Conformation, Length, and Speed Measurements of Electrodynamically Stretched DNA in Nanochannels

Christian H. Reccius,* Samuel M. Stavis,† John T. Mannion,* Larry P. Walker,† and H. G. Craighead*
*School of Applied and Engineering Physics and †Department of Biological and Environmental Engineering, Cornell University, Ithaca, New York 14853

ABSTRACT A method is presented to rapidly and precisely measure the conformation, length, speed, and fluorescence intensity of single DNA molecules constrained by a nanochannel. DNA molecules were driven electrophoretically from a nanoslit into a nanochannel to confine and dynamically elongate them beyond their equilibrium length for repeated detection via laser-induced fluorescence spectroscopy. A single-molecule analysis algorithm was developed to analytically model bursts of fluorescence and determine the folding conformation of each stretched molecule. This technique achieved a molecular length resolution of 114 nm and an analysis time of around 20 ms per molecule, which enabled the sensitive investigation of several aspects of the physical behavior of DNA in a nanochannel. λ-bacteriophage DNA was used to study the dependence of stretching on the applied device bias, the effect of conformation on speed, and the amount of DNA fragmentation in the device. A mixture of λ-bacteriophage with the fragments of its own *Hin*dIII digest, a standard DNA ladder, was sized by length as well as by fluorescence intensity, which also allowed the characterization of DNA speed in a nanochannel as a function of length over two and a half orders of magnitude.

INTRODUCTION

The demand for increased analytical ability in the biological sciences has spurred the development of submicrometer- and nanometer-scale structures for single-molecule analysis. These structures facilitate the manipulation and analysis of biological molecules with higher speed and precision than is possible with conventional technology. Such capabilities promise to be useful in applications ranging from genomic sequencing to pathogen detection and in fundamental research in fields such as molecular biology and biophysics.

An assortment of nanoscale structures for enhanced single-molecule analysis has recently been developed, including solid-state nanopores (1,2) entropic trap arrays (3–5), zero mode waveguides (6,7), metallic nanoslit near field scanners (8), and pillar arrays (9–11). Submicrometer- and nanometer-scale fluidic channels used in conjunction with fluorescence spectroscopy have also shown significant potential for the manipulation and analysis of single DNA molecules. Much of this work has focused on the implementation of rapid and sensitive analytical techniques such as fragment sizing (12,13), correlation spectroscopy (14), binding assays (15), identification of nucleic acid engineered labels (16), mobility measurements (17), and polymerase chain reaction analysis (18). The physical behavior and genetic analysis of single DNA molecules in nanochannels are subjects of particular interest as well (10,11,19,20). DNA has been hydrodynamically linearized in submicrometer fluidic devices, which can be used for genomic sequencing when combined with repeated fluorescence detection (21–25). DNA molecules have also been shown to become elongated to an extended equilibrium length when introduced into a nanochannel (26), which has been utilized for restriction mapping (27). Further work has investigated other aspects of the physics of elongated DNA strands in nanochannels, including entropically driven dynamics and compression against nanoscale constrictions (28–30). There has also been a growing interest in using nanochannels (26) and other methods (8,31,32) to analyze single molecules with spatial resolution beyond the optical diffraction limit while benefiting from the advantages of established fluorescence spectroscopy techniques.

In this study, a method is described to quickly and precisely measure the conformation, length, speed, and fluorescence intensity of single DNA molecules constrained by a nanochannel. DNA molecules were driven electrophoretically from a nanoslit into a nanochannel, as illustrated in Fig. 1 *A*, which confined and dynamically elongated the molecules beyond their equilibrium length in a nanochannel. The stretched molecules were then transported through two spatially separated focal volumes, defined by lasers focused sequentially on the nanochannel, which enabled speed and cross correlation measurements (8) and increased statistical validity and experimental throughput. The use of a nanochannel also reduced fluorescent background noise, increased excitation uniformity, and allowed single-molecule detection at higher concentrations when compared to measurements made in larger fluidic channels or free solution (14,15).

Submitted August 30, 2007, and accepted for publication February 12, 2008.

Address reprint requests to H. G. Craighead, Fax: +1 (607) 255-7658; E-mail: hgc1@cornell.edu.

Christian H. Reccius's present address is Philips Research, Redhill, Surrey RH1 5HA, UK.

Samuel M. Stavis's present address is Semiconductor Electronics Division, National Institute of Standards and Technology, Gaithersburg, MD 20899-8120.

Editor: Taekjip Ha.

0006-3495/08/07/273/14 $2.00

doi: 10.1529/biophysj.107.121020

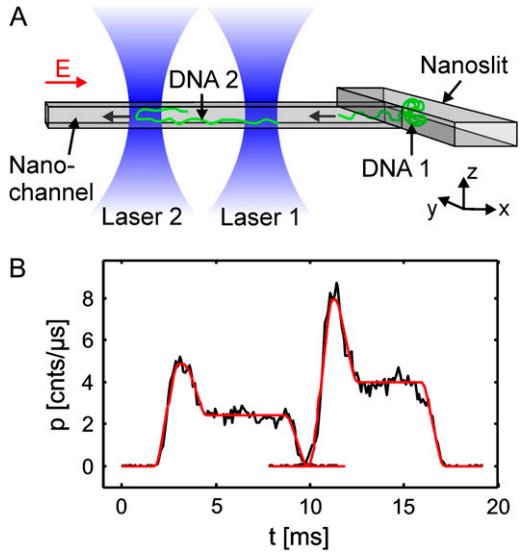

FIGURE 1 Experimental schematic. (*A*) Fluorescently labeled DNA molecules are dynamically elongated by driving them electrophoretically from a nanoslit into a nanofluidic channel (DNA 1). The nanochannel is probed by two sequentially focused lasers. DNA molecules driven through the resulting focal volumes generate two similar fluorescent signals, which are shifted in time relative to each other (DNA 2). These signals can be used to determine single-molecule statistics for the sample (such as speed, length, folding conformation, and fluorescence intensity) and are used in this work to illuminate several aspects of the physics of DNA translocation through nanofluidic channels. (*B*) Two photon count signals p resulting from the two focal volumes and fits versus time t with a 100 μs bin time at a device bias of 50 V. The fits result in an apparent DNA length $l_A = 9.10 \pm 0.05\,\mu m$, speed $v_S = 1315 \pm 3\,\mu m/s$, and folded length/apparent length ratio $l_L/l_A = 22.3 \pm 0.4\%$.

For each DNA molecule detected, photon bursts from the two fluorescent signals were matched and subsequently fit to analytical models describing the conformation, length, speed, and intensity of the DNA strands. An example of the signals and fits resulting from the analysis of a DNA molecule interpreted as folded at the front end is shown in Fig. 1 *B*. This analysis algorithm made possible a direct determination of molecular length and conformation with spatial resolution beyond the optical diffraction limit, established in the work presented here at 114 nm, with an analysis time of 20 ms per molecule. The measurements depicted in Fig. 1 *B* were also used to infer DNA length from total fluorescence burst intensity, a previously demonstrated analytical method (12,13), and the two methods were compared and combined. The confluence of these factors represents an important step forward toward applications that place a premium on spatial resolution and analysis time, such as single-molecule genomic sequencing.

The introduced methods were then used with the well-known system λ-bacteriophage DNA and its derivatives to study DNA fragmentation, folding, and dynamic stretching as a function of applied bias, and speed as a function of apparent length, folding, and contour length. The results of these experiments help elucidate several aspects of the electrophoresis and friction of DNA molecules in nanoscale environments. DNA speed was found to increase slightly with folding, and stretching was found to increase with applied device bias and electric field. For mixtures of DNA fragments, speed was found to be almost constant with length, showing a slight decrease only for short fragments.

