

Methal N. Albarghouthi¹
 Brett A. Buchholz¹
 Erin A. S. Doherty¹
 Felicia M. Bogdan¹
 Haihong Zhou²
 Annelise E. Barron¹

¹Department of Chemical Engineering,
 Northwestern University,
 Evanston, IL, USA

²Barnett Institute and
 Department of Chemistry,
 Northeastern University,
 Boston, MA, USA

Impact of polymer hydrophobicity on the properties and performance of DNA sequencing matrices for capillary electrophoresis

To elucidate the impact of matrix chemical and physical properties on DNA sequencing separations by capillary electrophoresis (CE), we have synthesized, characterized and tested a controlled set of different polymer formulations for this application. Homopolymers of acrylamide and *N,N*-dimethylacrylamide (DMA) and copolymers of DMA and *N,N*-diethylacrylamide (DEA) were synthesized by free radical polymerization and purified. Polymer molar mass distributions were characterized by tandem gel permeation chromatography - laser light scattering. Polymers with different chemical compositions and similar molar mass distributions were selected and employed at the same concentration so that the variables of comparison between them were hydrophobicity and average coil size in aqueous solution. We find that the low-shear viscosities of 7% w/v polymer solutions decrease by orders of magnitude with increasing polymer hydrophobicity, while hydrophilic polymers exhibit more pronounced reductions in viscosity with increased shear. The performance of the different matrices for DNA sequencing was compared with the same sample under identical CE conditions. The longest read length was produced with linear polyacrylamide (LPA) while linear poly-*N,N*-dimethylacrylamide (PDMA) gave ~ 100 fewer readable bases. Read lengths with DMA/DEA copolymers were lower, and decreased with increasing DEA content. This study highlights the importance of polymer hydrophilicity for high-performance DNA sequencing matrices, through the formation of robust, highly-entangled polymer networks and the minimization of hydrophobic interactions between polymers and fluorescently-labeled DNA molecules. However, the results also show that more hydrophobic matrices offer much lower viscosities, enabling easier microchannel loading at low applied pressures.

Keywords: DNA sequencing / Capillary electrophoresis / Polymer solutions / Matrices / Polymer hydrophobicity
 EL 4253

1 Introduction

The Human Genome Project (HGP) has successfully accomplished one of its primary goals through the completion of a “working draft” of the human genome, covering 85% of the clonable sequence with a 4 × depth of coverage [1]. This milestone was reached on an accelerated time schedule through the introduction of automated,

high-throughput capillary array electrophoresis (CAE) instruments, which are rapidly becoming the dominant tool in DNA sequencing centers [2, 3]. In addition to the HGP, a constantly increasing number of animal, plant and microbial genome projects [4] will continue to seek new advances in high-throughput and cost-effective sequencing technologies.

DNA sequencing is accomplished through the size-based separation of fluorescently-labeled, single-stranded DNA fragments that range in size from just a few to more than 1000 bases. In capillary electrophoresis (CE), DNA separation is typically performed in an entangled, replaceable, uncross-linked polymer solution. Compared to slab-gel electrophoresis, CE in polymer solutions eliminates gel preparation and pouring steps, and allows automated filling and replacement of sequencing matrices in capillary arrays. The necessity for a polymeric separation medium for DNA sequencing arises from the fact that electrophoretic size-based separation of DNA fragments cannot be achieved in free solution, as DNA molecules exhibit an almost constant charge-to-frictional coefficient ratio re-

Correspondence: Dr. Annelise E. Barron, Department of Chemical Engineering, Northwestern University, 2145 Sheridan Road, Room E136, Evanston, IL 60208, USA

E-mail: a-barron@northwestern.edu

Fax: +847-491-3728

Abbreviations: DEA, *N,N*-diethylacrylamide; DMA, *N,N*-dimethylacrylamide; FWHM, full width at half maximum; GPC, gel permeation chromatography; LPA, linear polyacrylamide; MALLS, multi-angle laser light scattering; PDEA30, PDEA50, PDEA70, linear poly(co-diethyl-co-dimethyl)acrylamide polymers with nominally 30%, 50%, and 70% content of the DEA monomer and the balance DMA, respectively; PDMA, linear poly-*N,N*-dimethylacrylamide

ardless of their size. Hence, the electrophoretic mobility of DNA in free solution is virtually independent of fragment size [5, 6]. However, in a polymeric separation medium, matrix chains serve as interactive, entangling “obstacles” to the electrophoretic migration of DNA molecules, breaking the symmetry of chain-length dependence between the net charge and the effective friction coefficient of DNA fragments. These physical DNA-polymer interactions endow DNA fragments with size-dependent electrophoretic mobilities, allowing the sequence of bases to be read.

According to both theory and observation, the performance of a polymer matrix in DNA sequencing is governed primarily by the ability of polymer chains to form a robust, entangled polymer network. It is useful to view a given polymer chain that is part of an entangled network as a succession of independent chain subunits, each defined by the collection of randomly-configured polymer segments that lie between two points of polymer-polymer entanglement. In the language of polymer physics, the average coil volume adopted by one of these polymer subunits is called a blob, and has a given average blob size, ξ_b [7]. As DNA molecules migrate through the network, they interact and entangle with these polymer ‘blobs’, and it is the time scale of these DNA-polymer interactions relative to those of polymer interactions that controls the ability of a network to separate DNA. Cottet *et al.* [8] describe the dynamics of an entangled polymer network by the ‘reptation time’, which is a measure of the lifetime of polymer chain entanglements as controlled by chain movements occurring by random, curvilinear diffusion. For a given polymer matrix, the average reptation time increases with increasing polymer concentration and molar mass, but decreases with increasing temperature. Size-dependent mobility and good resolution of electrophoresing DNA can be achieved when the reptation time of the polymer chains is greater than the residence time of a given DNA molecule in the given polymer blob [8, 9]. Moreover, polymer chain entanglements must be sufficiently strong and the entanglement density (number of entanglements per chain) sufficiently high to prevent extensive network rupture by migrating DNA molecules [10].

