

2000

High throughput sample preparation and analysis for DNA sequencing, PCR and combinatorial screening of catalysis based on capillary array technique

Yonghua Zhang
Iowa State University

Follow this and additional works at: <https://lib.dr.iastate.edu/rtd>

 Part of the [Analytical Chemistry Commons](#)

Recommended Citation

Zhang, Yonghua, "High throughput sample preparation and analysis for DNA sequencing, PCR and combinatorial screening of catalysis based on capillary array technique " (2000). *Retrospective Theses and Dissertations*. 12380.
<https://lib.dr.iastate.edu/rtd/12380>

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.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
800-521-0600

UMI[®]

High throughput sample preparation and analysis for DNA sequencing, PCR and
combinatorial screening of catalysis based on capillary array technique

by

Yonghua Zhang

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: Analytical Chemistry

Major Professor: Edward S. Yeung

Iowa State University

Ames, Iowa

2000

UMI Number: 9990507

UMI[®]

UMI Microform 9990507

Copyright 2001 by Bell & Howell Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

Bell & Howell Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

Graduate College
Iowa State University

This is to certify that the Doctoral dissertation of
Yonghua Zhang
has met the dissertation requirement of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program

Signature was redacted for privacy.

For the Graduate College

TABLE OF CONTENTS

ABSTRACT.....	v
CHAPTER 1. GENERAL INTRODUCTION	1
Dissertation Organization	1
Overview.....	1
Sample preparation for DNA sequencing.....	4
Sample preparation for PCR.....	8
Total system integration.....	10
Capillary array instruments.....	14
High throughput analysis for combinatorial screening of catalysis.....	18
References.....	23
CHAPTER 2. MULTIPLEXED AUTOMATED DNA SEQUENCING DIRECTLY FROM SINGLE BACTERIAL COLONIES	30
Abstract.....	30
Introduction.....	31
Experimental Section.....	33
Results and Discussion	34
Conclusions.....	42
Acknowledgement	44
References.....	44
CHAPTER 3. PCR ANALYSIS BY CAPILLARY ELECTROPHORESIS WITH UV DETECTION FROM CLINICAL SAMPLE.....	56
Abstract.....	56
Introduction.....	57
Experimental Section.....	59
Results and Discussion	63
Conclusions.....	66
Acknowledgement	66
References.....	67
CHAPTER 4. COMBINATORIAL SCREENING OF HOMOGENEOUS CATALYSIS AND REACTION OPTIMIZATION BASED ON MULTIPLEXED CAPILLARY ELECTROPHORESIS	74
Acknowledgement	79

Supporting Information.....	79
References.....	79
CHAPTER 5. GENERAL CONCLUSIONS.....	84
APPENDIX A. SUPPORTING INFORMATION FOR CHAPTER 2	86
APPENDIX B. SUPPORTING INFORMATION FOR CHAPTER 4	90
ACKNOWLEDGMENTS	97

ABSTRACT

Sample preparation has been one of the major bottlenecks for many high throughput analyses. The purpose of this research was to develop new sample preparation and integration approach for DNA sequencing, PCR based DNA analysis and combinatorial screening of homogenous catalysis based on multiplexed capillary electrophoresis with laser induced fluorescence or imaging UV absorption detection.

We first introduced a method to integrate the front-end tasks to DNA capillary-array sequencers. Protocols for directly sequencing the plasmids from a single bacterial colony in fused-silica capillaries were developed. After the colony was picked, lysis was accomplished in situ in the plastic sample tube using either a thermocycler or heating block. Upon heating, the plasmids were released while chromosomal DNA and membrane proteins were denatured and precipitated to the bottom of the tube. After adding enzyme and Sanger reagents, the resulting solution was aspirated into the reaction capillaries by a syringe pump, and cycle sequencing was initiated. No deleterious effect upon the reaction efficiency, the on-line purification system, or the capillary electrophoresis separation was observed, even though the crude lysate was used as the template. Multiplexed on-line DNA sequencing data from 8 parallel channels allowed base calling up to 620 bp with an accuracy of 98%. The entire system can be automatically regenerated for repeated operation.

For PCR based DNA analysis, we demonstrated that capillary electrophoresis with UV detection can be used for DNA analysis starting from clinical sample without purification. After PCR reaction using cheek cell, blood or HIV-1 gag DNA, the reaction mixture was injected into the capillary either on-line or off-line by base stacking. The protocol was also applied to capillary array electrophoresis. The use of cheaper detection,

and the elimination of purification of DNA sample before or after PCR reaction, will make this approach an attractive alternative to current methods for genetic analysis and disease diagnosis.

We further developed a new methodology, nonaqueous capillary array electrophoresis coupled with microreaction, to address the throughput needs of combinatorial approaches to homogeneous catalysis screening and reaction optimization. Nonaqueous CE was used because of the solvability of the products. Samples were injected directly from reaction vial without dilution and reaction quenching. Buffer compatibility was also found important for reliable 96-capillary array injection. By choosing different Pd and base, a combination of 88 different reaction conditions was quickly tested. The analysis time was less than one minute for one sample, which is a truly high throughput technique for catalysis where the information of regioselectivity and stereoselectivity is the motivation of combinatorial screening.

CHAPTER 1. GENERAL INTRODUCTION

Dissertation organization

This dissertation begins with a general introduction about high throughput technique and sample preparation in DNA sequencing, PCR analysis and combinatorial screening of catalysis. The following chapters are presented as three complete scientific manuscripts. General conclusions summarize the work. Appendices include supporting data for the articles and diagrams of the instrument used.

Overview

Life sciences took center stage virtually around the world on June 26, 2000, when President Bill Clinton announced the completion of the first working draft of the entire human genome.¹ The working draft consisted of 85-90% of 3 billion bases in the human genome. It is essentially the "blueprint " for the construction of man, which holds great potential in the discovery of disease genes, the detection of disease-causing mutations within the genes, and the development of diagnostic and therapeutic procedures to detect, treat, and prevent human diseases ranging from cancer to AIDS. DNA sequencing is one of the most important parts in the human genome project. It is the only way to obtain the necessary biological information for understanding of gene structure and functionality.²

Current DNA sequencing analysis involves (1) DNA library construction and template preparation; (2) sequencing reaction, separation and detection; and (3) information analysis, The first two parts are the most challenging and rate-limiting steps in the whole process of the human genome project.³ Since the inception of the Human Genome Project,

the individual steps of DNA sequencing technology have improved dramatically. These include the development of bacterial artificial chromosomes (BACs) for longer insert clones,⁴ new thermostable enzymes,⁵ high sensitivity dye sets,⁶⁻⁸ and perhaps most important highly multiplexed high-speed capillary array electrophoresis (CAE) instruments.⁹⁻¹² Assuming that all HGP goals can be met with the new techniques, all human genetic codes could be unveiled by the end of 2001. However, the human genome is only one of almost one hundred genomes being currently sequenced around the world. Also, de novo sequencing, sequencing of other organisms, comparative genomics, single nucleotide polymorphism (SNP), and other genome-related issues are examples of applications that will continue to advance the limits of DNA sequencing.² In order to realize the potential that genomics hold for us, the current techniques of DNA sequencing need to undergo further reduction in cost and increase in throughput by integration, automation and miniaturization of the front end of DNA sequencing-sample preparation with the back end.

The clinical use of genomic information is one of the main objectives of the Human Genome Project.² The entire human genome is estimated to comprise at least 100,000 genes, of which 4000 have already been identified as the cause of known heritable genetic diseases. Use of genomic information will also help pharmaceutical companies create drugs tailored to a patient's genetic profile, boosting effectiveness while drastically reducing side effects. It could even change our very concept of what a disease is, replacing broad descriptive categories with precise genetic definitions that make diagnosis sure and treatment swift. Advances in molecular biology have allowed the identification of genes and the specific mutations linked with a variety of human diseases, where associations between DNA polymorphisms and specific mutations have been established. Rapid and cost-effective

methods for detection of such mutations are the basis of genetic diagnosis, which is expected to play an increasingly important role in the fields of molecular pathology and genetics.¹³

A revolutionary technique for the detection of polymorphism is DNA amplification by the polymerase chain reaction (PCR), which allows the amplification of selected regions of DNA extracted from a variety of sample sources to a detectable level.¹⁴ It is also possible to perform multiplex amplification in cases where multiple mutations are present. Large-scale DNA polymorphism detection requires the development of new techniques which are fast, cost-effective and easily automated.¹⁵ Traditionally, DNA used for PCR-based diagnostic analysis has originated from blood, which involves labor intensive sample preparation. Considering the large number of samples required, a change in sample preparation is required.

Following the lead of genomics, chemists have developed combinatorial approaches to synthesize large arrays of related molecules or materials and screening them to identify the fittest members (e.g., the best inhibitor of a given enzyme, or the best catalyst for given transformation).¹⁶ While combinatorial chemistry has established its importance for drug screening, combinatorial methods are just starting to be applied increasingly in the field of materials research and catalyst development.¹⁷ One of the challenges, however, has been to develop fast and reliable high throughput analytical methods to support the high throughput synthesis activities of most medicinal and combinatorial chemistry departments. An analytical bottleneck exists principally because syntheses are performed in parallel, whereas analyses, are principally, conducted in a serial based manner. The elimination of bottleneck of the analysis by parallel separation is therefore high desirable.

Sample preparation for DNA sequencing

Plasmid purification

While cost effective and robust methods for purification of cloned DNA from bacteria impact the success of any DNA sequencing effort, they are especially crucial for high-throughput, large-scale projects where many thousands of DNA purification and subsequent sequencing reactions are performed daily. These activities require methods that are easily implemented, capable of the required throughput. The methods should also have a low level of technical difficulty, yield DNA in reproducible amounts, and are applicable of different bacterial host/vector combinations.¹⁸

Many procedures have been developed over the years for isolation of bacterial plasmids.¹² Further, several proprietary methods also give satisfactory results. One aspect that nearly all of these methods share is that they involved three basic steps: growth of bacteria, harvesting and lysis of the cells, and purification of the plasmid.¹⁹ All include centrifugation or several treatments of the colonies which are labor intensive and difficult to automate and interface with CAE. Some chromatographic columns (size exclusion,²⁰ ion-exchange,²¹ high-performance membranes²² hydrophobic interaction chromatography²³) have also been developed to avoid the use of centrifugation. All these methods, which may meet the stringent quality criteria for gene therapy, may introduce high cost and intensive labor in the separation. Another complication is that reagents used in purification of plasmids, such as ethanol and SDS,²⁴ might become interferences to the subsequent cycle-sequencing reaction.

Alternative methods also exist which require only heat-induced lysis of cells in bacterial colonies.^{25,26} Centrifugation or vigorous vortex are still needed to isolate cell debris. The resulting lysate is then used as the template in cycle sequencing using labeled primers.

Even though the performance and the ruggedness demonstrated so far by this method is still inferior to that of the standard protocol, it shows promise for significant savings in time and cost, particularly for large-scale sequencing. In fact a microwave protocol similar to above procedure has already become the preferred method for purification of double-stranded DNA at the Washington University Genome Sequencing center.²⁷ In their methods, the growth of bacterial cultures and subsequent DNA isolation took place in the same 96-well block and no further purification of DNA, by precipitation or other means, was necessary. The 96-well blocks can be reused indefinitely provided they were cleaned between uses. Furthermore the lysis solution was easily made and was stable at room temperature for a minimum of 3 months, allowing liter-quantity batches to be made and stored. The cost was estimated to be 3 cents per sample. This compared very favorably with the cost of commercial preparation methods, which can be 1 dollar per sample. Using this approach, the average high quality sequence length was 427 bases, while 70.52% of the sequences had at least 400 bases of high quality data.

DNA ladder purification and sample injection

Capillary gel electrophoresis (CGE) is an attractive technique for DNA analysis because the narrow-bore, gel filled capillaries provide high-speed, high-resolution separations, as well as automated gel and sample loading. The use of CGE for DNA sequencing was first demonstrated in 1990, when sequencing separations of ~350 bases were obtained on crossed-linked gels in ~ 80min.²⁸ Much progress has been made in the past decade, and sequencing read-lengths of more than 1000 bases can now be obtained using replaceable gel both in Dr Karger's group²⁹ and Dr. Yeung's group.³⁰ In fact, CGE has

eliminated the bottleneck involved in separation the DNA ladder produced by the Sanger reaction and has become the primary choice for DNA sequencing.

Compared with slab gel electrophoresis, the sample in CGE is injected into the separation capillary instead of loading and therefore dye labeled DNA fragments must compete with ions in the sample matrix to enter the capillary. It is well documented that the performance and reliability of DNA sequencing by capillary electrophoresis is sensitive to the quality of the DNA sample due to the electrokinetic injection method employed.³¹ Residual salt and dideoxynucleotides in the sequencing sample cause discrimination towards DNA because they have higher mobilities. As a result, a great variability in signal strength is often observed in CGE. Another problem associated with sample purification is the rehybridization of the single DNA strand injected. This causes the baseline bump and makes base calling difficult.

The benefit of stringent sample purification is demonstrated in articles by Ruiz et al.³² and Salas-Solano et al.'s³³ articles that described a novel sample purification method. A poly(ether sulfone) ultrafiltration membrane pretreated with linear polyacrylamide was first used to remove template DNA from the sequencing samples. Then, gel filtration in a spin column format (two columns per sample) was employed to decrease the concentration of salts below 10 μM in the sample solution. The method was very reproducible and increased the injected amount of the sequencing fragments 10-50 fold compared to traditional cleanup protocols. Using M13mp18 as template, the resulting cleaned-up single DNA sequencing fragments could routinely be separated to more than 1000 bases with a base-calling accuracy of at least 99% for 800 bases. A systematic study to determine the quantitative effects of the sample solution components such as high-mobility ions (e.g., chloride and dideoxynucleotides) and template DNA on the injected amount and separation efficiency of

the sequencing fragments revealed that in the presence of only 0.1 μg of template in the sample (one-third of the lowest quantity recommended in cycle sequencing) and at very low chloride concentration. ($\sim 5 \mu\text{M}$), the separation efficiency decreased by 70%. The deleterious effect of template DNA on the separation of sequencing fragments was not observed in ordinary cleaning sample because it was suppressed in the presence of salt in a concentration above 100 μM in the sample solution. The latest result with newly formulated matrixes showed that read length up to 1300 bases (av. 1250) with 98.5% accuracy can be achieved in 2 h for a single-stranded M13 template.³⁴ Thus the purified DNA ladder dramatically improved the result but at the expense of high cost and manual manipulation. In slab gel electrophoresis, unpurified sample has been demonstrated in DNA sequencing.³⁵ Swerdlow et. al.³⁶ first tried to perform DNA sequencing with unpurified DNA sequencing samples. They used a method called base stacking which allowed direct injection of unpurified products of dye-primer sequencing reactions onto capillaries without any pretreatment. Briefly on-column concentration of DNA fragments is achieved simply by electrokinetic injection of hydroxide ions. A neutralization reaction between these OH^- ions and the cationic buffer component Tris^+ results in a zone of lower conductivity, within which field focusing occurs. Without base stacking a drastic loss in signal was observed for the crude samples. This method can generate separation resolution of at least 0.5 up to 650 bp and the signal strength was excellent relative to conventional injection of highly purified samples. Furthermore no significant degradation of the capillary performance was observed over at least 20 sequencing runs using this new sample injection method. One shortcoming of this method is that the method did not yield satisfactory results with dye terminator chemistry due to the interference of large signals from unreacted dye terminators.

Sample preparation for PCR

From blood

Since its introduction by Nobel laureate Kary Mullis in the early 1980s, the polymerase chain reaction (PCR) has gained a strong following in DNA analysis research for rapid detection of very small amounts of specific DNA sequences. Current applications include identification of cancer-associated genes, viral subtyping for AIDS and hepatitis, bacterial drug susceptibility research, and determination of genetic disorders.³⁷

In a theoretical clinical situation a PCR-based test will proceed through the steps outlined in Figure 1. It is interesting that a new PCR method, FoLT (formamide low temperature) PCR has been developed for reactions directly from whole blood. Formamide solubilizes blood cells and frees the DNA for amplification. An important finding was that an alternative DNA polymerase, Tth polymerase, was less sensitive than Taq polymerase to the presence of blood. All these make PCR directly from blood possible.^{38,39}

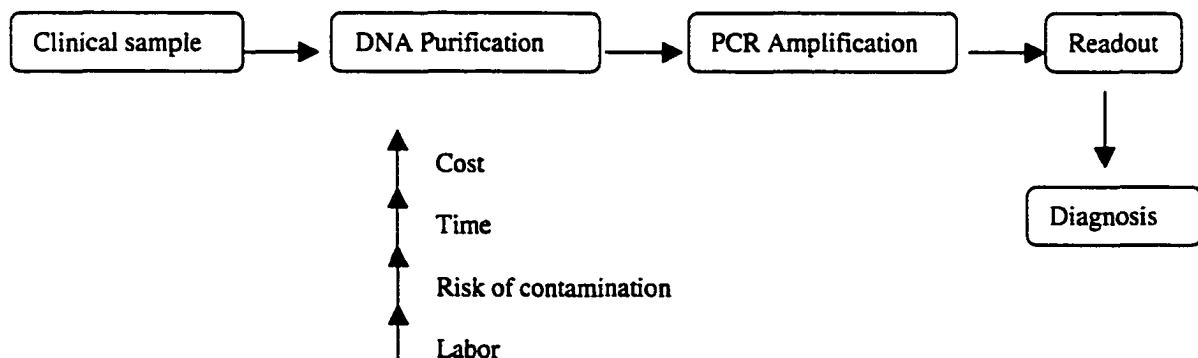


Figure 1. Flow diagram of a theoretical PCR test.

From cheek cell

Although blood can be used directly in PCR, there is clearly a need for simpler, non-invasive, and more cost-effective means of sample collection, DNA extraction and genetic diagnosis in general. There are several disadvantages of using blood. First blood collection can be very inconvenient, since genetic testing often involves analysis of multiple family members. Furthermore, drawing blood can be uncomfortable for the patient and most important, the handling of blood samples can increase the chance of infection of blood-borne pathogens such as HIV and hepatitis. To date a variety of alternate sources of DNA have been used for genetic testing including finger-prick blood samples,⁴⁰ hair roots⁴¹ as well as the use of cheek scrapings and oral saline rinses as a means of collecting buccal epithelial cells.⁴² The oral saline rinse is perhaps the most extensively used non-blood based sampling technique. The inconvenient practice of the mouthwash method is that it still involves liquid sample handling, and requires an additional centrifugation step to spin down the cells, which is difficult to automate and interface with subsequent analysis.

By avoiding centrifugation, a simpler method has been developed and validated by just using swab and brushes.⁴³ The buccal cells were collected on a sterile brush by twirling the brush on the inner cheek for 30 seconds. Although still involving a neutralizing step later, this method is generally easy and very reliable. In a blind study comparing the analysis of 12 mutations responsible for cystic fibrosis in multiplex products amplified with DNA from both blood and buccal cell samples from 464 individuals, there was 100% correlation of results for blood and cheek cell DNA. The success rate of PCR amplification on DNA prepared from buccal cells was 99%. This method has also been used by another group to

analyze DNA for genetic polymorphism by matrix-assisted laser desorption/ionization mass spectrometry.⁴⁴

Total system integration

The automation and integration of the front-end of DNA analysis such as DNA sample preparation smoothly with later separation part is extremely desirable especially with the array electrophoresis. This has to do with the small volumes and small amounts of samples that are needed for each electrophoretic run.⁴⁵ Typically, only nl sample is injected into the capillary. Normal sample preparation methods require μl range of sample to handle which wastes a substantial amount of reagent. Also lower costs will be achieved through a reduction in labor and instrument costs through automation and multiplexing. The attempts of integration have been demonstrated in three categories: robotics, capillary microfluidics and the lab-on-chip. Numerous endeavors have been made into developing robotic workstations to perform sequencing reaction, purification, pre-concentration and sample loading.^{46,47} Although robotics have shown advantages in repetitive operation with high precision, the adaptation to highly multiplexed capillary array separation interface suffers many incompatibilities in terms of the total reaction volume, purification by centrifugation and sample injection. On-line microfluidics system based on both capillary or microchip hold the promise for the next generation of total automated DNA sequencer.

Single and multiplexed capillary microfluidics systems

In the system of our group,⁴⁸ dye-labeled terminator cycle sequencing reactions are performed in a 250 μm i.d. fused-silica capillary which was placed into a hot-air thermal cycler. After PCR was completed, the reaction mixture was transferred on-line to a size-exclusion column to purify the reaction product from unreacted dye terminators. The purified

product was then injected through a cross onto a gel filled capillary to do the size separation. This system was closed and the operation was reliable since no moving part was involved. Cleaning of the system with 0.1M NaOH was required to remove cross-contaminants before reuse. The sequence could be called from 36 to 360 bases with an accuracy of 96.5% using in-house software; by manual editing, the accuracy improved to 98% for 370 bases. Later a multiplexed system based on above scheme was developed in which eight DNA sequencing samples could be processed simultaneously starting from DNA to called bases.⁴⁹ The major achievement in the instrument was the use of freeze/thaw switching valves instead of the previous rotary valves, which were unsuitable for multiplexed systems due to their size. For all eight processed samples, sequences could be called up to 400 bases with an accuracy of 98%. PCR analysis directly from blood was also demonstrated with a similar flow management concept.⁵⁰

A fully integrated single capillary instrument comparable in design to our group has also been designed and prototyped by Swerdlow et al.⁵¹ The reaction was performed inside Teflon tubing. Purification and separation columns were interfaced through a simple T-connector instead of a cross. The instrument was reliable and fast, performing PCR reaction cycling, purification and analysis all in 20 min. Adaptation of the instrument prototype for separation of DNA-sequencing reactions was described; cycle sequencing and electrophoresis of a single lane were complete in 90 min with bases calling beyond 600 bases.

Miniaturization of the on-line system will reduce the cost of DNA sequencing orders below current level since only one hundredth of reagent required for actual capillary electrophoresis. Soper et al.⁵² developed a miniaturized solid-phase cycle sequencing reactor

coupled with capillary electrophoresis. The nanoreactor consisted of a fused-silica column with a total volume of 62 nL. Biotinylated DNA template was bonded to the surface by biotin-streptavidin-biotin linkage. The main disadvantage of this scheme is the loss of template surface coverage. One solution for reactivation of the nanoreactor could be that by adding fresh streptavidin and new biotinylated target DNA. The read-length for a single color run was approximately 450 bases. The system is considered amenable to automation even though there is still manual operation with current design.

On-line microchip system

Microchips provide a new platform for integration with unique electroosmotic pumping and non-mechanical valves. A true nano-total analysis device was developed by Burns et. al.⁵³ which uses microfabricated fluidic channels, heaters, temperature sensors, and fluorescence detectors to analyze nanoliter-size DNA samples. The device is capable of measuring 100 nl reagent and DNA solution, on-line mixing the solutions together, amplifying or digesting the DNA to form discrete products, and separating and detecting those products on microfabricated channels. No external lenses, heaters, or mech. pumps are necessary for complete sample processing and analysis. The components have the potential for assembly into complex, low-power, integrated analysis systems at low unit cost due to the conventional photolithography used.

Microchips still need to overcome some technical difficulties before they come to mainstream. These include solvent evaporation and interface with outside world where the common working volume is μl in current bioanalysis lab. Litborn et. al. described using a closed humidity chamber to address a major problem, solvent evaporation.⁵⁴ Later they reported an improved technique for performing parallel reactions in open, 15 nL volume,

chip-based vials. The evaporation of solvent from the reaction fluid was continuously compensated by addition of solvent *via* an array of microcapillaries. Their results showed that the concept for continuous compensation of solvent evaporation should be applicable to reaction volumes down to 30 pL.⁵⁵

There are other approaches to avoid the problem of solvent evaporation. Soper also developed a hybrid system which coupled nanoliter sample preparation to PMMA-based microchips.⁵⁶ Compared with standard sample preparation are performed off-chip on a microliter-scale, a true integration was demonstrated on nl volume. An integrated system for rapid PCR-based analysis on a microchip has also been demonstrated recently by Ramsey's group.⁵⁷ The system coupled a compact thermal cycling assembly based on dual Peltier thermoelectric elements with a microchip gel electrophoresis platform. This configuration allowed fast (~1 min/cycle) and efficient DNA amplification on-chip followed by electrophoretic sizing and detection on the same chip. An unique on-chip DNA concentration technique based on adsorption and desorption has been incorporated into the system to further reduce analysis time by decreasing the number of thermal cycles to 10 cycles or 20 minutes for DNA amplification and subsequently detection.

