1999

A Novel Approach to DNA Sequencing Employing Near -IR Fluorescence Detection Coupled With Microscale Reaction Vessels With Capillary Gel Separations.

Daryl Clifford Williams

*Louisiana State University and Agricultural & Mechanical College*

Follow this and additional works at: [https://digitalcommons.lsu.edu/gradschool_disstheses](https://digitalcommons.lsu.edu/gradschool_disstheses)

**Recommended Citation**


[https://digitalcommons.lsu.edu/gradschool_disstheses/7033](https://digitalcommons.lsu.edu/gradschool_disstheses/7033)

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.
INFORMATION TO USERS

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

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

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

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

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

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

UMI
800-521-0800

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
A NOVEL APPROACH TO DNA SEQUENCING EMPLOYING NEAR-IR
FLUORESCENCE DETECTION COUPLED WITH
MICRO-SCALE REACTION VESSELS
WITH CAPILLARY GEL SEPARATIONS

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Chemistry

by
Daryl C. Williams
B.S., The Tulane University of Louisiana, 1992
August 1999
To everything there is a season and
a time to every purpose
under heaven.
Acknowledgments

I would first like to thank Dr. Steven A. Soper for his direction and support of my studies at Louisiana State University. His suggestions and advice have proven invaluable. I would also like to thank Drs. Robert P. Hammer, George G. Stanley, Robin L. McCarty, Robert Grodner and Isiah M. Warner for serving on my committee, reading my dissertation and their guidance and direction.

I would like to recognize the Louisiana State University chapters of the National Society of Black Engineers and the National Organization of Black Chemists and Chemical Engineers. Their efforts to spark the scientific interests of minority youth in the Baton Rouge area are truly noteworthy. I would also like to thank the Huel Perkins Foundation for their generous support during my graduate studies. Special thanks go to my many friends in the Chemistry department, especially to the students in Dr. Soper's group, James H. Flanagan Jr., Benjamin L. Legendre Jr., Yolanda Y. Davidson, Clyde V. Owens, Christopher S. McWhorter, Sean M. Ford, Dixie L. Moberg, Emanuell Waddell, Yichaun Xu, Bill Kar, Yuijin Dong, Sarah Romero, P. Andrew Utter, and Kattie Price. Special thanks go to Robert Willicutt even though he is not a group member.

I wish to thank my parents Nathan Williams Jr. and Thelma C. Williams for their unconditional love and support throughout my youth and adulthood. Thank you for teaching me the value of patience and the true meaning of manhood. Finally, I wish to thank my wife, Mrs. Audra D. Williams, without whose constant love, support, sacrifice and understanding this work would not have been possible.
# Table of Contents

Acknowledgments.................................................................................. iii

List of Tables........................................................................................ vi

List of Figures...................................................................................... vii

List of Equations.................................................................................. x

List of Abbreviations............................................................................ xi

Abstract................................................................................................ xii

1 Introduction........................................................................................ 1
  1.1 Origins of the Human Genome Project.................................. 1
  1.2 Dissertation Overview...................................................... 8
  1.3 References............................................................................. 10

2 Theory and Background..................................................................... 12
  2.1 Status of the Human Genome Project................................. 12
  2.2 Sanger Dideoxy-termination Protocol.................................. 14
  2.3 Gel Electrophoresis............................................................ 15
    2.3.1 DNA Migration Models in Hydrophilic Gel Matrices...... 15
  2.4 Slab Gel Electrophoresis in DNA Sequencing Applications... 23
  2.5 Capillary Gel Electrophoresis (CGE) in DNA Sequencing Applications................................................. 27
    2.5.1 Fundamental Relationships in CE............................... 27
    2.5.2 Current Technologies.................................................. 29
    2.5.3 Gel Immobilization to Fused-Silica Capillary Walls..... 30
    2.5.4 Gel Column Bubble Formation................................. 31
    2.5.5 Linear Polyacrylamides in CGE................................. 33
  2.6 References............................................................................. 35

3 Gel Electrophoresis of Dye-labeled DNA Fragments Using Near-IR Fluorescence Detection................................................. 38
  3.1 Introduction............................................................................. 38
  3.2 Experimental................................................................. 46
    3.2.1 LIF-CGE System....................................................... 46
    3.2.2 Capillary Gel Column Preparation............................. 48
    3.2.3 Near-IR Primer-Labeled Sequencing.......................... 48
    3.2.4 Near-IR ddNTP-Labeled Sequencing......................... 49
  3.3 Results and Discussion...................................................... 51
  3.4 Summary............................................................................... 65
  3.5 References............................................................................. 65

4 DNA Sequencing Reactions Performed in a Nanoliter Volume Reactor Directly Coupled to Capillary Gel Electrophoresis................................................. 69
  4.1 Introduction........................................................................... 69
  4.2 Experimental................................................................. 72
## List of Tables

Table 1.1  
**Goalset Proposed for Completion and Cost of the Human Genome Project**  ........................................... 8

Table 3.1  
**Summary of Data Collected from CGE Separation of Near-IR/Fluorescein-labeled Primers**  ........................................... 55

Table 3.2  
**Demonstration of Bias Due to Electrokinetic Injection**  ........................................... 59

Table 4.1  
**Calculation of the Available Number of DNA Binding Sites**  ........................................... 87

Table 4.2  
**Cost Projections for Genomic Sequencing using Standard Polymerase Reaction Protocols vs. Nanoscale Protocols**  ........................................... 105

Table 5.1  
**List of Migration Models**  ........................................... 108

Table 5.2  
**Evaluation of Temperature and Injection Volume Influences as a Function of Capillary Diameter**  ........................................... 118
List of Figures

Figure 1.1 The Four DNA Bases ................................................................. 2
Figure 1.2 Structure of a DNA Chain ....................................................... 3
Figure 1.3 Structure of DNA Double-Helix ............................................ 4
Figure 1.4 One Strategy for Performing Genomic Sequencing ................ 7
Figure 2.1 Autoradiogram of a SGE Separation of Sequencing Products .... 13
Figure 2.2 Sanger Dideoxy-termination Protocol ................................. 16
Figure 2.3 Sanger Sequencing using SGE and LIF Detection ................. 17
Figure 2.4 Mobility of DNA Fragments ................................................... 18
Figure 2.5 Ferguson Plot of ss or ds DNA Fragments ......................... 20
Figure 2.6 Gel Pouring Procedure using Glass Plate and Polyacrylamide Matrices .......................................................... 24
Figure 2.7 Schematic of LIF-SGE Detection System ......................... 25
Figure 2.8 Terminal Base Discrimination System ............................... 26
Figure 2.9 Schematic of Gel Anchoring Reaction .................................. 31
Figure 2.10 Bubble Formed in Polyacrylamide Gel-filled Capillary ............ 32
Figure 2.11 Resolution Loss in CGE ....................................................... 34
Figure 3.1 Raman Spectrum of Acrylamide ......................................... 44
Figure 3.2 Spacing Between Raman Bands ........................................... 44
Figure 3.3 Absorbance Spectrum of Filter System used for Near-IR-DNA Sequencing Experiments ........................................ 45
Figure 3.4 Block Diagram of CGE System ............................................ 47
Figure 3.5 Structure of Near-IR and Fluorescein Labeled Primers .......... 49
Figure 3.6 Structure of Near-IR-labeled Dideoxynucleotide ................. 50
Figure 3.7 Fluorescence Spectra of Near-IR Dyes in Various Solutions .... 52

vii

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 3.8  CGE Separation and Detection of Near-IR/Fluorescein Labeled Primers

Figure 3.9  Electropherogram of ddTTP Terminated Fragments

Figure 3.10  CGE Sequencing Run Using a 4:2:1:0 Peak Height Base Identification System

Figure 3.11  Chain Extension using Near-IR-ddATP terminators

Figure 3.12  Separation of Sequencing Fragments Generated Using $^{32}$P labeled Primers, ddATPs and Near-IR-ddATPs

Figure 4.1  Description of Polymerase Chain Reaction (PCR)

Figure 4.2  Biotinylated dATP

Figure 4.3  Purification of PCR Products

Figure 4.4  Schematic of Biotin-Streptavidin Anchor System

Figure 4.5  Design/Photograph of the Air Thermocycler

Figure 4.6  Low-dead Volume Connector

Figure 4.7  Zero-dead Volume Connector

Figure 4.8  Custom-built Laser-Induced Fluorescence Capillary Gel Electrophoresis Apparatus (LIF-CGE)

Figure 4.9  Evaluation of Different Enzymes used in the Nano-reactor

Figure 4.10  Diagram of Sampling System used to Perform Scintillation Measurements

Figure 4.11  Effect of Sample Orientation on Scintillation Intensity

Figure 4.12  Relative Loss of Immobilized Template Under EO Flow

Figure 4.13  Relative Loss of Immobilized Template Under Pressure-driven Flow

Figure 4.14  Efficiency of Reimmobilization of Biotinylated DNA Template

Figure 4.15  Effect of Denaturing Cycle on Immobilized dsDNA
Figure 4.16  Diagram and Picture of Microreactor using Liquid Sheathed Temperature Control Provided by Circulating Waterbaths .................................................... 95

Figure 4.17  Electropherogram of Sequencing Products Collected from Large-Scale Capillary Reactor .......... 96

Figure 4.18  Experimental Diagram of an Integrated Microreactor/CGE System .................................................. 99

Figure 4.19  Electropherogram Showing the Efficacy of Injecting Labeled DNA Through an LPA Coated Capillary ................................................................. 100

Figure 4.20  Electropherogram of the Direct Injection of ddATP Termination Sequencing Products onto a CGE Column ..................................................................... 104

Figure 5.1  Image of Fused-silica Capillary Columns ................................................................. 111

Figure 5.2  Structure of MAPS Anchor, Polyacrylamide and Bis-Acrylamide Matrix .......................... 112

Figure 5.3  Voids Formed in CGE Matrix .................................................................................. 113

Figure 5.4  High Pressure Vessel ......................................................................................... 115

Figure 5.5  Efficiency of Oligonucleotide Separation ......................................................... 117

Figure 5.6  Photo-initiated Polymerization System ................................................................ 120
List of Equations

Equation 2.1 The Extended Ogston Model............................. 19
Equation 2.2 Reptation Model................................................. 21
Equation 2.3 Number of Theoretical Plates............................ 27
Equation 2.4 Experimental Number of Theoretical Plates........... 27
Equation 2.5 Theoretical Plate Height................................. 28
Equation 2.6 Resolution............................................................. 28
Equation 2.7 Experimental Resolution..................................... 28
Equation 2.8 Total Variance..................................................... 28
Equation 2.9 Efficiency in CE................................................. 29
Equation 3.1 Raman cross-section........................................... 43
Equation 3.2 Apparent Electrophoretic Mobility..................... 55
Equation 3.3 Injection Quantity in CE.................................... 56
Equation 3.4 Signal-to-Noise Ratio......................................... 56
Equation 4.1 Biotin/Streptavidin Binding Expression.............. 88
Equation 4.2 Rearranged Binding Expression......................... 88
Equation 5.1 The Extended Ogston Model................................ 109
Equation 5.2 Reptation Model................................................. 109
Equation 5.3 Variance due to $\sigma_{\Delta r}^2$.............................. 109
Equation 5.4 Migration rate.................................................... 110
Equation 5.5 Beer’s Law............................................................ 110
Equation 5.6 Injection Quantity in CE..................................... 116
Equation 5.7 Power at Capillary Center in CGE..................... 117
Equation 5.8 $\Delta T$ Capillary wall-center axis..................... 118
Equation 5.9 Variance due to $\sigma_{t}^2$....................................... 118
Equation 5.10 Variance due to $\sigma_{n}^2$.................................... 119
List of Abbreviations

ABI Applied Biosystems Inc.
APS Ammonium Persulfate
APTS Aminopropyltrichloxy silane
CE Capillary Zone Electrophoresis
CGE Capillary Gel Electrophoresis
CPA Cross-linked Polyacrylamide
dATP Deoxyadenosine Triphosphate
DC Door-Corridor
dCTP Deoxycytosine Triphosphate
ddATP Dideoxyadenosine Triphosphate
ddCTP Dideoxyctytosine Triphosphate
ddGTP Dideoxyguanosine Triphosphate
ddNTP Dideoxyribonucleotide Triphosphate
ddTTP Dideoxytyrosine Triphosphate
gTTP Deoxyguanosine Triphosphate
DNA Deoxyribonucleic Acid
dNTP Deoxyribonucleotide Triphosphate
ds double stranded
dTTP Deoxytyrosine Triphosphate
EOF Electroosmotic Flow
FITC Fluorescien Isothiocyanate
HGP Human Genome Project
LDV Low dead volume
LIF Laser Induced Fluorescence
LPA Linear Polyacrylamide
MAPS Methacryloxypropyltrichloxy silane
PCR Polymerase Chain Reaction
RNA Ribonucleic Acid
RSD Relative Standard Deviation
SGE Slab Gel Electrophoresis
SNR Signal to Noise Ratio
ss single stranded
TEMED N, N, N', N'-tetramethylethylenediamine
UV Ultraviolet
ZDV Zero dead volume

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Abstract

Electropherograms of oligonucleotides labeled with near-IR fluorescent dyes, separated by capillary gel electrophoresis (CGE) and detected using an ultrasensitive near-IR fluorescence detection system will be presented. A Universal M13 sequencing primer was labeled on the 5' end with a near-IR dye containing an isothiocyanate functional group. Comparison of the on-column detection limits in capillary gel electrophoresis for the near-IR-labeled sequencing primer to that of a visible fluorescein-labeled primer indicated improved sensitivity for the near-IR case. The detection limit was found to be $3.4 \times 10^{-20}$ moles (SNR=3) for the near-IR-labeled primer while the on-column detection limits for the fluorescein analog was $1.5 \times 10^{-18}$ moles (SNR=3). The sequence of nucleotide bases in an M13mp18 template was determined using a single lane, single dye technique. The molar concentrations of the ddNTP's used during chain extension reactions were varied in order to achieve a ratio of 4:2:1:0 (A:C:G:T) which allowed the identification of each terminal base via fluorescence intensity measurements. Comparison of the known sequence of the M13mp18 plasmid to that obtained using this protocol yielded a base-calling accuracy of 84%.

A catfish gene DNA template was attached to the interior wall of an aminoalkylsilane derivatized (100 pm, 50 pm or 20 pm i.d.) fused-silica capillary tube via a biotin/streptavidin linkage. The DNA modified capillary reactor was then used to perform DNA sequencing reactions employing standard Sanger dideoxynucleotide termination protocols and the Vent® DNA polymerase enzyme. Preliminary analysis of nano-reactor products was performed using a commercial fluorescence-based DNA sequencing instrument, the DuPont Genesis 2000. The micro-reactor was then directly interfaced to a capillary gel electrophoresis system to take advantage of the small sample requirements of CGE and create an integrated near-IR LIF-based DNA analysis system.
An investigation of the effect of varying capillary internal diameter (i.d.) on the separation efficiency is presented. Gel-filled columns with i.d.s of 20\,µm, 50\,µm, 75\,µm, and 100\,µm were tested with an oligonucleotide standard, until failure, to determine the efficiency obtainable using these columns.
1 Introduction

1.1 Origins of the Human Genome Project

The human genome consists of 23 pairs of chromosomes. Each of these chromosomes contains polymeric strands of deoxyribonucleic acid (DNA), which is comprised of four distinct nucleotide bases which are adenosine (A), cytosine (C), guanosine (G) and tyrosine (T) (see Figure 1.1 and 1.2). The human genome encompasses approximately 3,000,000,000 bases in its entirety. The DNA strands are paired through the hydrogen bonding interaction of the nucleotide bases (A-T and C-G) which gives rise to its characteristic double helical conformation shown in Figure 1.3.

The DNA molecules present in each cell contain all the genetic information for an organism. This information controls all aspects of the development of the organism, such as the color, length and texture of a human's hair to the size and shape of their feet. About 100,000 segments of the polymeric nucleotide strand (chromosome) have been identified as genes. A gene is a region on a chromosome that directs the production of proteins, which are formed from 20 amino acids. The DNA sequence determines the connective order of amino acids in the protein molecules and it is the connective order of amino acid residues which determine the 3-dimensional conformation of the protein molecule and inevitably its function. The 3-dimensional structure of a protein defines its functionality and allows it to perform the complex reactions that sustain (and sometimes destroy) our bodies. Thus, it is clear that a detailed and thorough understanding of genomic material is of critical importance in the consideration of the form and function of genetic diseases in human beings. The human genome project has already had a profound effect on biomedical research. Recently, map produced by researchers have helped in finding gene...
others. In addition to the identification of many more disease genes, future developments may enable researchers to explore gene mutations and health effects induced by environmental agents.

![Molecular structures](image)

**Figure 1.1**
The structures of the bases and sugars that constitute DNA. Ribose is the major component of RNA while 2'-deoxyribose is found in DNA.
Figure 1.2
Structure of part of a DNA chain.
Figure 1.3
Structure of DNA double-helix. The diameter of the helix is 20 Å and adjacent bases are separated by 3.4 Å with a rotation of 36 degrees.

The Human Genome Project (HGP) is an international endeavor with the goal of completely deciphering the human genetic code. Initiated in 1987, this project has defined objectives and a projected timetable by which these objectives should be accomplished. The plan of action for the HGP, established by a National Research Council committee and chaired by Bruce Alberts, placed early emphasis on the creation of high resolution “maps” of the human genome using clone-based
techniques as depicted in Figure 1.4. These maps provide a “road map” for investigators to guide DNA sequencing projects. The second and current phase of the HGP places emphasis on actual sequencing of the individual clones, allowing the total reconstruction of the human genomic sequence. In addition, emphasis has been given to the development of faster, more accurate, higher throughput, lower cost sequencing tools during this phase.

DNA sequencing basically involves 5 steps:\(^2\) (1) preparation of the chromosomal (or template) DNA; (2) generation of fragments complementary to the DNA template; (3) separation of fragments, usually by a gel electrophoretic technique; (4) detection of fragments and (5) data analysis (base calling). The acceleration of these steps would facilitate genetic analysis, especially for large scale projects such as the Human Genome Initiative. At the onset of the HGP, expectations were high that the development of powerful new sequencing tools would replace the repetitive, labor intensive techniques such as manual colony isolation, recombinant clone amplification, slab gel electrophoresis (SGE) and autoradiographic analysis.\(^3\) Colony isolation is still primarily performed by scientists using non-automated techniques. Visually identifying individual colonies requires experience in order to be done correctly not to mention that a fair amount of manual dexterity is needed to extract the colonies from growth media. This operation requires a high skilled individual and requires a time frame on the order of day when colony growth time is included. The time frame needed and dependence on manual labor does not promote current colony isolation techniques as viable candidates for a high throughput DNA analysis system. Recombinant clone amplification uses a bacterial host, most often E.Coli, to amplify regions (vectors) of DNA for use in analysis. This system is dependent on reproduction of the host in order to amplify the vector. This system also requires a long time frame (1-2 days or more) to produce the desired amplification. Recombinant methods are also not easily automated since their products must be
located on growth media, isolated, excised, and further manipulated (including lysing, restriction cutting, etc) prior to use in analysis. Slab gel electrophoresis is a powerful separation technique used to separate DNA fragments with as little as a single base difference in length. SGE allows for parallel separation of multiple samples which increases throughput during analysis. However, SGE gels which commonly employ crosslinked polyacrylamide as a sieving matrix, can only be used for a single separation run before the separation efficiency is degraded. Precast SGE gels can be purchased commercially, but this is an expensive option. Most SGE gels are prepared “in-house”. A skilled, experienced researcher can prepare these gel fairly easily, but not with 100% efficiency. Autoradiography is the most common detection scheme used in DNA sequencing. This technique requires labeling of the DNA fragments with radioactive isotopes ($^{32}$P or $^{35}$S) then separation on a slab gel followed by exposure to X-ray film to produce a pattern on the film corresponding to the fragmentation pattern on the gel. The exposure time (1-3 days typically) is dependent on the activity of the isotope, the efficiency of isotopic DNA labeling, the amount of fragments loaded on the gel and other factors. Many advances in the areas of DNA amplification, gel electrophoresis and detection have occurred since the 1980s. Many of these can be considered noteworthy and in some cases revolutionary.

Several advances in the domain of molecular biology could be described as revolutionary, such as the introduction of the polymerase chain reaction (PCR) which provides a means of totally automated DNA amplification. Additionally, the introduction of Maxam-Gilbert chemical cleavage and Sanger dideoxynucleotide sequencing protocols provide two unique ways to produce nested sets of DNA fragments suitable for sequencing.
Total genomic DNA (23 pairs of chromosomes)

Single purified chromosome

Overlapping set of cosmids (40 kb)

1 kp fragments for sequencing

Figure 1.4
One strategy for performing genomic mapping and sequencing is shown. A single chromosome is isolated and used to prepare specific cosmid libraries. The cosmids are then fragmented into pieces suitable for sequencing (adapted from Biotechnology, Volume 5 (1987) pp. 934).

Progress in data processing, storage and retrieval is directly related to gains in computer processing power and has seen exponential growth over the past ten years. Computing advances could be best illustrated by one simple yet amazing fact: the information contained in the entire human genome would compromise 750 megabytes of digital information. Ten years ago, this information would have required an entire room of magnetic tape for storage, whereas today it would fit neatly on a single CD-ROM disk.

The third and final stage of HGP entails the application of the information obtained from the genetic code. This phase will include comparative sequencing which could facilitate the identification and correlation of genes with a number of
disease states such as Parkinson’s, Huntingdon’s and certain cancers. Table 1.1 shows in some detail the proposed goals and timeline for sequencing the human genome.8

Table 1.1
Goals proposed for completion and costs for the Human Genome Project in the United States.2 Total cost cost assumes sequencing of both strands to account for any mis-matched bases along the DNA strands.

