Anal. Chem. **2004,** *76,* 5249−5256

Articles

Sparsely Cross-Linked "Nanogel" Matrixes as Fluid, Mechanically Stabilized Polymer Networks for High-Throughput Microchannel DNA Sequencing

Erin A. S. Doherty,[†] Cheuk-Wai Kan,[†] Brian M. Paegel,[‡] Stephanie H. I. Yeung,[§] Shitong Cao,[†] Richard A. Mathies,[‡] and Annelise E. Barron*,[†]

Department of Chemical and Biological Engineering, Northwestern University, Evanston, Illinois 60208, and Department of Chemistry, and UCB/UCSF Joint Graduate Group in Bioengineering, University of California, Berkeley, California 94720

We have developed sparsely cross-linked "nanogels", subcolloidal polymer structures composed of covalently linked, linear polyacrylamide chains, as novel replaceable DNA sequencing matrixes for capillary and microchip electrophoresis. Nanogels were synthesized via inverse emulsion (water-in-oil) copolymerization of acrylamide and a low percentage (∼10^{-4} mol %) of *N*,*N*-methylene bisacrylamide (Bis). Nanogels and nanogel networks were characterized by multiangle laser light scattering and rheometry, respectively, and tested for DNA sequencing in both capillaries and chips with four-color LIF detection. Typical nanogels have an average radius of ∼230 nm, with ∼75% of chains incorporating a Bis cross-linker. The properties and performance of nanogel matrixes are compared here to those of a linear polyacrylamide (LPA) network, matched for both polymer weight-average molar mass (M_w) and the extent of interchain entanglements (c/c^*). At sequencing concentrations, the two matrixes have similar flow characteristics, important for capillary and microchip loading. However, because of the physical network stability provided by the internally cross-linked structure of the nanogels, substantially longer average read lengths are obtained under standard conditions with the nanogel matrix at a 98.5% accuracy of base-calling (for CE: 680 bases, an 18.7% improvement over LPA, with the best reads as long as 726 bases, compared to 568 bases for the LPA matrix). We further investigated the use of the nanogel matrixes in a high-throughput microfabricated DNA sequencing device consists of 96 separation channels densely fabricated on a 6-in. glass wafer. Again, preliminary DNA sequencing results show that the nanogel matrixes are capable of delivering significantly longer average read length, compared to an LPA matrix of comparable properties. Moreover, nanogel matrixes require 30% less polymer per unit volume than LPA. The addition of a small amount of low molar mass LPA or ultrahigh molar mass LPA to the optimized nanogel sequencing matrix further improves read length as well as the reproducibility of read length (RSD < 1.6%). This is the first report of a replaceable DNA sequencing matrix that provides better performance than LPA, in a side-by-side comparison of polymer matrixes appropriately matched for molar mass and the extent of interchain entanglements. These results could have significant implications for the improvement of microchip-based DNA sequencing technology.

Capillary electrophoresis (CE) of DNA through polymeric separation matrixes is presently the dominant technology for high-throughput sequencing. Hundreds of genome projects,[1] as well as individualized genomics, still require long DNA sequencing read lengths at low cost; continual improvement of polymeric materials used for DNA separation by microchannel electrophoresis is required to make the long-term goal of personalized genomics economically feasible. In this paper, we report the promising DNA sequencing performance of a novel polymeric material developed in our laboratory[2] for DNA sequencing in both capillary and chip electrophoresis systems. The "nanogel" networks we demonstrate here have the clear potential to combine the high-selectivity separations of a cross-linked slab gel with the replaceable nature of a separation matrix composed of linear polymer.

Originally, DNA sequencing was performed on highly cross-linked polyacrylamide slab gels.[3,4] Cross-linked polyacrylamide

* To whom correspondence should be addressed. Phone: (847) 491-2778. Fax: (847) 491-3728. E-mail: a-barron@northwestern.edu.

[†] Northwestern University.

[‡] Department of Chemistry, University of California, Berkeley.

[§] UCB/UCSF Joint Graduate Group in Bioengineering, University of California, Berkeley.

(1) Bernal, A.; Ear, U.; Kyrpides, N. *Nucleic Acids Res.* **2001**, *29*, 126−127.

(2) Doherty, E. A. S.; Kan, C. W.; Barron, A. E. *Electrophoresis* **2003**, *24*, 4170−4180.

(3) Chrambach, A.; Rodbard, D. *Science* **1971**, *172*, 440−450.

(4) Bishop, D. H.; Claybrook, L., Jr.; Spiegel, M. S. *J. Mol. Biol.* **1967**, *26*, 373−&.

10.1021/ac049721x CCC: $27.50
Published on Web 08/06/2004

yields excellent DNA separations, allowing for long reads under optimized conditions, but typically requires the use of low electric fields, which translates into long (5−8 h) run times.[5] The process of obtaining long DNA sequencing reads using ultrathin slab gels is time- and labor-intensive. For this reason, high-throughput genome sequencing centers largely abandoned slab gels in the late 1990s, in favor of automated capillary array electrophoresis (CAE).[6−10]

Although researchers created in situ-polymerized capillary gels capable of sequencing reads up to 350 bases,[11−18] the direct transfer of a "slab gel" technology to the micrometer-sized channels typical of fused-silica capillaries was not effective for a variety of reasons. First, voids left within the capillary gel due to the increased density of the polymer relative to its monomer are detrimental to efficient separations.[11,19] Second, an in situ-polymerized, highly cross-linked gel structure is difficult to remove from the capillary, making these prepared capillaries useful for only a small number of separations each. Finally, since there is no a priori knowledge of the final polymer properties, rigorous quality control is not possible for in situ-polymerized matrixes.

The use of physically entangled, linear polymer solutions for the separation of DNA sequencing fragments within capillaries allowed for faster run times as well as relatively facile loading and replacement of the separation matrix between runs,[20] enabling increased automation of DNA sequencing. Moreover, production and characterization of polymers ex situ has allowed researchers to correlate polymer physical and chemical properties with DNA separation performance. The chemical and physical properties of polymers used for microchannel DNA sequencing are critically important, as they control the time scale of polymer−polymer and polymer−DNA interactions within the entangled polymer network, which in turn influences the mechanism of DNA separation.[19,21] For high-throughput sequencing, the choice of matrix polymer is a crucial determinant of success.