Capillary array instruments

Commercial system

The current state of art technique for DNA sequencing is capillary array electrophoresis. In 1992, Mathies developed this approach to address the throughput requirement of genomic analysis.¹⁰ They constructed a confocal fluorescence scanner and demonstrated DNA sequencing in 25 parallel capillaries. Since then, constant improvement in optical design and separation matrixes has made the commercialization CAE sequencers

possible. Today, there are four commercial versions of CAE instruments. PE Biosystems has developed the ABI PRISM 3700 DNA analyzer.⁵⁸ This 96-capillary array instrument is based on the approaches of Kambara¹¹ and Dovichi.¹² In this instrument, DNA sequencing fragment are detected in sheath flow and spectrally resolved using a concave spectrograph and a cooled CCD camera. Bare capillaries are used with dynamic coating, which is stable for over 300 runs. The turn around time is roughly 2.6 hours with 600 bp in 120 min..

Molecular Dynamics' instrument MegaBACE 1000, is based on the confocal detection developed by Mathies.¹⁰ A microscope objective is used to focus the laser light inside the capillaries and, at the same time, collect the emitted light from the center of the column. The lifetime of the instrument can be limited by mechanical stress of moving parts of the scanner when fast sampling rates are required. One advantage of this system is four photomultiplier tubes instead of a CCD being utilized. The system uses LPA coated capillaries, which are stable for 200 runs. The average sequencing data is 500 bp and the turn around time is less than 2 hours.⁵⁹ Beckman Coulter has entered the market with an 8-capillary array design, the CEQ 2000 DNA analysis system.⁶⁰ The optical design of this instrument is similar to Molecular Dynamics' and features four-color IR dideoxy-terminator chemistry. On-column detection is the approach of the SpectruMedix instrument which was developed in our group.⁶¹ The laser beam crosses all 96 capillaries after the laser is expanded by a cylindrical lens. The fluorescent light is collected at a right angle from the laser axis and detected by a CCD camera. Since no moving parts are involved in detection, the optic design is very rugged. Bare fused silica capillaries are used with dynamic coating with hydrophobic PVP. The turn around time is 2 hours with average base calling of 500 bp.

Home-made systems

The new alternatives in system design are mainly concentrated on illumination and detection. Kambara et al. further tested side illumination with detection on column.⁶² For this design, the number of capillaries in an array is generally limited by laser-power attenuation along the array due to reflection and divergence. They overcame these problems by placing the capillaries in water and adding glass rod lenses between the capillaries. As a result, up to 45 capillaries could be simultaneously irradiated with a single laser beam and the fluorescence from all the capillaries could be detected with high sensitivity. Quesada⁶³ took another approach for a multiple capillary instrument by the use of optical fibers for illumination and collection of the fluorescence in a 90° arrangement. A more recent version of this instrument utilized cylindrical capillaries as optical elements in a waveguide, where refraction confined a focused laser beam to pass through 12 successive capillaries in a flat parallel array.⁶⁴ However, larger capillary arrays limited the refractive effects that spread the light along the length of the capillaries. Heller et al⁶⁵ has designed a new CAE instrument and they also compared it with the existing systems in term of detection limit. In their instrument, the illumination of all 96 capillaries was provided by a laser line generator to produce a uniform intensity profile over the array, yet using a low power laser (~30 mW). The instrument can also perform spectral analysis from 96 different channels with a single spectrograph. The sensitivity of 1×10^{-11} M (S/N=40), which can be extrapolated to a detection limit in the very low pM range, is comparable to that achieved with fiber optic detection (3.7×10^{-12} M)⁶⁴ or confocal scanning systems (2×10^{-12} M),¹⁰ and only sheath flow detection had a lower detection limit (1×10^{-13} M).¹²

Handling more than 96 capillaries is very challenging and several groups have attempted to address this problem by modifying existing design. Dovichi's group used sheath flow detection and a novel 2-D arrangement which can hold up to 576 capillaries.⁶⁶ A prototype 384-capillary array electrophoresis instrument has also been developed for higher throughput analysis by Spectrumedix.⁶⁷ Their instrument design is based on the 96 capillary platform followed by an exchange in the camera lens. Mathies has also continued to push the limit of the confocal system. They have developed a new system with the capillaries aligned in a circular array. The microscope objective spins inside a drum, illuminating each one of the capillaries at a time. They have shown sequencing data from 128 capillaries, but a larger number of capillaries could be easily accommodated in this geometry.^{68,69}

An automatic DNA fragment collector using capillary array gel electrophoresis has been developed in Dr. Kambara's group.⁷⁰ The array has the micro-preparative capability which is unique compared with other arrays. A sheath flow technique was used not only for detection but also for collection of DNA fragments. In a sheath flow cell, the DNA fragments separated by 16 capillaries flow independently into corresponding sampling capillaries. The fraction collector consisted of 16 sampling trays and each sampling tray was set beneath each end of the sampling capillaries to collect the flow-through DNA fragments. Certain DNA fragments were automatically sorted by controlling the movement of the sampling trays according to the signals from the system. Collected DNA fragments were even amplified by PCR and measured by electrophoresis. DNA fragments with base length differences of one (base lengths 363 and 364) were successfully separated. These results showed that the automatic DNA fragment collector is useful for gene hunting in research fields such as drug discovery and DNA diagnostics. Dr. Karger also designed a multi-capillary system which

permitted parallel separation and collection. The multi-well collection plate was preferably made of a solvent permeable gel.⁷¹

Microchip array

A totally different platform to perform CAE is by microchip. The first demonstration of CAE in microchip was by Mathies' group for genotyping.⁷² A microplate that can analyze 96 samples in less than 8 min have been produced by bonding 10-cm-diameter micromachined glass wafers to form a glass sandwich structure. The microplate had 96 sample wells and 48 separation channels with an injection unit that permitted the serial analysis of two different samples on each capillary. An elastomer sheet with an 8 by 12 array of holes was placed on top of the glass sandwich structure to define the sample wells. Samples are addressed with an electrode array that makes up the third layer of the assembly. Detection of all lanes with high temporal resolution was achieved using a laser-excited confocal fluorescence scanner as the authors used before. A single nucleotide polymorphism (SNP) typing assay has also been developed and evaluated on a microfabricated capillary array electrophoresis system by Mathies' group.⁷³ This study demonstrates the feasibility of using allele-specific PCR with covalently labeled primers for high speed fluorescent SNP typing.

DNA sequencing on microchip capillary array is very interesting and challenging. Earlier study of single channel on chip needed channel lengths comparable to capillaries. Making so many turns in chip also proved to be deleterious to separation performance and this may indicate a bigger diameter chip is needed for DNA sequencing.⁷⁴ Recent results show that there is still much room for improvement. Liu et. al. of Molecular Dynamic demonstrated DNA sequencing by 16 channel CAE in microchip format.⁷⁵ Samples are

loaded into sample reservoirs using an eight-tip pipetting device, and the chip is docked with an array of electrodes in the focal plane of a four-color scanning detection system. Under computer control, high voltage is applied to the appropriate reservoirs in a programmed sequence that injects and separates the DNA samples. An integrated four-color confocal fluorescent detector automatically scans all 16 channels. The system routinely yields more than 450 bases in 15 min in all 16 channels. In the best case using an automated base-calling program, 543 bases have been called at an accuracy of >99%. Separations, including automated chip loading and sample injection, normally are completed in less than 18 min. This demonstrates the potential of the microchip as the next generation CAE platform.

High throughput analysis for combinatorial screening of catalysis

There are few fields where combinatorial chemistry has had as long and as productive a history as catalysis. Most catalysts that are in commercial use, both homogenous and heterogeneous, have been developed via some form of combinatorial approach.⁷⁶ Combinatorial methods are also used to increase the efficiency of reaction optimization particularly in pharmaceutical development. Experimental variables will typically include monomer identity, solvent or solvent mixtures, temperature, reagent, catalyst, concentration and time. Given the complexity of this physicochemical space, parallel and combinatorial methods are attractive.⁷⁷ So far there are several methods for high throughput catalysis screening. They are IR thermography; resonance enhanced multiphoton ionization (REMPI); microplate UV-VIS absorption or fluorescence and mass spectrometry. Additionally, separation based techniques are often used as secondary screening to pinpoint the hit.

IR thermography

IR thermography has been used to probe the activity of catalysts since catalytic activity is indicated by the generation or absorption of heat in exo and endothermic reactions. The relative temperatures of library elements subjected to identical conditions can then be a gauge of their corresponding catalytic activity. Taylor and Morken⁷⁸ have used an IR camera to evaluate their library of nucleophiles for efficacy in catalyzing the acyl transfer reaction. The most active catalysts were identified as beads some 1°C hotter than the surrounding solvent and those beads without catalysts or with inactive catalysts. However, selectivity information is absent from the thermogram. This is a major disadvantage in partial oxidation catalysis, where the worst catalyst, the one which leads to total oxidation, will be the hottest. Nevertheless, thermograph will allow elimination of the totally inactive catalysts and might be an ideal method to prescreen the activities, but secondary screening is necessary to yield more information.

Resonance enhanced multiphoton ionization (REMPI)

Similar to IR thermography, resonance enhanced multiphoton ionization (REMPI), developed by Senkan,⁷⁹ is also optically triggered. Ionization of the target molecule via two-photon absorption with a laser generate ions that are captured by an electrode placed in the product plume behind or above the catalyst which is then detected as a charge signal. This principle has been shown to perform reliably in the case of the dehydrogenation of cyclohexane to benzene, where a 66-member library could be analyzed simultaneously. Broader application, however, is limited by the requirement of the target molecule having suitable electronic states with reasonable absorption cross-sections.

UV, colorimetric and fluorescence screening based on microplate reader

Optical properties are one of the most convenient testing methods because microplate readers have become increasingly popular and a wider variety of fluorescence assays has been available for a broader range of applications.⁸⁰

A novel catalyst for allylic alkylation in neutral medium was discovered with the aid of a fast, parallel colorimetric screening method.⁸¹ The strategy for detection of allylic alkylation catalysts relies on the fact that colorless 1-naphthol will undergo electrophilic aromatic substitution with Fast Red diazonium salt to give a bright orange azo product. Using the colorigenic azo coupling assay the authors had screened several 88 metal plus ligand combinations and 8 controls to ascertain that ligand alone does not generate color on treatment with Fast Red using a 96-well plate format. While simple visual analysis of the 96-well plate is sufficient to differentiate efficient from inactive catalysts, parallel UV analysis was used to differentiate catalysts of similar activity and provides a means to subtract background color due to ligand or metal salt. The colorigenic assay is simple and inexpensive, and should be applicable to other reactions of interest. Another interesting example is the enantioselective hydrolysis of chiral p-nitrophenol esters in which the course of the reactions of the R- and S-configured substrates was monitored in a parallel manner by UV/Vis spectroscopy. With the use of microtiter plates, crude screening of approximately 800 different enantioselective catalysts was possible per day.⁸²

Mass spectrometry

Mass spectrometry has proven to be the most widely utilized analytical tool to assess the selectivity or specific reaction products being generated even though optical methods allow the high-speed acquisition of data on the activities of catalysts in libraries without the

need to take samples.⁸³ Mass spectrometric analyses of combinatorial libraries are, in general, performed in one of three ways: (1) flow injection analysis -mass spectrometry (FIA-MS) ;(2) LC/MS with UV and /or evaporative light scattering detection; and (3) direct matrix-assisted laser desorption/ionization (MALDI) analysis to compounds. Flow injection-MS has proven to be the most widely incorporated technique for characterizing combinatorial libraries, due to its simplicity in design and more importantly, its throughput which is the highest of the mass spectrometric-based techniques currently used. Cycle times has been reported as short as 30 seconds/sample. The factors that limit FIA-MS speed have been related to the injection cycle, not the actual mass spectrometric acquisition time itself. To reduce the injection time, a multiple probe autosampler, capable of processing eight samples at a time has been demonstrated even though it is intrinsically a sequential method.⁸⁴

Mass spectrometry also can be readily used to analyze complex gaseous mixtures. Recently, Cong et al.⁸⁵ reported a mass spectrometer based system to screen heterogeneous catalyst libraries for the CO+O₂ and CO+NO reactions. The total time to heat and screen one catalyst site was reported to be about one minute. The results obtained are in complete agreement with the limited data reported previously in the literature. But it is also important to recognize that the sequential screening system developed by Cong et al. is slow and is only useful to determine the initial activities and selectivity of catalysts. Consequently, it will be of limited utility to identify leads for practical applications. The limitation arises from the fact that the activities and selectivity of heterogeneous catalysts significantly change (decrease and sometimes increase) with time on stream, which renders short-time data practically useless to assess the long-time performance of catalysts.

Senkan et. al.⁸⁶ proposed a very practical approach for screening by direct combination of an optical method with spatially resolved mass spectrometry. In large libraries with several hundred or thousand components, only active components were selected for mass spectrometric identification of selectivity. Furthermore, they proposed that the setup used can be readily modified. And by variation of the capillaries attached to the capillary bundle in the robot, a large number of other analytical techniques such as GC, GC-MS, IR, UV/ VIS, and Raman spectroscopy can be applied to libraries in a spatially resolved manner.

High throughput separation analysis

With relatively small combinatorial libraries, evaluation of the catalytic activity and selectivity can be performed in a sequential fashion using conventional means of separation or quantifying product concentrations such as HPLC, GC and CE. The throughput that can be achieved with serial separation schemes is low even with special techniques, such as sequential sample injection⁸⁷ and sample multiplexing.⁸⁸ The call for separation based screening originated in the shortcomings of the above methodologies for high throughput analysis. In these approaches, although the relative activity of the catalyst is determined swiftly, no information about the regioselectivity and stereoselectivity can be obtained easily. The product also needs exhibit very different measurable properties compared to the surrounding environment. Parallel separation is highly attractive in this aspect. Multiplexed HPLC is one interesting approach,⁸⁹ but achieving a high degree of multiplexing, such as 96 capillaries in CAE, is not trivial. CAE stands out compared with other techniques in term of speed, resolution and minimal sample requirement. Even though the detection for array such as UV-VIS absorption and fluorescence still not be considered information rich,⁹⁰ they are

useful for screening where, in most cases, a defined substrate and a desired product are already known.

References

1. Pennisi, E., *Science*, **2000**, 288, 2304.
2. Collins, F., Patrinos, A., Jordan, E.; Chakravarti, A., Gesteland, R., Walters, L., *Science*, **1998**, 282, 682.
3. Hunkapiller, T., Kaiser, R. J., Koop, B. F., Hood, L., *Science*, **1991**, 254, 59.
4. Shizuya, H., Birren, B., Kim, U.; Mancino, V., Slepak, T., Tachiiri, Y., Simon, M., *Proc. Natl. Acad. Sci. USA*, **1992**, 89, 8794.
5. Tabor, S. and Richardson, C. C., *Proc. Natl. Acad. Sci. USA*, **1995**, 92, 6339.
6. Rosenblum, B. B., Lee, L. G.; Spurgeon, S. L.; Khan, S. H., Menchen, S. M., Heiner, C. R., Chen, S. M., *Nucleic Acids Res.*, **1997**, 25, 4500.
7. Ju, J., Ruan, C; Fuller, C. W., Glazer, A. N.; Mathies, R. A., *Proc. Natl. Acad. Sci. USA*, **1995**, 92, 4347.
8. Lee, L. G., Spurgeon, S. L., Heiner, C. R., Benson, S. C., Rosenblum, B. B., Menchen, S. M., Graham, R. J., Constantinescu, A., Upadhy, K. G., Cassel, J. M., *Nucleic Acids Res.*, **1997**, 25, 2816.
9. Ueno, K., Yeung, E. S., *Anal. Chem.*, **1994**, 66, 1424.
10. Mathies, R. A., Huang, X. C., *Nature*, **1992**, 359, 167.
11. Dovichi, N. J., Swerdlow, H., Zhang, J. Z., Chen, D. Y., Harke, H. R., *Anal. Chem.*, **1991**, 63, 2835.
12. Takahashi, S., Murakami, K., Anazawa, T., Kambara, H., *Anal. Chem.*, **1994**, 66, 1021.
13. Sidransky, D., *Science*, **1997**, 278, 1054.

14. Bloch, W., *Biochem.*, **1991**, 30,2735.
15. Gao, Q., Yeung, E., *Anal. Chem.*, **2000**, 72, 2499.
16. *A Practical Guide to Combinatorial Chemistry*, Czarnih, A. and Dewitt, S., Ed., American Chemical Society, Washington DC, **1997**.
17. Jandeleit, B., Schaefer, D., Powers, T., Turner, H., and Weinberg, W., *Angew. Chem. Int. Ed.*, **1999**, 38, 2494
18. Venter, J.C., Adams, M.D., Sutton, G.G., Kerlavage, A.R., Smith, H.O., Hunkapiller, M., *Science*, **1998**, 280, 1540.
19. Mazza, P. G., *Microbiologica*, **1986**, 9, 113.
20. Edwardson, P. A., Atkinson, T., Lowe, C. R., Small, D. A., *Anal. Biochem.*, **1986**, 152, 215.
21. Hines, R., Cornor, K., *Biotechniques*, **1992**, 12, 430.
22. Giovannini, R., Freitag, R., Tennikova, T. B., *Anal. Chem.*, **1998**, 70, 3348.
23. Diogo, M.M., Queiroz, J.A., Monteiro, G.A., Martins, S.A., Ferreira, G. N., *Biotech. & Bioengineering*, **2000**, 68, 576.
24. Eun, H. M., *Enzymology Primer for Recombinant DNA Technology*, Academic Press, **1996**, 407.
25. Kilger, C., Krings, M., Poinar, H., Paabo, S., *Biotechniques*, **1997**, 22, 412.
26. Chen, Q., Neville, C., MacKenzie, A., Korneluk, R.G., *Biotechniques*, **1996**, 21, 453.
27. Marra, M. A., Kucaba, T.A., Hillier, L.W., Waterstion, R.H., *Nucl. Acids Res.*, **1999**, 27, e37.
28. Cohen, A. S., Najarian, D. R., Karger, B. L., *J. Chromatogr.*, **1990**, 516(1), 49.

29. Salas-Solano, O, Carrilho, E; Kotler, L, Miller, A. W, Goetzinger, W., Sosic, Z., Karger, B. L., *Anal. Chem.*, **1998**, 70(19), 3996.
30. Wei, W., Yeung, E.S., *J. Chromatogr., B: Biomed. Sci. Appl.*, **2000**, 745(1), 221.
31. Swerdlow, H., Dew-Jager, K. E., Brady, K., Grey, R., Dovichi, N., Gesteland, R. , *Electrophoresis*, **1992**, 13, 475.
32. Ruiz-Martinez, M.C., Salas-Solano, O., Carrilho, E., Kotler, L., Karger, B. L., *Anal. Chem.*, **1998**, 70, 1516.
33. Salas-Solano, O., Ruiz-Martinez, M., Carrilho, E., Kotler, L., Karger, B. L., *Anal. Chem.*, **1998**, 70, 1528.
34. Zhou, H., Miller, A. W., Sosic, Z., Buchholz, B., Barron, A. E., Kotler, L., Karger, B. L., *Anal. Chem.*, **2000**, 72, 1045.
35. Kawamoto, K., Okano, O., Kambara, H., *DNA res.*, **1994**, 1, 297.
36. Xiong, Y., Park, S. R., Swerdlow, H., *Anal. Chem.*, **1998**, 70, 3605.
37. A. Rolf, I. Schuller, U. Finckh and I. Wefer-Rolfs, *PCR: Clinical Diagnostics and Research*, **1992**, Springer-Verlag, Berlin
38. Panaccio, M., Lew, A., *Nucl. Acids Res.* **1991**, 19, 1151.
39. Panaccio, M., Lew, A., *PCR Technology: Current Innovations*, **1994**, CRC Press Inc., 151.
40. Janco, R.L., Word, B., Maness, K., Rheinboldt, M.C. and Phillips, J.A., *Am. J. Hum. Genet.*, **1989**, 45, A198.
41. Thomson, D.M., Brown, N.N., Clague, A.E., *Clinica Chimica Acta*, **1992**, 207, 169
42. Lench, N., Stanier, P. and Williams, R., *Lancet*, **1988**, 1356.

43. Richards, B., Skoletsky, J., Shuber, A., Balfour, R., Stern, R., Dorkin, H., Parad, R., Witt, D and Klinger, K., *Human Molecular Genetics*, **1993**, 2, 159.
44. Liu, Y., Bai, J., Zhu, Y., Liang, X., Siemieniak, D., Venta, P., and Lubman, D., *Rapid Commun. Mass Spectrom.* **1995**, 9(9), 735.
45. Drossman, H., Luckey, J.A., Kostichka, A.J., Dcunha, J., Smith, L.M., *Anal. Chem.*, **1990**, 62, 900.
46. Wilson, R.K., Yuan, A.S., Clark, S.M., Spence, C. Arakelian, P., Hood, L.E., *Biotechniques*, **1988**, 6, 776.
47. Zimmermann, J., Voss, H., Schwager, C., Stegemann, J., Angsorge, W., *FEBS Lett.*, **1988**, 233, 432.
48. Tan, H., Yeung, E. S., *Anal. Chem.*, **1997**, 69, 664.
49. Tan, H., Yeung, E. S., *Anal. Chem.*, **1998**, 70, 4044.
50. Zhang, N., Tan, H., E. S. Yeung, *Anal. Chem.*, **1999**, 71,1138.
51. Swerdlow, H., Jones, B. J., Wittwer, F. T., *Anal. Chem.*, **1997**, 69, 848.
52. Soper, S.A., Williams, D. C., Xu, Y., Lassiter, S.J., Zhang, Y., Ford, S.M., Bruch, R.C., *Anal. Chem.*, **1998**, 70, 4036.
53. Burns, M. A., Johnson, B. N., Brahmasandra, S. N., Handique, K., Webster, J. R., Krishanan, M., Sammarco, T. S., Man, P. M. , Jones, D., Heldsinger, D., Mastrangelo, C. H., Burke, D. T., *Science*, **1998**, 282, 484.
54. Litborn, E., Emmer, A., Roeraade, J., *Anal. Chim. Acta*, **1999**, 401, 11.
55. Litborn, E., Emmer, A., Roeraade, J., *Electrophoresis*, **2000**, 21, 91.
56. Soper, S. A., Ford, S. M., Xu, Y., Qi, S., McWhorter, S., Lassiter, S., Patterson, D., Bruch, R.C., *J. Chromatogr., A*, **1999**, 853(1 + 2), 107.

57. Khandurina, J., McKnight, T.E., Jascobson, S.C., Waters, L.C., Foote, R.S., Ramsey, J. M., *Anal. Chem.*, **2000**, 72, 2995.
58. William, E., PE Biosystem *HPCE'99* , 40.
59. Bashkin, J. et al. Molecular Dynamics *HPCE'99*, 39.
60. Pentoney, S., Lew, C. Rakestraw, D, Yang, D., *HPCE'99*, 114.
61. Li, Q., Kane, T., Liu, C., Zhao, H., Fields, R., Kernan, J., *HPCE'99*, 40.
62. Anazawa, T., Takahashi, S., Kambara, H, *Electrophoresis*, **1999**, 20, 539.
63. Quesada, M., Zhang, S., *Electrophoresis*, **1996**, 17, 1841.
64. Quesada, M., Dhadwal, H., Fisk, D., Studier, F., *Electrophoresis*, **1998**, 19, 1415.
65. Behr, S., Matzig, M., Levin, A., Eickhoff, H., Heller, C., *Electrophoresis*, **1999**, 20, 1492.
66. Dovichi, N.J., *Electrophoresis*, **1997**, 18, 2393.
67. Li, Q., Kane, T. E., Liu, C., Zhao H. et al., *Human Genome Program Contractor-Grantee Workshop VIII* ,US-DOE, 21.
68. Scherer, J.R., Kheterpal, I., Radhakris, A., Ja, W.W., Mathies, R.A., *Electrophoresis*, **1999**, 20, 1508.
69. Kheterpal, I., Mathies, R. A. , *Anal. Chem.*, **1999**, 71, 31A.
70. Irie, T., Oshida, T., Hasegawa, H., Matsuoka, Y., Li, T., Oya, Y., Tanaka, T., Tsujimoto, G., Kambara, H., *Electrophoresis*, **2000**, 21, 367.
71. Karger, B. L., Kotler, L., Foret, F., Minarik, M., Kleparnik, K. *PCT Int. Appl.* **1999**, 40 pp.
72. Simpson, P.C., Roach, D., Woolley, A. T., Thorsen, T., Johnston, R., Sensabaugh, G. F., Mathies, R. A., *Proc. Natl. Acad. Sci. U. S. A.*, **1998**, 95, 2256.