THEORY, METHODS, AND MATERIALS

Single-molecule burst theory

The two collected photon count signals are generated by projecting the fluorescent images of individual DNA molecules moving in the nanochannel on two different optical fibers, which are connected to separate avalanche photodiodes. To describe the general signal shape, a DNA molecule with N basepairs and contour length L in a nanochannel with a depth and width D is considered. The strand is homogenously stained with N_f fluorescent dye molecules, giving a dye/basepair ratio $g = N_f/N$, and uniformly stretched to an end-to-end length $l_R = sL$, where s is the stretching factor ($0 \le s \le 1$). Folded molecules, as illustrated in Fig. 1 *A*, are distinguished from unfolded molecules by one end of the strand showing a looped configuration, resulting in a higher fluorophore concentration along the length of the channel.

For an unfolded DNA molecule, the fluorophore concentration c can be written as a function of the position along the nanochannel axis x as

$$c(x) = c_0[\Theta(x - x_0 + l_R/2) - \Theta(x - x_0 - l_R/2)], \quad (1)$$

where $\Theta(x)$ is the Heaviside step function, x_0 is the position of the center of the molecule, and $c_0 = Ng/(D^2 l_R)$ is the average fluorophore concentration inside the channel volume occupied by the DNA strand.

The DNA strand is illuminated by a focused (but not diffraction-limited) laser beam of power P and frequency ν. Assuming a radial Gaussian shape with an $e^{-1/2}$ radius of σ_B, the laser intensity $I(x, y)$ in the focal plane of the objective is

$$I(x, y) = I_0 \exp\left(-\frac{x^2 + y^2}{2\sigma_B^2}\right) \quad (2)$$

with amplitude $I_0 = k_{Ex} P/(2\pi\sigma_B^2)$ and with the factor k_{Ex} used to account for dichroic absorption losses. The resulting fluorescence is collected by an objective of magnification m and imaged by the tube lens of the microscope on a fused silica fiber with diameter d. The investigated volume of the nanochannel is limited in the x direction from $-d/2m$ to $+d/2m$ by the fiber acting as an aperture and in the y and z directions to $-D/2$ to $D/2$ by the nanochannel. As $\sigma_B > d/m$, it is assumed that the DNA is illuminated uniformly by constant laser intensity I_0, and since $d/m \gg D$ the calculation is reduced to one dimension. The number of emitted photons per channel length and time $f(x)$ along the nanochannel x axis is therefore

$$f(x) = f_0[\Theta(x - x_0 + l_R/2) - \Theta(x - x_0 - l_R/2)] \quad (3)$$

with an amplitude $f_0 = c_0 Q\varepsilon D^2 I_0/(h\nu)$, Planck's constant h, quantum yield Q, and the natural molar extinction coefficient of the dye ε. The resulting image at the fiber position is a convolution of $f(x)$ with the intensity point spread function PSF(x) of the microscope objective and the tube lens which is modified to account for the objective magnification m:

$$i(x) = \frac{k_{Em}}{m} \int_{-\infty}^{+\infty} f(x')\mathrm{PSF}(x/m - x')dx' \quad (4)$$

with the factor k_{Em} accounting for the collection efficiency of the objective and absorption losses of the emission filters. Assuming a Gaussian point spread function with an $e^{-1/2}$ radius of σ:

$$\mathrm{PSF}(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad (5)$$

the image can be described by

$$i(x) = i_0\frac{1}{2}\left[erf\left(\frac{x/m - x_0 + l_R/2}{\sqrt{2}\sigma}\right) - erf\left(\frac{x/m - x_0 - l_R/2}{\sqrt{2}\sigma}\right)\right] \quad (6)$$

with an amplitude $i_0 = k_{Em}f_0/m$. In the microscope image plane, the image is focused on an optical multimode fiber and transmitted to the photodiode. Fiber coupling and absorption losses reduce the resulting signal by a factor k_{Fib}. The number of photons per time $p(x_0)$ arriving at the detector is therefore

$$p(x_0) = k_{Fib}\int_{-d/2}^{+d/2} i(x)dx = \frac{p_0 m}{d}\{pt(x_1) - pt(x_2) - [pt(x_3) - pt(x_4)]\} \quad (7)$$

with the signal amplitude

$$p_0 = \frac{k_{Fib}k_{Em}f_0 d}{m} = \frac{k_{Fib}k_{Em}c_0 Q\varepsilon D^2 k_{Ex}Pd}{2\pi\sigma_B^2 mh\nu}, \quad (8)$$

and the function parts ($i = 1,2,3,4$)

$$pt(x_i) = \frac{\sigma}{\sqrt{2\pi}}\exp\left(-\frac{x_i^2}{2\sigma^2}\right) + \frac{x_i}{2}erf\left(\frac{x_i}{\sqrt{2}\sigma}\right), \quad (9)$$

and $x_{1/2} = x_0 - l_R/2 \mp d/(2m)$, $x_{3/4} = x_0 + l_R/2 \mp d/(2m)$.

This simple integration was made assuming that the step-index multimode optical fiber used has a collection efficiency with a pulse function lateral dependence.

The fluorophore concentration of a molecule with a folded front end can be described in a manner similar to the unfolded case. Inside the channel it occupies only an apparent length l_A, but its real end-to-end length l_R can be calculated by adding the length of the folded over part l_L. The fluorophore concentration c_{fold} can be written as

$$c_{fold}(x) = c_0[\Theta(x - x_0 + l_A/2) + \Theta(x - x_0 - l_A/2 + l_L) - 2\Theta(x - x_0 - l_A/2)], \quad (10)$$

which leads to a photodiode signal:

$$p_{fold}(x_0) = \frac{p_0 m}{d}\{pt(x_1) - pt(x_2) + [pt(x_3) - pt(x_4)] - 2[pt(x_5) - pt(x_6)]\} \quad (11)$$

with $x_{1/2} = x_0 - l_A/2 \mp d/(2m)$, $x_{3/4} = x_0 + l_A/2 - l_L \mp d/(2m)$, and $x_{5/6} = x_0 + l_A/2 \mp d/(2m)$.

Higher orders of folding can be derived analogously.

In the described experiments, two lasers are used to define two focal volumes along the length of the nanochannel, and two photodiodes are then used to image the DNA strand at two different positions $x = 0$ and $x = sd$. The time-dependent signals can then be described by $p(v_S(t - t_0))$ and $p(v_S(t - t_0) - sd)$, respectively, with t_0 being the time when the center of the molecule is at $x = 0$ and v_S being the speed of the strand. It is assumed that different apparent DNA lengths (l_{A1}, l_{A2}) and amplitudes (p_{01}, p_{02}) are possible at $x = 0$ and $x = sd$ due to changes in molecular conformation and different optical loss factors.

Cross correlation theory

By analyzing the cross correlation function of the two measured fluorescent signals, fluorescence correlation spectroscopy (33–35) provides time-averaged information about the sample. A modified fluorescence correlation spectroscopy approach is used in this work to yield the average speed and length of DNA strands, which is then used to guide the subsequent single-molecule burst analysis algorithms. The cross correlation function is known to be (36)

$$G_C(\tau) = 1 + \frac{\int\int W_1(x)W_2(x')\rho(x,x',\tau)dxdx'}{\overline{C}^2\int W_1(x)dx\int W_2(x')dx'}, \quad (12)$$

with the molecular fluorescence detection efficiency functions $W_1(x) = W(x)$ and $W_2(x) = W(x - sd)$ of the two focal volumes, the concentration fluctuation function $\rho(x,x',\tau)$, and the mean concentration $\overline{C}$. A sample consisting of unfolded DNA molecules of a single length l_C is assumed. For the velocities v_C used for high-speed analysis, diffusion can be neglected (14), and the concentration fluctuation function becomes $\rho(x,x',\tau) = \overline{C}\delta(x - x' - v_C)$ (37). As the stretched DNA strands are longer than the investigated channel length ($l_C > d/m$), $W(x)$ becomes equal to $p(x)$ from Eq. 7. To analytically solve the integrals in Eq. 12, $W(x)$ is approximated by a modified sigmoidal function which matches the slopes of $p(x)$ at $x = \pm l_C/2$:

$$W(x) = p_0\left[1 - \frac{1}{1 + \exp(-4m/d(|x| - l_C/2))}\right]. \quad (13)$$

The result for $G_C(\tau)$ was derived analytically using Mathematica (Wolfram Research, Champaign, IL) but is cumbersome and is therefore not presented here. As $G_C(\tau)$ varies between experimental runs due to folding and fragmentation of the sample, cross correlation fits were imperfect but adequate to guide the subsequent single-molecule analysis algorithm. It is expected that the quality of cross correlation fits will improve as smaller nanochannels are used to reduce folding of the DNA molecules.

