Christopher P. Fredlake[1]
Daniel G. Hert[1]
Brian E. Root[2]
Annelise E. Barron[1,3]

[1]Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL, USA
[2]Department of Materials Science and Engineering, Northwestern University, Evanston, IL, USA
[3]Department of Bioengineering, Stanford University, Stanford, CA, USA

Received June 4, 2008
Revised August 29, 2008
Accepted August 29, 2008

Research Article

Polymer systems designed specifically for DNA sequencing by microchip electrophoresis: A comparison with commercially available materials

Electrophoresis-based DNA sequencing is the only proven technology for the *de novo* sequencing of large and complex genomes. Miniaturization of capillary array electrophoresis (CAE) instruments can increase sequencing throughput and decrease cost while maintaining the high quality and long read lengths that has made CAE so successful for *de novo* sequencing. The limited availability of high-performance polymer matrices and wall coatings designed specifically for microchip-sequencing platforms continues to be a major barrier to the successful development of a commercial microchip-sequencing instrument. It has been generally assumed that the matrices and wall coatings that have been developed for use in commercial CAE instruments will be able to be implemented directly into microchip devices with little to no change in sequencing performance. Here, we show that sequencing matrices developed specifically for microchip electrophoresis systems can deliver read lengths that are 150–300 bases longer on chip than some of the most widely used polymer-sequencing matrices available commercially. Additionally, we show that the coating ability of commercial matrices is much less effective in the borosilicate chips used in this study. These results lead to the conclusion that new materials must be developed to make high-performance microfabricated DNA-sequencing instruments a reality.

Keywords:
Band broadening / Capillary electrophoresis / DNA sequencing / Microchip electrophoresis / Polymer solution
DOI 10.1002/elps.200800352

1 Introduction

A wealth of biological information was unveiled by the sequencing of the first human genome [1, 2], and subsequent sequencing projects [3, 4] have added to this knowledge to give researchers insights into genome structure, genome function, and human evolution [5–9]. To further understand the basics of our genome, including studying genome diversity at scales larger than the single-nucleotide level, continued large-scale sequencing projects are needed [10, 11]. For example, while cancer may be characterized by the particular tissue where it originates, there does exist genome-scale diversity that must be characterized accurately to more effectively diagnose and treat these diseases [12].

Unfortunately, the costs associated with sequencing multiple individuals or cancerous tissue samples are currently prohibitive, with the cost of producing a copy of a human-sized genome estimated at ~$20 million [13]. Ultimately, technology development drives cost reductions, and DNA-sequencing technology has benefited previously from the invention of capillary array electrophoresis (CAE) instruments to replace the older slab gel methods used for Sanger fragment separation [14–17]. The increased rate of data acquisition by CAE was achieved through automated instrumentation, faster separation times, and parallel sample processing. Through the introduction of automated CAE instruments, determining the full sequence of gigabase-sized genomes became feasible.

While consumables (sample preparation kits, polymer matrices, *etc.*) make up a large fraction of the total sequencing costs at genome-sequencing centers, CAE instrument depreciation is the single largest cost associated with the process [18]. The potential to reduce the bottom line at

Correspondence: Professor Annelise E. Barron, Department of Bioengineering, 318 Campus Drive, W300B James H. Clark Center, Stanford, CA 94305, USA
E-mail: aebarron@stanford.edu
Fax: +1-650-723-9801

Abbreviations: CAE, capillary array electrophoresis; **DMA,** *N,N*-dimethylacrylamide; **GPC,** gel permeation chromatography; **LPA,** linear polyacrylamide; **MALLS,** multi-angle laser light scattering; **pDMA,** poly(*N,N*-dimethylacrylamide)

sequencing centers exists by developing technology that maintains a comparable capital investment to current instrumentation while generating sequencing data at rates much higher than the capabilities of current instruments.

While many alternatives to Sanger-based sequencing are under development [19–21], no other technology has proven capable of producing long and/or accurate enough read lengths to sequence a large and complex genome including large amounts of repetitive DNA. The miniaturization of Sanger sequencing, however, represents an avenue for increased sample throughput while obtaining data that are already usable by the current genome-sequencing infrastructure. Throughput is increased *via* shorter separation times, but overall sequencing costs can also be reduced since these systems also require fewer reagents *per* run and have the potential to be integrated into miniaturized sample preparation platforms (such as those demonstrated by the Mathies group [22, 23]). By obviating instrumentation involved in offline sample preparation, the total capital investment of the entire process is further reduced as well.

Miniaturized sequencing instruments, however, are still under development. While many challenges remain, creating materials for the high-quality separation of sequencing fragments is a problem that still needs to be addressed. Development of the DNA separation media was an important step in optimizing the performance of CAE instruments. The advent of polymers that serve as sequencing matrices and/or channel wall coatings was crucial to the development of robust CAE-sequencing instruments that deliver long and accurate DNA-sequencing read lengths.

Much work was done to optimize linear polyacrylamide (LPA) for capillary electrophoresis-sequencing applications to replace the crosslinked acrylamide gels that had been used in slab gels. In LPA solutions in water, physical entanglements between chains form the network of openings used to sieve the DNA as opposed to the chemical crosslinks that form pores in the gels. LPA is very hydrophilic and forms strongly entangled networks that have been shown to be excellent for high-quality DNA separations, with the Karger laboratory demonstrating that long sequencing read lengths (up to 1300 bases) can be obtained with highly optimized LPA-matrix formulations and highly pure DNA-sequencing samples [24]. Additionally, Doherty *et al*. showed that by adding very small amounts of crosslinking monomers back into the polymer (the polymer solution still flowed) to form stabilized LPA ''nanogels,'' sequencing performance could be improved compared with LPA synthesized in the absence of any crosslinker [25, 26].

LPA can also be used as a polymer wall coating for channel surfaces by using the method developed by Hjerten [27]. This method creates covalently linked polymer chains on the surface of the channels to form the coating. The synthesis of the polymer coating is performed within the capillary and requires up to 10 h to complete. This procedure can lead to capillary clogging and the formation of non-homogeneously coated surfaces making the use of these coatings non-ideal, although they have been previously shown to allow high-quality DNA-sequencing read lengths in some cases.