73. Medintz, I., Wong, W., Sensabaugh, G., Mathies, R. A., *Electrophoresis*, **2000**, 21(12), 2352.
74. Liu, S., Shi, Y., Ja, W.W., Mathies, R. A., *Anal. Chem.*, **1999**, 71, 566.
75. Liu, S., Ren, H., Gao, Q., Roach, D., Loder, R., armstrong, T.M., Mao, Q, Blaga, L., Barker, D., Jovanovich, S., *Proc. Natl. Acad. Sci. U.S. A.*, **2000**, 97, 5369.
76. Maier, W., *Angew. Chem. Int. Ed.*, **1999**, 38, 1216.
77. Ganesan, A., *Angew. Chem. Int. Ed.*, **1999**, 37, 2828.
78. Taylor, S. J, Morken, J. P., *Science*, **1998**, 280, 267.
79. Senkan, S. M., *Nature*, **1998**, 394, 350.
80. Zubritshy, E., *Anal. Chem.*, **1999**, 71, 39A.
81. Lavastre, O., Morken, J. P., *Angew. Chem. Int. Ed.*, **1999**, 38, 3163.
82. Reetz, M.T., Zonta, A., Schimossek, K., Liebeton, K., Jaeger, K., *Angew. Chem. Int. Ed*, **1997**, 36, 2830.
83. Siuzdak, G., Lewis, J.K., *Biotechnology and Bioengineering*, **1998**, 61, 127.
84. Wang, T., Zeng, L., Strader, T., Burton, L. Kassel, D.B., *Rapid. Commun. Mass Spectrom.*, **1998**, 12, 1123.
85. Cong, P, Doolen, R.D., Fan, Q., Giaquinta, D.M., Guan, S., McFarland, E. W., Poojary, D.M., Self, K., Turner, H., Weinberg. W. H., *Angew. Chem. Int. Ed.*, **1999**, 483.
86. Senkan, S., Krantz, K., Oztruk, S., Zengin, V., Onal, I. *Angew. Chem. Int. Ed.*, **1999**, 38, 2794.
87. Roche, M., Oda, R., Machacek, D., Lawson, G., and Landers, J., *Anal. Chem.*, **1997**, 69, 99.

88. Woodbury, C., Fitzloff, J., and Vincent, S., *Anal. Chem.*, **1995**, 67, 885.
89. Zeng, L., and Kassel, D., *Anal. Chem.*, **1998**, 70, 4380.
90. Yeung, E. S., *Anal. Chem.*, **1999**, 71, 522A.

CHAPTER 2. MULTIPLEXED AUTOMATED DNA SEQUENCING DIRECTLY FROM SINGLE BACTERIAL COLONIES

A paper published in *Analytical Chemistry*¹

Yonghua Zhang, Hongdong Tan and Edward S. Yeung

Abstract

Sample preparation has been one of the major bottlenecks for large-scale DNA sequencing projects in terms of time and cost. To improve sample throughput and to integrate the front-end tasks to capillary-array DNA sequencers, protocols for directly sequencing the plasmids from a single bacterial colony in fused-silica capillaries were developed. After the colony is picked, lysis is accomplished *in situ* in the plastic sample tube using either a thermocycler or a heating block. Upon heating, the plasmids are released while chromosomal DNA and membrane proteins were denatured and precipitate to the bottom of the tube. After adding enzyme and Sanger reagents, the resulting solution was aspirated into the reaction capillaries by a syringe pump and cycle sequencing was initiated. No deleterious effect was observed upon the reaction efficiency, the on-line purification system or the capillary electrophoresis separation even though the crude lysate was used as the template.

¹ Reprinted with permission from *Analytical Chemistry* 1999, 71, 5018

Multiplexed on-line DNA sequencing data from 8 parallel channels allowed base calling up to 620 bp with an accuracy of 98%. The entire system can be automatically regenerated for repeated operation. By the marriage of colony sequencing with the capillary array sequencer, both the front end and the back end of DNA sequencing are combined in a miniaturized format. This protocol will ultimately reduce the cost of sequencing to well below current levels.

Introduction

DNA sequencing is a multistage process that includes the preparation and manipulation of clone libraries, purification of the DNA, production and separation of the sequencing ladder, base analyses and interpretation of the resulting sequence information.¹ In order to achieve a high throughput with minimum cost and high accuracy, all of the steps involved in this process must be coordinated, and where possible integrated, automated and miniaturized.² Since the inception of the Human Genome Project, the individual steps of DNA sequencing technology have improved dramatically. These include the development of bacterial artificial chromosomes (BACs) for longer insert clones,³ high-throughput flow-through microcentrifuge methods for plasmid preparation,⁴ new thermostable enzymes,⁵ high sensitivity dye sets,⁶ and highly multiplexed high-speed capillary array electrophoresis (CAE) instruments.⁷ In fact, with current technology and hefty investment it should be possible to finish the Human Genome Project well ahead of 2005, the target completion date.⁸

The small diameter of capillaries coupled with ultrasensitive detection associated with laser-induced fluorescence in CAE can reduce the sample requirement from 1-10 fmol of DNA in a given band on the slab gel to 1-10 amol in capillary gel.⁹ This means we can

reduce the reaction volume from 20 μ l to 20 nl if the reaction efficiency is maintained and an efficient sample loading method can be found.¹⁰ However, present sequencing technology still leaves much to be improved if the potential for genomic science is to be fully realized. The front end of the sequencing process, especially sample preparation, which is typically labor intensive, repetitive, and time consuming, needs to interface smoothly with the back end.¹¹

There are many procedures developed over the years for the isolation of bacterial plasmids.¹² A lot of proprietary methods also give satisfactory results. One thing in common for nearly all of them is that they involved three basic steps: growth of bacteria, harvesting and lysis of the cells, and purification of the plasmid.¹³ All of them include centrifugation which is labor intensive and difficult to automate and interface with CAE. Some chromatographic columns (size exclusion,¹⁴ ion-exchange,¹⁵ high-performance membranes¹⁶) have also been developed to avoid the use of centrifugation. Another complication is that reagents used in purification of plasmids, such as ethanol and SDS,¹⁷ might become interferences to the subsequent cycle-sequencing reaction.

Alternative methods also exist which require only heat-induced lysis of cells in bacterial colonies.¹⁸ Centrifugation^{18a,d} or vigorous vortexing^{18a,c} are still needed to isolate cell debris. The resulting lysate was then used as the template in cycle sequencing using labeled primers. Even though the performance and the ruggedness demonstrated so far by this method is still inferior to that of the standard protocol,^{18c,d} it shows promise for significant savings in time and cost, particularly for large-scale sequencing. For routine sequencing applications, improvements are needed in terms of sequencing length, reproducibility, automation, adaptation to labeled terminators and elimination of the

centrifugation and vortex steps.

Recently, our group has demonstrated an automated, integrated, multiplexed on-line DNA sequencer which can start from purified templates to called bases with minimal human intervention.¹⁹ The microfluidic management was controlled by two sets of multiplexed freeze/thaw valves (MFTV).²⁰ The integrated sequencer also has the potential for miniaturization from μl to nl scale compared with the current automation protocols based on robotics.²¹ μTAS on chips, which can incorporate the operations from cell lysis to sequencing separation is also a promising approach for full integration.²² But practical obstacles such as sample introduction and purification still need to be overcome before multiplexed total integration for sequencing (as opposed to PCR) on chips can be demonstrated.²³

In this study, we will show that the implementation of colony sequencing in our multiplexed on-line reaction-purification-sequencing instrument is feasible. Bases can be called directly from a single colony without compromise in spite of the complex matrix of the cell lysate and the small size of the sample. This completes the long sought after goal of integration and automation of the front and back ends of sequencing in one machine.

Experimental section

The experimental set-up is similar to that described previously.¹⁹ *E. coli* DH5 α competent cells and plasmid pUC19 (Clontech, Inc., Palo Alto, CA), Plasmid Bluescript (Stratagene, Madison, WI) and pGEM (Promega, La Jolla, CA) were used. Cell transformation was according to the protocol provided by Clontech. Transformed cells were plated on 2 \times TY agar plates (Difco Laboratories, Detroit, MI) and incubated at 37 °C for 24 h. Transformed cells with different inserts from Dr. Johansen's group at Iowa State University were also used for testing. A colony was picked up by a wire loop and put into a vial with 15

μl autoclaved D.I. water inside. After heating at 96 °C for 11 min, sequencing premix (8 μl), BSA (2 μl, 2.5 mg/ml) (Idaho Technology, Idaho Falls, ID) and 3.2 pmol primer were added. 20 μl of the above solution was aspirated into the reaction capillary. Cycle sequencing began immediately by holding at 96 °C for 3 min, followed by 40 cycles (96 °C for 10 s, 50 °C for 30 s and 60 °C for 4 min or 2.5 min). Three different sets of dyes, Rhodamine, dRhodamine and Energy Transfer (Applied Biosystems, Foster City, CA) and two polymerases, AmpliTaq FS from ABI and ThermoSequenase from Amersham Life Science (Cleveland, OH) were used in the reaction. Performance in all cases are comparable. For the results presented here, dRhodamine terminators and AmpliTaq FS were used. After reaction the sequencing ladders were purified by homemade size-exclusion columns and injected on-line into the separation capillary, or purified by spin columns (Princeton Separations, Adelphia, NJ) and injected off-line for comparison.

Results and discussion

Comparison of different methods for cell lysing

The amount and concentration of DNA used for the sequencing reaction is critical. Amounts of pure DNA in solution can be determined by absorption measured at 260 nm. For the lysate, we found it is also possible to roughly estimate the amount of plasmid DNA by measuring absorption even in the presence of cell matrix by the ratio of absorption at 260 nm and 280 nm. However, since extremely precise determination was not the primary goal, there are other approaches to estimate the DNA amount quickly.^{18a} For a colony (1 mm diameter) developed after 24 h, we found by knowing the plasmid copy number in each cell (in our case, the cells generally have 100-500 copies) and the numbers of cells in a colony (10^7 to 10^8) that the amount of plasmid is around 100 ng. This is at the lower end of the plasmid

amount the manufacturer recommends for cycle sequencing. No further controls except for the colony size was needed in subsequent experiments to produce consistent results.

In an effort to increase sequencing efficiency and data quality we tested different methods for cell lysis. There are numerous methods to lyse cells including the use of detergents with or without heat.²⁴ DMSO used as the solvent,²⁵ glass-bead stirring,²⁶ sonication,²⁷ boiling in distilled water¹⁸ or PCR buffer,²⁸ proteases such as proteinase K (which must be inactivated before the sequencing reagent is added),^{18d} and NaOH with heat.²⁹ The best lysing protocol should break the cell wall and allow the plasmid to be easily accessed by the polymerase. Cell debris should be minimized. Another consideration is that lysing should be easily implemented and amenable to automation. Heat lysing is the obvious choice because of its simplicity and ability to deactivate some interfering materials such as RNA and proteins. In commercial DNA extraction kits, RNase treatment sometimes is used to degrade RNA that interferes with the reaction through nonspecific priming. This may be addressed by long annealing times to insure the correctness of annealing. It has been shown that the presence of small amounts of RNA in miniprep samples is not inhibitory to the reaction.³⁰ Heating can also cause some degree of RNA degradation³¹ and reduce the negative effect of RNA. Proteins can be denatured and inactivated by heating. The whole sequence of *E. coli* is already known.³² By choosing the appropriate primer, interference will be minimal. In our lysing protocol, only distilled water was used as the solvent. No detergents were added to assist in breaking the cell membrane because nonionic detergents such as NP-40, Tween-20 and Triton X-100 can stimulate polymerase activity and ionic detergents such as SDS are inhibitory to the enzyme at very low concentrations.¹⁷

Three slightly different heating protocols were studied to generate the lysate. One is

to put the colony in water and heat without vortex. The cell membrane proteins and chromosomes will denature and eventually settle on the bottom of the vial without centrifugation. Then, the supernatant (relatively clear) can be transferred for use in cycle sequencing. In this way, some cell debris will be kept in the lysing vial and does not enter the reaction capillary. When the lysate produced in this manner was used in on-line cycle sequencing the signal strength is similar to that using purified templates (Fig. 1A). The second protocol involves adding the reaction mixture to the lysing vial immediately after heating to eliminate the extra transfer step. It is interesting to note that cell debris can be allowed to enter the reaction capillary without seriously affecting the efficiency of the reaction (Fig. 1B). The signal strength of this protocol is generally one half of the first one but is still sufficient for sequencing. Subsequent multiplexed experiments were mostly performed in this way. In the third protocol, the combination of heat lysis and cycle sequencing in one step is attempted. It has been shown that for PCR, bacterial cells³³ and even blood cells³⁴ can be lysed in the reaction mixture. By adding the colony into the reaction mixture and lysing by the thermocycler and then cycling, we can get some signal but the efficiency is very low (Fig. 1C). It is generally thought that as long as the impurities associated with the target be adequately diluted, the purity of the DNA sample subjected to PCR amplification need not to be high.³⁵ This is supported by studies on single-cell PCR³⁶ and on-line PCR using blood.³⁷

Three features associated with the combined lysis/reaction protocol may account for the difference in efficiency. First, cells in the colony are still viable. The enzyme they produced if not deactivated may interfere with the polymerase. In PCR reactions, the cells can be lysed in the reaction buffer because the exponential amplification effect can

compensate for the loss caused by this kind of interference. This is not the case for the linear amplification character of cycle sequencing. Second, we suspect that the small inner diameter of the capillary reactor prevents uniform distribution of the plasmids. Cells in colonies tend to stick together even after being drawn into the capillary. After heating, the released plasmids localize in a confined region of the long capillary despite 10 min heating at 96 °C. Convection here is small, so the molecular movement is dominated by diffusion. For a diffusion-controlled process, mixing may be taken as nearly complete for Dt/l^2 from 0.1 to 1.³⁸ The diffusion constant for rhodamine-dGTP in water at room temperature is $D = 3-5 \times 10^{-6} \text{ cm}^2/\text{s}$.³⁹ Accounting for the increase in diffusion caused by temperature and viscosity at 96 °C,⁴⁰ the effective mixing distance in $t = 10 \text{ min}$ even for small molecules is thus less than $l = 1 \text{ cm}$. The aspiration speed used to load the mixture to the reactor was 2.4 cm/s which corresponds to a Reynold number of roughly 12. The flow will therefore be laminar and no extra mixing exists.³⁵ The third reason may relate to the thermal stability of the enzyme. Taq polymerase is thermostable with a half-life of 40 min at 95 °C and 10 min at 97.5 °C respectively.¹⁸ The extra heating time for lysis may cause the enzyme to lose some activity.

Further improvements on the lysing and cycling conditions mentioned above may be possible. Another method that has promise is the use of NaOH to lyse the cell and to use HCl or reaction buffer to neutralize the resulting lysate. This method may allow the lysis of cells in a high-density microtiter array without heating and can be easily automated. However, the problem in using this method was the high viscosity caused by chromosomal DNA, uncoiling from a compact form (nucleoid) into long strands, and by the saponification of the lipids. The resulting lysate is difficult to pipette.

Adjusting the parameters of cycle sequencing was vital to success. Compared to the standard recommended protocol, we use somewhat longer annealing times (5 s to 30 s) to assure the correctness and completeness of annealing between the primer and the plasmid to minimize interference from chromosomal DNA and RNA. The total signal strength may also benefit from this since there should be an increase in the amount of primed template.⁴¹ The signals among the 4 bases were balanced so no adjustment of ratio of ddNTP/dNTP was necessary. The sequencing reaction kinetic does not appear to be affected by the complex matrix. Reducing the extension time from 4 min to 2.5 min (Fig. 1D) was shown to be possible. Extension was also tested at temperatures higher than 60 °C, where Taq enzyme is supposed to have a faster polymerization rate. Similar results were obtained.

SEC purification and column regeneration

The capability for parallel and automatic processing is the requirement for any protocol used in DNA sequencing in order to achieve high throughput. We already demonstrated that pure DNA can be used as the template in the array on-line sequencer.¹⁹ By switching to lysates, which are complex mixtures, one concern is that the cell matrix affects the flow of the multiple size-exclusion columns. Fortunately, we found that the cell matrix does not change the uniform elution profile of the SEC columns, which is vital to synchronous injection. This may be due to the fact that temperature cycling helped to dissolve the cell debris. Another complication in using the lysate as the template is the regeneration of the size-exclusion columns. Some particulates especially cell membrane proteins may clog the column on repeated usage if they are adsorbed and settled in the column. We find that washing with 0.2 M NaOH is the solution. This is confirmed by the observation that in standard microvials, by heating the lysate mixture in 0.2 M NaOH at 50

°C, the mixture shows no precipitation even under centrifugation. The Sephadex beads used in the packing are still quite stable under these conditions.⁴² We tested the uniformity of the system after three months of daily operation using the crude lysate. Nearly identical signal-to-noise ratios and elution times were found (Fig. 2), showing that no harmful pressure buildup occurred in the individual channels.

Injection protocol and separation condition

On-line sequencing from lysate is a multistage process, in which each step should be considered in the context of the whole process to achieve the desired performance. A slight decrease in reaction efficiency can be tolerated if better stacking and separation can be achieved. Many stacking methods exist for on-column concentration of samples. They are field-amplification injection,⁴² base stacking,⁴³ sweeping in MEKC⁴⁴ and on-line hollow fiber extraction.⁴⁵ The easiest one for us to implement is to inject a water plug before the sample plug.⁴² However, the results were similar even without this step. This may be due to the use of 1×TE, which has 1/10 the ionic strength of the gel buffer and is already a good buffer for sample stacking. This is confirmed by the tolerance to long injection times (3 min) at the running voltage. Another parameter that needs to be considered is sample denaturation. Denaturation by heating is very efficient and has led to our implementation of hot injection.¹⁹ The PEEK cross was wrapped by heating tape to insure that the injection temperature is above 70 °C. However, the crudeness of wrapping of the heating tape caused unevenness in heating, which occasionally led to bubble formation thereby ruining the separation. More uniform heating is possible with better engineering of the cross. An alternative is to denature the sample at the reactor or somewhere along the transfer line.⁴⁶ The kinetics for denaturation at low ion strength is fast and it has been shown that highly purified DNA ladders can be

injected even without denaturation.⁴⁷ Fig. 3 shows the comparison of hot injection and room-temperature injection. The raw data produced by on-line colony sequencing is displayed around 440 bp. Slightly worse results were obtained without heating. This may be due to the delay time (2 min) needed to thaw the two liquid-nitrogen valves. In order to increase the denaturing power of gel, we increased the urea concentration from 3.5 M to 7 M. Although there is a 20 min increase in separation time, the 7 M urea gel shows better resolution (data not shown).

The lysate in *E. coli* is a very complicated matrix which contains salts, proteins and other biomolecules.⁴⁸ Surprisingly, we did not find any extra problems in injection. Presumably these are greatly diluted during the elution process. The Sephadex beads can retain some of the biomolecules such as degraded RNA, lipids, polyamine and lipopolysaccharide, as long as the molecular weight is less than 5000 daltons. There are still some high molecular weight molecules which may elute with the DNA ladder. In fact, the carrier protein BSA (10^{-11} mole) that is added to the reagents may co-elute also. Each *E. coli* cell may contain 2,350,000 protein molecules so one colony with 10^7 to 10^8 cells will result in another 10^{-11} mole protein being added to the reaction mixture. As with our earlier experiments, we did not observe any deleterious effect related to injection because of BSA. For BSA the isoelectric point is 4.7 and the MW is 67,000, which is roughly the MW of 200-bp ss-DNA. But the net charge per molecule at pH 7.4 is only -17, much smaller than a 200-bp ss-DNA. This results in a mobility of BSA that is at least one order of magnitude smaller than DNA.⁴⁹ More importantly, by long-time heating, a high fraction of BSA will be polymerized and aggregated and are thus further discriminated against in electrokinetic injection.⁵⁰ We also note that the observed interference of BSA to DNA separation is buffer

dependent.⁵¹ Urea used in the sequencing buffer may also help to reduce the interaction between the protein and DNA.

Performance of integrated system

Fig. 4 is the reconstructed image plot of the sequencing results from 6 experiments in a one-week period. The intensity was not uniform because different DNA samples were used. Generally, the intensities from purified DNA (40% of the channels) are two times higher than the intensities from lysate. But even the raw data starting directly from the lysate provides adequate signal-to-noise ratios for base calling. This is also evident in Fig. 3. Two clones, PGEM and M13mp18, were used in these experiments. The former has known problems with compressions, but the use of 7 M urea here provided excellent separation even at room temperature. Among all 48 runs, only two runs did not produce a useful sequence due to bubble formation at the cross and one run produced only a short sequence (150 bp). All three appeared in the same channel. These intensities are however still reasonable, indicating that injection rather than reaction was the culprit. Two runs provided sequence around 350 bp, and the electropherogram of all others provided resolution greater than 0.5 well beyond 430 bp.

The turn-around time for each run is 0.2 h for lysate preparation, 3.5 h for reaction, 1 h for regeneration of the system and 2.5 h for separation. With staggered operation of these individual steps analogous to an assembly line, the turn-around time will be essentially the reaction time, which can be reduced to 2.5 h even with lysates as the template (Fig. 1D). During these experiments, the reactors and SEC columns were regenerated by washing in-between runs with 1×TE and water. No memory effect was observed in the different channels. Fig. 5 shows the base-calling results (raw data, day 6, bottom channel) up to 620 bp

with 98% accuracy. The resolution after 620 bp is still above 0.5. Base calling was accomplished by software written in Labview using the 2-color scheme developed before.⁵² Longer reads can be expected with more sophisticated algorithms. Loss in the front part of the sequence may be due to the long injection time in diluted buffer. This can be solved by a better design of the cross assembly (to fully sweep the volume inside the cross) to allow more efficient injection. This is not a problem in shotgun sequencing because for the EcoR.I and the universal primer, which are widely used, the vector sequence is around 100 bp. For primer walking such a limitation will need to be considered.

Conclusions

Sequencing directly from a single colony in a closed, multiplexed and automated on-line capillary array instrument was demonstrated for the first time. Dealing with single colonies eliminates the need for overnight incubation in present approaches and potentially allows the use of the reaction capillary itself for picking clones. The optimized protocol was identified through systematic studies. This includes heat-induced lysis (to denature the cell proteins and chromosome, partially degrade the RNA and release the plasmid followed by a brief incubation period in place of centrifugation), and redesigned cycle conditions that are compatible with the crude lysate. Such colony sequencing can be coupled to the on-line CAE system and provide base calling up to 620 bp with 98% accuracy even with the use of a simple algorithm. There are no elution differences in the purification columns when using cells in colonies or using purified DNA as the template. This makes synchronous injection possible. The nearly identical S/N ratios among channels (using identical samples) after three months of operation with the crude lysate indicate that there is no adverse buildup among the SEC columns and that the columns can be regenerated repeatedly. The overall performance

from continuous operation in a one week period using lysates as templates shows similar reliability as the off-line CAE sequencer. DNA sequencing from lysates to bases in one machine with minimal human intervention is therefore achieved.

A major advantage of the present system is the potential for further multiplexing and miniaturization. Multiplexing 100 channels using freeze-thaw valves can be manageable even in the present assembly. By reducing the inner diameter yet keeping the same reagent concentrations, even smaller volume samples could be injected using the current protocol. Our instrumentation therefore promises to further reduce the reagent cost and labor requirement in high-throughput DNA sequencing.

The above automated, multiplexed flow management can also benefit other high-throughput analytical situations, e.g. proteomics and combinatorial chemistry. Parallel and automated analyses^{53,54} is the main approach, if not the only approach, to process large numbers of samples in a timely manner. With the features of low dead volume, adjustable injection volume and reliable microfluidics, we can envision coupling this system with other hyphenated techniques to speed up drug discovery and quality control.

