Review

Methal N. Albarghouthi
Annelise E. Barron

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

Polymeric matrices for DNA sequencing by capillary electrophoresis

We review the wide range of polymeric materials that have been employed for DNA sequencing separations by capillary electrophoresis. Intensive research in the area has converged in showing that highly entangled solutions of hydrophilic, high molar mass polymers are required to achieve high DNA separation efficiency and long read length, system attributes that are particularly important for genomic sequencing. The extent of DNA-polymer interactions, as well as the robustness of the entangled polymer network, greatly influence the performance of a given polymer matrix for DNA separation. Further fundamental research in the field of polymer physics and chemistry is needed to elucidate the specific mechanisms by which DNA is separated in dynamic, uncross-linked polymer networks.

Keywords: DNA sequencing / Capillary electrophoresis / Polymer solutions / Matrices / Review
EL 4173

Contents

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

Abbreviations: AAP, *N*-acryloylaminopropanol; **HEC**, hydroxyethylcellulose; **LOR**, length of read; **LPA**, linear polyacrylamide; **PDMA**, poly-*N,N*-dimethylacrylamide; **PEO**, polyethylene oxide; **PNIPA**, poly-*N*-isopropylacrylamide; **TAPS**, *N*-tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid; **TBE**, Tris-borate-EDTA buffer; **TTE**, Tris-TAPS-EDTA buffer

1 Introduction

Genome sequencing projects reveal the genetic makeup of an organism by reading off the sequence of the DNA bases that encode the fundamental information necessary for the life of the organism. Already revolutionizing biology, genome research provides a vital thrust to the increasing productivity and pervasiveness of the life sciences. Availability of the genome sequences of humans, animals, plants and microorganisms will have unprecedented impact on many disciplines, including health care, biological and biomedical research, forensics, biotechnology, agriculture, and the environment. To date, there are over 50 complete genomes available in public databases, and over 350 ongoing genome projects [1]. All genome projects and their applications will continue to seek new advances in high-throughput and cost-effective sequencing technologies. Therefore, the development of automated, high-throughput, cost-effective DNA sequencing technologies has gained huge momentum in both academic and industrial communities.

The focus of this review is upon polymeric materials that have been tested for DNA sequencing by capillary electrophoresis (CE). Hypothesized mechanisms of DNA separation in polymer solutions are briefly reviewed, along with a discussion of some experimental results that provide better insight into these mechanisms. The empirical approaches that have been taken towards the selection and performance optimization of polymers for DNA sequencing are discussed so as to provide guidelines for the selection of appropriate materials and conditions for resolving DNA sequencing fragments by CE.

0173-0835/00/1818-4096 $17.50+.50/0

2 DNA sequencing

The present state-of-the-art method for determining the base sequence of DNA relies upon the Sanger dideoxy-terminated reaction [2]. The method is based on controlled interruption of the enzymatic replication of ssDNA by DNA polymerase enzyme. The modern Sanger cycle sequencing reaction produces a fluorescently labeled, nested, single-base ladder of ssDNA fragments ranging in chain length from just a few bases to a few thousand bases. This set of DNA fragments is separated according to chain length using gel electrophoresis, allowing the order of bases to be read. Originally, the sequence was determined using autoradiography after electrophoresis was complete, and required up to 24 h. This type of "snapshot" detection severely limited the capacity and throughput of sequencing systems. High-throughput DNA sequencing became a reality with the commercial introduction of an apparatus for automated "finish line" detection of DNA fragments labeled with laser-induced fluorescence (LIF) dyes.

The majority of DNA sequencing has been carried out by slab-gel electrophoresis. Commercial sequencing instruments based on ultrathin cross-linked polyacrylamide slab-gel electrophoresis, such as the ABI PRISM 377™ (PE Biosystems, Foster City, CA, USA), produce about 600 bases per sample at a top rate of 200 bases/h/lane, and can only provide over 750 bases at the expense of extending run times to 10–12 h [3]. Modern sequencing gels have the advantage of running up to 96 samples in parallel, which increases the throughput of the instrument. However, slab-gel electrophoresis suffers from many limitations. Gel pouring is a manual process, and there is always a possibility of trapping air bubbles in the gel that can harm the separation efficiency. Furthermore, the Joule heating effect, which arises from the passage of electric current through the gel, places upper limits on the voltage that can be applied, resulting in long run times. Finally, with slab-gel systems, sample loading is done manually, obstructing complete automation of the sequencing process.

CE is an attractive alternative to slab-gel electrophoresis and is rapidly becoming the dominant technique in DNA sequencing centers [4, 5] due to its capacity for full automation. In CE, DNA separation is achieved in a fused-silica capillary, 25–100 μm in diameter. The high ratio of surface area to volume of the small capillary tube serves to efficiently dissipate the heat produced during electrophoresis, allowing the use of higher electric fields, which decreases the run time and improves DNA resolution. Although CE provides much faster separation than slab gels, it is necessary to operate multiple capillaries in parallel to compete with the throughput of slab-gel electrophoresis systems, which can run up to 96 samples in parallel. After the first study exploiting the potential of a multiple-capillary instrument was reported by Mathies *et al.* [6, 7], several investigators endeavored to produce a capillary array system that could significantly improve the sequencing rate relative to slab-gel-based instruments (see [8] for review). Today, there are two commercial versions of capillary array electrophoresis instruments, each featuring 96 capillaries and relying on LIF detection: the MegaBACE 1000™ from Molecular Dynamics (Sunnyvale, CA, USA) and the ABI PRISM 3700™ from PE Biosystems.

The MegaBACE 1000™ instrument echoes the original Mathies design [6, 7] that was based on confocal LIF detection consisting of a microscope objective to focus the laser light inside the capillaries and, at the same time, to collect emitted light from the center of the columns. To collect data from each and every capillary at the required frequency, a scanning system is used. The instrument uses polyacrylamide-coated capillaries with a covalent coating attachment and a high molar mass, linear polyacrylamide-based LongRead® matrix. Over 500 bases per capillary can be sequenced in 2 h, yielding a minimum throughput of 250 bases/h/capillary [9].

The ABI PRISM 3700™ employs postcolumn LIF detection with liquid sheath flow [10, 11]. In this detection system, the anodic end of the capillary bundle is aligned inside a quartz cuvette. Along the dead space between the columns and the walls of the cuvette, a polymer solution is pumped through the cell. The liquid sheath flowing outside the capillary entrains the DNA bands as they elute from the capillary, tapering them to a smaller diameter. A laser beam crosses all of the flow streams and excites the fluorescence of the dyes with which the DNA strands are labeled. Fluorescent light is dispersed through a diffraction grating and imaged onto a charge-coupled device (CCD) camera. The 3700™ instrument uses a separation matrix that is based on relatively low molar mass, linear polydimethylacrylamide, commercially known as POP®, for "performance-optimized polymer". The zero-shear viscosity of this matrix is relatively low compared to the aforementioned LongRead® matrix. The system can achieve a read length of 550 bases per capillary with a run time of approximately 4 h [12]. The same polymer is used for "dynamic coating" of the internal capillary wall by physical adsorption. Hence, arrays of bare fused-silica capillaries can be used in this instrument, reducing array cost. Typically, more than 100 consecutive runs can be carried out in such a "dynamically coated" capillary array before performance begins to degrade, after which time the capillaries can be "regenerated" with an acid rinse.

The next generation of automated, high-throughput DNA sequencing instruments most likely will be based on microfluidic devices or microchips (here, we are referring not to DNA hybridization chips, but to electrophoresis microchips). DNA sequencing on a microchip was first demonstrated in 1995 by Mathies and Woolley [13]. In the first unoptimized system of DNA sequencing on a microfluidic device, single-base resolution reached 150–200 bases in 10–15 min in an *in situ*-polymerized linear polyacrylamide (LPA) matrix, with an effective channel length of 3.5 cm. Later, optimized separation conditions that utilized a 3% LPA matrix, polymerized outside the chip, yielded read lengths of over 500 bases in about 20 min in a single channel run at 99.4% sequencing accuracy [14]. The shorter run time (20 min as compared to 2–4 h for capillaries) is a consequence of the narrow injection zones achievable on chips that have cross-flow or "T" injection geometries. Hence, even a 3.5 cm channel can deliver a 500-base read length, a surprising result for many people, providing a potential sequencing rate of 1500 bases/h/lane if the system could be automated. This work demonstrated the feasibility of high-speed DNA sequencing by CE on microchips and showed the tremendous potential for increased throughput that these systems offer. However, this breadboard system did not enable automated replacement of the separation matrix; automating that particular function may be more difficult to accomplish on microfluidic devices than it is in capillary arrays.