A range of linear polymers has shown good utility for use in DNA sequencing, including linear polyacrylamide (LPA),[22,23] poly(*N*,*N*-dimethylacrylamide),[24,25] poly(vinylpyrrolidone),[26] poly(*N*-hydroxyethylacrylamide) (polyDuramide),[27] and copolymers of *N*,*N*-dimethylacrylamide and *N*,*N*-diethylacrylamide.[28,29] Polyacrylamide is a near-ideal polymer for DNA sequencing due to its high hydrophilicity (hence, its excellent ability to entangle with other polyacrylamide chains in aqueous solution) as well as its facile production to high molar mass by free-radical polymerization chemistry. To date, high molar mass LPA gives the best sequencing performance and is able to produce (at the upper limit) a 1000-base read in about 1 h[30] and 1300 bases in 2 h[23] with highly optimized polymer molar mass distribution, matrix formulation, capillary wall coating, sample preparation and cleanup, and base-calling algorithms.

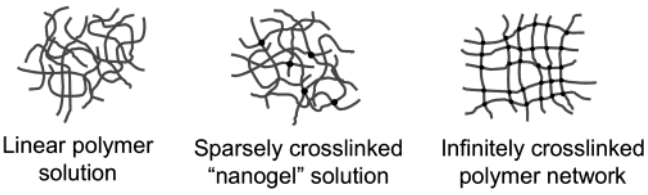

Figure 1. Schematic representation of a concentrated solution of sparsely cross-linked nanogels. Sparsely cross-linked nanogels occupy a middle ground, between an entangled solution of linear polymers and an infinitely cross-linked polymer network.

Chemical cross-links within slab gels produce an infinitely cross-linked separation medium and provide a mechanically stabilized pore structure for the migration of DNA. An abundance of cross-links within a slab gel decreases the effective pore size and limits sample diffusion and dispersion during separation,[31] which is desirable to produce narrow bands on the gel. Ideally, the presence of chemical cross-links in a high molar mass polymer for CE would provide the same benefits. The sparsely cross-linked nanogel networks we describe here occupy a new middle ground, between highly cross-linked, in situ-produced polyacrylamide capillary gels and the fluid, linear polymer networks that are now utilized for capillary array electrophoresis (CAE) (see Figure 1).

Sparsely cross-linked, discrete polymer nanogels must be synthesized in such a way as to avoid the production of large polymers of colloidal dimensions, which would scatter incident light and prevent DNA sequencing with laser-induced fluorescence (LIF) detection. The production of an infinitely cross-linked polymer gel ex situ must also be avoided, and the cross-link density of the nanogels should be limited, so that the final sequencing matrix retains good fluidity and so that individual polymer structures may physically entangle with each other.

Our preliminary work with sparsely cross-linked nanogel matrixes demonstrated that stabilization of a physically entangled

(5) Ansorge, W.; Voss, H.; Weimann, S.; Schwager, C.; Sproat, B.; Zimmermann, J.; Stegemann, J.; Erfle, H.; Hewitt, N.; Rupp, T. *Electrophoresis* **1992**, *13*, 616−619.
(6) Cheng, J. *Prog. Biochem. Biophys.* **1995**, *22*, 223−227.
(7) Huang, X. C.; Quesada, M. A.; Mathies, R. A. *Anal. Chem.* **1992**, *64*, 967−972.
(8) Huang, X. C.; Quesada, M. A.; Mathies, R. A. *Anal. Chem.* **1992**, *64*, 2149−2154.
(9) Kambara, H.; Takahashi, S. *Nature* **1993**, *361*, 565−566.
(10) Mathies, R. A.; Huang, X. C. *Nature* **1992**, *359*, 167−169.
(11) Heiger, D. N.; Cohen, A. S.; Karger, B. L. *J. Chromatogr.* **1990**, 33−48.
(12) Cohen, A. S.; Najarian, D. R.; Karger, B. L. *J. Chromatogr.* **1990**, *516*, 49.
(13) Guttman, A.; Cohen, A. S.; Heiger, D. N.; Karger, B. L. *Anal. Chem.* **1990**, *62*, 137−141.
(14) Harke, H. R.; Bay, S.; Zhang, J. Z.; Rocheleau, M. J.; Dovichi, N. J. *J. Chromatogr.* **1992**, *608*, 143−150.
(15) Smith, L. M. *Nature* **1991**, *349*, 812−813.
(16) Drossman, H.; Luckey, J. A.; Kostichka, A. J.; Dcunha, J.; Smith, L. M. *Anal. Chem.* **1990**, *62*, 900−903.
(17) Luckey, J. A.; Drossman, H.; Kostichka, A. J.; Mead, D. A.; Dcunha, J.; Norris, T. B.; Smith, L. M. *Nucleic Acids Res.* **1990**, *18*, 4417−4421.
(18) Baba, Y.; Matsuura, T.; Wakamoto, K.; Morita, Y.; Nishitsu, Y.; Tsuhako, M. *Anal. Chem.* **1992**, *64*, 1221−1225.
(19) Bae, Y. C.; Soane, D. *J. Chromatogr., A* **1993**, *652*, 17−22.
(20) Ruiz-Martinez, M. C.; Berka, J.; Belenkii, A.; Foret, F.; Miller, A. W.; Karger, B. L. *Anal. Chem.* **1993**, *65*, 2851−2858.
(21) Albarghouthi, M.; Barron, A. E. *Electrophoresis* **2000**, *21*, 4096−4111.
(22) 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.
(23) Zhou, H.; Miller, A. W.; Sosic, Z.; Buchholz, B.; Barron, A. E.; Kotler, L.; Karger, B. L. *Anal. Chem.* **2000**, *72*, 1045−1052.
(24) Madabhushi, R. S.; Vainer, M.; Dolnik, V.; Enad, S.; Barker, D. L.; Harris, D. W.; Mansfield, E. S. *Electrophoresis* **1997**, *18*, 104−111.
(25) Madabhushi, R. S. *Electrophoresis* **1998**, *19*, 224−230.
(26) Gao, Q. F.; Yeung, E. S. *Anal. Chem.* **1998**, *70*, 1382−1388.
(27) Albarghouthi, M. N.; Buchholz, B. A.; Huiberts, P. J.; Stein, T. M.; Barron, A. E. *Electrophoresis* **2002**, *23*, 1429−1440.
(28) Albarghouthi, M. N.; Buchholz, B. A.; Doherty, E. A. S.; Bogdan, F. M.; Zhou, H.; Barron, A. E. *Electrophoresis* **2001**, *22*, 737−747.
(29) Buchholz, B. A.; Doherty, E. A. S.; Albarghouthi, M. N.; Bogdan, F. M.; Zahn, J. M.; Barron, A. E. *Anal. Chem.* **2001**, *73*, 157−164.
(30) Salas-Solano, O.; Carrilho, E.; Kotler, L.; Miller, A. W.; Goetzinger, W.; Sosic, Z.; Karger, B. L. *Anal. Chem.* **1998**, *70*, 3996−4003.
(31) Ugaz, V. M.; Burke, D. T.; Burns, M. A. *Electrophoresis* **2002**, *23*, 2777−2787.