Acknowledgments

The authors thank Kristen Johansen, Anna Yeung, Wen-chy Chu and Gary Polking for providing the clones for sequencing. The Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under contract No. W-7405-Eng-82. This work was supported by the Director of Science, Office of Biological and Environmental Research, and by the National Institutes of Health.

References

1. (a) Hunkapiller, T.; Kaiser, R. J.; Koop, B. F.; Hood, L. *Science* **1991**, 254, 59-67.
(b) Trainor, G. L. *Anal. Chem.* **1990**, 62, 418-426.
2. Collins, F.; Patrinos, A.; Jordan, E.; Chakravarti, A.; Gesteland, R.; Walters, L. *Science* **1998**, 282, 682-689.
3. Shizuya, H.; Birren, B.; Kim, U.; Mancino, V.; Slepak, T.; Tachiiri, Y.; Simon, M. *Proc. Natl. Acad. Sci. USA* **1992**, 89, 8794-8797.
4. Marziali, A.; Willis, T. D.; Davis, R. W. *Proc. Natl. Acad. Sci. USA* **1999**, 96, 61-66.
5. (a) Tabor, S.; Richardson, C. C. *Proc. Natl. Acad. Sci. USA* **1995**, 92, 6339-6343.
(b) Reeve, M.; Fuller, C. W. *Nature* **1995**, 376, 796-797.
6. (a) Rosenblum, B. B.; Lee, L. G.; Spurgeon, S. L.; Khan, S. H.; Menchen, S. M.; Heiner, C. R.; Chen, S. M. *Nucleic Acids Res.* **1997**, 25, 4500-4504.
(b) Ju, J.; Ruan, C.; Fuller, C. W.; Glazer, A. N.; Mathies, R. A. *Proc. Natl. Acad. Sci. USA* **1995**, 92, 4347-4351.
(c) Lee, L. G.; Spurgeon, S. L.; Heiner, C. R.; Benson, S. C.; Rosenblum, B. B.; Menchen, S. M.; Graham, R. J.; Constantinescu, A.; Upadhy, K. G.; Cassel, J. M. *Nucleic Acids Res.* **1997**, 25, 2816-2822.
7. (a) Ueno, K.; Yeung, E. S. *Anal. Chem.* **1994**, 66, 1424-1431.
(b) Mathies, R. A.; Huang, X. C. *Nature* **1992**, 359, 167-169.
(c) Dovichi, N. J.; Swerdlow, H.; Zhang, J. Z.; Chen, D. Y.; Harke, H. R.; *Anal. Chem.* **1991**, 63, 2835-2841.
(d) Takahashi, S.; Murakami, K.; Anazawa, T.; Kambara, H. *Anal. Chem.* **1994**, 66, 1021-1026.

8. Waterston, R. Sulston, J. E. *Science* **1998**, 282, 53-54.
9. Drossman, H.; Luckey, J. A.; Kostichka, A. J.; D'Cunha, J.; Smith, L. M. *Anal. Chem.* **1990**, 62, 900-903.
10. Soper, S. A.; Williams, D. C.; Xu, Y.; Lassiter, S. J., Zhang, Y.; Ford, S. M.; Bruch, R. *C. Anal. Chem.* **1998**, 70, 4036-4043.
11. Venter, J. C.; Smith, H. O.; Hood, L. *Nature* **1996**, 381, 364-366.
12. (a) Wang, K.; Gan, L.; Boysen, C.; Hood, L. *Anal. Biochem.* **1995**, 226, 85-90.
(b) Ng, W.; Schummer, M.; Cirisano, F.; Baldwin, R. L.; Karlan, B. Y.; Hood, L. *Nucleic Acids Res.* **1996**, 24, 5045-5047.
(c) Alderton, R. P.; Eccleston, L. M.; Howe, R. P.; Christopher, A. R.; Reeve, M. A.; Beck, S. *Anal. Biochem.* **1992**, 201, 166-169.
(d) Andersoom, B.; Lu, J.; Edwards, K. E.; Muzny, D. M.; Gibbs, R. A. *Biotechniques* **1996**, 20, 1022-1027.
13. Mazza, P. G. *Microbiologica* **1986**, 9, 113-125.
14. (a) Edwardson, P. A.; Atkinson, T.; Lowe, C. R.; Small, D. A. *Anal. Biochem.* **1986**, 152, 215-220.
(b) Himmel, M. E.; Perna, P.J.; McDonell, M. W. *J. Chromatograph.* **1982**, 240, 155-163.
15. Hines, R.; Cornor, K.; *Biotechniques*, **1992**, 12, 430-434.
16. Giovannini, R.; Freitag, R.; Tennikova, T. B. *Anal. Chem.* **1998**, 70, 3348-3354.
17. Eun, H. M. *Enzymology Primer for Recombinant DNA Technology*, Academic Press, **1996**, 407.
18. (a) Krishnan, B. R.; Blakesley, R.W.; Berg, D. E. *Nuclei Acids Res.* **1991**, 19, 1153.

- (b) Hofmann, M.; Brian, D. A. *Biotechniques* **1991**, 11, 30-31.
- (c) Kilger, C.; Krings, M.; Poinar, H.; Paabo, S. *Biotechniques* **1997**, 22, 412-418.
- (d) Chen, Q.; Neville, C.; MacKenzie, A.; Korneluk, R.G. *Biotechniques* **1996**, 21, 453-457.
19. (a) Tan, H.; Yeung, E. S. *Anal. Chem.* **1997**, 69, 664-674.
- (b) Tan, H.; Yeung, E. S. *Anal. Chem.* **1998**, 70, 4044-4053.
20. Bevan, C. D.; Mutton, I. M. *Anal. Chem.* **1995**, 67, 1470-1473
21. <http://www.bio.net/hypermail/automated-sequencing/9704/0035.html>
22. (a) Woolley, A. T.; Hadley, D.; Landre, P.; J. deMello, A.; Mathies, R. A.; Northrup, M. A. *Anal. Chem.* **1996**, 68, 4081-4086.
- (b) Burns, M. A.; Johnson, B. N.; Brahmasandra, S. N.; Handique, K.; Webster, J. R.; Krishanan, M.; Sammarco, T. S.; Man, P. M.; Jones, D.; Heldsinger, D.; Mastrangelo, C. H.; Burke, D. T. *Science* **1998**, 282, 484-487.
23. (a) Mayer, G.; Tuchscheerer, J.; Laiser, T.; Wohlfart, K. *Microreaction Technology—Proceedings of the First International Conference on Microreaction Technology*, 112-119.
- (b) Litborn, E.; Roeraade, J. *HPCE* 99, 177.
24. Hutchison, W.; Dahl, H. *Biotechniques* **1995**, 19, 554-556.
25. Huynh, N. V.; Backer, O. D.; Declaire, M. Colson, C. *Anal. Biochem.* **1989**, 126, 464-467.
26. Song, D. D.; Jacques, N. A. *Anal. Biochem.* **1997**, 248, 300-301.
27. Buck, C. E.; O'Hara, L. C. L.; Summersgill, J. T. *J. Clin. Microbiol.* **1992**, 30, 1331-1334.

28. Woodford, K.; Usdin, K. *Nucleic Acids Res.* **1991**, 19, 6652.
29. Yie, Y.; Wei, Z.; Tien, P. *Nucleic Acids Res.* **1993**, 21, 361.
30. Yamamoto, T.; Horikoshi, M. *Nucleic Acid Res.* **1995**, 23, 3351-3352.
31. Sheibani, N.; Frazier, W. A. *Anal. Biochem.* **1997**, 250, 117-119.
32. Blattner, R. R.; Plunkerrt, G.; Bloch, C. A.; Perna, N. T.; Burland, U.; Riley, M.; Lollado-Vides, J.; Glasner, J. D.; Kode, F. K.; Mayhew, G. F.; Gregor, J.; Davas, N. W.; Kirkpatrick, H. A.; Goeden, M. A.; Rose, O. J.; Mau, B.; Shao, Y.; *Science* **1997**, 277, 1453-1462.
33. Waters, L. C.; Jacobson, S. C.; Kroutchinina, N.; Khandurina, J.; Foote, R. S.; Ramsey, J. M.; *Anal. Chem.*, **1998**, 70, 158-162.
34. Panaccio, M.; Lew, A. M. *PCR Technology: Current Innovations*, **1994**, CRC Press Inc., 150-157.
35. Kolmodin, L. A.; Williams, J. F.; *PCR Cloning Protocols*, **1997**, Humana Press Inc., 3-15.
36. Findlay, I.; Taylor, A.; Quirke, P.; Frazier, R.; Urguhart, A.; *Nature*, **1997**, 389, 555-556.
37. Zhang, N.; Tan, H.; Yeung, E. S. *Anal. Chem.* **1999**, 71, 1138-1145.
38. Robins, T.; Shaw, J.; Miller, B.; Tuner C.; *Microreaction Technology—Proceedings of the First International Conference on Microreaction Technology*, 35-46.
39. Xu, X. H.; Yeung, E. S.; *Science*, **1997**, 275, 1106-1109.
40. Kalinina, O.; Lebedeca, I.; Brown, J.; Silver, J. *Nucleic Acids. Res.* **1997**, 25, 1999-2004.

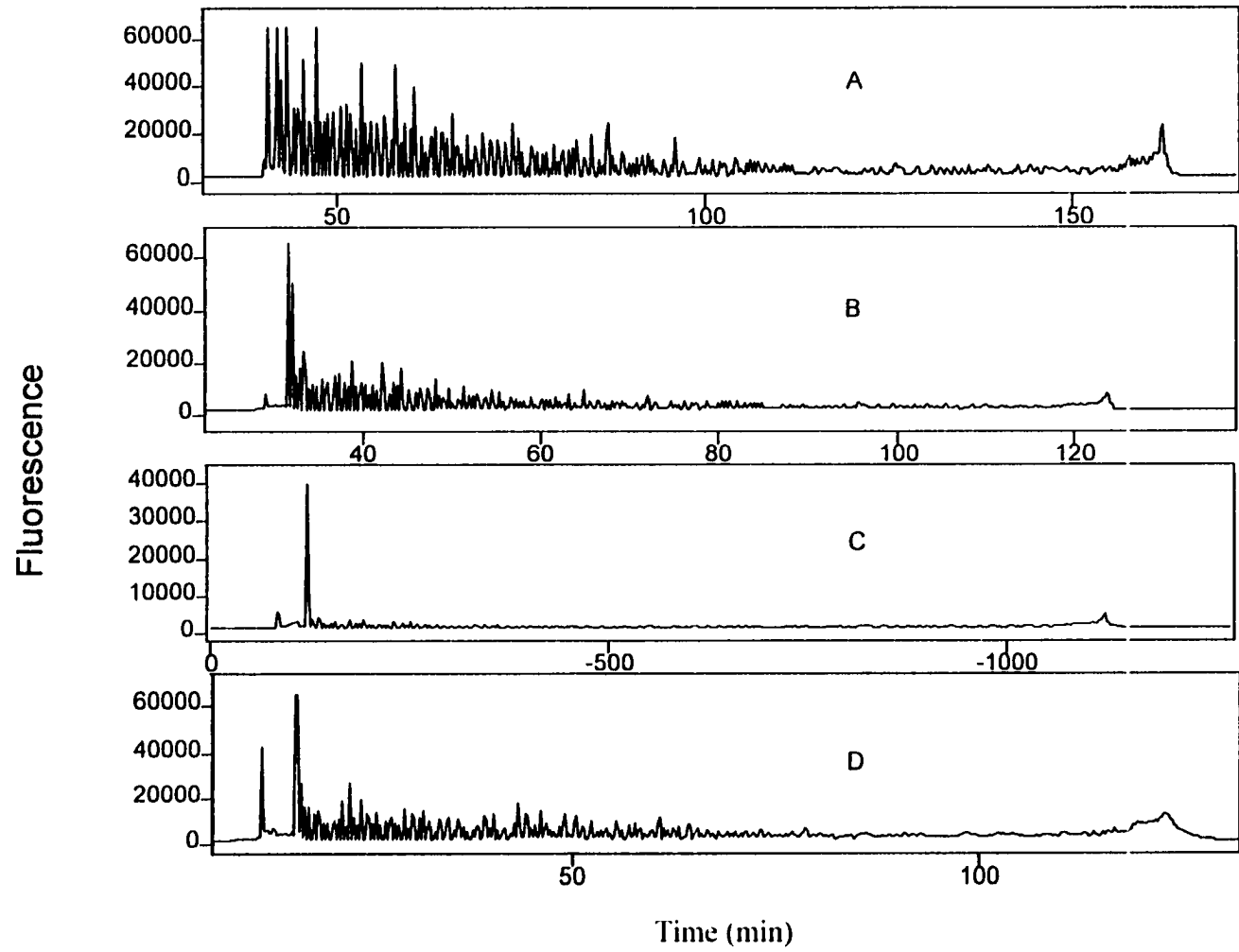
41. Lee, L. G.; Connell, C. R.; Woo, S. L.; Cheng, R. D.; McArdle, B. F.; Fuller, C. W.; Halloran, N. D.; Wilson, R. K. *Nuclei Acids Res.* **1992**, 20, 2471-2483.
42. Chien, R. L.; Burgi, D. S. *Anal. Chem.* **1992**, 489A-496A.
43. Xiong, Y.; Park, S. R.; Swerdlow, H. *Anal. Chem.*, **1998**, 70, 3605-3611.
44. Quirino, J. P.; Terabe, S. *Science*, **1998**, 282, 465-468.
45. Wu, X.; Hosaka, A.; Hobo, T. *Anal. Chem.* **1998**, 70, 2081-2084.
46. Swerdlow, H.; Jones, B. J.; Wittwer, F. T. *Anal. Chem.* **1997**, 69, 848-855.
47. Ruiz-Martinez, M. C.; Salas-Solano, O.; Carrilho, E.; Kotler, L.; Karger, B. L. *Anal. Chem.* **1998**, 70, 1516-1527.
48. Cooper, S. *Molecular Biology and Biotechnology, A Comprehensive Desk Reference*, VCH Publishers, Inc. **1995**, 57-65.
49. Perters, T. *All about Albumin, Biochemistry, Genetics and Medical Applications*, Academic Press. **1996**, 25.
50. Aoki, K.; Saro, K.; Nagaoka, S.; Kanada, M.; Hiramatsu, K. *Biochim. Biophys. Acta*, **1973**, 328, 323-333.
51. Landers, J. P.; Oda, R. P.; Spelsberg, T. C.; Nolan, J. A.; Ulfelder, K. J. *Biotechniques*, **1993**, 14, 98-115.
52. Li, Q.; Yeung, E. S. *Appl. Spectrosc.* 1995, **49**, 1528-1533.
53. Gang, X.; Pang, H.-M.; Yeung, E. S. *Anal. Chem.* **1999**, 71, in press.
54. Gong, X.; Yeung, E. S. *Anal. Chem.* **1999**, 71, in press.

Figure captions

- Figure 1. On-line colony sequencing results with different methods. A is the result from the clear lysate. B is the result using whole-cell lysate. C is the result of combined on-line lysing and reaction. D is the result using 2.5-min extension time in cycle sequencing of whole-cell lysate. All samples were prepared in fused-silica capillaries and on-line injected for 3 min at 160 V/cm.
- Figure 2. Electropherograms showing the uniformity of the array system after 3-month daily operation using the crude lysate. The 8 identical samples were separated on-line and injected into the gel capillaries. Nearly identical S/N was observed.
- Figure 3. Electropherograms showing the difference between on-line hot injection and room-temperature injection. The range shown is around 440 bp. A: hot injection (70 °C). B: room-temperature injection.
- Figure 4. Reconstructed DNA sequencing electropherograms from the on-line array sequencer. All experiments were done continually in one week. DNA used in each day's experiment are five lysates (two kinds of DNA cloned to pBluescript) and three pure DNA (PGEM or M13mp18 DNA). Capillaries (groups of 8) are ordered from top to bottom. Migration time is from left to right. SEC column regeneration during the run: 1×TE flushing. On-line injection: 3 min at 160 V/cm. Separation: 60-cm effective length of bare capillaries filled with PEO gel. Voltage: 160 V/cm. Buffer: 1×TBE with 7 M

urea. Laser: 15 mW 514-nm argon-ion laser. Detection: CCD camera.

Figure 5. Sequencing of a DNA insert in the pBluescript vector using the cell lysate from one colony as the template. The data used for base calling is the bottom channel of day 6 in Fig. 5. Primer: M13-40. The raw data from the blue and red channels are plotted. The miscalls are also corrected under the corresponding bases. The resolution is above 0.5 at 620 bp.



Time (min)

FIGURE 1

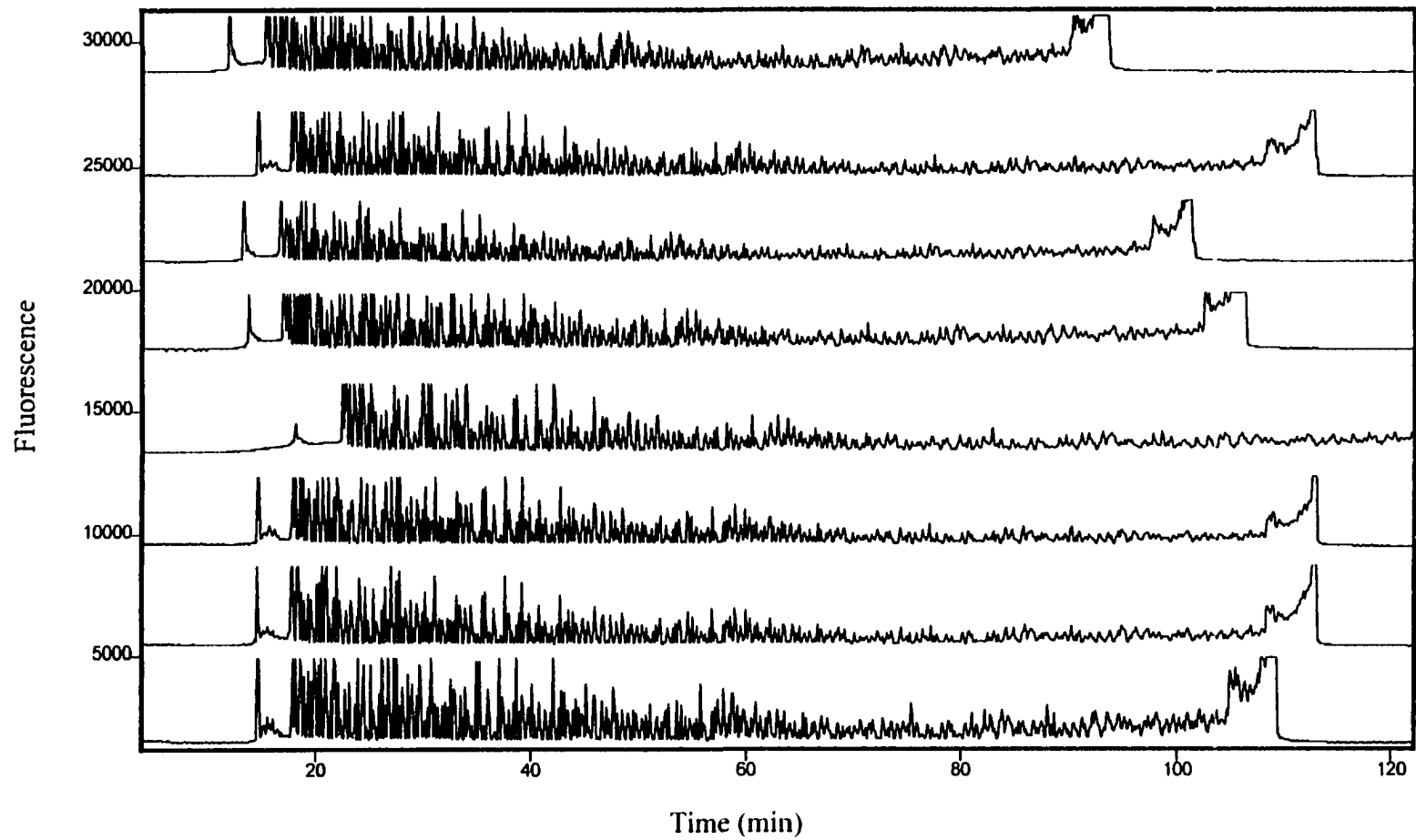


FIGURE 2

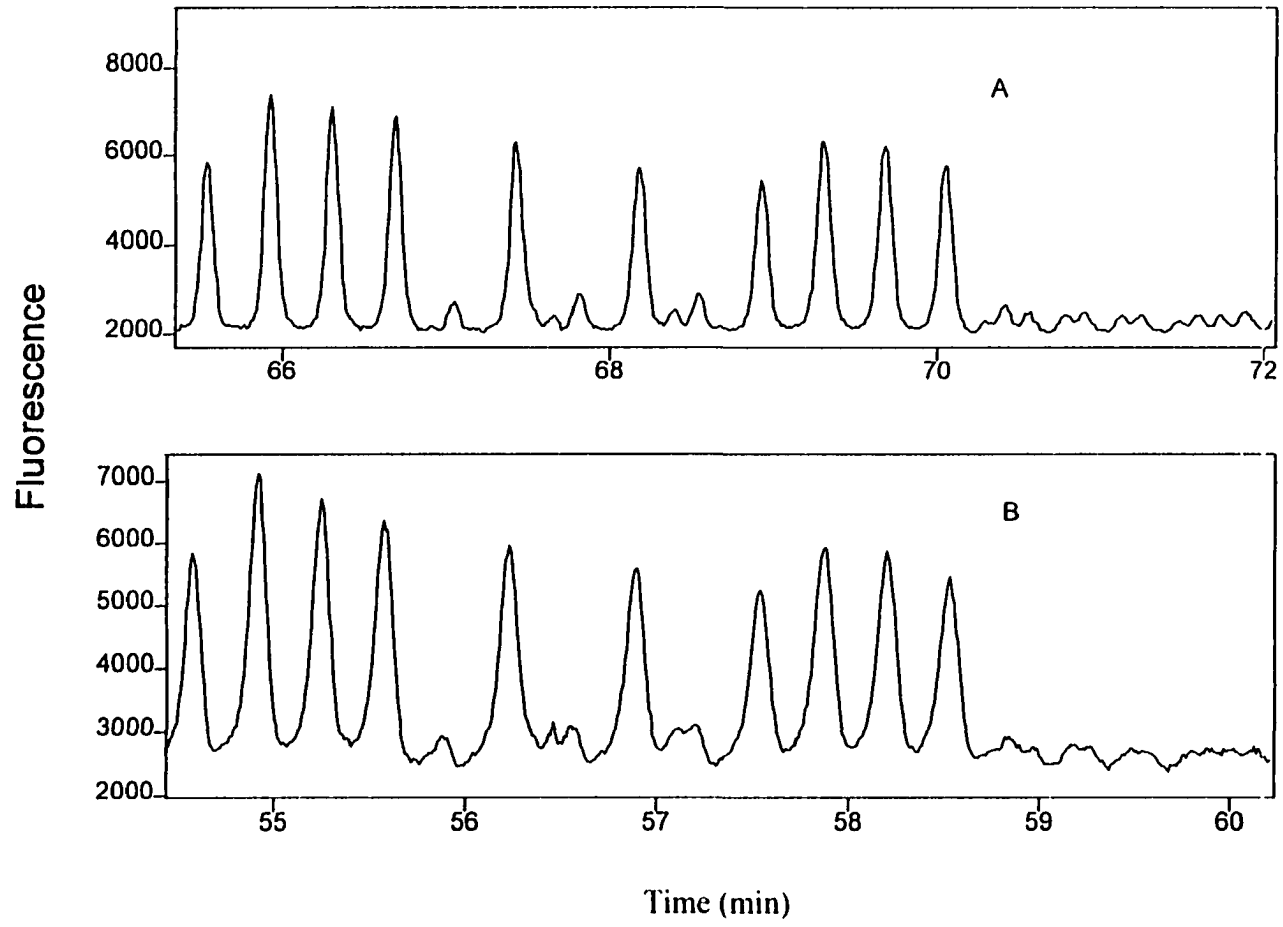
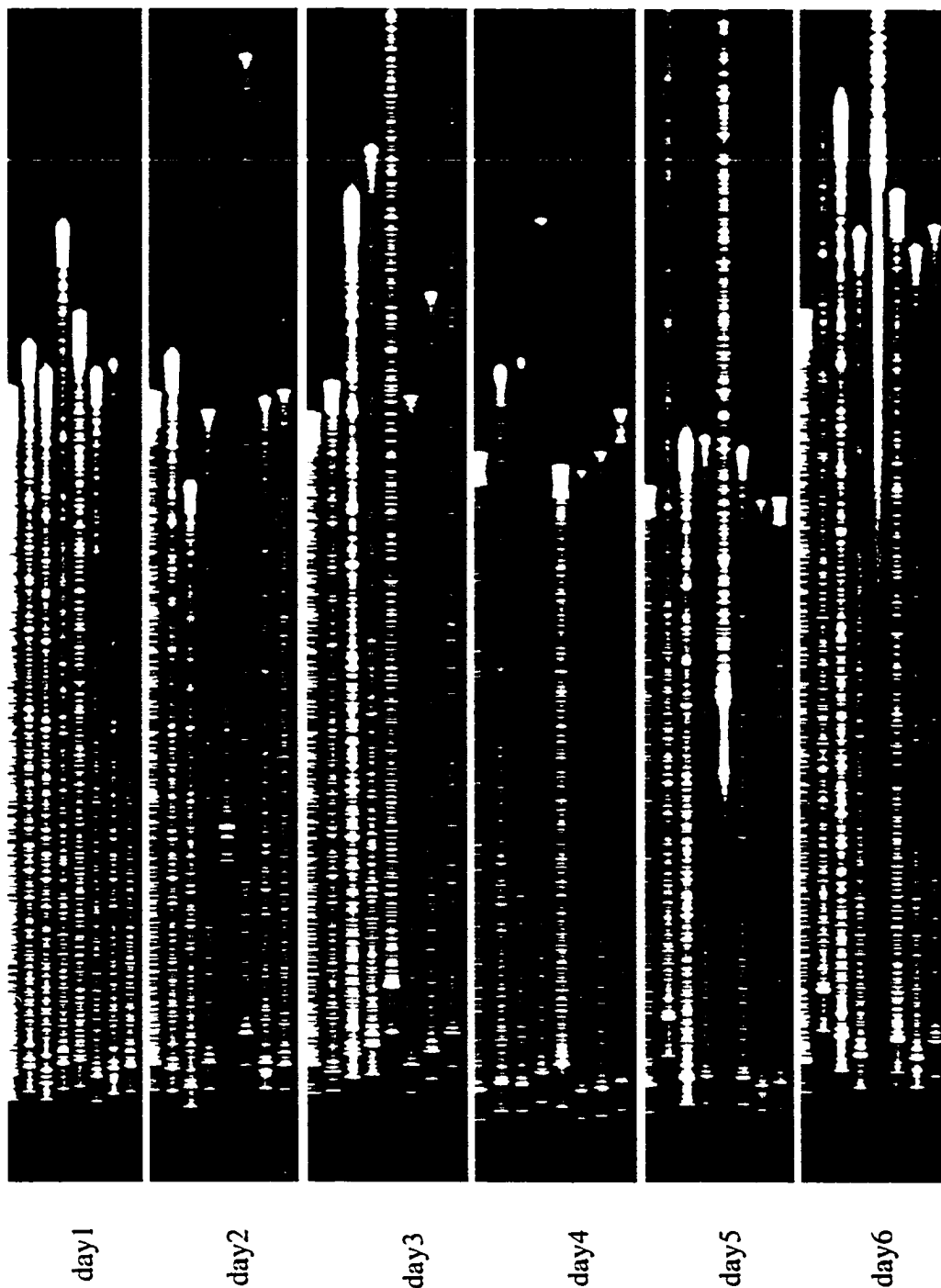