A range of water-soluble polymers has been studied for use in DNA sequencing including linear polyacrylamide (LPA) [11], linear poly-*N,N*-dimethylacrylamide (PDMA) [12], polyethylene oxide [13], hydroxyethylcellulose [14], polyvinyl pyrrolidone [15], polyethylene glycol with fluorocarbon tails [16], poly-*N*-acryloylaminopropanol [17] and copolymers of *N,N*-dimethylacrylamide (DMA) and *N,N*-diethylacrylamide (DEA) [18]. The best sequencing performance to date is that of LPA, producing 1000 bases in

about 1 h [19] and 1300 bases in 2 h [11] with highly optimized polymer molar mass distribution, matrix formulation, sample preparation and clean-up, and base-calling algorithms. Lower sequencing rates are more common in commercial CAE instruments such as the MegaBACE 1000™ (600 bases in 2 h) [20] and the ABI PRISM 3700™ (550 bases in 3–4 h) [21]. The shorter read lengths achieved by commercial instruments can be attributed to the use of less-optimized polymer matrices as well as the quality of real genomic DNA samples, in particular, a lower purity of sequencing reaction products and the tendency for anomalous migration of some fragments due to formation of DNA secondary structure. Moreover, high-molar-mass polymer solutions that provide very long DNA sequencing read lengths generally have high zero-shear viscosities, requiring the application of high pressures to initiate matrix flow into a capillary array or into chip microchannels. In automated CAE instruments and soon-to-appear microchip-based DNA sequencing devices, application of high pressure may or may not be practical for a given instrument design.

Performance optimization studies of DNA sequencing matrices have investigated variables such as polymer molar mass, matrix composition, electric field strength, and electrophoresis temperature [11, 13, 16, 17, 19, 22, 23]. These studies have converged in showing that a relatively low concentration of high-molar-mass polymer chains is needed to separate large DNA fragments, whilst a higher overall concentration of polymer chains helps to improve the resolution of small DNA fragments. The use of low electric field strength generally improves the resolution of DNA fragments by delaying the onset of biased reptation, thus extending the read length at the expense of a longer run time. High temperatures can reduce separation time and also help to minimize the occurrence of band compressions that result from DNA secondary structure formation, and can improve resolution provided that the entangled polymer network retains its robustness at elevated temperatures.

Most polymer matrices exhibit a similar dependence of DNA sequencing performance upon the variables discussed above. However, it is well known that different polymers exhibit different intrinsic capabilities to resolve DNA sequencing fragments, and that the read lengths and required run times vary widely from matrix to matrix. Since a given research laboratory typically focuses on only one class of polymer matrix in any given study, it remains unclear why different polymers exhibit such widely varying performances as DNA sequencing matrices. The lack of comparative data makes it difficult to predict the performance of untested polymers or to design improved polymer matrices. To our knowledge, there has

been no comprehensive report in the literature that compares the performance of different DNA sequencing matrices under similar conditions in the same instrument. Such a comparison will allow better interpretation of sequencing performance in relation to the physicochemical nature of the polymers, and will deepen the understanding of how DNA sequencing separation is impacted by polymer chemical and physical properties.

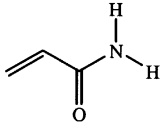
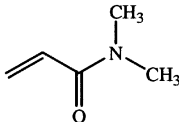
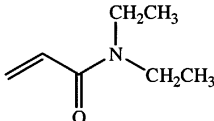
In this study, we compare the sequencing performance and rheological properties of five different polymer matrices of different chemical structures, yet similar average molar mass, including LPA, PDMA and different copolymer formulations of DMA and DEA. The structures of the monomers, shown in Table 1, indicate that acrylamide is more hydrophilic than DMA [24], which in turn is more hydrophilic than DEA. Polymer molar mass distributions were characterized, and the polymer mass and concentration were held constant for the comparison. The results reveal the dramatic impact of polymer hydrophobicity on DNA sequencing separations, and furthermore highlight the importance of the robustness of the entangled polymer network in providing high-resolution separation of DNA fragments and long read lengths. Additionally, the rheological properties of the polymer matrices are compared (*i.e.*, the dependence of matrix viscosity on shear rate), because the pressure-induced flow behavior of polymer solutions has an important impact on microchannel loading behavior, a matrix attribute that is in turn important for automated CAE and microchip sequencing instruments.

2 Materials and methods

2.1 Polymer synthesis

Ultrahigh purity (> 99.5%) DEA and DMA (Monomer-Polymer and Dajac Labs, Feasterville, PA, USA) were copolymerized by free radical polymerization in aqueous solution at 47°C for approximately 16 h. The polymerization was initiated by 0.03% w/v V-50 (2,2'-Azobis (2-amidinopropane) dihydrochloride; Wako Chemical USA, Richmond, CA, USA). Nitrogen gas was continuously bubbled through the solution for 3 h prior to initiation. Maintaining the total monomer concentration at 7.0% w/w, different formulations of the copolymers were obtained by varying the proportions of each monomer (0, 30, 50, and 70% w/v DEA) in the solution. The polymer was purified by dialysis against triply-distilled water using Spectra/Por cellulose ester dialysis membranes (Spectrum, Gardena, CA, USA), having a molecular mass cutoff of 1000 Da. The polymer was recovered from the dialyzed solution by lyophilization (Labconco, Kansas City, KS, USA). In this work, we will use the names of PDEA30, PDEA50 and

Table 1. Chemical structure of acrylamide, DMA and DEA

Monomer	Chemical structure
Acrylamide	
DMA	
DEA	

PDEA70 for the DMA/DEA copolymers prepared with 30, 50 and 70% w/v DEA, respectively. LPA was prepared using 7.0% w/v acrylamide (Amresco, Solon, OH, USA) and 0.03% w/v V-50 as the initiator at 47°C for about 16 h. LPA was purified by dialysis and recovered by lyophilization.

2.2 Copolymer composition

¹H-NMR was performed with a Varian INOVA 500 (Walnut Creek, CA, USA) to determine the composition of the DMA/DEA copolymers and the actual proportions of each monomer incorporated in each of the various copolymer formulations. ¹H-NMR (500 MHz, D₂O) δ (ppm): 1.12 (2 × CH₃ of DEA, br s), 1.43–1.73 (CH₃-CH₂, m), 2.66 (CH-CH₂, m), 2.95 (2 × CH₃ of DMA, br s), 3.32 (2 × CH₂ of DEA, m). Integration of NMR signals at 1.12 and 2.95 ppm was used to calculate the relative molar proportions of DEA and DMA, respectively, incorporated into the copolymer.