Single-molecule burst analysis

The analysis algorithm for the two photon count signals consisted of the localization of bursts of fluorescence resulting from the detection of DNA molecules, the matching of corresponding bursts in both photon count signals, and the combined fitting of the corresponding bursts. Partner bursts were analyzed simultaneously to enable the most sensitive determination of molecular speed from the time of travel between the two focal volumes, as well as from both individual bursts. Analysis parameters were modified with varying device bias and resulting molecular speed. The following parameter ranges correspond to device biases, U, of $1 - 100$ V. The photon count signals were originally recorded at a bin time of 10 μs and then rebinned to 10–500 μs to improve and accelerate analytic burst fitting. Rebinning was limited to preserve the shape of each fluorescent burst.

All bursts in both signals were localized by the use of a thresholding algorithm. Single-molecule detection thresholds were set at least seven standard deviations above the mean background noise values, which were determined by fitting the overall photon count distributions of the two signals to Poisson distributions. This eliminated virtually all noise from inadvertent analysis. Multiple bursts separated by $<\Delta t_{Gap} = 10 - 0.1$ ms were combined into one burst. The results of the cross correlation analysis were used to find matching bursts. For a burst in the first photon count signal at time t_0, the matching burst in the second signal was defined as the burst closest to the time $t_0 + t_P$, with $t_P = sd/v_C$ being the peak position of the cross correlation curve. A burst was considered for further analysis only if a corresponding burst was found between $t_1 = t_0 + t_P - \Delta t_P$ and $t_2 = t_0 + t_P + \Delta t_P$ with $\Delta t_P = (d + l_C)/v_C$. In the case of a heterogeneous sample, the correlation curves were not fitted, and the values for t_P and Δt_P were chosen manually from the cross correlation curve.

Partner bursts were fitted simultaneously using a modified Levenberg-Marquard algorithm. Six different folding models based on observations of existing burst shapes were used to account for up to three levels of molecular folding. The level of folding was determined by comparing the maximum photon count of the first partner burst to the maximum of the photon count distribution for all bursts in the measurement, neglecting the signal slopes at

the beginning and end of the bursts. This distribution maximum was assumed to be the average photon count resulting from the detection of unfolded DNA strands wholly in the focal volume and was assigned as the amplitude $p_{0,1}$ of the fitting formula for channel 1. The fitting amplitude $p_{0,2}$ for channel 2 was chosen analogously. Fitting amplitudes were fixed to analyze the bursts correctly. It was only in the case of very slow moving molecules and experimental run times exceeding 2 min (corresponding to a device bias of $U = 1 - 2$ V) that individually fitting the burst amplitudes for each molecule gave better results. This accounted for increasing residence times in the nanoslit and a resulting increase in dye loss to the nanoslit surfaces, as well as photobleaching from scattered light near the nanochannel array. If a molecule was assumed to be folded two or three times, two fitting models were compared and the one yielding a lower χ^2 was chosen. The photon count per molecule was determined by summing the counts between the start and end times of the pulse. These times were determined by the fitting algorithm, and not the thresholding algorithm, to increase their accuracy. This analysis was performed using custom software programmed in Labview (National Instruments, Austin, TX), which allowed correction of the automatically chosen model for every fit. A small number of detection events showing overlapping molecules or molecular laser damage were excluded from further analysis in all measurements, except for the one shown in Fig. 7 having over 16,000 molecules. Bursts from overlapping molecules were identified by the superposition of burst shapes of the type shown in Fig. 3 *B*. Photodamage was observed as a single photon burst from an intact molecule in the first focal volume, followed by two photon bursts from fragments in the second focal volume. Occasionally, the second burst remained intact but showed a partial decrease in fluorescence intensity.

Gaussian fitting of logarithmic distributions

In the experimental results shown in Fig. 7, the DNA length varied over 2.6 orders of magnitude, and therefore the color-coded plot in Fig. 7 *D* as well as the distributions in Fig. 7, *A* and *E*, were binned logarithmically. For Fig. 7 *A* it was assumed that each of the real length peaks is Gaussian. Peak *i* can then be described as $f_i(l_R) = a_{Ri} \exp\left[(l_R - l_{Ri})^2/2\sigma_{Ri}^2\right]$ with amplitude a_{Ri}, mean value l_{Ri}, and standard deviation σ_{Ri}. Logarithmic binning with bin size dy of this curve leads to

$$f_{logi}(l_R) = \int_{r_1}^{r_2} f_i(l'_R)dl'_R = \sqrt{\pi/2}a_{Ri}\sigma_{Ri}\left[erf\left(\frac{-10^{-dy/2}l_R + l_{Ri}}{\sqrt{2}\sigma_{Ri}}\right) - erf\left(\frac{-10^{dy/2}l_R + l_{Ri}}{\sqrt{2}\sigma_{Ri}}\right)\right] \quad (14)$$

using the integration boundaries $r_1 = 10^{(\log_{10}(l_R) - dy/2)}$ and $r_2 = 10^{(\log_{10}(l_R) + dy/2)}$. A superposition of nine functions of this type was used to fit the real length distribution in Fig. 7 *A*, and the photon count distribution in Fig. 7 *E* was fit in an analogous manner.

Optical setup

An Olympus IX71 inverted microscope (Olympus, Melville, NY) with a Prior Pro-Scan II stage (Prior Scientific, Rockland, MA) was used in conjunction with nanochannel devices for single-molecule detection. The microscope stage was equipped with integrated position sensors and servo motors to control lateral drift, and drift in the focal plane was measured using a stage height probe (Heidenhain, Traunreut, Germany) and controlled with an external focus servo motor (Prior Scientific). An argon-krypton mixed gas tunable laser (Melles Griot Laser Group, Carlsbad, CA) and a solid-state sapphire laser (Coherent, Santa Clara, CA) were used with Z488/10X excitation filters (Chroma Technology, Rockingham, VT) for fluorescence excitation at 488 nm. A variety of optical elements including mirrors and kinematic mounts (Newport, Irvine, CA) were used to guide the laser beams into the microscope, with their entry angles tuned to yield focal volumes separated along the length of the nanochannel by $11 - 12\,\mu$m. Both lasers were tuned to have a power of $250\,\mu$W, as measured after the circular polarizer, except in the experiment shown in Fig. 7 in which the laser power was 1.5 mW.

A Z488RDC dichroic mirror (Chroma) reflected the laser beams into the back of a UApo/340 water immersion objective (40×, 1.15 NA, Olympus). Collected fluorescence emission passed through a 535AF45 emission filter (Omega Optical, Brattleboro, VT). The signal was divided using a beam splitter (Omega) and focused by the internal tube lens of the microscope on two high power step-index fused silica multimode optical fibers ($50\,\mu$m inner diameter, $125\,\mu$m cladding diameter; OZ Optics Limited, Ottawa, ON, Canada). At the end of the fibers light was detected with two avalanche SPCM-AQR-14-FC photodiodes (Perkin Elmer, Fremont, CA). An overview of the optical setup is shown in Fig. 2 *A*. A Photometrics Cascade 512B camera (not shown in Fig. 2 *A*; Photometrics, Tucson, AZ) was used to align the focused lasers to the nanochannel, visualize the experiment, and image several hundred quantum dots (Qdot 525 streptavidin conjugate, Invitrogen, Carlsbad, CA) adsorbed to a fused silica coverslip for characterization of the point spread function of the microscope objective. The point spread function was described by a Gaussian with a standard deviation of $\sigma = 140 \pm 20$ nm. Further details of the optical setup can be found elsewhere (17).

