Ultrafast DNA sequencing on a microchip by a hybrid separation mechanism that gives 600 bases in 6.5 minutes

Christopher P. Fredlake*, Daniel G. Hert*, Cheuk-Wai Kan*, Thomas N. Chiesl*, Brian E. Root†, Ryan E. Forster†, and Annelise E. Barron*‡§

Departments of *Chemical and Biological Engineering, ‡Chemistry, and †Materials Science and Engineering, Northwestern University, Evanston, IL 60208

Edited by Robert H. Austin, Princeton University, Princeton, NJ, and approved December 3, 2007 (received for review June 1, 2007)

To realize the immense potential of large-scale genomic sequencing after the completion of the second human genome (Venter's), the costs for the complete sequencing of additional genomes must be dramatically reduced. Among the technologies being developed to reduce sequencing costs, microchip electrophoresis is the only new technology ready to produce the long reads most suitable for the *de novo* sequencing and assembly of large and complex genomes. Compared with the current paradigm of capillary electrophoresis, microchip systems promise to reduce sequencing costs dramatically by increasing throughput, reducing reagent consumption, and integrating the many steps of the sequencing pipeline onto a single platform. Although capillary-based systems require ≈70 min to deliver ≈650 bases of contiguous sequence, we report sequencing up to 600 bases in just 6.5 min by microchip electrophoresis with a unique polymer matrix/adsorbed polymer wall coating combination. This represents a two-thirds reduction in sequencing time over any previously published chip sequencing result, with comparable read length and sequence quality. We hypothesize that these ultrafast long reads on chips can be achieved because the combined polymer system engenders a recently discovered ''hybrid'' mechanism of DNA electromigration, in which DNA molecules alternate rapidly between reptating through the intact polymer network and disrupting network entanglements to drag polymers through the solution, similar to dsDNA dynamics we observe in single-molecule DNA imaging studies. Most importantly, these results reveal the surprisingly powerful ability of microchip electrophoresis to provide ultrafast Sanger sequencing, which will translate to increased system throughput and reduced costs.

DNA separation mechanism | microchip electrophoresis | entangled polymer solution | DNA imaging

The availability of a high-accuracy human genome sequence (1, 2) provides scientists and clinical researchers with an invaluable tool for discovering the underlying causes of genetically based diseases and for developing new treatments and cures. Because the cost of determining the complete sequence of a human-sized genome (6 billion base pairs counting both sets of chromosomes) now stands at approximately $20 million (NIH News Release, 10 October 2004, http://genome.gov/12513210), sequencing costs must be reduced by two orders of magnitude for projects such as the National Institutes of Health's Cancer Genome Atlas Project (Report of the Working Group on Biomedical Technology. February 2005, http://www.genome.gov/15015123). In addition, other scientific projects will benefit from reductions in cost for high-accuracy long-read DNA sequencing. For example, repetitive DNA elements in the human genome known as segmental duplications, which can be clearly revealed by highly accurate genome assembly, have been shown to influence an individual's susceptibility to disease (3) and to have contributed considerably to speciation events in human evolution (4).

Significant advances in sequencing technology must be achieved to allow the required cost reductions for future genome sequencing projects, similar to the development of capillary array electrophoresis (CAE), which spurred the early completion of the Human Genome Project. In moving from the slab gels originally used in DNA sequencing to capillaries, the separation medium was changed from cross-linked gels to linear polymer solutions. Because temperature is better controlled in capillaries than in slab gels, Joule heating effects are minimized so that higher electric fields can be used, resulting in faster separations. Additionally, the use of replaceable (fluid) linear polymer solutions allowed sequencing instruments to be automated, further increasing throughput. Highly optimized CAE systems, run under ideal conditions with ideal samples, are capable of providing up to 1,300-base reads in 2 hours (5) or 1,000-base reads in 1 hour (6), although 500–700 bases in 1–2 hours are more typical for commercial instruments with "real" (i.e., not highly purified or enriched) DNA samples. Miniaturization of electrophoresis to microchip platforms promises to further reduce sequencing times and costs. Through the use of a cross-injection scheme (7), DNA separation efficiency on microchips is vastly increased and requires much shorter separation distances. Additionally, the creation of totally integrated sequencing systems on microfluidic devices promises to reduce time and costs associated with off-line sample preparation (8, 9). The change from multiple instruments for performing sample preparation and analysis to a single integrated system is a great opportunity for cost savings in terms of both capital costs and total analysis time. Indeed, integrated sample preparation and analysis systems are being developed for DNA sequencing (8, 9) and other bioanalytical applications (10, 11). A recent paper has reported an integrated microchip-based sequencer that uses an "in-line injection" scheme to introduce a very small amount of sample in an ultranarrow injection zone, followed immediately by sequencing fragment separation and detection (12).

Alternative approaches to DNA sequencing other than the Sanger method (13) are also being explored. Recently, 454 Life Sciences demonstrated a nonelectrophoretic method of "pyrosequencing" (14). Individual read lengths in this report averaged ≈100 bases, a much lower read length than the 600- to 700-base reads that are needed for existing algorithms to assemble large repeat-rich genomes (15). Although longer average reads can be achieved in some cases (up to ≈230 bases), 454's raw sequences tend to be substantially less accurate than those of electrophoresis-

Author contributions: C.P.F., C.-W.K., T.C., and A.E.B. designed research; C.P.F., B.E.R., and R.E.F. performed research; C.P.F., D.G.H., C.-W.K., T.C., and R.E.F. analyzed data; and C.P.F. and A.E.B. wrote the paper.

