04	155.696	.8882016	4.479686
6	281	2.616108E-04	156.6085	.8934067	1.024276
7	310	2.577241E-04	158.9702	.9068798	2.623738
8	603	2.394249E-04	171.1203	.9761925	12.90453
9	872	2.334523E-04	175.4982	1.001167	4.42804
10	1078	2.308938E-04	177.4428	1.012261	1.931657
11	1353	2.286926E-04	179.1508	1.022004	1.679148

CRF = -80.89021 SQI =-81.53365 Total Time = 179.6618

Experiment Number 24
Capillary Length = 12.27393 Centimeters
Methyl Cellulose Concentration = .1502441 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.775087E-04	81.2824	.5358018	0
2	118	3.147672E-04	97.48417	.6426015	27.49783
3	194	2.762954E-04	111.058	.7320781	19.74833
4	234	2.660848E-04	115.3197	.7601706	5.711797
5	271	2.593235E-04	118.3264	.7799903	3.904385
6	281	2.578018E-04	119.0248	.7845944	.8928224
7	310	2.53944E-04	120.833	.7965137	2.287234
8	603	2.357804E-04	130.1415	.8578737	11.25306
9	872	2.298521E-04	133.4981	.8799998	3.862862
10	1078	2.273127E-04	134.9895	.8898309	1.685369
11	1353	2.251278E-04	136.2995	.8984668	1.465163

CRF = -81.23286 SQI =-83.51389 Total Time = 136.7488

Experiment Number 25
Capillary Length = 15.35962 Centimeters
Methyl Cellulose Concentration = .1427002 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.810053E-04	100.7835	.5938799	0
2	118	3.179153E-04	120.7839	.7117349	30.63753
3	194	2.792299E-04	137.5177	.8103409	21.98811
4	234	2.689625E-04	142.7673	.8412749	6.356931
5	271	2.621637E-04	146.4698	.863092	4.344667
6	281	2.606335E-04	147.3297	.8681592	.9934137
7	310	2.567542E-04	149.5557	.8812761	2.544801
8	603	2.384898E-04	161.0092	.9487674	12.5172
9	872	2.325286E-04	165.1369	.9730905	4.295547
10	1078	2.29975E-04	166.9705	.9838954	1.873936

11	1353	2.27778E-04	168.581	.9933856	1.629009

CRF = -80.67983 SQI =-81.68758 Total Time = 169.0777

Experiment Number 26
Capillary Length = 13.30249 Centimeters
Methyl Cellulose Concentration = .1477295 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.786543E-04	87.82741	.5561129	0
2	118	3.157986E-04	105.3083	.6668001	28.58899
3	194	2.772569E-04	119.9473	.7594926	20.52736
4	234	2.670276E-04	124.5423	.788587	5.93628
5	271	2.602541E-04	127.7837	.8091114	4.057618
6	281	2.587296E-04	128.5366	.813879	.9278584
7	310	2.548647E-04	130.4858	.8262207	2.376858
8	603	2.366681E-04	140.5184	.889746	11.69324
9	872	2.307291E-04	144.1354	.9126486	4.013571
10	1078	2.28185E-04	145.7424	.9228238	1.751042
11	1353	2.259961E-04	147.154	.9317618	1.522254

CRF = -80.20575 SQI =-82.02781 Total Time = 147.6199

Experiment Number 27
Capillary Length = 12.27393 Centimeters
Methyl Cellulose Concentration = .1502441 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.775087E-04	81.2824	.5358018	0
2	118	3.147672E-04	97.48416	.6426014	27.49782
3	194	2.762954E-04	111.058	.7320781	19.74834
4	234	2.660848E-04	115.3197	.7601706	5.711787
5	271	2.593235E-04	118.3264	.7799903	3.904395
6	281	2.578018E-04	119.0248	.7845944	.8928224
7	310	2.53944E-04	120.833	.7965137	2.287234
8	603	2.357804E-04	130.1415	.8578737	11.25306
9	872	2.298521E-04	133.4981	.8799998	3.862862
10	1078	2.273127E-04	134.9895	.8898309	1.685369
11	1353	2.251278E-04	136.2995	.8984668	1.465163

CRF = -81.23286 SQI =-83.51389 Total Time = 136.7488

Experiment Number 28
Capillary Length = 13.80835 Centimeters
Methyl Cellulose Concentration = .1100341 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution

1	72	3.986494E-04	86.59457	.5381696	0
2	118	3.338009E-04	103.4175	.6427214	28.492
3	194	2.940373E-04	117.4031	.7296387	20.38168
4	234	2.834837E-04	121.7737	.7568017	5.880736
5	271	2.764954E-04	124.8515	.7759296	4.016085
6	281	2.749226E-04	125.5658	.7803687	.9179206
7	310	2.709352E-04	127.4138	.7918534	2.350754
8	603	2.521617E-04	136.8997	.8508071	11.54953
9	872	2.460344E-04	140.3092	.871996	3.957996
10	1078	2.434096E-04	141.8222	.881399	1.725795
11	1353	2.411513E-04	143.1503	.8896529	1.49978

CRF = -80.10154 SQI =-82.08946 Total Time = 143.5951

*** Exp. Values 13.67346 Cm, 8.301997E-02 Percent Rejected

Experiment Number 29
Capillary Length = 12.52686 Centimeters
Methyl Cellulose Concentration = .1313964 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.866067E-04	81.00517	.5285561	0
2	118	3.229584E-04	96.96957	.6327233	27.4945
3	194	2.839307E-04	110.2985	.7196943	19.71129
4	234	2.735725E-04	114.4747	.746944	5.694951
5	271	2.667135E-04	117.4186	.7661529	3.891236
6	281	2.651698E-04	118.1022	.7706131	.8896226
7	310	2.612562E-04	119.8714	.7821568	2.278702
8	603	2.428302E-04	128.9673	.8415071	11.20415
9	872	2.368162E-04	132.2424	.8628772	3.843184
10	1078	2.3424E-04	133.6968	.8723672	1.676313
11	1353	2.320236E-04	134.974	.8807006	1.457058

CRF = -81.49633 SQI =-83.8362 Total Time = 135.4143

Experiment Number 30
Capillary Length = 14.0697 Centimeters
Methyl Cellulose Concentration = .1276245 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.885836E-04	90.51914	.5573106	0
2	118	3.247383E-04	108.3157	.6668809	29.07478
3	194	2.855898E-04	123.1636	.7582966	20.83651
4	234	2.751995E-04	127.8137	.7869265	6.018689
5	271	2.683193E-04	131.091	.8071048	4.112058
6	281	2.667708E-04	131.852	.8117898	.9400743

7	310	2.628451E-04	133.8212	.8239141	2.407836
8	603	2.44362E-04	143.9432	.8862333	11.83755
9	872	2.383295E-04	147.5867	.9086656	4.059814
10	1078	2.357453E-04	149.2045	.9186261	1.770706
11	1353	2.33522E-04	150.625	.9273722	1.539056

CRF = -80.27975 SQI =-81.96246 Total Time = 151.0887

Experiment Number 31
Capillary Length = 13.04114 Centimeters
Methyl Cellulose Concentration = .1301391 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.872593E-04	84.18867	.5383879	0
2	118	3.23546E-04	100.7673	.6444083	28.03289
3	194	2.844784E-04	114.6057	.7329053	20.09481
4	234	2.741096E-04	118.9409	.7606291	5.805317
5	271	2.672436E-04	121.9967	.780171	3.966522
6	281	2.656983E-04	122.7063	.7847085	.9068221
7	310	2.617807E-04	124.5426	.7964517	2.322739
8	603	2.433359E-04	133.9829	.8568227	11.42015
9	872	2.373157E-04	137.3817	.8785583	3.917104
10	1078	2.347369E-04	138.891	.8882101	1.7085
11	1353	2.325182E-04	140.2163	.8966855	1.485036

CRF = -80.58276 SQI =-82.69439 Total Time = 140.6647

Experiment Number 32
Capillary Length = 13.81256 Centimeters
Methyl Cellulose Concentration = .1282531 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.882501E-04	88.94113	.5526688	0
2	118	3.244381E-04	106.4345	.6613703	28.81846
3	194	2.853099E-04	121.0312	.7520723	20.65411
4	234	2.74925E-04	125.603	.7804808	5.966224
5	271	2.680484E-04	128.8253	.8005036	4.076289
6	281	2.665007E-04	129.5734	.8051524	.9318789
7	310	2.625771E-04	131.5096	.8171837	2.386932
8	603	2.441036E-04	141.4621	.8790271	11.73494
9	872	2.380742E-04	145.0447	.9012893	4.024742
10	1078	2.354914E-04	146.6355	.9111743	1.755406
11	1353	2.332692E-04	148.0324	.9198543	1.525788