Device fabrication

Nanochannels were fabricated in fused silica using a combination of electron beam and optical lithography. A 150-nm-thick film of poly(methyl-methacrylate) resist was spun onto a 500-μm-thick substrate wafer (Mark Optics, Santa Ana, CA) followed by an evaporated 25-nm-thick film of gold. A negative device pattern was exposed using a JBX-9300FS electron beam lithography system (JEOL, Peabody, MA). After removal of gold and development of the resist, the nanochannel pattern was transferred to a chromium mask by evaporation and liftoff. A microchannel etch mask was patterned using standard optical lithography procedures and the same liftoff process. An optical micrograph of the resulting chromium mask can be seen in Fig. 2 *B*. The micro- and nanochannels were then simultaneously etched into the substrate using a Plasmalab 80Plus RIE (Oxford Instruments, Eynsham, UK) with a CHF_3/O_2 mixture. Fig. 2, *C* and *D*, presents electron micrographs showing the interface between the microchannel and a nanochannel. Inlet and outlet holes were excised from the substrate by alumina powder blasting through the backside of the wafer. To enclose the channels, a 170-μm-thick fused silica cover wafer (Mark Optics) was bonded and annealed at a temperature of 1050°C to the substrate wafer. After device fabrication, sample reservoirs were glued to the inlet and outlet holes and sealed with reusable adhesive during use. Platinum wires provided electrical contacts with the sample solutions. Further details of the fabrication process can be found elsewhere (30).

Samples

DNA samples were labeled using the bis-intercalating fluorescent dye YOYO-1 (Invitrogen) with a dye/basepair ratio of 1:5. λ-bacteriophage DNA (New England Biolabs, Ispwich, MA) at a concentration of $25\,\mu g/m$L was heated to a temperature of 65°C for 5 min and added to a mixture of YOYO-1 dye in 5× Tris-Borate-EDTA buffer (TBE, pH 8.3, Sigma-Aldrich, St. Louis, MO) with 6% (v/v) β-mercaptoethanol (Sigma-Aldrich) as an antiphotobleaching reagent. A *Hin*dIII digest of λ-bacteriophage DNA (New England Biolabs) was prepared in a similar manner with the exception that it was heated to 60°C. The digest was mixed in equal parts with pure λ-bacteriophage at a total concentration of $50\,\mu g/m$L, with the appropriate heating temperature used for each constituent before mixing. The sample was then labeled with YOYO-1. In this case the buffer also contained 1% (w/w) poly(*n*-vinylpyrrolidone) (PVP, molecular mass 40 kD; Sigma-Aldrich) to

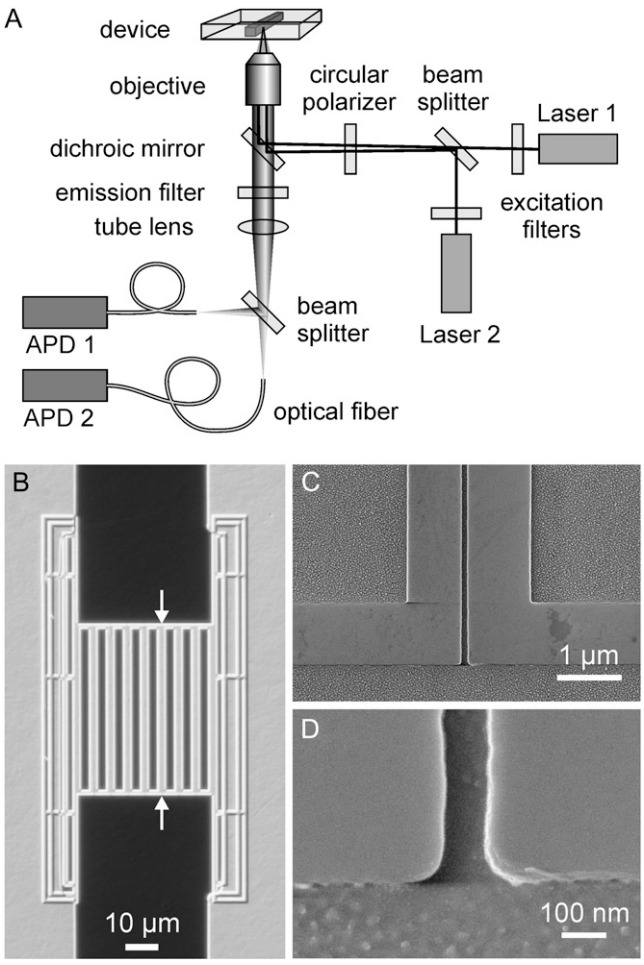

FIGURE 2 Optical setup and devices. (*A*) Two lasers are directed into the back of a microscope objective and focused at different positions along the length of a single nanochannel. The two fluorescent signals are collected, split, and focused on optical fibers coupled to avalanche photodiodes (APDs). (*B*) An optical micrograph of the chromium etch mask shows the structure of the nanochannel device. The white arrows mark the entrances to one of the eight nanochannels in the array. The mask consists of a coarse layer for the nanoslit made with optical lithography and a fine layer for the nanochannels made with electron beam lithography. The electron beam layer was designed to reduce write times and proximity exposure effects. (*C*) An electron micrograph shows the interface between the nanoslit and a nanochannel. The etch depth is 100 nm, and the nanochannels are 90 nm wide. (*D*) An electron micrograph shows the entrance to a nanochannel. The device floor roughness of $10 - 20$ nm is attributed to the etching process.

reduce both electroosmotic flow and nonspecific binding of DNA to channel walls (38–40). The contour length of λ-bacteriophage DNA (48.5 kbp) can be calculated from the basepair spacing of 0.34 nm to be $L_\lambda = 16.5\,\mu$m. Recent studies have shown that the dye TOTO-1, which is similar to YOYO-1, increases the contour length L_λ by 30–35% at a dye/basepair ratio of 1:4 (24,41). L_λ is therefore expected to rise by 23% to 20 μm for a dye ratio of 1:5.

RESULTS AND DISCUSSION

This work has two purposes. The first is to demonstrate a new method to dynamically elongate, rapidly detect, and carefully analyze single DNA molecules in a nanofluidic channel. The second is to use this method to investigate some of the physical phenomena occurring when DNA strands are manipulated in this manner, including speed as a function of folding, stretching and speed as a function of electric field, and speed as a function of length. Several experiments were conducted using variations of the well-known system λ-bacteriophage DNA, including random fragments thereof and controlled fragments from a mixture of λ-bacteriophage DNA with its own *Hin*dIII digest.

Length analysis of single λ-bacteriophage DNA molecules

Each experimental run yielded two photon count signal traces as well as their cross correlation function. Bursts of fluorescence in the two photon count signals were identified by a thresholding algorithm. Corresponding partner bursts were matched by estimating the time window of their occurrence based on preceding fits of the cross correlation function, yielding estimates for the speed v_C and average apparent length l_C of the molecules. The majority of bursts exhibited multiple levels of fluorescence intensity, which were interpreted as the result of folded DNA molecules. The observed burst shapes are similar to the electrical current changes observed during DNA translocation through artificial nanopores (2,42). An example of this phenomenon is shown in Fig. 1 *B*. Almost all molecules exhibited some degree of front end folding, which was also observed in previous experiments involving single DNA molecules in nanochannels (29,30), but to a lesser degree.