2.3 Polymer molar mass distribution

To determine the weight-average molar mass of the synthesized polymers, the samples were fractionated by gel permeation chromatography (GPC) prior to on-line multi-angle laser light scattering (MALLS) detection, using a Waters 2690 Alliance Separations Module (Milford, MA, USA) with Shodex (New York, NY, USA) OHPak columns SB-806 HQ, SB-804 HQ, and SB-802.5 HQ connected in series. In this tandem GPC-MALLS mode, the effluent from the GPC system flows directly into a DAWN DSP Laser Photometer and Optilab DSP Interferometric Refractometer connected in series (both from Wyatt Technology, Santa Barbara, CA, USA). One hundred μL of each

sample was injected into the tandem GPC-MALLS system at a concentration of ~ 0.5 mg/mL. For samples LPA, PDMA, PDEA30, and PDEA50 the flow rate was 0.35 mL/min and the mobile phase consisted of 100 mM NaCl, 50 mM NaH₂PO₄, and 200 ppm NaN₃. For sample PDEA70 the flow rate was 0.20 mL/min in a mobile phase consisting of 50% aqueous phase containing 100 mM NaCl, 50 mM NaH₂PO₄, and 200 ppm NaN₃, and 50% methanol. The tandem GPC-MALLS data were processed using ASTRA software from Wyatt Technology. ASTRA was used to calculate the weight-average molar mass, polydispersity index, and weight-average radii of gyration of the analyzed polymers. The error associated with each of the calculated values was 1% at most. All analyses were repeated three times and the standard deviation was less than 1%, indicating the high accuracy and the reproducibility of the estimated values. The measured weight-average molar mass and radius of gyration can then be used to estimate the threshold overlap concentration, c^* :

$$c^* = \frac{M_w}{\frac{4}{3}\pi N_A R_g} \quad (1)$$

where M_w is the polymer weight-average molar mass, N_A is Avogadro's constant, and R_g is the weight-average radius of gyration of the polymer.

2.4 Rheological characterization

A temperature-controlled rotational Bohlin VOR rheometer (Cranbury, NJ, USA) equipped with a cone-plate geometry (30 mm diameter, 2.5° angle) was used to determine the steady-shear viscosity of the DNA separation matrices at different shear rates in the range of 0.005–1470 s⁻¹. Measurements were taken at 25°C. Polymer solutions (7% w/v) were prepared in sequencing buffer consisting of 50 mM Tris, 2 mM EDTA (Amresco, Solon, OH, USA), 50 mM *N*-tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid (TAPS) (Sigma, St. Louis, MO, USA), and 7 M urea (Amresco).

2.5 DNA sequencing

DNA sequencing was performed on a MegaBACE 1000™ CAE instrument (Molecular Dynamics, Sunnyvale, CA, USA) equipped with 6 × 16 fused-silica capillary arrays (75 μm inner diameter, 64 cm total length, 40 cm effective length) covalently coated with LPA. DNA sequencing buffer was used to prepare a 7.0% w/v solution of each polymer as a DNA sequencing matrix. Mixing of polymer solutions was by slow rotation (Rotatorque, Cole-Palmer, Vernon Hills, IL, USA) for 48 h. MegaBACE Sequencing

Standards (Amersham Pharmacia Biotech, Piscataway, NJ, USA) consisting of M13mp18 sequencing reaction products, labeled with energy transfer dye primers and purified by ethanol precipitation, were used. Sequencing matrices were loaded into the capillaries under a pressure of 1000 psi for 200 s followed by a relaxation time of 20 min and a prerun electrophoresis for 5 min at 140 V/cm. After electrokinetic sample injection at 93.75 V/cm for 40 s, electrophoresis of DNA was performed at 140 V/cm and 44°C for 100 min. Laser-induced fluorescence (LIF) data were collected, analyzed and translated into called DNA sequence using the MegaBACE 1000 DNA Sequencing Software Version 2.0™.

2.6 Data analysis

DNA sequencing data were analyzed using the T-track of the four-color DNA sequencing reaction products. Raw LIF data were extracted from the MegaBACE sequencing software. Single T-peaks were fitted into Gaussian peaks using PeakFit™ 4.06 (SPSS, Chicago, IL, USA) from which the full width at half-maximum (FWHM) was estimated for each peak. Peak spacing was estimated as the average spacing between the centers of a given T-peak and the peaks on both sides of that T-peak. The plot of FWHM vs. DNA fragment size was modeled by an exponential function whilst the peak spacing curve was modeled by a polynomial of order 2. The selected functions best modeled the experimental data and yielded the lowest sum of squares of errors. Similar trends of peak spacing and FWHM have been observed in other sequencing matrices [12, 16]. The fitted functions were used to calculate the resolution, R_s , of the peaks using the following equation [12, 25]:

$$R_s = 0.59 \frac{x_2 - x_1}{FWHM} \quad (2)$$

where x_i is the center of peak i .

3 Results and discussion

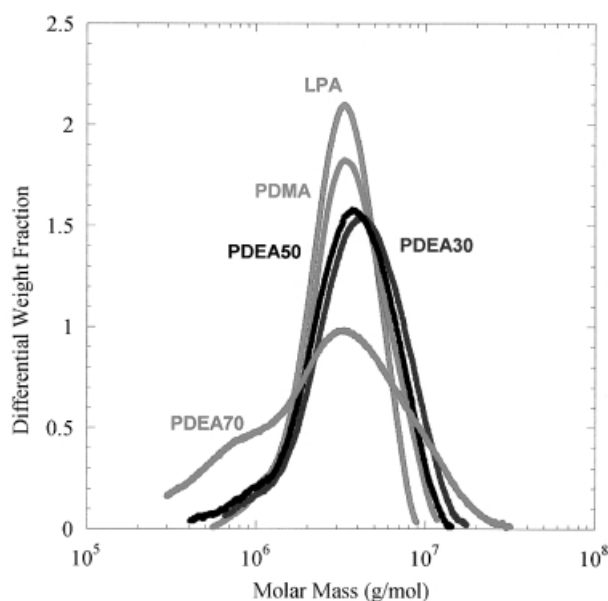
3.1 Polymer synthesis and characterization

Table 2 lists the composition of the DMA/DEA copolymers used in this study. The DEA content in the copolymer increases with increasing DEA composition in the initial monomer feed, and is close to the original monomer composition in the feed.

The molar mass distribution of the synthesized polymers is shown in Fig. 1 as determined by tandem GPC-MALLS. Table 3 lists the weight-average molar mass, polydispersity index and weight-average radius of gyration of each polymer used in this study. The weight-average molar mass of the polymers is in the range of 4.1 ± 0.6 MDa.