sequencing matrix is possible by including an extremely small amount of bifunctional monomer, which provides chemical cross-link points that fortify the final entangled polymer network and improve sequencing results.[2] In the present work, we show that through optimization of the nanogel cross-link density and blending with appropriate linear polymers to improve the matrix selectivity for small and large ssDNA fragments, a flowable, sparsely cross-linked DNA sequencing matrix formulation can provide routine read lengths of more than 680 bases at 98.5% accuracy in less than 120 min in a commercial CAE system. Moreover, we demonstrate the ability of these matrixes to allow ~600-base read lengths in a microfabricated capillary array electrophoresis (μCAE) sequencing system, in an analysis time of <25 min. This is the only reported nonlinear DNA sequencing polymer that outperforms an appropriately matched LPA matrix and can be easily replaced before each DNA sequencing run.

EXPERIMENTAL SECTION

Reagents. Tris(hydroxymethyl)aminomethane (Tris), ethylenediaminetetraacetic acid (EDTA), ultrapure grade *N*,*N*,*N*′,*N*′-tetraethylmethylenediamine (TEMED), ultrapure acrylamide, methylene bisacrylamide (Bis), and ammonium persulfate (APS) were purchased from Amresco (Solon, OH). *N*-[Tris(hydroxymethyl)methyl]-3-aminopropanesulfonic acid (TAPS) was obtained from Sigma (St. Louis, MO). Sorbitan monooleate (Span 80) was obtained from Fluka Chemical (St. Louis, MO). Isopar M (a C_{12}−C_{14} isoparaffinic mixture) was obtained from Exxon (Houston, TX). MegaBACE Sequencing Standards (Amersham Pharmacia Biotech, Piscataway, NJ) consisting of M13 DNA sequencing reaction products, labeled with energy-transfer dye primers and purified by standard ethanol precipitation by the manufacturer, were used without further purification. Beckman LongRead matrix was provided by Amersham Pharmacia Biotech.

Polymer Synthesis. Sparsely cross-linked nanogels were produced via inverse emulsion polymerization according to a protocol described in a previous report.[2] The emulsion was formed as follows: A mixture of 40.53 g of Isopar M and 2.47 g of Span 80 was mixed briefly before being poured into the reaction vessel, a 500-mL water-jacketed reaction flask with a four-neck top (Kontes, Vineland, NJ). The organic phase was immediately mixed by a high-torque overhead stirrer (Caframo Limited, Wiarton, Ontario, Canada) equipped with a homemade stainless steel shaft and a four-blade pitched-blade impeller, set at 600 rpm. The organic phase was degassed for 30 min using prepurified nitrogen (Air Products, Naperville, IL) that was further purified using an in-line oxygen/water vapor trap (Supelco, St. Louis, MO). A mixture of 22.8 g of acrylamide and 34.2 g of water was then added dropwise to the organic phase, resulting in a white, opaque emulsion. If sparsely cross-linked nanogels were desired, ~1 mL of an aqueous solution having the appropriate concentration of bifunctional monomer (Bis) was added immediately following the addition of the monomer solution. To produce the nanogels described in this work, Bis composed ~10^{-4} mol % of the total number of moles of acrylamide monomer (see Table 1).

The resulting emulsion was then degassed for 1 h or until the emulsion temperature (measured and recorded using a Jenco Electronics (Lazar Research Laboratory, Los Angeles, CA thermocouple and data logger) exceeded 35 °C, whichever was longer. Prior to initiation, a small aliquot of the emulsion was checked to ensure no autopolymerization had occurred. Polymerization was initiated with APS/TEMED, both at a concentration of 0.005 wt % (based on the mass of the aqueous phase). The reaction was allowed to proceed for 16 h under continuous mixing and degassing. The final product was precipitated by adding the product emulsion dropwise to a large volume of methanol with stirring. The precipitated polymer was washed copiously with methanol during filtration; product was dried in a room-temperature vacuum oven for at least 72 h.

Table 1. Polymer Physical Properties

polymer	mol % Bis ($\times 10^5$)	M_w ($\times 10^{-6}$ g/mol)	R_g (nm)	c^* (mg/mL)
LPA	0.00	8.435 ± 0.759	331 ± 12	1.21
LPA−LMM	0.00	0.34	51	n/a
LPA−HMM	0.00	13.85 ± 1.88	251 ± 10	1.03
LPA−UHMM	0.00	20.56 ± 4.63	343 ± 27	0.71
Nanogel-2×	15.72	10.17 ± 1.43	230 ± 16	0.80
Nanogel-4×	31.44	11.45 ± 1.82	220 ± 18	0.89
Nanogel-5×	39.30	10.20 ± 1.15	236 ± 10	0.81
Nanogel-6×	45.16	n/a[a]	n/a	n/a

[a] Nanogel-6× matrix did not completely dissolve; MALLS and rheometry did not yield useful physical property data.

Isolation of the Commercially Available Polymer. The commercially available LPA was recovered from prepackaged sequencing matrixes (LongRead matrix Lot 1681, Beckman Coulter, Inc., Fullerton, CA) by dilution with deionized, distilled water, and dialysis against deionized, distilled water using 1000 molecular weight cutoff (MWCO) cellulose ester membranes (Spectrum Laboratories, Inc., Rancho Dominguez, CA). The polymer solution was then frozen and lyophilized using a freeze-dry system (Labconco, Kansas City, MO), resulting in a white, foamlike mass. Polymer was redissolved at the desired concentration in a solvent or buffer of interest by slow rotation for at least 24 h (Roto-Torque, Cole-Parmer Instrument Co., Inc., Vernon Hills, IL).