The increased difference in the electric field strength between the nanoslit and the nanochannel in the device used here is a likely cause for this increased front end folding. It is hypothesized that the large electric field gradient leads to the random entrance of whatever part of the DNA molecule first comes close enough to the nanochannel, despite the higher entropic forces hindering the entrance of looped structures into the channels. The metaphor of ''sucking spaghetti'' (43) helps explain why folding is rarely observed in the middle or at the end of the molecules, which are stretched out at these locations. There are several other possible explanations for the observed front end folding. For the high speeds used in these experiments, molecules may become folded inside the channel due to hydrodynamic effects or interactions with the channel surface. Elongation and folding caused by interactions with the nanoslit surfaces before entrance into the nanochannel is possible, and interactions with the nanochannel surfaces may also lead to a compression of the leading strand end. None of these alternative explanations were observed at lower speeds (30); and the long persistence length of DNA as well as the existence of quantized burst shapes including 3, 6, and 7 in Fig. 3 makes compression unlikely, which is why the folding interpretation was preferred in this work. A mixture of the mentioned effects is also possible, and

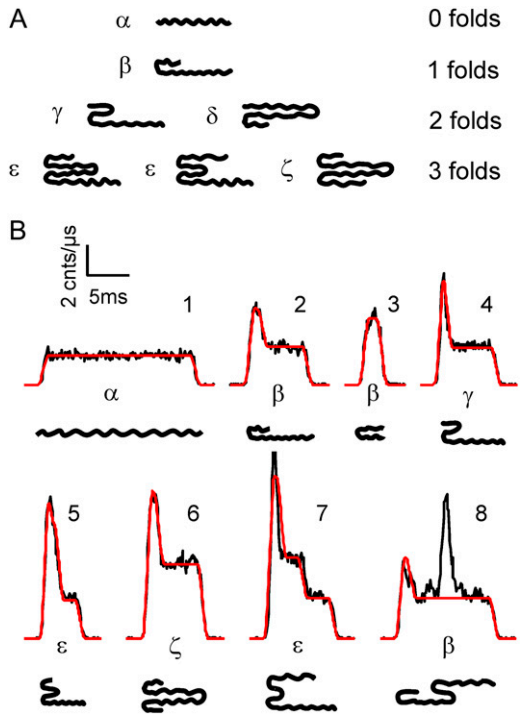

FIGURE 3 Single-molecule analysis. (*A*) Seven theoretical molecular conformations with up to three folds were used to describe electrodynamically stretched DNA molecules in a nanochannel. Six analytical models, denominated $\alpha - \zeta$, were developed from these conformations to fit bursts of fluorescence, with model ε covering two shapes. (*B*) Observed bursts of fluorescence and fits from the first focal volume (signal 1) are shown. Fluorescence bursts are interpreted with the folding model and conformation drawn underneath (fits shown in *red*). Burst 8 shows a rare fold, knot, or overlapping fragment in the middle of the molecule and is fitted using model β. The most frequently observed burst shapes were shapes 2, 4, and 5; the others were rare.

in this case molecules with a short initial peak could be modeled by a modified version of Eq. 10 using a variable initial step height. This would lead to fitting functions with shapes that differ from those used here only by different slopes in the leading signal edge. As analysis of the slopes is currently limited by fluctuations in fluorescence, the results for the real length would be very close to the results here. Moving to smaller nanochannel dimensions in future work will help to further clarify this topic.

Based on these observations, seven theoretical molecular conformations were used to describe up to three levels of front end folding, as shown in Fig. 3 *A*. For two and three folds, respectively, two and three conformations exist which differ in the length of the folded section of a molecule relative to neighboring nonfolded ends, when considering the molecule from right to left. Using these seven molecular conformations, six analytical models denominated $\alpha - \zeta$ were developed to fit measured fluorescence bursts. Differentiation of the fluorescence burst models comes from the mathematical description of the resulting fluorophore concentration profiles along the channel axis. For two folds, the apparent length of the molecules can be dominated either by the single strand end or by the loop, leading to different basic fluorophore concentration levels. The length of the next concentration level is then described as a fraction of the apparent length. For three folds, the apparent length of the molecules can again be dominated either by a single strand end or by the loop, leading to two different models.

The lengths of the two left shapes are both dominated by the single strand end, enabling them to be described by the same model ε. Almost all single-molecule detection events in the work here could be described by these various models. Several examples of fluorescent bursts, corresponding molecular conformations, and analytical fits are shown in Fig. 3 *B*. Although the fluorescence intensity was quite constant in the unfolded regions, as seen in burst shape 1, super-Poissonian signal fluctuations were still evident. This indicates that although the molecules were uniformly stretched, they were not fully elongated to their contour length, resulting in strand density fluctuations. For each DNA molecule investigated, a fitting model was selected based on burst height, and the two partner bursts were fitted together. An unfolded molecule, for example, was fitted with the function $p(v_S(t - t_0)) + p_{Off1}$ for signal 1 and $p(v_S(t - t_0) - sd) + p_{Off2}$ for signal 2, both based on the function $p(x_0)$ of Eq. 7. For a molecule with a single front end loop, $p_{fold}(v_S(t - t_0)) + p_{Off1}$ was used to fit signal 1 and $p_{fold}(v_S(t - t_0) - sd) + p_{Off2}$ was used to fit signal 2, both based on Eq. 11.

Analogous functions were created and used for higher degrees of folding. In all cases, the fitting parameters included the speed v_S, the apparent lengths l_{A1} and l_{A2} resulting from the two signals, and the initial burst time t_0. Previously determined parameters such as the laser spot distance sd, fiber diameter d, objective magnification m, background noise levels p_{Off1} and p_{Off2} for the two signals, and the Gaussian radius of the point spread function σ were kept constant for all fits. The signal amplitude prefactors $p_{0,1}$ and $p_{0,2}$ for the two signals were generally kept constant. For analysis times longer than 2 min, the prefactors were fitted to account for dye loss and photobleaching effects in the nanoslit. Depending on the degree of folding, additional folding factors were also fitted, such as $f_a = l_{L1}/l_{A1} = l_{L2}/l_{A2}$ for a molecule with a single fold. It was assumed that the folding factors for both signals were constant for this experiment, whereas changes in apparent length (l_{A1}, l_{A2}) and folded over (looped) length (l_{L1}, l_{L2}) were possible. Due to slightly different optical fiber alignments, the two signals gave molecule lengths differing up to 10%, and the higher average length was selected, as the smaller value was assumed to arise from a slight misalignment of that optical fiber. An actual variation of the strand length between the two detection volumes was unlikely as length changes were expected to happen on a timescale of 9 s (30).

A total of 752 intact λ-bacteriophage molecules were detected in three similar experiments with a device bias of 50 V.

Bursts of fluorescence described by the β-, γ-, and ε-folding models were observed most often (33%, 46%, and 20%, respectively), whereas only 1% of observed bursts were described by the other models. The automated fitting algorithm selected the δ-fitting model for short fragments only, although the γ-fitting model would have also worked, and never for intact λ-bacteriophage molecules. The eighth burst shape in Fig. 3 *B* is shown as an example of a rare detection event which was not described fully by one of the six basic models shown in Fig. 3 *A*. Model β came closest to describing the shape of the burst in this case, which could have been caused by overlapping molecules, a fold, or a knot. A similar phenomenon was observed previously (30); in this earlier study, a knot was detected in a DNA molecule undergoing entropic recoil from a nanochannel, which did not dispel over time as a simple fold would have. Although this detection event shows a limitation of the models used to fit the bursts of fluorescence, it is also an illustration of the ability in a single-molecule measurement to detect rare species or transient events.

Fig. 4 shows the results of a 1 min experiment in which 416 molecules from a λ-bacteriophage DNA sample were detected. The definitions of the different lengths used to describe a folded DNA strand with a single loop are shown in Fig. 4 *C*. The apparent length l_A neglects any folding of the strand and is determined by analyzing only the start to end length of the DNA strand. The apparent length is of interest, as it enables a more direct comparison of the results presented here with other methods of DNA analysis, including translocation of DNA through nanopores and hydrodynamic stretching. The looped length l_L is the length of the folded over segment of DNA, and the real length l_R is the sum of the apparent and looped lengths. The free length l_F describes the section of DNA that is not folded over. Although the apparent length could also be determined by a simple thresholding algorithm, the other length values result only from fitting the burst shapes. A comparison of these different length values is shown in Fig. 4, *A* and *B*. The distribution of real lengths in Fig. 4 *B* shows a much sharper peak, corresponding to intact λ-bacteriophage DNA molecules, than the apparent length distribution in Fig. 4 *A*.