In addition to LPA, the acrylamide derivative poly(*N,N*-dimethylacrylamide) (pDMA) has also been demonstrated as a high-quality polymer for sequencing DNA [28–30]. Sequencing matrices using pDMA generally have lower viscosities than LPA matrices and are also able to form adsorptive, or dynamic, coatings on the capillary surfaces, eliminating the need to perform chemical reactions at the capillary surface prior to polymer matrix loading and DNA fragment separation. Dynamic coatings are very advantageous as alternatives to covalent coatings because these coating can be applied relatively quickly (45 min or less) and are extremely stable [31].

While the dedicated development of polymer materials for DNA sequencing such as LPA and pDMA have been mainly completed in CAE devices, research focusing on the design of microfabricated systems for DNA have tended to focus on instrumentation while giving less attention to matrix and wall coating development [23, 32–38]. While instrumentation and integration challenges have been great, success in this field has been limited in part due to the assumption that the separation of sequencing fragments in microfluidic devices is not much different than in CAE instruments. Specifically, it has been assumed that the materials used in commercial CAE instruments can be plugged directly into microchip platforms with little or no change in sequencing performance. In previous work [39], we have demonstrated that polymer systems (comprising both a matrix and wall coating) developed exclusively using a microchip-sequencing instrument have delivered long sequencing read lengths (up to 600 bases) in short times (~6–7 min). This single-channel throughput is over three times higher than in previously published microchip-sequencing reports, which generally used polymer matrices and coatings based on formulations that had been developed for CAE instruments. In this paper, we show that commercially available CAE polymer matrices deliver much shorter read lengths in microchips than polymer matrices developed exclusively for microchip-sequencing platforms, and discuss the need for these new matrices and coatings to speed the development of microfabricated platforms for DNA-sequencing analysis.

2 Materials and methods

2.1 Polymer synthesis

Polymers were synthesized by aqueous-phase free-radical polymerization. The *N,N*-dimethylacrylamide (DMA) monomer (Monomer-Polymer & Dajac Labs, Feasterville, PA, USA) was polymerized by dissolving monomer into distilled and deionized water at a concentration of 4% w/w.

Additionally, isopropanol can be added to the solution to control the molar mass of the polymer. The DMA solution was placed in a jacketed flask where the temperature was controlled at 47°C by a re-circulating water bath. Oxygen was removed from the system by bubbling nitrogen gas through the solution for 30 min. The polymerization reaction was initiated using V-50 (2,2′-azobis(2-amidinopropane) dihydrochloride, Wako Chemicals, Richmond, VA, USA), and the reaction was allowed to proceed for 6 h.

For synthesis of the coating polymer, *N*-hydroxyethylacrylamide monomer (Cambrex, East Rutherford, NJ, USA) was dissolved at 0.5% w/w and prepared similar to the DMA solution with the re-circulating water bath set at a temperature of 25°C. The *N*-hydroxyethylacrylamide polymerization reaction was initiated using ammonium persulfate (APS, Amresco, Solon, OH, USA) and TEMED (Amresco). Purification of all synthesized polymers included dialysis using Spectrum Spectra/Por cellulose ester membranes with a molar mass cutoff of 100 000 Da against distilled and deionized water for 2 wk with frequent water changes. After dialysis, dry polymer was recovered by lyophilization.

2.2 Polymer molar mass characterization

Polymer molar mass distributions were determined using tandem gel permeation chromatography (GPC) (Waters, Milford, MA, USA)-multi-angle laser light scattering (MALLS) (Wyatt Technology, Santa Barbara, CA, USA). Polymer solutions were prepared at a concentration of 1.0 mg/mL in the GPC aqueous mobile phase (0.1 M NaCl, 50 mM NaH_2PO_4, and 200 mM NaN_3). The samples were fractionated on the GPC using Shodex (New York, NY, USA) OHpak columns SB-806 HQ, SB-804 HQ, and SB-802.5 HQ were connected in series. After fractionation, the samples flowed to the DAWN EOS MALLS instrument and then to the Optilab DSP refractive index detector. Data were processed using the ASTRA software from Wyatt Technology using the assumption of 100% mass recovery from the injected sample and the known instrument calibration constant. Light scattering data were analyzed using the Berry method with a first-order fit for the molar mass and coil radius distributions.

2.3 Rheology

The zero-shear viscosity of polymer matrices was determined using an Anton Paar Physica MCR 300 (Ashland, VA, USA) rheometer with the temperature maintained by a Peltier controller connected to a digitally controlled re-circulating water bath (Julabo USA, Allentown, PA, USA). Controlled shear stress and shear rate sweeps were performed with a cone-and-plate (model CP50-1) fixture and a double gap Couette (model DG26.7) fixture.

2.4 Application of polymer dynamic coating

Prior to adsorption of polymer chains onto the channel surface, a 1 M HCl solution was loaded into the channels and allowed to contact the surface for 15 min. Following acid treatment, the channels were flushed with water. A 0.1–0.25% w/v solution of the coating polymer was then introduced into the channels and allowed to contact the surface for 15 min. The polymer solution was then removed and water was flushed through the channels.

2.5 DNA sequencing by CAE

All DNA-sequencing separations by CAE were performed using the Applied Biosystems (ABI) 3100 Genetic Analyzer. Two separate 22 cm capillary arrays were used in these experiments. One array was dynamically coated with poly(*N*-hydroxyethylacrylamide) (pHEA) using the method described in Section 2.4 (this was performed ''offline'' to avoid problems with introducing 1 M HCl into the ABI 3100 instrument). A pDMA matrix formulated in our laboratory and ABI's POP-6™ media were used as sequencing matrices in order to compare matrix performance in this instrument using the pHEA-coated array. A second, uncoated array was used with POP-6™ (which coats the array upon loading) to compare sequencing read lengths obtained with the two different dynamic coatings. For these experiments, ABI Big Dye v3.1 four-color sequencing standard was used as the sequencing sample. Prior to sample injection, pre-electrophoresis was performed with an electric field of 250 V/cm. Sample was injected electrokinetically under an applied field of 45 V/cm for 22 s. Four-color DNA-sequencing separations were carried out at 55°C under an electric field strength of 250 V/cm. Basecalling was performed by the NNIM BaseCaller (NNIM, LLC, Salt Lake County, UT, USA) and Sequencher v 4.0.5 (Gene Codes, Ann Arbor, MI, USA).

2.6 DNA sequencing by microchip electrophoresis

Analysis of M13 sequencing fragments (Amersham Biosciences, Piscataway, NJ, USA) was carried out on a home-built microchannel electrophoresis system that has been described previously [40]. This system allows sensitive multi-color detection through LIF. The system comprises two subsystems: an electrical system that supplies voltage to the microfluidic device and an optical subsystem that allows detection of fluorescent molecules as they pass through a focused laser point from an 488-nm Argon-Ion laser. These two subsystems can both be controlled using a single program written in LabView software.

