Erin A. S. Doherty
Cheuk Wai Kan
Annelise E. Barron

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

Sparsely cross-linked "nanogels" for microchannel DNA sequencing

We have developed sparsely cross-linked "nanogels", sub-colloidal polymer structures composed of covalently linked, linear polyacrylamide chains, as novel DNA sequencing matrices for capillary electrophoresis. The presence of covalent cross-links affords nanogel matrices with enhanced network stability relative to standard, linear polyacrylamide (LPA), improving the separation of large DNA fragments. Nanogels were synthesized *via* inverse emulsion (water-in-oil) copolymerization of acrylamide and *N,N*-methylenebisacrylamide (Bis). In order to retain the fluidity necessary in a replaceable polymer matrix for capillary array electrophoresis (CAE), a low percentage of the Bis cross-linker ($< 10^{-4}$ mol%) was used. Nanogels were characterized by multi-angle laser light scattering and rheometry, and were tested for DNA sequencing by CAE with four-color laser-induced fluorescence (LIF) detection. The properties and performance of nanogel matrices were compared to those of a commercially available LPA network, which was matched for both weight-average molar mass (M_w) and extent of interchain entanglements (c/c^*). Nanogels presented in this work have an average radius of gyration of 226 nm and a weight-average molar mass of 8.8×10^6 g/mol. At concentrations above the overlap threshold, nanogels form a clear, viscous solution, similar to the LPA matrix ($M_w \sim 8.9 \times 10^6$ g/mol). The two matrices have similar flow and viscosity characteristics. However, because of the physical network stability provided by the internally cross-linked structure of the nanogels, a substantially longer read length ($\sim$ 63 bases, a 10.4% improvement) is obtained with the nanogel matrix at 98.5% accuracy of base-calling. The nanogel network provides higher-selectivity separation of ssDNA sequencing fragments longer than 375 bases. Moreover, nanogel matrices require 30% less polymer per unit volume than LPA. This is the first report of a sequencing matrix that provides better performance than LPA, in a side-by-side comparison of polymer matrices matched for M_w and extent of interchain entanglements.

Keywords: Capillary array electrophoresis / DNA sequencing / Linear polyacrylamide / Nanogels
DOI 10.1002/elps.200305702

1 Introduction

Capillary electrophoresis (CE) of DNA through polymeric separation matrices is presently the dominant technology for high-throughput sequencing. Although a final draft of the Human Genome was published recently [1, 2], hundreds of other genome projects [3], as well as individualized genomics, still require long DNA sequencing read lengths at low cost. Novel polymeric matrices that provide longer read lengths than commercially available sequencing matrices will be instrumental for the throughput increases and cost reductions that are still required to make the long-term goal of personalized genomics economically feasible. In the case of this work, we have developed a novel material for DNA sequencing with 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 [4, 5]. Cross-linked polyacrylamide yields excellent DNA separations, allowing for long reads under optimized conditions [6]. However, unless very low electric fields are used (which translates into long run times, typically 5–8 h), large DNA sequencing fragments ($>$ 300 bases) rapidly enter the "biased reptation with orientation" [7] migration mode in

Correspondence: Professor Annelise E. Barron, Northwestern University, Department of Chemical Engineering, 2145 Sheridan Road, Room E136, Evanston, IL 60208, USA
E-mail: a-barron@northwestern.edu
Fax: +847-491-3728

Abbreviations: APS, ammonium persulfate; c^*, polymer overlap threshold concentration; **CAE**, capillary array electrophoresis; **LCST**, lower critical solution temperature; **LPA**, linear polyacrylamide; **MALLS**, multiangle laser light scattering; R_g, *z*-average radius of gyration; **TAPS**, *N*-[tris(hydroxymethyl)-methyl]-3-aminopropanesulfonic acid; **TTE**, Tris-TAPS-EDTA buffer

these highly cross-linked media. 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) [8].

During the initial stages of the development of CE, researchers used *in situ*-polymerized, highly cross-linked polyacrylamide within the lumen of the capillary. Sequencing in cross-linked polyacrylamide capillary gels was shown by Karger and co-workers [9–11], Dovichi and co-workers [12], Smith and co-workers [13–15], and Baba *et al.* [16] for sequencing reads of up to 350 bases. Cross-linked polyacrylamide capillary gels were typically produced using a total monomer concentration of up to 5% and a concentration of Bis up to 5%; short-read sequencing separations typically required 60–70 min.

The direct transfer of a "slab gel" technology to the μm-sized channels typical of fused-silica capillaries was not effective for a variety of reasons. First, voids left within the capillary due to the increased density of the polymer relative to its monomer are detrimental to highly efficient separations [9, 17]. Second, an *in situ*-polymerized, highly cross-linked structure is difficult to remove from the capillary, making these prepared capillaries useful for 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 matrices.

The use of a replaceable DNA sequencing matrix, in particular a highly entangled solution of linear polyacrylamide (LPA), provided resolution of ssDNA fragments without the use of an infinitely cross-linked polymer network [9, 18]. A 6% solution of relatively low-molar-mass LPA ($\sim 1 \times 10^6$ g/mol) provided a read length of over 350 bases in close to 30 min, indicating that a highly cross-linked polymer network was not required for DNA sequencing within capillaries [18]. Also, compared to sequencing separations by CE using cross-linked gels, comparable sequencing reads could be achieved in a shorter time (*i.e.*, with a higher field), since a more open network shifts the "biased reptation with orientation" threshold to larger DNA sizes [7].

Importantly, the use of physically entangled, linear polymer solutions for the separation of DNA sequencing fragments within capillaries also allowed for relatively facile loading and replacement of the separation matrix between runs [18]. This enabled, for the first time, complete automation of DNA sequencing. Moreover, production and characterization of polymers *ex situ* has allowed researchers to correlate polymer physical and chemical properties, including weight-average molar mass [19], polydispersity [20, 21], and hydrophobicity [22–24], 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 [17]. An ideal polymer matrix for DNA sequencing should be hydrophilic, physically and chemically stable under sequencing conditions, and relatively low in viscosity (during loading and replacement). Typically, high-molar-mass polymers ($M_w > 2 \times 10^6$ g/mol) give the best performance because they form robust entangled networks [25].