Based on analysis accounting only for apparent length, many molecules in this distribution appear to be fragments. When the real length is calculated, however, the molecules are found instead to be intact and folded, correcting the distribution. This is readily observed when the number of photons collected from each molecule is plotted against the apparent length (Fig. 4 *D*) and the real length (Fig. 4 *E*) of the molecules. Although the apparent length distribution of molecules in Fig. 4 *D* is widely scattered, the real length distribution in Fig. 4 *E* is focused and fit to a single line. The sample also contained a small fraction of λ-DNA multimers, which were observed at higher photon counts and with longer length than visualized in Fig. 4. The distributions of the photon count per molecule and of the real length can be fitted with Gaussians yielding a mean photon count of $48,000 \pm 2,000$ photons and a real length of $10.7 \pm 0.3\mu$m for intact λ-bacteriophage DNA molecules. This is longer than the equilibrium length of $7 - 8\ \mu$m measured for λ-bacteriophage molecules in 100 nm $\times$ 200 nm channels (26,29), which is expected, as the measured value is not the equilibrium length but rather the dynamically stretched length.

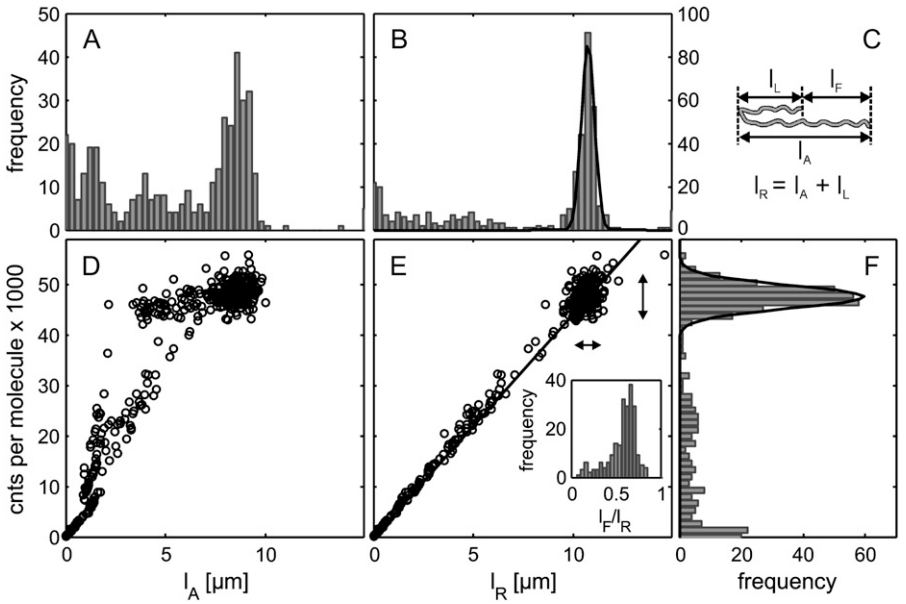

FIGURE 4 Photon count, length, and folding. In a 1 min run, 416 molecules from a λ-bacteriophage DNA sample are analyzed. The average molecule speed is $\bar{v}_S = 1.33 \pm 0.07\,\mathrm{mm/s}$ at a device bias of U = 50 V. (*A*) Distribution of the apparent length l_A. (*B*) Distribution of the real length l_R. (*C*) Schematic defining apparent length l_A, real length l_R, free length l_F, and folded (looped) length l_L for a molecule with a single loop. (*D*) Photon counts (cnts) per molecule versus apparent length l_A. (*E*) Photon counts per molecule versus real length l_R. The linear fit passing through the origin has a slope of $4,430 \pm 20\,\mathrm{cnts}/\mu\mathrm{m}$. (*F*) Distribution of the photon count per molecule. A comparison of *D* and *E* shows that folding explains the distribution of molecules with the same photon count over different apparent lengths. Fitting the photon count (*F*) and real length distributions (*C*) with Gaussian distributions leads to a mean photon count of $48,000 \pm 2,000$ photons and a real length of $10.7 \pm 0.3\ \mu$m for intact λ-DNA molecules. All molecules within two standard deviations of the mean photon count and the mean real length are considered to be intact (*horizontal* and *vertical arrows* indicate region). This accounts for 52% of the molecules analyzed, whereas the rest of the molecules are interpreted as fragments or concatemers. (*E*, *inset*) Distribution of the free length/real length ratio l_F/l_R for all intact molecules; 79% of the intact molecules have $l_F/l_R > 0.5$.

This stretching effect was previously found for T4-bacteriophage DNA molecules entering nanochannels. The ensuing relaxation showed a time constant of 9.3 s (30). The relative standard deviation of the Gaussian fits was found to be smaller for the real length distribution than for the photon count distribution (3.1% vs. 4.7%). All molecules within two standard deviations of the mean photon count and the mean real length were assumed to be intact λ-bacteriophage DNA molecules (*black arrows* mark region in Fig. 4 *D*), which accounts for 52% of all molecules detected. Assuming that the rest of the smaller molecules were fragments of λ-bacteriophage DNA molecules, and neglecting the few λ-bacteriophage concatemers detected, the sum of the length of these fragments divided by the mean real length of λ-bacteriophage DNA gives 60 as the original number of intact molecules.

The percentage of intact strands was then calculated to be 78%, which is somewhat lower than the 86% of intact λ-bacteriophage DNA molecules as determined by agarose gel electrophoresis. This measurement was performed by the manufacturer using samples from the same lot, and the resulting fluorescence image was analyzed with a rolling ball background correction using ImageJ (National Institute of Mental Health, Bethesda, MA), a Gaussian distribution to model the spatially averaged intensity of the DNA band, and a range of two standard deviations around the mean to differentiate intact molecules from fragments. Contributing factors to this increased fragmentation include hydrodynamic shear forces from manual sample loading, electrical forces in the nanochannel device, and photodamage from scattered laser light. The amount of fragmentation reported here compares favorably, however, to that observed using other methods of elongation and analysis (23).

To estimate the resolution of the method, the mean standard error of the fitted apparent length and the real length of the intact λ-bacteriophage molecules were calculated to be 84 nm and 114 nm, respectively, from 97 molecules analyzed using model β. For higher orders of folding, the mean standard error for the real length increases: 143 nm for model γ and 278 nm for model ϵ. These values can be indirectly compared with a resolution of 150 nm obtained in 1 min by Tegenfeldt et al. for a single λ-bacteriophage DNA molecule in a nanochannel at its equilibrium length of 8 μm (26). The values cannot be directly related, however, as the former is the average standard error of fits to single bursts and the latter results from a Gaussian fit to a distribution of length measurements taken over a longer period of time to average out thermal fluctuations for a single molecule.

The free length of the folded molecules is of particular interest, as this is the section of the strand where the position of a site-specific fluorescent label could be determined in future experiments without ambiguity. The inset in Fig. 4 *E* shows the distribution of the free length/real length ratio l_F/l_R for all intact molecules. Of the intact molecules, 79% show a $l_F/l_R > 0.5$, meaning that in half of these cases (39.5%) any position up to the middle of the strand can be determined. This is based, however, on the assumption that half of the molecules enter the channel in the right direction (5′-end first, for example). This reduces the number of λ-bacteriophage molecules that could yield information about the site-specific position of a fluorescent tag to 26% (0.67×0.39). The inset in Fig. 4 *E* also justifies the interpretation of the peak value of the photon count distribution of all bursts as the result of an unfolded DNA strand occupying the focal volume and the use of this value as a signal prefactor.