CRF = -80.22066 SQI =-82.00736 Total Time = 148.4923

Experiment Number 33

Capillary Length = 13.29828 Centimeters
Methyl Cellulose Concentration = .1295105 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.875879E-04	85.77588	.5432088	0
2	118	3.238419E-04	102.6603	.6501358	28.2976
3	194	2.847542E-04	116.7523	.7393787	20.28333
4	234	2.743801E-04	121.1666	.7673342	5.85956
5	271	2.675105E-04	124.2781	.7870389	4.003528
6	281	2.659645E-04	125.0005	.7916141	.9152647
7	310	2.620449E-04	126.8702	.8034548	2.344376
8	603	2.435905E-04	136.4819	.8643242	11.52628
9	872	2.375673E-04	139.9422	.886238	3.953389
10	1078	2.349872E-04	141.4788	.8959689	1.724334
11	1353	2.327673E-04	142.828	.9045135	1.498757

CRF = -80.14155 SQI =-82.14264 Total Time = 143.2803

*** Exp. Values 13.80414 Cm, 9.181514E-02 Percent Rejected

Experiment Number 34
Capillary Length = 13.4279 Centimeters
Methyl Cellulose Concentration = .1337509 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.854012E-04	87.1034	.548947	0
2	118	3.218731E-04	104.295	.6572926	28.50445
3	194	2.829191E-04	118.655	.7477927	20.44001
4	234	2.725804E-04	123.1554	.7761557	5.906312
5	271	2.657344E-04	126.3282	.7961516	4.03587
6	281	2.641935E-04	127.065	.8007948	.9227148
7	310	2.602874E-04	128.9719	.8128124	2.363503
8	603	2.418961E-04	138.7776	.8746102	11.62208
9	872	2.358935E-04	142.309	.896866	3.986943
10	1078	2.333222E-04	143.8773	.9067497	1.73906
11	1353	2.311099E-04	145.2545	.9154295	1.511644

CRF = -80.15607 SQI =-82.05618 Total Time = 145.7122

Experiment Number 35
Capillary Length = 13.67873 Centimeters
Methyl Cellulose Concentration = .1057937 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	4.013168E-04	85.21152	.5320774	0
2	118	3.362025E-04	101.7149	.635128	28.27851

3	194	2.962759E-04	115.4222	.7207189	20.2195
4	234	2.85679E-04	119.7036	.7474529	5.832301
5	271	2.78662E-04	122.7179	.7662745	3.982566
6	281	2.770828E-04	123.4173	.770642	.9101861
7	310	2.730791E-04	125.2268	.7819407	2.330927
8	603	2.542286E-04	134.5121	.8399197	11.45013
9	872	2.480761E-04	137.8481	.8607504	3.923176
10	1078	2.454406E-04	139.3283	.8699931	1.71049
11	1353	2.431731E-04	140.6275	.8781054	1.486391

CRF = -80.53132 SQI =-82.62583 Total Time = 141.0665

Experiment Number 36
Capillary Length = 13.49061 Centimeters
Methyl Cellulose Concentration = .1267616 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.890441E-04	86.69074	.5450751	0
2	118	3.251529E-04	103.7251	.65218	28.4557
3	194	2.859762E-04	117.9347	.7415239	20.39111
4	234	2.755785E-04	122.3844	.7695021	5.88971
5	271	2.686933E-04	125.5205	.7892204	4.023888
6	281	2.671437E-04	126.2486	.7937983	.919883
7	310	2.632152E-04	128.1329	.8056458	2.356156
8	603	2.447189E-04	137.8174	.8665382	11.5831
9	872	2.386819E-04	141.3032	.8884554	3.97243
10	1078	2.360959E-04	142.8509	.8981867	1.73254
11	1353	2.33871E-04	144.2099	.9067317	1.505895

CRF = -80.12849 SQI =-82.07195 Total Time = 144.6633

Experiment Number 37
Capillary Length = 14.00068 Centimeters
Methyl Cellulose Concentration = .1072853 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	4.003666E-04	87.42413	.5395803	0
2	118	3.35347E-04	104.3746	.6441983	28.63788
3	194	2.954784E-04	118.4577	.7311192	20.47985
4	234	2.84897E-04	122.8574	.7582739	5.907998
5	271	2.778902E-04	125.9551	.7773931	4.03439
6	281	2.763132E-04	126.674	.7818299	.9220706
7	310	2.723153E-04	128.5337	.793308	2.361321
8	603	2.534923E-04	138.078	.852215	11.60028
9	872	2.473487E-04	141.5075	.8733819	3.974879
10	1078	2.447171E-04	143.0292	.8827742	1.733057