Table 2. Comparison of the monomer ratio in the copolymer feed solutions and that in the synthesized copolymers

Polymer	DMA:DEA in feed	DMA:DEA in copolymer
PDEA30	70:30	70:30
PDEA50	50:50	49:51
PDEA70	30:70	28:72

**Figure 1.** Molar mass distributions of LPA, PDMA and copolymers of DMA and DEA.

With the exception of PDEA70, the polymers have similar polydispersity indexes ranging from 1.20 to 1.47. The polydispersity of PDEA70 is much higher (2.42) and this can be seen from the broad molar mass distribution of PDEA70 shown in Fig. 1, particularly in the low molar mass part of the curve.

With the exception of PDEA30, the radius of gyration decreases with increasing polymer hydrophobicity, as a result of interactions between side chain methyl groups in

DMA and ethyl groups in DEA. These hydrophobic moieties tend to collocate within the polymer coil, in order to minimize entropically unfavorable hydrophobic hydration interactions between 'greasy' alkyl groups and the aqueous solvent [26]. As a result, for polymers of similar molar mass, more hydrophobic polymer chains adopt a more compact coil conformation in water [27] leading to a decrease in the radius of gyration of the polymer chain with an increase in polymer hydrophobicity. The PDEA30 sample used in this study does not follow the expected trend, exhibiting a radius of gyration that is somewhat larger than that of the more hydrophilic PDMA sample. This stems from the greater average molar mass of the PDEA30 (4.7 MDa) as compared to that of PDMA (3.9 MDa). Here, the contribution to increased coil radius that is provided by an increase in chain length exceeds the relative reduction in coil size of PDEA30 that results from its greater overall hydrophobicity.

To account for the effect of molar mass variation between the polymers and to better observe the trend in material properties, the radius of gyration of each polymer was estimated for a given molar mass, using physical scaling laws that generally hold for a random-coil polymer in a good solvent [28]:

$$R_g = kM_w^{0.6} \quad (3)$$

where R_g is the radius of gyration, k is a constant of proportionality unique to a given polymer-solvent system, and M_w is the weight-average molar mass of the polymer. We neglect the fact that water is a somewhat 'less good' solvent for the more hydrophobic polymers, as we wish only to make a gross comparison. Using the experimentally determined values of molar mass and radius of gyration for each polymer, the radius of gyration for a polymer molar mass of 4.0 MDa was calculated from Eq. (3) and is given in Table 3. As expected, for a given weight-average molar mass, the radius of gyration of a polymer chain decreases with increasing polymer hydrophobicity.

To account for the difference in molar mass of monomer units incorporated into the copolymers as well as to allow

Table 3. Physical characteristics of LPA, PDMA and copolymers of DMA and DEA

Polymer	Weight-average molar-mass ^{a)} (MDa)	Polydispersity index ^{a)}	Measured R_g ^{a)} (nm)	Calculated c^* (% w/v)	c/c^*	Calculated R_g for 4.0 MDa polymer (nm)
LPA	3.5	1.21	129	0.06	108	140
PDMA	3.9	1.31	122	0.09	82	124
PDEA30	4.7	1.42	1.25	0.10	73	114
PDEA50	4.1	1.47	109	0.13	56	107
PDEA70	4.2	2.42	84	0.28	25	82

a) The standard deviation for three runs is less than 1.0%.

more direct comparison of the variation in coil size, the average radius of gyration was estimated for idealized polymers taken to be composed of 50 000 monomer units. Again, this was done with the use of scaling laws for the different polymers. The average molar mass of the monomer repeat unit was calculated based on the chemical composition of the polymer determined in Table 2. The predicted radii of gyration for the idealized polymers are listed in Table 4, which shows that the radius of gyration of LPA of a given contour length is actually somewhat smaller than that of PDMA. This is due to the smaller molar mass of the LPA composed of 50 000 monomer units (3.5 MDa) compared to that of PDMA (4.9 MDa). For DMA-containing copolymers, the molar mass of the polymer chain increases with increasing DEA content. Although the greater mass of the DEA side chains should increase the coil radius through steric effects, the radius of gyration is still predicted to decrease based on the use of the scaling equation as more DEA is incorporated into the copolymer chain. This confirms that the observed reduction in coil size is due to the hydrophobicity of the polymer chain leading to more compact, less-extended conformation in aqueous solution. Furthermore, Table 4 also shows that the polymer coil density, calculated from the polymer molar mass divided by the volume of the coil, increases with increasing polymer hydrophobicity for the polymers considered. Thus, in agreement with the results in Table 3, for a polymer coil of a given molar mass, the coil volume and radius of gyration decrease with increasing polymer hydrophobicity.

Figure 2 shows the viscosity of polymer matrices as a function of the shear rate. All of the polymer solutions that were studied behave as Newtonian fluids at near-zero shear rates, where the viscosity is almost independent of shear rate. At higher shear rates, the polymer matrices exhibit non-Newtonian, shear-thinning behavior in which the viscosity decreases with increasing shear rate. This behavior has been widely observed in polymer solutions [29, 30]. The results show that for any given shear rate, the LPA solution has the highest viscosity in comparison to the solutions of other, less hydrophilic polymers. The zero-shear viscosity of the DEA-containing polymer matrix decreases markedly with increasing DEA content in the polymer.

For concentrated polymer solutions, the viscometric flow behavior of the solution is governed by the extent, or density, of chain entanglements [30, 31]. Table 3 lists the estimated ratio of d/c^* for each polymer matrix, providing a quantitative measure of the extent of entanglement of polymer chains in each matrix, and showing that the extent of chain entanglement is highest for the LPA matrix and decreases progressively with increasing DEA content in DMA-containing polymer matrices. LPA is a linear,

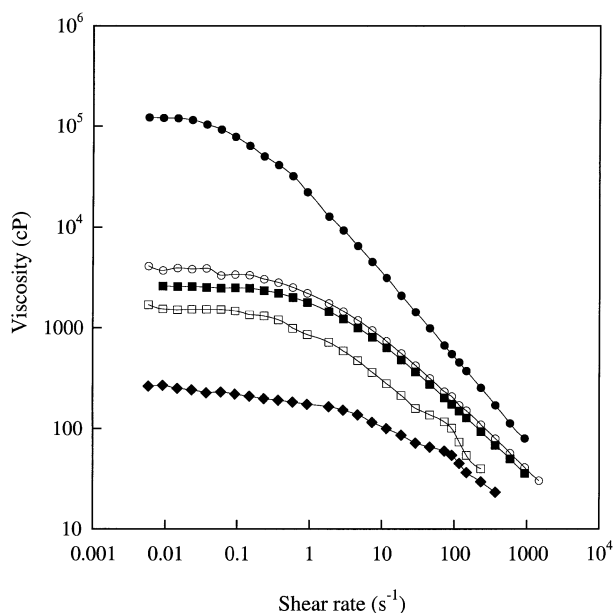


Figure 2. Effect of shear rate on the viscosity of 7% w/v polymer solution in sequencing buffer (50 mM Tris, 50 mM TAPS, 2 mM EDTA and 7 M urea). (●) LPA, (○) PDMA, (■) PDEA30, (□) PDEA50, and (◆) PDEA70.