Rheological Characterization. Rheometry was performed using an Anton Paar Physica (Glen Allen, VA) MCR 300 equipped with a Julabo USA, Inc. (Allentown, PA) digitally controlled recirculating water bath. Steady-shear rheometry was performed with a double-gap Couette fixture (model DG26.7) at 288 s^{-1}. Shear-dependent rheometry was performed with a 50-mm, 2° cone-and-plate fixture (model CP50-2).

Multiangle Laser Light Scattering (MALLS). The weight-average molar mass and radius of gyration of high molar mass polymer samples were determined by batch MALLS using a DAWN DSP Laser Photometer-Optilab DSP Interferometric Refractometer system (both, Wyatt Technology, Santa Barbara, CA). Data collection and analysis were performed as described in previous reports.[32]

DNA Sequencing and Data Analysis with Capillary Array Electrophoresis. DNA sequencing was performed on a MegaBACE 1000 CAE instrument (Molecular Dynamics, Sunnyvale, CA) equipped with 6 × 16 fused-silica capillary arrays (75-μm inner diameter, 64-cm total length, 40-cm effective length) covalently coated by the manufacturer with LPA. Polymers to be tested as DNA sequencing matrixes were dissolved at the

(32) Buchholz, B. A.; Barron, A. E. *Electrophoresis* **2001**, *22*, 4118−4128.

concentration of interest in a 1× TTE (50 mM Tris, 50 mM TAPS, 2 mM EDTA) buffer with 7 M urea. Sequencing matrixes were loaded under a pressure of 1000 psi, followed by a polymer 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, separation of DNA was performed at 140 V/cm and 44 °C for 120 min. LIF data were collected, analyzed, and translated into called DNA sequence using the MegaBACE 1000 DNA Sequencing Software Version 2.0.

Raw LIF data were extracted from the MegaBACE sequencing software and fitted into Gaussian peaks using PeakFit 4.06 (SPSS, Chicago, IL) from which the full width at half-maximum (fwhm) and the migration time were calculated for each peak. An equation for migration time as a function of base number was determined by fitting the data provided by PeakFit to a third-order polynomial function, a trend observed within a high molar mass LPA matrix by Karger and co-workers[22] and Barron and co-workers.[27] This equation was used to calculate the selectivity of the sequencing matrix as

$$S_n = 2(\mu_n - \mu_{n-1})/(\mu_n + \mu_{n-1}) \quad (1)$$

where μ_n is the apparent mobility of the nth DNA fragment and μ_{n-1} is the apparent mobility of the $(n-1)$th DNA fragment. Peak width (fwhm) as a function of base number was plotted and fit to a second-order polynomial. This function best modeled the experimental data; a similar empirical fitting function has been successfully employed to characterize other sequencing matrixes.[25,33]

Microdevice Fabrication and Design. The microfabrication procedure and chip design are described in previous reports.[34,35] Briefly, Borofloat glass wafers (150-mm diameter; Schott, Yonkers, NY) coated with a 2000-Å amorphous silicon film and spin-coated with photoresist (S1818, Shipley, Marlborough, MA) were used. The channel pattern was exposed and developed using standard photolithography processes and then transferred to the silicon film by reactive ion etching with SF_6 plasma to expose the glass layer. The wafer was then isotropically etched in concentrated HF for 3.5 min to create channels with a depth of 25 μm. Reservoirs were diamond-drilled using a CNC mill, and after cleaning, this glass wafer was thermally bonded to a blank glass substrate at 670 °C to form a closed-channel sandwich structure.

The 96 separation channels were arrayed around a common central anode. Adjacent pairs of lanes are grouped to share cathode and waste reservoirs. The widths of the separation channels are 200 μm after etching and the effective separation length is 15.9 cm.

Microdevice Operation. The microchannels were precoated with LPA according to a modified Hjerten procedure.[36] Before matrix loading, the separation channels and reservoirs at the cathode end of the array are filled with deionized water. The linear polyacrylamide sieving matrix is then forced in through the anode access port at 300 psi for 3 min by using a high-pressure loading system.[37] Excess matrix is removed from the sample wells and 1.8 μL of sample introduced. The cathode and waste buffer moats are filled with 3 mL of 5× TTE each. An elastomer buffer reservoir is affixed above the central anode access port and filled with 3 mL of 5× TTE. The filled μCAE device is placed on a stage heated to 67 °C on the four-color confocal rotary scanner. An electrode ring array is placed over the chip to provide electrophoresis voltages to the various reservoirs. A PDMS ring was applied to the chip to create a continuous buffer reservoir to electrically address the cathode and waste reservoirs simultaneously.

The sample is injected electrokinetically from the sample reservoir to the injection intersection at 190 V/cm for 35 s. During this period, the anode and cathode reservoirs are floated. The plug defined by the 250-μm intersection is driven down the separation column at 150 V/cm while a back-biasing electric field of 200 V/cm is applied at the sample and waste reservoirs to withdraw excess sample and prevent leakage from the reservoirs onto the column during the separation. Electrophoretic analysis is complete in 24 min.

Data Acquisition and Reduction. The scanning system involves a four-color rotary scanner.[38] Briefly, excitation at 488 nm from an argon ion laser is coupled into the optical path with a dichroic beam splitter. The laser light is passed through the hollow shaft of a stepper motor, displaced from the rotation axis by 1 cm with a rhomb prism, and focused on the microchannels through a 60× objective. Fluorescence is collected by the objective and passed back through the hollow shaft and dichroic beam splitter and sorted into four spectral channels. Data acquisition procedures have been described elsewhere.[35] After electrophoretic analysis, the four-color electropherograms are used as input for the MegaBACE Sequence Analyzer.

RESULTS AND DISCUSSION

Nanogels were synthesized via inverse emulsion polymerization, precipitated in methanol, washed, filtered, and dried, resulting in a free-flowing white powder that is easily handled. Resulting nanogels could be dissolved in sequencing buffer in 48−72 h by slow rotation and/or slow mechanical stirring. Nanogels have a molar mass of (10−12) × 10^6 g/mol and a Z-average radius of gyration of 220−240 nm. These polymer structures are discrete and sparsely internally cross-linked and are able to form highly entangled matrixes for DNA sequencing. Although other researchers have produced nonlinear polyacrylamide-based structures for dsDNA separations,[39,40] these relatively low molar mass (~1 × 10^6 g/mol) copolymers were designed to combine the properties of their respective linear homopolymers and were not tested as DNA sequencing matrixes.