11	1353	2.424529E-04	144.365	.8910183	1.506065

CRF = -80.12148 SQI =-82.05883 Total Time = 144.8105

*** Exp. Values 14.31842 Cm, 9.055781E-02 Percent Rejected

Experiment Number 38
Capillary Length = 13.69756 Centimeters
Methyl Cellulose Concentration = .1177106 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.940721E-04	86.89756	.5422322	0
2	118	3.296798E-04	103.8702	.6481397	28.51651
3	194	2.901959E-04	118.0027	.7363254	20.41588
4	234	2.797166E-04	122.4236	.7639111	5.893541
5	271	2.727774E-04	125.5379	.7833441	4.025614
6	281	2.712156E-04	126.2608	.787855	.9201894
7	310	2.672563E-04	128.1313	.7995268	2.356729
8	603	2.486149E-04	137.7387	.8594762	11.58216
9	872	2.425306E-04	141.1942	.8810377	3.970567
10	1078	2.399243E-04	142.7279	.8906083	1.731478
11	1353	2.37682E-04	144.0745	.8990106	1.504837

CRF = -80.12009 SQI =-82.06932 Total Time = 144.524

Experiment Number 39
Capillary Length = 13.50523 Centimeters
Methyl Cellulose Concentration = .1204595 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.925053E-04	86.01943	.5405613	0
2	118	3.282691E-04	102.8518	.646339	28.36363
3	194	2.88881E-04	116.8754	.7344657	20.31219
4	234	2.784271E-04	121.2636	.7620421	5.864626
5	271	2.715047E-04	124.3554	.7814713	4.006141
6	281	2.699467E-04	125.0731	.7859815	.9157691
7	310	2.65997E-04	126.9303	.7976523	2.345454
8	603	2.474008E-04	136.4712	.8576089	11.52795
9	872	2.413313E-04	139.9034	.8791779	3.952432
10	1078	2.387313E-04	141.4271	.8887529	1.723672
11	1353	2.364944E-04	142.7648	.8971592	1.498066

CRF = -80.16049 SQI =-82.16438 Total Time = 143.2134

Experiment Number 40
Capillary Length = 13.87682 Centimeters
Methyl Cellulose Concentration = .1105788 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.983142E-04	87.09718	.5399562	0
2	118	3.334992E-04	104.0244	.6448958	28.57269
3	194	2.93756E-04	118.0982	.7321457	20.4406
4	234	2.832079E-04	122.4967	.7594146	5.897971
5	271	2.762232E-04	125.5943	.7786176	4.027889
6	281	2.746512E-04	126.3131	.7830741	.9206227
7	310	2.706658E-04	128.173	.7946042	2.357712
8	603	2.51902E-04	137.7204	.8537931	11.58387
9	872	2.457778E-04	141.1521	.8750678	3.969881
10	1078	2.431544E-04	142.675	.8845088	1.730971
11	1353	2.408973E-04	144.0117	.8927962	1.504291

CRF = -80.11496 SQI =-82.06693 Total Time = 144.4581

Experiment Number 41
Capillary Length = 13.98761 Centimeters
Methyl Cellulose Concentration = .1029023 Percent□

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	4.031977E-04	86.72921	.5355414	0
2	118	3.378959E-04	103.4905	.6390401	28.54001
3	194	2.978543E-04	117.4031	.7249486	20.39987
4	234	2.87227E-04	121.747	.7517715	5.883146
5	271	2.801898E-04	124.8047	.7706528	4.016974
6	281	2.78606E-04	125.5142	.7750337	.9180132
7	310	2.745907E-04	127.3496	.7863668	2.350912
8	603	2.55686E-04	136.7655	.8445086	11.54702
9	872	2.495158E-04	140.1475	.8653923	3.955825
10	1078	2.468727E-04	141.6479	.8746574	1.724615
11	1353	2.445987E-04	142.9648	.882789	1.498645