Table 4. Comparison of estimated radii of gyration of idealized polymer chains composed of 50 000 monomer units

Polymer	Average monomer molar mass (Da)	Molar mass of 50 000 monomers (MDa)	Calculated R_g (nm)	Polymer coil density (Da/nm^3)
LPA	71	3.5	130	0.36
PDMA	99	4.9	141	0.42
PDEA30	109	5.5	137	0.51
PDEA50	115	5.7	133	0.58
PDEA70	120	6.0	104	1.27

hydrophilic polymer that adopts an open and extended coil conformation in aqueous solution, as evidenced by its relatively large radius of gyration (Table 3) and low coil density (Table 4). Entanglements between open and extended polymer chains are generally frequent along the polymer backbone, leading to a high entanglement density and to the high zero-shear viscosity of the LPA solution (*i.e.*, its strong resistance to flow at low shear). For DMA- and DEA-containing polymers, hydrophobic interactions between alkyl-modified side chains produce a decrease in radius of gyration of the polymer coil, which decreases the frequency of entanglements along the polymer chain in a semi-dilute polymer solution. Moreover, the increase in polymer coil density with increasing poly-

mer hydrophobicity may increase the steric hindrance or excluded volume effect between polymer chains, decreasing the overall extent of interchain penetration and entanglement, thus producing a lower solution viscosity at zero shear. The large difference in viscosity between PDEA50 and PDEA70 can be explained by a consideration not only of their difference in hydrophobicity, but also of the greater polydispersity of the PDEA70, which has a larger fraction of low-molar mass chains.

At high shear rates, LPA exhibits shear-thinning behavior that is significantly more pronounced than that of the DMA and DEA-containing polymers. Furthermore, the viscosity-reducing effect of the increased shear rate on the viscosity of DMA- and DEA-containing polymer solutions generally decreases with increasing DEA content in the polymer. This can be seen in a comparison of the negative slopes of the non-Newtonian viscosity vs. shear rate plots in Fig. 2. Generally, increasing the shear rate on an entangled polymer solution increases the shear force that in turn deforms and extends the polymer coils and disrupts entanglements between polymer chains, thus reducing entanglement density and lowering the resistance to flow [30]. The extent of shear-thinning is more significant in LPA than in the DMA or DMA/DEA matrices because in LPA, there are initially more polymer chain entanglement points to be disrupted by the application of shear. Furthermore, the side chains on DMA and DEA monomers are more bulky than those of LPA. Bulky side chains make the movement of polymer chains under the influence of the applied shear more difficult, reducing the effect of increasing shear rate on solution viscosity.

3.2 DNA sequencing

The DNA sequencing performances of these different matrices was compared for polymer and copolymer samples of comparable weight-average molar mass distributions and at the same concentration (7% w/v). Although it is known that high molar mass polymer chains (> 10 MDa) generally are needed to achieve long read lengths [11, 19], our polymers have smaller molar masses (4.1 ± 0.6 MDa) which were greatly governed by the method of polymerization adopted in this work (solution polymerization as opposed to inverse emulsion polymerization [32]). Thus, the matrices that were studied were not expected to perform as well as fully optimized polymer matrices [11]. The concentration of the polymers was selected to be 7% w/v based on our observation that these five polymer matrices produced the longest read lengths at concentrations in the range of 6.0–7.5% w/v. A 7% w/v concentration was then selected to eliminate the effect of variation of polymer concentration while still maintaining reasonably good performances for all the polymer matrices,

Table 5. Effect of polymer chemistry on DNA sequencing read length

Polymer	Read length at 98.5% accuracy	Base number at resolution = 0.59
LPA	620	750
PDMA	512	600
PDEA30	460	530
PDEA50	378	480
PDEA70	122	140

and to retain polymer hydrophobicity as the only major variable in the experiments.

Table 5 summarizes the read length, at 98.5% base-calling accuracy, produced in the various DNA sequencing matrices. Not surprisingly, LPA produced the longest read length (620 bases), followed by PDMA (512 bases). This was the case even though the PDMA sample had a higher average molar mass than the LPA. For DMA/DEA copolymers, the read length decreased monotonically with increasing DEA content in the copolymer. Table 5 shows a similar trend for the read length estimated from analysis of “crossover” plots of peak spacing and FWHM as a function of DNA fragment size (Fig. 3). The point of intersection of peak spacing and peak FWHM curves corresponds to the read length at which resolution of the peaks is approximately 0.59 [12, 25, 33], the point to which a sequence can generally be called with high confidence by an average, well-trained base-caller. We note that some advanced base-calling programs [11] can accurately call bases at resolution as low as 0.25–0.30, and that generally, for best results a base-calling program should be optimized for a given matrix. The base-calling program used for all of the matrices tested for this study was optimized for LPA matrices, so it is possible that the read lengths are somewhat underestimated for the non-LPA matrices. However, we do not expect this effect to be dramatic.