A range of linear polymers have shown good utility for use in DNA sequencing, including LPA [26, 27], poly(*N,N*-dimethylacrylamide) (PDMA) [28, 29], poly(ethylene oxide) (PEO) [30], poly(vinylpyrrolidone) (PVP) [31], poly(*N*-hydroxyethylacrylamide) (polyDuramide™) [32], and copolymers of *N,N*-dimethylacrylamide (DMA) and *N,N*-diethylacrylamide (DEA) [22, 33]. To date, high-molar-mass LPA gives the best sequencing performance, able to produce a 1000-base read in about 1 h [21] and 1300 bases in 2 h [27] with highly optimized polymer molar mass distribution, matrix formulation, sample preparation and cleanup, and base-calling algorithms.

The long reads demonstrated by Karger and co-workers using high-molar-mass LPA were accomplished using blends of high- and low-molar-mass LPA. The 1000-base read was performed in a matrix blend composed of 2.0 wt% 9×10^6 g/mol and 0.5 wt% 5×10^4 g/mol LPA [21]; 1300 bases were sequenced in a matrix blend composed of 2 wt% 1.7×10^7 g/mol and 0.5 wt% 2.7×10^5 g/mol [27]. The inclusion of a low percentage of low-molar-mass polymer increases the total polymer concentration of the matrix, allowing smaller ssDNA fragments to be separated without significantly decreasing the selectivity for the large ssDNA fragments provided by the highly entangled, high-molar-mass polymer [34–36]. Significantly shorter read lengths are common in commercial CAE instruments such as the ABI PRISM 3700™ (550 bases in 3–4 h) and the MegaBACE 1000™ (600 bases in 2 h), due to the use of lower-viscosity, less entangled matrices for practical reasons and the lower quality of actual genomic DNA samples, among other factors.

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, and its facile production to high molar mass using standard free-radical polymerization chemistry. Poly-

acrylamide is also relatively easy to purify, as it readily precipitates from water with the addition of acetone or methanol. A highly entangled LPA matrix suitable for long-read sequencing has an extremely high zero-shear viscosity (60 000–120 000 cP); hence, high pressure (*i.e.*, 6895 kPa (1000 psi)) is required to initiate flow into the microchannel.

The use of branched copolymer structures for DNA sequencing matrices has been explored as a way to improve the performance of lower-molar-mass LPA by modifying its network properties. Viovy and co-workers [37] have produced a relatively low-molar-mass, branched copolymer with a polyacrylamide backbone. Poly(acrylamide-*g*-*N*-isopropylacrylamide), when heated above the lower critical solution temperature (LCST) of *N*-isopropylacrylamide (NIPA), forms micelle-like aggregates of NIPA grafts, stabilizing the branched polymer network with "transient cross-links" and increasing the matrix viscosity by nearly two orders of magnitude (from 100 to 10 000 cP). This polymer displayed excellent utility for the separation of dsDNA in published work, but was not tested as a DNA sequencing matrix. Although this class of branched polymers improves the loading properties of LPA (since lower-molar-mass, less viscous solutions can be used), they do not provide the highly entangled network presented by solutions of ultrahigh-molar-mass LPA ($M_w > 9 \times 10^6$ g/mol) previously demonstrated by Karger and co-workers to give long read lengths.

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 [38], 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. Occasional cross-links within a physically entangled matrix composed of high-molar-mass polyacrylamide should decrease sample diffusion as well as provide a more robust network for migrating DNA, allowing for significantly longer read lengths when compared to a commercially available linear polyacrylamide having a similar molar mass and extent of physical entanglements.

Two major challenges exist when attempting to incorporate cross-links into a polymer to be used as a sequencing matrix for CAE. First, the formation of an infinitely cross-linked polymer gel during the polymerization process must be avoided. In addition, large polymer structures of colloidal dimensions must be avoided, as these particles would scatter incident light, preventing sequencing using LIF detection. Finally, the cross-link density should be limited so that the final sequencing matrix retains good fluidity, and so that individual polymer structures may physically entangle with each other (see Fig. 1). To fulfill these requirements, we have produced sparsely cross-linked "nanogels" using inverse emulsion polymerization (polymerization within small water droplets stabilized by surfactant within an organic medium). Inverse emulsion polymerization of acrylamide and other water-soluble monomers is a common method for the production of highly cross-linked, monodisperse latex particles [39] as well as high-molar-mass linear polymer [40–42]. Advantages of inverse emulsion polymerization over solution polymerization include isolation of the domains of chemical reaction, better temperature control, relatively low polydispersity of the polymer product, and low-viscosity reaction products containing a high mass fraction of polymer (20–40 wt%). Karger and co-workers [19, 21, 27] have used this approach to produce high-molar-mass LPA for DNA sequencing.

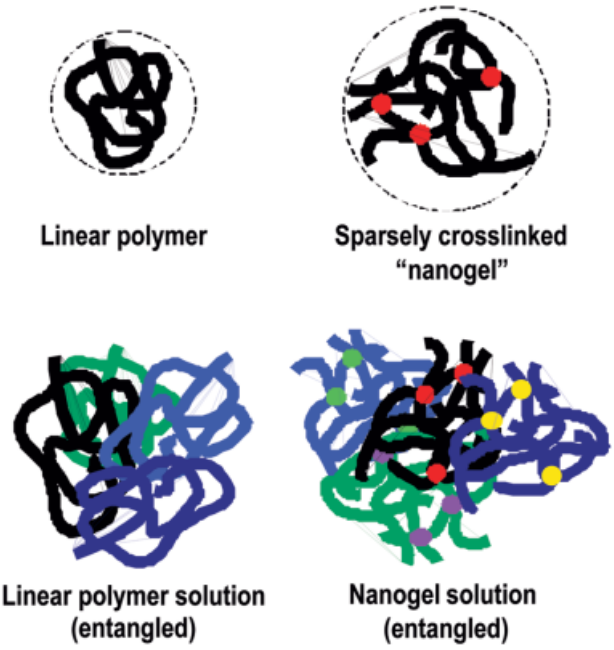


Figure 1. Schematic representation of nanogels (right) and linear polymer chains (left). Nanogels entangle with each other, forming an network similar to a solution of entangled linear polymer, but stabilized by the presence of chemical cross-links.