This article is a PNAS Direct Submission.

§To whom correspondence should be addressed at the present address: Stanford University, Department of Bioengineering, W300B James H. Clark Center, 318 Campus Drive, Stanford, CA 94305. E-mail: aebarron@stanford.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0705093105/DC1.

based methods. Although the throughput for this instrument is extremely high because of the highly parallel nature of the device [25 million bases of raw sequencing data can be collected in 4 hours (14)], short individual read lengths so far limit the usefulness of this technology to smaller bacterial and viral genomes (<2 Mbp) that have very low, if any, repetitive DNA sequence content (16). For *de novo* sequencing of large repeat-rich genomes, microchannel electrophoresis methods, especially those based on integrated devices, will strongly contend to be the method of choice.

Woolley and Mathies (17) were the first to demonstrate four-color Sanger sequencing on a microfluidic chip. Since then, many other papers reporting four-color DNA sequencing in glass and plastic chips have been published (10, 20–24), with Salas-Solano *et al.* (19) achieving 580 base reads in 18 min and 640 base reads in 30 min in single-channel fused silica glass chips. In these studies, read lengths are reported either with a certain accuracy when obtained read lengths are compared with the known DNA sequence (usually using a cutoff accuracy of 98.5%) or by using a quality assessment program called *phred*, which assigns a quality score based on peak shapes, where a called base with a *phred* score of 20 (Q20) has a 99% probability of being correct. To demonstrate the potential for increased throughput of microfluidic sequencing devices, Liu *et al.* (20) sequenced DNA in a 16-channel glass chip, obtaining 543 bases per channel in 15 min (20). Similarly, Paegel *et al.* (9) obtained read lengths of 430 bases per channel in 96 parallel lanes, in 24 min.

Here, we report a microfluidic chip-based system that provides ultrafast DNA sequencing, providing read lengths up to 600 bases in 6.5 min using a high-performance polymer system comprised of both a separation matrix and a physically adsorbed microchannel wall coating. Previously published microchip sequencing results have exclusively used linear polyacrylamide (LPA) as the sequencing matrix as well as a covalently bonded LPA coating for the channel walls, with analysis times in these systems ranging from 15 to 35 min. In this study, the properties of the acrylamide derivatives we use enable a two-thirds decrease in sequencing time while maintaining comparable read lengths. In addition, the polymer wall coating we used is physically adsorbed to the channel wall to reduce both the electroosmotic flow (EOF) from charged glass surfaces and interactions of the DNA sequencing fragments with channel surfaces. No covalent chemistry is required to attach the coating, which makes it easier to use. Furthermore, we investigate the dynamics of DNA migration in entangled polymer networks to gain a better understanding of the mechanisms behind these ultrafast separations. Based upon an analysis of the electrophoretic mobilities of DNA fragments in the polymer matrices, as well as observations of DNA migrating through the polymer network at the single-molecule level, we propose a hybrid separation mechanism within these matrices that yields faster migration and extremely narrow DNA peaks and accounts for the tremendous speed increase we obtain. Our results show that microfluidic chip electrophoresis is an extremely efficient way to separate DNA sequencing fragments and, moreover, that the ultimate throughput and read length limits of chip platforms have yet to be reached.

Results

Polymer Properties and DNA Ladder Separations. For accurate long-read DNA sequencing, hydrophilic high-molar mass polymers are needed to form a highly entangled polymer network that can efficiently separate the wide range of sequencing fragment sizes (DNA size ranges from 20–1,000 bases in a sequencing sample) (23). Poly(*N,N*-dimethylacrylamide) (pDMA) and poly(*N*-hydroxyethylacrylamide) (pHEA) polymers used in this study were synthesized by aqueous-phase solution polymerization targeting molar masses in excess of 1×10^6 Da, which is typically the threshold for achieving longer sequencing reads and for forming stable adsorbed coatings (24). In some cases, the chain transfer agent isopropanol was added to DMA polymerization reactions to target lower molar mass polymers used for preparing matrices with molar mass "blends" of pDMA. Polymers purified (by dialysis) of unreacted monomer, initiator, and low-molar mass chains were characterized for molar mass and radius of gyration distributions using tandem GPC-MALLS (25). These properties are listed in Table 1 for all polymers used in this study. Zero-shear viscosities of the pDMA matrices used in this report range from 1,000 to 10,000 cP, compared with ≈100,000 cP for the LPA-based matrices (26) [see supporting information (SI)].

Table 1. Properties of pDMA and pHEA polymers synthesized for microchip sequencing and channel surface coating

Polymer	M_r, MDa	R_g, nm	PDI
pDMA	3.4	125	1.6
pDMA*	0.28	31	1.9
pHEA	4.0	134	3.2

R_g, radius of gyration; PDI, polydispersity index.
*Five milliliters of isopropanol was added to the reaction.