<table>
<thead>
<tr>
<th>Table 1.1</th>
<th>Timeline of Human Genome Project</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stage I (5-10 years)</strong></td>
<td></td>
</tr>
<tr>
<td>1. Technology Development</td>
<td></td>
</tr>
<tr>
<td>2. Physical and genetic mapping</td>
<td></td>
</tr>
<tr>
<td>3. Regional sequencing</td>
<td></td>
</tr>
<tr>
<td>4. Regional comparisons (polymorphisms)</td>
<td></td>
</tr>
<tr>
<td>5. Small genome analysis (E. coli and yeast)</td>
<td></td>
</tr>
<tr>
<td><strong>Stage II (5-20 years)</strong> (cont.)</td>
<td></td>
</tr>
<tr>
<td>1. Sequencing the human genome</td>
<td></td>
</tr>
<tr>
<td>2. Sequencing of other genomes</td>
<td></td>
</tr>
<tr>
<td><strong>Stage III (&gt;20 years)</strong></td>
<td></td>
</tr>
<tr>
<td>1. Application of sequencing tools</td>
<td></td>
</tr>
<tr>
<td>2. Comparative human genome sequencing</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Economic Goals</th>
<th>Raw Data</th>
<th>Total Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. technology (1987)</td>
<td>$1.00/base</td>
<td>$6 x 10^9</td>
</tr>
<tr>
<td>Japan technology (1987)</td>
<td>$0.17/base</td>
<td>$1 x 10^9</td>
</tr>
<tr>
<td>Current technology (1996)*</td>
<td>$0.10/base</td>
<td>$6 x 10^8</td>
</tr>
<tr>
<td>Proposed technology</td>
<td>$0.01/base</td>
<td>$6 x 10^7</td>
</tr>
</tbody>
</table>

1.2 Dissertation Overview

Chapter 2 will review the current technologies used in the Human Genome Project beginning with a discussion of the molecular biology employed in performing Sanger DNA sequencing. This chapter will also describe the sieving technologies employed in DNA separation including slab gel electrophoresis (SGE) and capillary gel electrophoresis (CGE) techniques. The details of several models of DNA motion

* Current price quoted from GENELABS, Louisiana State University, Jan. 1997.
in a sieving medium under an applied potential will be discussed with particular emphasis placed on the Ogston and Reptation theories. Finally, a brief discussion of the instrumentation utilized in DNA sequencing analyses will be presented.

Chapter 3 will describe the integration of near-IR fluorescence detection into DNA sequencing applications. The motivation for using fluorescence monitoring based in the near-IR region includes the smaller backgrounds observed during signal collection and the simple instrumentation required to carry out detection. Backgrounds are lower, in part, because few species fluoresce in the near-IR, thus interferences are reduced. The $1/(\lambda^4)$ dependence of the Raman cross section also provides a lower scatter contribution at these longer excitation wavelengths, especially when detection must occur in highly scattering gel systems. The result should be improved signal-to-noise in gel systems using near-IR fluorescence detection. The instrumentation required for near-IR fluorescence can consist of simple and inexpensive diode lasers and avalanche photodiodes. Since both of these components are solid-state, the detector can be run for extended periods of time requiring little maintenance, operator expertise, or cost.

Chapter 4 will discuss the use of a miniaturized reaction vessel to perform DNA sequencing reactions. A $^{32}$P labeled DNA molecule was covalently attached to the interior wall of an aminomethylsilane derivatized fused-silica capillary tube via a biotin/avidin/avidin linkage. Stability of the template anchor to the capillary wall under rinsing (gravity and electroosmotic) was investigated, as well as stability of template immobilization under temperature cycling conditions similar to those encountered in cycle sequencing protocols. The DNA modified capillary reactor was then used to perform DNA sequencing reactions employing standard Sanger dideoxynucleotide termination protocols and the Vent® DNA polymerase enzyme. The sequencing fragments generated in the reactor were removed using electroosmotic rinsing and directly injected onto a capillary gel column for near-IR-LIF detection. The
concentration of termination fragments in the reactor volume was in the μM range. Separation of reaction products was performed using a custom design LIF-CGE device which employed polyacrylamide (3%T / 3%C) as the sieving medium.

Chapter 5 will discuss the effects of capillary diameter in CGE on separation efficiency determined by number of theoretical plates. The enhancement of Joule heat dissipation afforded by capillary inner diameter (i.d.) reduction and the subsequent increase in surface to volume ratio is a well understood phenomena. The advantages provided by using small capillaries in free solution electrophoresis (CZE) have been described⁹. Capillary reduction in CGE separations should provide analogous benefits; however, the fabrication of highly viscous gel-filled columns with id < 50 μm has proven to be nontrivial. This chapter presents initial studies of capillary diameter effects observed when attempting to perform separations in capillaries with i.d.s ranging from 20-100 μm using high viscosity, polyacrylamide sieving matrices. A number of gel-column fabrication techniques are evaluated for effectiveness in consistently producing usable separation columns.

Chapter 6 draws conclusions from the results and discussions in previous chapters and proposes avenues to guide future studies in these areas.

1.3 References


1.8. **Smith, L. and Hood, L.** *Biotechnology 1987, 5, 933-939.*

2 Theory and Background

2.1 Status of the Human Genome Project

The Human Genome Project is a massive endeavor to reveal the sequence of the human genetic code. This effort was initiated in the late 1980's in the United States. Having completed approximately 1/3 of its strategic plan, the Human Genome Project has begun to focus on increasing sequencing capacity and developing technologies which will allow greater data throughput at a lower cost. Traditional sequencing techniques, namely autoradiographic analysis and slab gel electrophoresis, do not possess the necessary throughput to successfully sequence the human genome in a reasonable time frame. Low data throughput is generated from SGE techniques using isotopic detection in part because the system requires banding patterns present on the gel to be transferred to X-ray film. Exposing or “transferring” images to X-ray films is accomplished by placing the gel containing fractionated, radioactively labeled DNA fragments in close proximity to the film. Emission from the radioactive isotope (\(^{32}\)P, \(^{33}\)P, or \(^{35}\)S) which has been incorporated into each DNA fragment, exposes the X-ray film which produces a banding pattern on the film identical to that on the gel. The film is then developed to allow visualization of the pattern and reconstruction of the sequence. The time needed to generate a suitable pattern on the film can be in excess of 72 hours, after which the film must be “developed” and “fixed”, which involves submerging the film in various chemicals in a photographer’s dark room. The visible pattern on the film can then be “read” in order to obtain a sequence. An example of the pattern produced using SGE and autoradiography is shown in Figure 2.1. In spite of its labor intensity and tedious nature, analysis by autoradiographic techniques is still currently the most widely used means by which DNA sequencing is performed.

The development of fluorescence-based detection has reduced the amount of time required to determine a DNA sequence. The performance of current
fluorescence-based sequencing technologies can be benchmarked using the commercially available SGE sequencer from Applied Biosystems, Inc. (ABI). The ABI 377, which employs laser-induced fluorescence (LIF) as its mode of detection, is one of the most commonly used automated SGE systems. This device is capable of fractionating 36 lanes simultaneously with read lengths ranging from 450-500 bases. The development time for the electrophoresis of 500 bases using this instrument is approximately 3.5 hours which translates to an overall throughput of 5100 bases/hr or 123,000 bases/day. Direct extrapolation of this sequencing rate to the human genome, which contains over 3 billion base pairs, would suggest that one could determine its sequence in about 67 years using a single ABI 377 instrument.

Figure 2.1
Autoradiogram of SGE separation of sequencing products generated from an M13mp18 DNA template. Separation was performed using a 6%T / 5%C polyacrylamide gel matrix. M13 universal sequencing primers were labeled with $\gamma^{32}$P. The four reaction mixtures were electrophoresed at 2500 V on a 16'' slab gel.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
However, these rates do not account for time consuming actions such as gel pouring and sample loading. A more realistic portrait of genomic sequencing is obtained when one considers the work performed at The Institute of Genomic Research (TIGR). Fleischmann and co-workers in Venter's group were the first to successfully sequence a free living organism, *Haemophilus influenza*, which contains 1.8 million base pairs. This project required 14 ABI 377 sequencers and 13 months to complete. This sequencing rate translates to 4600 finished bases/day (~ 330 bases/day-instrument) and would require 2000 years if applied to the analysis of the human genome.

### 2.2 Sanger Dideoxy-termination Protocol

Dideoxynucleotide termination reactions were introduced by Fred Sanger in 1977. This procedure has been embraced by the sequencing community due to its simplicity and amenability toward automation. Detection of Sanger termination products is accomplished by using labeled primers (short oligonucleotides), deoxynucleosides (dNTPs), or dideoxynucleosides (ddNTPs). These labels typically consist of either radioisotopes ($^32$P, $^33$P or $^35$S) or fluorescent probes. The nascent action in this procedure involves heating the reaction mixture containing the single stranded (ss) DNA template and primers to a temperature above 75°C. Double stranded (ds) templates must be denatured at temperatures above 90°C prior to use. The primer is a short strand of nucleotides (10-20 bp) with a sequence complementary to the template DNA at a specific site. While this mixture is allowed to cool slowly, the primer and template combine (or anneal) to form a primer-template construct as depicted in Figure 2.2a. The subsequent action taken on the reaction mixture will vary depending upon the particular DNA polymerase enzyme used, however, all share the commonality in that dNTPs and ddNTPs are added by the action of the enzyme which attaches itself to the primer-template construct, then catalyzes the addition of...
dNTPs complementary to the corresponding bases on the template DNA (see Figure 2.2b). The polymerase action is repeated until a dideoxynucleoside, which is a nucleoside lacking 2' and 3' hydroxyl groups (ddNTP), is incorporated into the strand. No further chain extension occurs once a ddNTP is incorporated into the growing strand; thus, a distribution of every possible fragment length is generated (see Figure 2.2c). The fragments are denatured, then fractionated using gel electrophoresis and the information used to reconstruct the sequence of the unknown DNA template (see Figure 2.3).

2.3 Gel Electrophoresis

2.3.1 DNA Migration Models in Hydrophilic Gel Matrices

Supports were first employed to provide anticonvective properties during electrophoretic analysis. A number of materials have been used as supports in electrophoresis including paper, silica beads, glass powder, polyurethane foam and Sephadex. More recently, hydrophilic polymeric gels have also shown to be an effective sieving medium in which to perform the separation of biomolecules and have become the primary sieving matrix for analysis using slab and capillary gel electrophoresis. Agarose gels are commonly used to separate large DNAs (kilobases), while polyacrylamide is used to separate smaller DNA molecules (hundreds of bases). Extremely high separation efficiencies can be achieved by using sieving gels to separate DNA molecules.

It is necessary to employ a sieving matrix in potential driven DNA separations because the electrophoretic mobility \( \mu \) of DNA fragments longer than 10 bases is essentially constant.\(^5\) The behavior of ssDNA under an applied potential can be described by the free-draining coil model. Motion of the coil is governed by two dominant factors, charge and friction.
Figure 2.2
Sanger sequencing strategy. (a) Anneal primer to template DNA. (b) Extension and denaturation of complementary strand of DNA formed by polymerase enzyme. (c) Fragment distribution generated by the incorporation of dideoxy-terminators (ddNTPs). T* represents ddTTPs.
Figure 2.3
The procedure used to reconstruct a DNA sequence using Sanger dideoxynucleotide termination protocols using dye-labeled primers, SGE and laser-induced fluorescence detection.

As the coil is elongated (longer DNA fragments), the frictional factor is increased which tends to slow migration; however, since each sub-unit (nucleotide base) adds charge, which speeds migration, the net effect is cancellation of both influences leading to a single $\mu_{em}$ for DNA fragments >10 bases. This effect renders free solution electrophoretic based separations ineffective for DNA fragments. (See Figure 2.4)

Electrophoretic separations performed in a hydrophilic gel matrix, usually agarose or polyacrylamide, have been utilized for many years primarily in the
biological sciences for the purification of enzymatic digestion products and protein analysis. The products of many enzymatic digestion reactions can be quickly and easily resolved using an agarose gel as the sieving medium. Large DNA (>2,000 bp) fragments process rapidly through agarose polymer systems due to the relatively large average pore size (100 nm in a 2% gel solution) which allows fast analysis. However, the agarose matrix does not possess the sieving power, due to its large pore size, necessary to resolve oligonucleotide fragments with small differences in total molecular weight. In order to achieve the extremely high efficiency and selectivity necessary to affect the separation of oligonucleotide fragments that differ in length by only one base, such as sequencing fragments, one must employ a matrix which possesses greater sieving power, such as polyacrylamide.

The nature of the interaction between the polyacrylamide gel matrix and DNA species during separation has been studied extensively by a number of scientists and there are a host of models predicting the behavior of oligonucleotides in a polyacrylamide gel matrix.6,7

Figure 2.4
Graph of the electrophoretic mobility of DNA versus the number of bases. (Taken from Capillary Electrophoresis: Theory and Practice, 1992, Grossman, Acad. Press. San Diego, CA.)
The multitude of models can be divided into two distinct categories, namely, those which describe fixed pores within the polyacrylamide gel matrix and those which describe fractionation based upon interaction of biomolecules with mutable pore structures within the gel medium.8

The most mature theory describing the motion of a molecule in the presence of fixed, straight, long fibers is the Ogston model9, which was expanded to describe motion in networks of points and planes in 1970.10,11 The extended Ogston model predicts that the apparent electrophoretic mobility, $\mu_{\text{app}}$, of a large molecule is a function of the volume fraction of fixed gel pores large enough to accommodate passage of the molecule. The extended Ogston model has been used extensively to describe the motion of large DNA molecules (2-50 kilobase pairs) in agarose medium and is most effectively applied when the size of the migrating species is smaller than the average pore size of the gel. This model assumes the sieving medium is fixed, the migrating species is spherical, and there is no attractive or repulsive interaction between migrating species in the gel matrix during electrophoresis. The mathematical relation given by the extended Ogston model is:

$$\log \mu_{\text{app}} = \log \mu_o - \pi L (r + R)^2 T$$

(2.1)

where $\mu_{\text{app}}$ is the apparent electrophoretic mobility of the molecule, $\mu_o$ is the electrophoretic mobility of the molecule migrating in free solution, $L$ is the gel polymer length per unit volume, $r$ is the polymer radius, $R$ is the radius of the migrating molecule, and $T$ is the total polymer concentration. This model is supported by the linear relationship observed when gel concentration is plotted vs. $\mu_{\text{app}}$ of DNA fragments. Graphs depicting the relationship between DNA mobility and gel concentration are known as Ferguson plots (see Figure 2.5).12
Figure 2.5
Ferguson plot for sample DNA fragment mobility vs. hydroxyethylcellulose (HEC) concentration: 118 bp, 194 bp, 234 bp, 281 bp, 310 bp, 603 bp, 872 bp, 1078, 1353 bp. Adapted from Capillary Electrophoresis: Theory and Practice, 1992, Grossman and Colburn, Acad. Press. Inc. San Diego, CA

Although the Ogston model can be used to effectively describe the electrophoresis of DNA molecules in low concentration gels (large pore radius), obvious limitations in the model are shown when it attempts to predict behavior in a system that does not possess fixed, large pores structures, such as sequencing ladder separations in high concentration polyacrylamide media which possess small pore diameters. The extended Ogston model treats the migrating species as an immutable sphere traveling through defined openings in the gel; thus, it predicts that the mobility of migrating species will quickly approach zero when the molecular radius approaches the pore size of the sieving matrix. But it is well known that many large, flexible molecules will migrate through sieving media with mesh sizes smaller that the radius of the molecule. These phenomena are explained by the reptation model of electrophoresis.

The reptation model is based on the assumption that the migrating molecule does not behave as an immutable sphere with fixed radius, but instead behaves as a free draining coil. The reptation theory, proposed by deGennes\textsuperscript{13} and later applied to
gel electrophoresis by Lerman and Lumpkin, suggested that elongated molecules, such as DNA, would exhibit migratory behavior much like that of a snake ("reptilelike"-reptation) while migrating in a gel matrix under the influence of an electric field. Due to the elongated shape and the "snake-like" motion proposed for long DNA under an electric potential, $\mu_{app}$ becomes solely dependent on the length of the DNA fragment. The relation given by the reptation theory is shown by

$$\mu_{app} = \frac{q}{3\zeta} \left( \frac{1}{N} + \frac{(E')^2}{3} \right)$$

where $\mu_{app}$ is the mobility, $q$ is the total net charge, $\zeta$ is the frictional resistance for the translational motion along the tube, $N$ is the number of tube segments (direct relationship to DNA length), and $E'$ is the normalized electric field. In an extension of the reptation model, it was proposed that during electrophoresis DNA molecules in polyacrylamide gels migrate through spaces connected by narrow passages thus nicknamed the "lakes and straights" addendum. The proposed "lakes" and "straights" exist as intrinsic structures in gels and their structures are not influenced by the passage of macromolecules. This model was used to address anomalous experimental findings obtained using field inversion gel electrophoresis which could not be described using the classical reptation model.

A number of new theories describing molecular migration in gel medium were introduced after experimental results showed that the migration of proteins and nucleic acids in systems containing noncross-linked polyacrylamide (LPA) was similar to that observed in cross-linked polyacrylamide (CPA). The idea that covalent links between polymer chains in the gel were not necessary to resolve biomolecules had been proposed as early as 1964. The workers who introduced electrophoresis in polyacrylamide gels had originally suggested a "mutable" pore theory to explain their experimental results. The conclusion reached was that finite pores were not
necessary to provide a selective retardation of large molecules, since no pores of any
finite size should be present in a solution of randomly distributed flexible polymers.

Recently, a mutable pore theory describing the behavior of DNA in
polyacrylamide has been reported by Calladine. This theory is a mechanical model
which describes the migration of DNA molecules through a network of cells. It
suggests that migrating DNA "piles up" against the wall of a cell until the cell either
deforms to allow passage or ruptures. This description, termed the "gel-cell" model,
is suitable for the explanation of DNA motion through media which possesses flexible
and elastic polymer fibers such as LPAs.

Recently a new theory, which challenges the reptation model's description of
DNA motion through gel media, has surfaced. This recent model, the
"Door/Corridor" (DC) theory, employs a narrow definition to describe gels as
blocks which contain randomly distributed, flexible polymer chains in water. This
definition would exclude media such as silica gels and macrorecticular
polystyrene/divinylbenzene gels because they do not possess random, hydrated
polymers with motional freedom. While previous theories have differed regarding
the nature of gel pores as fixed or mutable, the DC model is anomalous from almost
all other migration models because it proposes that there are no preexisting pores in
the gel matrix. A migrating molecule is predicted to push aside polymer chains under
an electric field, thereby clearing the space in which it resides. Molecules are also
predicted to move in discrete increments or steps in which each step represents a gel
layer. The terms "door" and "corridor" refer to the different modes by which a
macromolecule can move through the gel matrix. A door is created when a processing
molecule pushes through only one gel layer at a time. Motion through doors is
predicted to provide optimal resolution of macromolecules such as DNA. The
formation of corridors is described as the result of the macromolecules pushing aside
multiple gel layers simultaneously. This phenomenon is theorized to occur in
situations where the gel matrix has low flexibility, for example near crosslinking sites, in matrices with high total gel concentration (>15% T) or when the force generated by the molecule is substantially greater than the resistance provided by the gel matrix (i.e. under high applied potential > 300 V/cm). Corridors are also predicted to occur when the sample capacity of the gel is exceeded (sample overloading). Corridors are described as vacant paths through the gel that allow DNA molecules of different sizes to migrate at the same rate. The result of this phenomenon is the loss of resolution.

2.4 Slab Gel Electrophoresis in DNA Sequencing Applications

The most commonly used apparatus for DNA sequencing is the vertical slab gel system. The primary action in SGE is the formation of a cross-linked polyacrylamide gel between two glass plates (see Figure 2.6). Polymerization of a gel solution is commonly initiated using an ammonium persulfate (APS) and N,N,N',N'-tetramethylethylenediamene solution. The solution is poured between the plates with special attention not to form bubbles in the gel sheet. Until the mid 1980's, detection of DNA fragments on SGE apparati could only be accomplished by the use of autoradiography. In 1986 an important advance was made when three research groups reported on the integration of fluorescence detection with SGE. Smith, et al.\textsuperscript{27} Ansorge, et al.\textsuperscript{28} and Prober, et al.\textsuperscript{29} demonstrated the replacement of radioisotopes in Sanger dideoxynucleotide termination methods with fluorescent labels and LIF detection. The LIF-SGE systems reported by these authors have been made commercially available by Applied Biosystems, Inc., Pharmacia, and DuPont, respectively; however, as of this writing, the DuPont system has been discontinued.
An example of an LIF-SGE system is shown in Figure 2.7. The Genesis 2000® is designed to perform vertical DNA fractionation with visible LIF detection of migrating oligonucleotides labeled with four distinct fluorescent probes. The four succinylfluorescein dyes used to label the DNA fragments have emission maxima centered at 505(G), 512(A), 519(C), 526(T) nm. All four of the probes can be excited by the 488 nm line of a small frame Ar ion laser.
Figure 2.7
Schematic of the LIF-SGE detection system used in the DuPont Genesis 2000® Genetic Analyzer system. (Schematic taken from the operations manual.)

