**Daniel G. Hert**[1]
**Christopher P. Fredlake**[1]
**Annelise E. Barron**[1,2]*

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

Received June 20, 2008
Revised August 29, 2008
Accepted August 29, 2008

Research Article

# DNA sequencing by microchip electrophoresis using mixtures of high- and low-molar mass poly(*N,N*-dimethylacrylamide) matrices

Previous studies have reported that mixed molar mass polymer matrices show enhanced DNA sequencing fragment separation compared with matrices formulated from a single average molar mass. Here, we describe a systematic study to investigate the effects of varying the amounts of two different average molar mass polymers on the DNA sequencing ability of poly(*N,N*-dimethylacrylamide) (pDMA) sequencing matrices in microfluidic chips. Two polydisperse samples of pDMA, with weight-average molar masses of 3.5 MDa and 770 kDa, were mixed at various fractional concentrations while maintaining the overall polymer concentration at 5% w/v. We show that although the separation of short DNA fragments depends strongly on the overall solution concentration of the polymer, inclusion of the high-molar mass polymer is essential to achieve read lengths of interest (>400 bases) for many sequencing applications. Our results also show that one of the blended matrices, comprised of 3% 3.5 MDa pDMA and 2% 770 kDa pDMA, yields similar sequencing read lengths (>520 bases on average) to the high-molar mass matrix alone, while also providing a fivefold reduction in zero-shear viscosity. These results indicate that the long read lengths achieved in a viscous, high-molar mass polymer matrix are also possible to achieve in a tuned, blended matrix of high- and low-molar mass polymers with a much lower overall solution viscosity.

**Keywords:**
DNA sequencing / Microchip electrophoresis / Mixed molar mass / Poly(*N,N*-dimethylacrylamide)
DOI 10.1002/elps.200800389

## 1 Introduction

The completion of the Human Genome Project [1, 2] has led to tremendous advances in several biotechnology fields, including medicine [3–6] and genetics [7, 8]. However, the era of personalized medicine based on the human genome cannot be fully realized until the cost of full human genome sequencing is significantly reduced from the current cost of approximately $5 million *per* genome [9].

One of the reasons for the early completion of the Human Genome Project was the technological advancement in DNA sequencing from slab gels to capillary instruments [10]. The large surface to volume ratio of capillaries reduces Joule heating by allowing for more efficient heat transfer, which thereby allows higher electric field strengths to be used. Higher electric field strengths allow for much faster sample analysis times. These shorter run times were then complemented by the bundling of many capillaries together into arrays and overall system automation to create capillary array electrophoresis instruments [11]. A similar technological advancement was made when electrophoretic separations were transferred from capillaries to microchips [12], which offer the promise of lower required reagent volumes, much faster analysis times, and the possibility of integration with other processes common to the sequencing pipeline.

Replacing capillary-based sequencing instruments with an as-yet-undeveloped microchip-based sequencing platform has been theorized to be able to reduce the cost of microchannels, one of the major recurring costs in capillary array electrophoresis, by up to 90% [13], which should significantly reduce the overall costs of Sanger-based sequencing. Integration of the DNA sequencing step with other sample preparation steps could deliver further cost and time reductions [14, 15]. However, only a limited

**Correspondence:** Associate Professor Annelise E. Barron, Department of Bioengineering, Stanford University, 318 Campus Drive, Clark Center, Room S170, Stanford, CA 94305-5444, USA
**E-mail:** aebarron@stanford.edu
**Fax:** +1650-723-8544

**Abbreviations: LPA,** linear polyacrylamide**; pDMA,** poly(*N,N*-dimethylacrylamide)

*Current address: Department of Bioengineering, Stanford University, Stanford, CA 94305, USA.

number of research groups are focused on developing new polymer materials specifically designed for microchip-based DNA sequencing separations.

The use of linear polyacrylamide (LPA) as a DNA sequencing matrix on microfluidic chips was a natural progression due to the widespread use of this material in capillaries and, before that, the use of crosslinked polyacrylamide in slab gels [16, 17]. Further studies showed that this material was adequate for DNA sequencing on microfluidic chips, achieving significant reductions in separation time (70 min down to 30 min) by using shorter separation lengths while seeing only moderate reductions in sequencing read length [18, 19]. However, the recent work of Fredlake *et al*. [20] demonstrated similarly long read lengths as previously published microchip sequencing results that utilized LPA matrices, but in separation times of less than 7 min using a high-molar mass poly(*N,N*-dimethylacrylamide) (pDMA) formulation.