Comparing Figs. 3a–d, we observe that the point of crossover moves to shorter base numbers with increasing matrix hydrophobicity. The plots of peak spacing and FWHM vs. DNA fragment size appear to be fundamentally different in shape for the PDEA70 polymer matrix (Fig. 3e), suggesting a significant difference in the properties of the entangled network formed by these relatively hydrophobic polymers. In all polymer matrices studied, peak spacing decreases exponentially with increasing DNA base number (Fig. 3). The decrease in peak spacing is primarily due to the exponential decrease in electrophoretic mobility of DNA fragments with increasing DNA size, as shown in Fig. 4. The curves indicate that the difference in migration velocity of successive peaks decreases with

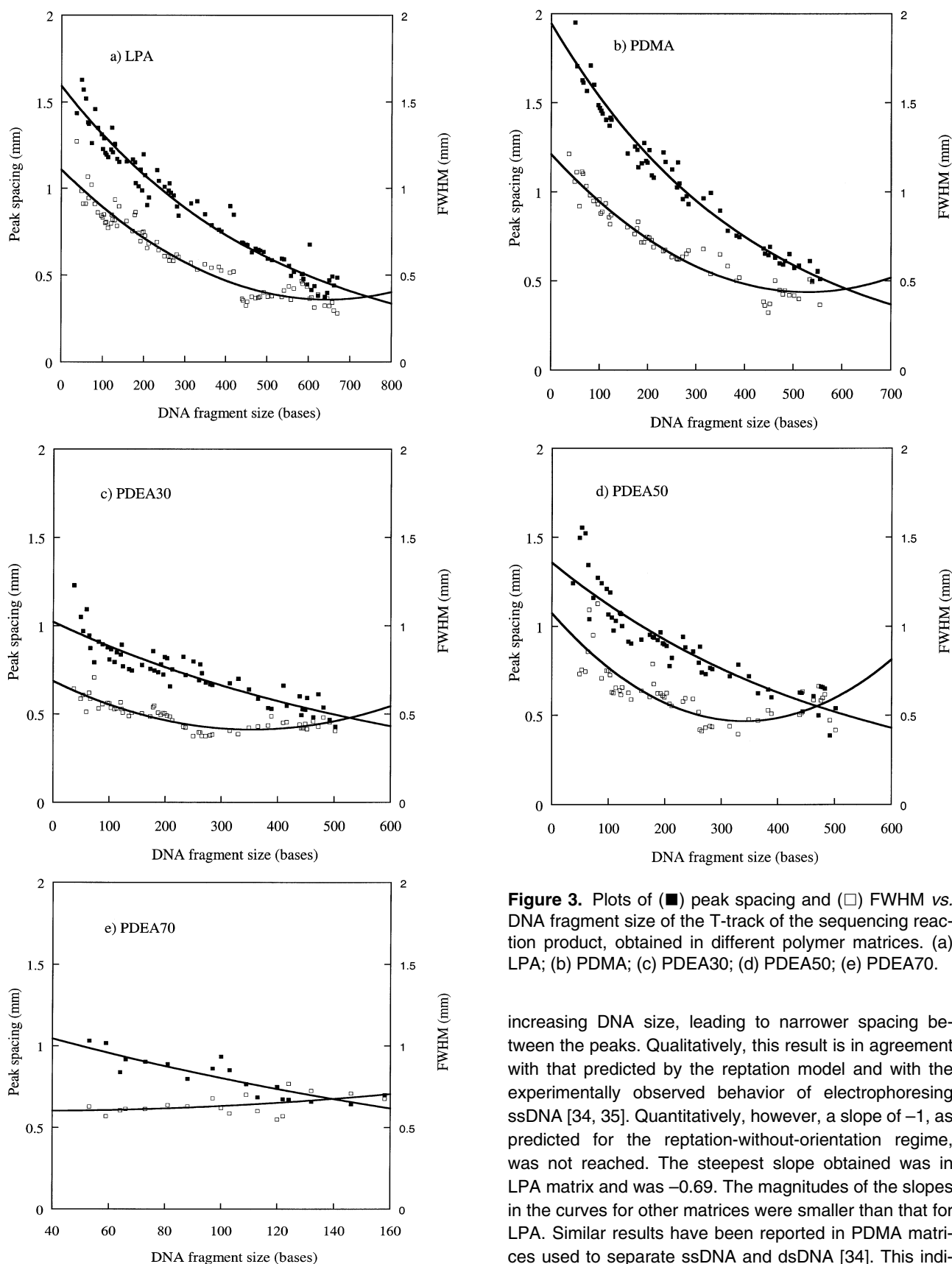


Figure 3. Plots of (■) peak spacing and (□) FWHM vs. DNA fragment size of the T-track of the sequencing reaction product, obtained in different polymer matrices. (a) LPA; (b) PDMA; (c) PDEA30; (d) PDEA50; (e) PDEA70.

increasing DNA size, leading to narrower spacing between the peaks. Qualitatively, this result is in agreement with that predicted by the reptation model and with the experimentally observed behavior of electrophoresing ssDNA [34, 35]. Quantitatively, however, a slope of -1 , as predicted for the reptation-without-orientation regime, was not reached. The steepest slope obtained was in LPA matrix and was -0.69 . The magnitudes of the slopes in the curves for other matrices were smaller than that for LPA. Similar results have been reported in PDMA matrices used to separate ssDNA and dsDNA [34]. This indi-

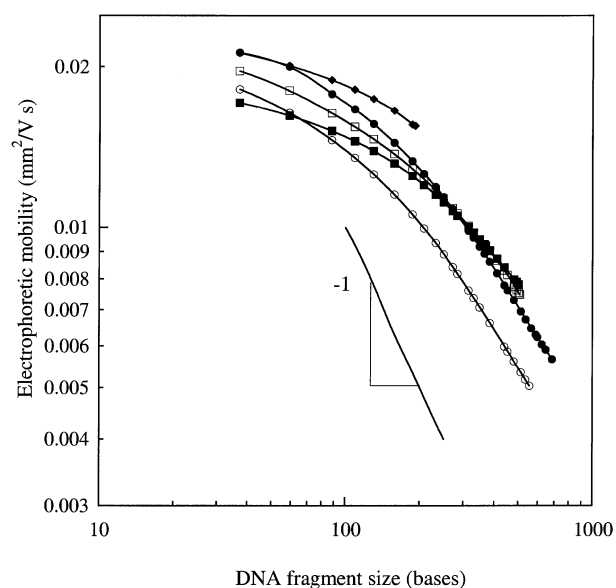


Figure 4. Electrophoretic mobility of DNA fragments as a function of DNA fragment size, obtained in different polymer matrices and presented as a log-log plot. (●) LPA, (○) PDMA, (■) PDEA30, (□) PDEA50, and (◆) PDEA70. The straight line illustrates a slope of 1.

cates that the use of high electric field strengths results in some degree of biased reptation of DNA fragments, rather than pure unbiased reptation. It is possible that with optimization of the applied field strength for each individual matrix, results could be somewhat improved for PDMA and the more hydrophobic matrices at the expense of longer run times.