The fluorescence generated from these four probes is collected by two separate photodetectors placed in close proximity to the gel slab. Data analysis, which includes identification of the terminal base on a DNA fragment, is accomplished by discrimination of the spectral emission collected from each electrophoretic band during analysis. The intensity of fluorescent light impinging on each photodetector is used to determine the identity of the fluorophore, and consequently the identity of the terminal base present on the labeled fragment during electrophoresis (see Figure 2.8).
Figure 2.8
Terminal base discrimination is accomplished using four dye-labeled dideoxynucleotides in the Genesis 2000® system. (Taken from the operations manual.) The relative intensity of the signals detected on each photomultiplier (A or B) tube is used to identify the terminal base present on that fragment as it passes the detection region.

One very attractive aspect of the SGE format is that it is inherently multiplex-capable. Several samples can be analyzed simultaneously and independently of each other. Multiplexing is of critical importance for large scale sequencing systems to generate the high data throughput (megabases/day) needed to perform genomic sequencing. However, there are a number of disadvantages associated with SGE-based DNA sequencing. While high throughput is facilitated by the fractionation and detection of numerous samples on this type of system, the separation time in SGE is long (100 bases/hr-lane) relative to that achieved using miniaturized separation systems, such as capillary gel electrophoresis (CGE) (800-1000 bases/hr-lane). Attempts have been made to enhance the separation speed achieved in SGE by employing the use of ultrathin gels as the separation medium.30,31 The thickness of this gel is determined by the dimensions of the spacers used along the sides of the gel.
during polymerization. Sequencing gels are typically between 0.25 mm and 0.40 mm in thickness, but some investigators have been able to employ slab gels with thickness' ranging from 10-100 μm. Ultrathin SGE takes advantage of the greater surface-to-volume ratio which enables more efficient dissipation of Joule heat that can cause thermal gradients, convection and increase diffusion, all of which reduce separation efficiency. The use of ultrathin gels and higher electric fields during electrophoresis result in shorter analysis times and higher efficiency. Some problems associated with ultrathin SGE include bubble formation during gel pouring. Also, sample overloading is not uncommon since the volume of the separation matrix is greatly reduced; thereby, lowering the sample capacity of the gel system.

2.5 Capillary Gel Electrophoresis (CGE) in DNA Sequencing Applications

2.5.1 Fundamental Relationships in CE

In order to discuss the separation of species using potential driven systems in capillaries, it is important to understand some fundamental expressions used to describe separation parameters such as resolution, efficiency and theoretical plate height. The number of theoretical plates is defined by:

\[ N = \frac{L^2}{\sigma_T^2} \quad (2.3) \]

where \( L \) is the total length traveled by the zone and \( \sigma_T^2 \) is the total variance of the zone including contribution from all dispersive processes. Experimentally, one can use Equation 2.4 to calculate the number of theoretical plates obtained during an analysis;

\[ N = 2\pi \left( \frac{h_p L}{A_p} \right) \quad (2.4) \]

Equation 2.4 is particularly handy since peak height \( h_p \) and peak area \( A_p \) are readily available using an integrator or chromatography software. The theoretical plate height, \( H \), is calculated using Equation 2.5.
It is noteworthy to mention that in the strictest definition, it is inappropriate to refer to electromigratory separations in terms of theoretical plates. The concept of theoretical plates relies on the existence of partitioning processes between inhomogenous phases, which is not the case in electrophoresis (excluding micellar electrokinetic chromatography (MEKC) and capillary electrochromatography (CEC) which are not discussed here). Thus, in this discussion, theoretical plates serve as a convenient tool to describe Gaussian peaks. Resolution ($R$), the spacing between adjacent zones, can be described by Equation 2.6.

$$R = \frac{1}{4} \frac{\Delta \mu_{app}}{\mu_{app}} N \left( \frac{1}{2} \right)$$

Equation 2.6 is comprised of a selectivity term ($\Delta \mu_{app}/\mu_{app}$) which is media dependent and an efficiency term ($N^{1/2}$) which is primarily dependent on instrumental parameters. Experimentally, $R$ can be calculated using Equation 2.7;

$$R = \frac{x_2 - x_1}{\frac{1}{2}(w_1 + w_2)} \quad (2.7)$$

where $x_2$ and $x_1$ are the peak elution times (s) and $w_2$ and $w_1$ are the base widths of the peaks (s).

The total variance, $\sigma_T^2$, contains contributions from all dispersive phenomena in CE;

$$\sigma_T^2 = \sigma_D^2 + \sigma_{AT}^2 + \sigma_i^2 + \sigma_w^2 + \sigma_{sh}^2 + \sigma_{sk}^2 \quad (2.8)$$

where $\sigma_D^2$ is the variance due to longitudinal diffusion, $\sigma_{AT}^2$ is the variance due to a parabolic temperature profile, $\sigma_i^2$ is the variance due to injection, $\sigma_w^2$ is the variance associated with solute-wall interactions, $\sigma_{sh}^2$ is the variance due to siphoning caused by different heights of the capillary tube ends and $\sigma_{sk}^2$ is the variance due to sample-
buffer conductivity differences. In an ideal CE system, $\sigma_w^2$, $\sigma_{\Delta t}^2$, $\sigma_r^2$ and $\sigma_{\Delta t}^2$ can be disregarded. Therefore, for CE $\sigma_{\Delta t}^2 \approx \sigma_D^2$ since $\sigma_{\Delta t}^2$ is small for capillaries. Theoretical plates in CE can then be expressed by Equation 2.9:

$$N = \frac{L^2}{2Dt}, \quad (2.9)$$

where $L$ is the effective length of the column, $D$ is the diffusion coefficient of the species and $t$ is the elution time.

### 2.5.2 Current Technologies

CGE has been applied to the separation of fluorescently-labeled DNA sequencing fragments. Comparisons between the CGE and SGE separation of sequencing mixtures, produced from a salmonella tRNA gene, has shown the resolution of CGE to be greater than that in SGE by as much as 2.7. In addition, the advantages in separation speed using high field strengths in CGE have been demonstrated. Drossman, et al., described a capillary-based system that produced sequencing rates as high as 1000 bases/hr, which is approximately 25 times greater than that achieved with traditional SGE. Karger's lab has reported DNA separations using a single-fluor system yielding sequencing rates of 450 bases/hr. More recently, authors have reported single lane sequencing rates up to 1000 bases/hr using capillary-based sequencing systems. Enhancements in data throughput are achieved by using sieving matrices which provide high efficiency with short analysis times and highly accurate base discrimination.

Most of the reports using CGE-based sequencers have employed custom designed apparati. This is in contrast to the tremendous commercial success of SGE and free solution capillary electrophoresis (CZE) systems, which have proven easily transferable to the commercial arena. Capillary gel-based systems have found only limited success in becoming commercial entities. The most prominent hindrance in the commercialization of automated CGE systems is the lack of robustness of the...
capillary gel column. Extrusion and bubble formation in the separation column lead to the loss of resolution and are common occurrences in capillary gel matrices under sequencing conditions. Many strategies have been utilized in an attempt to eliminate or minimize the occurrence of these phenomena.

2.5.3 Gel Immobilization to Fused-Silica Capillary Wall

A tendency exists for the polyacrylamide gel matrix to extrude from the end of the capillary during electrophoresis because of the extremely high field strengths applied across the capillary during the separation of oligonucleotides through the gel. Extrusion has been attributed to an electroosmotic force acting upon the gel matrix under an applied potential. Electroosmotic flow (EOF) is minimized by the presence of the gel matrix; however, the forces that generate EOF (excess (-) charge at the capillary surface due to the presence of SiO2 groups) are still present in the capillary and can manifest themselves in the occurrence of extrusion. This phenomenon is also encountered in SGE; however, extrusion in SGE is minimal because of the lower applied field strengths employed in that format which results in lower electroosmotic forces being generated. In the event of capillary gel extrusion, cutting the capillary tip can restore column performance, but this approach is limited because the capillary will eventually be shortened beyond a useful length.

Gel extrusion in CGE can be virtually eliminated by the covalent attachment of the gel matrix to the interior capillary wall. The covalent attachment of CPAs to fused silica capillaries was first demonstrated by Hjerten and several patents have been issued regarding the use of functionalized silane compounds to bind polyacrylamide to fused-silica capillaries. The covalent attachment of the gel, using Hjerten's procedure, is accomplished using a functionalized alkylsilane compound, γ-methacryloxy-trimethoxysilane (MAPS), to modify the interior wall of the fused silica capillary. A linear polyacrylamide (LPA) coating is then attached to the silane anchor. The LPA coating is necessary to prevent bubble formation in the cross-linked
polyacrylamide which serves as the sieving medium. Direct attachment of the CPA to the wall does not afford the flexibility needed to accommodate the shrinkage that occurs during gel polymerization (16%) and will result in bubble formation. The LPA coating imparts enough flexibility to the anchor to compensate for gel shrinkage and allows for the efficient production of gel columns. An illustration of the anchoring reaction is shown in Figure 2.9.

2.5.4 Gel Column Bubble Formation

While the formation of voids during the polymerization of immobilized polyacrylamide can be counteracted by the implementation of a non-crosslinked coating on the interior surface of the capillary, this action has no effect on bubble formation during electrophoresis.

![Schematic of gel anchoring reaction using a MAPS anchor and a polyacrylamide matrix.](image)

The typical lifetime of a CGE column is determined by the development of voids or bubbles along the column. The random formation of void areas in the cross-linked polyacrylamide matrices used in CGE presents a major problem for analysis using these columns. These gaps, shown in Figure 2.10, represent an area of high resistance to electrical current which leads to resistive heating and will cause convective currents to flow across the temperature gradient formed. This will increase the variance contribution due to parabolic temperature profile, $\sigma_{AT}$, leading to loss of
separation efficiency. A possible mechanism for the formation of these bubbles is that the high pH of the buffer solution employed (8.5 - 9.5) catalyzes the hydrolysis of the polyacrylamide bonds to monomer units which have lower density than the polymeric material. This hypothesis, while speculative, is supported by evidence that the degradation of gel matrices is slightly accelerated in buffer solutions with high pH.42

![Figure 2.10](image)

**Figure 2.10**
Picture of "bubble" formed in capillary filled with a polyacrylamide gel matrix. Capillary diameter is 75 µ and gel matrix consists of 6%T / 5%C polyacrylamide.

Susceptibility to hydrolysis can account for the "random" formation of bubbles in the column; however, experimentally there is a tendency to form voids at the injection end of the column. Voids present at the capillary tip can increase variance due to injection, σ, resulting in poor separation efficiency and reproducibility. These voids can also propagate through the column leading to premature column failure. The incidence of void formation at the column injector has been related to the conductivity differences between the gel matrix and sample buffer solution.41 Variations in buffer/gel conductivity also account for a decreasing current profile for a given gel column over time. Reductions in the occurrence of void formation at the injection tip can be achieved by closely matching the conductivity of the buffer solution to that of the gel matrix.
2.5.5 CGE Column Degradation

Sequential separations performed on gel-filled capillaries exhibit a bias toward poorer selectivity and decreasing efficiency after multiple runs have been performed on the same column. Decreasing efficiency can be attributed to three major factors: (1) electroosmosis, (2) bubble formation, and (3) capillary inlet fouling. When a cross-linked polyacrylamide filled column begins to deteriorate, very few options are available to correct the problem. Since bubble formation most commonly occurs at the injection tip of a column, clipping the first few millimeters can restore performance. However, this will practice can lead to variation during the injection of multiple samples since the cleaving process may not produce identical surfaces after each application. This process will also eventually create a capillary that is too short to be useful in analysis, and it provides no solution in the event of a bubble forming at the midpoint of a column. Another option is to simply replace the entire column after a predetermined number of runs. Cross-linked gel-filled columns are best suited for this type of application. They will typically allow between 10-15 runs before single-base resolution of sequencing fragments is lost. This option may be practical in a small scale sequencing application, but the effort required to produce, replace, realign and monitor (to prevent the occurrence of gross errors) in large-scale sequencing would likely be prohibitive. An example of the degradation encountered when performing separations on a polyacrylamide matrix is shown in Figure 2.11. A linear polyacrylamide gel-filled column is used for this example, since the degradation is more rapid and pronounced in linear polyacrylamide gels. The electropherograms show the loss in efficiency (calculated using Equation 2.4) exhibited when three sequential separations are performed without replacing the sieving media.
Sequential DNA separations performed on a linear polyacrylamide gel-filled capillary column. Near-IR-labeled primer (5 x 10⁻¹⁰ M) was injected for 5 seconds at 250 V/cm. Plate numbers obtained during electrophoresis were; 155K plates/m (A), 45K plates/m (B) and < 5K plates/m (C). Separations in A, B, and C were performed sequentially on an 6% linear polyacrylamide filled 75 μm i.d. capillary column.

To alleviate the problem of degrading matrices, investigators have been able to replace certain types of sieving matrices between electrophoretic separations.⁴⁴,⁴⁵,⁴⁷ These
sieving media have relatively low viscosity when compared to traditional cross-linked polyacrylamide, yet they possess the sieving power necessary to provide single base resolution during DNA separations in sequencing applications. The low viscosity of these media facilitates their removal and reinsertion after each electrophoretic run providing a pristine environment for each separation and increasing run-to-run reproducibility. Reports have shown that migration times using LPA can be controlled to within 2.4% RSD.\(^4\)\(^6\) Because of their ease of use and replaceability, low viscosity matrices greatly enhance the feasibility of constructing automated CGE analysis systems. Low viscosity matrices can be replaced without removal or realignment of the capillary which would facilitate high throughput analysis. In fact, there is a commercial CGE genetic analyzer available from ABI which employs replaceable matrix technology.\(^4\)\(^7\)

### 2.6 References

2.1 Collins, F. and Galas, D. Science 1993, 43.


2.12 Ferguson, K. A. *Metabolism* 1964, 13, 985.
2.15 Lumpkin, O. J. and Zimm, B. H. *Biopolymers* 1982, 21, 2315-2316.


2.44 Sudor, J; Foret, F. Bocek, P. Electrophoresis 1991, 12, 1056-1058.


2.47 ABI PRISM 310 Genetic Analyzer®, Perkin Elmer Applied Biosystems Division, 850 Lincoln Centre Dr. Foster City, CA 94404 (www.perkin-elmer.com)
3 Gel Electrophoresis Of Dye-Labeled DNA Fragments Using Near-IR Fluorescence Detection

3.1 Introduction

The use of fluorescent probes for detection is a viable alternative to radiolabeling; thus, a great deal of work has been performed using fluorescent probes in DNA sequencing applications. Many detection systems employ four visible dyes with each fluorescent dye corresponding to a particular terminal base. Applied Biosystems, DuPont, and Pharmacia have produced automated sequencers based on dye-primer and/or dye-labeled dideoxynucleotide technologies. One difficulty that must be overcome, when using a four fluor labeling approach to DNA sequencing, is the requirement that the labels be structurally similar yet spectrally distinct. The labels must be structurally similar in order to minimize the perturbations on the intrinsic electrophoretic mobilities of the oligonucleotide fragments imparted by the dye. Failure to regulate perturbations in the oligonucleotide mobilities can potentially lead to “flip flopping” especially for longer DNAs which possess similar μ∞. This phenomenon occurs when a short fragment exhibits a longer migration time than a larger fragment due to the influence of the label, which can result in errors in sequence reconstruction. Therefore, using dyes which possess similar structures is necessary to minimize the effect on migration perturbations. However, this presents a “catch-22”, since dyes which possess similar structures tend to possess similar spectral properties. Overlap in the fluorescence emission (leading to cross-talk) causes difficulties in distinguishing between the different fluorophores as they migrate through the gel matrix. This can lead to base misidentification and errors in sequence reconstruction during analysis. Multi-fluorophore protocols have been used in multi-lane formats. Multi-lane formats can relieve the cross-talk problem; however, this format reduces throughput because four lanes must be used to determine the sequence of a single DNA strand.
In order to eliminate the necessity for multiple separation lanes and/or multiple fluorescent probes in DNA sequencing applications, Richardson and Tabor\textsuperscript{8-10} and independently Ansorge\textsuperscript{11} developed a single label, single lane sequencing protocol. In this method, a modified T7 DNA polymerase enzyme is used in conjunction with manganese ions to uniformly terminate the reaction products to within 15-25%. Determining the identity of the terminating base, when using fluorescence detection, is accomplished by adjusting the molar ratios of the ddNTPs in the chain extension reaction and comparison of the resulting peak amplitudes using steady-state measurements. Using a molar ratio for T:G:C:A of 8:4:2:1, successful sequencing of an M13mp18 template with a single dye, tetramethylrhodamine isothiocyanate, and a single capillary has been demonstrated.\textsuperscript{12-13} Difficulties with this technique result from several factors: (1) Non-uniformity in the incorporation efficiencies of the different nucleotides. In order to observe the desired peak heights, it is necessary to modify the ddNTP concentrations to account for differences in the incorporation efficiencies. (2) \textit{The use of dye-labeled ddNTPs is not viable}. Modified terminators are not incorporated uniformly by DNA polymerase enzymes and cannot be adjusted to generate peaks with specific intensities; therefore, this approach is restricted to the use of dye-labeled primers. Dye-primers present in solution at high concentration (necessary for extension) will reduce detection sensitivity for low molecular weight oligonucleotides during the early stages of electrophoresis. (3) “Ghost” peaks, which result from pauses in the polymerization reaction, can appear causing variation in peak amplitudes. (4) Identification of electrophoretic peaks resulting from the lowest concentration of ddNTPs is difficult when they occur between bands which contain the highest concentrations of ddNTPs. (5) Inadequate resolution for longer fragments (>250 bp) can make identification difficult; thus, long reads are increasingly inaccurate. (6) This protocol is restricted to the use of the modified T7 DNA polymerase enzyme, which uniformly incorporates unmodified dNTP/ddNTPs, but
does not allow the use of cycle sequencing protocols and limits the size of templates that can be analyzed. The accuracy of DNA sequence determination using the four-level, single dye, single lane approach has been estimated to be approximately 90% for oligonucleotides shorter than 250 bases. Therefore, the template must be sequenced several times in order to reduce the error rate, inevitably reducing throughput.

To increase the accuracy of sequence determination using peak intensities, an alternative technique has been suggested. The method uses a two-level approach, in which two different dyes are used to label primers with the molar ratios of the ddNTPs adjusted to 2:1. Two separate chain extension reactions are performed and combined prior to electrophoresis. The accuracy of the sequence determination was found to be 97.5% for the first 350 bases of the template read. But, as with the four-level system, the two level system is limited to the use of the modified T7 DNA polymerase enzyme and dye-labeled primers. In addition, this approach requires two detection channels which adds to instrumental costs and complexity. In another report, a two level, binary coding system was used to fluorescently label oligonucleotides produced using the T7 enzyme. These authors employed two different fluorescent probes to label sequencing fragments. Three sets of fragments were generated. The first set (terminated in A) contained DNA labeled with both (525 nm “JOE” and 590 nm “FAM”) fluorophores and was coded [1,1], the second set was terminated with G and was labeled with JOE only and was coded [0,1]. The third set of fragments (T) was labeled with FAM and coded [1,0]. C-terminated fragments were not generated in this experiment and are represented by gaps in the electropherogram and are thus coded [0,0]. The accuracy of base-calling using this system was 94.6% for 280 bases, which is comparable with other reports in the literature. However, the accuracy of this system suffers from factors such as label-induced mobility shifts of the migrating DNA which can cause errors in sequence.
reconstruction and must be compensated for using sophisticated software algorithms. Spectral “cross-talk” between the probes is also encountered in this system which can produce erroneous signals during detection.

It is generally accepted that large scale genomic sequencing will require parallel running of multiple separation lanes in an automated fashion to increase overall throughput. If it is assumed that the rate limiting step in obtaining raw sequencing data is the electrophoretic separation and that capillary gel electrophoresis with laser-induced fluorescence detection can produce approximately 600 bases/hour of raw data per lane, sequencing rates approaching 1 megabase (1 million bases) per day would require 70 independent lanes. Several approaches have been suggested for the construction of a high throughput device. In one approach, referred to as “single channel scale up”, a focusing optic, collection optic, spatial filter, bandpass filter and photodetector are simply repeated for each required capillary channel with laser excitation provided by one or more lasers distributed to each lane. Instrumental complexity and optical realignment along with the extensive hardware necessary for mounting optics would make this approach challenging and may reduce the net throughput due to time required for system reoptimization when capillary replacement is required.

Mathics and co-workers have proposed a confocal microscope arrangement for monitoring the sequencing data in multiple capillary columns requiring only two detection channels per device using a binary coding system for base identification.15,16 The excitation light and collection optic scans over the capillary array and monitors the sequencing data in each column sequentially. Increasing the number of channels in the array reduces the sampling time per channel, resulting in less signal accumulated and a potential loss in signal-to-noise.

A capillary array system using four fluorescent labels and a sheath flow detector has been reported.17,18 The two lasers, required to excite the four probes,
were directed down the length of a series of sheath flows. The sheath flow system was employed to eliminate scattering contributions from the capillary used for separation. It also provides a post-column concentration chamber which facilitates the detection of dilute species. Also, the rate of the sheath flow must be precisely controlled due to its influence on detector volume during analysis. Failure to regulate this flow can have detrimental effects on detection and separation efficiency.\textsuperscript{18} For multiple columns, alignment of the sheath flows, attenuation of the laser beam over long propagation distances along with the sophisticated imaging optics and multiple lasers can present difficulties in this approach.