That study and others have shown that inclusion of two polymers of different average molar masses can yield better separations of DNA fragments in both capillaries and microchips [18, 21–23]. Bünz *et al*. [21] showed that the blending of low- and high-molar mass batches of hydroxyethylcellulose in dilute polymer solutions enables much higher resolution separation of both small and large dsDNA restriction fragments than in either of the single molar mass matrices by themselves. Studies by the Karger group later showed that mixed molar mass LPA polymer matrices could be used for ssDNA sequencing on capillaries, which resulted in 1000-base reads in 80 min of sequencing time [22] and later, 1300-base reads in 120 min of sequencing time [23]. This idea was then transferred to microchips in another study that used blends of differently sized LPA in an 11.5 cm long microchip to achieve 580-base reads in 18 min and 640-base reads in 30 min [18]. In this study, we investigate the DNA sequencing performance of several different formulations of mixed molar mass, or blended, pDMA matrices. This is the first time that pDMA blends have been systematically investigated. We show that, by tailoring the composition of the pDMA network and using a matrix comprised of 3% 3.5 MDa pDMA and 2% 770 kDa pDMA, read lengths of up to 555 bases are possible in less than 7 min in a separation distance of 7.5 cm, while lowering the overall solution viscosity, indicating that matrix viscosity can be controlled, to a degree, independently of the sequencing performance.

## 2 Materials and methods

### 2.1 Polymer synthesis and characterization

High-purity *N,N*-dimethylacrylamide (99.5%, Monomer-Polymer & Dajac, Feasterville, PA, USA) was dissolved in 100 mL deionized water at a concentration of 5% w/w in a jacketed flask maintained at 50°C. To synthesize lower molar mass polymer, 2 mL of the chain transfer agent 2-propanol (VWR, West Chester, PA, USA) was included in the reaction flask. After removing the oxygen in the system by bubbling nitrogen for at least 30 min, 0.03 g 2,2′-azobis(2-methylpropionamidine)dihydrochloride (V-50, Wako Chemicals, Richmond, VA, USA) was added to initiate the polymerization. Each polymerization was carried out for 5 h. After polymerization, the solutions were transferred to cellulose ester dialysis membranes with a molecular weight cutoff of 100 kDa (Spectrum Labs, Rancho Dominguez, CA, USA) and dialyzed against distilled, deionized water for 2 wk with frequent water changes. The polymers were recovered by freeze-drying.

Polymer molar mass was determined by tandem gel permeation chromatography-multi-angle laser light scattering. Polymers were fractionated by gel permeation chromatography using a Waters 2690 Alliance Separation Module (Milford, MA, USA) and OHpak columns SB-806 HQ, SB-804 HQ, and SB-802.5 HQ (Shodex, New York, NY, USA). The fractions were then analyzed by an online multi-angle laser light scattering system using a DAWN DSP laser photometer and an Optilab DSP interferometric refractometer (both Wyatt Technology, Santa Barbara, CA, USA). The collected data were analyzed using WTC ASTRA 4.73 software (Wyatt Technology) to determine the weight-average molar mass, polydispersity index (PDI), and *z*-average radius of gyration ($R_g$).

Rheological testing of the polymers was performed with a Paar Physica (Ashland, VA, USA) modular compact rheometer, MCR 300, using the CP50-1 fixture. The CP50-1 is a cone-and-plate fixture with diameter of 50 mm and 1° cone. The temperature was maintained at 25°C by a Peltier system connected to an external water bath. Polymer solutions were prepared at concentrations of 5.0% w/w in buffer comprised of 49 Tris, 49 mM TAPS, 2 mM EDTA, and 7 M urea (all from Fisher Scientific, Pittsburgh, PA, USA). Shear stress sweeps were conducted from 0.1 to 1000 Pa using a logarithmic ramp and 45 data points were collected during each sweep.