DNA Sequencing Results. Although pDMA has been explored as an effective sequencing matrix in CAE instruments (27, 28), this polymer has not been used for sequencing in microchip devices. Sequencing of an M13 standard was carried out in entangled pDMA matrices at 3–5% (wt/vol) concentration in pHEA-coated microchips, and raw fluorescence data were processed by the basecaller from NNIM. Both the average read length (for three runs) and the longest read length obtained are shown in Table 2 (both reported at 98.5% accuracy in comparison with the known DNA sequence). Clearly, the 4% pDMA matrix represents the optimal polymer concentration for long-read sequencing. Resolution of larger DNA fragments is higher in the 4% than in the 5% matrix, and this results in average read lengths that are 20 bases longer for the 4% matrix, whereas the 3% matrix performed much more poorly than the higher concentrations. At this lower concentration, resolution of DNA fragments shorter than 100 bases was much poorer than at higher polymer concentrations, a phenomenon that has been seen in LPA matrices during CAE separations (6). Thus, the 4% pDMA matrix gives the best resolution of both smaller and larger fragments and is the best of the three sequencing matrices. Table 2 also reports sequencing times in these matrices, which are up to 66% lower than in any other chip sequencing report published (8, 18–21, 29). Because most microchip-based DNA sequencing studies have used longer channels as well as different electric field strengths and temperatures in efforts to optimize each individual system, it is difficult to compare migration times in our system directly with their published sequencing times. However, all LPA matrices used previously for microchip sequencing have resulted in much longer separation times.

In separations of a 25-base ssDNA ladder (SI), the 3% pDMA matrix separates large DNA fragments quickly and with high resolution but does not provide sequencing read lengths comparable to the higher pDMA concentrations. Because higher polymer concentration is the most important parameter for single-base

Table 2. DNA sequencing results for high molar mass pDMA matrices with a pHEA dynamic coating

Sequencing matrix	Average read length* (n = 3)	Longest read length*	Time, min
3% pDMA	349 ± 40	377	3.8
4% pDMA	512 ± 54	550	5.7
5% pDMA	489 ± 69	530	6.5

All runs were conducted at 50°C and 235 V/cm (≈3 μA current).
*At 98.5% accuracy.

Table 3. Sequencing comparison of blended molar mass pDMA matrices using pHEA coatings

Sequencing matrix*	Average read length† (n = 3)	Longest read length†	Time, min
3% high M_r pDMA + 1% low M_r pDMA	560 ± 29	587	6
3% high M_r pDMA + 2% low M_r pDMA	542 ± 37	601	6.5

Run conditions were identical to Table 2.
*High molar mass is 3.4 MDa; low molar mass is 240 kDa.
†Read length at 98.5% accuracy.

resolution of short DNA fragments, and lower concentrations (larger mesh size) favor separating larger fragments, polymer matrices were formulated by blending polymers with high- and low-average molar mass. Table 3 shows the sequencing results obtained in two different blends of pDMA polymers (at 98.5% basecalling accuracy). The pDMA with high molar mass (3.4 MDa) was used at a concentration of 3% (wt/vol), whereas the lower molar mass pDMA (240 kDa) was used at either 1% or 2% (wt/vol), so that the total pDMA concentrations for the two matrices were 4% and 5% (wt/vol), respectively. Although the matrices with a single average molar mass perform very well, the mixed molar mass matrices perform even better. The 4% blended matrix gave the highest average read length at 560 bases (with a long read of 587 bases in 6 min), whereas the 5% blended matrix achieved the longest individual read at 601 bases, requiring only 6.5 min to achieve this long read. The four-color sequencing electropherogram for that longest sequencing run is presented in Fig. 1.

Interestingly, these pDMA matrices yield considerably longer sequencing read lengths on a microchip than a commercially available LPA matrix from Amersham, which delivers <300 bases under the same electrophoresis conditions as the pDMA matrices. (The LPA matrix can deliver extremely long reads in a CAE system.) Also, DNA fragments in the 4% mixed molar mass pDMA matrix show greatly reduced band-broadening on a chip relative to the LPA sequencing matrix, as shown in Fig. 2. Here, the peak widths, measured in units of time, of DNA fragments are normalized by the elution times of the fragments, to account for differences in DNA mobilities between the two matrices. In LPA, the peaks become broader for DNA molecules larger than ≈250 bases, corresponding well to the DNA size where resolution drops dramatically and basecalling becomes less accurate. No such transition

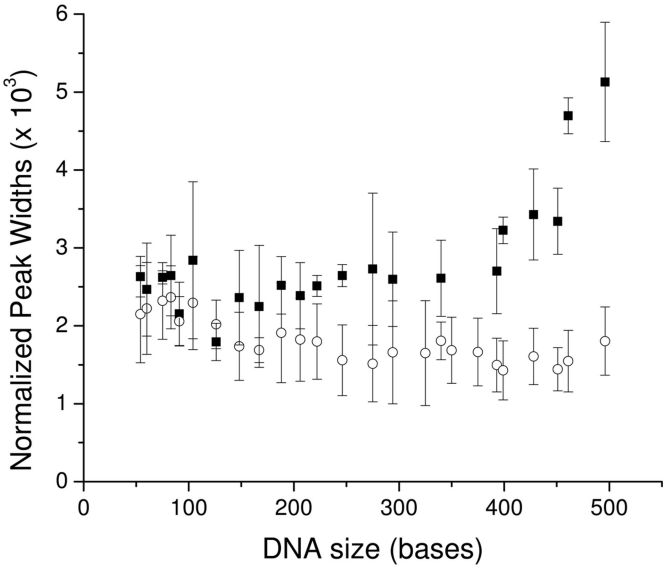

Fig. 2. Analysis of normalized peak widths from T-terminated fragments of M13 sequencing sample moving through LongRead (closed squares) and 4% mixed molar mass pDMA (open circles). Peak widths were measured in units of time and normalized by the elution time of the fragment from the microchannel.