CRF = -80.13557 SQI =-82.13139 Total Time = 143.4062

Experiment Number 42
Capillary Length = 13.77007 Centimeters
Methyl Cellulose Concentration = .1140086 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.962411E-04	86.87939	.5406896	0
2	118	3.316326E-04	103.8052	.6460264	28.52541
3	194	2.920161E-04	117.8879	.7336698	20.41431
4	234	2.815016E-04	122.2912	.7610734	5.89169
5	271	2.745392E-04	125.3926	.7803747	4.023982
6	281	2.729722E-04	126.1124	.7848544	.9197509

7	310	2.689996E-04	127.9749	.7964453	2.355584
8	603	2.502956E-04	137.5381	.8559618	11.57496
9	872	2.441909E-04	140.9765	.8773606	3.967427
10	1078	2.415758E-04	142.5026	.8868579	1.729998
11	1353	2.393259E-04	143.8423	.8951952	1.503502

CRF = -80.11355 SQI =-82.07251 Total Time = 144.2899

Experiment Number 43
Capillary Length = 13.7016 Centimeters
Methyl Cellulose Concentration = .1134639 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.965661E-04	86.37655	.5389016	0
2	118	3.319253E-04	103.198	.6438501	28.44455
3	194	2.922889E-04	117.1923	.7311605	20.35523
4	234	2.817692E-04	121.5676	.7584581	5.874437
5	271	2.748032E-04	124.6492	.7776843	4.012143
6	281	2.732354E-04	125.3645	.7821465	.9170353
7	310	2.692608E-04	127.215	.7936919	2.348642
8	603	2.505474E-04	136.7167	.8529727	11.54051
9	872	2.444397E-04	140.1328	.8742858	3.955525
10	1078	2.418233E-04	141.6489	.883745	1.724821
11	1353	2.395723E-04	142.9799	.8920486	1.498972

CRF = -80.13023 SQI =-82.12523 Total Time = 143.4259

Experiment Number 44
Capillary Length = 13.83301 Centimeters
Methyl Cellulose Concentration = .1113001 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.97873E-04	86.91854	.5397012	0
2	118	3.331019E-04	103.8197	.6446451	28.5409
3	194	2.933857E-04	117.874	.731912	20.41947
4	234	2.828447E-04	122.2669	.7591886	5.892125
5	271	2.758647E-04	125.3605	.7783978	4.024005
6	281	2.742938E-04	126.0785	.7828559	.9197375
7	310	2.703112E-04	127.936	.7943901	2.355466
8	603	2.515601E-04	137.4723	.8536032	11.57314
9	872	2.4544E-04	140.9001	.8748879	3.966335
10	1078	2.428184E-04	142.4214	.8843338	1.729464
11	1353	2.405628E-04	143.7568	.8926254	1.502969

CRF = -80.10932 SQI =-82.07189 Total Time = 144.2031

Experiment Number 45

Capillary Length = 13.87129 Centimeters
Methyl Cellulose Concentration = .1073257 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	4.00341E-04	86.62173	.5371156	0
2	118	3.35324E-04	103.4171	.6412587	28.50599
3	194	2.954569E-04	117.3715	.7277859	20.38565
4	234	2.848759E-04	121.731	.7548177	5.880826
5	271	2.778695E-04	124.8004	.7738504	4.015849
6	281	2.762925E-04	125.5127	.7782671	.9178371
7	310	2.722948E-04	127.3555	.7896934	2.350477
8	603	2.534725E-04	136.8126	.8483342	11.54697
9	872	2.473292E-04	140.2108	.8694057	3.95664
10	1078	2.446976E-04	141.7187	.8787556	1.725097
11	1353	2.424335E-04	143.0422	.8869624	1.499143

CRF = -80.12054 SQI =-82.11304 Total Time = 143.4857

Experiment Number 46
Capillary Length = 13.79538 Centimeters
Methyl Cellulose Concentration = .1123378 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.972431E-04	86.8195	.5398211	0
2	118	3.325347E-04	103.7138	.6448655	28.52115
3	194	2.92857E-04	117.7655	.7322351	20.4076
4	234	2.823263E-04	122.1581	.7595473	5.889105
5	271	2.753531E-04	125.2517	.7787826	4.022041
6	281	2.737837E-04	125.9697	.7832468	.9193003
7	310	2.698049E-04	127.8273	.7947972	2.354362
8	603	2.51072E-04	137.3648	.8540984	11.56826
9	872	2.449579E-04	140.7934	.8754166	3.964833
10	1078	2.423388E-04	142.315	.8848777	1.728827
11	1353	2.400854E-04	143.6507	.893183	1.502467