The sparsely cross-linked nanogels we have created occupy a new middle ground, between the highly cross-linked, *in situ*-produced polyacrylamide capillary gels and the fluid, linear polymer networks that are now utilized for CAE. Nanogels were characterized by batch multiangle laser light scattering (MALLS), steady-shear and oscillatory shear rheometry, and tested as DNA sequencing

matrices. We demonstrate that stabilization of a physically entangled sequencing matrix is possible by including an extremely small amount of tetrafunctional monomer, which provides chemical cross-link points that fortify the final entangled polymer network and improve sequencing results.

2 Materials and methods

2.1 Reagents

Tris, EDTA, ultrapure grade TEMED, ultrapure acrylamide, Bis, and ammonium persulfate (APS) were purchased from Amresco (Solon, OH, USA). *N*-[Tris(hydroxymethyl)-methyl]-3-aminopropanesulfonic acid (TAPS) was obtained from Sigma (St. Louis, MO, USA). Sorbitan monooleate (Span 80) was from Fluka Chemical (St. Louis, MO), Isopar M (a C_{12}-C_{14} isoparaffinic mixture) from Exxon (Houston, TX, USA). 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.

2.2 Polymer synthesis

Sparsely cross-linked nanogels were produced *via* inverse emulsion polymerization according to a protocol developed by Baade and Reichert [42] and later modified by Karger and co-workers [19]. Briefly, a spherical, 500 mL, water-jacketed reaction flask with a four-neck top (Kontes, Vineland, NJ, USA) was used with a homemade stainless steel shaft and a circular, four-blade, pitched-blade impeller centered and lowered to the lowest possible position within the vessel. The impeller was driven by an overhead, high-torque stirrer (Caframo Limited, Wiarton, Ontario). Reaction temperature was measured and recorded using a Jenco Electronics (Lazar Research Laboratory, Los Angeles, CA, USA) thermocouple and data logger. Based on a total emulsion mass of 400 g, the emulsion was formed as follows. 40.53 wt% Isopar M and 2.47 wt% Span 80 were mixed briefly before being poured into the reaction vessel. The organic phase was immediately mixed by the overhead stirrer set at 600 rpm. The organic phase was degassed for 30 min using prepurified nitrogen (Air Products, Naperville, IL, USA), further purified using an oxygen/water vapor trap (Supelco, St. Louis, MO, USA). A mixture of 22.8 wt% acrylamide and 34.2 wt% water was then added dropwise to the organic phase, resulting in a white, opaque emulsion. If sparsely cross-linked nanogels are desired, approximately 1 mL of an aqueous solution having the appropriate concentration of tetrafunctional monomer was added immediately following the addition of the monomer solution. To produce the nanogels described in this work, Bis composed less than 10^{-4} mol% of the total number of moles of monomer. The resulting emulsion was then degassed for 1 h, or until the emulsion temperature exceeded 35°C, whichever was longer. Prior to initiation, a small aliquot of the emulsion was added to approximately 30 mL of methanol, shaken, and centrifuged at low speed for 2–3 min to check for autopolymerization (indicated by pelleting of polymer powder). 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 numerous times with methanol during filtration; product was dried in a room-temperature vacuum oven for at least 72 h.

2.3 Isolation of the commercially available polymer

The commercially available LPA was recovered from prepackaged sequencing matrices (LongRead matrix, Beckman Coulter, Fullerton, CA, USA) by diluting the sequencing matrix in deionized, distilled water, then pouring the resulting diluted mixture into 1000 Da molecular weight cutoff (MWCO) cellulose ester membranes (Spectrum Laboratories, Rancho Dominguez, CA, USA). Diluted sequencing matrices were dialyzed against deionized, distilled water for 10 days with frequent water changes. The polymer solution was then frozen and lyophilized using a freeze-dry system (Labconco, Kansas City, MO, USA), resulting in a white, stiff, foam-like mass. Polymer was redissolved in a solvent or buffer of interest by slow rotation for at least 24 h (Roto-Torque, Cole-Parmer Instrument Company, Vernon Hills, IL, USA).

2.4 Rheological characterization

Rheometry was performed using an Anton Paar Physica (Glen Allen, VA, USA) MCR 300 equipped with a Julabo USA (Allentown, PA, USA) 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 CP50-1) over a range of 0.1–100 s^{-1}.

2.5 MALLS

Weight-average molar mass and radius of gyration of high-molar-mass polymer samples was determined by batch MALLS using a DAWN DSP Laser Photometer-Optilab DSP Interferometric Refractometer system (both, Wyatt Technology, Santa Barbara, CA, USA) (see [43] for more information on the utility of MALLS for analysis of high-molar-mass polyacrylamides). Data collection and analysis were performed as described in previous reports [43].

2.6 DNA sequencing and data analysis

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. Polymers to be tested as DNA sequencing matrices were dissolved at the concentration of interest in a $1 \times$ TTE (50 mM Tris, 50 mM TAPS, 2 mM EDTA) buffer. Sequencing matrices were loaded under a pressure of 6895 kPa (1000 psi) , 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, separation of DNA was performed at 140 V/cm and 44°C for 120 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™. Raw LIF data were extracted from the MegaBACE sequencing software and fitted into Gaussian peaks using PeakFit™ 4.06 (SPSS, Chicago, IL, USA) 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 high-molar-mass LPA matrix by Karger and co-workers [26]. 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 n-th 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 matrices [29, 44].

3 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. Assuming complete conversion of all monomer to a monodisperse polymer having a molar mass of 1.0×10^7 g/mol (see below) containing 8×10^{-5} mol% Bis, approximately 15% of the initiated polymers contain one or more Bis cross-links. Inverse emulsion polymerization limits initiation events by distributing radicals among the aqueous droplets [40, 41], hence, the production of low-molar-mass polymer should be minimal. The final matrix may be considered as a high-molar-mass linear polymer matrix with occasional cross-links, between and within chains.