3 Mechanism of DNA electrophoresis in polymer solutions

An understanding of the mechanism of size-based separation of DNA chains in polymeric media is critical for improving and optimizing the performance of sieving matrices and electrophoresis conditions. By some, the mechanism of DNA separation in uncross-linked polymer solutions has been described as being essentially the same as that in traditional slab-gel electrophoresis, as modeled by the Ogston model and the reptation models, recently reviewed by Heller [15]. In the Ogston model, a DNA molecule is assumed to migrate as a rigid spherical particle through a network of rigid, linear, infinitely long rods, and to diffuse laterally until it encounters a "pore" large enough to allow its passage. Thus smaller molecules migrate faster through the matrix as they have access to a larger fraction of the pores. The Ogston model predicts a linear dependence of log (μ), where μ is the electrophoretic mobility, on matrix concentration for small DNA fragments in the limit of low electric fields [16, 17]. According to this model, however, molecules with radii of gyration larger than the average pore size should not penetrate the matrix at all, in contrast with the actual behavior of DNA molecules in gel electrophoresis. Hence, the assumptions underlying the Ogston model are invalid for larger DNA [18]. Furthermore, fundamental assumptions of this model about the nature of the separation matrix are extremely far from the reality of a dynamic entangled polymer network.

In the reptation model, the DNA molecule travels through a matrix as a flexible chain that reptates "snake-like" through a fictitious "tube" in the polymer network. The model predicts, in the limit of zero electric field, that the electrophoretic mobility is inversely proportional to DNA size. This dependence is indeed observed under some conditions for slab gels, over a certain "medial" range of DNA size. However, experimental observations that DNA larger than 20 kbp migrate with a constant electrophoretic mobility led to a modification to the reptation model, the biased reptation model (BRM) [19, 20]. This model postulates that when the electric field strength is large, DNA chains will orient in the direction of the electric field, strongly biasing the random walk in the forward direction so that DNA assumes a stretched, rod-like conformation. DNA molecules in this oriented conformation no longer take a tortuous path through the gel, thus the size-based separation ability of the gel is lost for large DNA and at high field strengths. The biased reptation with fluctuations (BRF) model was proposed to improve the original BRM by incorporating fluctuations in the theoretical tube length [21].

Polymer solutions differ from cross-linked gels in that the physical entanglements between the chains have a finite lifetime. The group of Viovy and Duke [17] adopted the physical concept of "constraint release" to describe this phenomenon through a modification of the BRF model. Such constraint release occurs when the polymers, which form the tube, are allowed to reptate away even while the DNA remains within the tube. Cottet *et al.* [22] described the dynamics of the entangled polymer network by the reptation or relaxation time which, for a given polymer, increases with increasing polymer concentration or molar mass but decreases with increasing temperature. In order to achieve uniform mobility and good resolution of electrophoresing DNA, the relaxation time should be greater than the residence time of the DNA molecule segment in the blob [22, 23].

The Ogston model and reptation models are applicable at low electric field strengths such as those typically used in slab-gel electrophoresis, but would seem to be inapplicable for CE in uncross-linked polymer solutions, where much higher field strengths are commonly applied. This points to a fundamental inconsistency of the application

of the reptation model to DNA separation by CE. Polymer motion by "reptation" is postulated to prevail in a fixed polymer network, in which chains are sufficiently constrained, each in a reptation "tube", that end-on motion is their only possible mode of movement besides small, lateral segmental motions (the formation of "chain defects") [24]. This type of motion has been seen by microscopy to occur in entangled DNA solutions under equilibrium conditions of zero applied field [25].

Mathematically, the initial distribution of chain conformations for computer simulation of reptative polymer motion is assumed to be either Gaussian or self-avoiding, depending upon solvent conditions, and hence is derived by equilibrium-based, statistical mechanical arguments [26, 27]. These assumptions are only appropriate in the case of zero external applied force, and hence are violated in the presence of a strong electric field [28, 29]. Under an external field, the initial distribution of chain conformations (for the DNA) must be heavily influenced by kinetics, and hence cannot be determined accurately from statistical mechanical arguments. Hence, the reptation model *per se* cannot be applied meaningfully to high-field electrophoresis. Although it has had significant success for interpreting slab-gel electrophoresis, it may not be a useful framework with which to consider the mechanism of separation for CE in uncross-linked polymer solutions. While these mechanisms may be applicable at extremes of low electric field and high polymer concentrations, these are generally not conditions of interest for genomic sequencing.

The difference in stiffness of ssDNA (4 nm persistence length) and dsDNA (45 nm persistence length) molecules, as well as the availability of bases in ssDNA molecules to hydrogen bond with the surrounding water molecules, urea (when used as denaturant) and polymer chains have been shown to affect the electrophoretic migration of DNA and have led to the identification of new separation regimes [15, 30]. Furthermore, the separation mechanisms postulated by both the Ogston and reptation models require polymer concentration to be in the entangled regime for DNA separation to be achieved. However, Barron *et al.* [31–33] have shown that separation of DNA fragments is possible in dilute polymer solutions at concentrations well below the entanglement threshold. To explain this experimental observation, the authors proposed a transient entanglement coupling mechanism for DNA separation. It was hypothesized that DNA fragments migrating through the polymer solution entangle or collide with individual polymer molecules and are forced to drag them through the solution, resulting in a decrease in DNA electrophoretic mobility. Larger DNA molecules were postulated to have a higher probability of encountering and entangling polymer molecules, hence to experience a greater reduction in mobility. Variations of this mechanism are likely to apply to both dilute and entangled polymer solutions, as the CE separations of DNA fragments obtained at concentrations above and below the polymer entanglement threshold were similar [32, 34].

Direct observations of dsDNA molecules electrophoresing through polymer solutions [35–38] shed new light on the mechanism of separation and the interactions between DNA and polymer molecules during electrophoresis. DNA molecules electrophoresing through high molecular mass polymer solutions, both above and below the entanglement threshold, undergo conformational changes from a compact, globular conformation to U-shaped conformations that appear to drag the entangling polymer molecules. However, in low molecular mass polymer solutions DNA migrates primarily in a globular conformation, not forming entanglements. Three types of DNA-polymer interactions have been proposed to account for this behavior [37, 38]: (i) U-shape collisions, in which the DNA entangles with the polymer obstacle and releases itself in a pulley-like motion; (ii) brief collisions, in which the DNA entangles with the polymer for a brief period of time before the polymer releases itself from the DNA; and (iii) transient nonentangling collisions, where the DNA in globular form pushes polymer molecules out of its path. Nonentangling collisions, modeled by Sunada and Blanch [38], occur between small DNA and small polymer molecules and explain the separation of small DNA molecules, whilst the entangling collisions, modeled by Hubert *et al.* [39], are necessary to separate large DNA molecules by entangling with large polymer molecules. Based on these mechanisms of separation, one can explain the experimentally observed improvements of separation of a wider size range of DNA molecules in solutions containing high and low molecular mass polymer, both above and below the entanglement threshold [40–43]. DNA was not observed to reptate, but to cycle in conformation between a spherical glob and a deformed U-shape.

It is clear that the mechanism of DNA electrophoretic separation in uncross-linked polymer solutions is more complicated than that in cross-linked gels, due to the dynamic nature of the network and due to the fact that separation is possible even in the absence of the polymer network in dilute solutions. At present, theoretical work on the mechanism of DNA separation is most useful for providing a physical framework with which to qualitatively interpret experimental results. None of these models are quantitatively predictive at this time. However, considering the entangling and nonentangling collisions which have been seen to take place over a wide range of polymer concentrations, it should be possible for theorists to formulate a

single model of DNA electrophoresis in polymer solutions, taking into account the properties of DNA, both in single- and double-stranded form, and polymer molecules under the various electrophoretic conditions. This is an exquisitely complicated problem in polymer physics and will, no doubt, keep theorists occupied for a long while.

4 Quantitative analysis of DNA separation

A quantitative measure of electrophoretic separation performance is the resolution, R_s, which can be expressed as:

$$R_s = 0.59\frac{x_2 - x_1}{FWHM} = \frac{1}{4}\frac{\Delta\mu}{\mu_{av}}N_P^{1/2} \quad (1)$$

where x_i is the center of peak *i*, *FWHM* is the peak full width at half maximum (this is derived based on the assumption that both peaks have the same width), $\Delta\mu$ is the difference in electrophoretic mobility for two fragments, μ_{av} is the average mobility between two fragments, and N_P is the number of theoretical plates. The term $\Delta\mu/\mu_{av}$ is a measure of the selectivity or peak spacing in a separation and reflects the sieving power of the matrix. The term N_P is a measure of separation efficiency or peak width due to the dispersion characteristics of the electrophoresis system.

An alternative quantitative measure of the performance of a sequencing separation matrix and instrument is the length-of-read, LOR. The LOR is defined as the point at which the peak spacing is equal to the peak width giving a resolution value of 0.59 [16]. This point can be determined graphically from the point of intersection of the plots of peak width and peak spacing *versus* DNA size. Another definition of the LOR is associated with the use of base-calling software used with DNA sequencing instruments. It is defined as the number of bases that can be called accurately, and requires single-base resolution of DNA fragments. Expert computer programs that are able to call bases accurately at resolution as low as 0.25 have been reported [44].