Crossover plots in Fig. 3 also show that the FWHM curve exhibits an apparent minimum value, and that this apparent minimum shifts to smaller DNA fragment size with increasing polymer hydrophobicity. The minimum FWHM corresponds to a maximum efficiency of separation, which has also been observed by other groups [11, 13, 16, 19, 36, 37]. Beyond the apparent minimum point, FWHM increases slightly with increasing DNA fragment size. This can be explained by considering the fact that the hydrophobicity of DNA increases with DNA size [11]. Hydrophobic DNA molecules interact with the hydrophobic moieties on the polymer chain resulting in increased band dispersion and band-broadening. The length of the DNA at which hydrophobic interactions with the polymer matrix start to effect FWHM appears to decrease with increasing polymer hydrophobicity.

Peak resolution as a function of DNA fragment size for each sequencing matrix is given in Fig. 5. For all matrices, high resolution values (> 0.8) are maintained at low DNA base numbers, and then the resolution drops drastically

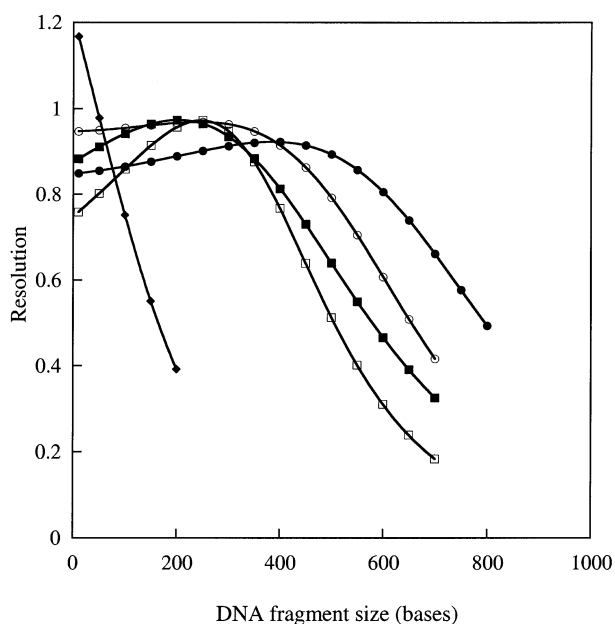


Figure 5. Resolution of DNA sequencing fragments as a function of DNA fragment size, obtained in different polymer matrices. (●) LPA, (○) PDMA, (■) PDEA30, (□) PDEA50, and (◆) PDEA70.

with increasing DNA size. The downturn in resolution is shifted to smaller DNA fragment size with increasing polymer hydrophobicity. Here again, the shape of the curve for PDEA70 differs significantly from those of the other, less hydrophobic matrices.

The mechanism of separation of DNA in uncross-linked polymer matrices has been described by a number of models, reviewed by Heller [34], such as the Ogston model and the various permutations of the reptation model. At present, theoretical work on the mechanism of DNA separation is most useful for providing a physical framework with which to qualitatively interpret experimental results. Typically, each of these models can be applied accurately only within a certain range of DNA fragment sizes and electrophoretic separation conditions, generally for low electric fields. Interpreting our results more phenomenologically, we hypothesize that good resolution of electrophoresing DNA molecules is obtained as long as the DNA-polymer interactions responsible for the size-based separation of DNA molecules do not substantially disrupt the polymer-polymer entanglements and hence locally destroy the polymer network [9].

The extent and strength of entanglements between polymer chains in solution dictate the mechanical robustness of the network. Hydrophilic, high molar mass polymer chains such as LPA adopt an open conformation in aqueous solution, allowing more frequent entanglements along

the chains, which is reflected in the high zero-shear viscosity of the LPA solution. For polymers of comparable molar mass, an increase in polymer hydrophobicity, such as in PDMA and DMA/DEA copolymers, results in the adoption of more dense and compact coil sizes in aqueous solution and the formation of less robust, less entangled polymer networks that do less well in resolving large DNA fragments in comparison to LPA. Thus, as implied by the c/c^* ratio in Table 3, the mechanical strength of the entangled polymer matrix and hence its effective resistance to DNA migration decreases with increasing DEA content in the copolymer.

Microscopic observations have shown that individual DNA molecules can hook around one or more polymer chain(s) during electrophoresis and drag the polymer(s) along before disengaging by a pulley-like action [38]. In our view, short DNA fragments do not generate sufficient electrophoretic momentum to succeed in 'ripping' chains out of the entangled polymer network, and hence are well-separated. However, the polymer network is easily disrupted by facile dragging of poorly-entangled polymer chains by large DNA chains migrating with more momentum. The less robust the entangled polymer network, the more easily it can be disrupted mechanically by DNA migration, and hence the smaller the minimum DNA chain length required to adversely affect the mechanical integrity of the polymer network and the shorter the attained read length.

Figure 4 shows that there is no major, systematic difference in the electrophoretic mobilities of DNA molecules in the different polymer matrices, except for the PDEA70 matrix in which the electrophoretic mobility of DNA molecules was the highest in comparison with that in other matrices. The high electrophoretic mobility of DNA fragments in the PDEA70 matrix is attributable to the fact the PDEA70 polymer chains have the smallest coil size among the studied polymers, and thus the PDEA70 network provides the weakest resistance to DNA migration.

Finally, according to our observations of these and related polymer matrices, some loss of resolution with increasing polymer hydrophobicity may also be attributable to additional band-broadening effects stemming from hydrophobic interaction between alkyl groups on the polymer chains and the fluorescently-labeled DNA molecules. Hence, as polymer hydrophobicity is increased, combined effects of the easier disruption of hydrophobic polymer networks by electrophoresing DNA molecules and hydrophobic polymer-DNA interactions together result in the observed reduction in resolution and read length. However, increases in matrix hydrophobicity also result in dramatic

reductions in solution viscosity at low shear rates, making these matrices potentially easier to load into microchannels under low applied pressure.