Experiments were conducted using single-channel borosilicate glass microchips (chip T8050) purchased from Micronit Microfluidics BV (Enschede, The Netherlands). These chips have an effective separation length of 7.5 cm.

Channels were dynamically coated with pHEA as described in Section 2.4. Separations of sequencing fragments were carried out in pDMA synthesized in our laboratory, which was dissolved in 1 × TTE (49 mM Tris, 49 mM TAPS, and 2 mM EDTA) and 7 M urea. POP™ matrices from ABI (POP-5™, POP-6™, and POP-7™) and Beckman LongRead™ sequencing matrix were used as purchased. For each run, pre-electrophoresis at 235 V/cm electric field was performed for 60 s prior to sample injection. The sample was injected for 40 s at 400 V/cm using an offset T injector with a 100 μm offset. Separation was carried out at 235 V/cm with 150 V/cm back biasing applied to the sample and sample waste wells to ensure that no excess DNA migrated into the separation channel. The chip was maintained at 50°C for the pre-electrophoresis, sample injection, and fragment separation steps using a copper plate interfaced with a temperature controller. Basecalling was completed using NNIM Basecaller and Sequencher v 4.0.5.

3 Results and discussion

3.1 Polymer synthesis and characterization

The physical and chemical properties of polymer chains greatly influence the properties of both sequencing matrices and channel wall coatings [24, 31, 41]. The average molar mass of a polymer is critical for a polymer matrix to deliver longer read lengths and for polymers physically adsorbed to the channel surface to produce a robust and stable coating. Generally, the average molar mass must be in excess of 1 MDa for both applications, although higher average molar masses in separation matrices have been shown to lead to better sequencing performance [42, 43].

The physical properties of the polymers synthesized for this study as measured by GPC-MALLS are given in Table 1. Careful control of various parameters of the synthesis such as the monomer concentration, reaction temperature, and initiator concentration is required to produce materials with high molar mass. Additionally, isopropanol can be added to the reaction in order to reduce the average molar mass of the synthesized polymers [28]. These lower molar mass polymers can be included in matrix formulations for increasing sequencing read lengths. All polymers were extensively purified after synthesis by dialyzing the reaction products for 2 wk against distilled and deionized water followed by lyophilization. The polymer matrix formulations were made by dissolving the sequencing polymer into the 1 × TTE+7 M urea buffer.

Table 1. Physical properties of polymers synthesized for this study as measured by GPC-MALLS

Polymer	M_w (MDa)	R_g (nm)	PDI
pDMA	5.9	98	1.5
pDMA[a)]	0.28	31	1.9
pHEA	3.0	98	1.5

a) Isopropanol added to reaction.

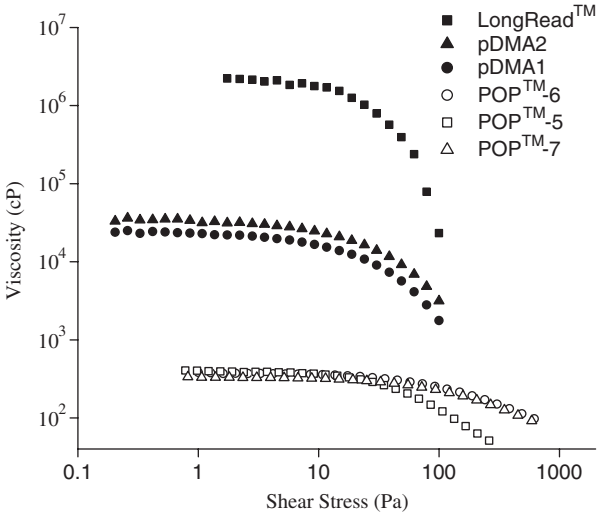

Figure 1. This plot shows the change in viscosity for the different matrices used in this study with shear stress. While viscosities can be high for some of the matrices at low shear stress, the viscosity can often drop by an order of magnitude or more at higher shear stresses making filling the microchip easier.

Table 2. Zero-shear viscosity and channel loading times for sequencing matrices

Matrix[a)]	Zero-shear viscosity (cP)	Chip loading time[b)] (min)
POP-5™	329	0.26
POP-6™	374	0.27
POP-7™	395	0.13
PDMA1	24 000	1.16
PDMA2	34 700	2.09
LongRead™	2 138 000	1.88

a) pDMA1 and pDMA2 were formulated in our laboratory using polymers from Table 1.
b) Time for matrix to fill chip with total length 8.5 cm at 190 psi.

The polymer matrices are entangled polymer solutions that can have a wide range of viscosities, which depend on both polymer molar mass and concentration as well as the polymer chemistry. The viscosities at different shear stresses are shown in Fig. 1. The matrices exhibit shear-thinning behavior typical of entangled polymer solutions [41]. The zero-shear viscosities along with the loading time of the matrices into the 8.5-cm long microchip channels at 190 psi are given in Table 2. The viscosity of these solutions can also be affected by the presence of the salts and other additives in the buffers. Urea, which is used as a denaturant in our matrices, is known as a solution thickener although its

contribution to the matrix viscosity is much lower than the contributions from the polymer chain entanglements. The viscosities of these solutions range over several orders of magnitude from ~400 cP for the POP^{TM} polymers to over 2 million cP for the LongRead^{TM} LPA matrix. Two of our pDMA formulations have viscosities at an intermediate value between the POP^{TM} matrices and the LongRead^{TM} matrix. With higher matrix viscosities, in general, the filling of the microchips takes longer time. However, shear-thinning behavior is more pronounced in the LongRead^{TM} matrix as shown in Fig. 1, so that the loading times are the same as the pDMA1 and pDMA2 matrices (1–2 min at 190 psi), even though the zero-shear viscosity is two orders of magnitude larger.

3.2 DNA sequencing by CAE

Previously, we have demonstrated that increased read lengths can be obtained in sequencing matrices where two pDMA polymers with different average molar masses are blended into a matrix relative to a matrix where just a single average molar mass polymer is used to formulate the matrix [39]. Blends were formulated using a 3:1 and a 3:2 ratio of high molar mass polymer to low molar mass polymer with total polymer concentration of 4 and 5% w/v, respectively. Here, we evaluate the sequencing performance of the 4% w/v blended pDMA matrix on the ABI 3100, a commercially available CAE-sequencing instrument, and compare the read lengths with those obtained in ABI's POP-6^{TM} matrix. In these experiments, the hydrophilic polymer pHEA was used to dynamically coat the capillaries in the array. In order to coat the capillaries with pHEA, a new array was removed from the instrument and connected to a syringe pump. Using the pump, acid was first flushed through the capillaries, followed by pumping the dilute polymer coating solution through the capillaries (the full method is described in Section 2).