5 ssDNA *versus* dsDNA

DNA molecules in aqueous solution can exist in native double-stranded form or in denatured, single-stranded form, depending upon conditions such as solvent and temperature. DNA sequencing is necessarily performed under denaturing conditions, because the sequencing reactions deliver ssDNA molecules with a partly elongated complementary DNA strand attached. Also, ssDNA strands have the propensity to form intrastrand base pairs and secondary structure. In particular, genomic DNA often has "trouble regions" of repetitive sequence that are difficult to fully denature and hence difficult to sequence. Full denaturation is needed to enable correct and reliable size-based separation depending strictly upon chain length. For DNA sequencing applications, DNA denaturation is typically achieved by using high concentrations of urea (> 4 M, more typically 7–8 M), formamide (~ 10%) and high temperatures (up to 70°C), or any combination of the three.

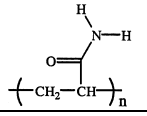
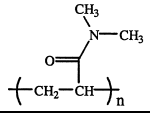
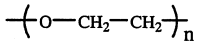
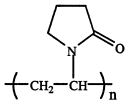
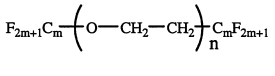
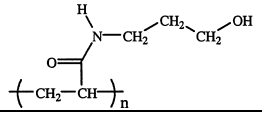
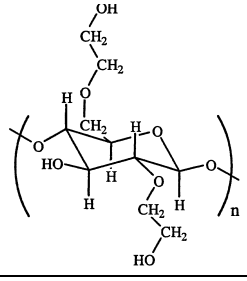
For the separation of PCR products and restriction fragments, DNA molecules are separated under nondenaturing conditions in the double-stranded form. For such separations, single-base resolution is not a critical requirement, allowing the use of less concentrated, and hence less viscous, polymer solutions, and permitting easier filling and replacement of polymer matrix in the capillaries. Van der Schans *et al.* [45] reported that when DNA is separated in single-stranded form, the selectivity of separation is increased, which results in significantly higher resolution of DNA fragments up to 500 bases. Using 4% linear polyacrylamide solution, baseline resolution of fragments differing in size by four base pairs in the 200 base pair range was achieved under denaturing conditions, while the separation failed under native conditions.

In a series of articles, Heller [15, 46, 47] compared the separation of ssDNA and dsDNA (< 700 bases) using similar conditions except that electrophoretic runs under denaturing conditions were carried out at 50°C, in the presence of 4 M urea, whilst for nondenaturing conditions the temperature was 25°C. In all cases, peak spacing decreases with increasing DNA size, but increases with increasing polymer concentration. However, separations of ssDNA fragments give larger peak spacings than separations of dsDNA fragments (< 700 bases or base pairs). Peak widths are smaller under denaturing compared to nondenaturing conditions. As a result, the resolution of ssDNA fragments is much higher than that for dsDNA of the same length. The use of strongly alkaline conditions (pH 11) has been reported to denature DNA fragments in hydroxethylcellulose (HEC) solution and to improve the resolution of DNA fragments compared to that of dsDNA at neutral pH [48]. Although separation of DNA has been shown to be possible in dilute solutions, concentrated, entangled polymer solutions gave superior separations of DNA as compared to dilute solutions [48]. Therefore, for high-resolution applications, such as DNA sequencing, relatively concentrated, entangled polymer solutions are needed.

6 Polymer matrices used in DNA sequencing

Several different types of water-soluble polymers have been used in DNA sequencing by CE. The choice of polymer has often been arbitrary and empirical, primarily because the mechanism of DNA separation in uncrosslinked polymer solutions is not fully understood. The chemical structures of the polymers that have been tested for this application are shown in Table 1. In addition to chemical properties, the physical properties of polymers used for CE are critically important, as they control the attributes of the entangled polymer network and hence influence the predominant mechanism and time-scale of DNA-polymer and polymer-polymer interactions [23]. To compare the performance of different polymers as DNA sequencing matrices, a number of important issues need to be evaluated simultaneously: resolving power, speed of separation, viscosity of polymer solution and suppression of electroosmotic flow. Resolving power reflects the ability of the polymer matrix to achieve long LORs, which is a highly desired feature for genome sequencing projects, though less important for many other biological and medical applications of DNA sequencing. Long LORs achieved within reasonable time limits, minimize sample preparation and computational effort required assembling the sequencing data into finished sequence. Thus, long LORs increase the throughput and reduce the cost of sequencing.

Table 1. Structures of polymers used for DNA sequencing by CE

Polymer	Chemical structure
LPA	
PDMA	
PEO	
PVP	
PEG (end-capped)	 m = 6 or 8
poly(AAP)	
HEC	

The viscosity of a polymer matrix is an important factor considered in designing and engineering automated CE sequencing instruments. Low-viscosity polymer solutions may be pumped into the capillary and replaced using a practically achievable and robust pressurizing system, and the requisite matrix loading step may significantly contribute to the turnaround time and cost of the instruments. More subtly, the dependence of polymer solution viscosity on the rate of applied shear is an important attribute of DNA sequencing matrix, as it controls the rate of flow of the solution through the microchannel under a given applied force. Another factor that contributes to cost is the ability of the polymer to adsorb on the internal capillary wall and suppress electroosmotic flow, thus eliminating the need for a covalently bound capillary coating. The lifetime of the coating, as well as the ease of its regeneration, impact the overall performance and cost of sequencing.

6.1 Linear polyacrylamide

Replaceable LPA, introduced for DNA sequencing in 1993 [49], has become one of the most widely employed separation media in DNA sequencing by CE due to its excellent performance in terms of LOR and separation time. Carrilho *et al.* [50] investigated the effect of molecular mass and concentration of LPA on DNA sequencing. For a given high LPA molecular mass, it was found that higher solution concentrations (4% LPA) improved the resolution of DNA fragments smaller than 450 bases, while a lower concentration (2% LPA) gave better resolution of DNA fragments larger than 450 bases. Others reported similar results on the effect of LPA concentration on separation performance [51, 52]. Selectivity for low base numbers was higher for more concentrated polymer solutions, and decreased with base number for all polymer solutions. In addition, the rate of change of selectivity with base number was significantly greater for high-concentration LPA solutions. For any polymer concentration, peak efficiency decreased with increasing DNA fragment size; however, there was no significant difference in efficiency for a specific fragment with changing LPA concentration, suggesting that simply increasing the polymer

concentration will not serve to improve the resolution of DNA peaks.

Using a 4% LPA solution, the selectivity of bases up to 500 bases was higher with matrices formulated with high-molecular mass LPA as opposed to low-molecular mass LPA. For fragments larger than 500 bases, the selectivity merged to a common low value for all matrices. Migration time was found to be more dependent on polymer concentration than on polymer molecular mass. For each polymer solution, a maximum in efficiency was observed, and this maximum was linearly shifted to longer DNA fragments with increased LPA molecular mass. Moreover, the absolute value of the maximum efficiency increased with LPA molecular mass. Also, lowering the electric field strength from 200 V/cm to 150 V/cm improved the selectivity and shifted the onset of biased reptation to longer DNA fragments, extending the sequencing LOR at the expense of run time. Thus, the use of a low concentration of high-molecular mass LPA extended the LOR. Using 2% high-molecular mass LPA (> 5.5 MDa) solution, a column temperature of 50°C, and 150 V/cm field strength, sequencing analysis of more than 1000 bases was achieved with 96.8% accuracy in 80 min. A similarly spectacular performance was obtained by high-molecular mass (9 MDa) LPA prepared by inverse emulsion polymerization [53].

The performance of LPA sequencing performance was further improved by fine-tuning of polymer molecular mass distribution, LPA solution composition, electric field strength, run temperature, internal capillary diameter, dye chemistry, sample clean-up protocols [41], and base-calling software. The final conditions comprised a 2.5% LPA mixture consisting of 2% high-molecular mass (9 MDa) and 0.5% low-molecular mass (50 kDa) LPA operated at 60°C and 200 V/cm. The new conditions produced an LOR of 1000 bases in 55 min at 98–99% accuracy. An increase of LPA concentration by 0.5% with low-molecular mass LPA improved the separation of DNA fragments below 100 bases, without causing a significant reduction in the resolution of the large fragments. This was explained by the fact that short-chain polymers interact preferentially with small DNA molecules, whereas long-chain polymers interact mainly with large DNA fragments. More likely, small-DNA separation is insensitive to polymer molecular mass, but instead is dependent on overall polymer concentration, whereas large DNA can only be separated by large polymers [34].

An increase in the run temperature shifts the onset of biased reptation to longer DNA fragments [20], and this counterbalances the effect of increased electric field strength on the onset of biased reptation (in terms of DNA chain length). Moreover, the use of elevated temperature helps to minimize compressions that result from DNA secondary structure, improving resolution and reducing separation time [51, 54, 55]. Table 2 summarizes the effect of LPA solution composition, temperature and electric field strength on electrophoretic separation of DNA fragments as determined by Karger's group [44]. Optimal results were obtained using an LPA matrix comprising 0.5% 270 kDa and 2% 17 MDa LPA. This formulation delivered an LOR of 1300 bases (average 1249 bases) at 98.5% accuracy in 2 h at 70°C and 125 V/cm. The final resolution was 0.25. This work also involved some improvements in the base-calling software and the DNA Sanger cycle sequencing reaction.