A limiting factor of the method presented in Fig. 4 is the length of the focal volume $l = d/m = 1.25\ \mu$m. When the length of a DNA molecule in the nanochannel was shorter than this value, the resulting fluorescent bursts did not exhibit quantized intensity levels, and the signal amplitude was determined only by the number of dye molecules which are intercalated in the strand. Eq. 7 predicts that although molecular folding would not affect the amplitude of these signals, it would be visible in a sudden variation in the slopes at the beginning and the end of the fluorescent burst. These slope changes were not detectable, however, as signal fluctuations were too high. The α-fitting model was used in this case to describe signal intensities lower than the average burst height, resulting in an apparent length value determined by peak amplitude and as a result photon count. The burst-fitting and photon-counting methods therefore converged for molecules shorter than 1.25 μm, which can be seen in Fig. 4 *D*. The same limitation affected folded molecules with a folded over length of $<1.25\ \mu$m.

Additional information, such as molecular conformation, was theoretically present as variations in the signal slopes, but its analysis was again obscured by signal fluctuations as well as any deviations from assumptions made about the optical system, such as laser illumination and fiber collection efficiency profiles. In the case of conformation, different conformations of the folded strands were described by the same fitting model, and it was not possible to determine if a burst of fluorescence with a short initial peak was the result of a strand folded over at the end or a molecule showing a random coil configuration at the end. Although the latter configuration has been found in microchannel structures for DNA stretching (8), it was not described in DNA translocation experiments using 10–20 nm pores (2) nor was it found in optical investigations of longer DNA molecules in 100 nm nanochannels (30). Further insight into the conformation of smaller molecules could be gained by reducing the length of the focal volume. This could be achieved by using smaller fiber diameters, by using an objective with higher magnification (with the resulting reduction in spot distance for the optical setup used here), or by combining nanochannels with optical nanoslits for near field microscopy (8). Velocity information was also present in the signal slopes, and its analysis was similarly limited. This limit was overcome, however, by fitting partner bursts simultaneously, in which case the time of travel between the focal volumes dominated

the determination of velocity due to its large effect on the combined fit.

Previous work involving DNA stretching and analysis has made use of hydrodynamic forces to elongate DNA and subsequently determine the position of fluorescent peptide nucleic acid tags bound to specific motif sites, yielding sequence information about the strand (21,22). Although this method is able to fully stretch DNA strands, the fluidic channels used were inhomogeneously illuminated. As shown here, dynamic electrophoretic stretching in nanochannels is not yet optimized to yield fully stretched strands but achieves uniform illumination across the nanochannel. This increases excitation uniformity and enables the determination of the DNA conformation by analytic fitting, theoretically permitting even the use of folded strands for the determination of site-specific information. Uniform illumination also yields more accurate fragment sizing based on photon counting, which leads to a higher percentage of molecules that can be confirmed as intact. This is achieved without the velocity correction of the photon count as often used in hydrodynamic stretching (22,23).

Electrodynamic stretching was also observed to be uniform along the length of the nanochannel, although the molecules were not completely elongated. This means that even at speeds an order of magnitude smaller than in hydrodynamic stretching, tag positions could be determined using the stronger fluorescent signals resulting from longer illumination times. Further electrodynamic stretching of the strands could most likely be achieved by the creation of stronger electric field gradients at the entrance of the nanochannel. Gradients of at least an order of magnitude stronger can be achieved by the use of smaller nanochannels, the connection of fewer parallel nanochannels to the nanoslit, or increasing the depth of the nanoslit. The fabrication of 10 nm nanochannels (44) and the repetitive translation of DNA molecules from nanoslits to microchannels (3,4) have already been shown.

Dependence of molecular speed on degree of folding

To investigate the effects of DNA folding on speed, the results of three subsequent 1 min runs of λ-bacteriophage DNA (including the experiment shown in Fig. 4) at a device bias of $U = 50$ V were combined. Fig. 5 *A* shows the distribution of the speed v_S versus the apparent length/real length ratio l_A/l_R for 752 intact λ-bacteriophage DNA molecules. Molecules with increased folding and therefore smaller l_A/l_R had a slightly higher speed v_S. The speed distribution v_S is also plotted as a function of the number of folds in the strands in Fig. 5 *B*. The number of folds for each molecule was determined by the analytical model chosen to describe its resulting fluorescent burst. The overall influence of folding was found to be weak. Unfolded molecules were rare, resulting in just two data points, and molecules with three folds were observed

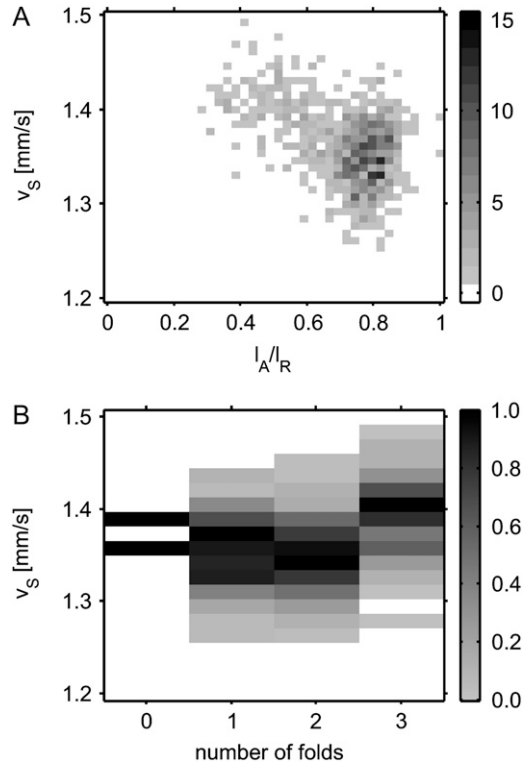

FIGURE 5 Speed and folding. (*A*) Distribution of speed v_S versus apparent length to real length ratio l_A/l_R for 752 λ-bacteriophage DNA molecules. Intact λ-bacteriophage molecules from three subsequent 1 min runs (including the experiment in Fig. 4) at a device bias of $U = 50$ V are presented. DNA molecules with increased folding and smaller l_A/l_R have a slightly higher speed v_S. (*B*) Normalized distribution of speed v_S versus number of folds for the same molecules shown above. The number of folds is determined by the analytical model (Fig. 3 *A*) chosen to analyze each molecule and is a measure of the number of parallel strands in the channel. DNA molecules with three folds in the molecule confirmation have slightly higher speed v_S.

to have a slightly higher speed v_S than molecules with one or two folds. This could be the result of increased hydrodynamic interactions between the parallel strands of the folded molecules. As electrophoretically driven DNA molecules are supposed to be free draining, these hydrodynamic interactions might take place only when the strand segments are very close, explaining the weakness of the effect. But molecular dynamics simulations may be necessary to fully explain this result.

Dependence of molecular speed, mobility, and friction on device bias

To investigate the effects of device bias on molecule speed and length, molecules from a λ-bacteriophage DNA sample were driven through a nanochannel at several different device biases U. Five continuous runs were performed for each device bias, and the cross correlation curves as well as the single bursts of fluorescence were analyzed by use of the

respective models. The average single-molecule speed $\bar{v}_S$ is plotted as a function of device bias U in Fig. 6 *A*. A linear fit of $\bar{v}_S(U)$ yields a slope of $m_S = 60.9 \pm 1.4\,\mu m/(Vs)$, confirming that speed increases linearly with device bias and electric field, as expected. For comparison, a plot of the speed $\bar{v}_C(U)$ resulting from the cross correlation functions is added to the figure and shows a very similar slope of $m_C = 61.0 \pm 0.7\,\mu m/(Vs)$.

The ratio of the electric field E inside one of the nanochannels to the device bias U was calculated from device dimensions to be 31 cm^{-1}. Dividing the slope of the linear fit of $\bar{v}_S(U)$ by this ratio gives the mobility of a DNA molecule as $\mu_S = \bar{v}_S/E = 2.0 \times 10^{-4}\,cm^2/(Vs)$. Assuming a charge per unit length of $\lambda = 1.1\,e_0/nm$ for the labeled DNA molecules (30) and considering that folding seems to have only a slight effect on speed, the friction coefficient per unit contour length was estimated to be $\xi_S = \lambda/\mu_S = 9.3\,fNs/\mu m^2$. This is close to the 10.0 $fNs/\mu m^2$ found in similar nanochannel devices (30). This calculation can be viewed only as an approximation, however, as some parallel channels to the investigated channel showed unintended constrictions with unknown influence on the real electric field in this particular device.