3.1 Batch MALLS characterization

MALLS is an absolute method for the characterization of macromolecules that are too large to be fractionated by gel permeation chromatography (GPC); MALLS is particularly suited for the analysis of nanogels and high-molar-mass LPA [43]. In MALLS "batch mode", several dilutions of a stock polymer solution ($c \sim 0.05\ c^*$) are analyzed in series in order to extrapolate light scattering and refractive index data to infinitesimally low scattering angle and concentration. Batch MALLS allows for the determination of the weight-average molar mass (M_w) and the z-average radius of gyration (R_g) [43]. Unlike GPC and tandem GPC-MALLS, however, batch MALLS is unable to provide data on the molar mass distribution of a polymer sample. Zimm plots [45, 46] for the commercially available polymer (LPA) and the sparsely cross-linked nanogels are shown in Fig. 2.

In the Zimm plots, "horizontal" lines indicate light scattering of several polymer concentrations at a single detection angle; "vertical" lines indicate light scattering of a single polymer concentration at several different detection angles. By combining light-scattering data with refractive index data and extrapolating to zero concentration and to the 0° scattering angle, M_w and R_g may be determined. We find that the commercially available LPA has an M_w of $8.909 \pm 0.209 \times 10^6$ g/mol and an R_g of 167.8 ± 2.8 nm. The nanogels have an M_w of $8.495 \pm 0.836 \times 10^6$ g/mol and an R_g of 235.4 ± 10.2 nm. Although the two polymer samples have a similar weight-average molar mass, the sparsely cross-linked nanogels have a significantly larger z-average radius of gyration, which may seem, at first, counterintuitive. In general, the pres-

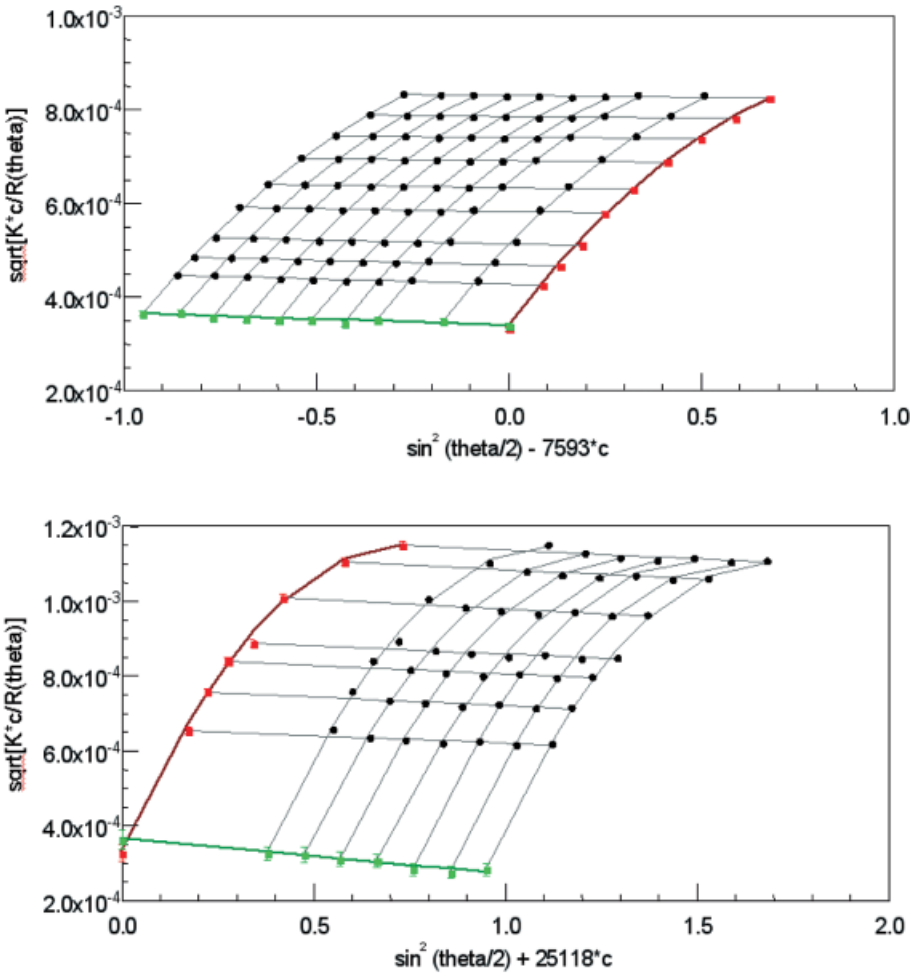


Figure 2. Zimm plots (Berry formalism, with a first-order angle fit and a second-order concentration fit) for the commercially available LPA (top) and sparsely cross-linked nanogels (bottom).

ence of chemical cross-links within a polymer having a similar molar mass should result in a more compact polymer structure. However, very little cross-linker is included in these nanogels. Moreover, batch MALLS does not provide information on molar mass distribution, which can strongly affect these average values. The larger average radius of gyration may indicate the presence of a larger fraction of very high-molar-mass structures for nanogels than for linear polymer.

3.2 Rheological characterization

3.2.1 Overlap concentration

The overlap concentration (c^*), or the concentration of polymer in solution at which polymer chains interact with each other in solution, is a critical measure of the extent of physical entanglements within a polymer solution, which in turn is critical to the DNA sequencing performance of a matrix. 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 matrices [34, 36, 47]. This then allows the performance of these matrices to be compared on a fair basis. The extent of polymer-polymer entanglements controls the lifetime of the virtual polymer "tube" that the DNA migrates through while under the influence of 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. Overlap concentrations for both LPA and nanogels (as experimentally determined by steady-shear rheometry [48]) are shown in Table 1. The LPA has a higher overlap concentration, consistent with its smaller average R_g as determined by light scattering.