The commonality in all fluorescence-based sequencing methods and multiplexing systems is the need for ultrasensitive detection using visible fluorescence, especially in capillary formats where the amount of material loaded onto the column is commonly in the low attomole range. The ability to detect molecules in solution using fluorescence at very low levels (\(< 2000\) molecules) will offer important benefits to DNA sequencing. For example, single ultrasensitive detection in DNA sequencing can significantly reduce the amount of template molecules needed for analysis, which could eliminate the need for PCR or recombinant amplification steps. In addition, ultrasensitive detection can reduce the amount of expensive reagents used in chain extension reactions. One difficulty associated with visible ultrasensitive fluorescence measurements is the presence of a large number of interferences, especially when detecting dyes in complex biological matrices, which can degrade SNR.

Near-IR fluorescence has recently been shown to be a viable alternative to visible fluorescence detection in many bioanalytical applications such as fluoroimmunoassays,\textsuperscript{19-21} high performance liquid chromatography,\textsuperscript{22-25} the analysis of intact proteins\textsuperscript{19} and free solution capillary electrophoresis.\textsuperscript{27-29} As a demonstration of the impressive detection sensitivity that can be achieved in this
region, several groups have demonstrated the efficient detection of single molecules in the near-IR.$^{30,31}$ The motivation for using fluorescence monitoring in the near-IR includes the smaller backgrounds observed during signal collection and the simpler instrumentation required to carry out detection. The lower background is a consequence of the fact that few species fluoresce in the near-IR, thereby reducing these interferences. The $1/\lambda^4$ dependence of the Raman cross-section also provides a lower scatter contribution at these longer excitation wavelengths and is shown in Equation 3.1;

$$\eta = \frac{2\pi c k_B T}{\lambda^4} \quad (3.1)$$

where $\eta =$ emittance $(J/m^3s)$, $c =$ speed of light $(m/s)$, $k_B =$ Boltzman's constant $(J/K)$, $T =$ absolute temperature $(K)$ and $\lambda =$ wavelength of impinging light $(m)$. The reduction of scattered light generated when using near-IR excitation (795 nm GaAlAs beam) versus visible (488 nm Ar ion beam) is 7 fold (calculated using Equation 3.1). Near-IR detection also provides a larger Raman free observation window when compared to that observed using 488 nm excitation. The Raman spectra for polyacrylamide using 514 nm excitation is shown in Figure 3.1$^{32}$ The position of these Raman bands are then calculated for 488 nm and 795 nm excitation (see Figure 3.2). As can be seen in Figure 3.2, near-IR excitation allows greater exclusion of scattering photons from the detection system. This results in a lower background which enhances the detection of ultradilute samples. The spectral filter system used in this work, which takes advantage of the larger Raman-free region, is shown in Figure 3.3.

The instrumentation required for near-IR fluorescence detection can consist of simple and inexpensive diode lasers and avalanche photodiodes. Since both of these components are solid-state, the detector can be run for extended periods of time requiring little maintenance or operator expertise. The use of a visible red-diode laser

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
as an excitation source for detection of a cyanine dye-labeled oligonucleotide in CGE has recently been demonstrated.\textsuperscript{33} Near-IR detection has also been demonstrated in DNA sequencing applications using SGE.\textsuperscript{34,35} These authors were able to obtain a detection limit of 2000 molecules in an unpolymerized polyacrylamide gel solution.

![Raman spectrum of acrylamide](image)

**Figure 3.1**
Raman spectrum of acrylamide.\textsuperscript{32}

![Spacing between Raman bands](image)

**Figure 3.2**
Spacing between Raman bands when using 488 nm (green) or 795 nm (red) laser excitation. Each dark vertical line denotes the position of a Raman signal which was calculated from data given in Figure 3.1. Most near-IR dyes have emissions near 810 nm while visible dyes have emissions near 500 nm.
Figure 3.3
Absorbance spectrum of filter system used for near-IR DNA sequencing experiments. Position of dye emission, laser excitation and Raman lines are shown in red. Absorbance measurements taken on a Perkin Elmer spectrophotometer.

Here, I will present my results demonstrating the first use of near-IR fluorescence in DNA sequencing in conjunction with CGE in a single-lane, single-fluor base identification protocol using a tricarbocyanine-labeled sequencing primer. Terminal base identification was accomplished using fragment band intensities which had been adjusted by modifying the ddNTP concentrations during Sanger extension to yield a 4:2:1:0 (A:C:G:T) intensity pattern. A critical comparison of on-column detection limits for a visible and near-IR labeled primer in CGE will be made as well. I will also present results demonstrating the use of near-IR labeled dideoxynucleotide terminators in DNA sequencing reactions.
The integration of near-IR labeled terminators is an integral step in the development of a near-IR based sequencing system that can take advantage of multi-labeled, single-lane sequencing techniques, because DNA extensions using labeled ddNTP terminators allow all chain extensions to be performed in a single reaction vessel unlike multi-fluor primer-labeled sequencing reactions which require separate reaction mixtures to be pooled before analysis. Pooling the reaction mixtures will cause dilution of the sequencing fragments making detection difficult. I will also present preliminary data which demonstrates the feasibility of using near-IR labeled ddNTPs in Sanger termination reactions.

3.2 Experimental

3.2.1 LIF-CGE System

Electrophoresis and detection were performed using a custom-designed CE system which is shown in Figure 3.4. The 10 mW excitation beam (795 nm), supplied by a Ti:Sapphire laser (Coherent Lasers Mira 900-F, Palo Alto, CA), was pumped by the all-lines output of a small frame Ar ion laser (Coherent Lasers, Innova 310) and focused to a 10 \( \mu \text{m} \) diameter spot inside the capillary (1/e²). For visible excitation, the 488 nm line of the Ar ion laser was used. The fluorescence was collected at right angles with a 40X high numerical aperture microscope objective (Nikon, Natick, MA; NA=0.85) and spatially filtered through a slit with a width set at 2 mm giving a viewing distance of 50 \( \mu \text{m} \) within the capillary. For near-IR fluorescence, a filter stack consisting of a Corning long pass filter (50% T at 820 nm) and a bandpass interference filter (CWL = 830 nm, HBW = 15 nm, Omega Optical, Brattleboro, VT) were used. In the visible case, a Corning longpass filter (50% T at 510 nm) and a bandpass interference filter (CWL = 530 nm, HBW = 10 nm, Omega Optical) served as the spectral filters. The light was then focused onto the photodetector by a 10X microscope objective. The detector was a single photon avalanche diode (SPAD, EG&G Electrooptics Canada, Vaudreuil, Canada) mounted...
on a thermoelectric cooler with a photoactive area of $1.77 \times 10^{-4} \text{cm}^2$ (i.d.=150 \(\mu\text{m}\)). The collected signal from the photodetector was conditioned using a discriminator (Tennelec Nucleus, TC 754) and sent to a multichannel scalar resident in a 486 PC for displaying the electropherograms (Tenelec PCA II, Oak Ridge, TN). Photocounts were typically integrated for 1 second during electrophoresis. The displayed data were not subjected to a filtering algorithm and is presented as raw data.

**Figure 3.4**
Block diagram of the capillary electrophoresis system. M = mirrors, L = focusing lens, C = capillary, BL = beam block, MO\(_1\) and MO\(_2\) = microscope objectives, SF = spatial filter, BF = bandpass filter, SPAD = single photon avalanche diode, DISC = discriminator.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
3.2.2 Capillary Gel Column Preparation

The column used for the separations was a 75 μ i.d., 375 μm o.d. fused-silica capillary (Polymicro, Phoenix, AZ). Before the column was filled with the polyacrylamide gel solution, a small portion of the polyimide coating was removed to provide an optical window by use of a resistively heated coil. The interior wall of the capillary was pretreated with 1M NaOH (10 min.), double distilled water (10 min.), 1M HCl (10 min.) and finally, double distilled water (10 min.). A mixture of [3-(methacryloxy)propyl] trimethoxysilane (Aldrich Chemical Company) with methanol (50/50) was placed into the column and allowed to react overnight. The column was then evacuated using low pressure and the gel forming solution injected into the capillary. The gel forming solution (3%T / 3%C or 6%T / 5%C) consisted of 0.1 M EDTA, 8M urea and riboflavin (Sigma Chemicals, St. Louis, MO), which was used as the photo-polymerization initiator. Preparation of the polyacrylamide gel column was accomplished by vacuum injection of the gel forming solution into the pre-treated capillary and followed by photo-initiated polymerization. The ends of the capillary were capped and the column submerged in an ice bath which was exposed to a high energy UV light source overnight. After polymerization, the column was trimmed to a total length of 60 cm with an effective length of 50 cm (injection to detection). The column was pre-run for approximately 30 minutes at 6 kV prior to sample injection.

3.2.3 Near-IR Primer-Labeled Sequencing

Extension of the near-IR dye-labeled primer was accomplished using a Sequenase® Version 2.0 DNA sequencing kit (United States Biochemical, Cleveland, OH). Two pmoles of template DNA were annealed to 1 pmole of near-IR labeled M13mp18 primer (LiCor, Lincoln, NE) at 65°C for 2 minutes. Figure 3.5 shows the
structure of the near-IR labeled primer as well as a fluorescein-labeled primer.

![Chemical structures of near-IR labeled and fluorescein-labeled primers.](image)

**Figure 3.5**
Structure of near-IR and Fluorescein labeled primers

This mixture was allowed to cool slowly to room temperature over a period of thirty minutes. After cooling, 0.2 μmol dithiothreitol (DTT), 0.1 μmol MnCl₂, 2 μL of 1:5 diluted dNTP solution, 0.004 units pyrophosphatase, and 0.4 units of Sequenase enzyme were added. The mixture was allowed to sit at room temperature for 5 minutes. A 3.5 μL aliquot of this mixture was placed in a tube containing 3.2 pmols ddATP, 1.6 pmols ddCTP, 0.8 pmols ddGTP and 0 pmols ddTTP to achieve a ratio of 4:2:1:0 for A, C, G and T terminators. Additional volumes of dNTPs were added to achieve an overall dNTP:ddNTP ratio of 1200:1. This mixture was incubated at 37°C for 30 minutes before the addition of 4 μL of stop solution (95% formamide).

### 3.2.4 Near-IR ddNTP Labeled Sequencing

Near-IR ddATP labeled extension reactions were performed using standard Sanger sequencing protocols and the Vent® DNA polymerase enzyme. To begin, 1
μL of a $^{32}$P labeled-M13mp18 primer (180 μM) was added to a (1 μM) 1 μL volume of M13 control DNA. This mixture was denatured by raising the temperature to 95°C and cooled slowly to room temperature (5-10 minutes) to allow annealing. After annealing, 4 μL of ddH₂O, 3 μL of 1X buffer solution, 1 μL (80 μM) of each dNTP, 1 μL of dATP and 1 μL of near-IR labeled-ddATP (Figure 3.6) were then added to the solution.

![Figure 3.6](image)

**Figure 3.6**
Structure of near-IR labeled dideoxynucleotide (A) terminator

This mixture was extended at 72 °C for 2 minutes. The reaction was stopped by the addition of 4 μL of a 95% formamide solution. The products of this reaction were denatured at 95 °C for 3 minutes before loading onto a SGE system for separation. Detection in this experiment was performed using autoradiographic techniques.
Autoradiography was employed in this part of the experiment because it allowed the analysis of multiple reaction mixtures on a single gel which minimizes matrix variations when evaluating the quality of various sequencing reaction products. Autoradiography also allowed a comparison of extension reactions performed with near-IR labeled and unlabeled ddATP terminators.

### 3.3 Results and Discussion

The fluorescence emission of many tricarbocyanine dyes intimately depends on the solvent and has been shown to increase when placed in less hydrogen bonding or more non-polar solvents.\(^3^8\) Recently, it has been shown that several red-emitting rhodamine dyes display enhanced fluorescence when dissolved in 7M urea as compared to pure aqueous solutions.\(^3^9\) In Figure 3.7, the emission spectra of the near-IR dye-labeled primer is shown in methanol, water and unpolymerized denaturing gel solutions containing urea or formamide. In methanol, a quantum yield of 0.17 was determined for the near-IR labeled-primer construct and in pure water, the quantum yield was reduced to less than 0.01. When the dye-primer was placed in the denaturing gel solutions, the quantum yield was found to be approximately 0.09, significantly better than that observed in the pure aqueous solvent. The presence of the high organic content resulting from the formamide or urea helps in reducing fluorescence quenching that was observed in the pure aqueous solvent. Since many tricarbocyanine dyes display quenching effects resulting from photoisomerization, the increased viscosity of the unpolymerized gel solution may result in the increased quantum yield. However, our previous research has demonstrated that the bridged tricarbocyanine dyes show negligible solvent viscosity-dependent photophysics.\(^4^0\) These results indicate that increased viscosity associated with the gel solution does not result in the observed fluorescence enhancement.
Figure 3.7
Fluorescence spectra of near-IR dye labeled primers in methanol, 30% formamide, 8M urea and water. Dye-primer concentration was 1x10^-6 M and the excitation wavelength was set at 725 nm in all cases. The emission spectra were acquired on a conventional spectrofluorometer (SPEX Fluorolog) using a Xe lamp for excitation.
In order to compare the limits of detection of our near-IR labeled primer to that of a primer labeled with the visible fluorescent dye, fluorescein, the near-IR and visible primers were electrophoresed in a capillary column containing a 6% T / 5% C polyacrylamide gel. The resulting electropherograms are shown in Figure 3.8 and a summary of the data is presented in Table 3.1. In order to achieve comparable backgrounds levels, the spatial filter slit width was reduced to 0.2 mm and the laser power was dropped to approximately 3 mW in the visible case (488 nm). When the laser powers and slit widths were set to approximately the same values, the background level generated from the capillary gel matrix was found to be greater than 200,000 cps for visible excitation while in the near-IR case, the background was only 10,000 cps. The 20-fold reduction in background signal exceeds the 7-fold difference projected by Equation 3.1. This deviation from the theoretical prediction can be accounted for by the additional influence of fluorescent impurities present in the sample and sieving matrix (polyacrylamide), which are efficiently excited in the visible, but not in the near-IR case. Since fewer species fluoresce in the near-IR region, the contribution to the background signal from contaminants is minimal when near-IR fluorescence is employed. Detection in the visible region is greatly hindered by fluorescent contaminants, especially when analyzing biological samples such as DNA, proteins or amino acids. The observed difference in the background counting rate did not result from differences in the efficiency of the photodetector at the monitored wavelengths, since the silicon avalanche diode has a larger single-photon detection efficiency at 800 nm than it does at 550 nm. The lower background level in the near-IR was observed in spite of the larger spectral window that was associated with the fluorescence bandpass filters (30 nm for near-IR and 10 nm for visible). In addition, there was a 42 nm shift between the center wavelength of the interference filter and the excitation wavelength in the visible case, while for the near-IR, only 34
nm separated the excitation wavelength and the center wavelength of the interference filter.

Figure 3.8
CGE separation and on-column detection of a $5.0 \times 10^{-9}$ M solution of FITC-labeled M13mp18 universal primer and a $1.3 \times 10^{-9}$ M solution of near-IR labeled M13mp18 universal primer. The laser power was set at 10mW for the near-IR case and 3 mW for visible excitation. The field strength in both cases was 250 V/cm. The samples were electrophoretically injected onto the gel column for 10 seconds at 5 kV. The gel column consisted of a 6%T / 5%C polyacrylamide matrix.
Table 3.1
Data from electropherograms in Figure 3.8. *The slit width on the LIF detector used in these experiments was set at 2.0 mm (near-IR) and 0.2 mm (visible) with a laser power of 10 mW (near-IR) and 3 mW (visible) in order to give comparable backgrounds. bNet signal was calculated as peak height (cps) minus the average background (cps). *Mobility was calculated using Equation 3.2

<table>
<thead>
<tr>
<th></th>
<th>near-IR Primer</th>
<th>FITC Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Concentration (M)</td>
<td>1.3 x 10⁻⁹</td>
<td>5.0 x 10⁻⁹</td>
</tr>
<tr>
<td>Migration Time (s)</td>
<td>2555</td>
<td>660</td>
</tr>
<tr>
<td>Mobility (cm²/Vs)*</td>
<td>5.6 x 10⁻⁵</td>
<td>2.1 x 10⁻⁴</td>
</tr>
<tr>
<td>Injection volume (L)</td>
<td>2.9 x 10⁻⁹</td>
<td>1.1 x 10⁻⁸</td>
</tr>
<tr>
<td>Moles Injected</td>
<td>3.8 x 10⁻¹⁹</td>
<td>5.8 x 10⁻¹⁷</td>
</tr>
<tr>
<td>Net Signal (cps)a,b</td>
<td>33,500</td>
<td>15,920</td>
</tr>
<tr>
<td>Background (cps)</td>
<td>10,000</td>
<td>19,490</td>
</tr>
<tr>
<td>SNR</td>
<td>335</td>
<td>114</td>
</tr>
<tr>
<td>Detection limit</td>
<td>3.4 x 10⁻²⁰ moles</td>
<td>1.5 x 10⁻¹⁸ moles</td>
</tr>
<tr>
<td>Quantum yield of Dye</td>
<td>0.07</td>
<td>0.90</td>
</tr>
</tbody>
</table>

As can be seen in Figure 3.8, a significant change in the apparent mobility of the near-IR labeled primer was observed when compared to the fluorescein primer. Both electropherograms were run using the same capillary gel column and a field strength of 250 V/cm. The apparent electrophoretic mobility can be calculated from experimental data using Equation 3.2 which is shown here:

\[ v = \frac{L_{\text{eff}} L_{\text{tot}}}{V \cdot t} \]  

(3.2)

where \( L_{\text{eff}} \) is the capillary effective length, \( L_{\text{tot}} \) is the total capillary length, \( V \) is the applied voltage and \( t \) is the elution time. The apparent mobilities were found to be \( 5.6 \times 10^{-5} \text{ cm}^2/\text{Vs} \) for the near-IR primer and \( 2.1 \times 10^{-4} \text{ cm}^2/\text{Vs} \) for the visible primer. A
reduction in the mobility of the near-IR labeled primer may result from the extended length of the near-IR dye-label. Molecular modeling of this dye has indicated an overall length of approximately 17 Å.

One can determine the number of moles (Q) loaded onto the capillary column by using Equation 3.3;

\[ Q = \frac{L_{inj}}{t} V_{inj} A C t_{inj} \left( \frac{\text{liter}}{1000 \text{cm}^3} \right), \]  \hspace{1cm} (3.3)

where \( V_{inj} \) is the injection voltage, \( A \) is the cross-sectional area of the capillary column, \( C \) is the sample concentration, \( t_{inj} \) is the time allowed for injection. Comparison of the on-column detection limits, calculated using Equation 3.4;

\[ SNR = \frac{S_{net}}{\sqrt{N}}, \]  \hspace{1cm} (3.4)

where \( S_{net} \) is the net signal (maximum signal minus background) and \( N \) is the average background, (SNR = 3, see Table 3.1), which indicated a significant improvement for the near-IR primer with a mass detection limit of \( 3.4 \times 10^{-20} \) moles (34 zmol) for the near-IR primer and \( 1.5 \times 10^{-18} \) moles (1.5 amols) for the visible primer. While the laser power used in the near-IR case is near the saturation point of the electronic transition producing optimal signal-to-noise,\(^2\) it is not in the visible case due to the smaller absorption cross section associated with fluorescein and the lower irradiance used. Improvement in the detection limit would be expected with higher laser powers. In spite of the lower quantum yield exhibited by the near-IR dye, we were able to realize an approximate 40-fold improvement in detection sensitivity, resulting primarily from the significantly smaller background observed in the near-IR region. Based on the size of the effective sampling zone used in these experiments (10 μ x 50 μ) and the cross-sectional area of the capillary tube (4,418 μm²), the detection limit for only those near-IR labeled molecules entering the sampling zone is 3.7 zmol (2250 molecules), comparable to state-of-the-art detection limits using off-column detection.
with a sheath-flow cuvette system. These detection limits are comparable to our previously reported results for several near-IR tricarbocyanine dyes separated by free-solution CE in methanol/water buffer systems. The present results indicate minimal degradation of SNR arising from the complex gel matrix when using near-IR monitoring.

Preliminary attempts to generate sequencing fragments using near-IR labeled primers yielded favorable results, as depicted in Figure 3.9. This electropherogram, produced using a near-IR labeled M13 universal primer (see Figure 3.5 for structure), shows extension of the primer out to approximately 200 bases. This demonstrates that neither the size nor the slight hydrophobic character associated with the dye-label is detrimental to the action of the DNA polymerase. The absence of polymerase retardation by the near-IR probe is not surprising due to its spatial distance from the enzymatic binding site (minimum 17 bases). The declining intensity of the later eluting bands is a function of two parameters. First, the lower relative population of long fragments in the sample with respect to the concentration of short fragments is partly due to the processive action of the DNA polymerase enzyme. After annealing, each primer is extended by the enzyme until a terminating nucleotide (ddNTP) is incorporated into the growing strand which ceases extension. The probability of primer extension without encountering a terminating species decreases as the fragment length increases. The population of these fragments in an electrophoretic band is directly related to the intensity of that band when using direct detection methods. This phenomenon is also encountered in autoradiographic detection schemes, but is not a critical parameter in those techniques since band intensities are not used for terminal base identification. Second, a peak intensity bias is introduced by the use of electrokinetic injection which is the only injection mode currently available to CGE. The total volume of short fragments loaded onto the column is greater than that of long fragments due to their faster migration rate in the gel matrix.
The injection bias incurred when injecting a 20 base fragment vs. a 200 base fragment is shown in Table 3.2. Intensity regulation is of significant importance in spectral base identification and is critical in intensity-based discrimination. The population of longer fragments can be regulated by adjusting the ratio of dNTPs to ddNTPs during extension. Higher dNTP/ddNTP ratios (1000-1500) promote the generation of longer sequencing fragments; however, this increase comes at the expense of the shorter fragment population. A dNTP/ddNTP ratio of 1200:1 proved optimal for sequencing reactions in this work and was used for subsequent reactions.
Table 3.2
An assessment of the bias associated with the electrokinetic injection of DNA fragments in capillary gel electrophoretic separations. Concentrations of 20mer and 200mer are assumed to be $1 \times 10^{-12}$ M. *Calculated from electropherogram in Figure 3.9.