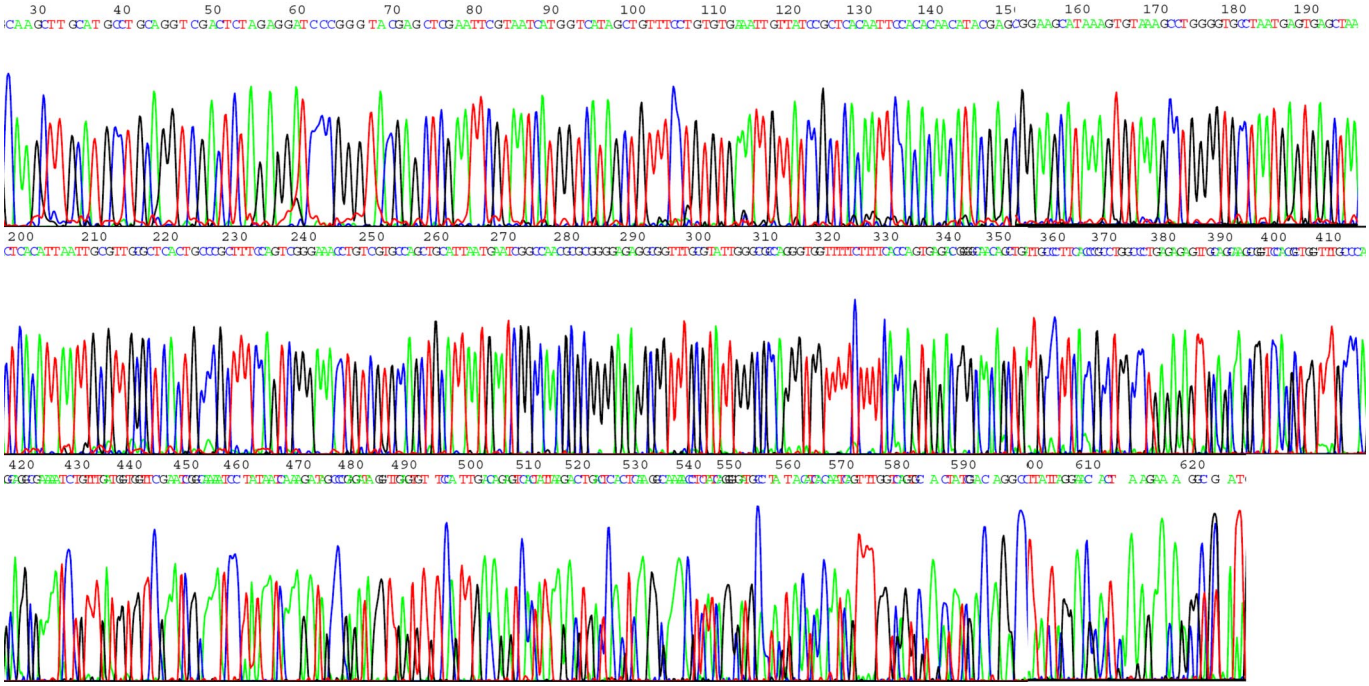

Fig. 1. Electropherogram of the 601-base read (at 98.5% accuracy) in 5% mixed molar mass pDMA matrix (3% 3.4 MDa pDMA + 2% 240 kDa pDMA). Other conditions: electric field, 235 V/cm; temperature, 50°C; current, ≈3 μA; buffer, 1× TTE + 7 M urea; effective channel length, 7.5 cm; injector, 100-μm offset; basecalling, NNIM Basecaller and Sequencer, Ver. 4.0.5.

in peak width is seen in pDMA. To our knowledge, such a significant difference in the band-broadening behavior of DNA in different sequencing matrices is a previously undescribed observation. In part, it is the ability of these pDMA matrices to provide reduced peak widths that leads to the faster, more efficient DNA separations we observe.

Discussion

Chemical and Physical Properties of the Polymer System Beneficial to Microchip-Based DNA Sequencers. Polymer matrices that have low viscosities are much easier to load into microchannels at pressures compatible with thermally bonded chips (these chips fail at matrix filling pressures >200 psi). Although highly viscous polymer solutions sometimes can be loaded into chips manually with a syringe, lower-viscosity matrices greatly facilitate automated matrix filling, which will be required of commercial chip electrophoresis systems. Additionally, polymers that physically adsorb to channel walls to eliminate EOF and reduce DNA–wall interactions reduce costs associated with producing the covalently bonded wall coatings prevalent in most microchip sequencing systems demonstrated to date. Covalently bonded coatings can take up to 10 h to synthesize, often produce inhomogeneous surfaces, and can clog channels, rendering the chips unusable. Adsorbed coatings, on the other hand, can be formed on the channel surface in 15–30 min, are reproducible and stable, and result in almost 100% yield of well coated channels with no clogging.

DNA Separation Mechanisms. Theory on electrophoretic DNA migration in entangled polymer networks predicts that the separation can occur via two distinct mechanisms. Smaller DNA fragments ($<\approx$200 bases) tend to be "sieved" through the polymer network by a "percolation" mechanism similar to that postulated by Ogston for spheres migrating although a dense network of fibers (30, 31). DNA coils that cannot fit into and through the dynamic "pores" of the polymer network must migrate by uncoiling and moving end-on by a mechanism related to polymer reptation, with a bias in the field direction (32).