FIGURE 3



400bp

FIGURE 4

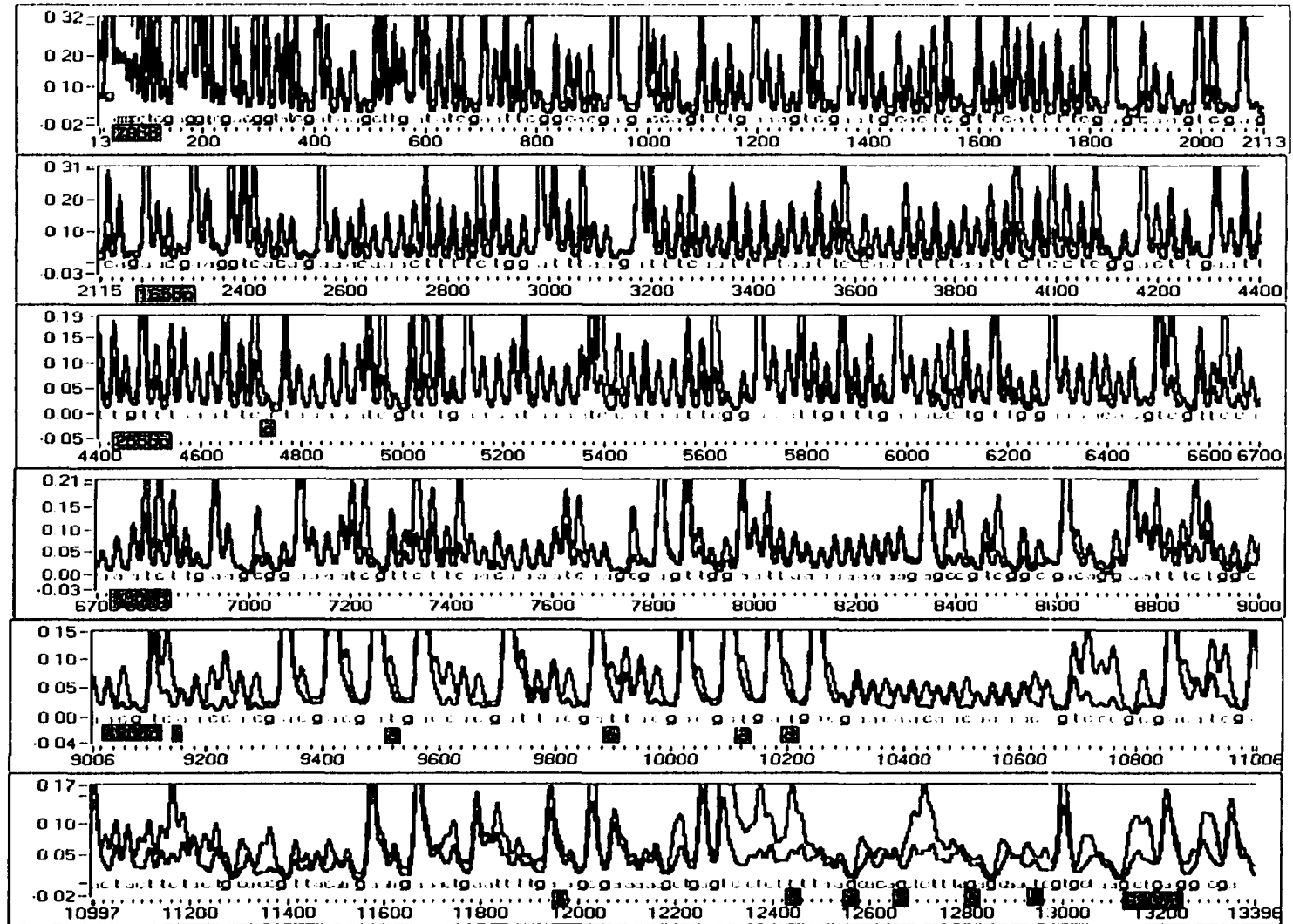


FIGURE 5

**CHAPTER 3. PCR ANALYSIS BY CAPILLARY
ELECTROPHORESIS WITH UV DETECTION FROM
CLINICAL SAMPLE**

A manuscript prepared for Journal of Chromatography B

Yonghua Zhang and Edward S. Yeung

Abstract

Routine genetic analysis of large numbers of individuals by PCR using capillary electrophoresis is often restricted by the high volume sample demand of DNA and high cost of laser induced fluorescence detection. Here we demonstrated that capillary electrophoresis with UV detection can be used in PCR based DNA analysis starting from clinical sample without purification. After PCR reaction using cheek cell, blood or HIV-1 gag DNA, the reaction mixtures were injected into the capillary either on-line or off-line by base stacking. Primers for both genomic and mitochondrial DNA were tested. This protocol was also performed in capillary array electrophoresis to increase the throughput. The use of multiplexed imaging absorption, and the elimination of purification of DNA sample before or after PCR reaction, will make this approach attractive alternative to current methods for genetic analysis.

Introduction

A revolutionary technique for DNA analysis is DNA amplification by the polymerase chain reaction (PCR), which allows the amplification of selected regions of DNA extracted from a variety of sample sources to a detectable level.^{1,2} In the last fifteen years, PCR has stimulated tremendous endeavor in genetic analysis, and so far there are many genes have been identified as the cause of a known heritable genetic disease.³ PCR also has been used to detect various infectious disease such as HIV⁴ and hepatitis^{5,6} because of its sensitivity and specificity. In post-PCR analysis, capillary electrophoresis has proved itself as the primary choice in term of speed, high resolution, and minimal sample requirement.⁷ In fact, capillary array electrophoresis has paved the way for the possible early finish of the Human Genome Project⁸ and should also play a important role in obtaining the genetic information for clinical use.⁹

One of the major obstacles to the rapid adoption of the use of PCR for routine diagnosis has been the high cost associated with the sample preparations and laser induced fluorescence (LIF) detection for capillary array electrophoresis. In the clinical and forensic laboratories, conventional PCR diagnosis, in most cases, starts from DNA purified from blood.² Even though the methods is well established, the labor input and the risk of sample contamination by extensive blood handling make people search for simple DNA purification methods to allow a faster turnaround time. Further improvement showed the blood can used as the template for PCR.^{10,11} Several alternative sources of DNA such as cheek cell,¹²⁻¹⁴ hair root,¹⁵ Guthrie card¹⁶ have also been reported. These sampling and preparation method are more cost effective than drawing blood, which also reduce exposure to harmful pathogens. But all these procedures still used slab gel electrophoresis for DNA analysis, which is quite

different with capillary electrophoresis in term of sample loading and injection.¹⁷ Also in these methods, either centrifugation or neutralization step was need for sample treatment, which obviously should be avoided for an ideal automated high throughput testing.

For laser induced fluorescence detection (LIF) in capillary electrophoresis, as few as 10^5 molecules are generally detectable in most laboratories and with state-of -the-art equipment, single molecule detection is attainable.¹⁸ But the analytes need derivation, which involves expensive reagents and raises the concerns for waste disposal because of the toxic nature of these labeling reagents. These make LIF for general application expensive. Recently, a 96-capillary array electrophoresis system based on imaging absorption detection has been invented in our lab¹⁹ and used in DNA analysis for genotyping and HIV diagnosis.²⁰ Even though starting with pure DNA and the PCR product purified by Microcon membrane, the whole procedure of analysis was cheaper because a low-cost and rugged high throughput instrument was used. Here we show that a new PCR sample preparation starting from cheek cell or blood can be used in capillary electrophoresis with UV detection and furthermore, no purification of PCR product was necessary by just using base stacking.²¹ The basis for this approach is that for every PCR reaction, if successful, the concentration of product will eventually reach a plateau which is around sub μM range.²² And this is sufficient for UV detection since each product has more than hundreds absorption unit. High throughput analysis using the new protocol also was demonstrated using a 13-capillary array electrophoresis. The overall integration provides a very cost efficient way to be used in clinical and forensic lab for PCR- based DNA analysis.

Experimental section

Reagents and materials

1×TE buffer (50 mM Tris/HCl, 2 mM EDTA, pH 8.0) was prepared by dissolving proportional Trizma⁺base, Trizma⁺Hydrochloride and EDTA disodium (all from Sigma, St Louis, MO) in deionized water. The coating matrix was made by dissolving 1% (w/v) of 1.300.000 MW Poly(vinylpyrrolidone) (PVP) from Sigma into the buffer, shaking for 2 min and letting it stand for 1 h to get rid of bubbles. Poly(ethylene oxide) (PEO) was obtained from Aldrich Chemical (Milwaukee, WI). The sieving matrix was made by putting 1% (w/v) 8.000.000 MW PEO into the buffer. The solution was stirred vigorously for 2 hours until all the material was dissolved and then slowly stirred for another 10 hours. 25-bp, 50-bp and 100-bp DNA ladders were purchased from Life Technologies (Gaithersburg, MD).

Sample preparation

1. On-line Human β -globin gene amplification from blood

5 ml blood samples were collected from volunteers into the containers which has 0.057 mL, 0.34M K₃EDTA as preservative. The final concentration of EDTA in blood is 3.9 μ M, which has no effect to the PCR reaction. The blood was stored in refrigerator to prevent the hemolysis. 6 μ l blood sample was mixed thoroughly with 90 μ l formamide. The mixture was incubated at 95°C for 10 min and then ready for PCR reaction. 3 μ l above mixture was used in a 20 μ l reaction to amplify a 110 bp fragment of the β -globin gene with specific primers (Roche Boehringer Mannheim Corp, Indianapolis, IN). In the 20 μ l reaction mixture, it had the following final composition: standard 1x PCR buffer from Promega (50 mM KCl and 10 mM Tris/HCl, pH 8.3; 3.5 mM MgCl₂; 250 μ M each of dNTP, 1 μ M of each primer; 0.25

$\mu\text{g}/\mu\text{l}$ BSA (Idaho Technology, Salt Lake city, UT); $0.5\mu\text{g}/\mu\text{l}$ T4 gene 32 protein; 0.25 unit/ μl Tth enzyme (Promega). The reaction was performed in a $360\mu\text{m}$ O.D., $250\mu\text{m}$ I.D. capillary (Polymicro Technologies, Phoenix, AZ) using Rapid Cyclor (Idaho Technology). The PCR protocol is listed in Table 1.

TABLE 1. PCR protocol using blood as template

β -globin from blood	Temp ($^{\circ}\text{C}$)	Time second	Cycle #
predenature	85	120	1
denature	80	60	3
anneal	45	60	
extension	60	60	
denature	80	15	40
anneal	45	30	
extension	60	60	
hold	60	600	1

The on-line set up and operation is similar to what we used before.²³ Briefly, PCR solution was aspirated by syringe pump to the reaction capillary and sealed by freeze-thaw valve. After reaction, the reaction mixture was on-line transfer to the injection cross and injected at $50\text{V}/\text{cm}$ for six min. Then a step of injection of 0.1 M NaOH is followed to insure the stacking of the DNA sample. Finally $1\times\text{TE}$ buffer was switched back to run the separation.

2. On-line HIV-1 gag amplification

A 115 bp fragment from the gag region of the HIV-1 DNA was amplified using the

HIV testing kit from Perkin-Elmer. The individual components were added according to the manufacture suggestion, except 0.25 $\mu\text{g}/\mu\text{l}$ BSA was added to prevent the adsorption of the Taq enzyme on the capillary. The protocol is 94°C for 2 min and followed by 40 cycles of denaturation at 95 °C for 15 s, annealing and extension at 60 °C for 1 min.. The annealing and extension temperatures were the same for this amplification.

3. PCR amplification from cheek cell

Buccal epithelial cells were collected by twirling a sterile swab (Solon manufacturing company, solon, ME) on the inner cheek for 30s. The swab was immersed into 400 μl formamide in a microfuge tube. The tube with the swab was heated at 95°C for 15 min. and the swab was removed. A 20 μl reaction consists of 3 μl of the buccal cell DNA solution, 1 μM of each of the primers, 200 μM of dNTP; 3 mM of MgCl_2 and 2.5 units fo Taq DNA polymerase (Promega) and 1xTris/HCl buffer with 0.25 $\mu\text{g}/\mu\text{l}$ BSA (Idaho Technologies). The primers used are shown in Table 2 and they were synthesized by DNA facility, Iowa State University. Cheek cells from six people in our group were tested and PCR reaction were. taking place in a Perkin-Elmer GeneAmp PCR system 2400. The temperature protocol was predenaturing at 80°C for 2 min. and followed by 35 cycles of denaturation at 80°C for 20 s, annealing at 53°C for 1 min. and extension at 60°C for 1 min..The primer sets used are listed in Table 2.

TABLE 2. PCR primer sequence

Primer name	Sequence of the primer	Product size
M69	5'GGTTATCATAGCCCACTATACTTTG 5'-ATCTTTATTCCCTTTGTCTTGCT	256
Cyto	5'-GCTGCTGGCATCACTATACT 5'-GGAAGCCAATTGATATCATA	241

Capillary array electrophoresis

The optics of the 13-capillary array electrophoresis system with photodiode array absorption detection is similar to the 96-capillary array which has been detailed before.²⁰ The capillaries used here are 75 μ m I.D. and 360 μ m O.D. and so the imaging pixels for each capillary is around 20 and the centered three pixels all gave good signal. Considering packing capacity and stability of the capillary, 250 μ m O.D. capillary will be the best choice. A DC-powered mercury lamp (UVP Inc., Upland, CA) was used as the light source for its lower noise levels compared with the AC-powered mercury lamp. The absorption wavelength was set at 254 nm by an interference filter (Oriel, Stamford, CT). The total length of the capillaries was 70 cm, with 50-cm effective length. The capillary array was first flushed with deionized water and then with 1 ml of 2% PVP. Finally 0.5 ml of 1% PEO (8,000,000 MW) sieving matrix was pushed into the capillary bundle by syringe. The system was then pre-run for 10 min with the electric field strength at 150 V/cm. After the 10-min. pre-run, the samples were injected electrokinetically at 50 V/cm for 360s. The 0.1 M NaOH was injected at 150 V/cm for 1 min., then the same field strength was employed for the separation. The total current was about 240 μ A during the separation process. 12 different samples plus a

baseline subtraction buffer sample were used in the 13-capillary array electrophoresis system.

Results and discussion

Base stacking for PCR analysis

PCR has the exponential amplification power for DNA, which is even able to start from a single copy.²⁴ The concentration of DNA after reaction is generally around sub μM and should be sufficient for UV detection. But for CE with UV detection, sample pretreatment is required for DNA separation because excess salts in the sample matrix interfered with effective injection.^{25,26} There are several stacking methods which can be used to overcome the low signal-to-noise problem.^{27,28} A possible solution may use size-exclusion column to purify the sample and on-line inject the sample into the capillary. As we already demonstrated in on-line DNA sequencing, sample in low ion strength buffer can be injected up to 3 min. at running voltage without sacrificing the resolution.²⁹ But a more simple way exists for the analysis of crude PCR mixture. Swerdlow et al.²¹ described a base stacking method for DNA sequencing from unpurified product. On column concentration of DNA fragments was achieved by electrokinetic injection of hydroxide ions. A low conductivity zone was produced by the neutralization reaction between OH^{-1} ions with the cationic buffer and DNA was concentrated. We use this stacking method to inject PCR reaction mixtures which contain 50mM KCl although the KCl is not always necessary.²⁸ As shown in Fig1, the signal is better in the case of base stacking (Fig 1.B) than in ordinary injection(Fig 1.A). Also there is no interference of dNTP since they move faster than the PCR product. In support of base stacking mechanism, the 0.1 M NaOH should be kept closed after stacking. Otherwise CO_2 will neutralize some OH^{-1} and ruin the stacking.

PCR from clinical sample

It is always desirable to amplify DNA directly from clinical material since the high volume testing involved for genetic analysis. Although blood is very complicated and using blood brings extra caution about contamination, the amount of DNA in blood is relatively consistent since 1 μ l blood generally contains $4.1-10.3 \times 10^3$ white blood cells.³¹ So as long as drawing blood is still necessary in lab for other testings, we will still see it will also be used in genetic diagnosis. It was found that the major problem for PCR from blood was due to the inability of the DNA polymerase to access the target DNA.¹⁰ The solution, then is to find conditions that simultaneously release DNA from larger numbers of cells in a form suitable for PCR while preserving the activity of Taq DNA polymerase. We tried to use water to lyse the blood, but the PCR was not successful. Due to the observed "cell debris", the DNA most likely was trapped by coprecipitation with proteins. So, FoLT (formamide low temperature) PCR which was demonstrated before for slab gel electrophoresis was also used here for UV-CE analysis.¹¹ The result is shown in Fig2.A.

For cheek cell, it is cleaner than blood. And cheek cell lysed in water has been used in PCR. Common method is using 15ml NaCl solution to wash mouth, which produces a yield of total genomic DNA about 2-5 μ g and sufficient for many PCR reactions.¹² One thing inconvenient is that centrifugation is needed to precipitate the cells. The use of cytology brushes and swabs as a more efficient means to cell collection/DNA extraction has also been demonstrated and validated in several groups.^{13,14} This method is easily performed in a 96-well format and is compatible with high volume testing. But in their process, cheek cells were lysed in 0.1 M NaOH and later neutralized with Tris/HCl. Here we show that we can avoid the neutralization step by using formamide and heating to lyse the cells. Compared

with using normal PCR from pure DNA (Fig 2.B), PCR from cheek cells (Fig 1.B) or blood (Fig 2.A) gave just slightly worse but decent results. One thing we need to mention is that UV is almost universal detection, and complex lysate may bring some interference. But control experiments show a clear window for DNA detection.

On-line and high throughput analysis

Several on-line analysis systems have been reported for DNA analysis and all of them used laser induce fluorescence detection since it is extremely sensitive.³²⁻³⁴ We are interested in on-line UV system for handling infectious virus and blood since automated procedure can avoid the manual operation and reduce the chance of contamination. Also such automated UV system may be useful for the measurements of other clinical chemicals, which current methods are also based on different absorption properties of the analyte. In Fig 3, PCR amplification of β -globin from blood and HIV analysis were demonstrated. Fig 4 shows the results of analyses using capillary array electrophoresis. 1 \times TE buffer was injected into capillary 13 and was used as the base-line subtraction capillary to cancel out the flicker noise from the mercury lamp, as reported before.¹⁹ 11 different samples using six people's cheek cells with two different primer sets were simultaneous amplified in PE cycler and analyzed by the array (from capillary 1 to 11). The electropherogram from capillary 12 shows the negative PCR result, which no cheek cell was added but only formamide and every other reagent for PCR reaction. This shows no DNA contamination in any reagent. Amplification from mitochondrial DNA generates more DNA product since mitochondria DNA are present in hundreds in each human cell.³⁵ To increase the confidence level for identification, the PCR products were coinjected with 100 bp ladder for capillary 5, 6, 9, 10, 11. The electropherograms show the two products which are round 240bp for cyto primer and 250bp

for the M69 primer. The capillary array was flushed with water between runs and did not show degradation over tens of runs in a month.

Conclusions

New sample preparation starting from clinical samples such as cheek cell and blood was developed using low temperature formamide PCR for capillary electrophoresis with UV detection. The cheek cells and blood were heated in formamide for 15 min and used directly as the DNA source. No centrifugation, vortex or neutralization steps was involved. After reaction, the PCR product can be injected into the capillary without sample purification by base stacking. The protocol was also demonstrated feasible in on-line format for HIV analysis and PCR from blood. Different samples from six people using two primer sets were successfully analyzed by the array electrophoresis. The capillaries in the array are stable for base stacking. Since the sample preparation step and sample injection step can be easily integrated and automated by this protocol, we demonstrated a cost efficient and high throughput way to perform genetic analysis or disease diagnosis at very low cost.

Acknowledgement

The authors thank Dr. Xiaoyi Gong for assistance in setting up the capillary array and Dr. Yan He for many helpful discussions about PCR. Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract No. W-7405-Eng-82. The funding for the instrument and protocol development was from the DOE and National Institutes of Health.