In the pHEA-coated array, separation of DNA-sequencing fragments was performed with both the 4% pDMA and the POP-6^{TM} matrices. Figure 2 shows unprocessed sequencing electropherograms from typical runs in both of the matrices. The separation is much faster in the 4% pDMA formulation, with the last of the sequencing fragments eluting from the capillary in 22 min, compared with the 60 min required in POP-6^{TM}. The mobility of the DNA fragments is a function of the polymer matrix itself, and is not affected by the coating. This was confirmed by separating DNA-sequencing fragments in a second array in which POP-6^{TM} was used as both the separation matrix and the dynamic wall coating. For this second capillary array, the POP-6^{TM} polymer was loaded into the bare capillaries prior to sample injection and separation, which is the normal operating procedure for the instrument. Figure 3 shows the unprocessed fluorescence data for DNA fragment in a typical run for separations in POP-6^{TM} using both the pHEA- and POP^{TM}-coated arrays. The required run time

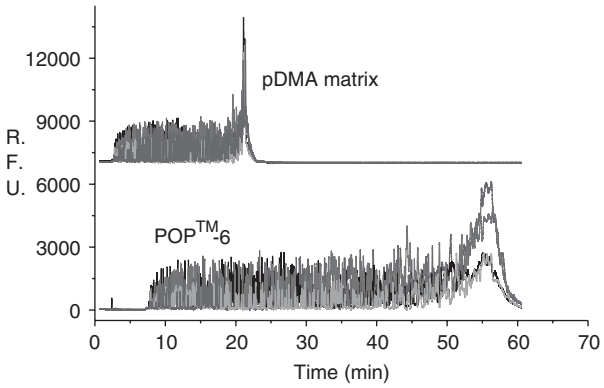

Figure 2. These unprocessed sequencing electropherograms show that the injected DNA-sequencing sample elutes in much shorter times from the pDMA matrix formulated in our laboratory than in POP-6^{TM} in the ABI 3100. For these experiments, the electric field was set at 250 V/cm and the temperature was set at 55°C. The capillaries had an effective length of 22 cm and were dynamically coated with pHEA polymer prior to matrix loading and DNA injection. The DNA in the pDMA matrix elutes in ~22 min while the DNA takes 60 min to move through the POP-6^{TM} matrix.

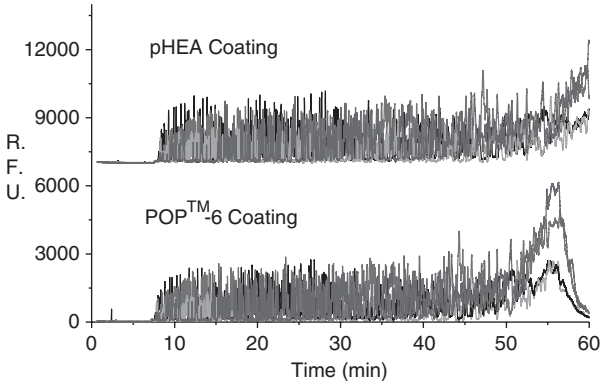

Figure 3. These unprocessed electropherograms show the effect of dynamic polymer coating on DNA elution time in the ABI 3100. In 22-cm capillaries, dynamically coated with either pHEA or POP-6^{TM}, DNA elution time is 60 min through the POP-6^{TM}-sequencing matrix. Thus there is no effect of the coating polymer on DNA speed through the matrix. Other experimental conditions are the same as in Fig. 2.

was 60 min for separations in arrays used with either coating polymers.

Sequencing read lengths in the two matrices using the pHEA-coated capillaries are presented in Table 3, with all read lengths given at 98.5% accuracy. POP-6^{TM} provided longer read lengths on average (677 bases) as well as the longest read of 757 bases, but the 4% pDMA delivered only slightly shorter read lengths on average (650 bases), and had a lower standard deviation of the data. The longer read lengths in POP-6^{TM} derive, in part, from better separation of the smaller DNA fragments. Figure 4 shows the first 70 called bases in a typical electropherogram

Table 3. Comparison of sequencing matrices on ABI 3100 with pHEA dynamically coated capillary arrays

Polymer	Average read length[a] ($n = 4$)	Long read length[a]	Time[b] (min)
4% pDMA[c]	650 ± 10	660	22
POP-6™	677 ± 68	757	60

a) All read lengths given at 98.5% accuracy.
b) Time required for last DNA sequencing fragment to elute from column.
c) 3% high molar mass (3.4 MDa) + 1% low molar mass (280 kDa).