The initial proof-of-principle results were obtained using nanogels with 7.86 × 10^{-5} mol % Bis (Nanogel-1×)[2]. Assuming complete conversion of all monomer to a monodisperse polymer

(33) Menchen, S.; Johnson, B.; Winnik, M. A.; Xu, B. *Electrophoresis* **1996**, *17*, 1451−1459.
(34) Simpson, P. C.; Roach, D.; Wooley, A. T.; Thorsen, T.; Johnston, R.; Sensabaugh, G. F.; Mathies, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 2256−2261.
(35) Paegel, B. M.; Emrich, C. A.; Wedemayer, G. J.; Scherer, J. R.; Mathies, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 574−579.
(36) Hjerten, S. *J. Chromatogr.* **1985**, *347*, 191−198.
(37) Scherer, J. R.; Paegel, B. M.; Wedemayer, G. J.; Emrich, C. A.; Lo, J.; Medintz, I. L.; Mathies, R. A. *Biotechniques* **2001**, *31*, 1150.
(38) Shi, Y. N.; Simpson, P. C.; Scherer, J. R.; Wexler, D.; Skibola, C.; Smith, M. T.; Mathies, R. A. *Anal. Chem.* **1999**, *71*, 5354−5361.
(39) Sudor, J.; Barbier, V.; Thirot, S.; Godfrin, D.; Hourdet, D.; Millequant, M.; Blanchard, J.; Viovy, J.-L. *Electrophoresis* **2001**, *22*, 720−728.
(40) Barbier, V.; Buchholz, B. A.; Barron, A. E.; Viovy, J. L. *Electrophoresis* **2002**, *23*, 1441−1449.

having a molar mass of 10×10^6 g/mol (see below) containing 8×10^{-5} mol % Bis, ~15% of the initiated polymers in these previously studied materials contain one or more Bis cross-links at the 1× cross-link density. In this work, optimization of these structures with respect to cross-link density was investigated using nanogels with 6.29×10^{-5}–4.72×10^{-4} mol % Bis (0.8×–6×). For the optimized cross-link density presented in this work, ~75% of the polymer chains in the nanogels incorporate at least one point of cross-linking.

The average molar mass and *Z*-average radius of gyration of the nanogels were measured using MALLS in batch mode.[32] Overlap threshold concentrations, which correlate with the onset of polymer–polymer entanglement interactions, were determined using steady-shear rheometry to measure specific viscosity of the polymer solution as a function of concentration. Physical properties of the linear polymers and sparsely cross-linked polymer nanogels presented in this work are shown in Table 1.

Comparison of linear polymer matrixes and matrixes composed of sparsely cross-linked nanogels must be done on an equal basis; i.e., the properties of the linear polymer matrix and the sparsely cross-linked polymer matrix should be as similar as possible, notwithstanding the presence or absence of chemical cross-links. It has been shown that the ratio of polymer concentration to polymer overlap concentration (c/c^*) can be used to "match" different, highly entangled linear polymer solutions used as DNA sequencing matrixes. This then allows the performance of these matrixes to be compared on an equal basis.[41–43] The extent of polymer–polymer entanglements controls the lifetime of the virtual polymer "tube" that the DNA migrates through while being driven by the electric field. Although scaling laws governing the mesh size of the network differ with the exact chemical structure of the polymer (linear vs cross-linked), the use of the c/c^* matching to match mesh size when comparing an entangled linear polymer network to an entangled polymer network containing very sparse cross-links should still provide an accurate comparison of the two networks. In this work, matrixes were matched according to the extent of interchain entanglements (c/c^*) and according to average molar mass.

Preliminary work with nanogels having a cross-link density of $\sim 8 \times 10^{-5}$ mol % indicated that the presence of chemical cross-links at that density improves CE sequencing read lengths by 10.4% compared to the appropriately matched LPA matrix. This improvement in read length was provided by the improved selectivity of ssDNA fragments longer than 350 bases in the nanogel sequencing matrix.[2] Our present work shows that the optimization of nanogel cross-link density provides further improvement in sequencing performance compared with LPA (see Table 2). In particular, sparsely cross-linked acrylamide matrixes showed substantial increases in read length at 98.5% accuracy; the Nanogel-5× matrix provided the longest read lengths at 98.5% accuracy as well as an 18.7% improvement in read length (>100 bases) over the c/c^*-matched LPA matrix obtained in a CAE instrument (Figure 2). Using CAE, a routine average read length of 674 bases could be obtained in the nanogel matrix; in the best case, high-accuracy read lengths as long as 726 bases were achieved. By comparison, the c/c^*-matched LPA matrix gave 568-base reads at best.

Table 2. DNA Sequencing Read Lengths at 98.5% Accuracy[a]

polymer	concn (w/v)%	av read length[b]	best read length	% improvement (average)
LPA	4.47		512	
Nanogel-2×	3.00	629 ± 28 (n = 14)	673	22.9
LPA	3.70		626	
Nanogel-4×	2.75	655 ± 25 (n = 14)	685	4.6
LPA	4.05		568	
Nanogel-5×	2.75	674 ± 32 (n = 14)	726	18.7
LPA				
Nanogel-6×	3.00	549 ± 55 (n = 9)[c]	618	

[a] LPA is c/c^* matched for each nanogel cross-link density. [b] Error indicates the standard deviation in the data. [c] Nanogel-6× matrix did not completely dissolve, resulting in a cloudy, inhomogeneous solution that was unable to provide routine sequencing of >600 bases.

Further performance testing of the nanogel networks by microchip electrophoresis suggests that they will be useful for high-throughput DNA sequencing in microfluidic devices. While the development of microfabricated, multichannel electrophoresis systems represents a major breakthrough for high-throughput and low-cost genotyping, the need for a practical DNA sequencing matrix that enables long-read DNA sequencing has so far been one factor preventing their widespread practical application. A DNA sequencing read length of ~480 bases in 25 min with base-calling accuracy of 98.5% was reported previously with the use of a homemade, difficult-to-load, high molar mass LPA in the same microfabricated device described in this paper, with optimized electrophoresis conditions.[35] Within a sequencing chip, the Nanogel-5× matrix provides read lengths of ~500 bp with 98% accuracy in 25 min without any optimization of electrophoresis conditions and is easier to load into chip microchannels. For the chip-based sequencing separation with nanogels, this represents a major improvement when compared with commercially available, high-performance DNA sequencing matrix based on high molar mass LPA. The electrophoretic conditions will be further optimized to improve the DNA sequencing performance.