References

1. R. K. Saiki, S. J. Scharf, F. A. Faloona, K. B. Mullis, G. T. Horn, H. A. Erlich and N. Arnheim, *Science*, **1985**, 230, 1350.
2. M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White. Eds *PCR protocols*, **1990**, Academic Press, San Diego
3. H. A. Erlich and N. Arnheim, *Anal. Rev. Genet.* , **1992**, 26, 479.
4. C. Y. Ou, S. Kwok, S. W. Mitchell, D. H. Mack, J. J. Sninsky, J. W. Krebs, P. Feorino, D. Warfield and G. Schochetman *Science*, **1988**, 239, 295.
5. S. Kaneko, R. H. Miller, S. M. Fernsont, M. Unoura, K. Kobayashi, N.Hattori and R.H. Purcell, *Proc.Natl. Acad. Sci. U.S.A.*, **1989**, 86,312.
6. N.Bianchi, C.Mischiati, G. Feriotto, R. Gambari, *Int. J. Oncol.*, **1994**, 4, 903
7. N. J. Dovichi, *Electrophoresis*, **1997**, 18, 2393
8. E. Pennisi, *Science*, **2000**, 288, 2304
9. Q. Gao and E. S. Yeung, *Anal. Chem.*, **2000**, 72, 2499
10. M.Panaccio and A. M. Lew, *Nucl. Acids. Res*, **1991**, 19, 1151.
11. M. Panaccio and A.M. Lew, *PCR Technology: Current Innovations*, **1994** CRS Press Inc., 151
12. N. Lench, P. Stamer and R. Williamson, *Lancet*, **1988**, 1356
13. B. Richards, J. Skoletsky, A. P. Shuber, B. Rosemary, R.C. Stern, H.L. Dorkin, R. B. Parad, W. David and W.K. Katherine, *Hum.Mol. Genet.*, **1993**, 2, 159.
14. Y. Liu, B. Jian, Y. Zhu, X. Liang, D. Siemieniak, J.V. Patrick and D. M. Lubman, *Rapid Commun. Mass. Spectrom.*, **1995**, 9, 735.
15. D.M. Thomson, N. N. Rrown and A. E. Clague, *Clinic. Chim. Acta.*, **1992**, 207, 169.

16. B. C. Harty, M. Garcia-Closas, N. Rothman, Y.A. Reid, M. A. Tucker, P. Hartge. *Cancer Epidemiol. Biomarkers Prev.* **2000**, 9, 501.
17. H. Swerdlow, K.E. Dew-Jager, K. Brady, R. Grey, N. Dovichi and R. Gesteland. *Electrophoresis.* **1992**, 13, 475.
18. C. E. MacTaylor and A.G. Ewing, *Electrophoresis.* **1997**, 12-13, 2279.
19. X. Gong and E. S. Yeung, *Anal. Chem.*, **1999**, 71,4989.
20. X. Gong and E. S. Yeung, *J. Chromatorgr. B.* **2000**,741,15.
21. Y. Xiong, S .R. Park and H. Swerdlow, *Anal. Chem.*, **1998**, 70,3605.
22. W. Block, *Biochem.*, **1991**, 30, 2735.
23. N. Zhang, H. Tan and E. S. Yeung, *Anal.Chem.*, **1999**,71,1138.
24. H. Li, U. B. Gyllensten, X. Cui, R. K. Saiki, H. A. Erlich and N. Arnheim, *Nature*, **1988**, 335, 414.
25. X.C. Huang, S. G. Stuart, P.F. Bente, T.M. Brennan, *J. Chromatogr.*, **1992**, 600, 289.
26. P.E. Williams, M.A. Marino, S.A. Del Rio, L. A. Turni, J. M. Devaney, *J. Chromatogr. A.* **1994**, 680, 525.
27. R. L. Chien and D. S .Burgi, *Anal.Chem.*, **1992**, 64, 489A
28. J. P. Quirino and S. Terabe, *Science*, **1998**, 282, 465.
29. Y. Zhang, H. Tan, and E. S. Yeung, *Anal.Chem.*, **1999**, 71, 5018.
30. H. M. Eun, *Enzymology Primer for Recombinant DNA Technology*, **1996**, Academic Press, New York..
31. *Directory of Service and Interpretive Guide*, Labcorp, **1997**
32. S. A. Soper, D.C. Williams, Y. Xu, S. J. Lassiter, Y .Zhang, S. M.Ford and R.C. Brunch, *Anal.Chem.*, **1998**, 70, 4036.

33. H. Swerdlow, B.J. Jones and F. T. Wittwer. *Anal.Chem.*..1997, 69, 848.
34. H. Tan and E. S. Yeung . *Anal.Chem.*.. 1997. 70. 4044.
35. J. S. Hanekamp, W.G. Thilly, M.A. Chaudhry, *Hum. Genet.*. 1996. 98. 243.

Figure captions

- Figure 1. Sample analysis of the crude PCR product from cheek cells spiked with 100 bp DNA ladder using different injection method (A) 150 V/cm, 30 sec. normal injection: (B) 50V/cm. 6 min., then 0.1 M NaOH, 150V/cm, 1 min.
- Figure 2. Comparison of PCR product starting from (A) Blood; (B) Purified DNA; Also see Fig.1 (A) form cheek cells.
- Figure 3. On-line PCR analysis (A) a 110 bp fragment of β -globin gene amplification from blood spiked with 25 bp ladder: (B) a 115 bp fragment of HIV gag is clearly identified by the peak between 100 bp and 150 bp.
- Figure 4. Electropherograms from capillary array electrophoresis: products peaks amplified from mitochondrial DNA are marked with M; products peaks amplified from genomic DNA are marked with G.

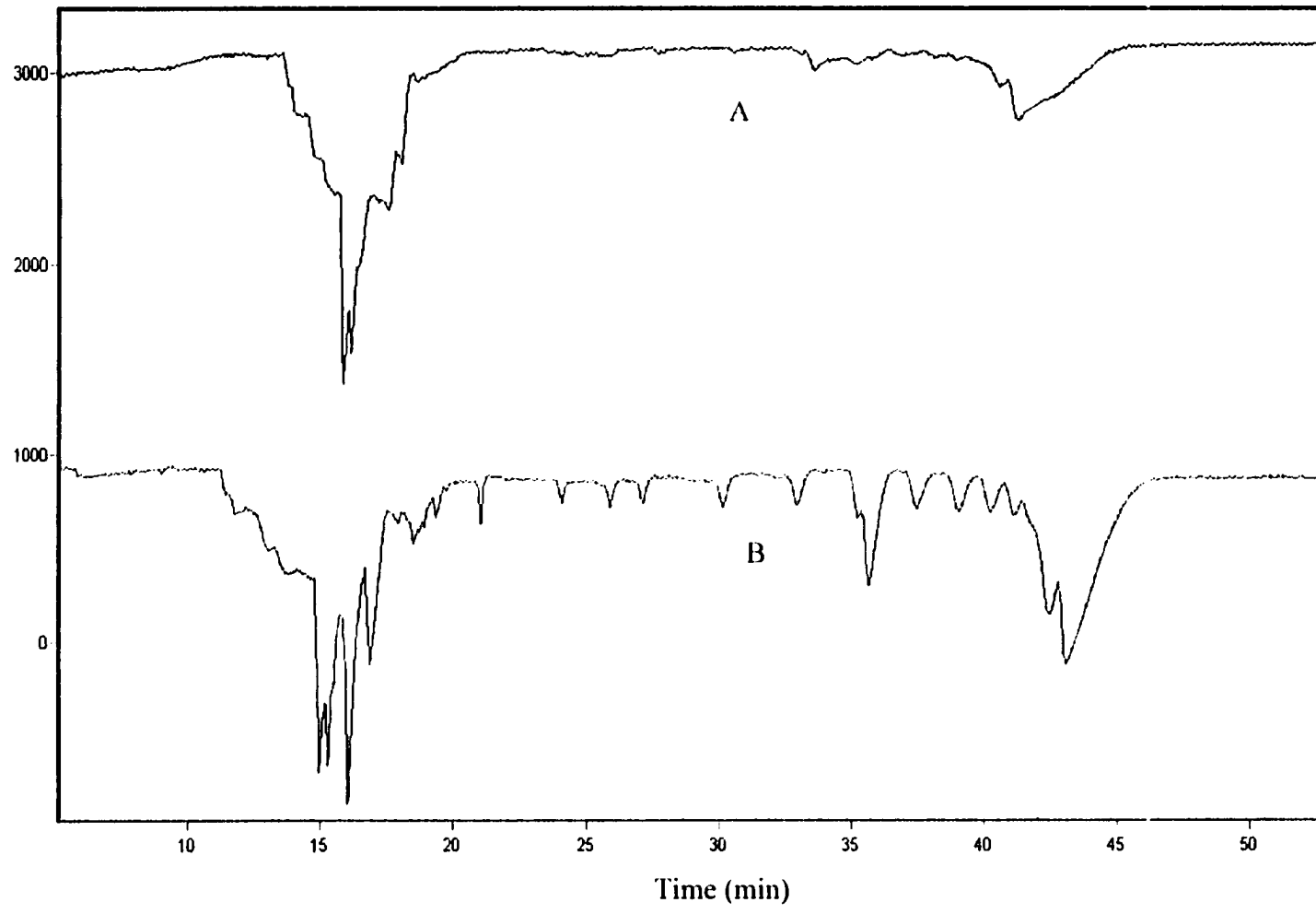


FIGURE 1

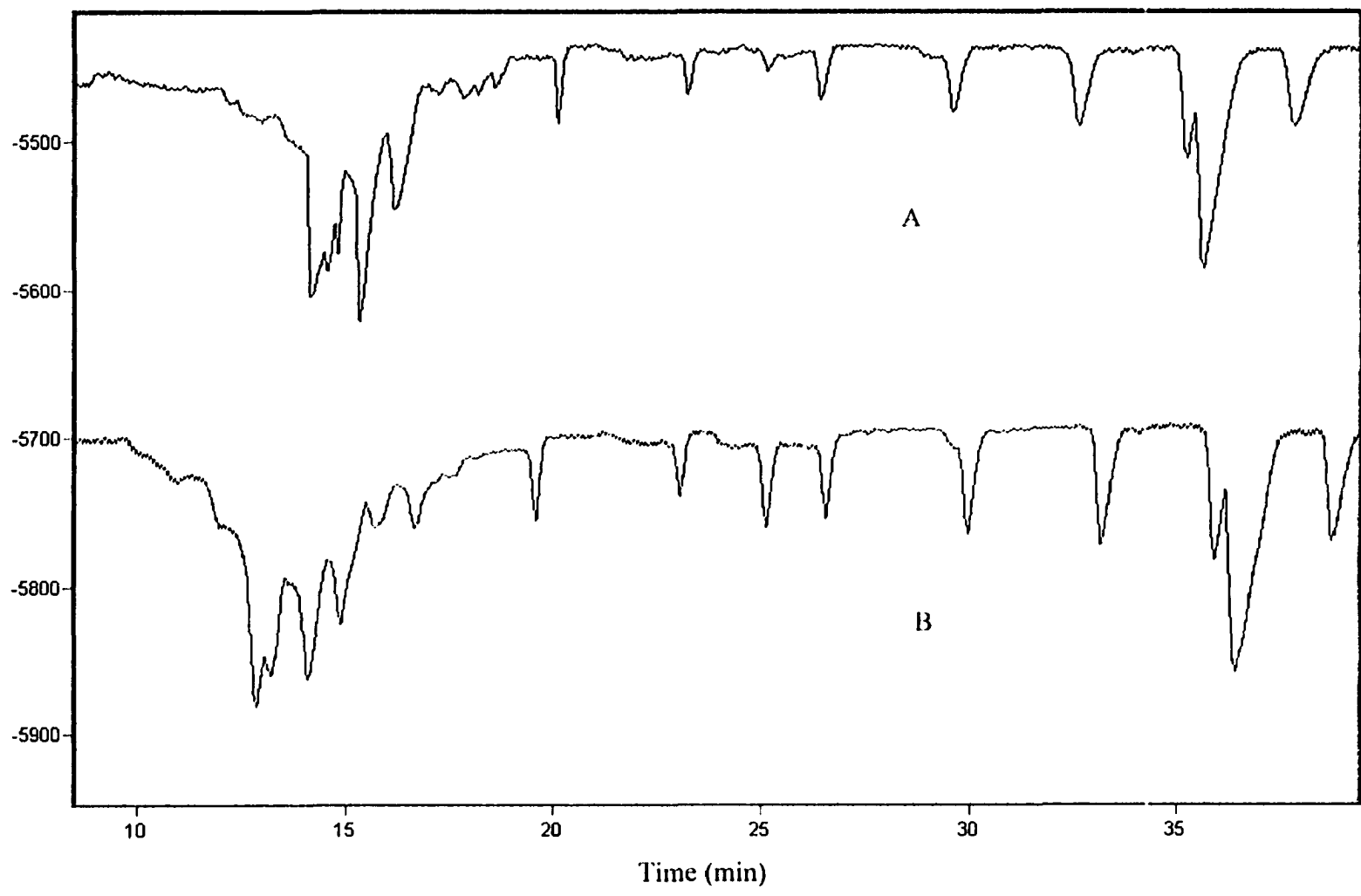


FIGURE 2

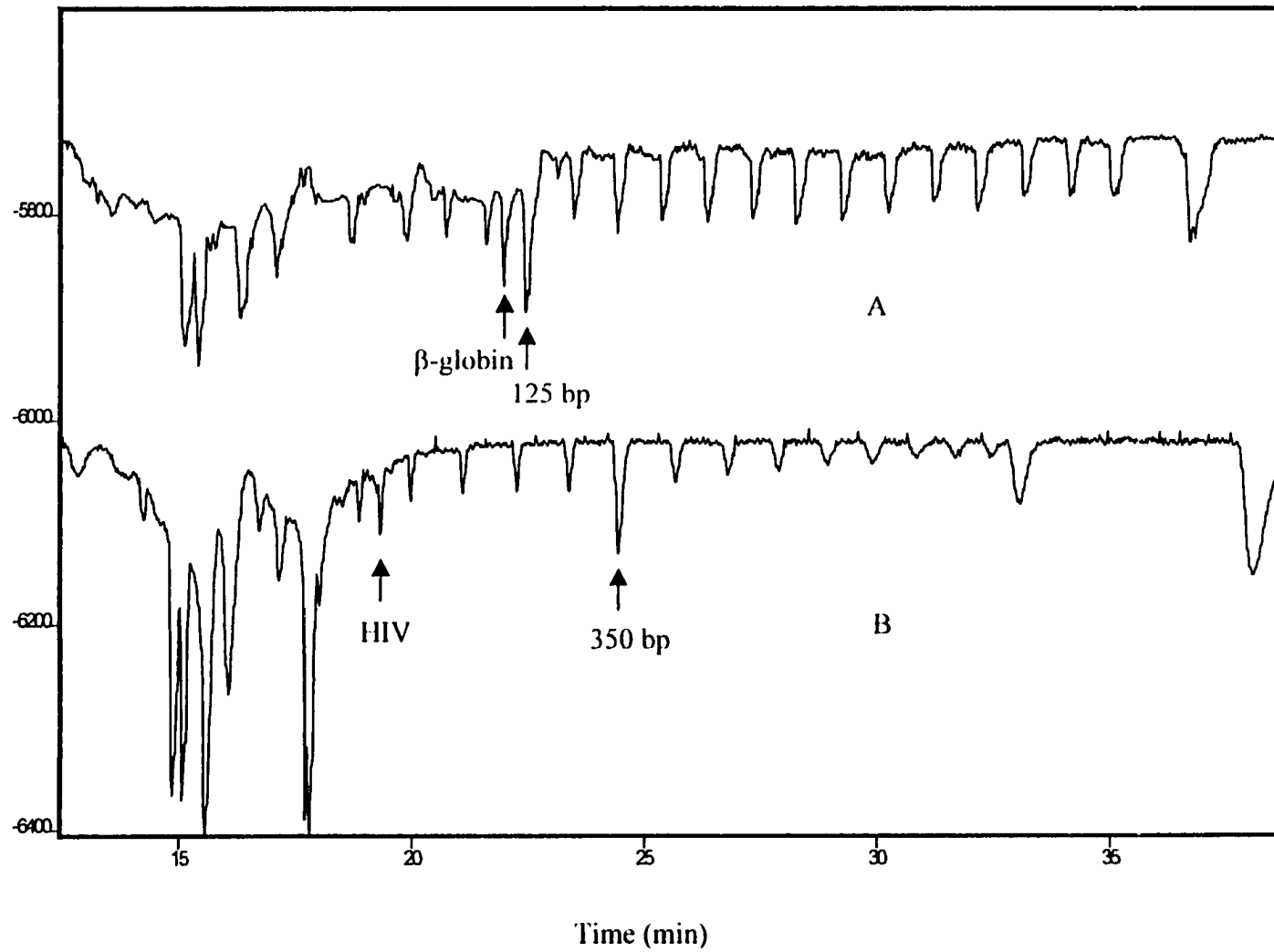
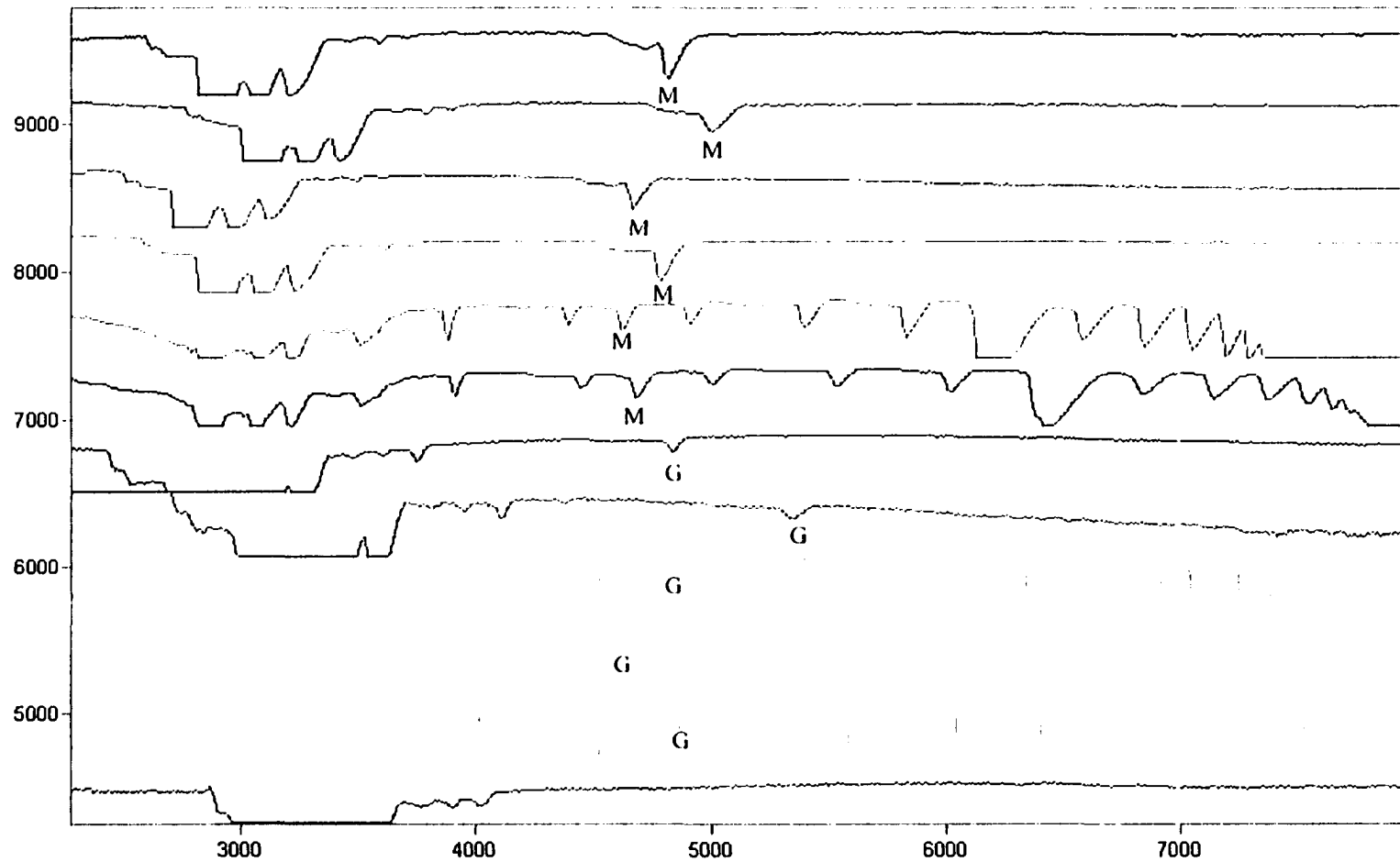


FIGURE 3



Frame number

FIGURE 4

CHAPTER 4. COMBINATORIAL SCREENING OF HOMOGENEOUS CATALYSIS AND REACTION OPTIMIZATION BASED ON MULTIPLEXED CAPILLARY ELECTROPHORESIS

A paper accepted by Journal of Combinatorial Chemistry

Yonghua Zhang, Xiaoyi Gong, Haiming Zhang, Richard C. Larock and Edward S. Yeung

Combinatorial chemistry is revolutionizing the discovery of new drugs,¹ novel materials and efficient catalysts² by scanning and testing vast numbers of possibilities. In order to fully realize the potential of combinatorial chemistry, general and powerful schemes for high-throughput screening (HTS) are essential.³ Capillary array electrophoresis (CAE), a high-throughput technique driven by the Human Genome Project, has taken a key role in genomic analysis⁴ and potentially will contribute to proteomics as well.⁵ We report here the use of CAE for the rapid screening of a homogeneous catalytic reaction in a combinatorial manner. This approach has allowed the effective optimization of a homogeneously catalyzed synthetic organic reaction and the discovery of conditions that produce yields superior to those obtained previously by a less systematic approach.

So far there are several parallel assays for screening homogeneous catalysts. Modifications in UV absorption,⁶ fluorescence,⁷ color⁸ or temperature⁹ induced by the catalytic reactions are indicators of catalytic activity. In these approaches, although the relative activity of the catalyst is determined quickly, no quantitative information about the overall yield or the regioselectivity and stereoselectivity of the process can be obtained. It is also necessary that the product exhibit very different measurable properties compared to the solvent or the re-

agents. Most of the time, secondary screening is necessary. Mass spectrometry (MS),¹⁰ which has also been widely used to screen catalysts, can provide selective detection. However, to address stereoselectivity, these procedures still tend to be laborious.¹¹ So far, MS is still a serial, rather than a parallel, approach although the analysis time is reasonably short. Application to the optimization of synthetic organic reactions will require the development of a high-throughput interface.

Separation-based techniques can solve the above problems. Serial methods, which include fast high performance liquid chromatography (HPLC) and capillary electrophoresis (CE), have been used to analyze asymmetric catalysis¹² and alkylation reactions.¹³ The throughput that can be achieved with serial separation schemes is low even with special techniques, such as sequential sample injection¹⁴ and sample multiplexing.¹⁵ Multiplex HPLC is another interesting approach,¹⁶ but achieving a high degree of multiplexing, such as 96 capillaries in CAE, is not trivial. Thin-layer chromatography and gel electrophoresis, on the other hand, are difficult to completely automate. The uniqueness of CAE with absorption detection¹⁷ is the easy adaptation to large numbers of samples and near-universal UV detection. Also the injection volume is minimal (10-100 nL) and thus microreactors can be used to save reagent cost.

The model reaction we have used for demonstration of the unique capabilities of multiplexed CAE (Scheme 1) is a new palladium-catalyzed annulation reaction,¹⁸ which readily affords γ -carboline, noteworthy for their biological activity. The optimal reaction conditions and the regiochemistry for this type of annulation are generally highly dependent on the nature of the palladium catalyst and the base employed. Previous efforts to optimize this process employed 5% Pd(OAc)₂, 10% PPh₃ and Na₂CO₃ as base and afforded a 1:1 ratio of isomers A/B in essentially a quantitative yield.

The nature of this and other catalytic reactions is that a lot of parameters can affect the yield and "optimum" conditions are often found by trial and error. The general scale on which

we have run the above reaction is 0.25 mmole in 5 mL of DMF. We have reduced the volume to 120 μ L by using 6 mm O.D. glass tubes sealed at one end arranged in a 96-well format. The individual components were added as a DMF solution or as a slurry by pipetting. Septums were used to cap the reaction tubes to prevent evaporation. All reactions were thus run on a 5 μ mole scale. Heating was provided by a dry heat bath kept at 110 $^{\circ}$ C. As an internal standard, 1 μ mole of norharman was added to the reaction mixture. We did not observe any catalytic effect on the system from the addition of the norharman in control experiments. The CAE experiment is similar to what we have reported before.⁵ Organic-based buffers,¹⁹ which are more appropriate for organic synthesis, have been used because of the low solubility of the products in water. Complete separation of the reactant, products (isomer A and B) and internal standard was achieved in 40 mM NH_4OAc , 0.75% formic acid and 80% DMF with 20% H_2O at an applied field of 140 V/cm in a fused-silica capillary with effective/total length of 50/75-cm, 50- μ m I.D. and hydrodynamic injection 15 s at 8-cm height (Fig. S1, Supporting Information).