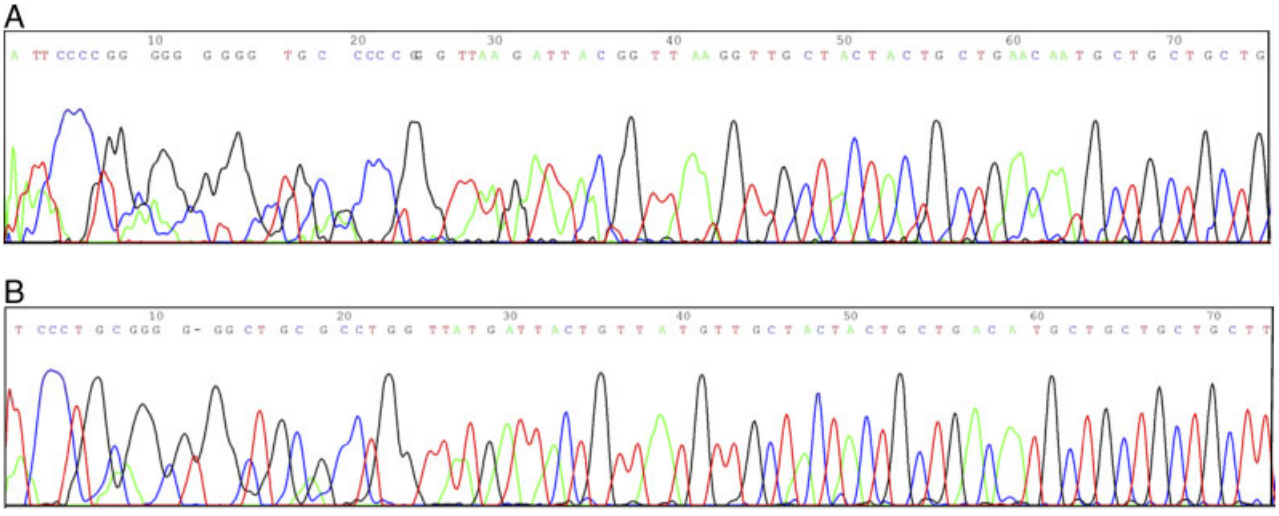

Figure 4. Sequencing traces for small DNA fragments through (A) the pDMA matrix and (B) the POP-6™ matrix on the ABI 3100 show the different abilities of the matrices to resolve DNA fragments in this size range. The double T peaks around base 31, 37, and 42 in the POP-6™ matrix (trace b) are better resolved than the same peaks in the pDMA trace. Note that in the pDMA trace, these same peaks are 28, 33, and 39 because basecalling begins later in this matrix due to poorer resolution. Peaks are better resolved to baseline even to the end of the shown sequence (~base 70).

for each of the two matrices after basecalling. It can be seen that the peaks eluting from the 4% pDMA matrix (Fig. 4A) are more poorly resolved than peaks eluting from the POP-6™-filled capillaries (Fig. 4B) in this DNA fragment size range.

POP-6™ might be expected to perform better in these experiments since this sequencing matrix was developed specifically for the ABI instrument. Both the data processing software and the sample injection conditions are tuned for the POP™ matrices. In our analysis of the data, we used a basecaller (from NNIM, LLC) that was not developed for any one particular matrix or system (we did not use ABI's basecaller) so that specific adjustments for POP™ polymers were minimized. However, the injection conditions programmed into the instrument depend largely on the DNA mobility in the matrices, and the voltage and injection time were not changed from the conditions in the standard POP-6™ sequencing program that was used to run the separations on the instrument. Because DNA mobility is higher in the 4% pDMA matrix, injection plugs were most likely wider than in POP-6™ for these injection conditions and could lead to lower read lengths relative to a specifically optimized injection protocol for the pDMA matrix.

Even though read lengths were slightly lower, it may be beneficial to use the 4% pDMA if one were to choose between using one of these two matrices in this instrument since the separation time was much faster. The sequencing throughput is approximately three times faster as shown in Fig. 2. Therefore, with the 4% pDMA matrix, 1950 bases *per* capillary can be acquired in about 1 h of electrophoresis, while only a single separation run can be completed with the POP-6™ in that same hour, yielding 677 bases *per* capillary.

We also compared sequencing performance using either the pHEA coating or the POP-6™ coating when POP-6™ was used as the separation matrix, as given in Table 4. The average read lengths and the longest read length using the POP™ coating are lower than the read lengths in pHEA-coated capillaries. The pHEA is a more hydrophilic polymer coating than POP™, resulting in a reduction of interactions between the wall and the ssDNA fragments than would be observed in a more hydrophobic coating. While the pHEA coating slightly increases read lengths, the application of the coating requires a more time-consuming offline coating procedure than the application of the POP™ coating, which only requires the capillaries to be

Table 4. Sequencing performance of POP-6™ matrix with different dynamic polymer coatings

Coating	Average read length[a] ($n = 4$)	Long read length[a]	Time[b] (min)
pHEA	677 ± 68	757	60
POP-6™	633 ± 42	692	60

a) All reads given at 98.5% accuracy.
b) Time required for last DNA sequencing fragment to elute from capillary.

filled with matrix polymer. The offline coating method, however, is very simple to perform.

3.3 Comparison of polymer matrices on microchip-sequencing system

While the sequencing performance of our pDMA formulations and a commercially available POP™ matrix was comparable on the ABI 3100, the main goal of the study was to compare sequencing performance on a microchip electrophoresis system. The microchip separations were performed on a system that was custom-built in our laboratory, and separations in all matrices were performed at the same temperature, electric field strength, and separation distance. Additionally, the injection conditions were held constant for every matrix. Because of the offset T injection scheme employed in the chips [44], the sample injection is not greatly affected by the choice of separation matrix. In these injections, the initial sample plug width is roughly defined by the length of the offset between sample and waste arms in the chip, which is approximately 100 μm for the Micronit chips used in these experiments. We formulated two pDMA matrices, matrices pDMA1 and pDMA2, from the synthesized polymers described in Section 3.1, and compared sequencing performance of these two matrices with the commercially available sequencing matrices POP-5™, POP-6™, and POP-7™ from ABI and the LongRead™ matrix from Amersham/GE Healthcare.

The average read length, longest read length, and sequencing time for the longest read length for these matrices are reported in Table 5. The sequencing electropherogram from the longest read in the pDMA2 matrix is shown in Fig. 5. Both custom formulations, pDMA1 and pDMA2, deliver average read lengths of 570 bases, which is 150 bases or more greater than any of the commercial matrices. The three POP™ matrices perform similar to each other with average read lengths ranging from 380 to 420 bases. The POP™ matrices never gave a read length greater than 500 bases, with the longest read of 460 bases occurring in POP-7™. The Beckman LongRead™ matrix delivered the shortest reads of all the matrices with average read lengths less than 300 bases and with no reads longer than 320 bases.

Table 5. Comparison of sequencing read lengths and sequencing times between custom pDMA formulations and commercially available matrices[a]

Sequencing matrix	Average read length[b] ($n = 5$)	Longest read length[b]	Time[c] (min)
pDMA1	576 ± 10	583	6.7
PDMA2	573 ± 16	591	6.9
POP-5™	378 ± 13	378	6.1
POP-6™	417 ± 16	440	8.0
POP-7™	434 ± 16	454	4.5
LongRead™ LPA	< 300	318	5.2

a) Conditions: temperature: 50°C; electric field: 235 V/cm; separation length: 7.5 cm; channel coating: pHEA; current: ~3 μA.
b) Read lengths given at 98.5% accuracy.
c) Time is given for average read length except for Long Read™ matrix, for which the time for the longest read length is given.