Wu *et al.* [56] studied the separation of ssDNA sequencing reaction products up to 393 bases in size in polyacrylamide solutions. It was found that the electrophoretic mobility of DNA molecules up to 393 bases in size increased with decreasing polyacrylamide concentration, leading to a faster separation of DNA fragments. The molecular mass and polydispersity of the polymer did not seem to have a significant effect on the mobility. However, increasing the average polyacrylamide chain length reduced the dispersion of electrophoresing DNA and increased the sharpness of the bands, implying an increase in efficiency of separation.

Using low-viscosity (150 cP), 6.2% low-molecular mass LPA (339 kDa) solution, Grossman [57] resolved DNA sequencing fragments up to 580 bases long in 110 min at 218 V/cm. Under the run conditions, the selectivity decreased with fragment size, whilst the efficiency of separation increased, and did not exhibit a maximum value

Table 2. Comparison of migration times and LORs for different compositions of LPA matrices at different temperatures and electric field strengths [44]

LPA	Temp. (°C)	Electric field (V/cm)	Migration time for base 1019 (min)	LOR at at 98.5% accuracy (bases)
2% 17 MDa + 0.5% 270 kDa	70	125	105.4	1249
2% 10 MDa + 0.5% 270 kDa	70	125	101.0	1190
2% 17 MDa + 0.5% 50 kDa	70	125	100.0	1083
2% 10 MDa + 0.5% 50 kDa	70	125	99.5	965
2% 10 MDa + 0.5% 50 kDa	60	200	55.6	1013
2% 10 MDa	50	150	81.0	951
2% 10 MDa + 0.5% 270 kDa	70	250	44	927
2% 10 MDa + 0.5% 270 kDa	70	200	55.6	1042
2% 10 MDa + 0.5% 270 kDa	70	150	80.5	1127
2% 10 MDa + 0.5% 270 kDa	70	100	131.0	1172

as was observed by Karger's group using high-molecular mass LPA [50, 58]. Empirical models, derived by Manabe *et al.* [52], predicted that a high electric field and a high LPA concentration are needed to achieve high separation efficiency, whilst a low electric field and a low polymer concentration are required to achieve long LORs. Using 9%T *in situ* polymerized LPA and an electric field strength of 100 V/cm, single-base resolution of ssDNA sequencing fragments up to 520 bases was obtained. Zhang *et al.* [51] separated DNA sequencing fragments up to 640 bases in 2 h using a 5%T polyacrylamide solution at 60°C and 150 V/cm.

Although LPA matrices have a high sieving capacity for DNA fragments, they have a few drawbacks. One of the main problems of high-molecular mass LPA is the high viscosity of its solutions, a result of the high hydrophilicity of the polymer. A saving grace is that LPA solutions behave as non-Newtonian (shear-thinning) fluids, such that the viscosity of high-molecular mass LPA comes close to that of concentrated low-molecular mass LPA when the shear force is increased [53]. Nonetheless, the viscosity of the 2%, 9 MDa LPA was *ca.* 260 000 cP at zero shear and about 27 000 cP at a shear rate of 1.32 s^{-1} [53]. High-molecular mass LPA also requires careful preparation and handling, because the method of dissolution can affect the mass distribution through chain scission, which can adversely affect the subsequent separation of DNA sequencing fragments. Chemical instability towards alkaline hydrolysis and the neurotoxicity of acrylamide monomer are other drawbacks of LPA [59, 60]. Another problem with using LPA for CE is that it cannot be used with bare fused-silica capillaries. The capillary has to be precoated to suppress electroosmotic flow. Capillary coatings have the propensity to become fouled after a number of DNA sequencing runs, and this leads to an eventual deterioration of the separation performance.

6.2 Polyethylene oxide

Polyethylene oxide (PEO) is another polymer that has been extensively studied for DNA sequencing [42, 43, 61, 62]. It was first introduced in 1995 by Fung and Yeung [61]. In 1997, Kim and Yeung [42] reported the separation of a single-color DNA sequencing ladder up to 1000 bases (resolution of raw data = 0.5 at 966 bases) in a mixture of 1.5% high-molecular mass PEO (8 MDa) and 1.4% low-molecular mass PEO (0.6 MDa) at 75 V/cm. A separation time of 7 h was needed to achieve this separation, which is very long compared to the 1 h run time needed for LPA to achieve the same LOR [58]. Increasing the concentration of the high-molecular mass PEO to 2.0% decreased the size of the largest DNA fragments that could be resolved. The influences of separation voltage, column length, and PEO composition were also studied. In agreement with findings for other polymers [50, 63], PEO performance increased with lower field strengths and longer separation distances. The selectivity up to 320 bases was independent of the electric field strength and decreased with increasing field strength for DNA larger than 320 bases. The efficiency of the separation increased to a maximum at around 250 bases and then decreased with increasing base number. Regarding the PEO composition, it was found that when low-molecular mass PEO was blended with high-molecular mass PEO, the separation of smaller DNA fragments improved significantly, a result similar to that obtained for other polymers [40, 44, 58]. For best sequencing performance, the total PEO concentration was found to be between 2.5 and 3.0%.

PEO has the interesting and favorable property of being "self-coating" through adsorption to the capillary wall [61]. This potentially eliminates the need for covalent wall passivation, as is required for LPA matrices, thereby reducing the material cost of DNA sequencing. PEO-coated capillaries had to be flushed with HCl and PEO solutions before being reusable and this required about 2 h. It was further observed that PEO capillaries could not be used for more than 30 runs without reconditioning. This was tentatively attributed to the formation of PEO aggregates on the wall surface, and was increased by increasing the PEO concentration. The adsorbed PEO coating was shown to reduce the electroosmotic flow velocity by more than one order of magnitude compared to the bare capillary at pH 7.0. However, at pH 8.2, the reduction in electroosmotic flow was only 30–50% [64].

6.3 Poly-*N,N*-dimethylacrylamide

Madabhushi [65] introduced the acrylamide derivative poly-*N,N*-dimethylacrylamide (PDMA) for DNA sequencing. The author considered low viscosity of the sieving matrix to be a higher priority than the attainment of a long LOR. Low viscosity translates to the ability to use lower pressure for filling and flushing the capillary, thus providing faster and safer operation. An LOR of 600 bases was achieved in 125 min using a 6.5% low-molecular mass (98 kDa) PDMA, the viscosity of which was claimed to be just 75 cP, many orders of magnitude lower than that of LPA solutions. The increase in PDMA molecular mass beyond 98 kDa slightly improved the resolution but adversely increased the viscosity of the solution. PDMA also exhibits two advantages over LPA: (i) it is resistant to hydrolysis [59]; (ii) it possesses self-coating properties, *i.e.*, it physically adsorbs to the capillary wall and virtually eliminates electroosmotic flow. The dynamic coating process required no special conditioning, *e.g.*, no HCl

wash, and capillaries could be used directly after dynamic coating. The capillary was typically ready for use after 1 min of sieving matrix replacement by low pressure. It was reported that at least 100 sequencing runs could be performed in a single capillary using this simple procedure without any adverse effect. In four-color sequencing, 600 bases could be analyzed with a final resolution of 0.59. The accuracy achieved was more than 98% for the first 450 bases and greater than 96.5% for the 600 bases. The run was performed at 42°C and 160 V/cm.

Heller [15, 30, 47, 63] investigated the use of PDMA for ssDNA separation. Peak spacing was found to decrease with increasing DNA size, but increased with increasing polymer concentration. Moreover, for a given PDMA concentration, the peak spacing increased with increasing PDMA chain length. The peak width decreased, reflecting an increase in the efficiency of separation, with increasing polymer concentration and polymer chain length, but it did not change as a function of DNA size. 5% solutions of low-molecular mass (216 kDa) or high-molecular mass (1.15 MDa) PDMA achieved single-base resolution up to 500 bases, but beyond that point the resolution failed dramatically. The use of lower PDMA concentration (1 and 2%) resulted in lower resolution of DNA but the rate of decrease of resolution with DNA base number was smaller than that using 5% solution. As with other polymers, a reduction in electric field strength improved the resolution of the DNA fragments.

The main disadvantage of PDMA is that it is relatively hydrophobic compared to LPA [59]. This increases the extent of hydrophobic interaction between the polymer and the hydrophobic DNA-labeling dyes, resulting in changes in DNA mobility. This can lead to peak shifting and band broadening, thus affecting the resolution of DNA fragments and increasing the chances of errors in base calling. This increased hydrophobicity also causes the PDMA chains to adopt more compact coils in aqueous solution at a given chain length, reducing the average number of chain entanglements and producing a less robust entangled network than, for example, LPA, which is more hydrophilic.