One important feature of the experimental protocol is that we injected the reaction mixture into CAE without diluting or quenching before analysis. At predetermined times during the reaction, the reaction block was removed from the heating platform, quickly cooled and put under the injection ends of the capillary array. No deleterious effect on the catalytic system was observed by this operation. By avoiding sample manipulation (e.g. by pipetting out of the reaction vials), we can reduce errors associated with transfer and contamination. We also noted that the CAE running buffer should be compatible with the reaction buffer for hydrodynamic injection. When using methanol as the buffer, injection was not uniform. Only about half of the 96 capillaries had adequate signal. It was not possible to increase the injection time, because some capillaries then became overloaded. When DMF-based buffer was used, all 96 channels had uniform signal over three consecutive runs. This buffer compatibility issue for CAE may be attributed to the differences in solution properties, such as viscosity and

surface tension, and was not observed in single-capillary experiments. The total analysis time is typically 60 min. plus 30 min for capillary cleaning. Judging from the resolution, the capillaries could have been shortened to 25% to provide analysis times of 15 min.

By choosing 8 different Pd catalysts and 11 different bases, 88 different combinations have been tested (Fig. S2, Supporting Information). We can determine the total yield (Fig. 1), selectivity (Fig. 2) and reaction kinetics (see Supporting Information) from the electropherograms. Some of the conditions have been tested previously.¹⁸ Our results agree well with those. One example is that by using Pd(OAc)₂ with the ligand PPh₃ as catalyst and Na₂CO₃ as the base, a total yield of 84% was achieved with virtually no regioselectivity in the microreactor, compared with a quantitative conversion (90% after 17 hours) with no selectivity under the protection of N₂ in a 5 mL reaction. Among all of the bases, inorganic bases proved to be more effective in promoting the reaction. When pyridine or other organic bases were used, the yield was low and some side products appeared. The ability to detect side products is clearly an advantage of CAE. Our preliminary results also reveal several new conditions which are quite effective in this annulation reaction. They are Pd(PPh₃)₄ with Na₂CO₃ (C9, 74%), Pd(dba)₂ with K₂CO₃ (E10, 72%), PdBr₂ plus 2PPh₃ with Na₂CO₃ (G9, 88%) and PdBr₂ plus 2PPh₃ with K₂CO₃ (G10, 96%). The latter two are in fact superior to our previous best catalytic condition.¹⁸ Complete regioselectivity is not observed in any of the test conditions (Fig. 2), even though some prove to be better than other systems.²⁰ The conditions G2, H2, B1 and C1 have some selectivity, but unfortunately their yields are low. There are significant differences in the rates and the shapes of the rate plots (Fig. S3, Supporting Information). This illustrates the need to monitor the reactions at several points in time.

In summary, a new methodology, nonaqueous capillary array electrophoresis coupled with microreaction, is developed to address the throughput needs of combinatorial approaches to homogeneous catalysis and reaction optimization. Catalytic activity, selectivity and kinetics of the various combinations are determined quickly. Other combinatorial applications that can

be envisioned based on this method are screening for asymmetric catalysts and drugs, as well as combinatorial library synthesis.⁵

Acknowledgement

The Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract No. W-7405-Eng-82. This work was supported by the Director of Science, Office of Basic Energy Sciences, Division of Chemical Sciences. The Larock group gratefully acknowledges partial financial support from the Petroleum Research Fund, and Kawaken Fine Chemicals Co., Ltd. for some of the palladium reagents.

Supporting Information

Figures of separations, result of CE separation and kinetics of reactions are shown in Appendix A of this dissertation.

References

1. *A Practical Guide to Combinatorial Chemistry*, Czarnih, A. and Dewitt, S., Ed., American Chemical Society, Washington DC, 1997.
2. (a) Brocchini, S., James, K., Tangpasuthadol, V., and Kohn, J., *J. Am. Chem. Soc.*, 1997, 119, 4553-4554. (b) Jandeleit, B., Schaefer, D., Powers, T., Turner, H., and Weinberg, W., *Angew. Chem. Int. Ed.*, 1999, 38, 2494-2532.
3. Kyranos, J., and Hogan, J., *Anal. Chem.*, 1998, 389A-395A.
4. Collins, F., Patrinos, A., Jordan, E., Chakravarti, A., Gesteland, R., and Walters, L., *Science*, 1998, 282, 682-689.
5. (a) Kang, S., Gong, X., and Yeung, E. S., *Anal. Chem.*, accepted. (b) Ma, L.; Gong, X., and Yeung, E. S., *Anal. Chem.*, accepted.
6. (a) Wagner, J., Lerner, R., and Barbas, C., *Science*, 1995, 270, 1797-1800. (b) Menger, F., Ding, J., Barragan, V., *J. Org. Chem.*, 1998, 63, 7578-7579.

7. (a) Cooper, A., McAlexander, L., Lee, D., Torres, M., and Crabtree, R., *J. Am. Chem. Soc.*, **1998**, 120, 9971-9972. (b) Shaughnessy, K., Kim, P., and Hartwig, J., *J. Am. Chem. Soc.*, **1999**, 121, 2123-2132. (c) Copeland, G., and Miller, S., *J. Am. Chem. Soc.*, **1999**, 121, 4306-4307.
8. Lavastre, O., and Morcken, J., *Angew. Chem. Int. Ed.*, **1999**, 38, 3163-3165.
9. (a) Taylor, S., and Morcken, J., *Science*, **1998**, 280, 267-270. (b) Reetz, M., Becker, M., Kuhling, K., and Holzwarth, A., *Angew. Chem. Int. Ed.*, **1999**, 37, 2647-2650.
10. Orschel, M., Klein, J., Schmidt, H., and Maier, W., *Angew. Chem. Int. Ed.*, **1999**, 38, 2791-2794.
11. Reetz, M., Becker, M., Klein, H., and Stockigt, D., *Angew. Chem. Int. Ed.*, **1999**, 38, 1758-1761.
12. (a) Porte, A., Reibenspies, J., and Burgess, K., *J. Am. Chem. Soc.*, **1998**, 120, 9180-9187. (b) Ding, K., Ishii, A., and Mikami, K., *Angew. Chem. Int. Ed.*, **1999**, 38, 497-501.
13. Gaus, H., Kung, P., Brooks, D., Cook, D., and Cummins, L., *Biotech. & Bioeng.*, **1998/1999**, 61, 169-177.
14. Roche, M., Oda, R., Machacek, D., Lawson, G., and Landers, J., *Anal. Chem.*, **1997**, 69, 99-104.
15. Woodbury, C., Fitzloff, J., and Vincent, S., *Anal. Chem.*, **1995**, 67, 885-890.
16. Zeng, L., and Kassel, D., *Anal. Chem.*, **1998**, 70, 4380-4388.
17. Gong, X., and Yeung, E. S., *Anal. Chem.*, **1999**, 71, 4989-4996.
18. Zhang, H., and Larock, R., manuscript in preparation. For previous annulation chemistry of this type, see Larock, R. C., *J. Organometal. Chem.*, **1999**, 576, 111-124.
19. Sahota, R., and Khaledi, M., *Anal. Chem.*, **1994**, 66, 1141-1147.
20. Larock, R. C., Yum, E., and Refvik, M., *J. Org. Chem.*, **1998**, 63, 7652-7662.

Figure captions

FIGURE 1. Total yield of the reaction after 17 hr at 110 °C. dppe = bis(diphenylphosphino)ethane, TBAC = tetra-n-butylammonium chloride, DABCO = 1,4-diazabicyclo[2.2.2]octane. dba = trans, trans-dibenzylideneacetone.

FIGURE 2. Selectivity plot of the two isomers produced in the reactions. P1/P2 is the ratio of the two isomers A and B respectively.

SCHEME 1. The model reaction

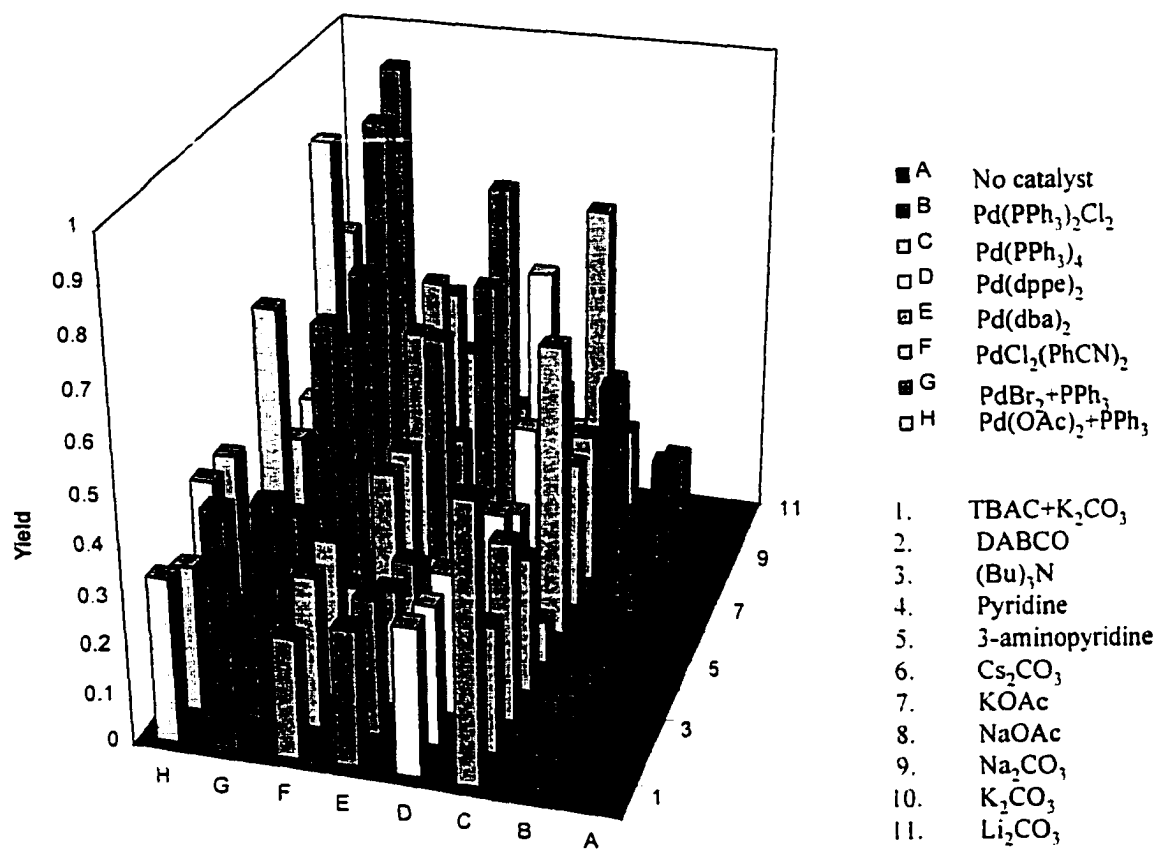


FIGURE 1

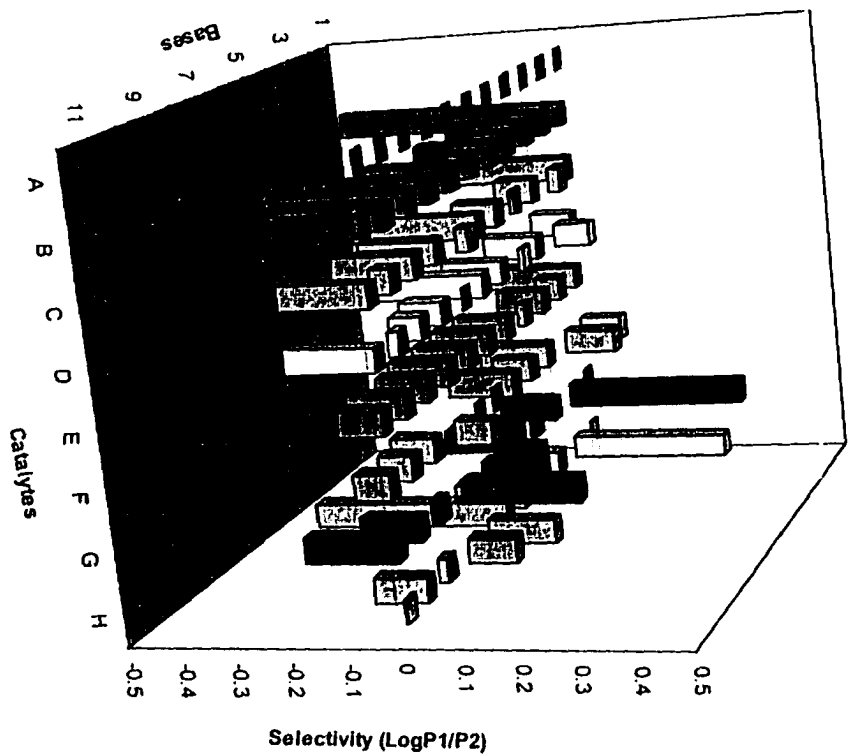
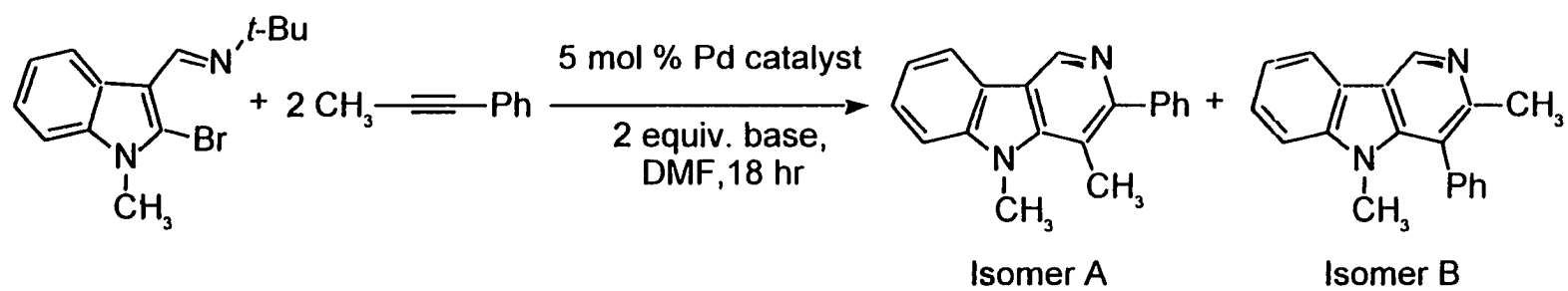


FIGURE 2



SCHEME 1

CHAPTER 5. GENERAL CONCLUSIONS

Capillary array electrophoresis (CAE), a high throughput technique driven by the Human Genome Project, has taken a key role in genomic analysis and potentially will contribute to proteomics and combinatorial chemistry as well. The unique properties of capillary electrophoresis such as high speed, high resolution separation and nl sample requirement coupled with the high parallel power of the capillary array impose a great demand on the both the throughput and approach of sample preparation. Obviously, both the sample preparation and the separation must be balanced, integrated and multiplexed to the same level to fully realize the potential of capillary array.

The work in this dissertation has described the developments of sample preparation and their subsequent analysis by capillary array electrophoresis. First for DNA sequencing, sequencing directly from a single colony in a closed, multiplexed, and automated on-line capillary array instrument was demonstrated for the first time. Such procedure can provide base calling up to 620 bp with 98% accuracy even with the simple algorithm. The overall performance from continuous operation in a one-week period using lysates as templates showed similar reliability to that of the off-line sequencer. Secondly, similar to colony sequencing, new sample preparation was also developed starting from front-end samples such as cheek cell and blood. After reaction, the PCR mixture can be injected into the capillary by base stacking without purification. The coupling of this PCR procedure with imaging UV absorption detection capillary array electrophoresis provide a cost efficient and high throughput way to perform genetic analysis or disease diagnosis at very low cost. Finally, the same microreaction technique widely used in biology was successfully transferred to the combinatorial screening of homogeneous catalysis. Nonaqueous CE was

first time demonstrated in capillary array electrophoresis and expanded the use of CAE to water insoluble compounds.

For the future efforts, several other interesting sample preparation applications can be envisioned based on multiplexed capillary electrophoresis. With the miniaturization of high throughput drug screening, we may see a dramatic change in the amount of sample prepared, and CAE with bigger I.D. capillary is well positioned for the preparative use of combinatorial chemistry. Since each capillary in the array can be independent and separation optimization for CE analysis will become trivial when we optimize separation conditions in parallel, and this is especially true for chiral separation. The further development of more information rich detectors for CAE will bring CAE to a new height for accurate identification of combinatorial library compounds.

APPENDIX A. SUPPORTING INFORMATION FOR CHAPTER 2

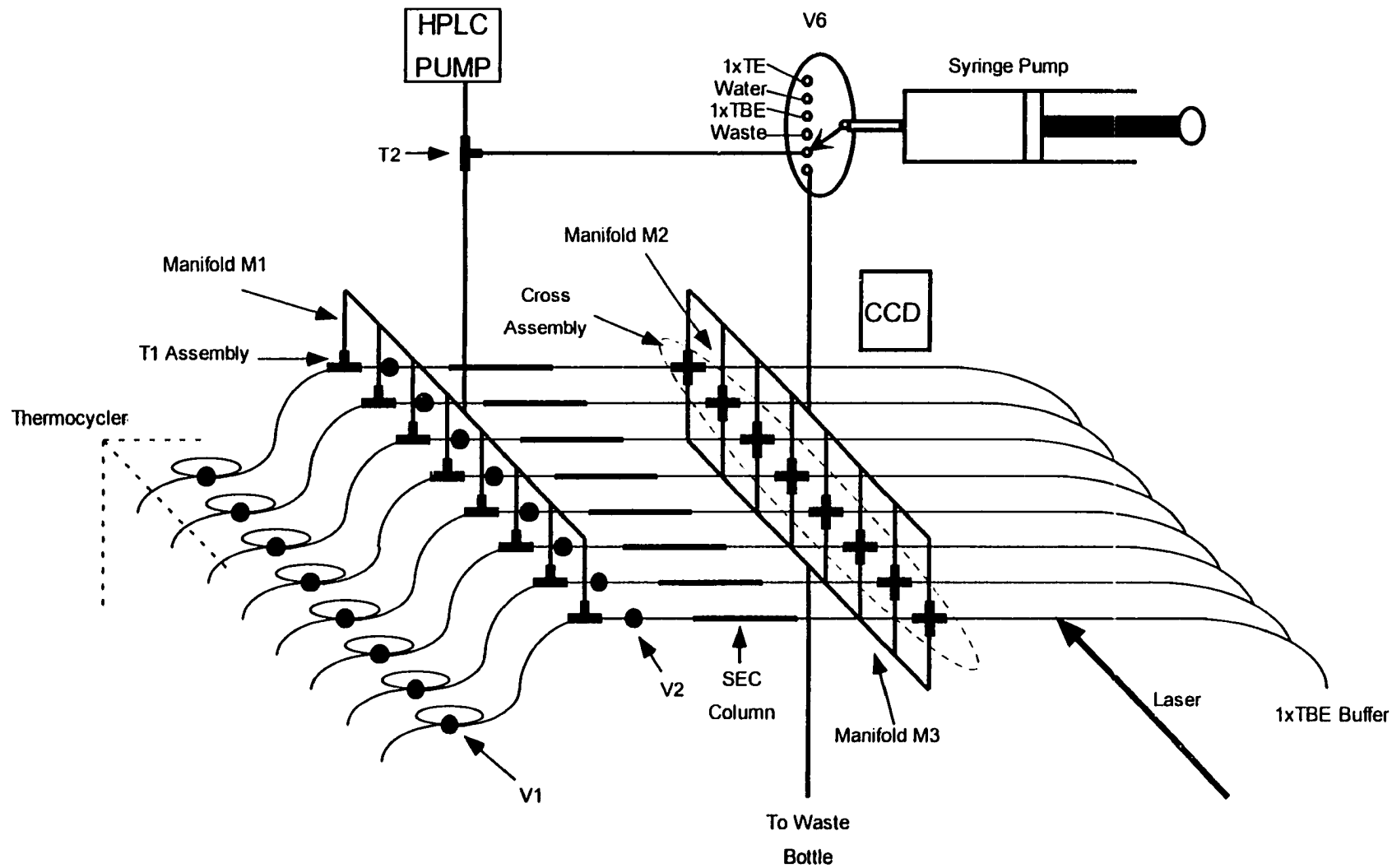


FIGURE A1. Multiplexed on-line sequencer set up

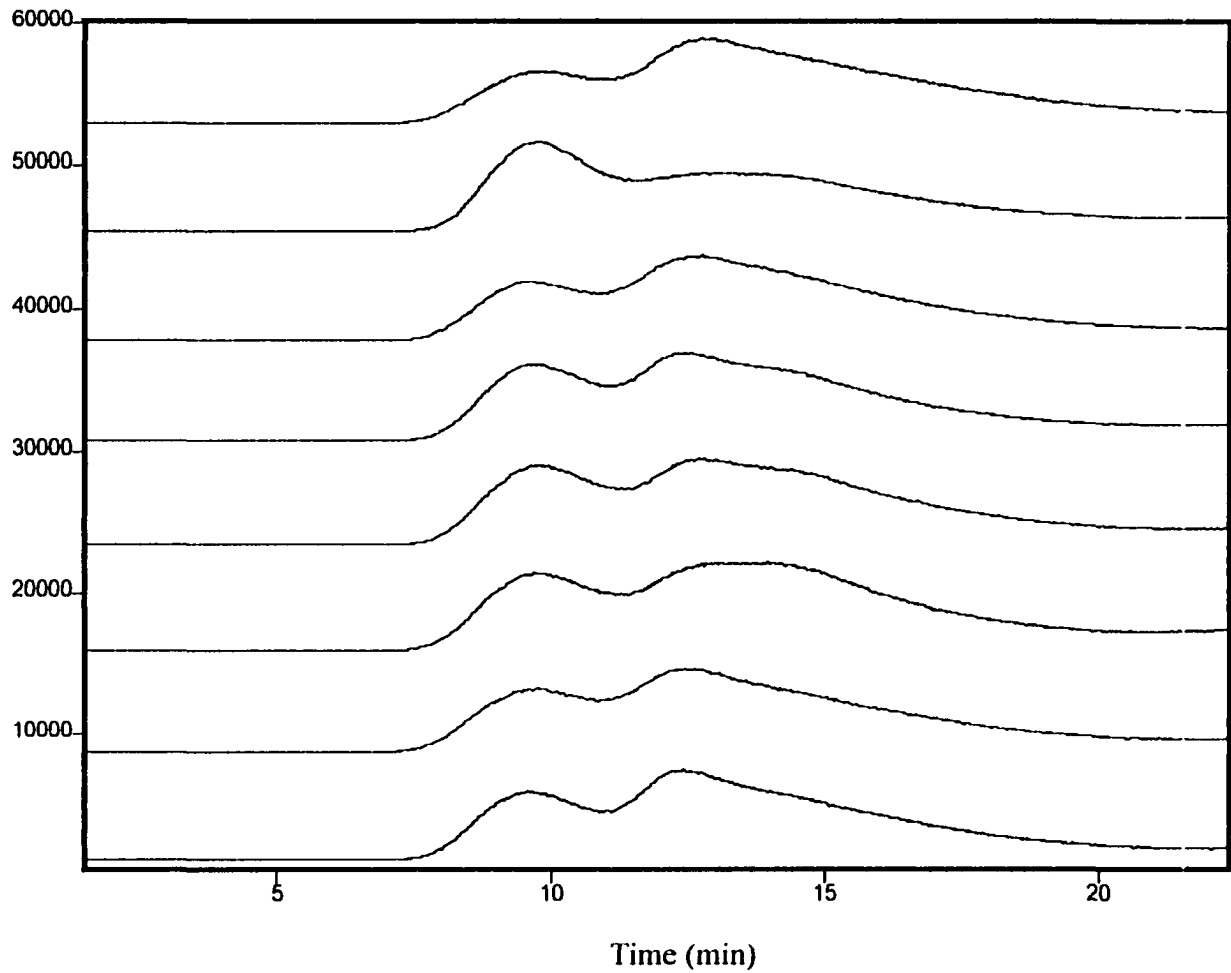


FIGURE A2. Elution profile of DNA ladder produced from cell lysate (first 4 channels) and pure DNA (last 4 channels).