The results presented in Table 5 also show the sequencing times for the longest read lengths. This is the time required for the final called base to move from the injection cross to the detection point. It is difficult to make comparisons between matrices from these times because the read lengths are all different. Figure 6 shows the migration velocity for the 300-, 425-, 575-, and 700-base fragments in the matrices. There is a wide range of DNA fragment velocities between the matrices with the DNA peaks having the highest velocity in POP-7™ and the lowest velocity in POP-6™. The two pDMA matrices have values in between these two POP™ matrices.

Without optimizing other parameters, an increase in resolution and read length can be achieved by increasing the channel length. Thus, to achieve comparable read lengths in the commercial matrices to the read lengths achieved in the pDMA matrices at this channel length, longer separation distances will be needed resulting in longer sequencing times. For dispersion limited systems, the resolution increases as the square root of the separation length, while the separation time increases linearly with the length assuming a constant applied electric potential over the course of the separation [45]. Thus, to improve resolution (or read length) by increasing the separation distance by some factor, the separation time will be increased as the square of that factor. For example, in the microchip used in this study with a 7.5-cm separation distance, the pDMA1 matrix sequenced ~31% more bases than POP-7™. For POP-7™, the extrapolated separation distance required to achieve 570 bases would be 12.9 cm, requiring approximately 8.6 min compared with 6.6 min in the pDMA1 matrix used in the 7.5 cm chip, resulting in a 32% increase in separation time. Over 24 h of electrophoresis time, then, 218 runs in the pDMA1 matrix *per* separation channel would be possible *versus* 167 runs in POP-7™, resulting in greater than 29 000 more bases sequenced *per* channel over that time period.

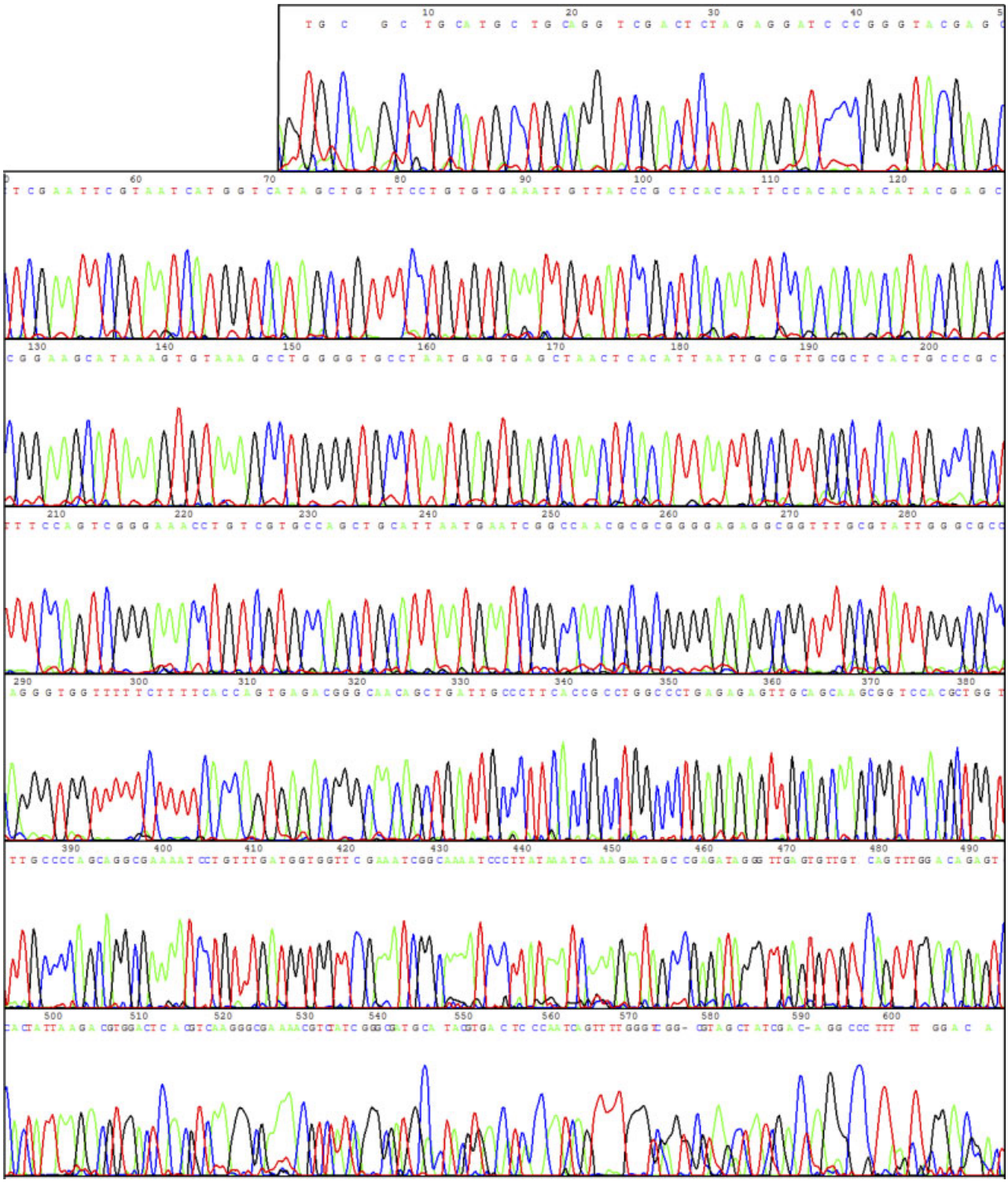

Figure 5. This processed sequencing electropherogram for the pDMA2 matrix on the microchip provides a read length of 591 bases at 98.5% accuracy. This trace and all microchip-sequencing experiments were run at a temperature of 50°C, an electric field of 235 V/cm, separation distance of 7.5 cm, and pHEA dynamic wall coating.

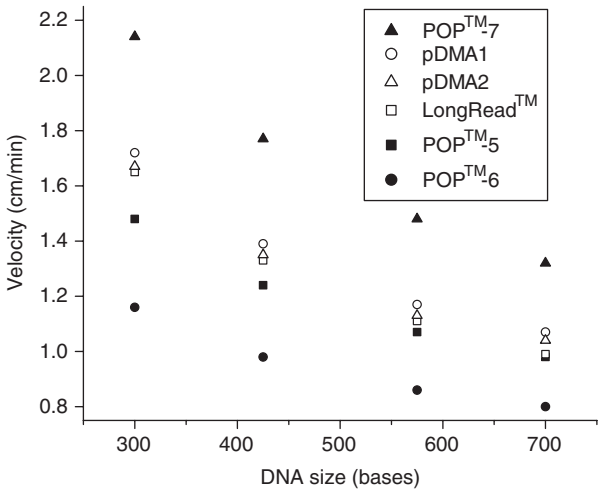

Figure 6. The speeds of ssDNA fragments are shown for the different matrices used for sequencing in this study. All conditions are identical to those in Fig. 5. Since the DNA velocity is assumed to be constant throughout the separation, the required sequencing time can be calculated for a given separation distance, assuming that the distance is long enough to resolve DNA fragments differing in size by one base.