<table>
<thead>
<tr>
<th></th>
<th>20 mer</th>
<th>200 mer</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{app}$ (cm²/sV)*</td>
<td>$1.7 \times 10^{-4}$</td>
<td>$0.72 \times 10^{-4}$</td>
</tr>
<tr>
<td>Elution time (sec)</td>
<td>1150</td>
<td>2750</td>
</tr>
<tr>
<td>$t_{inj}$ (sec)</td>
<td>60</td>
<td>60</td>
</tr>
<tr>
<td>V$_{inj}$ (V/cm)</td>
<td>250</td>
<td>250</td>
</tr>
<tr>
<td>Q$_{inj}$ (molecules)</td>
<td><strong>83,000</strong></td>
<td><strong>35,000</strong></td>
</tr>
</tbody>
</table>

In order to demonstrate the efficacy of using near-IR fluorescence in actual CGE DNA sequencing applications, we performed sequencing experiments with the M13mp18 plasmid and the single lane, single fluor peak height identification protocol using ddNTP concentration ratios 4:2:1:0 (A:C:G:T). The result of a sequencing run is shown in Figure 3.10. The gel concentration was reduced to 3%T / 3%C in order to increase the apparent mobility of the dye-labeled sequencing fragments and to shorten the analysis time. The efficiency for the “*” labeled A-terminated peak was determined to be $2.0 \times 10^6$ plates with a resolution of 0.8 between the adjacent A-terminated oligonucleotide using Equations 2.3 and 2.7.

Intensities of the peaks in the 2800-3200 time frame averaged $A = 26951$ cps ± 809, $C = 19770$ cps ± 791, $G = 16320$ cps ± 1469. The sequence of the bases, determined from the unconditioned electropherograms, were compared to the known sequence of the M13mp18 template using the Genetics Computer Group (GCG Madison, WI) program Pileup®. Pileup® compares experimental DNA sequence data with the known data. Overlap is performed by realigning the given data set in a manner to create the greatest possible number of matches between the two data sets.
Penalties are assigned when gaps must be added to either the input data or known set to enhance overlap. Application of Pileup® analysis to the data generated in these experiments yielded an overlap rate of 84% (gap creation penalty = 0.10, gap extension penalty = 1) reading up to 250 bases from the primer annealing site. The sequencing data could be read to 400 bases per run, but the base-calling accuracy deteriorated significantly (60%). One can predict the maximum obtainable accuracy by considering the standard deviations of the peak heights. Using a 2σ window (95% confidence for each called base) overall base-calling accuracy should approach 83% (95.5% for A x 95.5% for C x 95.5% for G x 95.5% for T) for a single run. This level of base-calling accuracy was achieved for short read lengths; however accuracy
deteriorated quickly for long reads. The lower accuracy encountered for long reads can be explained by considering a few additional sources of error during base-calling. One source of error in the 4:2:1:0 ratio base-calling scheme arises from the necessity to use null signals to identify T-terminated bases. Ambiguity is present when multiple null signals are encountered, which can lead to insertions or deletions. An additional problem occurs when a null signal appears between two large peaks causing the null signal to be missed during sequence reconstruction. This will increase the error rate during DNA sequence analysis. Another source of error results from the variance in the activity of the T7 DNA polymerase enzyme used to incorporate the ddNTPs at the desired 4:2:1:0 ratios. The incorporation efficiency of the various ddNTPs using this enzyme can vary by as much as 25%. This variation is evident in the data obtained during the separation performed in this work and can lead to the misidentification of terminal bases. However, the data obtained does show that near-IR fluorescent labeling and detection can be used as effectively as visible fluorescence for primer labeled sequencing application. Near-IR detection provides lower detection limits than visible detection and allows comparable base-calling accuracy when using peak-height discrimination techniques.

My interests next focused on the integration of near-IR labeled ddATPs into Sanger DNA sequencing protocols. The integration of near-IR ddNTPs into Sanger chain extension protocols is an integral step in the development of near-IR based sequencing techniques. The ability to use near-IR terminators would allow multiprobe, single-batch chain extension followed by single lane analysis, analogous to that used with visible, multi-fluor terminators. Near-IR terminators would also allow the development of a near-IR based, multiprobe sequencing system that does not require pooled reaction products, such as in primer labeled protocols, which degrade detection limits through dilution. Since near-IR labeled terminators are not commercially available, they were synthesized by workers in our laboratory. Since
the exact concentration of the labeled near-IR ddATP was not known, a battery of experiments to determine the viability and optimal concentration of near-IR ddATP for chain extension were performed. The products of chain extension reactions using varying dilutions of a near-IR ddATP stock solution were electrophoresed and the results are shown in Figure 3.11. The banding patterns shown in lanes 2-5 exhibit behavior that is consistent with what would be expected when adjusting terminator concentration. Lane 3 shows a greater population of long fragments, consistent with a low terminator concentration during extension. This is evidenced by their slow migration through the gel matrix creating a dark band near the loading side of the gel. Lanes 4 and 5 show progressively shorter, faster moving fragments, which are formed during extension using high terminator concentrations which promote a higher probability of termination. An SGE autoradiogram of DNA polymerase products generated using near-IR ddATP and $^{32}$P-labeled primers along with products from a reaction employing unlabeled terminators is shown in Figure 3.12. First observations should note that the banding patterns generated during electrophoresis do not align with each other. There are a number of possible explanations for this behavior. The near-IR probe on the terminating nucleotide may induce misincorporations during polymerase extension. This would lead to erroneous sequence determinations. Another possible reason could be the influence of the near-IR probe on the mobility of the DNA fragments. Severe perturbations in fragment mobility will cause difficulties when aligning the banding patterns generated using labeled and unlabeled terminators. Further interpretation of the data in Figure 3.12 reveals that the incorporation efficiency (insertion during chain extension) of near-IR ddATP by the Vent enzyme is of similar magnitude to that of the non-labeled ddATP terminators. This statement is based on the observation that electrophoreic bands of similar intensity are generated using near-IR ddATP and ddATP in chain extension reactions. This suggests that similar quantities of terminated fragments are present in each electrophoreic band.
Figure 3.11
Autoradiogram of sequencing products generated using near-IR labeled ddATP and $^{32}$P labeled-primers. Samples were generated using various dilution of ddATP stock solution. Separation was performed using a 6%T / 5%T polyacrylamide gel on a vertical Bio-Rad SGE sequencer. Lanes 1 & 7 were control experiments using standard dNTP/ddATP. Lane 2 used a neat near-IR labeled ddATP termination solution. Lane 3 used a 1:100 dilution of near-IR labeled ddATP. Lane 4 used a 1:50 dilution of near-IR ddATP. Lane 5 used a 1:10 dilution and Lane 6 used a 1:2 of near-IR labeled ddATP. All other reaction conditions were held constant.
Figure 3.12
Separation of sequencing fragments generated using $^{32}$P labeled primers, ddATPs and near-IR ddATPs. Separation was performed on a Bio-Rad vertical SGE apparatus employing a 6%T / 5%C polyacrylamide gel matrix which contained 8M urea.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
3.4 Summary

The use of near-IR fluorescence in DNA sequencing applications using CGE can result in significant advancements in detection and instrumentation. The background resulting from scattering and fluorescence interferences are much smaller in the near-IR region, which results in impressive sensitivity without the need for off-column detection which can add to the complexity of the instrument. Since many near-IR dyes have absorption maxima which match fundamental lasing lines of inexpensive semiconductor diode lasers, sensitive fluorescence detectors can be constructed in multiplexing devices appropriate for handling large-scale sequencing projects, such as those encountered in the Human Genome Project at a fraction of the cost associated with Argon laser systems. In addition, these diode lasers can be constructed to deliver milliwatts to several watts of laser power and run continuously for long periods of time without the need for replacement of ion tubes. I have demonstrated that near-IR dye-labeled primers can be effectively used in a single-lane, single-fluor, peak height identification protocol in CGE applications. Effective reads could be obtained up to 400 bases from the primer annealing site in a single run. However, the error rates arising from misidentifications of the terminal nucleotide bases were significant at these read lengths. Accuracy rates of 84% using a 4:2:1:0 intensity ratio were achieved up to 250 bases. Significant advances must still be made to improve the base-calling accuracy in a single-lane, single-fluor format. I have also shown the presence of a near-IR probe covalently attached to a nucleotide terminator does not exclude the modified base from enzymatic incorporation during DNA sequencing reactions. An assessment of the effect of near-IR modification on the fidelity of terminator incorporation must be studied.

3.5 References


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


DNA Sequencing Reactions Performed in a Nanoliter Volume Reactor Directly Coupled to Capillary Gel Electrophoresis

4.1 Introduction

DNA sequencing basically involves 5 steps: (1) preparation of the chromosomal (or template) DNA; (2) generation of fragments complementary to the DNA template; (3) separation of fragments, usually by a gel electrophoretic technique; (4) detection of fragments and (5) data analysis (base calling). The acceleration of any of these steps would facilitate genetic analysis, especially for large scale projects such as the Human Genome Initiative.

The use of laser-induced fluorescence (LIF) detection of dye-labeled primers or nucleotide bases, instead of labor intensive and time consuming autoradiographic imaging has allowed an enormous decrease in the time needed to evaluate sequencing data and has reduced the number of lanes required to perform base-calling, thus increasing system throughput. Various LIF-based detection and base-calling schemes have been integrated into a number of commercially available DNA sequencing instruments. In fact, there are commercial devices which can provide essentially automated separation, detection and reconstruction of a particular DNA sequence, for example, the ABI 377 DNA sequencer, which employs a slab gel format and LIF detection using a CCD camera with base-calling accomplished using a four-color approach.

In addition to improvements in detection, several research groups have made tremendous advances in increasing the speed of DNA separations using capillary gel and recently, microchip electrophoresis. The principle advantages associated with capillary gel and microchip electrophoresis result from the fact that the surface-to-volume ratio is much larger than in conventional slab gel electrophoresis. As a result, higher field strengths can be used to drive the separation, thus decreasing the time required to effectively fractionate the DNA. Improvements in resolution are also encountered when comparing separations performed in capillaries versus slab...
gels. Reports have detailed a nearly 3-fold improvement in resolution for the CGE separation of sequencing fragments compared to SGE.\textsuperscript{13} Recently, micro-electrophoresis devices, which are prepared in glass substrates using photolithography, have been shown to reduce the separation time for DNA analysis compared to conventional capillary gel electrophoresis resulting from the use of shorter columns and better heat dissipation capabilities. In both of these micro-separation techniques, the amount of sample required for the separation ranges in the nanoliter regime.

While advances in separation, detection and base-calling have been numerous, they have outpaced the development of methodologies for preparing DNA polymerase reactions in a volume more commensurate with micro-column separation techniques. Standard DNA sequencing reaction protocols generate products in the \( \mu \text{L} \) volume range, therefore not exploiting the nanoliter (nL) sample requirements for capillary and microchip separations. To take full advantage of the extremely small sample volume requirements of capillary gel and other microscale separation techniques, sequencing protocols must produce volumes in the pL to nL regime. Miniaturized sequencing systems would benefit from the substantial reduction in the amounts of reagents and expendables employed in fragment generation. Low consumable costs for any sequencing protocol are of critical importance, especially if it is to be financially feasible for large scale sequencing projects, such as the Human Genome Initiative. For example, to sequence the human genome, six million sequencing reactions are required (assuming a 500 base average read length per electrophoretic run) for complete coverage. At $12.50 per reaction the cost of reagents alone would exceed $75 M using conventional scale preparation.

Solid-phase DNA sequencing methods possess several advantages which could potentially facilitate their integration into a miniaturized DNA sequencing system and facilitate the production of sub-\( \mu \text{L} \) sequencing products. Solid-phase reactors
have been described for a number of biochemical reactions such as enzymatic cleavage,\textsuperscript{14,15} oligonucleotide purification\textsuperscript{16} and DNA sequencing reactions.\textsuperscript{17,18} DNA sequencing using a solid support is advantageous in that it affords an easy method by which the products can be purified leading to a higher efficiency separation due to the removal of excess template, primers and salts which can degrade resolution.\textsuperscript{19} Solid-phase DNA sequencing protocols have typically employed the use of magnetic beads as the support media with species attached via a streptavidin-biotin\textsuperscript{20} or glutaraldehyde linkage.\textsuperscript{19,20} The streptavidin-biotin anchor is usually accomplished by attaching the streptavidin protein to the support then allowing a biotin modified molecule to bind with the immobilized streptavidin. The biotin-streptavidin linkage is a suitable choice for sequencing applications due to the strength of the streptavidin-biotin couple ($K_d = 10^{-15} \text{ M}^{-1}$)\textsuperscript{19} which can withstand the large and rapid temperature changes associated with sequencing reactions. While sequencing using magnetic beads as the support medium allows for simple purification of the reaction products, the system is not well suited for producing small volumes of material (nLs). A very effective solid-phase strategy has been described by M. Ulhen and coworkers which involves the immobilization of the DNA template onto an agarose support via a streptavidin-biotin linkage.\textsuperscript{21} While very effective in producing small volumes of sequencing fragments, this procedure does not provide a means for direct (on-line) connection to a separation medium as required for an autonomous sequencing system.

This work describes the use of a miniaturized solid-phase DNA polymerase reactor which can provide a reduction in the amounts of reagents and other consumables used in the preparation of Sanger, dideoxynucleotide-terminated DNA sequencing fragments. In addition, this system will be directly coupled to a capillary gel column for DNA fractionation, providing automated analysis of sequencing fragments on a nanoliter scale. The reactor consists of a conventional fused-silica capillary tube (volume = 30-200 nL), in which biotin is covalently attached to the
wall, with the subsequent addition of streptavidin to provide an anchoring point for the 5' biotinylated DNA template. Long term stability of the immobilized template under constant flow conditions and temperature variations will be reported. The preparation of sequencing fragments directly in the nano-reactor will also be demonstrated. Reagents needed to prepare these fragments are pumped in and out of the reactor using electropumping with the sequencing products injected directly onto a CGE column via low dead volume capillary interlock devices. In order to provide the necessary temperature conditions for denaturation, annealing and polymerase chain extension, the nano-reactor is placed in a thermocycler in order to take advantage of the rapid temperature transitions possible with such apparatus.  

The development of a nanoscale system to perform sequencing reactions will facilitate the creation of a completely autonomous sequencing system which can perform biochemical reactions, analyte fractionation and sample detection in a single unit. Complete DNA sequencing systems have recently been reported, however, these systems require μL sample volumes and do not exploit the cost and efficiency advantages of nanoscale analysis.

4.2 Experimental Section

4.2.1 Preparation of Biotinylated DNA

The carrier vector, pCRII (3.9 kb) (Invitrogen, San Diego, CA) was cleaved at a single restriction site using EcoRI. A 900 bp actin gene was inserted into the carrier vector using a T4 DNA ligase (Promega, Madison, WI). The DNA/pCRII construct was then amplified by the polymerase chain reaction (PCR) in a Perkin Elmer 2400 series thermocycler. PCR, developed in 1984 by Mullis, is a method by which specific DNA sequences can be amplified exponentially. Prior to the development of PCR, DNA amplification could only be accomplished using recombinant (bacterial) methods which are labor intensive and not amenable to automation. A description of PCR is shown in Figure 4.1.
The PCR mixture contained 1 μL pCRII vector, 10 μL 1X PCR buffer (20 mM Tris-HCl (pH=8.4), 50 mM KCl), 2 μL dNTPs, 85 μL double distilled water, 1 μL SP6 forward primer with biotinylated dATP (see Figure 4.2) present on the 5' end (1 μM) and 1 μL M13 reverse primer (1 μM). PCR was performed under "hot start" conditions, which entails addition of 1 μL of the Thermus aquaticus (Taq) DNA polymerase (Gibco BRL, Gaithersburg, MD) after the reaction temperature reached 80°C.
Figure 4.2
Biotinylated dATP used to immobilize the DNA template.
(L = (CH2)6NHC-O-(CH2)5)

The “hot start” procedure ensures high fidelity during DNA amplification. Twenty-five PCR cycles were performed using the following program: 1) denature dsDNA at 94°C for 45 seconds; 2) anneal primers at 55°C for 30 seconds; 3) extend primers at 72°C for 90 seconds. The reaction products were separated by agarose gel electrophoresis (0.8 % agarose, 25V/cm, 1 μM EtBr), visualized under a UV lamp and excised from the gel matrix using a scalpel (see Figure 4.3). The DNA was removed from the agarose using a gel removal centrifugal apparatus obtained from Amicon (Beverly, MA), then stored at -20°C until use.

4.2.2 Radioactive Labeling of Biotinylated DNA

The PCR amplified gene was radiolabeled at the 5' terminus with [γ-32P]ATP (DuPont NEN, Wilmington, DE). Two pmoles of DNA was suspended in 29 μL of ddH2O and 5 μL of 10X kinase buffer. To the DNA-buffer solution, 15 μL of [γ-32P]ATP (3000Ci/mmol, 10mCi/mL) and 1 μL of the T4 polynucleotide kinase labeling enzyme (Gibco-BRL) were then added. The reaction mixture was incubated
at 37°C for 30 minutes. The action of the kinase enzyme was halted by the addition of 2 μL of 0.5M EDTA and heating the mixture to 65°C for 2 minutes.

Figure 4.3
Agarose purification of PCR products generated from the amplification of the actin gene. 0.8% agarose gel and 90 V potential was used for separation. 1 μM EtBr was added to the gel matrix to allow visualization under UV light. The ΦX174 Hae III ladder serves as a marker by which the size of the PCR product can be judged. The modified protocol attempted to use a 2X increase in gene starting material concentration. This negatively effected DNA amplification efficiency. It is known that enzymatic activity can show a strong sensitivity to concentration. The root cause of this enzymatic sensitivity was not studied in this work.

Unincorporated ATP was removed by performing a selective precipitation of labeled DNA >20 bp in length. Precipitation was accomplished by the addition of 25 μL of 7.5M ammonium acetate and 150 μL of cold (4°C) 100% ethanol. The solution was allowed to sit 30 minutes at -20°C and then centrifuged at 10 krpm for 20 minutes at -20°C to form a DNA pellet. The supernatant liquid was decanted after centrifugation was complete. Scintillation measurements of the 32P labeled DNA in the nano capillary reactor core were performed by submerging the sealed reactor into a
scintillation liquid (3 mM p-terphenyl in toluene). Measurements were performed on a LS6000IC series scintillation counter (Beckman Instruments, Fullerton, CA).

4.2.3 Immobilization of DNA Template to Reactor Wall

A 10 cm length fused-silica capillary tubing (20 μm, 50 μm or 100 μm i.d., 360 μm o.d.) was cut and rinsed successively with 1N NaOH, water, 1N HCl respectively for 10 minutes each using vacuum suction. The capillary was purged with air (10 minutes) and oven baked at 80°C for 10 minutes to remove any residual water. The capillary was then filled with a 2% solution of 3-aminopropyltriethoxysilane (APTS) in acetone, purchased from Sigma Chemical (St. Louis, MO), for 30 minutes and air purged for 5 minutes and finally incubated for 24 hours at 45°C. After incubation, the tube was filled with a bicarbonate solution (50 mM, pH 8.3) containing 5.0 mg/mL of NHS-LC-biotin (Sigma) for 4 hours at room temperature. Following biotinylation, the reactor was exposed to a 4.0 mg/mL solution of streptavidin (Sigma) in 50 mM sodium phosphate buffer (pH 7.3). The streptavidin solution was allowed to incubate for 24 hours at 4°C followed by the removal of any free streptavidin by rinsing the capillary tube with distilled water, then filling with a 1 μM solution of biotinylated DNA and incubating the tube for 24 hours at 4°C. Excess template was removed by rinsing with distilled water. The column was capped with quartz caps and stored at 4°C until use. This method of bimolecular immobilization was initially described by W. Kuhr.28 A diagram of the streptavidin-biotin-streptavidin anchoring system is shown in Figure 4.4.
Figure 4.4
Schematic of the biotin-streptavidin-biotin linkage used to immobilize template DNA to the interior wall of the fused silica reactor core.
Figure 4.5
The design/photograph of the thermocycler used to perform polymerase reaction is shown. Thermocycler features insulated walls, heating provided by coil and cooling provided by Peltier plate and an internal fan to circulate air (not visible). This system is capable of performing 24 simultaneous reactions.
4.2.4 Air Thermocycler

The nano-reactor was housed in a double-walled insulated stainless steel box (see Figure 4.5). A peltier plate/heat sink was used to cool the temperature of the inner cavity of the steel box while heat was generated via resistive heating of an element. Fans were used to circulate air inside the thermocycler in order to provide uniform temperature. The temperature was monitored using a thermocouple and active feedback was used to control the temperature to within ± 1°C. On-line direct injection of the reaction products were performed by flanking the reactor on one side with a 20 cm length of fused-silica tubing and the other by the CGE separation column. The linking and gel-filled capillaries were attached to the reactor through the thermocycler housing using low dead volume (LDV) glass capillary connectors (MicroQuartz, Phoenix, AZ) shown in Figure 4.6.

![Diagram](image)

**Figure 4.6**
Low-dead volume capillary connector used to link nano-reactor vessel inside thermocycler with CGE separation column. Inter-capillary volume was calculated to be approximately 15 nLs.