### 2.2 Microchip electrophoresis

Analysis of single-stranded M13mp18 DNA sequencing fragments (Amersham/GE Healthcare, Piscataway, NJ, USA) was carried out on a custom-built system using multi-color LIF detection, as described previously [24]. Experiments were carried out in BF4-TT100 borosilicate glass microchips (Micralyne, Edmonton, Alberta, CA, USA) dynamically coated with poly(*N*-hydroxyethylacrylamide), synthesized in our laboratory, to reduce electroosmotic flow, which has been discussed previously [25, 26]. The microchips have a standard 100 μm offset T injector, a 7.5 cm separation distance, and 50 μm wide by 20 μm deep channels. Polymer matrices were loaded into the microchannel under 180 psi nitrogen pressure using a modified hydraulic press and a lab-fabricated rubber gasket to seal one end of the chips under nitrogen pressure. Samples were injected electrokinetically at 400 V/cm by grounding the sample well and applying 400 V on the sample waste well

 

for 40 s while floating the anode and cathode. After the 40 s injection, the voltages were switched such that the cathode was grounded and 2000 V were applied to the anode, yielding separation field strengths of 235 V/cm. A pullback voltage of 150 V was applied to both the sample and sample waste wells to prevent any leakage of excess sample into the separation channel. Basecalling was performed using the NNIM Basecaller (NNIM, Salt Lake County, UT, USA) and Sequencher v 4.8 (Gene Codes, Ann Arbor, MI, USA). Read lengths were calculated by determining the longest contiguous, accurately called sequence when compared with the known M13 DNA sequence at 98.5% accuracy, a commonly used sequencing accuracy threshold [18, 23, 27, 28].

## 3 Results and discussion

The individual pDMA samples used are labeled as pDMA1 and pDMA2 and are used as such throughout this discussion, with pDMA1 having a weight-average molecular weight ($M_w$) of approximately 3.5 MDa and pDMA2 having an $M_w$ of approximately 770 kDa. The physical characteristics of each polymer sample are given in Table 1 and the molar mass distributions are shown in Fig. 1. The relatively high-molar mass of pDMA1 was targeted based on a previous study that showed microchip separation of longer sequencing fragments using pDMA [20] and the lower molar mass was targeted to be in the range of $10^5$ Da, which was in the order of molecular weights used in previous studies using blended matrices [21, 23].

For the electrophoresis experiments performed in this study, the composition of the polymer separation matrix was changed while keeping the overall concentration of polymer in solution constant at 5.0% w/v. Six different matrix formulations were created with systematically varied amounts of the pDMA1 and pDMA2 polymers in each matrix, as given in Table 2. Each polymer sample was dissolved at concentration fractions of 0, 0.2, 0.4, 0.6, 0.8, and 1, with the remaining fraction being of the other polymer sample. For example, a blend to examine the effect of including a small amount of larger molar mass polymer included a final concentration of 1% w/v pDMA1 and 4% w/v pDMA2, for a total overall polymer concentration of 5% w/v.

Electrophoresis was used to perform four-color sequencing of M13 phage ssDNA in each of the six matrices. The longest and average read lengths for each matrix are presented in Table 3. In the matrix containing only the higher molar mass pDMA1, Matrix 1, average read lengths of 529 bases were obtained. As small amounts of the lower molar mass pDMA2 are included in the matrices, the read

**Table 1.** Physical characteristics of each polymer sample used in these experiments

| Polymer | $M_w$ (MDa) | $R_g$ (nm) | Polydispersity index |
|---|---|---|---|
| pDMA1 | 3.50 | 91 | 1.6 |
| pDMA2 | 0.77 | 51 | 1.4 |