3.4 Comparing pHEA and POP™ coatings in microchips

In addition to comparing the sequencing performance of the matrices, read lengths in one of the commercial matrices, POP-5™, were determined using both the pHEA and POP™ dynamic coatings. As discussed above, the POP™ polymers are sold as self-coating matrices such that when the matrices are loaded into the chip the polymer should dynamically coat the channels. All POP™ matrices are used as self-coating matrices on the ABI instrument as discussed earlier, and for these experiments, POP-5™ was loaded into an uncoated Micronit chip for sequencing fragment separation. Table 6 presents the read lengths obtained in the POP-5™ matrix both when POP-5™ itself is used as the coating and when the chip is pre-coated with pHEA. The results show that POP-5™ is completely ineffective as a self-coating matrix in these chips, whereas the application of pHEA prior to matrix loading leads to read lengths close to 400 bases.

Glass chemistries in the chips and capillaries used in this study are different. The capillaries are made from fused silica while the glass used for the chips contain a significant amount of salts such as sodium borate. In the chip fabrication process, the salts lower the glass melting temperature so that bonding of the chips to the cover plate can be completed at lower temperatures. The presence of salts and other impurities can affect the interaction between the glass surface and a polymer, potentially altering its efficacy as a dynamic coating. Wall coatings compatible with microchip glass chemistries are obviously required for robust sequencing performance in these chips. Borosilicate glass chips are easier to produce and thereby cheaper than fused-silica chips, and are very commonly used for many microchip separations. The issue of chip/coating compatibility must be kept in mind as other glass chemistries, such as soda lime, are introduced, and as plastic chips are further developed as DNA separation and sequencing substrates [37, 38, 46]. Depending on chip material chemistry, microchip surface coatings will have to be developed specifically to the chip substrate that is used, and current commercially available materials may not be the best solution.

Table 6. Comparison of sequencing performance using different dynamic coatings on borosilicate chips with POP-5™ sequencing matrix

Coating	Average read length[a] ($n = 5$)	Longest read length[a]
pHEA	384 ± 18	430
POP-5™ [b]	< 50	N/A

a) Read lengths given at 98.5% accuracy.
b) No precoating of chip; matrix loaded onto uncoated chip.

3.5 Comparing matrix performance in chips with matrix performance in capillaries

The relative sequencing performances for these matrices differ greatly on the capillary and microchip platforms. In either the ABI 3100 or the MegaBace 1000, the commercial matrices tested here easily deliver 600–700 base reads in an hour or two. Interestingly, the pDMA matrices formulated in our laboratory also deliver long sequencing reads in a little over 20 min. In the chips, however, these pDMA matrices deliver much higher sequencing read lengths than the commercial matrices. These results suggest that, much like surface coatings, the materials developed for sequencing in capillary arrays will not be the optimal materials for DNA sequencing in microchip systems. Thus, new sequencing matrices developed and optimized specifically for microchip-based sequencing systems are needed.

As mentioned above, the main difference between capillary and microchip systems is the sample injection scheme. The injection method determines the initial sample plug width, which contributes to the overall efficiency of the separation [47]. In microchips, the geometry of the cross injector or offset T injector approximately defines the width of the sample plug resulting in plug widths of the order of 100–200 μm. CAE instruments inject the sample DNA *via* direct electrokinetic injection from the sample well into the capillary. The injection must be long enough to reach a concentration in the capillary so that the fluorescence can be detected with high signal to noise, but the sample plug width, w_{inj}, also increases with time, t_{inj}, as

$$W_{inj} = \mu \times E_{inj} \times t_{inj} \quad (1)$$

where E_{inj} is the injection electric field strength and μ is the DNA mobility at E_{inj}. Equation (1) shows that the sample

plug width, and thus the separation efficiency, depends on the mobility of the DNA during injection, which is directly related to the properties of the polymer matrix. In capillary-based separations for DNA sequencing, the sample plug width can be of the order of 1 mm and is generally the dominant factor contributing to eluted peak widths for the separations. In microchip-based separations, the matrix does not greatly affect the plug width during injection. Since the initial plug widths are smaller, the separation efficiency is dominated by other factors such as field-induced dispersion or analyte–wall interactions. Thus, different parameters related to the separation matrix and wall coating may limit the efficiency of the separations depending on the platform, so that the dependence of resolution on a matrix will not follow the same trends in capillary and microchip separations.

4 Conclusions

The development of high-performance polymer matrices and coatings has been one of the major obstacles in the development of a microchip-based commercial system for Sanger-based DNA sequencing. A long-held assumption was that the polymer materials available commercially for capillary-based systems could be easily transferred to microfabricated systems. However, we have shown here that this assumption is invalid and the development of materials for use exclusively on microchip electrophoresis platforms is necessary to make these instruments a practical reality.

In this paper we also have demonstrated that sequencing matrices and channel wall coatings developed on microchip electrophoresis platforms can deliver longer read lengths in less time than the matrices and coatings that are currently commercially available for CAE instruments. At the same channel length, our matrices deliver over 150 more sequenced bases than the POP™ matrices from ABI and over 250 more sequenced bases than the LongRead™ matrix from Amersham. For development of new polymer matrices for microchannel systems, optimization of the materials must be completed on microfabricated systems, as matrix-sequencing performance is not identical in capillaries and microchips. In addition to matrix performance, the difference in glass chemistries between the fused silica used in capillaries and the borosilicate glass used for chip fabrication can lead to varying wall coating–channel surface compatibility. Thus, coatings must be developed specifically for the chip substrate chemistry as well.

This work was supported by grant # 2 R01 HG001970-07 from the National Human Genome Research Institute of the National Institute of Health. However, the views expressed in this paper are solely the responsibility of the authors and do not necessarily represent the official views of the National Human Genome Research Institute or the National Institute of Health.

The authors have declared no financial or commercial conflict of interest.

5 References

[1] Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C. *et al.*, *Nature* 2001, *409*, 860–921.