The flanking capillary was coated with polyacrylamide in order to essentially eliminate electroosmotic flow. This allowed the reagents to be delivered into the reaction vessel from the cathodic end of the system by using electropumping. Polymerase

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
reactions were then performed on the immobilized template present in the reactor. Attempts were also made to directly inject reaction products onto a CGE column after it had been removed from the thermocycler apparatus. In this case the reactor was attached directly to the CGE column using a zero dead volume connector purchased from MicroQuartz and shown in Figure 4.7. This allowed the investigation of any influences due to the LDV capillary connector.

![Diagram](image)

**Figure 4.7**
The interface used to connect the reactor core to the CGE column for fractionation. Provides a zero dead volume connection between reactor system and separation system.

The sequencing fragments were then denatured using heat and directly injected onto the gel column for fractionation by applying a negative potential at the injection end of the system.

**4.2.5 LIF/CGE System**

DNA fragments were separated and detected using a custom built laser-induced fluorescence capillary gel electrophoresis apparatus (LIF-CGE) (see Figure 4.8). The matrix used during separation was a 3%T/3%C covalently anchored polyacrylamide gel containing 8M urea as a denaturant. The field strength (E) was 200 V/cm during electrophoretic separation and the running buffer consisted of 1X TBE solution (pH= 8.6).
Figure 4.8
Schematic of Near-IR LIF CGE system used to analyze fluorescently labeled reaction products. M= mirror, L= lens, BL= beam block, C= separation capillary, SF= spatial filter (slit, 0.6 mm), BF= spectral filters (bandpass, longpass), MO = microscope objective, SPAD= single photon avalanche diode.

The LIF detector was designed to operate in the near-IR for excitation and detection of the extension fragments which were labeled with a near-IR dye at the 5' end of the sequencing primer. Basically, the system consisted of a GaAlAs diode laser producing up to 20 mW of laser light at 780 nm which was focused onto the capillary using a singlet lens to a spot size of approximately 10 μ (1/e^2). The fluorescence was collected with a 60X microscope objective (NA=0.85) and imaged onto a spatial filter with a slit width set to 0.2 mm. This produced a sampling volume in the gel column of ~1.05 pL. After spatial filtering, the emission was spectrally filtered with an 8-cavity bandpass filter (CWL = 820 nm; HBW = 30 nm; Omega Optical, Brattleboro, VT) and a longpass glass filter (Corning, Ithaca, NY) which allowed light with λ > 820 nm to effectively pass. The emission was then focused onto the face of the photodetector using a 20X microscope objective. The
photodetector was a single photon avalanche diode (SPAD) that was passively quenched and interfaced to a PC which contained a counting board (CT101, Cyber Research, Branford, CT) with the data acquisition software written using LabTech Notebook® (Cyber Research).

4.2.6 Solid Phase Nano-Polymerase Reactions

Denaturation of the PCR amplified dsDNA was accomplished by heating the reactor to 95°C for 2 minutes, immediately cooling the reactor to 4°C for 2 minutes, and then rinsing any free DNA with 1X TBE buffer using pressure or electropumping flow. The extension of the immobilized gene was performed directly inside the reactor. The reactor was filled with a 1 x 10^{-6} M solution of IRD41 labeled M13 primer (LiCor, Lincoln, NE) and the temperature was raised to 95°C and then immediately cooled to 4°C for 20 minutes to allow annealing of dye-primer to the immobilized template. The system was then filled with a mixture containing Vent® polymerase (2 units), the four deoxynucleotides (1.4 μM each) and a single dideoxynucleotide (ddATP, 8 μM). The Vent® thermophilic enzyme was chosen for this application because it allowed all materials necessary for primer extension to be added simultaneously. This characteristic is very important since the nano-reactor is designed to perform “batch” type reactions. Enzymes, such as T7, that require the periodic addition of reactants are not well suited for use in this system since the inflow of any additional solution will cause the outflow of materials already present in the reactor. The temperature of the reactor was raised to 72°C for 30 minutes to allow extension of the annealed primer to form the complementary strands of the template.
Figure 4.9
Evaluation of different enzymes used in the nano-reactor. Lanes 1-2 used Vent® with varying [ddATP] (standard and 1:10 dilution), Lanes 3-4 used a T7 enzyme and varying [ddATP] (standard and 1:10 dilution), Lanes 5-6 used Bst® enzyme with varying [ddATP] (standard and 1:10 dilution), Lanes 7-8 used Bst® system in a standard microtube vessel (no nano-reactor) with varying [ddATP] (standard and 1:10 dilution). Standard reactant concentrations were used for each reaction protocol. All reaction were performed in capillary tubes each with a 7.85 μL volume. 32P labeled dATP was used for detection. Biotinylated DNA templates were immobilized using the procedure previously described in this chapter. Samples were denatured at 95°C and collected for SGE separation and autoradiographic detection.
Finally, the reactor temperature was raised to 95°C for 2 minutes in order to denature the template and rinsed using electropumping to directly inject the terminated DNA products onto a gel column for fractionation.

### 4.3 Results and Discussion

Figure 4.9 shows the result of performing DNA sequencing reactions in the 'large scale' nano-reactor using various DNA polymerase enzymes. Sequencing using the Bst® enzyme in a standard microtube protocol was used as a control in lane 7. Lanes 1 and 2 exhibit the most prominent sequencing fragments. These lanes were filled with the products generated using the Vent® DNA polymerase. Lanes 3 and 4 show the products of T7 polymerase reaction and 6 shows the product of Bst® polymerase reactions performed inside a nano-reactor. The results of the T7 reaction, which showed no banding on X-ray film, were not encouraging. This is due to a low terminated fragment population. Poor fragment production may be caused by the modifications of the standard T7 protocol which are necessary to adapt this system for use in the nanoreaction vessel. Standard T7 protocols call for the periodic addition of solutions to aid DNA polymerization. This is problematic in the nano-reactor format since the addition of material will lead to the removal of solutions already present in the reactor. The Bst® system suffered from similar reactant loading problems. The Bst® also produced faint banding patterns compared to other lanes on the gel, albeit to a lesser extent. The Vent® system allows for the addition of all necessary materials in a single step. This format is more easily adapted to meet the requirements of the nano-reactor format and the Vent® was the most efficient in producing DNA fragments in the reactor system of the enzymes tested.

The efficiency of template immobilization onto the capillary wall was then studied. The degree of coverage was determined by incorporating a $^{32}$P label into the template which contained a biotin linkage at the 5' terminus. The amount of
radioactivity radiating from the reactor, after rinsing, was then measured by submerging the entire reactor in a vial containing a scintillation cocktail, which allowed the β emission from the $^{32}\text{P}$ labeled DNA template to excite the scintillant, and monitoring the system using a scintillation counter. The energy of the $^{32}\text{P}$ decay (1.71 meV) is great enough to penetrate the 265 μ thick wall of the fused-silica tube, thus it is possible to obtain accurate counting rates using this procedure. Figure 4.10 diagrams the system used for scintillation measurements.

![Diagram of sampling system used to perform scintillation measurements. Scintillation cocktail is 3mM p-terphenyl in toluene. Capillary length is 4 cm.]

Since the $^{32}\text{P}$ labeled DNA is confined by the dimensions of the capillary tube and is not uniformly distributed in the scintillant, experiments were performed to determine any orientational dependence present in the scintillation measurements. An assessment of orientational bias was performed by rotating the sample vial 1/4 turn between sequential measurements. The results of this experiment are shown in Figure 4.11. The small variance obtained indicated that sample orientation had little effect on scintillation intensity.
Figure 4.11
The effect of sample orientation on scintillation intensity. Samples were rotated 1/4 turn between measurements. Three full rotations were used to determine the mean and standard deviation.

An additional benefit of this system is that it allows for the use of a very simple scintillation cocktail (3 mM p-terphenyl in toluene). Many scintillation cocktails contain a number of drying agents intended to remove water from the solution and thus prevent any aqueous quenching of the scintillant. The addition of drying ingredients to prevent aqueous quenching of scintillant in this system is unnecessary because the radioactive DNA (and all associated buffers) are sealed inside the reactor when it is submerged in the cocktail.

Experiments were performed to determine the effect of removing any buffer solutions from the reaction prior to scintillation measurement. The concern in this instance was that the presence of buffer solution may block some of the β emission from the ³²P and reduce the amount of activity measured from the reactor tube. To
investigate this, measurements were performed on multiple capillary reactors with and without buffer present in the reactor. Quantitative observation indicated that no significant differences were found in the intensity measured from reactors which were filled with 1X TBE and those which had been purged with air.

The theoretical scintillation intensity for complete monolayer coverage was determined by calculating the surface area available for immobilization inside the reactor and assuming that the degree of coverage is restricted by the area occupied by the streptavidin protein (2000 Å²) used to complex the biotinylated DNA template. The coupling ratio of the biotinylated template to the streptavidin anchor was assumed to be 1:1 based on the steric interactions which would result when multiple 1 kbp templates were bound to a single immobilization site. Based on these assumptions the efficiency of template immobilization (or percent coverage) of the interior reactor wall was determined to be 77% (± 20%). Based on the physical dimensions of this reactor (100 μm x 5 cm), the average amount of DNA immobilized in this system is ~ 1x10⁻¹² moles (1.3 x 10⁻¹² x 0.77) yielding an effective DNA concentration of 5 x 10⁻⁶ M in the 390 nL volume. Table 4.1 reports the theoretical considerations used to determine the percent coverage of the template on the reactor wall.

### Table 4.1
Number of binding sites on interior reactor wall. *Theoretical signal intensity was adjusted to account for half-life (t₁/₂ = 14.7 days) deterioration for 72.3 days. The reactor in this case was 4.0 cm long with an i.d of 100 μm. (Calculations shown in Appendix A)*

<table>
<thead>
<tr>
<th>Reactor Surface Area</th>
<th>1.57 x 10⁻⁸ m²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume</td>
<td>390 nL</td>
</tr>
<tr>
<td>Streptavidin Surface Area</td>
<td>2000 Å²</td>
</tr>
<tr>
<td>Binding Sites</td>
<td>7.85 x 10¹¹</td>
</tr>
<tr>
<td>³²P-labeled Template (³²P labeling efficiency optimal = 60 %)</td>
<td>3.93 x 10¹¹ molecules (0.65 pmoles)</td>
</tr>
<tr>
<td>Specific Activity ⁴²P</td>
<td>4.99 x 10⁻¹⁸ Ci/molecule</td>
</tr>
<tr>
<td>Conversion from Ci to dpm</td>
<td>2.22 x 10¹² dpm/Ci</td>
</tr>
<tr>
<td>Theoretical Signal*</td>
<td>126,000 cpm</td>
</tr>
<tr>
<td>Actual Signal</td>
<td>90,115 cpm</td>
</tr>
<tr>
<td>Coverage</td>
<td>72%</td>
</tr>
</tbody>
</table>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
One can predict the loss rate of the immobilized species under constant flow conditions by considering the binding coefficient, the flow rate and the amount of surface bound species. Once all unbound species have been removed by buffer flow, the equilibrium of the system must be reestablished. The time required to flush the system can be calculated using Equation 4.1:

\[ t_{\text{flush}} = \frac{V_{\text{col}}}{f} \]  

(4.1)

where \( t_{\text{flush}} \) is the time needed to flow a single column volume through the system; \( V_{\text{col}} \) is the system volume and \( f \) is the volumetric flow rate. The concentration of material in the system can be calculated using Equation 4.2:

\[ [BA] = \frac{S_{\text{surface}} \text{(moles)}}{V_{\text{col}}} \]  

(4.2)

where \( S_{\text{surface}} \) is the amount of surface-bound streptavidin. Using the equilibrium expression and substituting for \([BA]\), it is trivial to calculate the amount of free streptavidin \( S_{\text{free}} \). Complete calculations are shown in Appendix B. Two assumptions are made when calculating the theoretical loss rate; (1) initial surface coverage by streptavidin is 100% and; (2) equilibrium between the bound and unbound streptavidin is established in a time less than that needed to flush one column volume through the system. The results depicted in Figure 4.12 show the effects of long term rinsing on the immobilized DNA for extended time periods using electropumping (11.3 nL/sec) of 1X TBE buffer solution (pH 8.6) through the reactor.
Stability of the template anchor to the capillary wall under electropumping was favorable, requiring > 150 hours of continuous rinsing to reduce initial coverage by 50%. However, depletion of the immobilized DNA from the reactor wall is slightly accelerated over the rate predicted using the dissociation constant for biotin/streptavidin at 250 hours of continuous rinsing. The surface coverage was reduced to 38% whereas based on the known binding constant for the biotin/streptavidin couple and the volumetric flow rate one would predict a surface coverage of 84%. In order to determine if the accelerated loss was due to the
influence of the applied potential across the reactor during electropumping, experiments using pressure driven flow were undertaken. The loss rate using gravity flow is shown in Figure 4.13. These results also showed an increased depletion rate versus the theoretical prediction, but had a smaller deviation from the predicted loss rate. At this volumetric flow rate, the predicted surface coverage would be 38% after 250 hours of rinsing and the actual value was 16%.

![Graph showing surface coverage over time for theoretical and pressure driven flow.](image)

**Figure 4.13**
The immobilized template loss encountered when using pressure driven flow of 1X TBE to flush the capillary reaction vessel.

The higher loss encountered in both experiments may be partially explained by the alkaline nature of the running buffer (1X TBE, pH=8.6) which could expedite
denaturation of the streptavidin protein or loss of biotin from the reactor wall. Accelerated loss of the streptavidin protein or the biotin anchor would allow greater amounts of DNA template to be washed away in both cases. The accelerated loss may also be due to hydrolysis of the silane anchor on the biotin molecule. The potential applied across the reactor during electropumping may also have a negative effect on the stability of the biotin/streptavidin anchor which would explain the larger deviation from theoretical predictions seen when EO rinsing is used to flush the reactor system.

Figure 4.14 shows the efficiency of reactivation of the nano-reactor by reimmobilization of biotinylated $^{32}$P-DNA template. After extensive gravity driven rinsing, streptavidin was added to the nano-reactor followed by biotin-DNA. After incubation at 4°C, the coverage of reimmobilized DNA was found to be restored. The percent coverage calculated after reimmobilization falls within the statistical deviation seen during initial immobilization (± 20%). Addition of biotinylated $^{32}$P labeled DNA without first rejuvenating the streptavidin layer resulted in negligible increases in reactor wall surface coverage. These results indicate that the covalent linkage of biotin to the wall remains essentially intact under long term use and that simple steps can be taken to reactivate the reactor without requiring tube replacement. This result is encouraging because it implies that a series of reactions may be performed in a single reactor without replacement by simply restoring the streptavidin linkage and biotinylated DNA. Regeneration of the streptavidin coating after rinsing could be accomplished by binding of the streptavidin protein to unhydrolyzed biotin since there is an excess of biotin molecules on the reactor surface relative to streptavidin linkages based on steric considerations.
To demonstrate the utility of this system for sequencing DNA templates which are commonly double stranded, we next focused on the efficacy with which the dsDNA could be denatured inside the reactor core to provide ssDNA suitable for polymerase reactions. Figure 4.15 shows the result of raising the reactor temperature...
to 95°C for 5 minutes then immediately cooling the reactor to 5°C for 3 minutes, followed by pressure rinsing (33 nL/sec). An average 36% loss of scintillation intensity was seen from the immobilized, γ³²P-labeled template after the initial denaturing step.

![Figure 4.15](image)

**Figure 4.15**
Effect of heating reactor to 95°C followed by rapid cooling to 5 °C and pressure rinsing with 1X TBE. (DNA denaturing conditions)

A second denaturing step averaged a total 43% loss in initial intensity. Additional denaturing steps lead to only minimal decreases in scintillation intensity. Since both strands are labeled with ³²P, the loss of one strand would be expected to yield roughly a 50% reduction in scintillation intensity. This result shows that an immobilized dsDNA template can be denatured by rapid heating and cooling with no adverse effect on the biotin/strepavidin anchor. No further loss in scintillation signal after repeated denaturing steps demonstrate that the biotin/strepavidin anchor can withstand the
temperature changes associated with denaturation. I also attempted to denature the template using 0.5N NaOH solution in the absence of heat, but observed that after three rinses an 81% reduction in scintillation signal occurred. This would indicate a loss of immobilized template from the reactor wall most likely caused by attack of hydroxide ions at the Si-O-Si bond of the MAPS anchor. Rinsing procedures using high pH to denature dsDNA are deemed unsuitable for use in this reactor system due to a loss of the surface bound biotin species.

In order to demonstrate the efficacy of performing DNA polymerase reactions using the nano-reactor system, a section of an immobilized gene was extended inside a 100 cm long, 100 μm i.d. reactor (Volume = 7.8 μL). It is important to establish the feasibility of performing polymerase reactions in this capillary system since it is known that enzyme-wall interactions can effect the activity of certain enzymes during reactions.31 The large scale of this reactor (7.8 μL) was necessary to provide manageable product volume, since off-line detection was used for this experiment and manual manipulation of solutions was required. Temperature control in this system was accomplished using temperature regulated water baths linked through a tubing network which sheathed the capillary reactor (see Figure 4.16). A fluorescien labeled M13 universal primer was used to generate complementary fragments of the gene. The products of this reaction were collected and analyzed using an off-line LIF-based SGE analyzer. The resulting electropherogram is shown in Figure 4.17.
Figure 4.16
Diagram and actual picture of nano-reactor system using liquid sheathed temperature control provided by water baths.
Figure 4.17
Electropherogram of DNA products from sequencing reactions performed using immobilized DNA inside a capillary reactor. The Vent® enzyme system was used with a dNTP/ddNTP ratio of 1. The products were collected and analyzed off-line using a commercial SGE-LIF sequencer. A 6%T / 5%C polyacrylamide gel was run at 50 V/cm using 1X TBE running buffer for separation. Template was extended approximately 150 bases from the priming site.

The results shown in this electropherogram demonstrate the feasibility of performing LIF-based DNA sequencing reactions using immobilized templates and a polymerase enzyme. The dNTP/ddNTP ratios were not optimized for this reaction which may have caused premature termination of complementary fragments resulting in declining peak intensities for the longer fragments. Adjusting the dNTP/ddNTP ratio should result in the production of a more uniform distribution of fragment lengths as shown previously (Chapter 3). Optimization of reaction parameters was not performed at this time because the focus of this research was the use of dye-labeled ddNTP during DNA sequencing reaction. Optimization of this system would provide little insight.
into the behavior of a dye-modified ddNTP system since this system used FITC-labeled primers for detection. Another possible reason for the short extension products could be an adverse effect of the reactor wall on the polymerase enzyme if negatively charged Si-O molecules at the capillary surface are not uniformly deactivated through bonding. Enzyme-wall interactions could detrimentally effect the activity of the DNA polymerase creating non-ideal reaction conditions which result in poor sequencing results. A number of problems due to the use of water temperature control were also encountered. The most prominent problem was the long time period required to change temperatures in the reactor by heating and cooling a single water bath; thus, in order to achieve rapid temperature changes, multiple baths at various temperatures were employed. This system was quite cumbersome and difficult to manipulate. For these reasons, the development of an air-based thermocycler system was initiated. The use of air thermocyclers for sequencing has been reported in the literature. These devices can provide temperatures in the ranges needed for sequencing (4°C to 95°C). In addition they have been reported to provide rapid temperature transitions which are critical for high enzymatic efficiency and activity during DNA sequencing.

In order to demonstrate the feasibility of performing DNA polymerase reactions using the nano-volume system, a section of DNA was extended to produce T-terminated fragments. The reactor dimensions were 10 cm long and 50 μm i.d. (Volume = 196 nL). The calculated DNA fragment concentration (assuming 77% coverage immobilized template, 100% chain extension and complete denaturing of the fragments) is 6.6 x 10^6 M. This projected concentration can only be used as a crude guide to predict the actual concentration of the sequencing fragments in the reactor. Complete extension will only be achieved under conditions optimized to promote DNA polymerase activity. Many DNA polymerase enzymes (Vent® included) demonstrate activity based upon template, primer and nucleotide concentration among
other factors such as temperature and pH. However, one factor which compensates for the reduced number of fragments is the small volume of the reactor, which will cause only minimal dilution of the reaction products and should allow the injection of sufficient amounts of material to permit detection. The products of the sequencing reaction were directly injected onto an LIF-CGE analysis system.

The efficiency of performing direct injections onto CGE columns using functionalized “linking” capillaries attached to the nano-reactor was investigated. Linking capillaries will be necessary to serve as conduits through which reactants and products will be transferred in an integrated DNA sequencing system. Thus, understanding the effect of linking capillaries on efficiency during analysis is important. A near-IR labeled primer was injected through a linking capillary, the DNA coated nano-reactor and directly onto a CGE column. Measurement of the electroosmotic flow (opposite the direction of the migrating DNA fragments) yielded a $\mu_{eo}$ value of $-4 \times 10^{-4} \text{cm}^2/\text{Vs}$ in free solution which was determined by the migration of a neutral marker (riboflavin) in a bare fused-silica capillary using 1X TBE running buffer and a field strength of 250 V/cm. This flow is detrimental to rapid and efficient injection of products through the nano-reactor onto the separation column since it is opposite to the migratory direction of negatively charged species such as DNA fragments (see Figure 4.18 for experimental setup). Experiments were performed using silanized capillaries coated with biotin which were connected to the reaction capillary. The biotin coating was used in an attempt to abate electroosmotic flow in the flanking capillaries. The presence of electroosmotic flow would have a negative effect during DNA sequencing causing poor efficiency separation. Injection of near-IR primers onto the system using a biotin coated flanking capillary yielded no measurable peak from the dye-labeled primer.
Negatively charged oligonucleotides are separated inside the gel capillary based on their size.