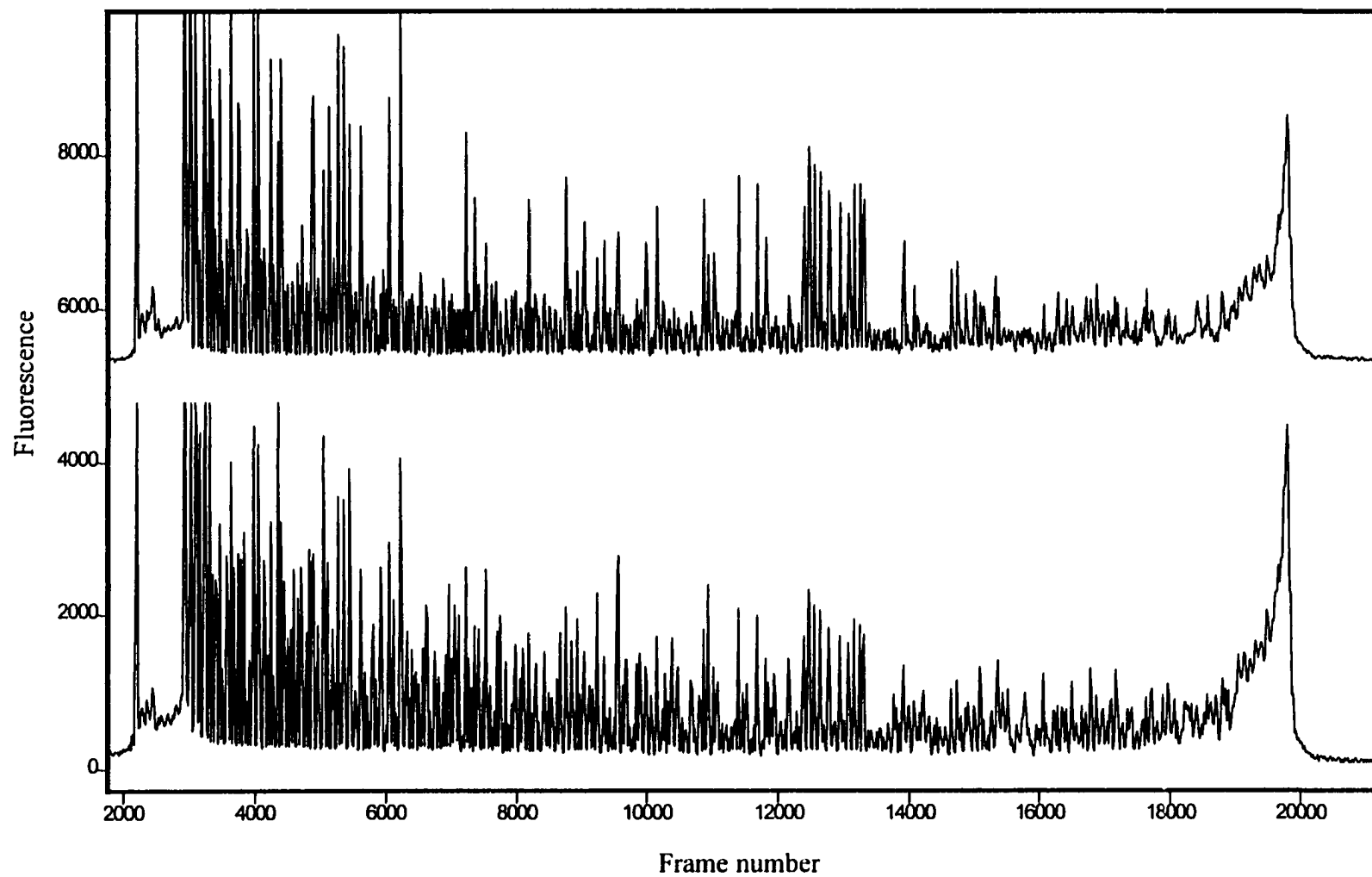


FIGURE A3. The raw data of base calling up to 620 bp from clone. Upper: Blue channel
Lower: Red channel.

APPENDIX B. SUPPORTING INFORMATION FOR CHAPTER 4

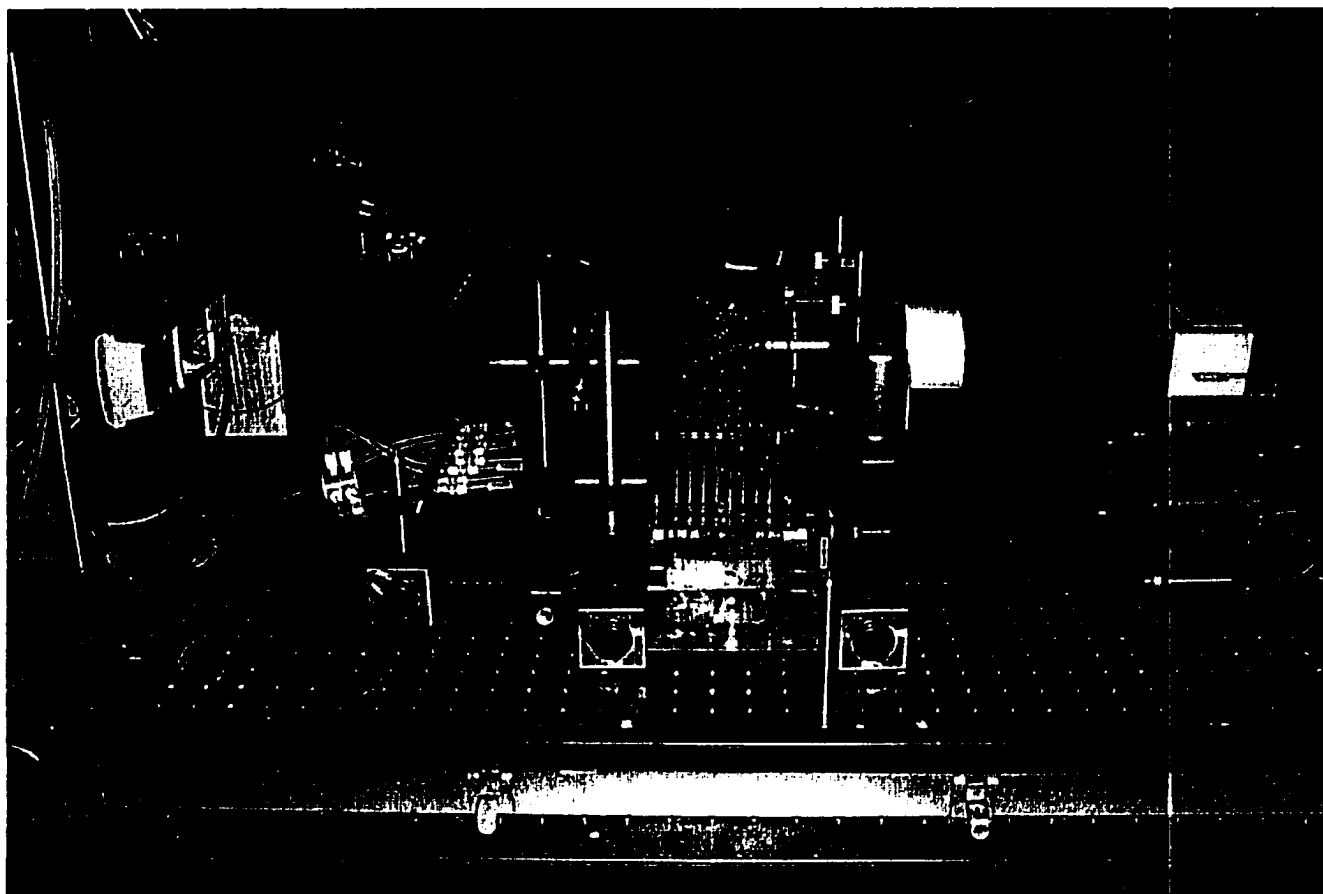


FIGURE B1. Photograph of the 96-capillary array experimental set up with imaging UV absorption



FIGURE B2. Injection ends of capillary array electrophoresis

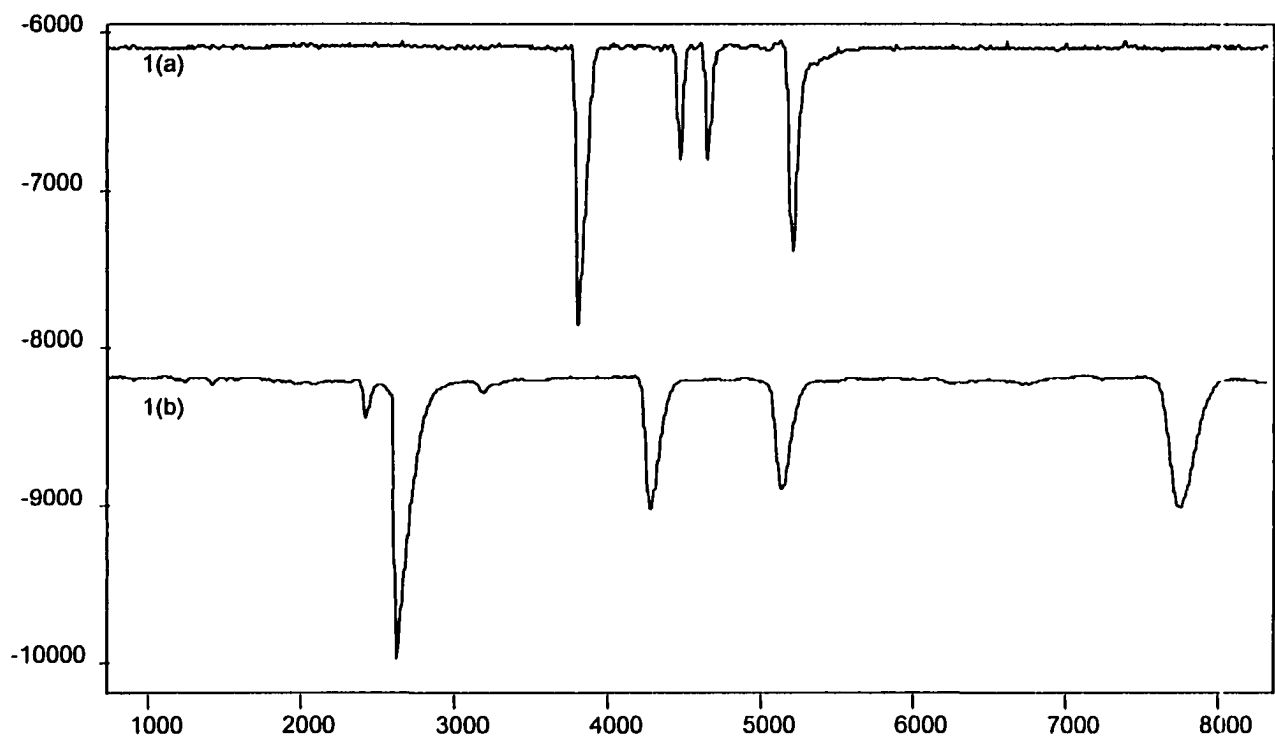


FIGURE B3. Separations of reactant, products (isomer A and B) and internal standard (Norharman) by two different solvents. Buffer, 40mM NH_4Ac , 0.75% formic acid in (a) MeOH (b) 80% DMF with 20% H_2O . Applied electric field, 140V/cm. Column, bare fused-silica capillary with effective/total length of 50/75-cm and 50- μm I.D. Hydrodynamic injection 15s at 8-cm height.

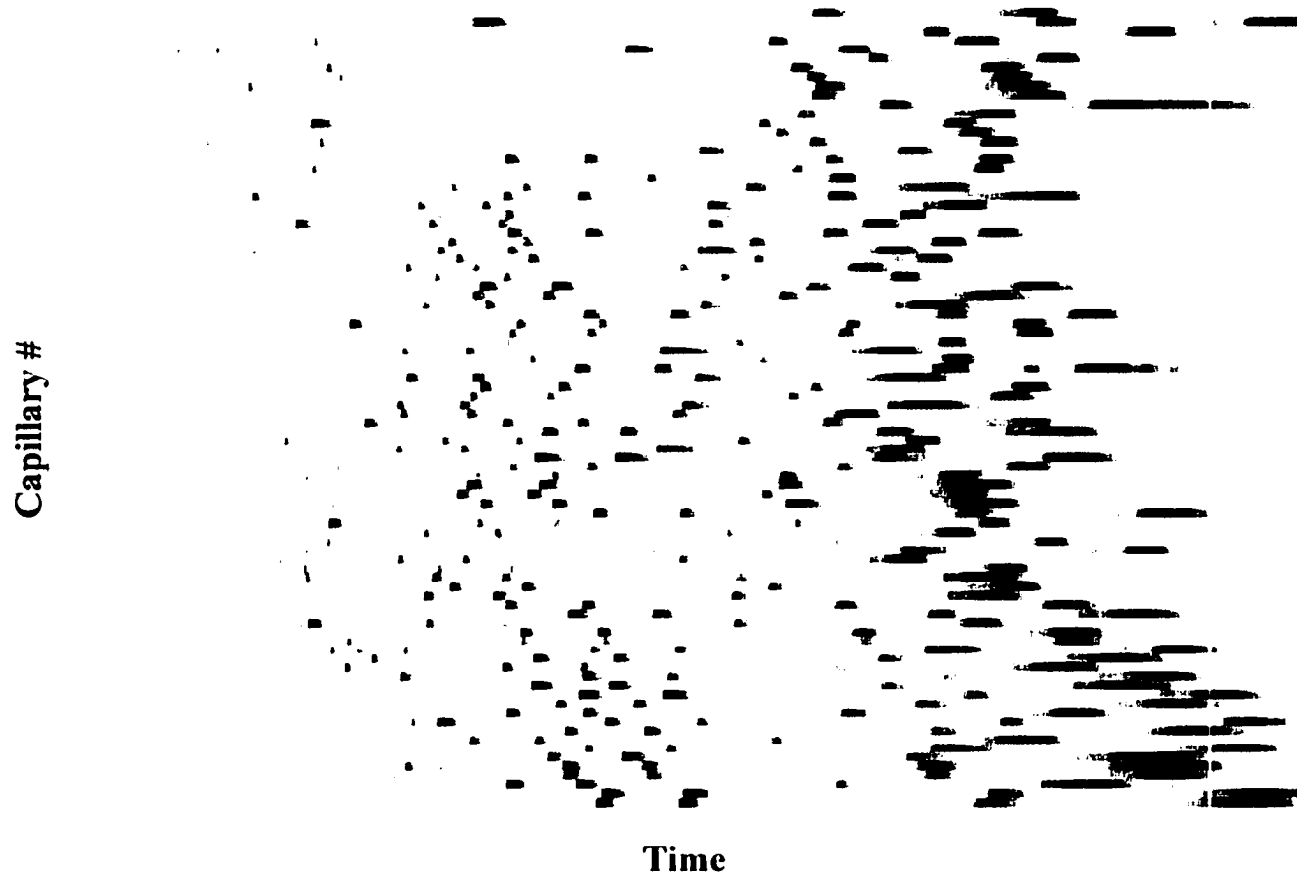


FIGURE B4. Result of CE separation of reaction mixtures in the 96-capillary array. Separation condition as in FigureA1 (b). Hydrodynamic injection 1 min.

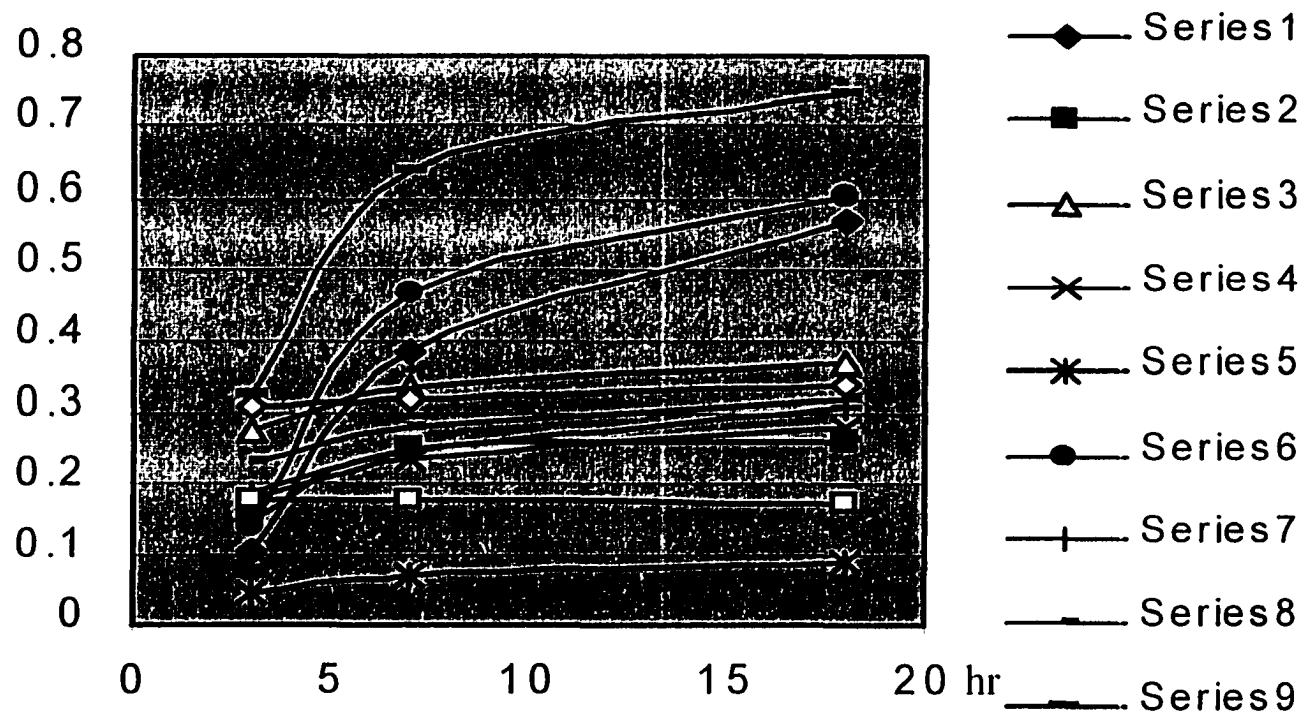


FIGURE B5. The kinetics of reactions using $\text{Pd}(\text{PPh}_3)_4$ as catalyst and 11 different bases. Combination of bases can be found in FIGURE 1, Chapter 4 in this dissertation.

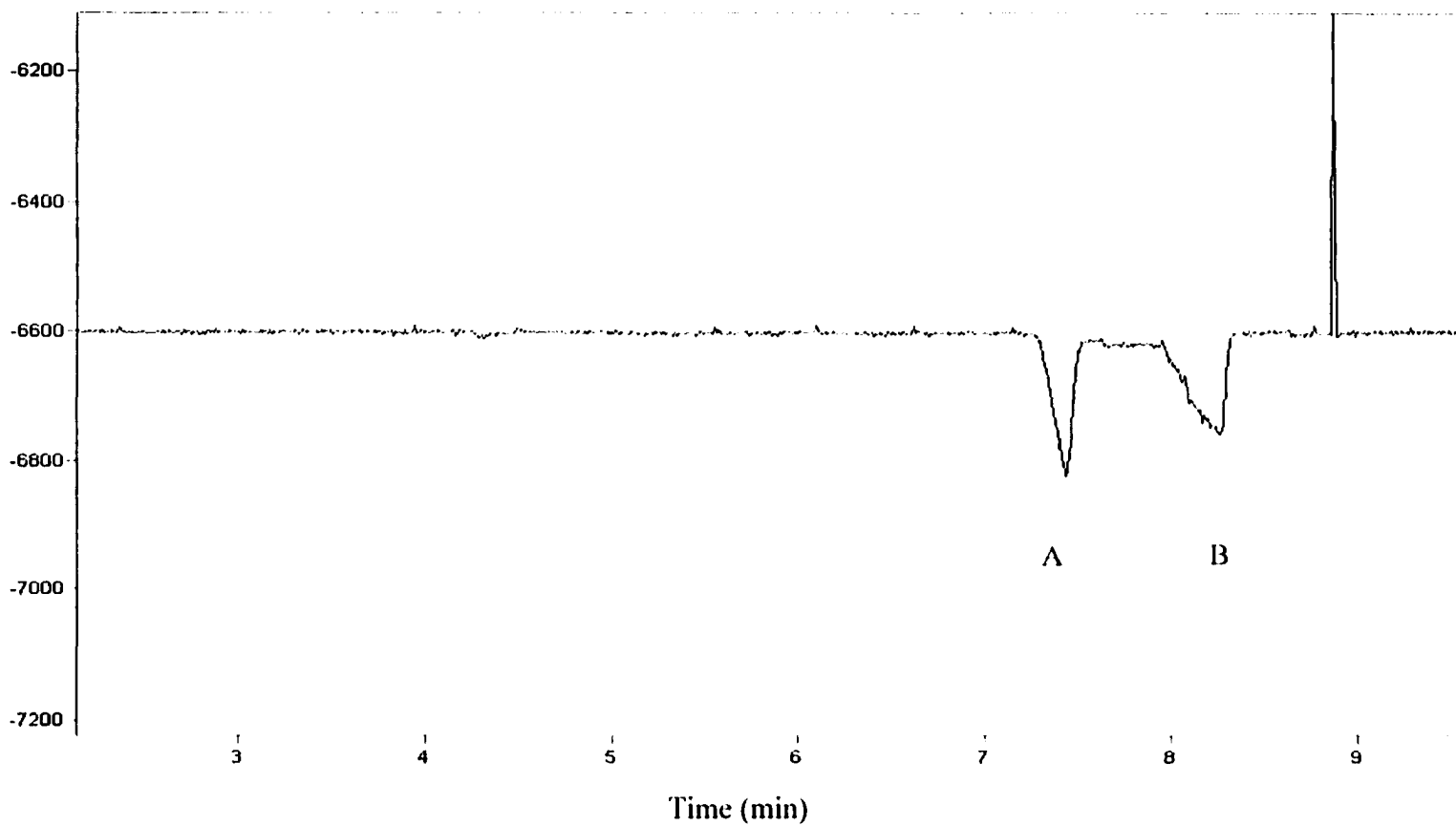


FIGURE B6. CD-MEKC separation of styrene (A) and ethylbenzene (B). Buffer: phosphate and borate, pH=7; 20% Acetonitrile; 15 mM β -CD, 50mM Lithium dodecyl sulfate; hydrodynamic injection: 5 sec at 5 cm height; capillary: 50 μ m I.D. and 365 μ m O.D., 60 cm total with 40cm effective length; 150 V/cm.

ACKNOWLEDGEMENTS

First, I feel indebted so much to Dr. Yeung, my major professor. I would like to thank him for his continuous advice, support and encouragement throughout past four years. His creativity, insight on important research issues, hard-working spirit, knowledge and effective presentation skills have taught me about what comprises a great scientist. The precious training I got from his lab in many-disciplines research projects will benefit my whole career.

Second, I am grateful to my graduate committee, Drs. Dennis C. Johnson, Robert S. Houk, Walter S. Trahanovsky and Richard C. Seagrave, for their willingness to investigate precious time and provide advice whenever I need it.

My sincere thanks go to Drs. R. C. Larock, Xiaoyi Gong, Hongdong Tan, Yan He, Homing Pang, Micheal Shortreed and Yinfu Ma for their valuable suggestions and cooperation on my experiments.

I also want to thank all the Yeung's group members, both the previous and present. David, Hui Su, Nanyan, Wei Wei, Fenglei, Gang Xue, Hanlin, Jason, Michael, Craig, Nicholas, Kang, Park, Jinjian, Wenwan have provided me their help in all aspects. They are always there whenever I have questions. I would like to thank Ms. Vickie Johnson for editing my papers' formats for publication.

Special thanks go to my wife, Xuemei Lu, who always believes in me and puts up with my numerous late nights, working stress with a sweet sunshine smile. I thank her for her encouragement when I was stressed, for her endurance when I was busy on work. I also thank my parents who have provided me the guidance of the first steps of the study and encouraged the later ones. They are the ones who showed me never give up and taught me

always keep optimistic attitude. They brought me to life and never hesitate to give their unconditional love to me. I dedicate this dissertation to them.

This work was performed at Ames Laboratory under contract no. W-7405-Eng-82 with the U. S. Department of Energy.