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	8.909 ± 0.209	168 ± 3	1.66
Nanogels	7.86	8.495 ± 0.836	235 ± 10	1.17

3.2.2 Shear-thinning behavior of DNA sequencing matrices: implications for capillary loading

Loading of sequencing matrices composed of linear polymers into microchannels can be difficult, due to the high zero-shear viscosity of the matrix. Such difficulties can be mitigated by the shear-thinning properties of the matrix [19]. Matrix viscosity as a function of shear rate for the sequencing matrices employed in this work is shown in Fig. 3. As the shear rate applied to the highly entangled solution increases, polymer chains align with the flow, and the viscosity of the solution drops with a power-law dependence of viscosity on shear rate, allowing highly concentrated polymer solutions to flow into microchannels. The power-law exponents listed in the legend of Fig. 3 indicate the extent of shear thinning. As expected, sparsely cross-linked nanogels are slightly less able to shear-thin than linear polymers because the cross-linked structures are less able to align in flow. However, even with sparse cross-linking, the nanogel matrix still shows significant shear-thinning and flows well under the applied pressures available in commercial sequencing instruments for loading of capillary arrays (~1000 psi).

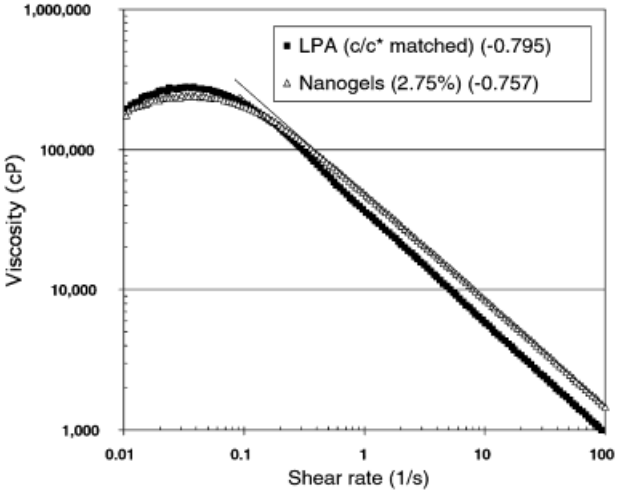


Figure 3. Flow curves of LPA and nanogel DNA sequencing matrices (both dissolved in 1 × TTE, 7 M urea) at 20°C. LPA matrices (■) shear-thin to a greater extent than nanogel matrices (△). However, both show strong shear-thinning behavior. Power-law exponents of the best-fit through the shear-thinning region are listed in the legend.

3.2.3 Shear-thinning behavior of DNA sequencing matrices: effect of cross-link density

Use of a very small molar percentage of the tetrafunctional monomer Bis as a comonomer for acrylamide produces the chemically cross-linked nanogel structure that stabilizes the final DNA sequencing matrix. To verify the incorporation of the cross-linker, as well as the relative percentage of Bis in the polymer nanogels, shear-rate-dependent viscosity data were collected for three nanogel batches having varying cross-link densities. Nanogels having similar overlap concentrations were used at the same concentration, in order to more accurately compare the power-law regions of the shear-dependent viscosity data. As expected, the magnitude of the power law exponent of the shear-thinning region of each flow curve decreases as the chemical cross-link density increases (data not shown). This result indicates that the amount of cross-linker incorporated into the nanogels is proportional to the amount that is included in the comonomer polymerization reaction mixture.