[2] Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W. *et al.*, *Science* 2001, *291*, 1304–1354.

[3] Gibbs, R. A., Rogers, J., Katze, M. G., Bumgarner, R. *et al.*, *Science* 2007, *316*, 222–234.

[4] Mikkelsen, T. S., Hillier, L. W., Eichler, E. E., Zody, M. C. *et al.*, *Nature* 2005, *437*, 69–87.

[5] Bussell, J. J., Pearson, N. M., Kanda, R., Filatov, D. A., Lahn, B. T., *Gene* 2006, *368*, 94–100.

[6] Harris, R. A., Rogers, J., Milosavljevic, A., *Science* 2007, *316*, 235–237.

[7] Newman, T. L., Tuzun, E., Morrison, V. A., Hayden, K. E. *et al.*, *Genome Res.* 2005, *15*, 1344–1356.

[8] Perry, G. H., Tchinda, J., McGrath, S. D., Zhang, J. J. *et al.*, *Proc. Natl. Acad. Sci. USA* 2006, *103*, 8006–8011.

[9] Zhao, S. Y., Shetty, J., Hou, L. H., Delcher, A. *et al.*, *Genome Res.* 2004, *14*, 1851–1860.

[10] Check, E., *Nature* 2005, *437*, 1084–1086.

[11] Gonzalez, E., Kulkarni, H., Bolivar, H., Mangano, A. *et al.*, *Science* 2005, *307*, 1434–1440.

[12] Recommendation for a Human Cancer Genome Project. Report of the Working Group on Biomedical Technology. February 2005. http://www.genome.gov/15015123

[13] NIH News Release. NHGRI Seeks Next Generation of Sequencing Technologies. 10 October 2004. http://genome.gov/12513210

[14] Cohen, A. S., Najarian, D. R., Karger, B. L., *J. Chromatogr.* 1990, *516*, 49–60.

[15] Dovichi, N. J., *Electrophoresis* 1997, *18*, 2393–2399.

[16] Hunkapiller, T., Kaiser, R. J., Koop, B. F., Hood, L., *Science* 1991, *254*, 59–67.

[17] Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P. *et al.*, *Nature* 1986, *321*, 674–679.

[18] Mardis, E. R., Personal Communication 2006.

[19] Chan, E. Y., *Mutat. Res.* 2005, *573*, 13–40.

[20] Margulies, M., Egholm, M., Altman, W. E., Attiya, S. *et al.*, *Nature* 2005, *437*, 376–380.

[21] Shendure, J., Porreca, G. J., Reppas, N. B., Lin, X. X. *et al.*, *Science* 2005, *309*, 1728–1732.

[22] Blazej, R. G., Kumaresan, P., Mathies, R. A., *Proc. Natl. Acad. Sci. USA* 2006, *103*, 7240–7245.

[23] Paegel, B. M., Yeung, S. H. I., Mathies, R. A., *Anal. Chem.* 2002, *74*, 5092–5098.

[24] Zhou, H. H., Miller, A. W., Sosic, Z., Buchholz, B. *et al.*, *Anal. Chem.* 2000, *72*, 1045–1052.

[25] Doherty, E. A. S., Kan, C. W., Barron, A. E., *Electrophoresis* 2003, *24*, 4170–4180.

[26] Doherty, E. A. S., Kan, C. W., Paegel, B. M., Yeung, S. H. I. *et al.*, *Anal. Chem.* 2004, *76*, 5249–5256.

[27] Hjerten, S., *J. Chromatogr.* 1985, *347*, 191–198.

[28] Chiari, M., Riva, S., Gelain, A., Vitale, A., Turati, E., *J. Chromatogr. A* 1997, *781*, 347–355.

[29] Heller, C., *Electrophoresis* 1998, *19*, 3114–3127.

[30] Madabhushi, R. S., *Electrophoresis* 1998, *19*, 224–230.

[31] Doherty, E. A. S., Berglund, K. D., Buchholz, B. A., Kourkine, I. V. *et al.*, *Electrophoresis* 2002, *23*, 2766–2776.

[32] Koutny, L., Schmalzing, D., Salas-Solano, O., El-Difrawy, S. *et al.*, *Anal. Chem.* 2000, *72*, 3388–3391.

[33] Liu, S. R., Shi, Y. N., Ja, W. W., Mathies, R. A., *Anal. Chem.* 1999, *71*, 566–573.

[34] Paegel, B. M., Emrich, C. A., Weyemayer, G. J., Scherer, J. R., Mathies, R. A., *Proc. Natl. Acad. Sci. USA* 2002, *99*, 574–579.

[35] Salas-Solano, O., Schmalzing, D., Koutny, L., Buonocore, S. *et al.*, *Anal. Chem.* 2000, *72*, 3129–3137.

[36] Schmalzing, D., Adourian, A., Koutny, L., Ziaugra, L. *et al.*, *Anal. Chem.* 1998, *70*, 2303–2310.

[37] Shi, Y. N., *Electrophoresis* 2006, *27*, 3703–3711.

[38] Shi, Y. N., Anderson, R. C., *Electrophoresis* 2003, *24*, 3371–3377.

[39] Fredlake, C. P., Hert, D. G., Kan, C. W., Cheisl, T. N. *et al.*, *Proc. Natl. Acad. Sci. USA* 2008, *105*, 476–481.

[40] Chiesl, T. N., Shi, W., Barron, A. E., *Anal. Chem.* 2005, *77*, 772–779.

[41] Albarghouthi, M., Buchholz, B. A., Doherty, E. A. S., Bogdan, F. M. *et al.*, *Electrophoresis* 2001, *22*, 737–747.

[42] Goetzinger, W., Kotler, L., Carrilho, E., Ruiz-Martinez, M. C. *et al.*, *Electrophoresis* 1998, *19*, 242–248.

[43] Salas-Solano, O., Carrilho, E., Kotler, L., Miller, A. W. *et al.*, *Anal. Chem.* 1998, *70*, 3996–4003.

[44] Jacobson, S. C., Hergenroder, R., Koutny, L. B., Ramsey, J. M., *Anal. Chem.* 1994, *66*, 1114–1118.

[45] Heller, C., *Electrophoresis* 2000, *21*, 593–602.

[46] Llopis, S. D., Stryjewski, W., Soper, S. A., *Electrophoresis* 2004, *25*, 3810–3819.

[47] Luckey, J. A., Norris, T. B., Smith, L. M., *J. Phys. Chem.* 1993, *97*, 3067–3075.