6.4 Polyvinyl pyrrolidone

Using a mixture of PDMA and polyvinyl pyrrolidone (PVP), Madabhushi *et al.* [66] demonstrated the feasibility of four-color sequencing in a bare silica capillary. Later, Gao and Yeung [67] explored the use of PVP as a DNA sieving matrix. The polymer is soluble in water and aqueous buffers, forming a homogenous, low-viscosity (27 cP at 4.5%) solution within less than 1 h. The polymer exhibited a self-coating property and reduced electroosmotic flow to negligible levels. The low viscosity of the polymer solution reduced the time needed for regeneration of the column surface to 3 min, compared to 2 h for PEO. Single-color sequencing with 7% PVP of 1 MDa molecular mass showed separation of up to 350 bases. The separation was extended to more than 500 bases when a 5% solution of high-molecular mass PVP extracted from the 1 MDa material was used. The sequencing time was 83 min in a 50 cm effective column length at 150 V/cm and room temperature.

6.5 Polyethylene glycol with fluorocarbon tails

A novel class of material was reported by Menchen *et al.* [68] who modified polyethylene glycol (PEG) by end-capping it with a fluorocarbon tail. This material self-assembles in water into equilibrium network structures with a well-defined mesh size. Above a critical concentration, the hydrophobic end groups associate into micelles in water to form extended networks. These micellar systems form flowable gel-like networks that permit electrophoretic DNA sequencing in a capillary column. In addition, the network structure is disrupted under high shear, allowing relatively easy replacement of the sieving matrix. DNA electrophoretic mobility appeared not to be adversely affected by the hydrophobicity of the polymer tails. The use of longer PEG chains decreased the band width and improved the resolution of DNA sequencing fragments. Polymer concentration influenced the peak-to-peak spacing. Longer LORs were obtained by using the polymer at concentrations above the critical micelle concentration. Increasing the temperature from 22°C to 55°C destabilized the polymer network and resulted in severe band broadening and loss of resolution. Optimum sequencing results were obtained from a 6% solution of a 1:1 mixture of C_6F_{13} end-capped and C_8F_{17} end-capped PEG of 35 kDa molecular mass. An LOR of 450 bases was obtained at 200 V/cm. The run time was not reported.

6.6 Hydroxyethylcellulose

The use of HEC as a sieving matrix of ssDNA was explored by Bashkin *et al.* [69]. The best performance was obtained at 190 V/cm using 2% HEC of 90–105 kDa molecular mass. Sharp, well-resolved peaks of one-color sequencing reaction products were observed to around 570 bases in 70 min. Higher voltages shifted the onset of the biased reptation to smaller DNA fragments. Purification of the HEC was needed to remove oxidized forms of cellulose and other charged impurities that affected the performance of the crude HEC sieving matrix. The use of entangled HEC solution at pH 11 significantly improved the efficiency and hence the resolution of ssDNA frag-

ments in denaturing buffer [48]. The drawback of these HEC matrices that prevented their use in a commercial high-throughput sequencing instrument such as the MegaBACE 1000^{TM}, was that in a practical setting with realistic DNA sequencing samples, LORs of *ca.* 450 bases were typical. Inconsistent performance of HEC batches stems from the fact that this is a naturally derived, not synthetic, polymer.

Using one-color DNA sequencing data, Dolník and Gurske [70] derived empirical equations for the relationship between selectivity per base and sieving matrix concentration and between the mobility slope and matrix concentration using HEC (300 kDa). The same authors proposed the use of the inflection slope, *i.e.*, the slope of the log-log mobility curve at its inflection point, as a quantitative parameter of sieving performance. With increasing polymer concentration, the inflection slope decreased and moved to shorter DNA length and eventually reached a plateau. Kheterpal and Mathies [71] reported a comparison of the performance of replaceable LPA, HEC, and a mixture of PEO and HEC under identical conditions for DNA sequencing using identical separation conditions. Although experimental details were not given, LPA was said to provide the longest LOR (~ 1000 bases) compared with ~ 600 bases in HEC and < 300 bases in the HEC-PEO mixture. Mixtures of PEO and HEC were reported by Kim and Yeung [42] to give unreadable resolution of DNA sequencing fragments.

6.7 Poly-*N*-acryloylaminopropanol

Righetti's group [60, 72, 73] explored the utility of a novel acrylamido monomer, namely *N*-acryloylaminopropanol (AAP). The resistance to alkaline hydrolysis of poly(AAP) was shown to be 500 times higher than that of polyacrylamide. Moreover, AAP is more hydrophilic than acrylamide. Single-color DNA sequencing was performed at 50°C and 200 V/cm using a 10.9%T poly(AAP). An LOR of 385 bases was achieved in about 100 min at a final resolution of 0.5. An increase in run temperature from 25°C to 50°C was efficient in resolving compression of DNA zones. The use of formamide was detrimental to gel stability and sieving performance. The performance of an *in situ* cross-linked poly(AAP) gel, with linear poly(AAP) as an additive, was also evaluated. The gel was formed by introducing into the capillary a mixture of the 5.45% AAP monomer, 1.1% dihydroxyethylenebisacrylamide cross-linker and 5.45% prepolymerized poly(AAP). A rise in temperature from 25 to 50°C reduced the LOR from 400 bases to 303 bases and adversely affected the efficiency of the separation and the resolution of the DNA fragments.

6.8 Poly(*N,N*-dimethylacrylamide-co-*N,N*-diethylacrylamide)

Sassi *et al.* [74] observed that for uncross-linked, thermoresponsive polymer solutions, the reversible volume-phase transition at the lower critical solution temperature (LCST) results in a precipitous viscosity drop, by at least one order of magnitude. Taking advantage of this property exhibited by hydrophobically modified polymers, the authors introduced a new approach towards solving the problem of obtaining a single-base resolution of DNA fragments in entangled polymer solutions and ease of handling and loading of polymer solutions into the capillary. Thermoresponsive polymers can be loaded into the capillary at high temperatures (≥ LCST) at which the viscosity is low, and the viscous, entangled network can be regained by cooling the polymer solution below its LCST at which DNA separation can be achieved.

Using copolymers of *N,N*-dimethylacrylamide (DMA) and *N,N*-diethylacrylamide (DEA), it was shown that these copolymers could show single-base resolution of DNA sequencing products up to 150 bases in a solution of 6% copolymer. The copolymer had a viscosity of 10 cP at 70°C, making it easy to load and replace in the capillary. We have developed different formulations of linear DEA/DMA copolymers for use in DNA sequencing (Buchholz *et al.*, in preparation). Our results show that 420 bases can be sequenced at 98.5% accuracy in 75 min, using a copolymer of 53% DEA, 47% DMA copolymer (4.0 MDa molecular mass). The electrophoretic run was performed at 44°C and 140 V/cm. Furthermore, it was found that increasing the DEA composition decreased the LOR due to increasing polymer hydrophobicity, which caused band broadening and peak shifting. A comparison of the sequencing performance, rheological properties and ability to suppress electroosmotic flow of the various polymers used for DNA sequencing by CE is given in Tables 3, 4 and 5, respectively.

7 Polymers for separation of dsDNA

The separation of dsDNA fragments has a number of important biological applications, such as the separation of PCR products, restriction fragments, and plasmids. For such applications, single-base resolution is not a necessity. Thus, many of the polymers that have been used for DNA sequencing have provided excellent separation of dsDNA [32, 34, 40, 43, 62, 67, 73–77]. A variety of polymers has also been successfully explored for the separation of dsDNA, and may potentially be applicable for DNA sequencing under the appropriate conditions. For several of these polymers, the challenge has been to achieve sufficiently high-molecular mass polymers to form robust, entangled networks as needed for long-read DNA sequencing.

Table 3. Comparison of performance of polymer solutions for DNA sequencing by CE

Polymer	Mol. mass (kDa)	Conc. (wt%)	Electric field (V/cm)	Effective length (cm)	Temp. (°C)	Buffer[a)]	LOR (bases)	Time (min)	Resolution at LOR	Reference
LPA	> 5500	2	150	30	50	TTE, 7 M urea	1000	80	0.4	[50]
LPA	9000	2	150	30	50	TTE, 7 M urea	1000	80	0.4	[53]
LPA	9000 + 50	2 0.5	200	30	60	TTE, 7 M urea	1000	55	0.25	[58]
LPA	17 000 + 270	2 0.5	125	30	70	TTE, 7 M urea	1300	120	0.25	[44]
LPA	1000	8	100	12	Room	TBE, 3.5 M urea, 30% formamide	393	90	nr[a)]	[56]
LPA	339	6.2	218	40	Room	180 mM Tris-H_3PO_4, 8 M urea	580	109	0.75	[57]
LPA	nr[b)]	9 T	100	40	Room	100 mM Tris-borate, 7 M urea	520	430	0.55	[52]
LPA	nr[b)]	5 T	150	39	60	TBE, 3.5 M urea	640	118	nr[b)]	[51]
PEO	8000 + 600	1.5 1.4	75	70	Room	TBE, 3.5 M urea	1000	420	0.5	[42]
PDMA	98	6.5	160	40	42	100 mM TAPS, 8 M urea	600	125	0.59	[65]
PDMA	1145	5	210	36	50	0.5 × TBE, 4 M urea	500	34[c)]	0.5	[47]
PVP	1000	5	150	50	Room	TBE, 3.5 M urea	500	83	nr[b)]	[67]
PEG[d)]	35	6	200	36	22	100 mM TAPS, 6.6 M urea	450	nr[b)]	0.59	[68]
HEC	97	2	190	41	Room	TBE, 6 M urea, 10% formamide	570	70	0.5	[69]
Poly(AAP)	nr[b)]	10.9 T	200	40	50	TBE, 7 M urea	385	100	0.5	[55]
Poly(DMA/DEA)	nr[b)]	6 T	170	45	Room	TBE, 7 M urea	150	60	nr[b)]	[74]

a) TBE buffer: 89 mM Tris, 89 mM borate, 2 mM EDTA; TTE buffer: 50 mM Tris, 50 mM TAPS (*N*-tris(hydroxymethyl)methyl-3-aminopropanesulfonic acid) and 2 mM EDTA
b) No value reported
c) Estimated from electrophoretic mobility data from [15]
d) 1:1 mixture of C_6F_{13} end-capped and C_8F_{17} end-capped PEG 35 kDa