It is notable that above the 5× cross-link density, the nanogels did not fully dissolve in sequencing buffer, forming a cloudy, inhomogeneous polymer solution that was unable to provide routine read lengths over 600 bases at 98.5% accuracy for CAE.

High molar mass LPA networks have been shown to provide read lengths much greater than 600 bases in a single run, when several parameters, including sample preparation, capillary wall coating, matrix formulation (blends of low and high molar mass LPA), sequencing temperature, and base-calling algorithms, are completely optimized.[22,23,44] Note that, since the purpose of our study is a comparison of linear and sparsely cross-linked matrixes, we have made no special effort to obtain ultralong reads using CAE (using a commercial instrument and commercial base-calling software, standard DNA samples with no subsequent purification, and performing sequencing at 44 °C, the upper limit of the MegaBACE 1000, instead of at elevated temperature[44]).

(41) Cottet, H.; Gareil, P.; Viovy, J.-L. *Electrophoresis* **1998**, *19*, 2151–2162.
(42) Heller, C. *Electrophoresis* **1999**, *20*, 1962–1977.
(43) Viovy, J.-L.; Duke, T. *Electrophoresis* **1993**, *14*, 322–329.
(44) He, H.; Buchholz, B. A.; Kotler, L.; Miller, A. W.; Barron, A. E.; Karger, B. L. *Electrophoresis* **2002**, *23*, 1421–1428.

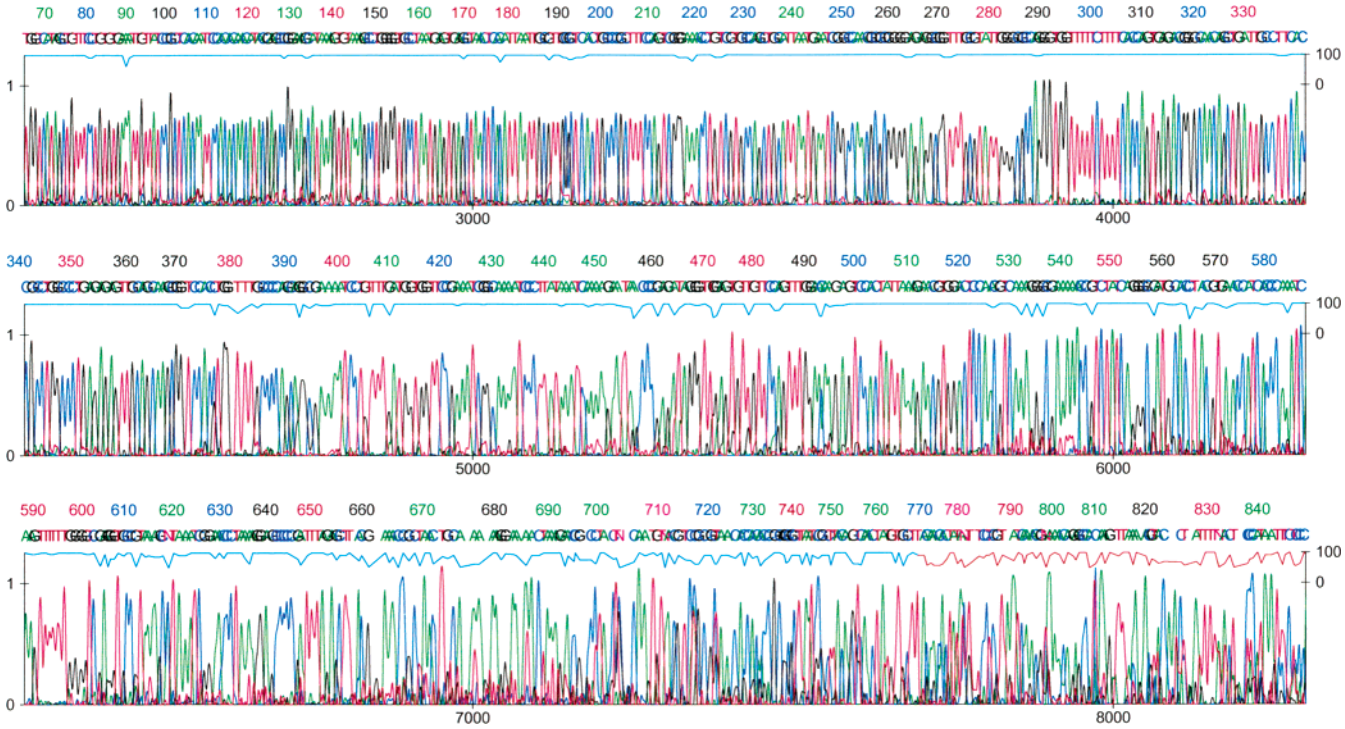

Figure 2. Electropherogram representing the average DNA sequencing performance of the Nanogel-5× matrix obtained on the MegaBACE 1000 sequencing system.

Improvements in read length can stem from an increased selectivity of the matrix (increased peak spacing), decreased sample peak widths, or a combination of both factors. Figure 3 shows plots of the matrix selectivity and sample peak width as a function of DNA base number for nanogel matrixes with three different cross-link densities. The Nanogel-5× matrix has a lower selectivity for small DNA (<250 bases), but commercial base-calling software is able to easily distinguish sample peaks at this selectivity. For large DNA (>600 bases), the results shown in Figure 3a indicate the selectivities of the three nanogel matrixes are similar. Instead, as seen in Figure 3b, the improvement provided by increasing the cross-link density (up to 5×) may be attributed to decreased peak widths provided by increasing the cross-link density. A direct comparison of each nanogel matrix with an appropriately matched LPA matrix is presented and was discussed in an earlier communication.[2]

Blends of high molar mass and low molar mass LPA have been employed by Karger and co-workers to improve DNA sequencing read lengths.[23,30] We expected the use of low and high molar mass linear polymer within a nanogel matrix containing sparse cross-links to follow trends that are similar to those seen in linear polymer blends. In particular, low molar mass polymer should improve the resolution of small DNA fragments by creating a range of smaller mesh sizes for the separation of small ssDNA, while high molar mass polymer should provide a more stable entangled network for the improved separation of large ssDNA fragments. The Nanogel-5× DNA sequencing matrix showed the most promising DNA sequencing performance. Using this matrix as a starting point, four blends, each primarily composed of nanogels and having a total polymer concentration of 2.75 (w/v)%, were formulated and tested (see Table 3).