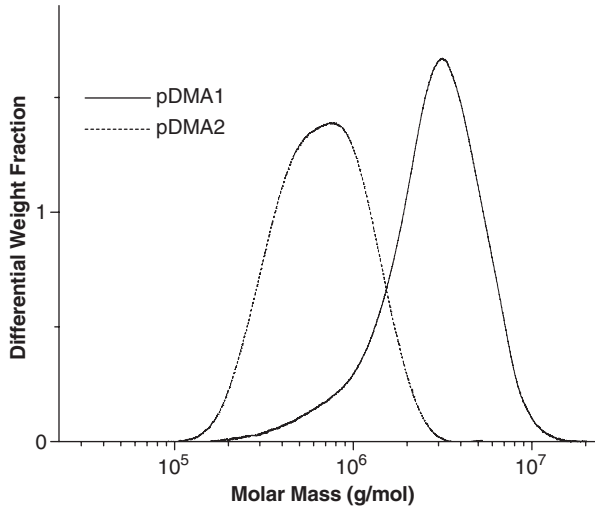


**Figure 1.** Molar mass distributions of each polymer sample used in this study. For pDMA1 (solid line), $M_w = 3.5$ MDa, PDI = 1.6, and $R_g = 91$ nm. For pDMA2 (dashed line), $M_w = 770$ kDa, PDI = 1.4, and $R_z = 51$ nm.

**Table 2.** Compositions of each polymer matrix formulation studied

| Matrix | Concentration of pDMA1 (% w/v) | Concentration of pDMA2 (% w/v) | Total polymer concentration (% w/v) |
|---|---|---|---|
| 1 | 5 | 0 | 5 |
| 2 | 4 | 1 | 5 |
| 3 | 3 | 2 | 5 |
| 4 | 2 | 3 | 5 |
| 5 | 1 | 4 | 5 |
| 6 | 0 | 5 | 5 |

**Table 3.** Sequencing read lengths, separation times, and zero-shear viscosities for each polymer blend

| Polymer | Long read length (bases) | Average read length[a)] (bases) | Separation time (min) | Zero-shear viscosity (cP) |
|---|---|---|---|---|
| Matrix 1 | 568 | 529 ± 28 | 6.9 | 41 000 ± 2000 |
| Matrix 2 | 529 | 478 ± 33 | 6.9 | 22 600 ± 400 |
| Matrix 3 | 555 | 522 ± 45 | 6.7 | 8400 ± 400 |
| Matrix 4 | 500 | 457 ± 34 | 6.6 | 2800 ± 40 |
| Matrix 5 | 440 | 400 ± 42 | 6.6 | 860 ± 25 |
| Matrix 6 | 235 | 159 ± 69 | 6.4 | 230 ± 5 |

a) Average read lengths were the result of at least four electrophoresis runs, and listed errors are the standard deviations of the data.

 

lengths are only slightly reduced. As the concentration of pDMA2 is increased to 4%, in Matrix 5, much lower read lengths are observed, with an average read length of only 400 bases. When only pDMA2 is used in the sequencing matrix, read lengths are dramatically reduced to an average of 159 bases. While the exact dependence of sequencing read length on matrix composition cannot be quantified from these data, the data indicate that the inclusion of high-molar mass polymer, even at concentrations as low as 1%, is critical to enable the separation of larger ssDNA fragments and to achieve long DNA sequencing read lengths.

Details of the differences in performance of the various blended pDMA matrices can be elucidated by examining the electropherograms. At short DNA fragment lengths (~50–100 bases), very little difference in separation quality is seen from one separation matrix to another, as shown in Fig. 2. While there is slightly lower resolution between repeated bases when Matrix 6 is used, adding just a small amount of the higher molar mass polymer (Matrix 5) yields a matrix that is able to separate these peaks. At longer DNA fragment lengths (beyond those shown in Fig. 2), the recorded peaks begin to broaden and overlap, causing a loss of resolution and the inability to call bases correctly, which limits the achieved read lengths. This loss of resolution occurs at shorter DNA fragment lengths in matrices rich in pDMA2, which accounts for the observed trend in read lengths with matrix properties given in Table 2. By looking at the electropherograms and observing the differences in read length between matrices, we can conclude that the resolution of the very short (40–150 bases) ssDNA fragments is highly dependent on the overall concentration of the polymer matrix, but not polymer molar mass *per se*, while the separation of longer (>150 bases) ssDNA fragments is highly dependent on the amount of higher molar mass polymer (pDMA1) incorporated in the sequencing matrix, which agrees in principle with previous separations