Table 4. Comparison of viscosities of polymer solutions used for DNA sequencing

Polymer	Composition (wt%)	Mol. mass (kDa)	Buffer[a)]	Temp. (°C)	Viscosity (cP)	Reference
LPA	6.2	339	180 mM Tris-H_3PO_4, 8 M urea	25	150	[57]
LPA	2	9000	TTE, 7 M urea	25	27 000[b)]	[53]
LPA	2	9000	TTE, 7 M urea	25	3700[c)]	[53]
LPA	6	nr[d)]	TBE, 3.5 M urea	25	4900	[61]
PEO	1.5 + 1.4	600 8000	TBE, 3.5 M urea	Room	1200	[61]
PDMA	6.5	98	100 mM TAPS, 8 M urea	30	75	[65]
PDMA	6.5	200	100 mM TAPS, 8 M urea	30	1200	[65]
PEG[e)]	6	35	100 mM TAPS, 6.6 M urea	nr[d)]	10 000	[68]
HEC	2	97	TBE, 6 M urea, 10% formamide	25	5000	[69]
PVP	4.5	1000	TBE	20	27	[67]
Poly(DMA/DEA)	6	nr[d)]	TBE, 7 M urea	70	10	[74]

a) Buffer compositions as in Table 3
b) Viscosity measured at 1.32 s^{-1} shear rate
c) Viscosity measured at 22 s^{-1} shear rate
d) No value reported
e) 1:1 mixture of C_6F_{13} end-capped and C_8F_{17} end capped PEG 35 kDa

Table 5. Suppression of electroosmotic flow by different DNA sequencing polymers

Polymer	Mol. mass (kDa)	Conc. (wt%)	Buffer[a)]	Temp. (°C)	Field strength (V/cm)	Electrophoretic mobility (10^{-4} cm^2/V s)	Reference
None	–	–	100 mM glycine	30	200	6.29	[65]
LPA	150	0.01	100 mM glycine	30	200	2.43	[65]
PDMA	98	0.01	100 mM glycine	30	200	0.20	[65]
PVP	40	0.01	100 mM glycine	30	200	0.87	[65]
PVP	360	0.01	100 mM glycine	30	200	0.47	[65]
PEG	35	0.01	100 mM glycine	30	200	1.44	[65]
PNIPA	50	0.01	100 mM glycine	30	200	0.34	[65]
PDMA	nr[b)]	0.01	TTE	nr[b)]	nr[b)]	0.17	[80]
PDMA	nr[b)]	0.05	TTE	nr[b)]	nr[b)]	0.107	[80]
PVP	1000	0.1	TBE	Room	250	0.27	[67]
PVP	1000	1.0	TBE	Room	250	0.06	[67]
PEO	8000	0.2	Phosphate, pH 7.0	Room	nr[b)]	0.2	[64]
PEO	8000	0.2	Phosphate, pH 8.2	Room	nr[b)]	2.5	[64]

a) Buffer composition as in Table 3
b) No value reported

Several novel *N*-substituted acrylamides, which combine two important properties needed for DNA fragment separation, hydrophilicity and resistance to alkaline hydrolysis, have been proposed by Righetti and co-workers [59, 60, 72]. Chiari *et al.* [59] introduced *N*-acryloylethoxyethanol (AAEE) as a monomer that could be polymerized *in situ* yielding a sieving matrix for dsDNA that gave excellent performance in separation of DNA fragments ranging from few hundreds to 23 000 bp [78]. Furthermore, poly(AAEE) offered excellent performance as a covalent coating of the silica capillary wall [79]. However, while the polymer offered good performance in CE, the monomer suffered from two main limitations. First, the yield of the monomer synthesis was low, at less than 18% [59]. Second, it was found that the monomer exhibited a unique instability. Upon storage, even in the presence of polymerization inhibitors, it could suddenly auto-polymerize [60], which greatly limited its shelf life.

N-(acrylaminoethoxy)ethyl-β-D-glucopyranose (AEG) is another monomer bearing hydrophilic glucose units on *N*-substituted acrylamido derivatives, synthesized by Chiari *et al.* [75] *via* a chemo-enzymatic pathway. Radical polymers of AEG were found to have low viscosity in solution relative to polyacrylamide. Moreover, the viscosity of poly(AEG) solution increases weakly with polymer concentration, allowing the replacement in the capillary of more concentrated solution. Poly(AEG) was reported to have a dsDNA sieving capacity fully comparable with that of polyacrylamide, but migration was faster in the former, possibly due to the formation of a more flexible network with less robust entanglements.

Chiari *et al.* [80] proposed another novel low-viscosity glucose-containing copolymer of acrylamide and allyl-β-D-glucopyranose (AG). The viscosity of the copolymer solution decreased as the sugar content of the copolymer increased. In spite of its low molecular mass (209 kDa), the copolymer obtained by polymerizing a 1:1 molar ratio of AG:acrylamide showed the ability to sieve oligonucleotides and dsDNA ranging in size from a few tens up to 600 bases. The matrix showed good resolving power of DNA restriction fragments. However, single-base resolution needed for DNA sequencing application was not demonstrated.

Another class of material that has been extensively studied for the separation of dsDNA fragments is the group of polysaccharides, such as glucomannan [81], dextran [46, 82], agarose solutions [83], cellulose derivatives, including methylcellulose (MC) [84], HEC [16, 31, 32], hydroxypropylcellulose (HPC) [34, 84, 85] and hydropropylmethylcellulose (HPMC) [84, 86, 87], and composite agarose/HEC [88]. HEC polymers in water assume a stiff and extended conformation [31]. Cellulose derivatives are typically used to separate DNA fragments at concentrations of 0.2–1%. Due to the fact that stiff and extended molecules become entangled more easily than flexible, random-coil polymers and, furthermore, form more robust entanglements, the operative concentration of cellulosic polymers is considerably lower than that generally used for LPA. Baba *et al.* [84] studied the performance of cellulose derivatives, MC, HPC and HPMC, as sieving networks for DNA separation by CE. Solutions of MC, HPC, and HPMC at a concentration of 0.5% provided separa-

tion of large fragments (from 1000 to 23000 bp) of a 1 kbp DNA ladder. However, a higher concentration (0.7%) was needed to improve the resolution of smaller fragments from 75 to 500 bp. Using HPC (1 MDa) solution, Mitnik *et al.* [85] showed that increasing the electric field strength increased the separation speed and improved the resolution of small DNA fragments. Resolution of large fragments was lost at electric fields greater than 500 V/cm. In agreement with other findings, the resolution of DNA fragments larger than a certain size (600 bp) degraded with increasing polymer concentration whilst the resolution of smaller fragments improved.

Barron *et al.* [34] evaluated the separation of dsDNA in HEC, HPC and LPA in the dilute and entangled solution regimes. HEC, which is a stiffer polymer than HPC and LPA, provided better resolution of large DNA fragments (> 603 bp) than the other two polymers. This result pointed out the importance of polymer chain stiffness (or persistence length) for the resolution of large DNA fragments. Moreover, it was shown that the three polymers, HEC, HPC and LPA, show similar trends with respect to the effect of molecular mass and polymer concentration on the sieving performance. The authors also investigated the effect of polymer polydispersity on DNA separation performance using 1 MDa LPA of different polydispersities. It was concluded that polymer polydispersity had only a small effect on the separation of small DNA fragments whereas the separation of large DNA fragments is significantly improved by the use of low-polydispersity, high-molecular mass polymer solutions. All three of these polymers separate large dsDNA fragments (> 600 bp) well in the unentangled regime, if countermigration CE in uncoated capillaries in employed. In coated capillaries, lower resolution was obtained [89] because the effective separation distance is shortened in this mode of separation.