Electrophoretically driven DNA reptation through polymer networks has been described theoretically by biased reptation models (33–35). These models predict that at low electric field strength, DNA fragments migrate through the network with a snake-like undulating motion, following a virtual "tube" formed by interlinked matrix polymer chains, and with mobilities inversely proportional to the DNA size in bases ($\mu_{DNA} \sim 1/N$). For larger DNA fragments, the mobility ceases to depend on DNA fragment length, and separation is no longer possible, because reptating DNA becomes too strongly oriented with the electric field (36). This "critical DNA size" depends on both the polymer concentration and the electric field strength.

A log-log plot of DNA electrophoretic mobility vs. fragment size (Fig. 3) can be used to interpret such dependencies. The separation of DNA is achieved in portions of the plot where the mobility changes with DNA size. For small DNA fragment sizes, the relationship on the log-log scale is nonlinear, and in gels and some separations by CE, this region is referred to as the Ogston-like sieving regime. For larger DNA sizes (to the right of the first left-most dashed line in Fig. 3), the plot becomes linear, and this regime is usually referred to as the "unoriented biased reptation regime" (37). The second dashed line marks a transition to the regime of "oriented biased reptation" and the ensuing loss of size-based separation for large DNA molecules. Data in Fig. 3 qualitatively agree with the general assumptions of reptation theory, because the shape of the plot is similar, but it is difficult to make quantitative comparisons because the theory predicts a slope of -1 in the unoriented biased reptation regime at zero or very low electric field, whereas slopes of our plots (for pDMA networks) range from -0.45 to -0.60 in that central linear regime. Although the assumptions of a zero or low field contribution are not valid for

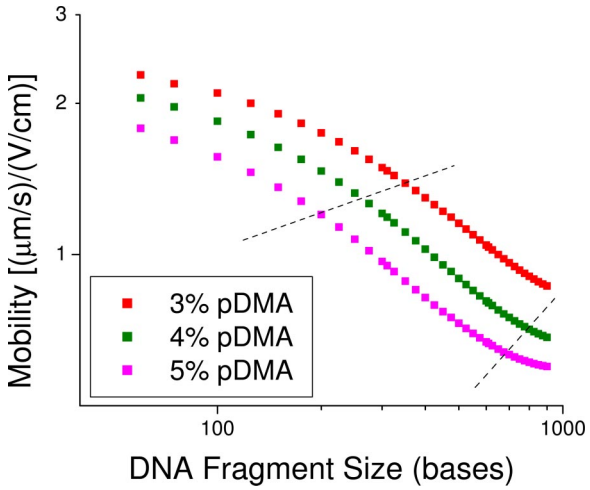

Fig. 3. Mobility data for ssDNA ladder in the pDMA matrices. Conditions are identical to Fig. 1. The dashed lines mark the DNA sizes where the plot is linear.

our experimental data, other mechanisms in addition to reptation may also provide size-based separation, while maintaining trends similar to those of Fig. 3. Indeed, DNA separations in ultradilute polymer solutions, where reptation theory is certainly not valid, showed qualitatively similar shapes in plots of log μ vs. log DNA size (38, 39). Therefore, one cannot conclude that the traditional mechanisms of gel electrophoresis can describe our system, especially given the greatly increased speed and quality of our sequencing separations.

For long DNA sequencing reads, the reptation mechanism (without strong molecular orientation with the electric field) is preferred for large DNA sizes, because separation conditions that tend to orient the DNA fragments result in shorter read lengths. In Fig. 3, the transition to oriented reptation begins at smaller DNA sizes for the 5% matrix relative to the 4% network, so read lengths for the 5% matrix might be expected to be lower. The transition to molecular orientation of DNA is a theoretical limit, however, and does not consider loss of resolution by band broadening. In any practical sequencing matrix, read lengths are lower than this theoretical limit, because of broadening of the detected peaks. Yet, differences in detected peak widths, such as shown in Fig. 2, can strongly affect the sequencing performance of the matrix. Studies are currently underway to investigate the underlying reasons for the narrower peaks in our matrices, especially for larger DNA fragments. We believe that these narrower bands are related to different DNA electromigration dynamics in the different networks.

Hypothesis for a "Hybrid" Separation Mechanism for DNA Sequencing in pDMA Networks. An important factor influencing the separation performance for larger DNA sizes is the strength of the entanglements of the polymer network. Generally, at a given molar mass, the strength of the entangled network increases as the polymer concentration is increased, reflecting the degree of chain overlap and the physical interactions of the individual polymer chains. LPA is more hydrophilic than pDMA and is better able to entangle in solution (at the same molar mass and concentration) (40). So in general, LPA-based entangled networks are physically stronger than pDMA networks, as reflected in their much higher zero-shear viscosities (40).

Strongly entangled networks separate DNA by reptation, and disruptions of the polymer network structure by migrating DNA would be expected to reduce separating power and, thereby, reduce read lengths (41). In weakly entangled networks, disruptions by electromigrating DNA are more frequent (41), especially by larger DNA, and more hydrophobic matrices cannot form robust enough