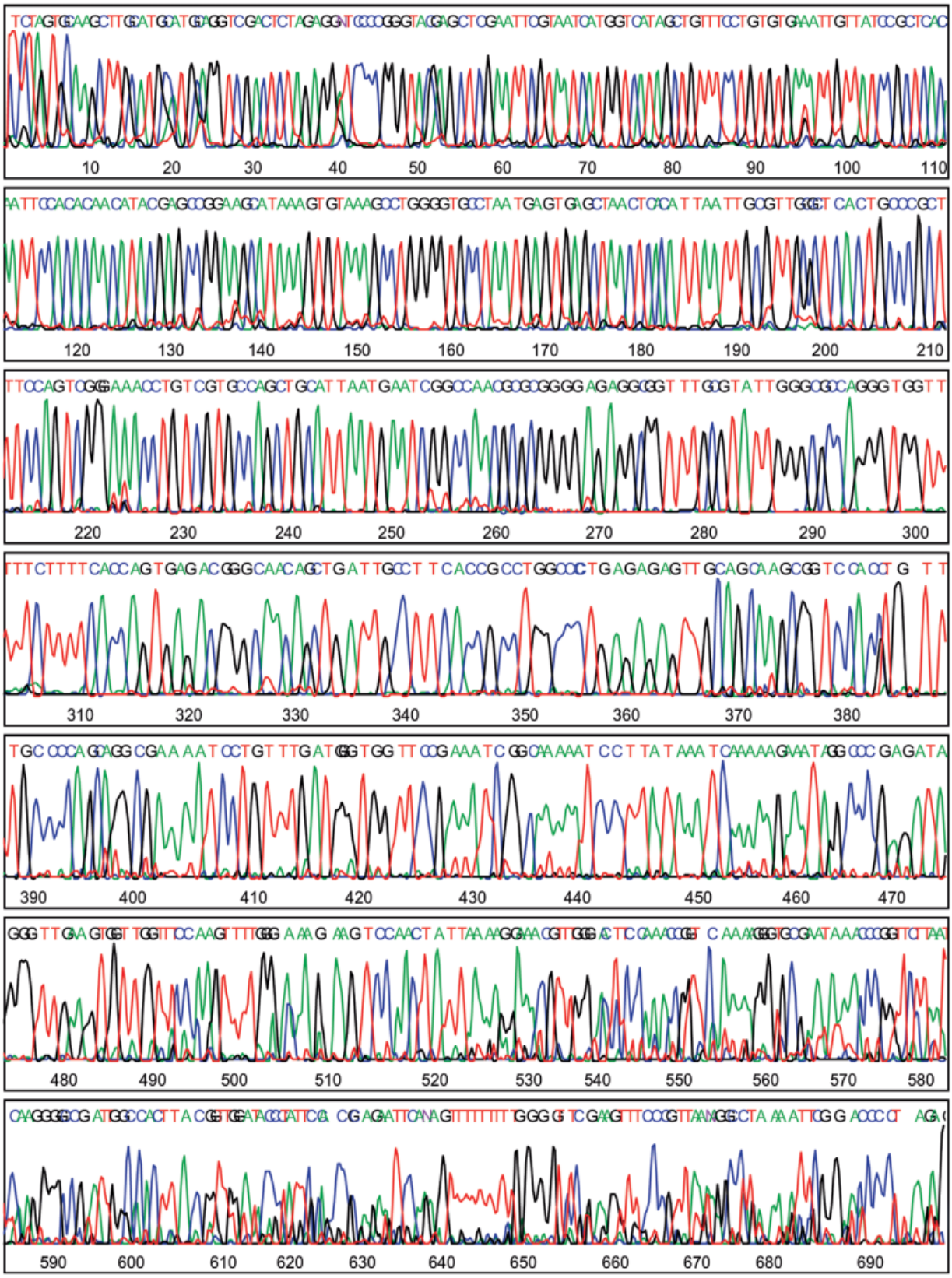
3.3 DNA sequencing

Table 2 lists the read lengths obtained at 98.5% base-calling accuracy for LPA and sparsely cross-linked polyacrylamide nanogel matrices. When the LPA concentration is c/c^* matched with the concentration of nanogels, the matrix composed of nanogels provides a 10.4% extension of read length at 98.5% base-calling accuracy and a significantly more reproducible performance (666 ± 10 bases for nanogels *vs.* 603 ± 26 bases for LPA; representative electropherograms are shown in Fig. 4). Note that, since the purpose of our study is a comparison of matrices, we have made no special effort to obtain ultralong reads (using commercial base-calling software, standard DNA samples with no subsequent purification, and performing sequencing at 44°C [49]). In addition, we have not employed blended matrices, which could serve to increase the resolution of short DNA fragments [26, 27]. We are unsure why the nanogel matrix provides a more reproducible read length; this matter is under investigation.

Table 2. DNA sequencing read lengths (120 min run)

Polymer sample	Conc. (w/v %)	98.5% Accuracy	
		Read length[b)]	% Improvement
LPA[a)]	3.87	603 ± 26	–
Nanogels[a)]	2.75	666 ± 10	10.4

a) Matched for c/c^*
b) Error indicates the standard deviation in the data ($n = 7$).



(a)

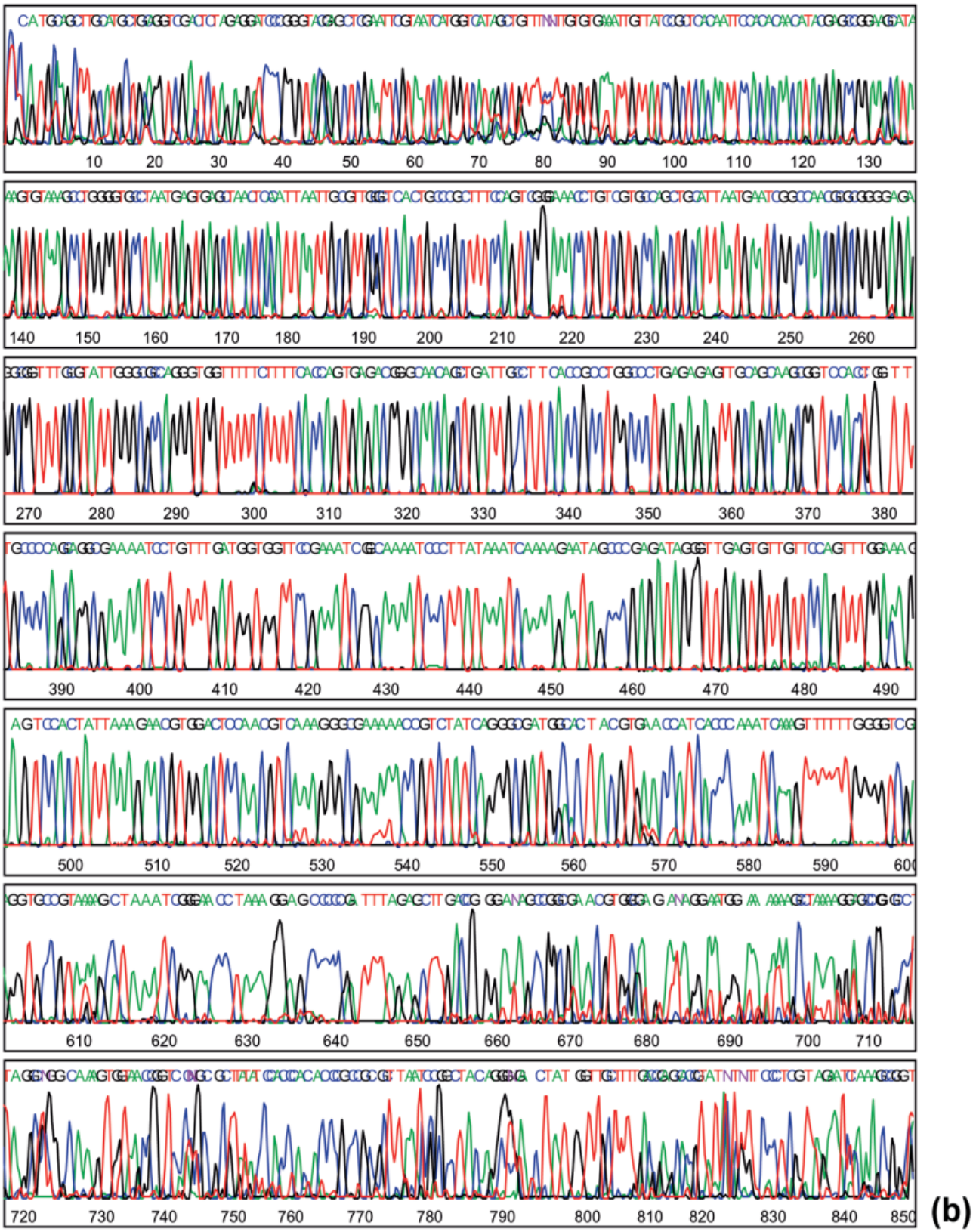


Figure 4. (a) Electropherogram representing the average performance of the LPA (*c*/*c** matched) matrix. (b) Electropherogram representing the average performance of the nanogel matrix. Electrophoresis conditions are listed in Section 2.6.