4 Concluding remarks

This study demonstrates the dramatic impact of polymer hydrophobicity/hydrophilicity on matrix performance for DNA sequencing by capillary electrophoresis. According to our interpretation of the data, hydrophilic polymers form more robust, entangled polymer networks and have minimal DNA-polymer hydrophobic interactions and hence minimal band broadening. The extent of entanglement in polymer networks also impacts the viscometric flow behavior of a separation matrix, an important attribute for pressurized microchannel loading. Although more hydrophilic polymers improve the overall resolution of electrophoresing DNA fragments and deliver longer read lengths, the high zero-shear viscosities of these matrices require the initial application of high pressures and potentially longer times to replace matrix from capillary arrays and/or from the channels of microfluidic devices. Thus, with the information provided by this study, one may select a polymer chemistry that delivers a suitable compromise between these contrary requirements, depending on the desired read length and turnaround time of the electrophoresis run, as well as on the pressure limitations of the electrophoresis instrument or device. The criteria demonstrated here should be equally applicable to the selection of optimized polymeric separation matrices for electrophoretic separation of double-stranded DNA fragments.

The authors gratefully acknowledge Professor Barry L. Karger at Northeastern University for technical support in the preliminary phases of this work. We thank Dr. Alexander P. Sassi for useful discussions. Financial support and equipment for this work were provided by the National Institutes of Health (AEB, grant # R01HG01970-01), by the Department of Energy (BLK, grant # DE-FG02-90ER60985), and by Molecular Dynamics (Sunnyvale, CA). Some experimental work for this study was performed in the Keck Biophysics Facility at Northwestern University (www.biochem.northwestern.edu/keck/keckmain.html).

Received August 3, 2000

5 References

- [1] Marshall, E., *Science* 2000, 288, 2294–2295.
- [2] Marshall, E., Pennisi, E., *Science* 1998, 280, 994–995.
- [3] Brennan, M., Zurer, P., *Chem. Eng. News* 2000, 78, 11.
- [4] <http://geta.life.uiuc.edu/~nikos/genomes.html>.

- [5] Olivera, B. M., Baine, P., Davidson, N., *Biopolymers* 1964, 2, 245–257.
- [6] Stellwagen, N. C., Gelfi, C., Righetti, P. G., *Biopolymers* 1997, 42, 687–703.
- [7] Broseta, D., Leibler, L., Lapp, A., Strazielle, C., *Europhys. Lett.* 1986, 2, 733–737.
- [8] Cottet, H., Gareil, P., Viovy, J. L., *Electrophoresis* 1998, 19, 215–226.
- [9] Bae, Y. C., Soane, D., *J. Chromatogr.* 1993, 652, 17–22.
- [10] Carlsson, C., Larsson, A., Jonsson, M., Norden, B., *J. Am. Chem. Soc.* 1995, 117, 3871–3872.
- [11] Zhou, H., Miller, A. W., Sosic, Z., Buchholz, B., Barron, A. E., Kotler, L., Karger, B. L., *Anal. Chem.* 2000, 72, 1045–1052.
- [12] Madabhushi, R. S., *Electrophoresis* 1998, 19, 224–230.
- [13] Kim, Y., Yeung, E. S., *J. Chromatogr. A* 1997, 781, 315–325.
- [14] Bashkin, J., Marsh, M., Barker, D., Johnston, R., *Appl. Theor. Electrophor.* 1996, 6, 23–28.
- [15] Gao, Q., Yeung, E. S., *Anal. Chem.* 1998, 70, 1382–1388.
- [16] Menchen, S., Johnson, B., Winnik, M. A., Xu, B., *Electrophoresis* 1996, 17, 1451–1459.
- [17] Lindberg, P., Righetti, P. G., Gelfi, C., Roeraade, J., *Electrophoresis* 1997, 18, 2909–2914.
- [18] Sassi, A. P., Barron, A., Alonso-Amigo, M. G., Hion, D. Y., Yu, J. S., Soane, D. S., Hooper, H. H., *Electrophoresis* 1996, 17, 1460–1469.
- [19] Salas-Solano, O., Carrilho, E., Kotler, L., Miller, A. W., Goetzinger, W., Sosic, Z., Karger, B. L., *Anal. Chem.* 1998, 70, 3996–4003.
- [20] <http://www.mdyn.com/products/MegaBACE/default.html>.
- [21] <http://www.pebio.com/ab/about/dna/377/377a1a.html>.
- [22] Kleparnik, K., Foret, F., Berka, J., Goetzinger, W., Miller, A. W., Karger, B. L., *Electrophoresis* 1996, 17, 1860–1866.
- [23] Zhang, J. Z., Fang, Y., Hou, J. Y., Jiang, R., Roos, P., Dovichi, N. J., *Anal. Chem.* 1995, 67, 4589–4593.
- [24] Gelfi, C., Debesi, P., Alloni, A., Righetti, P. G., *J. Chromatogr.* 1992, 608, 333–341.
- [25] Manabe, T., Chen, N., Terabe, S., Yohda, M., Endo, I., *Anal. Chem.* 1994, 66, 4243–4252.
- [26] Inomata, H., Saito, S., *Fluid Phase Equilibria* 1993, 82, 291–302.
- [27] Barron, A. E., Sunada, W. M., Blanch, H. W., *Electrophoresis* 1996, 17, 744–757.
- [28] Sun, S. F., *Physical Chemistry of Macromolecules*, John Wiley and Sons, New York 1994.
- [29] Bird, R. B., Armstrong, R. C., Hassager, O., *Dynamics of Polymeric Liquids*, John Wiley & Sons, New York 1987.
- [30] Bohdanecky, M., Kovar, J., *Viscosity of Polymer Solutions*, Elsevier Scientific Publishing Company, Amsterdam 1982.
- [31] Volpert, E., Selb, J., Candau, F., *Macromolecules* 1996, 29, 1452–1463.
- [32] Goetzinger, W., Kotler, L., Carrilho, E., Ruiz-Martinez, M. C., Salas-Solano, O., Karger, B. L., *Electrophoresis* 1998, 19, 242–248.
- [33] Grossman, P. D., *J. Chromatogr. A* 1994, 663, 119–227.
- [34] Heller, C., *Electrophoresis* 1999, 20, 1962–1977.
- [35] Heller, C., *Electrophoresis* 1999, 20, 1978–1986.
- [36] Carrilho, E., Ruiz-Martinez, M. C., Berka, J., Smirnov, I., Goetzinger, W., Miller, A. W., Brady, D., Karger, B. L., *Anal. Chem.* 1996, 68, 3305–3313.
- [37] Wu, C. H., Quesada, M. A., Schneider, D. K., Farinato, R., Studier, F. W., Chu, B., *Electrophoresis* 1996, 17, 1103–1109.
- [38] Mitnik, L., Salome, L., Viovy, J. L., Heller, C., *J. Chromatogr. A* 1995, 710, 309–321.