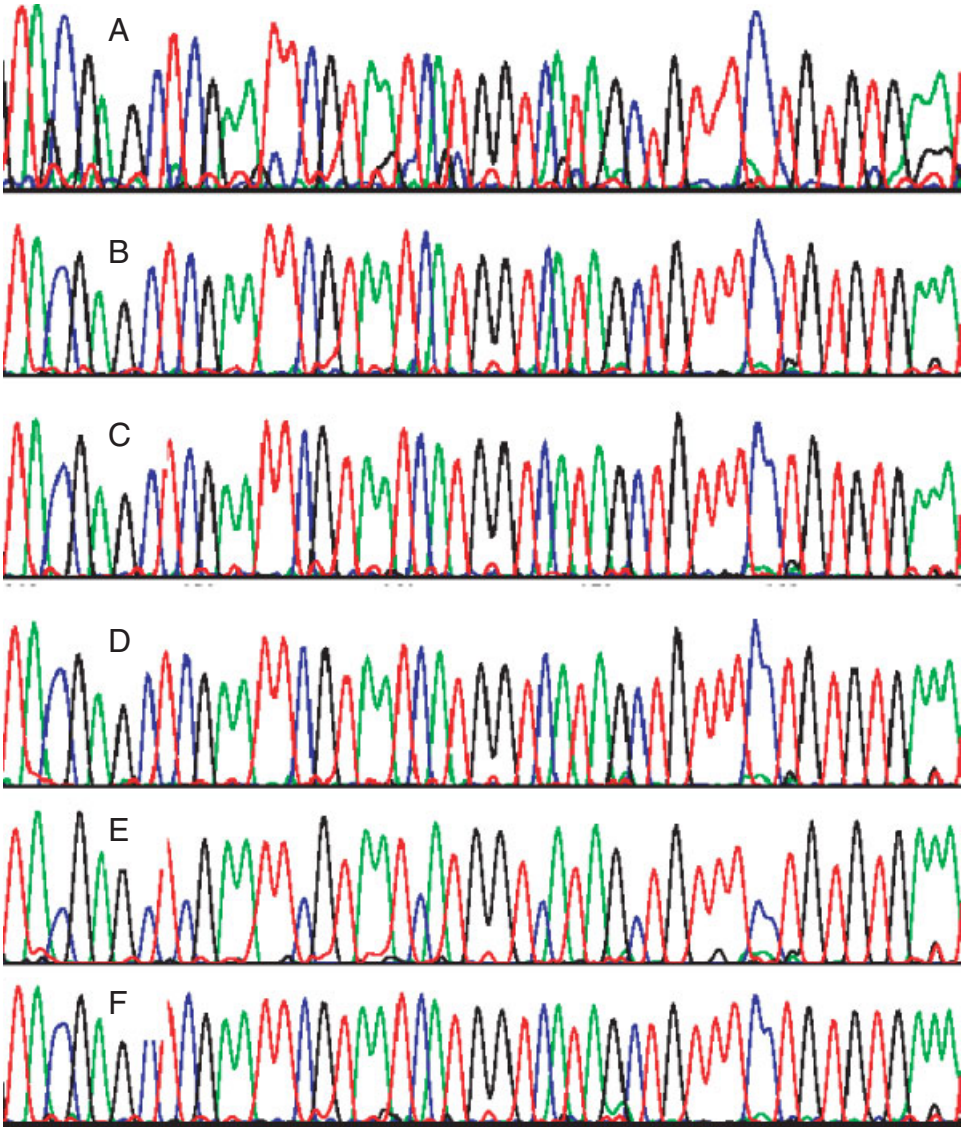


**Figure 2.** Electropherograms of the same DNA sequence in each of the six matrices tested. These fragments are approximately 50–100 bases in length. The matrices that correspond to each trace are: (A) Matrix 6, (B) Matrix 5, (C) Matrix 4, (D) Matrix 3, (E) Matrix 2, and (F) Matrix 1. With the exception of a few repeats in Matrix 6, these fragments show very similar separation quality, indicating that the greatest differences in read length are at the larger fragment sizes.

 

of dsDNA [21], as well as of ssDNA sequencing fragments [18, 22]. However, we have also shown that the separation of the larger ssDNA fragments is possible even at pDMA1 concentrations as low as 1% w/v, indicating that a matrix need not be composed primarily of high-molar mass chains to achieve separation of these fragments. This separation performance is not seen in Matrix 6, comprised of only pDMA2, which indicates that at least some minimum amount of higher molar mass pDMA must be present as an element of the entangled polymer network, to separate larger ssDNA fragments.