The electrode and capillary are transferred into a running buffer after the electrokinetic injection of the oligonucleotide.

Figure 4.18
Diagram of experimental setup used to probe the effectiveness of injecting DNA through the nano-reactor apparatus directly onto a CGE column for separation and detection. LPA coated linking capillary was 75 μ i.d. with a 2% LPA coating. A linking capillary coated with only biotin did not provide sufficient EO flow reduction to allow injection of DNA.

These experiments suggest that biotinylation alone does not suppress the EO flow to a great enough extent to allow the capillaries to be used as conduits for injection of DNA onto a CGE column. The connecting capillaries were then coated with a 2% linear polyacrylamide gel. The presence of a linear polyacrylamide monolayer on the capillary has been reported to drastically reduce the magnitude of EO flow; thus, allowing injection of the negatively charged DNA fragments onto the gel column.32
Injection of a near-IR labeled primer onto a gel column through the reactor system using an LPA coated linking capillary and the biotinylated DNA coated reactor capillary is shown in Figure 4.19.

![Electropherogram showing the efficacy of injecting labeled DNA (A) directly onto a capillary gel column from a microtube and (B) through an LPA coated capillary (20 cm) linked to the miniature reaction vessel (10 cm). Low dead volume connectors were used to connect linking capillary, reactor and CGE column. The capillary gel column was filled with 3% T, 3% C polyacrylamide gel with a total length of 60 cm and an effective length of 50 cm. Injection time was 30 seconds. Plate numbers calculated using full width at half height of the first peak.](image)

Notice the efficiency obtained when injecting a sample of fluorescently labeled primer onto a gel column through the LPA coated linking capillaries from a microtube,
~23,500 plates (47,000 plates/meter) versus the efficiency obtained from the direct injection of labeled primer onto a gel column from the nano-reactor 23,300 plates (46,000 plates/meter). Only a 1% efficiency loss is encountered when using the linear polyacrylamide coated ‘linking’ capillary and low dead volume connectors to inject dye primer onto a gel column. The air thermocycler system’s design requirements, such as the need for double walled insulation for thermal isolation, did not allow the integration of true “zero” dead volume unions between the thermocycler and CGE. The “low” dead volume connectors used introduced a dilution volume of ~ 15 nL (see Figure 4.5). An additional source of broadening could be provided by any heating that may occur at the low-dead volume connector position. The larger diameter at the center of the capillary connector will have a detrimental effect on the efficiency of Joule heat dissipation by reducing the surface to volume ratio at that point. A large surface to volume ratio is desired in capillary systems, because it provides a more efficient removal of heat generated by current flow in the system. However, it is clearly evident that the efficiency of DNA injection using these capillary interlock devices only minimally reduced when compared to direct injection. Another notable fact is that the mobility of the dye-primer is greater than any residual EO flow that may remain in this system after the fused silica linking capillaries have been coated with linear polyacrylamide. Dye-primer injection can overcome the convective influences of the integrated low-dead volume connectors to allow direct injection onto a CGE separation system without a significant effect on efficiency.

Once convinced of the feasibility of direct, on-line injection of DNA and capillary reactor based sequencing reactions, I attempted to link the two systems in order to create a completely integrated DNA analysis system. The products of sequencing reactions performed in an air thermocycler, using a 196 nL reactor, were directly injected onto a LIF-CGE analysis system for 2 minutes using a low dead
volume capillary connector and a linear polyacrylamide linking capillary. After injection from the nano-reactor the CGE column was placed in buffer solution.

Several problems were encountered in these experiments. The first issue was the difficulty in maintaining fluid contact throughout the system when loading the capillary reactor with sequencing reagents. After loading, the reactor was disconnected from the CGE column which allowed the introduction of air bubbles. The presence of air bubbles in the system will impede fractionation of reaction products on the separation column. Loading reagents into the reactor prior to sequencing was performed manually using a gas-tight syringe and required a tremendous level of dexterity in order to accomplish. This operation was tedious and highly irreproducible. Contamination, spillage and the presence of bubbles in the reactor were major problems. Another issue was the inadequate temperature control provided by the air thermocycler apparatus. This system, while providing adequate temperature accuracy, did not provide the rapid temperature changes required to produce high efficiency sequencing products. Commercial cycle sequencers can produce temperature transitions from 100°C to 4°C in 10-15 seconds. The air thermocycler employed here required in excess of 300 seconds to change from 100°C to 10°C. Slow temperature transitions translate into long cycle times. Long cycle times reduced the number of cycles that could be performed before the activity of the polymerase enzyme degraded, leading to low efficiency sequencing fragment generation. Since the reaction temperature could not be optimized, DNA polymerase activity could also be negatively influenced. After multiple attempts this experiment garnered only moderated success. An electropherogram from the most successful attempt is shown in Figure 4.20. The nonuniformity of the peak-heights seen in Figure 4.20 are partially attributable to electrokinetic injection bias, which favors the injection of shorter fragments, and the unoptimized reaction conditions. Since the injection of the total contents of the reactor (196 nL) would remove the effect of any
preferential injection bias, this result implies that the entire amount of sequencing products were not injected onto the CGE column. Another noticeable attribute of this electropherogram is the lack of the large peak due to unextended primers. The primer peak is usually present due to the large excess of primer present during DNA sequencing reactions. In this experiment, the excess primers have been removed by rinsing prior to chain extension. This is an advantage afforded by immobilizing the template DNA which allows removal of excess primers, dNTP, enzymes before fractionation. The absence of the large unextended primer peak will facilitate sequencing near the priming site, a region that is usually obscured due to the intensity of the primer peak.

4.4 Summary

The use of a solid-phase nanoscale reactor-based DNA polymerase reaction system with on-line detection will result in significant advancements in sequencing. The ability to interchange the DNA template easily and efficiently will facilitate the use of this reactor system in a seemingly unlimited number of applications. I have established that the activity of the Vent® enzyme in the nanoreaction vessel is adequate to perform DNA polymerase reactions. It is expected that a number of other DNA polymerase enzymes will demonstrate a compatibility toward use in this format, which is of integral importance if this system is to be useful in numerous highly varied approaches toward DNA analyses. Favorable temperature stability coupled with a resistance to template removal under rinsing conditions indicates that this system may perform well as a reusable cycle sequencing tool which will allow the user to perform multiple sequencing reactions with minimal sample preparation and purification procedures needed. Primer-walking protocols, which use information obtained from an initial extension cycle to design primers for subsequent extensions, should prove easily adaptable for use in a nano-reactor format as well.
A 196 nL reaction volume is approximately 1/100th the size of those encountered in common sequencing reaction protocols. Off-line manipulation of such minute volumes is impractical. However, using a custom designed reactor/CGE interface,
we have successfully manipulated this minute reaction volume generated in the reactor core. The direct analysis of a nanoliter reaction volume illustrates the feasibility of minimizing reaction protocols and represents a possibility for tremendous savings in genetic analysis. Costs, in large-scale sequencing projects, when projected using a miniaturized reactor system show staggering reductions in reagent cost when compared to macro-scale techniques (see Table 4.2).

Table 4.2
Cost projections for genomic sequencing using standard polymerase reaction protocols vs. nanoscale protocols. *Assumes the use of capillary gel separation in standard case and microchip based separation in nanoscale system.

<table>
<thead>
<tr>
<th>Table 4.2</th>
<th>Standard</th>
<th>Nanoscale</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Vol./rxn</td>
<td>Cost</td>
</tr>
<tr>
<td>Enzyme</td>
<td>1,000 nL</td>
<td>$12,375</td>
</tr>
<tr>
<td>Dye-Primer</td>
<td>1,000 nL</td>
<td>$8,750</td>
</tr>
<tr>
<td>Column*</td>
<td>50 cm</td>
<td>$750</td>
</tr>
<tr>
<td>Gel Soln.</td>
<td>1,000 μL</td>
<td>$2,500</td>
</tr>
<tr>
<td>Total/day</td>
<td></td>
<td>$24,375</td>
</tr>
<tr>
<td>Human Genome</td>
<td></td>
<td>$73x10⁶</td>
</tr>
</tbody>
</table>

Further reduction in the volume of material produced in this system should be easily accomplished by simply reducing the diameter of the reactor core. The terminal focus of this research is further reduction of reactor size and direct integration into a microchip based separation and detection system to produce a completely autonomous DNA analysis unit. DNA sequencing applications will be integrated into this system by incorporating multi-fluor or temporal discrimination detection systems which have been described elsewhere.33-36

4.5 References


Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


The Effects Of Capillary Column Diameter On Efficiency For Electrophoretic Separation of ssDNA molecules

5.1 Introduction

Gel electrophoresis is currently one of the most widely used methods for the analysis of biological macromolecules, such as proteins and nucleic acids. The widespread use of gel electrophoresis is plainly justified by the unparalleled efficiencies obtained during these types of separations. Gel electrophoretic techniques also offer a variety of detection options (e.g. autoradiography, densitometry, absorbance and LIF). Low sample requirements, which range from μL to nL, are also an attractive attribute of gel electrophoretic techniques. The gel electrophoresis format was first reported in the literature in the 1950s; however, the description of the mechanism by which gels separate biomolecules, such as DNA, continues to spark some of the most vigorous debate in the scientific community. To date, numerous theories and theoretical revisions have been reported in the literature, a number of which have been listed in Table 5.1.

### Table 5.1
List of migration models in gel electrophoresis. (Adapted from *Anal. Biochem.* 1995, 231, 1-12.)

<table>
<thead>
<tr>
<th>Model</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extended Ogston*</td>
<td>1970</td>
</tr>
<tr>
<td>Viscosity*</td>
<td>1977-1980</td>
</tr>
<tr>
<td>Reptation*</td>
<td>1982-1985</td>
</tr>
<tr>
<td>Bias Reptation*</td>
<td>1986-1992</td>
</tr>
<tr>
<td>&quot;Modified&quot; Reptation*</td>
<td>1986-1994</td>
</tr>
<tr>
<td>Deformable DNA*</td>
<td>1991</td>
</tr>
<tr>
<td>Gel-cell*</td>
<td>1991</td>
</tr>
<tr>
<td>Matrix Interaction*</td>
<td>1983</td>
</tr>
<tr>
<td>Transient Coupling*</td>
<td>1994</td>
</tr>
<tr>
<td>Door-Corridor*</td>
<td>1994</td>
</tr>
</tbody>
</table>

The two most commonly used models are the extended Ogston model, which provides a description of proteins and globular DNA in agarose matrices, and the reptation model, which describes the motion of ssDNA in small pore gels such as polyacrylamide. However, when attempting to evaluate a relationship between
capillary diameter and efficiency both of these models fall short because neither provides a term which relates the influence of capillary diameter on the mobility of electromigratory species such as DNA. The extended Ogston equation is shown in Equation 5.1 and the reptation model equation is shown in Equation 5.2.

\[
\log \mu = \log \mu_e - \pi \ell' (r + R)^2 T \quad (5.1)
\]

\[
\mu = \frac{q}{3 \zeta} \left( \frac{1}{N} + \frac{E^2}{3} \right) \quad (5.2)
\]

The importance of capillary column inner diameter on the efficiency (theoretical plates) obtained during separations using free solution electrophoresis (CZE) has been well documented in the literature. Analysts choose to increase or decrease the diameter of the column employed based upon the requirements of a particular analysis. When high efficiency and short analysis times are paramount, a capillary with the smallest i.d. available would be used. This would allow the greatest degree of Joule heat dissipation and minimal thermal gradients across the column when high electric fields are employed. The efficiency enhancement observed in smaller i.d. columns in CGE can be rationalized by applying theoretical considerations used to describe zone variance contributions in CZE. Variance due to a parabolic temperature profile within the separation capillary will show an \( R^6 \) dependence on capillary inner diameter\(^{18} \);

\[
\sigma_{\Delta T}^2 = \frac{R^6 E^6 \kappa_e^2 \Omega_T^2 \mu_{app} t}{1536 D k^2} \quad (5.3)
\]

where \( R \) is the interior radius of the capillary (m), \( E \) is the electric field (V/m), \( \kappa_e \) is the electrical conductivity of the buffer solution (\( \Omega^{-1} \ m^{-1} \)), \( \Omega_T \) is the temperature coefficient of electrophoretic mobility (\( K^1 \)), \( \mu_{app} \) is the apparent mobility of the species (\( m^2/\text{Vs} \)), \( D \) is the diffusion factor (\( m^2/\text{s} \)), and \( k \) is the thermal conductivity of the buffer solution. The migration time of analyte is dependent upon the applied field strength. This relationship is shown in Equation 5.4;
\[ \text{velocity (cm/s)} = \mu E_{\text{field}} \] (5.4)

where \( \mu \) is mobility (cm\(^2\)/sV) and \( E_{\text{field}} \) is the electric field strength (V/cm). Therefore, higher fields will yield shorter analysis times assuming all other factors are held constant. However, if the analyst is faced with samples which are not easily detectable or a detection scheme with limited sensitivity, it may be advisable to employ the use of a slightly larger i.d. capillary to allow the injection of a greater analyte mass and would provide a longer path length (in the case of absorbance detection) to facilitate detection according to Beer’s Law (Equation 5.5):

\[ A = ebc \] (5.5)

where \( A \) is absorbance, \( e \) is the extinction coefficient, \( c \) is the concentration and \( b \) is the path length.

The advantages of using a larger column to aid in analyte detection are easily demonstrated; however, the demonstration of small capillary efficiency enhancements, in this work, have been impeded by the difficulties associated with producing uniform gel-filled capillary columns with i.d.s \( \leq 50 \) \( \mu \text{m} \).

This work attempts to demonstrate the influence of reducing the column diameter on efficiency in capillary gel systems and relate this influence to a gel electrophoresis model. Attempts to use gel filled columns with diameters of 20, 50, 75 and 100 \( \mu \text{m} \) (as shown in Figure 5.1) to separate oligonucleotide ladders are discussed. In addition, an evaluation of several gel column fabrication protocols will be given.

5.2 Experimental Section

5.2.1 CGE Absorbance System

Absorbance measurements were performed using a Waters Quanta 4000\textsuperscript{®} CE system equipped with a 20-well autosampler. This system provided for both hydrodynamic and electrokinetic injection modes, however since these analyses
employed gel columns, only electrokinetic injection was employed. The light source used was a Hg lamp with a replaceable filter for wavelength selection.

![Image of fused-silica capillary columns with varying diameters.](image)

**Figure 5.1**
Image of fused-silica capillary columns with varying diameters.

Detection was performed at 254 nm to exploit the native absorptivity of oligonucleotides at this wavelength. The Quanta 4000® is capable of providing both negative and positive potential to the separation capillary up to a maximum of 25 kV.

### 5.2.2 Oligomer Standards

Oligonucleotide standards used in CGE experiments with absorbance detection were obtained from Sigma Chemicals (St. Louis, MO). These standards consisted of poly-dT nucleotides ranging in size from 20-25 and 40-60 bases.
5.2.3 Capillary Gel Column Preparation

Chemical Initiation Method. Before the column was filled with the polyacrylamide gel solution, a small portion of the polyimide coating was removed by use of a heated coil to provide an optical window. The interior wall of the capillary was pretreated with 1M NaOH (10 min.), double distilled water (10 min.), 1M HCl (10 min.) and finally, double distilled water (10 min.). A mixture of [3-(methacryloxy)propyl] trimethoxysilane (MAPS) (Aldrich Chemical Company) and methanol (50/50) was placed into the column and allowed to react overnight. MAPS serves as an anchoring point for the polyacrylamide matrix during polymerization and reduces gel extrusion under high electric fields. After MAPS treatment, the column was then evacuated using low pressure and a 2% non-crosslinked polyacrylamide (LPA) gel forming solution was injected into the capillary using aspiration and allowed to polymerize for at least 3 hours. The LPA solution bonds with the MAPS anchor and effectively coats the interior wall of the capillary. The LPA serves as a flexible linker between the crosslinked polyacrylamide matrix and the MAPS species.

![Diagram of MAPS anchor and polyacrylamide matrix](image)

**Figure 5.2**
Structure of MAPS anchor and polyacrylamide matrix used in CGE separations.

The LPA linker is necessary to compensate for the shrinkage which occurs when crosslinked gels polymerize. Failure to use a flexible coating when forming
immobilized gels results in the formation of tears and voids in the gel matrix as shown in Figure 5.3.

![Capillary Gel Column](image)

**Figure 5.3**
Voids formed in capillary gel matrices due to the shrinkage of the gel during polymerization. Voids formed due to shrinkage are usually evenly spaced.

A gel solution containing monomer (acrylamide), crosslinker (bis-acrylamide), 50 μL of a 10% by volume ammonium persulfate (APS) solution (radical propagator) and 5 μL of tetraethylmethylenediamine (TEMED) (radical initiator) is then introduced into the column. Filling the capillary column with the crosslinked gel solution must be done quickly since polymerization will begin upon the addition of APS and TEMED. If the column is filled too slowly, the viscosity of the solution will increase rapidly becoming too viscous to flow; however, too rapid an introduction will result in turbulent flow of the solution in the capillary causing the formation of bubbles in the column. It is possible to control the rate of gel polymerization by modifying the concentrations of TEMED and APS. Modification of the TEMED/APS concentration can effect the sieving properties of the gel matrix and have detrimental effects on separation efficiency.\(^\text{20}\) Temperature baths can also be used to regulate the rate of polymerization, but they can become cumbersome and difficult to use during gel fabrication. The crosslinked gel column is allowed to polymerize for 24 hours. The crosslinked gel solutions consisted of 5% total acrylamide, 5% bis-acrylamide, 1X TBE, and 8M urea (denaturant).
**UV Initiated Method.** Before the column was filled with the polyacrylamide gel solution, a small portion of the polyimide coating was removed to provide an optical window. The interior wall of the capillary was pretreated with 1M NaOH (10 min.), double distilled water (10 min.), 1M HCl (10 min.) and finally, double distilled water (10 min.). A mixture of MAPS and methanol (50/50) was placed into the column and allowed to react overnight. The column was then evacuated using low pressure and the gel forming solution injected into the capillary using aspiration. The gel solution (3%T / 3%C or 6%T / 5%C) contained 0.1 M EDTA, 8M urea and 0.008% (w/v) riboflavin (Sigma Chemicals, St. Louis, MO), which was used as the photo-polymerization initiator. Preparation of the polyacrylamide gel column was accomplished by vacuum injection of the gel forming solution into the pre-treated capillary and followed by photo-initiated polymerization. The ends of the capillary were capped and the column was submerged in an ice bath. The chilled column was then exposed to a high energy UV light source overnight. After polymerization, the column was pre-run for approximately 30 minutes at 5-6 kV prior to sample injection.

**Linear Polyacrylamide Gel-filled Columns.** In the LPA system all gels were fully polymerized for 24 hours prior to insertion into the capillary column. LPA gels consisted of 8% T, 0%C, 1XTBE, 3.5M urea and 30% formamide (co-denaturant). The addition of formamide reduces the viscosity of the polymerized solution without substantially reducing efficiency during electrophoresis. The viscosity of the LPA solution was measured using a Brookfield viscometer and was found to be 2100 cP, which is comparable with values previously reported for polyacrylamide gel matrices containing formamide and urea. Polymerization was accomplished using APS and TEMED as described in the chemical initiation method. This solution was inserted into capillaries using a custom-built high pressure gel injection vessel (shown in Figure 5.4).
Figure 5.4
High pressure vessel used to fill small diameter capillary columns with viscous polyacrylamide gel solution. System was rated to handle 1500 psi. Usual operating pressures ranged from 200-800 psi.

By applying a pressure of 600 psi that was provided by a nitrogen tank, this unit allowed the replacement of 2000 cP gels inside a 60 cm, 50 μm i.d. capillary in approximately 3 minutes.
5.3 Results and Discussion

Detection of samples was performed by absorbance in most cases, however, analysis employing capillaries with i.d. \( \leq 50 \mu m \) provided poor limits of detection (LOD). The high LODs obtained during analysis using 50 \( \mu m \) or smaller capillaries is a consequence of the short path length available for absorbance detection and the smaller loading volumes encountered. Mass loading (moles) in capillary electrophoresis is proportional to the square of the capillary radius given by Equation 5.6:

\[
Q = \frac{LV_{inj} A C t_{inj}}{t V_{run}} \left( \frac{1L}{1000 cm^3} \right) \quad (5.6)
\]

where \( L \) is capillary length (cm), \( V_{inj} \) is the injection voltage (V), \( A \) is the cross sectional area (cm\(^2\)), \( C \) is the sample concentration (M), \( t_{inj} \) is the time of injection (s), \( t \) is elution time (s) and \( V_{run} \) is the voltage applied during electrophoresis (V). As can be seen from Equation 5.3, \( Q \propto A \) with \( A = \pi r^2 \) where \( r \) is the capillary radius.

The efficiency obtained during oligonucleotide ladder separation using chemically polymerized CPA-filled capillary columns is shown in Figure 5.5. Figure 5.5 does show an increase in separation efficiency as column diameter is reduced. A notable result is that efficiency does not increase proportionally for different polymer concentrations. Capillary diameter reductions induce a small, almost linear increase in efficiency when a 3\%T/3\%C polyacrylamide matrix is employed as the separation medium (100 \( \mu m \)-250,000 plates, 75 \( \mu m \)-323,000 plates, 50 \( \mu m \)-410,000 plates). Capillary diameter reductions induce large, nonlinear increases in efficiency when a 5\%T/5\%C polyacrylamide matrix is used (100 \( \mu m \)-312,000 plates, 75 \( \mu m \)-565,000 plates, 50 \( \mu m \)-2,100,000 plates). From this data, there appears to be a nonlinear increase in efficiency as capillary diameter decreases for columns filled with a 5\%T/5\%C polyacrylamide matrix.
Figure 5.5
Efficiency obtained during electrophoretic separation of oligonucleotide standards using cross-linked gel matrices. Separations were performed using a Waters CE instrument equipped with an absorbance detector.