The electrophoresis time for all separations was 120 min. Hence, DNA migration is not appreciably slowed due to the presence of the sparse chemical cross-links.

DNA-polymer and polymer-polymer interactions dictate DNA separation mechanisms. For large ssDNA that migrates *via* reptation within the polymer matrix, network rupture, or dragging of polymer chains by migrating DNA, can destroy resolving power by temporarily changing the network structure of the entangled matrix [22]. Chemical cross-links make the network more stable to matrix disruption caused by DNA motion, possibly explaining the longer read lengths. In particular, the improved sequencing performance of the nanogel matrix could be attributed to improved selectivity, decreased peak widths, or a combination of both factors. In order to extend DNA sequencing read lengths, the selectivity, or peak spacing, of long ssDNA separation must be increased, or the peak widths of the long ssDNA fragments must be decreased. The selectivity and peak width as a function of base number provided by LPA and nanogel sequencing matrices are shown in Fig. 5.

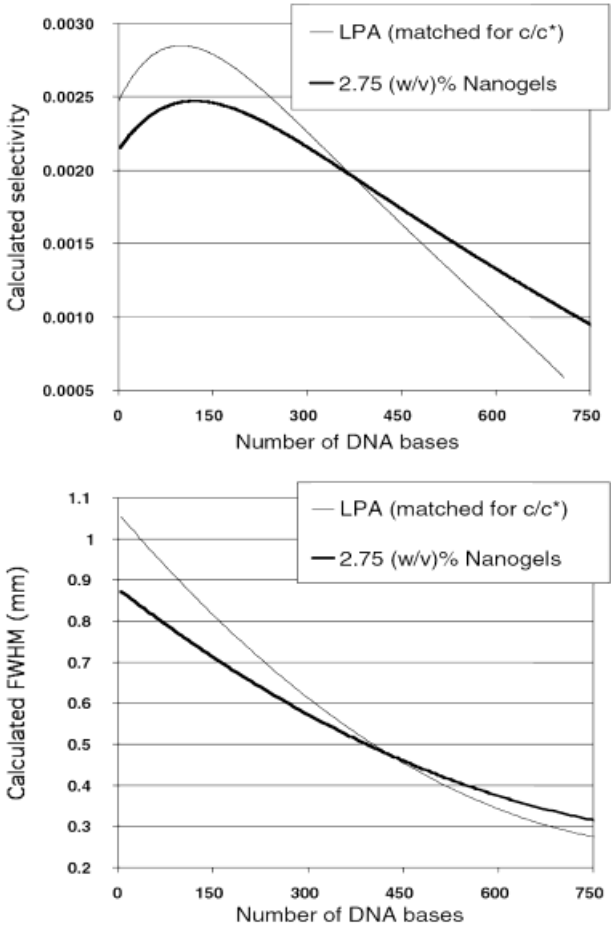


Figure 5. Selectivity (top) and peak width (bottom) as a function of DNA base number for LPA and nanogel DNA sequencing matrices.

We find that in general, DNA sequencing matrices composed of nanogels provide higher-selectivity separations than the LPA matrices for large sequencing fragments. Not surprisingly, lower selectivity is observed for ssDNA shorter than 375 bases, due to the lower overall concentration of the nanogels [21]. However, this does not compromise base-calling for shorter fragments (they are still well-separated). In particular, we find that the 2.75 w/v % nanogel matrix provides improved selectivity for ssDNA longer than 375 bases when compared to the LPA matrix that was matched for c/c^*.

The improvement in selectivity of large sequencing fragments provided by nanogel sequencing matrices is somewhat offset by ssDNA peak widths that are similar to or wider than those seen in LPA matrices for sequencing fragments longer than approximately 450 bases. This is unexpected, since chemical cross-links are thought to reduce diffusion within microchannel *in situ* polymerized cross-linked gels [38]. It is possible that the comparison of an entangled linear polymer network with an entangled network of polymer containing sparse cross-links solely on the basis of the extent of entanglements (c/c^*), which led to a moderate disparity in polymer concentration between the two matrices, may have a slight effect on the mesh size. This could explain the similar or wider ssDNA peak widths.

A higher nanogel cross-link density may be needed to impact ssDNA peak widths for large DNA. It is possible that an appropriate blend of linear polymer and nanogels will improve the DNA sequencing read length by further improvements in selectivity or sample peak width, both for small and large DNA. It should be noted that other nanogel batches produced under identical conditions also out-performed the matched LPA matrix; here we present our most promising results to date.

4 Concluding remarks

We have shown that we can improve the DNA sequencing performance of highly entangled polyacrylamide matrices through network stabilization with chemical cross-links in "nanogel" structures. Inverse emulsion polymerization may be used to produce discrete, sparsely cross-linked nanogels of ultrahigh molar mass (greater than 8×10^6 g/mol), which were characterized *via* batch MALLS and rheometry. Use of a tetra-functional acrylamide monomer to create occasional chemical cross-links results in significant increases in read length at 98.5% accuracy ($\sim$60 bases, a 10.4% longer read). The improvement in DNA sequencing read length is provided by increased selectivity for sequencing fragments longer than approximately 375 bases. This polymer network stabilization was

provided without negatively impacting matrix loading behavior or sequencing run time. This is the first report of a sequencing matrix that provides better performance than LPA, in a side-by-side comparison of polymer matrices matched for M_w and extent of interchain entanglements.

Nanogel matrices reported in this work have the potential to increase the read length (per microchannel per run) by 10% and decrease polymer consumption (per unit volume of matrix) by approximately 30%. Optimization of cross-link type and density, polymer molar mass, and matrix formulation may lead to further improvements in read length; other major improvements in loading properties may be realized through the use of LCST-exhibiting acrylamide-based monomers [33] in the production of thermoresponsive nanogels.