Blend 1 was designed to improve the selectivity of the matrix for small ssDNA separation due to the presence of a distribution of smaller mesh sizes within the sequencing matrix and would be expected to have the lowest overall viscosity because of the inclusion of low molar mass LPA. Blend 2 should improve the selectivity of the matrix for large sequencing fragments, since with the inclusion of 0.25% high molar mass LPA ($M_w \sim 14 \times 10^6$ g/mol), the entangled network should be more robust than blend 1. Blend 3 should have some combination of the characteristics of blends 1 and 2. Blend 4, which includes 0.25% of an ultrahigh molar mass LPA ($M_w \sim 21 \times 10^6$ g/mol) along with 2.50% nanogels, should have the greatest network stability and should give good separation of large DNA fragments.

Table 3 shows the read lengths at 98.5% accuracy provided by the four blends. The addition of a low percentage of low molar mass polymer to the nanogels (blend 1) improved the matrix selectivity for short ssDNA as expected (see Figure 4a). The addition of a small fraction of high molar mass linear polymer to the nanogels (blend 2) decreased peak widths for large sequencing fragments. Blend 3, a mixture of nanogels, high molar mass polymer, and low molar mass polymer, gave intermediate performance as expected, showing improved resolution of most DNA sequencing fragments (up to ~550 bases; see Figure 4a), but the DNA peak widths in this matrix were greater than in the other two blends (Figure 4b). We hypothesize that the replacement of 0.50 (w/v)% of sparsely cross-linked polymer with linear polymer led to broader DNA peaks, by allowing increased DNA diffusion.

When comparing the sequencing performance of blends 1−4 to the performance of the Nanogel-5× matrix, we find that only blend 4 can provide DNA sequencing performance similar to that of the Nanogel-5× matrix; however, the performance of blend 4 was not a substantial improvement over the matrix composed solely of nanogels. This similarity in sequencing performance may indicate that high molar mass, sparsely cross-linked nanogels form networks that are essentially as stable as an ultrahigh molar mass

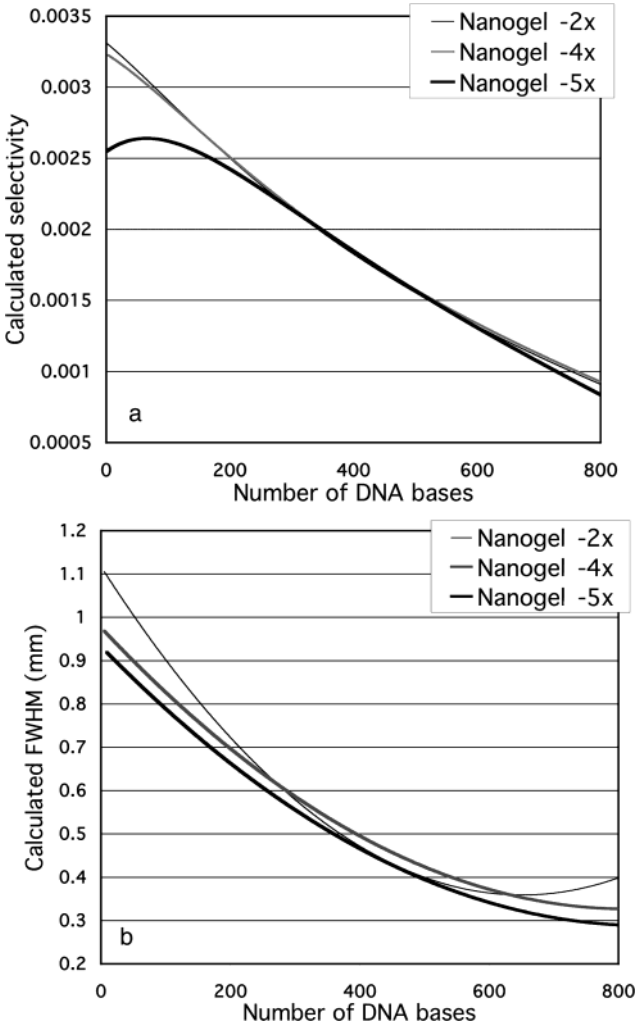

Figure 3. Selectivity (panel a) and peak width (panel b) as a function of DNA base number for Nanogel-2×, Nanogel-4×, and Nanogel-5× DNA sequencing matrixes. Each curve represents data obtained from a capillary showing the average sequencing performance for the specified matrix.

Table 3. DNA Sequencing Read Lengths at 98.5% Accuracy for LPA−Nanogel Blends (Total Polymer Concentration 2.75 (w/v)%)

blend	composition	average read length[a]	best read length
1	2.50% Nanogel-5× + 0.25% LPA−LMM	588 ± 74 (n = 13)	677
2	2.50% Nanogel-5× + 0.25% LPA−HMM	658 ± 28 (n = 14)	697
3	2.25% Nanogel-5× + 0.25% LPA-LMM + 0.25% LPA−HMM	635 ± 28 (n = 14)	683
4	2.50% Nanogel-5× + 0.25% LPA−UHMM	678 ± 20 (n = 10)	717

[a] Error indicates the standard deviation in the data.

(>20 × 10^6 g/mol) LPA network. Note that the average molar mass of the nanogels in this case is only ~10 × 10^6 g/mol; hence, the matrix viscosity is substantially lower than with 20 × 10^6 g/mol LPA.

Again, using the Nanogel-5× matrix as a starting point, two LPA−nanogel blends having a total polymer concentration of 3.00 (w/v)% were formulated and tested as DNA sequencing matrixes (see Table 4). These blends performed similarly to the Nanogel-5× matrix, with blend 5, a low molar mass LPA−nanogel blend, showing slight improvement with an average read length of 683 ± 8 bases (n = 14) at 98.5% accuracy. One interesting observation is that the blends having a higher total polymer concentration provide highly reproducible read lengths (RSD < 1.6%).

Highly cross-linked polymer networks, such as those developed in the early 1990s as in situ-polymerized capillary gels, were impractical for high-throughput sequencing due to the difficulty or impossibility of replacing the gel between sequencing runs.