A recent evaluation of the influence of thermal gradients in CGE has stated that parabolic temperature profiles have a minimal contribution to band broadening in typical CGE applications. Indeed, a calculation of the heat generated at the center of the capillary during separations using Equation 5.7;

\[
Q = \frac{EI}{r^2} = \frac{250Vcm^{-1} \cdot I}{r^2}
\]

where Q is power, E is the field strength (V/cm), I is the current (A) and r is the column radius (cm), yields interesting results which are tabulated in Table 5.2.
Table 5.2
Evaluation of the temperature and injection volume influences during gel electrophoresis as a function of capillary diameter. *Injection time was increased for the 50 μ column to aid in detection.

<table>
<thead>
<tr>
<th>Column Diameter (μm)</th>
<th>50</th>
<th>75</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>I @ 250V/cm (μA)</td>
<td>4.7</td>
<td>8.8</td>
<td>12.3</td>
</tr>
<tr>
<td>Q (W/cm²)</td>
<td>59.8</td>
<td>49.8</td>
<td>37.6</td>
</tr>
<tr>
<td>ΔT center/wall</td>
<td>0.02K</td>
<td>0.04K</td>
<td>0.07K</td>
</tr>
<tr>
<td>Vol_inj</td>
<td>8 nL</td>
<td>7 nL</td>
<td>18 nL</td>
</tr>
<tr>
<td>Vol_inj/Vol_total x100</td>
<td>0.88 %</td>
<td>0.34 %</td>
<td>0.40 %</td>
</tr>
</tbody>
</table>

According to Knox,\textsuperscript{25} the temperature difference between the center and inner wall (θ) of a capillary is given by Equation 5.8;

\[
θ = \frac{Qr^2}{4k} \tag{5.8}
\]

where \( k \) is the thermal conductivity of the solution. The value of this term for 7M urea, 1X TBE, 4% acrylamide gel solution is estimated to be 4x10\(^3\) W cm K\(^{-1}\). The temperature difference between the capillary wall and the capillary center was calculated using Equation 5.8 and is shown in Table 5.2. The calculated temperature differences range from 0.02-0.07 °C. The band broadening contribution from such small temperature gradients in gel electrophoresis is insignificant\textsuperscript{25} It seems unlikely that such small temperature differences <0.1°C would account for the large efficiency increases shown in Figure 5.5.

In addition to thermal variance we must also consider the variance associated with the injection volume of the analyte species which also exhibits a 6\(^{th}\) order radius dependence (see Equation 5.9):

\[
σ_{t,χ}^2 = \frac{R^6 E^6 k^2 \Omega^2 \mu^2 t_{inj}}{1536Dk^2} \tag{5.9}
\]
where \( t_{inj} \) is the injection time. While the 6th order relation is theoretically reasonable, it is experimentally known that variance due to injection volume is essentially insignificant if the injection volume is kept below 1% of the total column volume. A calculation of the injection volumes used in these experiments shows that in all cases the injection volume was well below 1% of the total capillary volume (see Table 5.2).

Since the walls of the separation capillary were coated with polyacrylamide and both the 5%T/5%C and 3%T/3%C gels are solid matrices, variance due to electroosmotic flow, solute-wall interactions and siphoning are ignored in this evaluation. Differences in conductivity between the separation matrix and sample solution do contribute to band broadening in each of the cases. However, the influence of conductivity has no relation to capillary diameter, therefore, it is also ignored in this evaluation.

The longitudinal diffusion contribution, which is shown in Equation 5.10:

\[
\sigma_d^2 = \frac{2D}{L} t 
\]

is also not influenced by capillary diameter. Application of a regressive fit to the data obtained from separations using the 5%T/5%C gel in 100, 75 and 50 \( \mu \)m i.d. capillaries (Figure 5.4) yields only a 2nd order dependence upon the capillary diameter. Further attempts were made to observe the efficiency increases expected from separations performed in 20 \( \mu \)m i.d. capillaries. However, fabrication of 20 \( \mu \)m columns filled with crosslinked media has proven difficult. The problems include the high pressures (1000-3000 psi) needed to fill the 20 \( \mu \)m i.d. columns in approximately 1 minute with viscous gel solutions (approximately 2000 cp.) before the onset of polymerization. Also, the fact that after polymerization, all tubing and pressure vessels must be cleared of all solidified polyacrylamide material or replaced. An alternative method of capillary gel fabrication was developed to produce a large number of cross-linked filled capillaries and minimize problems associated with gel
forming in capillaries with i.d.s of 50 μm or less. This system is a modification of the one first used by Poppe,\textsuperscript{26} which employed riboflavin and UV irradiation to initiate polymerization of the polyacrylamide. The reported system used acrylate coated capillaries to allow sufficient bombardment of UV light on the gel matrix. The polyimide coating present on most fused-silica capillaries does not allow the efficient passage of UV light; however, the polyimide coating imparts a mechanical stability to fused-silica tubing that is not present with the acrylate coating. Our initial attempts to use light initiated polymerization employed polyimide coated capillaries with a small section of the coating removed to allow exposure of the unpolymerized material inside the tube (see Figure 5.6a). This system was successful in causing gel polymerization, but because of the non-uniform UV light exposure, gross inhomogeneities are found in the gel matrix. The irregular nature of sieving matrices produced in this manner make them unsuitable for use.

Figure 5.6
Diagram of photo initiated polymerization systems used to fabricate cross-linked gel filled capillary columns. UV source was a Xe lamp. Capillaries were chilled in an ice bath during UV exposure.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Attempts to increase the uniformity of the gel matrix included forming multiple windows along the column to allow more uniform light distribution as shown in Figure 5.6b. This system seemed to improve gel uniformity, as evidenced by monitoring the incidence of bubble formation in the matrix. Unfortunately, the presence of multiple windows so severely weakened the mechanical strength of the columns that they were no longer usable.

Small diameter capillary columns filled with crosslinked polyacrylamides were difficult to fabricate; Therefore, I attempted to study the influence of small capillary diameter using columns filled with LPA gels. Capillary columns with i.d. = 20 \( \mu \text{m} \) were then filled with LPA solutions to separate the DNA ladders. Column fabrication was successful using this approach, however the small path length encountered was detrimental to absorbance detection. No peaks were observed during electrophoresis when capillaries with 20 \( \mu \text{m} \) i.d. were used.

5.4 Conclusions

I have designed a custom-built high pressure injection system to insert the linear polyacrylamide solution into a capillary column prior to electrophoresis. The replaceable matrix will allow continuous operation of the gel electrophoresis without having to replace the capillary tube. I have also demonstrated that when using 5\%T/5\%C polyacrylamide gels a substantial improvement is realized with column diameter is reduced from 100 \( \mu \text{m} \) to 50 \( \mu \text{m} \) and it is suspected that additional gains will be realized from further reductions in separation column diameter. The efficiency enhancements seem to be dependent upon the choice of sieving medium, which is an unexpected result.

5.5 Future Studies

Dye-labeled sequencing ladders should be used to probe the efficiency of the system. The dye-labeled DNA will allow the use of a LIF detection system to provide
adequate sensitivity to detect the small analyte mass injected onto the column with 20 μm i.d.s.

A comparison reduced diameter efficiency enhancements achieved when using highly flexible LPA solutions versus those achieved when using cross-linked gels which possess less motional freedom may also provide insight into this phenomenon. In addition, an evaluation of elution times would also yield some important information on the mode of migration. If gel resistance is a factor in migration, then identical gel solutions in progressively smaller capillary columns (which have been reported to augment this resistive force) should retard DNA fragment mobility.

5.6 References

5.5. Lumpkin, O. J.; Zimm, B. H. Biopolymers, 1982, 21, 2315-2316.


6 Conclusions and Future Studies

6.1 Conclusions

The use of near-IR fluorescence in DNA sequencing applications using CGE can result in significant advancements in detection and instrumentation. The background resulting from scattering and fluorescence interferences is much smaller in the near-IR region, which results in impressive sensitivity without the need for off-column detection that can add to the complexity of the instrument. Since many near-IR dyes have absorption maxima which match fundamental lasing lines of inexpensive semiconductor diode lasers, sensitive fluorescence detectors can be constructed in multiplexing devices appropriate for handling large-scale sequencing projects, such as the Human Genome Project, at a fraction of the cost associated with Ar-ion laser systems. In addition, these diode lasers can be constructed to deliver milliwatts to several watts of laser power and run continuously for long periods of time without the need for replacement of ion tubes. We have also shown that near-IR dye-labeled primers can be effectively used in a single-lane, single fluor peak height identification protocol in CGE applications. Effective reads could be obtained up to 400 bases from the primer annealing site in a single run. However, the error rates arising from misidentifications of the terminal base were significant at those read lengths. Accuracy rates of 84% using a 4:2:1:0 intensity ratio were achieved up to 250 bases. Significant advances must still be made to improve the base-calling accuracy in a single lane format for near-IR use.

The use of solid-phase nanoscale reactor-based DNA polymerase reaction system with on-line detection could result in significant advancements in sequencing. The amenability of the solid-phase system to analyze varying templates and allow the use of multiple enzyme types will facilitate genetic analysis. The ability to interchange the DNA template easily and efficiently will facilitate the use of this reactor system in a seemingly unlimited number of applications. We have established that the activity of
the Vent® enzyme in the microreaction vessel is adequate to perform DNA polymerase reactions. I expect that a number of other DNA polymerase enzymes will demonstrate a compatibility toward use in this format, which is of integral importance if this system is to be useful in the numerous and highly varied approaches to DNA analysis. Favorable temperature stability coupled with resistance to template removal under rinsing conditions indicate that this system may perform well as a reusable cycle sequencing tool which will allow the user to perform multiple sequencing reactions with minimal sample preparation and purification procedures needed. Primer-walking protocols, which use information obtained from an initial extension cycle to design primers for subsequent extensions, should prove easily adaptable for use in a microreactor format. Reduction of the amounts of consumables will provide dramatic cost savings for large scale sequencing projects, such as the Human Genome Project. Further reduction in the volume of material produced in this system should be easily accomplished by simply reducing the radius of the reactor core. The terminal focus of this research is the further reduction of reactor size and direct integration into a microchip based separation and detection system to yield a completely autonomous DNA analysis unit. DNA sequencing applications could be integrated into this system by incorporating multi-fluor or temporal discrimination detection systems which have been described.1-4

The use of ultra-small capillaries or channels will increase separation speeds and efficiency in DNA sequencing and other electrophoretic analyses. The benefits of smaller separation columns in gel-based electro-driven separations on efficiency have been demonstrated. This benefit in efficiency is not due to reductions in thermal gradients (σ_d2) or decreased contributions from injection volume (σ_l2). A greater understanding of the nature of the interaction of the gel matrix and the migrating analyte will aid in the development of new sieving media which can be designed to possess desirable characteristics.
6.2 Future Studies

The development of a multi-probe, nanoscale, autonomous near-IR sequencing system is the terminal goal of this research. Increased sequencing accuracy is an immediate need for the near-IR LIF detection/base-calling system. Near absolute (100%) accuracy is required for total human genomic sequencing, therefore, base-calling strategies must be improved to address this requirement.

Reduction in the costs of DNA sequencing is also paramount. Costs can be reduced by simply reducing the amounts of material needed to perform a DNA analysis. Further development of nanoscale reaction systems will facilitate the reduction of consumable costs thereby, reducing the costs of sequencing in general. Also the interfacing of nanoscale reactors with fractionation and detection may make the autonomous DNA sequencing feasible. A completely self-governing DNA sequencer would reduce operation costs by reducing the need for manpower. This system would also speed analysis due to its ability to operate continuously with minimal operator intervention. An understanding of the effects of column miniaturization on fragment separation efficiency will be important to the development of a nanoscale autonomous sequencing system. This will promote the development of sieving media which will enhance fractionation and will facilitate detection and base-calling.

6.3 References


Appendix A

Calculation of Theoretical Signal from reactor

Reactor surface area $1.57 \times 10^5 \text{m}^2$
Streptavidin surface area $2000 \text{Å}^2 (2 \times 10^{-17} \text{m}^2)$

Thus:

$$\frac{1.57 \times 10^{-5} \text{m}^2}{2 \times 10^{-17} \text{m}^2} = 7.85 \times 10^{11} \text{sites}$$

Assumptions:
(1) Area of the streptavidin molecule is the limiting factor in calculating the number of available binding sites.
(2) Binding of biotin-labeled DNA to immobilized streptavidin is 100%.

Therefore:

$$7.85 \times 10^{11} \text{sites} = 7.85 \times 10^{11} \text{DNAs}_{(ss)}$$

DNA templates were labeled with $^{32}\text{P}$ by T4 kinase at 50% efficiency.
Thus:

$$7.85 \times 10^{11} \times 50\% = 3.93 \times 10^{11} \text{P} \text{DNA}$$

Calculation of expected radiation yields:

$$3.93 \times 10^{11} \text{P} \text{DNA} \times 4.99 \times 10^{-18} \text{Ci/}^{32}\text{P DNA} = 1.96 \times 10^{6} \text{Ci}$$

Conversion of Ci to cpm or dpm yields:

$$1.96 \times 10^{6} \times 2.22 \times 10^{12} \text{cpm/Ci} = 4.350,000 \text{cpm}$$

The 4.35M cpm total must be adjusted for $t_{1/2}$ deterioration for 72 days.

$$\frac{0.049 \text{day}^{-1}(72 \text{days})}{2.303} = 1.54$$

Antilog$_{10}$ 1.54 = 34.67

and finally.

$$\frac{4,350,000}{34.6} = 126,000 \text{cpm}$$
Appendix B

Calculation of loss rate immobilized template using biotin/streptavidin binding coefficient.

The time required to completely flush the reactor is;

\[ t_{\text{flush}} = \frac{V_{\text{col}}}{f} \]  \hspace{1cm} (B1)

The amount of material removed with each flush can be calculated using the binding coefficient and equilibrium expression;

\[ K_b = \frac{[\text{BSA}]}{[\text{B}][\text{SA}]} = 10^{15} \text{M}^{-1} \]  \hspace{1cm} (B2)

and;

\[ [\text{BSA}] = \frac{SA_{\text{surface}}}{N_{\text{col}}} \]  \hspace{1cm} (B3)

where \([\text{BSA}]\) is the amount of biotin-streptavidin complex in the system. If we assume 100% coverage and \([\text{B}]\) biotin = \([\text{SA}]\) streptavidin then;

\[ [\text{SA}] = \sqrt{\frac{\text{BSA}}{K_b}} \]  \hspace{1cm} (B4)

Each flush volume will remove \([\text{SA}]\) thus;

\((\text{Rinse time}/t_{\text{flush}}) \times V_{\text{col}} \times [\text{SA}] = \# \text{ moles removed.}\)

For example, using the values in Table 4.1 and Equations B3 and B4 the \([\text{SA}]\) in the reactor is \(5 \times 10^{11}\text{M}\). Assuming a flow rate of 11 nL/s (\(f\)) (Figure 4.12) and reactor volume of 390 nL (Table 4.1), the entire volume of the reactor can be rinsed in 35 seconds \(t_{\text{flush}}\). Thus, the moles of SA are reduced by;

\begin{align*}
2 \times 10^{-17} \text{ moles at 35 seconds} & \quad (5 \times 10^{11} \text{ M} \times 390 \times 10^{-9} \text{ L}) \\
4 \times 10^{-17} \text{ moles at 1 minute} & \quad (5 \times 10^{11} \text{ M} \times 780 \times 10^{-9} \text{ L}) \\
2 \times 10^{-15} \text{ moles at 1 hour} & \quad (5 \times 10^{11} \text{ M} \times 470 \times 10^{-7} \text{ L})
\end{align*}

and so forth. Thus, at 240 hours \(= 5.6 \times 10^{-13}\) moles of SA have been removed. The total amount of SA in the reactor initially was \(= 1 \times 10^{12}\) moles. Therefore one would expect an 80% removal of SA from the reactor after 240 hours of rinsing.
Appendix C

EDUCATION

Louisiana State University, Baton Rouge, LA
Doctorate of Philosophy, in Chemistry, Fall 1998

The Tulane University, New Orleans, LA
Bachelor of Science, May 1992
Major: Chemistry

EXPERIENCE

Research Assistant August 1994-Present
Louisiana State University Baton Rouge, LA
Department of Chemistry
Experienced in nucleic acid sequencing including DNA synthesis, PCR and recombinant amplification. I have also designed and constructed novel instrumentation for DNA polymerase reactions. Possess strong expertise in electrophoretic separation of biomolecules (CE, CGE, SDS-PAGE) and laser induced fluorescence detection. Also familiar with amino acid and protein separation. Experienced in directed mutagenesis and transformation. Have performed immunoassays (homogenous and heterogeneous) and ELISA.

Teaching Assistant August 1992-August 1994
Louisiana State University Baton Rouge, LA
Department of Chemistry
Responsible for the instruction of undergraduate students for laboratory course Chem. 1212 (introductory chemistry lab). Responsibilities included lecturing, grading and enforcement of laboratory safety regulations.

Summer Chemist May-August 1992
Albemarle Corporation Baton Rouge, LA
Process Development Center
Specialized in method development for the determination of Mn in environmental air and ground samples using Atomic Absorption spectroscopy.

Summer Chemist May-August 1991
Albemarle Corporation Baton Rouge, LA
Process Development Center
Specialized in method development for the determination of various transition metals in nonroutine environmental and fuel samples using x-ray fluorescence.

Summer Chemist May-August 1990
Albemarle Corporation Baton Rouge, LA
Process Development Center
Trained in the operation of the Karl Ficher 636 Titroprocessor, Mitsubishi water determination oven, and the Phillips 1490 x-ray fluorescence instrument for the determination of Si, Cl, and P in nonroutine samples. Experienced in numerous methods of halogen determination. Specialized in the determination of Pb in gasoline, soil and concrete. Formally oriented with OSHA 77 and the TSCA.
Summer Chemist
Albemarle Corporation

Researeh and Development Center
Assisted Senior Research Chemists in various experiments and project reports. Trained in the operation of NMR, GC, GC-MS, FT-IR, HPLC, Parr Hydrogenation and high pressure reaction vessels. Formally trained in the operation of self-contained breathing apparatus.

ACTIVITIES
Vice President, National Organization of Black Chemists and Chemical Engineers (L.S.U. Chapter 1995-96)
Member National Society of Black Engineers
Member ACS, Analytical Division
President L.S.U. Chemistry Graduate Student Council (1993-1994)

HONORS
Huel Perkins Fellowship (1995)
Phi Lambda Upsilon Travel Grant Recipient
Procter & Gamble Technology Management Conferee
Phi Lambda Upsilon (1993)
L.S.U. Graduate School Tuition Award
Tulane University Scholarship
National Merit Semifinalist
Superior Performance in Research, LSU Chemistry Department (1997)

PRESENTATIONS
Automated DNA Sequencing Using a Microreaction Vessel, Daryl C. Williams and Steven A. Soper, Workshop on Bioanalysis, Lawrence, KS June, 1996. (poster presentation)

DNA Sequencing Using A Microreactor, Daryl C. Williams and Steven A. Soper, High Performance Capillary Electrophoresis Conference, Orlando, FL January, 1996. (poster presentation)


Use of Near-IR Probes on Oligonucleotides for DNA Sequencing, Daryl C. Williams, James H. Flanagan Jr., ACS Regional Meeting, Birmingham, AL October, 1994. (oral presentation)

PUBLICATIONS


130

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.


Vita

Daryl Clifford Williams was born in Montgomery, Alabama on March 25, 1969. He attended the St. Jude Educational Institute where he was named a National Merit Semifinalist and received his H. S. Diploma in 1987. He entered The Tulane University of Louisiana for his undergraduate studies, where his undergraduate research project was directed by Dr. Harry E. Ensley. Daryl earned his Bachelor of Science degree in Chemistry in May 1992. In August, 1992, after an internship with the Albemarle Corporation, he began his graduate study at Louisiana State University, where he worked in the field of Analytical Chemistry under the supervision of Dr. Steven A. Soper. He was selected to participate in the Proctor & Gamble Technology Management Conference and was initiated into Phi Lambda Upsilon in 1993. In 1995, Daryl was named as one of the inaugural Huel Perkins Fellows. Daryl also received an award for Superior Performance in Chemistry from the Louisiana State University Chemistry Department. He is a member of the American Chemical Society and Alpha Phi Alpha Fraternity Inc. Daryl has served as President of the Chemistry Graduate Student Council in 1993 and V. President of the National Organization of Black Chemists and Chemical Engineers in 1995. He received the degree of Doctor of Philosophy at Louisiana State University in 1999. Daryl is currently employed as a Senior Chemist by the General Electric Company in Louisville, KY.
Candidate: Daryl C. Williams

Major Field: Chemistry

Title of Dissertation: A Novel Approach to DNA Sequencing Employing Near-IR Fluorescence Detection Coupled with Micro-scale Reaction Vessels with Capillary Gel Separations

Approved:

[Signatures]

Major Professor and Chairman

Dean of the Graduate School

EXAMINING COMMITTEE:

[Signatures]

Date of Examination:

25 April 1997