Thermoresponsive copolymers of poly-*N*-isopropylacrylamide (PNIPA), densely grafted with short PEO chains, were applied for dsDNA separation by CE [90, 91]. The copolymer had a random-coil conformation at low concentrations and low temperatures (< 31°C). At temperatures above 31°C, the copolymer underwent a broad coil-to-globule transition accompanied by a small decrease in viscosity (~ 10%), with the PNIPA chains inside the core and the hydrophilic PEO on the surface. With a further increase in temperature (> 35°C), the PNIPA backbone chains became more hydrophobic so that the interchain aggregation became very strong at high concentrations, resulting in a dramatic increase in viscosity. As a separation medium for dsDNA restriction fragments at temperatures less than 31°C, the copolymer provided a resolution comparable to that of LPA with the advantage of a shorter migration time. Moreover, the copolymer PNIPA-g-PEO showed a self-coating ability for the capillary inner wall. At temperatures above 31°C, the copolymer chains collapsed to the globular state, destroying the chain network and thus losing their DNA-separating ability. These matrices are likely to be too hydrophobic in nature (due to the PNIPA portion of the chains) to be useful for high-efficiency DNA sequencing.

Triblock copolymers of PEO_{99}-PPO_{69}-PEO_{99} (PPO being polypropylene oxide), also known as Pluronic F127®, showed a thermoresponsive behavior and capillary coating ability similar to that of PNIPA-g-PEO [92–96]. Increasing the temperature from 4 to 25°C, 18–30% F127 aqueous solution was transformed from a low-viscosity solution (40 cP at 4°C) to a transparent, self-supporting, gel-like liquid crystalline phase. The attainment of this gel-like state is thermally reversible. The crystalline phase consists of micelles, each composed of a dense core of PPO and a hydrated PEO shell. DNA fragments in the size range of 50–3000 bp were separated in a 20% Pluronic solution. Increasing the polymer concentration increased the number of micelles formed, resulting in slower DNA migration. This improved the resolution of the fragments but reduced the separable DNA size range to an upper limit of 1000 bp. Increasing the temperature led to the formation of more compact micelles, and reduced the interaction between the DNA fragments and the copolymer resulting in loss of sieving power. DNA separation in Pluronic liquid crystals is envisioned to occur by mechanisms other than the presently hypothesized ones for CE.

8 Conclusions

The intensive search for and optimization of polymer matrices applied for DNA sequencing seems to indicate that the optimal matrix has yet to be found. The systematic approach that has been undertaken to optimize the performance of these matrices have revealed that important properties for an “ideal” matrix for DNA sequencing by CE are: (i) high resolving power; (ii) high hydrophilicity; (iii) high average molecular mass to separate large DNA fragments; (iv) sufficient concentration to separate small DNA fragments; (v) hydrolytic stability at pH 7–8.5; (vi) relatively low viscosity for easy capillary loading and matrix replacement; and (vii) dynamic adsorption onto a bare capillary wall for suppression of electroosmotic flow.

For DNA sequencing applications, the use of concentrated, entangled polymer solutions is needed to provide single-base resolution. The delivery of long sequencing LORs is greatly dependent upon the stability and robustness of the entangled polymer network in solution. Good

resolution of electrophoresing DNA molecules can be obtained as long as the DNA-polymer interactions responsible for the size-based separation of DNA molecules do not disrupt the polymer-polymer entanglements and hence the polymer network. Hence, the extent and strength of entanglements between polymer chains in solution control the robustness of the network. Hydrophilic, high-molecular mass polymer chains such as LPA adopt an extended conformation in aqueous solution, allowing more frequent entanglements along the chains. An increase in polymer hydrophobicity, such as in PDMA, PVP and PEG, results in the adoption of more compact coil sizes in aqueous solution and the formation of less robust polymer networks, that do less well in resolving large DNA fragments than LPA. Side chains pendant on the polymer backbone also boost the entanglements between the polymer chains, as in the case of HEC. For PEO, the necessity for low field strengths (75 V/cm) and long run times most likely reflects the inherent flexibility of the unsubstituted PEO chains and hence the "delicate" nature of its entangled polymer networks. DNA molecules, traveling under powerful electric fields such as 150 V/cm, are simply not sufficiently hindered by PEO-PEO entanglements to be separated by chain entanglement mechanisms. A good analogy is to American football: large DNA molecules are like huge linebackers, shoving polymer chains out of their way as they barrel towards the anode. Strong points of entanglement that force DNA chains to pause in their headlong rush provide the necessary change in dependence of DNA molecular friction coefficient for separation. Although the presence of side chains on the matrix polymers has a positive effect on the robustness of the polymer network, bulky side chains, such as aminopropanol in AAP, can increase the difficulty of obtaining high-molecular mass polymers due to steric hindrance in the polymerization.

High polymer concentrations improve the separation of small DNA sequencing fragments (< 500 bases), whilst low concentrations of high-molecular mass polymers are better in resolving large DNA fragments (> 500 bases). Small DNA fragments are not long enough to interact sufficiently by entangling with the polymer chains of the matrix. Increasing the polymer concentration increases the frequency of small DNA-polymer collisions that lead to separation. For large DNA fragments, large polymer chains are needed to maintain the robustness of the entangled polymer network and prevent facile dragging of the polymer chains by migrating DNA chains.

Thermoresponsive polymers are potentially suitable candidates as DNA sequencing matrices for CE and microchips, allowing easy filling and replacement in the capillary at one temperature and providing high resolution of DNA sequencing reaction products at another, though typically at the cost of somewhat shorter LORs due to the hydrophobicity of these polymers. Further experimental and theoretical investigations are needed to elucidate the specific mechanisms by which polymer chemical and physical properties influence matrix performance for DNA sequencing separation. In light of the new insights researchers have gained into the impact of polymer properties on DNA sequencing performance, the door is still open to exploit and invent new polymers that will improve overall DNA sequencing throughput and reduce the cost-per-base rate.

Financial support was provided by the National Institutes of Health (NIH), Grant R01HG01970-01. We are grateful to Professor Monica Olvera de la Cruz (Northwestern University) for helpful discussions of DNA separation mechanisms.

Received July 5, 2000

9 References

[1] http://geta.life.uiuc.edu/~nikos/genomes.html
[2] Sanger, F., Nicklen, S., R., Coulson, A. R., *Proc. Natl. Acad. Sci. USA* 1977, *74*, 5463–5467.
[3] http://www.pebio.com/ab/about/dna/377/377a1a.html
[4] Marshal, E., Pennisi, E., *Science* 1998, *280*, 994–995.
[5] Brennan, M., Zurer, P., *Chem. Eng. News* 2000, *78*, 11.
[6] Huang, X. H. C., Quesada, M. A., Mathies, R. A., *Anal. Chem.* 1992, *64*, 2149–2154.
[7] Huang, X. H. C., Quesada, M. A., Mathies, R. A., *Anal. Chem.* 1992, *64*, 967–972.
[8] Carrilho, E., *Electrophoresis* 2000, *21*, 55–65.
[9] http://www.mdyn.com/products/MegaBACE/default.htm
[10] Swerdlow, H., Zhang, J. Z., Chen, D. Y., Harke, H. R., Grey, R., Wu, S., Dovichi, N. J., *Anal. Chem.* 1991, *63*, 2835–2841.
[11] Kambara, H., Takahashi, S., *Nature* 1993, *361*, 565–566.
[12] Mullikin, J. C., McMurray, A. A., *Science* 1999, *283*, 1867–1868.
[13] Woolley, A. T., Mathies, R. A., *Anal. Chem.* 1995, *67*, 3676–3680.
[14] Liu, S. R., Shi, Y. N., Ja, W. W., Mathies, R. A., *Anal. Chem.* 1999, *71*, 566–573.
[15] Heller, C., *Electrophoresis* 1999, *20*, 1962–1977.
[16] Grossman, P. D., Soane, D. S., *Biopolymers* 1991, *31*, 1221–1228.
[17] Viovy, J. L., Duke, T., Caron, F., *Contemp. Phys.* 1992, *33*, 25–40.
[18] Slater, G. L., Mayer, P., Drouin, G., *Methods Enzymol.* 1996, *270*, 272–295.
[19] Slater, G. W., Noolandi, J., *Biopolymers* 1986, *25*, 431–454.
[20] Lumpkin, O. J., Dejardin, P., Zimm, B. H., *Biopolymers* 1985, *24*, 1573–1593.
[21] Duke, T., Viovy, J. L., Semenov, A. N., *Biopolymers* 1994, *34*, 239–247.

[22] Cottet, H., Gareil, P., Viovy, J. L., *Electrophoresis* 1998, *19*, 2151–2162.

[23] Bae, Y. C., Soane, D., *J. Chromatogr.* 1993, *652*, 17–22.

[24] de Gennes, P. G., *Phys. Today* 1983, *36*, 33–39.

[25] Smith, D. E., Perkins, T. T., Chu, S., *Phys. Rev. Lett.* 1995, *22*, 4146–4149.

[26] deGennes, P. G., *J. Chem. Phys.* 1971, *55*, 572–579.

[27] Doi, M., Edwards, S. F., *The Theory of Polymer Dynamics*, Clarendon Press, Oxford 1986.

[28] Olivera de la Cruz, M., Deutsch, J. M., Edwards, S. F., *Phys. Rev. A* 1986, *33*, 2047–2055.

[29] Shaffer, E. O., Olivera de la Cruz, M., *Macromolecules* 1989, *22*, 1351–1355.

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

[31] Barron, A. E., Soane, D. S., Blanch, H. W., *J. Chromatogr.* 1993, *652*, 3–16.