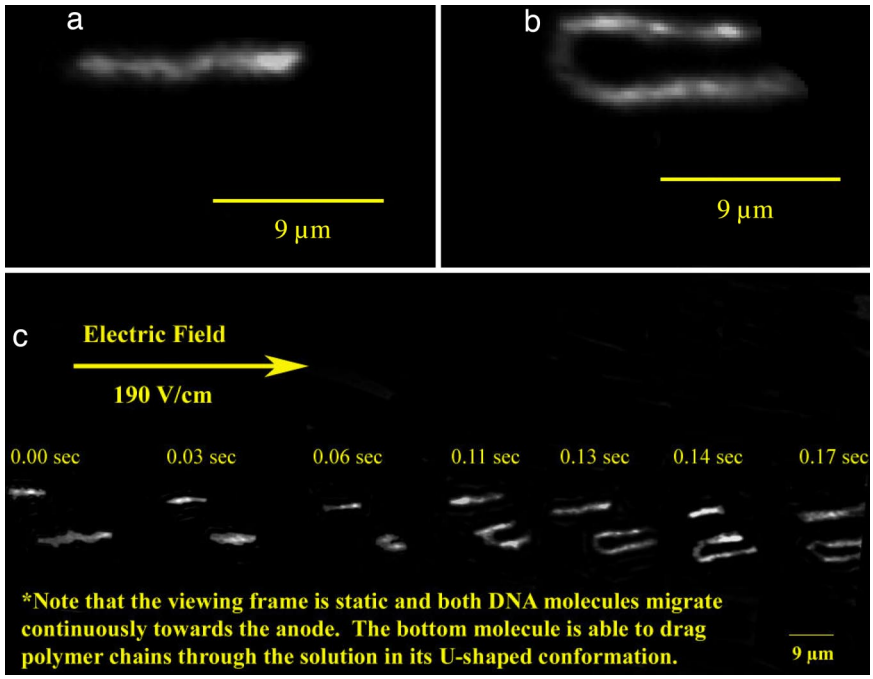

Fig. 4. Images captured from DNA imaging videos. (*a*) λ-DNA is reptating through the polymer network. (*b*) This is an image of a λ-DNA molecule that has hooked around the polymer matrix in a U-shaped conformation and is dragging the disentangled matrix polymers. (*c*) A series of frames at the shown time intervals show two DNA molecules moving through the network (same molecules in *a* and *b*). The top molecule reptates through the entire viewing frame in the given time. The lower molecule is reptating at first and then hooks and drags the polymer network through the viewing frame.

networks to provide long read lengths (40). However, the phenomenon of DNA molecules entangling with and dragging polymer chains and causing the network to be broken apart still imparts a size-based mobility to the DNA, in a mechanism related to transient entanglement coupling (TEC) as discovered by Barron *et al.* (38, 39) for the separation of dsDNA in dilute polymer solutions by CE. We propose that although unoriented reptation is the dominant mechanism of DNA separation in these pDMA matrices, network disruption and DNA-polymer chain dragging result in a greatly increased speed of migration and also contribute significantly to the separation.

We looked for additional, circumstantial evidence of this proposed hybrid mechanism of DNA separation (reptation plus polymer chain dragging) using single-molecule imaging of electrophoresing dsDNA in pDMA matrices. Fig. 4 shows the time evolution of the dynamic chain configurations of two electromigrating, fluorescently labeled, dsDNA molecules in a 3% pDMA matrix at room temperature. One molecule is migrating by reptation (with molecular conformation shown in Fig. 4*a*), whereas the other molecule has broken the network and is dragging pDMA polymer chains through the solution (with molecular conformation shown in Fig. 4*b*). The lower DNA molecule in Fig. 4*c* is initially reptating through the network and then disrupts the physical interactions of the polymer chains by entangling and dragging a small portion of the network along whereas the nearby upper DNA molecule continues to reptate for the duration of the frame series.

A recent paper from our group (41) quantitatively demonstrates that λ-DNA migration by either reptation or TEC depends strongly on the extent of polymer matrix entanglement. Although such a large (48.5-kbp) dsDNA molecule labeled with an intercalating fluorescent dye is clearly an imperfect model for shorter ssDNA molecules, there is still some value in the use of dsDNA to probe polymer network structure. The dsDNA is a stiff polymer with a persistence length of ≈50 nm (44), which corresponds to ≈150 bp, whereas the ssDNA that is separated for sequencing is flexible on a much smaller scale with a persistence length of ≈5 nm (43) or ≈15 bases. The flexibility of the molecule is an important parameter for the interaction of the DNA molecules and the polymer chains in the network during migration. Thus, in terms of the number of Kuhn lengths, a ssDNA molecule that is 600 bases long is similar to a dsDNA molecule 6,000 bp in size. Therefore, a λ-DNA roughly approximates the flexibility of ssDNA at the sizes typical for sequencing, and the images in Fig. 4 represent a proof of principle that these two mechanisms can prevail simultaneously, in the same polymer matrix.

Conclusions

We have shown that electrophoretic DNA sequencing of up to 600 contiguous bases can be achieved in 6.5 min on a microfluidic chip using a pDMA matrix with a pHEA dynamic coating, reducing sequencing time by two-thirds over any previously published sequencing study on chips and by one order of magnitude over CAE instruments. Both the high-molar mass pDMA matrix and the pHEA dynamic coating are necessary elements for excellent sequencing performance in chips.

These pDMA matrices benefit from having an intermediate entanglement strength, between a strongly entangled network, which promotes reptation, and a weakly entangled network, which allows physical network disruption with polymer chain dragging by electromigrating DNA. Although network disruptions can lead to lower separation ability, electromigrating DNA can interact with and drag disentangled polymer chains, leading to a migration mechanism that gives size-dependent mobilities. Although we have directly observed dsDNA migrating in this manner, no current theory for DNA separation addresses these mechanisms concurrently.