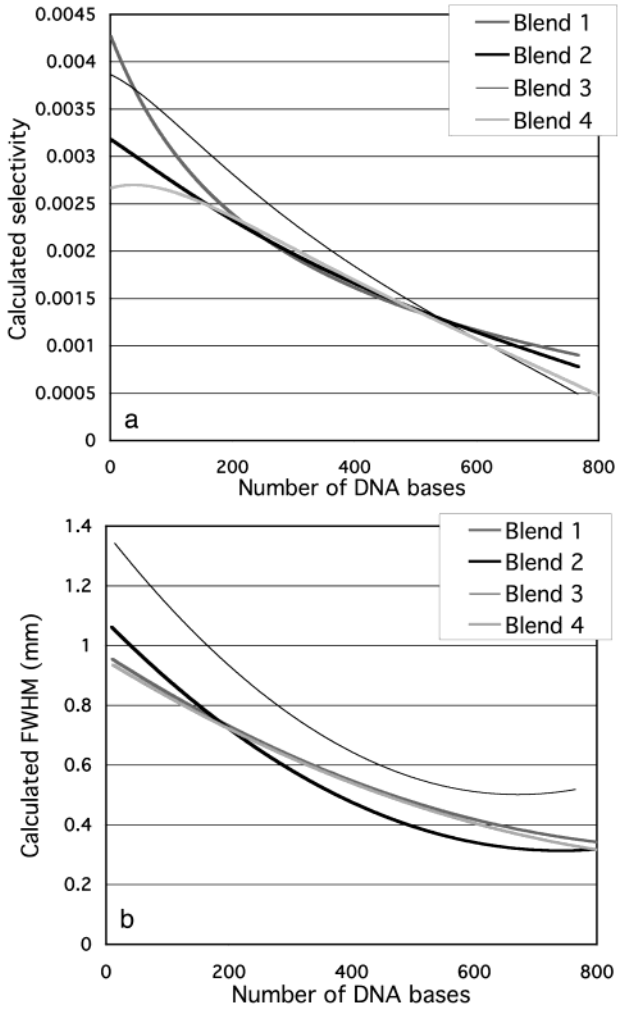

Figure 4. Selectivity (panel a) and peak width (panel b) as a function of DNA base number for LPA−nanogel blended DNA sequencing matrixes. Each curve represents the capillary showing the average sequencing performance for the specified matrix.

Table 4. DNA Sequencing Read Lengths at 98.5% Accuracy for LPA−Nanogel Blends (Total Polymer Concentration 3.00 (w/v)%)

blend	composition	average read length[a]	best read length
5	2.75% Nanogel-5× + 0.25% LPA−LMM	683 ± 8 (n = 14)	694
6	2.75% Nanogel-5× + 0.25% LPA−UHMM	674 ± 10 (n = 14)	691

[a] Error indicates the standard deviation in the data.

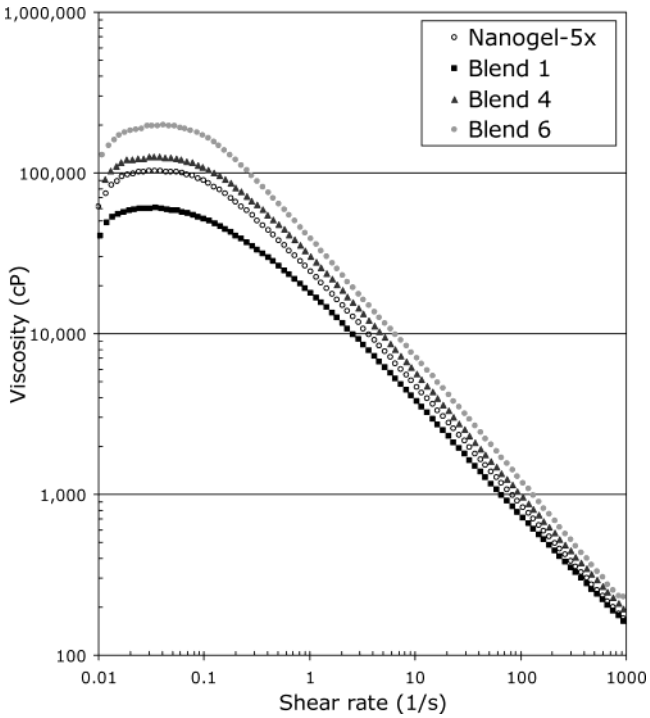

Figure 5. Flow behavior of nanogel and LPA−nanogel blended DNA sequencing matrixes. All matrixes show a dramatic drop in viscosity with increasing shear rate, indicating that they will flow into microchannels under applied shear. Data for blends 2, 3, and 5 were omitted for clarity.

Our novel use of sparse cross-linking in a highly entangled network of nanogels provides a shear-thinning polymer solution (see Figure 5), which allows for the replacement of the matrix between runs within a commercial CAE instrument as well as in a chip-based sequencing instrument with a custom-built matrix loading device.[37] As expected, substitution of a small proportion of the Nanogel-5× matrix with low molar mass LPA (blend 2) lowers the matrix viscosity; replacement of a small proportion of the Nanogel-5× matrix with very high molar mass LPA (blend 4) increases the viscosity slightly. Both blends 5 (not shown) and 6, with a total polymer concentration of 3.00 (w/v)%, are more viscous than the Nanogel-5× matrix.

On the basis of these results, we conclude that the use of sparse cross-linking for entangled polymer network stabilization shows significant utility for DNA sequencing for conventional and chip-based electrophoresis. The presence of sparse cross-links improves the selectivity of the matrix for large ssDNA sequencing fragments and, at the optimized cross-link concentration, further improves read lengths by reducing sample peak widths. This highly entangled network of acrylamide-based nanogels is the first material that outperforms highly entangled LPA as a sequencing matrix, when the linear and sparsely cross-linked polymers are matched for molar mass as well as the extent of interchain entanglements (c/c^*). Polymer network stabilization through occasional chemical cross-linking may aid in the stabilization of other polymers used for DNA sequencing. In particular, this approach to polymer network stabilization may improve the performance of thermothinning[29] or thermothickening[45] polymer networks, which may be loaded in a low-viscosity state and subsequently used for DNA sequencing as robust, highly entangled matrixes.

ACKNOWLEDGMENT

The authors gratefully acknowledge financial support and equipment for this work provided by the National Institutes of Health (Grant NHGRI R01HG1970-01) and by Amersham Biosciences (Sunnyvale, CA).

Received for review February 18, 2004. Accepted June 15, 2004.

AC049721X

(45) Kan, C. W.; Doherty, E. A. S.; Barron, A. E. *Electrophoresis* **2003**, *24*, 4161−4169.