CRF = -80.10745 SQI =-82.07442 Total Time = 144.0973

Experiment Number 47
Capillary Length = 13.77071 Centimeters
Methyl Cellulose Concentration = .1110719 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.980123E-04	86.49678	.5382959	0
2	118	3.332273E-04	103.3132	.6429497	28.47237
3	194	2.935026E-04	117.2963	.7299711	20.3699
4	234	2.829594E-04	121.6669	.7571701	5.877759

5	271	2.759779E-04	124.7447	.7763245	4.014155
6	281	2.744066E-04	125.459	.7807698	.9174808
7	310	2.704232E-04	127.3071	.7922709	2.349667
8	603	2.51668E-04	136.7944	.8513135	11.5447
9	872	2.455467E-04	140.2047	.8725364	3.956531
10	1078	2.429245E-04	141.718	.8819548	1.725169
11	1353	2.406684E-04	143.0465	.8902223	1.499259

CRF = -80.11925 SQI =-82.11149 Total Time = 143.4916

Experiment Number 48
Capillary Length = 13.81744 Centimeters
Methyl Cellulose Concentration = .111243 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.979078E-04	86.81307	.53935	0
2	118	3.331332E-04	103.693	.6442215	28.52379
3	194	2.934149E-04	117.7295	.731427	20.40708
4	234	2.828734E-04	122.1168	.7586843	5.888541
5	271	2.75893E-04	125.2065	.7778798	4.021549
6	281	2.74322E-04	125.9236	.7823346	.9191668
7	310	2.703391E-04	127.7788	.7938606	2.354025
8	603	2.51587E-04	137.3028	.8530311	11.56604
9	872	2.454667E-04	140.7262	.8743003	3.963874
10	1078	2.428449E-04	142.2455	.8837393	1.728388
11	1353	2.405892E-04	143.5792	.8920249	1.502056

CRF = -80.10544 SQI =-82.07542 Total Time = 144.0252

Experiment Number 49
Capillary Length = 13.83041 Centimeters
Methyl Cellulose Concentration = .1089393 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.993281E-04	86.58552	.537684	0
2	118	3.34412E-04	103.3935	.6420592	28.49432
3	194	2.946069E-04	117.3633	.7288095	20.38091
4	234	2.840423E-04	121.7285	.7559167	5.880105
5	271	2.770467E-04	124.8022	.775004	4.015512
6	281	2.754722E-04	125.5155	.7794335	.9177728
7	310	2.714807E-04	127.3609	.7908934	2.350375
8	603	2.526876E-04	136.8331	.8497143	11.54715
9	872	2.465539E-04	140.2372	.8708535	3.957001
10	1078	2.439264E-04	141.7478	.8802341	1.725316
11	1353	2.416658E-04	143.0738	.888468	1.49934

CRF = -80.11555 SQI =-82.1067 Total Time = 143.518

Experiment Number 50
Capillary Length = 13.80414 Centimeters
Methyl Cellulose Concentration = .1114882 Percent

Peak	Base Pairs	Mobility	Retention Time	Base Width	Resolution
1	72	3.977584E-04	86.76208	.5392929	0
2	118	3.329987E-04	103.6351	.6441714	28.51457
3	194	2.932895E-04	117.6665	.7313875	20.40104
4	234	2.827504E-04	122.0523	.7586488	5.886892
5	271	2.757716E-04	125.141	.7778474	4.020435
6	281	2.74201E-04	125.8578	.782303	.9189209
7	310	2.702191E-04	127.7125	.7938309	2.353381
8	603	2.514713E-04	137.2337	.853013	11.56305
9	872	2.453523E-04	140.6563	.8742868	3.962904
10	1078	2.427311E-04	142.1752	.8837278	1.727944
11	1353	2.40476E-04	143.5085	.8920153	1.501696

CRF = -80.10361 SQI =-82.07654 Total Time = 143.9545

Vita

Michael Allen Stebbins was born in La Mirada, California on April 17, 1970. Shortly thereafter his family moved to Charleston, South Carolina where he attended the Hanahan public school system. Upon graduation from Hanahan High School, he entered The College of Charleston, where in May of 1992 he received the Bachelor of Science degree in Chemistry. In August of 1992 he enrolled in the graduate school at The University of Tennessee, Knoxville and pursued research in the area of micro-scale based strategies for biochemical analysis. He earned a Ph.D. in Chemistry in December of 1997 and is presently employed with Alcon Laboratories in Ft. Worth, Texas.