[32] Barron, A. E., Blanch, H. W., Soane, D. S., *Electrophoresis* 1994, *15*, 597–615.

[33] Barron, A. E., Sunada, W. M., Blanch, H. W., *Biotechn. Bioengineer.* 1996, *52*, 259–270.

[34] Barron, A. E., Sunada, W. M., Blanch, H. W., *Electrophoresis* 1996, *17*, 744–757.

[35] Shi, X., Hammond, R., Morris, M. D., *Anal. Chem.* 1995, *67*, 1132–1138.

[36] Carlsson, C., Larsson, A., Jonsson, M., Norden, B., *J. Am. Chem. Soc.* 1995, *117*, 3871–3872.

[37] Sunada, W. M., Blanch, H. W., *Biotechnol. Progr.* 1998, *14*, 766–772.

[38] Sunada, W. M., Blanch, H. V., *Electrophoresis* 1998, *19*, 3128–3136.

[39] Hubert, S. J., Slater, G. W., Viovy, J. L., *Macromolecules* 1996, *29*, 1006–1009.

[40] Bunz, A. P., Barron, A. E., Prausnitz, J. M., Blanch, H. W., *Ind. Engineer. Chem. Res.* 1996, *35*, 2900–2908.

[41] Ruiz-Martinez, M. C., Salas-Solano, O., Carrilho, E., Kotler, L., Karger, B. L., *Anal. Chem.* 1998, *70*, 1516–1527.

[42] Kim, Y., Yeung, E. S., *J. Chromatogr. A* 1997, *781*, 315–325.

[43] Madabhushi, R. S., Vainer, M., Dolnik, V., Enad, S., Barker, D. L., Harris, D. H., Mansfield, E. S., *Electrophoresis* 1997, *18*, 104–111.

[44] Zhou, H., Miller, A. W., Sosic, Z., Buchholz, B., Barron, A. E., Kotler, L., Karger, B. L., *Anal. Chem.* 2000, *72*, 1045–1052.

[45] van der Schans, M. J., Kuypers, A. W., Kloosterman, A. D., Janssen, H. J., Everaerts, F. M., *J. Chromatogr. A* 1997, *772*, 255–264.

[46] Heller, C., *Electrophoresis* 1998, *19*, 1691–1698.

[47] Heller, C., *Electrophoresis* 1999, *20*, 1978–1986.

[48] Liu, Y., Kuhr, W. G., *Anal. Chem.* 1999, *71*, 1668–1673.

[49] Ruiz-Martinez, M. C., Berka, J., Belenkii, A., Foret, F., Miller, A. W., Karger, B. L., *Anal. Chem.* 1993, *65*, 2851–2858.

[50] 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.

[51] Zhang, J. Z., Fang, Y., Hou, J. Y., Jiang, R., Roos, P., Dovichi, N. J., *Anal. Chem.* 1995, *67*, 4589–4593.

[52] Manabe, T., Chen, N., Terabe, S., Yohda, M., Endo, I., *Anal. Chem.* 1994, *66*, 4243–4252.

[53] Goetzinger, W., Kotler, L., Carrilho, E., Ruiz-Martinez, M. C., Salas-Solano, O., Karger, B. L., *Electrophoresis* 1998, *19*, 242–248.

[54] Kleparnik, K., Foret, F., Berka, J., Goetzinger, W., Miller, A. W., Karger, B. L., *Electrophoresis* 1996, *17*, 1860–1866.

[55] Lindberg, P., Righetti, P. G., Gelfi, C., Roeraade, J., *Electrophoresis* 1997, *18*, 2909–2914.

[56] Wu, C. H., Quesada, M. A., Schneider, D. K., Farinato, R., Studier, F. W., Chu, B., *Electrophoresis* 1996, *17*, 1103–1109.

[57] Grossman, P. D., *J. Chromatogr. A* 1994, *663*, 119–227.

[58] Salas-Solano, O., Carrilho, E., Kotler, L., Miller, A. W., Goetzinger, W., Sosic, Z., Karger, B. L., *Anal. Chem.* 1998, *70*, 3996–4003.

[59] Chiari, M., Micheletti, C., Nesi, M., Fazio, M., Righetti, P. G., *Electrophoresis* 1994, *15*, 177–186.

[60] SimoAlfonso, E., Gelfi, C., Sebastiano, R., Citterio, A., Righetti, P. G., *Electrophoresis* 1996, *17*, 723–731.

[61] Fung, E. N., Yeung, E. S., *Anal. Chem.* 1995, *67*, 1913–1919.

[62] Chang, H. T., Yeung, E. S., *J. Chromatogr. B* 1995, *669*, 113–123.

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

[64] Preisler, J., Yeung, E. S., *Anal. Chem.* 1996, *68*, 2885–2889.

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

[66] Madabhushi, R. S., Menchen, S. M., Efcavitch, J. W., Grossman, P. D., *US Patent No. 5,567,292,* 1996.

[67] Gao, Q., Yeung, E. S., *Anal. Chem.* 1998, *70*, 1382–1388.

[68] Menchen, S., Johnson, B., Winnik, M. A., Xu, B., *Electrophoresis* 1996, *17*, 1451–1459.

[69] Bashkin, J., Marsh, M., Barker, D., Johnston, R., *Appl. Theor. Electrophor.* 1996, *6*, 23–28.

[70] Dolník, V., Gurske, W. A., *Electrophoresis* 1999, *20*, 3373–3380.

[71] Kheterpal, I., Mathies, R. A., *Anal. Chem.* 1999, *71*, 31A–37A.

[72] SimoAlfonso, E., Gelfi, C., Sebastiano, R., Citterio, A., Righetti, P. G., *Electrophoresis* 1996, *17*, 732–737.

[73] Gelfi, C., SimoAlfonso, E., Sebastiano, R., Citterio, A., Righetti, P. G., *Electrophoresis* 1996, *17*, 738–743.

[74] Sassi, A. P., Barron, A., AlonsoAmigo, M. G., Hion, D. Y., Yu, J. S., Soane, D. S., Hooper, H. H., *Electrophoresis* 1996, *17*, 1460–1469.

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

[76] Braun, B., Blanch, H. W., Prausnitz, J. M., *Electrophoresis* 1994, *15*, 1994–1997.

[77] Gelfi, C., Perego, M., Libbra, F., Righetti, P. G., *Electrophoresis* 1996, *17*, 1342–1347.

[78] Chiari, M., Nesi, M., Righetti, P. G., *Electrophoresis* 1994, *15*, 616–622.

[79] Chiari, M., Nesi, M., Sandoval, J. E., Pesek, J. J., *J. Chromatogr. A* 1995, *717*, 1–13.

[80] Chiari, M., Damin, F., Melis, A., Consonni, R., *Electrophoresis* 1998, *19*, 3154–3159.

[81] Izumi, T., Yamaguchi, M., Yoneda, K., Isobe, T., Okuyama, T., Shinoda, T., *J. Chromatogr.* 1993, *652*, 41–46.

[82] Ganzler, K., Greve, K. S., Cohen, A. S., Karger, B. L., Guttman, A., Cooke, N. C., *Anal. Chem.* 1992, *64*, 2665–2671.

[83] Boček, P., Chrambach, A., *Electrophoresis* 1992, *13*, 31–34.

[84] Baba, Y., Ishimura, N., Samata, K., Tsuhako, M., *J. Chromatogr.* 1993, *653*, 329–335.

[85] Mitnik, L., Salome, L., Viovy, J. L., Heller, C., *J. Chromatogr. A* 1995, *710*, 309–321.

[86] Shihabi, Z. K., *J. Chromatogr. A* 1999, *853*, 349–354.

[87] Han, F., Huynh, B. H., Ma, Y., Lin, B., *Anal. Chem.* 1999, *71*, 2385–2389.

[88] Siles, B. A., Anderson, D. E., Buchanan, N. S., Warder, M. F., *Electrophoresis* 1997, *18*, 1980–1989.

[89] Barron, A. E., Sunada, W. M., Blanch, H. W., *Electrophoresis* 1994, *15*, 64–74.

[90] Liang, D. H., Song, L. G., Zhou, S. Q., Zaitsev, V. S., Chu, B., *Electrophoresis* 1999, *20*, 2856–2863.

[91] Liang, D. H., Zhou, S. Q., Song, L. G., Zaitsev, V. S., Chu, B., *Macromolecules* 1999, *32*, 6326–6332.

[92] Rill, R. L., Locke, B. R., Liu, Y., Van Winkle, D. H., *Proc. Natl. Acad. Sci. USA* 1998, *95*, 1534–1539.

[93] Wu, C. H., Liu, T. B., Chu, B., *Electrophoresis* 1998, *19*, 231–241.

[94] Wu, C., Liu, T., Chu, B., Schneider, D., Graziano, V., *Macromolecules* 1997, *30*, 4574–4583.

[95] Liang, D. H., Chu, B., *Electrophoresis* 1998, *19*, 2447–2453.

[96] Rill, R. L., Liu, Y., Van Winkle, D. H., Locke, B. R., *J. Chromatogr A* 1998, *817*, 287–295.