Another important factor in developing a polymer matrix for microchip DNA sequencing is formulating a solution that can be loaded easily into a microchannel. As reported in Table 3, the average sequencing read length of Matrix 1 is only slightly higher than the average sequencing read length enabled by Matrix 3. However, the viscosity of the blended Matrix 3 is much lower than that of Matrix 1. Also reported in Table 3 are the zero-shear viscosities for each matrix that was formulated. As the concentration of pDMA1 in the matrix increases, the zero-shear viscosity of the polymer matrix increases exponentially. In this context, Matrix 3 becomes much more attractive because it has a very similar average read length to the 5% pDMA1, but has a zero-shear solution viscosity that is 80% lower than that of Matrix 1, which then lowers the loading requirements imposed on microchips and the equipment that is used to fill their channels. These results indicate that it is possible to control polymer matrix viscosity somewhat independently of DNA sequencing read length performance by carefully controlling the sequencing matrix composition at a given overall polymer concentration.

## 4 Concluding remarks

In this study, we have shown the advantages of blended pDMA matrices, comprised of both high- (3.5 MDa) and low- (770 kDa) molar mass polymer at specific concentrations, to perform DNA sequencing by electrophoresis in a microfluidic device. The concentration of all matrices was held constant at 5% w/v and the fractional concentrations of the polymer samples were varied. We have shown that similar resolution of short sequencing fragments was achieved in all matrices, indicating that separation of these fragments is strongly dependent on the overall concentration of the matrices. We have also shown that one particular matrix, comprised of 3% 3.5 MDa pDMA and 2% 770 kDa pDMA, achieves nearly the same sequencing read length at 80% lower solution viscosities than the unblended 5% 3.5 MDa pDMA. Our data show that an engineered mixture of pDMA of low- and high-average molecular weights permits the separation of both small and large ssDNA sequencing fragments at much lower viscosities, which could facilitate the creation of an automated, integrated microchip sequencing device due to lower pressure microchip loading requirements.

*This work was supported by the National Human Genome Research Institute, of the NIH (Grant 2 R01 HG001970 and Grant 5 R01 HG003583* via *Microchip Biotechnologies) and C. P. F. was also partially supported by the Malkin Family Foundation.*

*The authors have declared no 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*, 1304–1351.

[3] Collins, F. S., McKusick, V. A., *J. Am. Med. Assoc.* 2001, *285*, 540–544.

[4] Kurian, S., Grigoryev, Y., Head, S., Campbell, D. *et al.*, *Int. Immunopharmacol.* 2007, *7*, 1948–1960.

[5] Sanders, J., Gill, M., *Ir. J. Med. Sci.* 2007, *176*, 5–9.

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

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

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

[9] NIH News Release. New Grants Bolster Efforts to Generate Faster and Cheaper Tools for DNA Sequencing. August 2007. http://www.genome.gov/25522229.

[10] Guttman, A., Cohen, A. S., Heiger, D. N., Karger, B. L., *Anal. Chem.* 1990, *62*, 137–141.

[11] Huang, X. H. C., Quesada, M. A., Mathies, R. A., *Anal. Chem.* 1992, *64*, 967–972.

[12] Woolley, A. T., Mathies, R. A., *Anal. Chem.* 1995, *67*, 3676–3680.

[13] Fredlake, C. P., Hert, D. G., Mardis, E. R., Barron, A. E., *Electrophoresis* 2006, *27*, 3689–3702.

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

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

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

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

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

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

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

 

[21] Bünz, A. P., Barron, A. E., Prausnitz, J. M., Blanch, H. W., *Ind. Eng. Chem. Res.* 1996, *35*, 2900–2908.

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

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

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

[25] Albarghouthi, M. N., Buchholz, B. A., Huiberts, P. J., Stein, T. M., Barron, A. E., *Electrophoresis* 2002, *23*, 1429–1440.

[26] Albarghouthi, M. N., Stein, T. M., Barron, A. E., *Electrophoresis* 2003, *24*, 1166–1175.

[27] Kotler, L., He, H., Miller, A. W., Karger, B. L., *Electrophoresis* 2002, *23*, 3062–3070.

[28] Liu, S. R., *Electrophoresis* 2003, *24*, 3755–3761.

 