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 July 17, 2003

5 References

[1] Marshall, E., *Science* 2000, *288*, 2294–2295.
[2] Smaglik, P., *Nature* 2000, *404*, 111–111.
[3] Bernal, A., Ear, U., Kyrpides, N., *Nucleic Acids Res.* 2001, *29*, 126–127.
[4] Chrambach, A., Rodbard, D., *Science* 1971, *172*, 440–450.
[5] Bishop, D. H., Claybrook Jr., L., Spiegel, M. S., *J. Mol. Biol.* 1967, *26*, 373–378.
[6] 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.
[7] Slater, G. W., Noolandi, J., *Biopolymers* 1986, *25*, 431–454.
[8] Cheng, J., *Prog. Biochem. Biophys.* 1995, *22*, 223–227.
[9] Heiger, D. N., Cohen, A. S., Karger, B. L., *J. Chromatogr.* 1990, 33–48.
[10] Cohen, A. S., Najarian, D. R., Karger, B. L., *J. Chromatogr.* 1990, *516*, 49.
[11] Guttman, A., Cohen, A. S., Heiger, D. N., Karger, B. L., *Anal. Chem.* 1990, *62*, 137–141.
[12] Harke, H. R., Bay, S., Zhang, J. Z., Rocheleau, M. J., Dovichi, N. J., *J. Chromatogr.* 1992, *608*, 143–150.
[13] Smith, L. M., *Nature* 1991, *349*, 812–813.
[14] Drossman, H., Luckey, J. A., Kostichka, A. J., Dcunha, J., Smith, L. M., *Anal. Chem.* 1990, *62*, 900–903.
[15] 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.
[16] Baba, Y., Matsuura, T., Wakamoto, K., Morita, Y., Nishitsu, Y., Tsuhako, M., *Anal. Chem.* 1992, *64*, 1221–1225.
[17] Bae, Y. C., Soane, D., *J. Chromatogr.* 1993, *652*, 17–22.
[18] Ruiz-Martinez, M. C., Berka, J., Belenkii, A., Foret, F., Miller, A. W., Karger, B. L., *Anal. Chem.* 1993, *65*, 2851–2858.
[19] Goetzinger, W., Kotler, L., Carrilho, E., Ruiz-Martinez, M. C., Salas-Solano, O., Karger, B. L., *Electrophoresis* 1998, *19*, 242–248.
[20] Barron, A. E., Sunada, W. M., Blanch, H. W., *Electrophoresis* 1996, *17*, 744–757.
[21] Salas-Solano, O., Carrilho, E., Kotler, L., Miller, A. W., Goetzinger, W., Sosic, Z., Karger, B. L., *Anal. Chem.* 1998, *70*, 3996–4003.
[22] Albarghouthi, M. N., Buchholz, B. A., Doherty, E. A. S., Bogdan, F. M., Zhou, H., Barron, A. E., *Electrophoresis* 2001, *22*, 737–747.
[23] Chiari, M., Micheletti, C., Nesi, M., Fazio, M., Righetti, P. G., *Electrophoresis* 1994, *15*, 177–186.
[24] Gelfi, C., Simo-Alfonso, E., Sebastiano, R., Citterio, A., Righetti, P. G., *Electrophoresis* 1996, *17*, 738–743.
[25] Albarghouthi, M., Barron, A. E., *Electrophoresis* 2000, *21*, 4096–4111.
[26] 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.
[27] Zhou, H., Miller, A. W., Sosic, Z., Buchholz, B., Barron, A. E., Kotler, L., Karger, B. L., *Anal. Chem.* 2000, *72*, 1045–1052.
[28] Madabhushi, R. S., Vainer, M., Dolník, V., Enad, S., Barker, D. L., Harris, D. W., Mansfield, E. S., *Electrophoresis* 1997, *18*, 104–111.
[29] Madabhushi, R. S., *Electrophoresis* 1998, *19*, 224–230.
[30] Fung, E. N., Yeung, E. S., *Anal. Chem.* 1995, *67*, 1913–1919.
[31] Gao, Q. F., Yeung, E. S., *Anal. Chem.* 1998, *70*, 1382–1388.
[32] Albarghouthi, M. N., Buchholz, B. A., Huiberts, P. J., Stein, T. M., Barron, A. E., *Electrophoresis* 2002, *23*, 1429–1440.
[33] 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.
[34] Heller, C., *Electrophoresis* 1999, *20*, 1962–1977.
[35] Duke, T., Viovy, J.-L., Semenov, A. N., *Biopolymers* 1994, *34*, 239–247.
[36] Cottet, H., Gareil, P., Viovy, J.-L., *Electrophoresis* 1998, *19*, 2151–2162.
[37] Sudor, J., Barbier, V., Thirot, S., Godfrin, D., Hourdet, D., Millequant, M., Blanchard, J., Viovy, J.-L., *Electrophoresis* 2001, *22*, 720–728.
[38] Ugaz, V. M., Burke, D. T., Burns, M. A., *Electrophoresis* 2002, *23*, 2777–2787.
[39] Arshady, R., *Colloid Polym. Sci.* 1992, *270*, 717–732.
[40] Hunkeler, D., *Polym. Int.* 1992, *27*, 23–33.
[41] Hunkeler, D. J., *Macromolecules* 1991, *24*, 2160.
[42] Baade, W., Reichert, K. H., *Eur. Polym. J.* 1984, *20*, 505–512.
[43] Buchholz, B. A., Barron, A. E., *Electrophoresis* 2001, *22*, 4118–4128.
[44] Menchen, S., Johnson, B., Winnik, M. A., Xu, B., *Electrophoresis* 1996, *17*, 1451–1459.
[45] Wyatt, P. J., *Anal. Chim. Acta* 1993, *272*, 1–40.
[46] Zimm, B. H., *J. Chem. Phys.* 1945, *13*, 141–145.
[47] Viovy, J.-L., Duke, T., *Electrophoresis* 1993, *14*, 322–329.
[48] Barron, A. E., Soane, D. S., Blanch, H. W., *J. Chromatogr.* 1993, *652*, 3–16.
[49] He, H., Buchholz, B. A., Kotler, L., Miller, A. W., Barron, A. E., Karger, B. L., *Electrophoresis* 2002, *23*, 1421–1428.