This microchip system is net yet fully optimized, and analysis of electrophoretic mobility vs. DNA size data suggests that longer read lengths should be possible, for example by minimizing sources of band-broadening in the system. The short (7.5-cm) separation length of the chips is almost certainly an important factor limiting read length. Sequencing on microchips has been demonstrated in

glass microchips with both 11.5- (19) and 15.9-cm (21) channel lengths. For a dispersion-limited system, the minimum resolution where the basecaller can accurately call bases scales with the square root of the channel length (44). Therefore, read lengths should be expected to increase when these matrices are used in chips with longer separation distances. Our data indicate that the transition to molecularly oriented reptation in these matrices may begin near 850 bases, which could alter the dependence of resolution on length, so that somewhat shorter reads may be expected. Nevertheless, the ability of this optimized pDMA/pHEA polymer matrix/coating system to provide long high-quality sequencing reads in unprecedented short times represents a significant step forward in the development of microchip sequencing systems.

These high-performance materials can be used with any integrated chip system and will likely find the greatest use in a microfabricated device that combines sample preparation with multichannel separation, which is the ultimate goal of this field. The immense potential of integration indicates that electrophoresis will continue to be a powerful tool for future genomic technology, especially for certain types of *de novo* DNA sequencing projects aimed at complex repeat-rich genomes.

Methods

Polymer Synthesis and Characterization. Polymers were synthesized by aqueous-phase free-radical polymerization. The DMA monomer (Monomer-Polymer & Dajac Laboratories) was polymerized according to the procedure described in Doherty *et al.* (24). For high molar masses (>1 MDa), only monomer (5% wt/wt), water, and polymerization initiator were added to the solution, whereas for synthesis of lower molar masses (<500 kDa), 5 ml of the chain transfer agent isopropanol was added to the solution. The HEA monomer (Cambrex) was polymerized at the conditions reported in Albargouthi *et al.* (45) with the exception that a 0.5% (wt/wt) monomer solution was used to minimize polymer crosslinking during polymerization. Additionally, LongRead LPA was purchased from Amersham/GE Healthcare and used to compare separation performance with the pDMA polymer matrices.

Polymer molar mass and root mean square radius distributions were determined by tandem gel permeation chromatography (GPC) (Waters)-multiangle laser light scattering (MALLS) (Wyatt Technologies). The detailed procedure is described by Buchholz and Barron (25).

Microchip DNA Sequencing. Analysis of M13 sequencing fragments and ssDNA ET-900 ladder (both from Amersham) was carried out on a custom-built four-color laser induced fluorescence-based sequencing system. The system has been described in detail by Chiesl *et al.* (46). Briefly, this system consists of an electrical subsystem and an optical subsystem along with a temperature-control setup for the microchips.

Single-channel borosilicate glass microchips purchased from Micronit Microfluidics were used for DNA separations. The chips have an offset T injector where the offset is 100 μm and an effective separation distance of 7.5 cm spanning the distance from the injection cross to the detection point. The microchips were coated with a pHEA dynamic coating, as described previously (46).

DNA sequencing and ssDNA separations were carried out in pDMA solutions with concentrations ranging from 3% to 5% (wt/vol) in 1×TTE buffer (49 mM Tris, 49 mM *N*-(Tris(hydroxymethyl)methyl)-3-aminopropanesulfonic acid, and 2 mM EDTA) with 7 M urea (see SI for details on polymer loading into the chip). For each run, a 235-V/cm electric field is applied for 60 seconds before sample injection. The sample was injected for 40 seconds at 400 V/cm. Separation was carried out at 235 V/cm with 150 V/cm back-biasing applied to the sample and sample waste wells to eliminate sample leakage during the separation. The chip was maintained at 50°C by using a programmable heated stage. Basecalling was completed using the NNIM Basecaller (NNIM) and Sequencher v 4.0.5 (Gene Codes).

DNA Imaging. Lambda DNA (Invitrogen) fluorescently labeled with YOYO-1 (Molecular Probes) was visualized during microchannel electrophoresis using a homebuilt system described previously (47). Briefly, the imaging system consists of an inverted epifluorescence microscope outfitted with a ×100-N.A. 1.4 oil immersion microscope objective. DNA fluorescence was achieved using a 100-watt mercury lamp light source focused with a blue light excitation filter cube (460–500 nm). Emitted fluorescence from the DNA was collected with a 0.5-inch CCD camera through a 510-nm long-pass filter and an image intensifier. All videos were captured at 30 frames per sec by using the XCAP-STD software (EPIX). Electrophoresis voltages were achieved by using 9-V batteries in series to produce electric field strengths of ≈100 to 190 V/cm.

ACKNOWLEDGMENTS. We thank Dr. Karl Putz for help with rheology measurements. This publication was made possible by National Institutes of Health Grants 2 R01 HG001970-07, from the National Human Genome Research Institute, and 5 U01 AI0161297-07, from the National Institute of Allergy and Infectious Diseases. Additional support was provided by the National Science Foundation through Northwestern University Nanoscale Science and Engineering Center Grant EEC-0647560 and by the Department of Homeland Security (DHS) under the DHS Scholarship and Fellowship Program.

1. Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W, *et al.* (2001) *Nature* 409:860–921.
2. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, *et al.* (2001) *Science* 291:1304–1354.
3. Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, Nibbs RJ, Freedman BI, Quinones MP, Bamshad MJ, *et al.* (2005) *Science* 307:1434–1440.
4. Cheng Z, Ventura M, She XW, Khaitovich P, Graves T, Osoegawa K, Church D, DeJong P, Wilson RK, Pääbo S, *et al.* (2005) *Nature* 437:88–93.
5. Zhou HH, Miller AW, Sosic Z, Buchholz B, Barron AE, Kotler L, Karger BL (2000) *Anal Chem* 72:1045–1052.
6. Salas-Solano O, Carrilho E, Kotler L, Miller AW, Goetzinger W, Sosic Z, Karger BL (1998) *Anal Chem* 70:3996–4003.
7. Jacobson SC, Hergenroder R, Koutny LB, Ramsey JM (1994) *Anal Chem* 66:1114–1118.
8. Blazej RG, Kumaresan P, Mathies RA (2006) *Proc Natl Acad Sci USA* 103:7240–7245.
9. Paegel BM, Yeung SHI, Mathies RA (2002) *Anal Chem* 74:5092–5098.
10. Skelley AM, Scherer JR, Aubrey AD, Grover WH, Ivester RHC, Ehrenfreund P, Grunthaner FJ, Bada JL, Mathies RA (2005) *Proc Natl Acad Sci USA* 102:1041–1046.
11. Easley CJ, Karlinsey JM, Bienvenue JM, Legendre LA, Roper MG, Feldman SH, Hughes MA, Hewlett EL, Merkel TJ, Ferrance JP, *et al.* (2006) *Proc Natl Acad Sci USA* 103:19272–19277.
12. Blazej RG, Kumaresan P, Cronier SA, Mathies RA (2007) *Anal Chem* 79:4499–4506.
13. Sanger F, Nicklen S, Coulson AR (1977) *Proc Natl Acad Sci USA* 74:5463–5467.
14. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen ZT, *et al.* (2005) *Nature* 437:376–380.
15. Chaisson M, Pevzner P, Tang HX (2004) *Bioinformatics* 20:2067–2074.
16. Rogers YH, Venter JC (2005) *Nature* 437:326–327.
17. Woolley AT, Mathies RA (1995) *Anal Chem* 67:3676–3680.
18. Liu SR, Shi YN, Ja WW, Mathies RA (1999) *Anal Chem* 71:566–573.
19. Salas-Solano O, Schmalzing D, Koutny L, Buonocore S, Adourian A, Matsudaira P, Ehrlich D (2000) *Anal Chem* 72:3129–3137.
20. Liu SR, Ren HJ, Gao QF, Roach DJ, Loder RT, Armstrong TM, Mao QL, Blaga I, Barker DL, Jovanovich SB (2000) *Proc Natl Acad Sci USA* 97:5369–5374.
21. Paegel BM, Emrich CA, Weyemayer GJ, Scherer JR, Mathies RA (2002) *Proc Natl Acad Sci USA* 99:574–579.
22. Shi YN (2006) *Electrophoresis* 27:3703–3711.
23. Albarghouthi MN, Barron AE (2000) *Electrophoresis* 21:4096–4111.
24. Doherty EAS, Berglund KD, Buchholz BA, Kourkine IV, Przybycien TM, Tilton RD, Barron AE (2002) *Electrophoresis* 23:2766–2776.
25. Buchholz BA, Barron AE (2001) *Electrophoresis* 22:4118–4128.
26. Goetzinger W, Kotler L, Carrilho E, Ruiz-Martinez MC, Salas-Solano O, Karger BL (1998) *Electrophoresis* 19:242–248.
27. Chiari M, Riva S, Gelain A, Vitale A, Turati E (1997) *J Chromatogr A* 781:347–355.
28. Madabhushi RS (1998) *Electrophoresis* 19:224–230.
29. Shi YN, Anderson RC (2003) *Electrophoresis* 24:3371–3377.
30. Ogston AG (1958) *Trans Faraday Soc* 54:1754–1757.
31. Lunney J, Chrambach A, Rodbard D (1971) *Anal Biochem* 40:158–173.
32. de Gennes PG (1979) *Scaling Concepts in Polymer Physics* (Cornell Univ Press, Ithaca, NY).
33. Lumpkin OJ, Dejardin P, Zimm BH (1985) *Biopolymers* 24:1573–1593.
34. Slater GW, Noolandi J (1986) *Biopolymers* 25:431–454.
35. Duke T, Viovy JL, Semenov AN (1994) *Biopolymers* 34:239–247.
36. Ueda M, Oana H, Baba Y, Doi M, Yoshikawa K (1998) *Biophys Chem* 71:113–123.
37. Slater GW, Kenward M, McCormick LC, Gauthier MG (2003) *Curr Opin Biotechnol* 14:58–64.
38. Barron AE, Blanch HW, Soane DS (1994) *Electrophoresis* 15:597–615.
39. Barron AE, Soane DS, Blanch HW (1993) *J Chromatogr A* 652:3–16.
40. Albarghouthi M, Buchholz BA, Doherty EAS, Bogdan FM, Zhou HH, Barron AE (2001) *Electrophoresis* 22:737–747.
41. Chiesl TN, Forster RE, Root BE, Larkin M, Barron AE (2007) *Anal Chem* 79:7740–7747.
42. Viovy JL (2000) *Rev Mod Phys* 72:813–872.
43. Tinland B, Pluen A, Sturm J, Weill G (1997) *Macromolecules* 30:5763–5765.
44. Heller C (2000) *Electrophoresis* 21:593–602.
45. Albarghouthi MN, Buchholz BA, Huiberts PJ, Stein TM, Barron AE (2002) *Electrophoresis* 23:1429–1440.
46. Chiesl TN, Shi W, Barron AE (2005) *Anal Chem* 77:772–779.
47. Chiesl TN, Putz KW, Babu M, Mathias P, Shaikh KA, Goluch ED, Liu C, Barron AE (2006) *Anal Chem* 78:4409–4415.