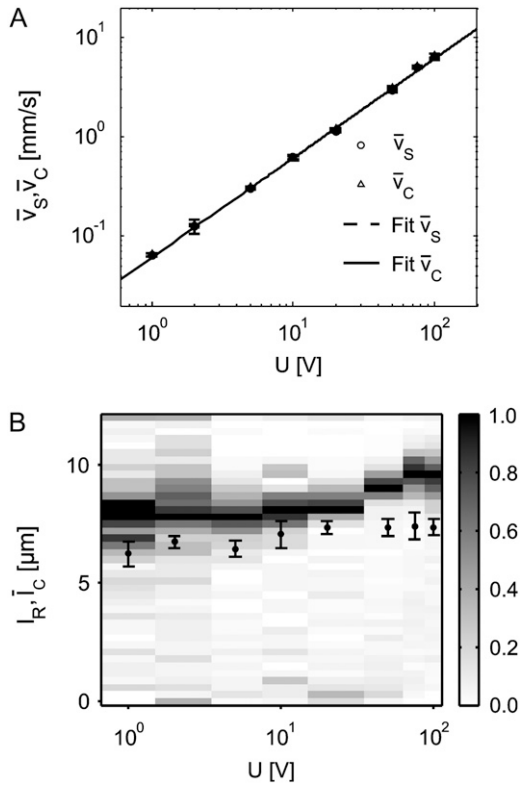

FIGURE 6 Speed, length, and device bias. The dependence of DNA speed and length on device bias for a λ-bacteriophage sample is shown. Five continuous runs were performed for each bias, and the cross correlation functions as well as the single-molecule photon bursts are analyzed. (*A*) Average single-molecule speed $\bar{v}_S$ and average cross correlation speed $\bar{v}_C$ versus device bias U. Linear fitting of $\bar{v}_S(U)$ yields a slope of $m_S = 60.9 \pm 1.4\,\mu m/(Vs)$, and a linear fit of $\bar{v}_C(U)$ gives a slope of $m_C = 61.0 \pm 0.7\,\mu m/(Vs)$. The plots overlap with each other. (*B*) Normalized distribution of real length l_R versus device bias U. The maximum of the length distribution shows up in black and corresponds to the majority of intact DNA molecules, which increase in real length l_R as the device bias increases. This behavior is not as evident for the average length $\bar{l}_C$ (*black dots*) resulting from cross correlation curves, probably because this analysis neglects folding and fragmentation. The number of molecules n for the different device biases U was 263 at 1 V, 381 at 2 V, 226 at 5 V, 144 at 10 V, 270 at 20 V, 417 at 50 V, 1217 at 75 V, and 868 at 100 V.

Dependence of stretching on device bias

Fig. 6 *B* illustrates the variation of real length of the DNA strands with increasing device bias. Because of the fragmentation of the λ-bacteriophage DNA molecules, the results for the single-molecule length are plotted as grayscale intensity distributions instead of the mean values. As the number of molecules per distribution varied between ~140 and ~1200 molecules, the real length distributions are shown normalized for each device bias. In rough terms, the maximum of the real length distributions, corresponding to the majority of intact λ-bacteriophage DNA strands, remains constant until ~10 V. An increase in real length is then observed until ~75 V, followed by a plateau at higher device biases. A complete explanation for this trend is beyond the scope of this study and must account for DNA stretching as a result of entrance effects and entropic confinement in the nanochannel, as well as hydrodynamic and channel surface friction. The average length $\bar{l}_C$ resulting from cross correlation analysis is plotted as an overlay to the single-molecule distributions. The increasing trend from the single-molecule distributions is not as evident, as the single-molecule analysis routine accounts for folding and fragmentation of the molecules neglected by the correlation analysis.

Dependence of length and speed on DNA size

An experiment was performed to test the sensitivity of the method, compare it to previously demonstrated techniques based on intensity measurements, and determine the dependence of DNA speed on the number of basepairs in a nanochannel. This dependence is of interest for the development of theory regarding the electrophoresis of DNA confined to nanoscale environments and for applications that exploit DNA size dependence to achieve biomolecular separation. A sample with a broad size range of DNA molecules was prepared by mixing λ-bacteriophage DNA in equal parts with its own *Hin*dIII digest. The resulting mixture contained nine primary types of DNA molecules (in basepairs: 125, 564, 2,027, 2,322, 4,361, 6,557, 9,416, 23,130, and 48,502). More than 16,000 molecules were detected in a 2 min experiment and automatically analyzed, the results of which are presented in Fig. 7. Almost all the scattered data points were the result of overlapping molecules following Poisson focal volume occupation statistics (12).

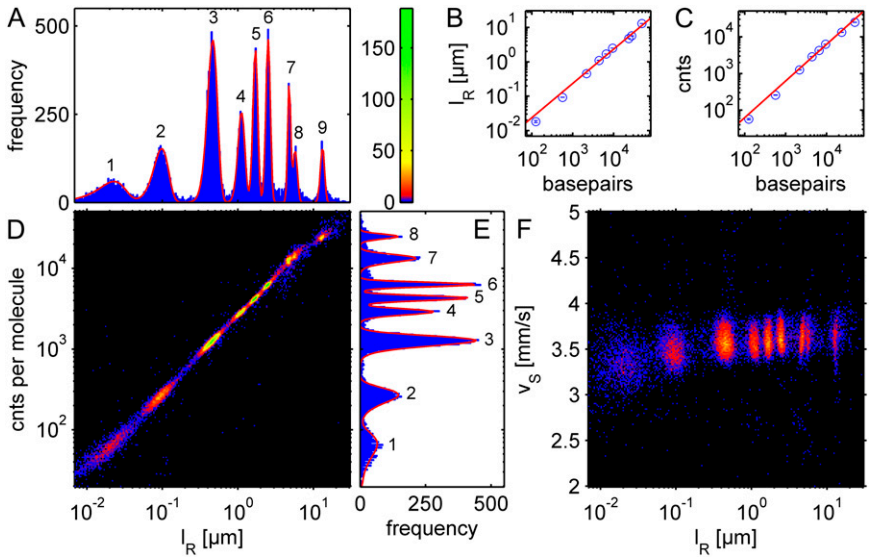

FIGURE 7 Photon count, length, speed, and size. Analysis of a sample containing a mixture of λ-bacteriophage DNA and its *Hin*dIII digest. In 2 min, 16,315 molecules were detected with an average speed $\bar{v}_S = 3.6 \pm 0.6$ mm/*s*. (*A*) Distribution of the real length l_R. The peaks are shown fitted to nine modified Gaussians (*red*). (*B*) Real length l_R versus number of basepairs. (*C*) Photon count (cnts) per molecule versus the number of basepairs. A linear fit yields a slope of $m_C = 0.63 \pm 0.02$ cnts/bp. (*D*) Distribution of photon counts cnts per molecule versus real length l_R. (*E*) Distribution of the photon counts per molecule. The peaks are fitted to eight modified Gaussians (*red*). The peaks in *A* and *E* are interpreted as the following DNA strand sizes, in number of basepairs: 125, 564, (2027 + 2322)/2, 4361, 6557, 9416, 23,130, 23,130 + 4361, 48,502. Whereas peaks 7 and 8 are resolved separately in *A*, they are combined in *E*. The values plotted in *B* and *C* are the mean values resulting from the Gaussian fits from *A* and *E*, respectively. (*F*) Distribution of molecule speed v_S versus real length l_R. DNA speed is essentially constant with length, showing only a slight decrease for the shortest fragment.

Two other phenomena were observed in this experiment which required modifications to the analysis algorithm, as described previously. The second photon burst consistently exhibited increased folding, which was addressed by allowing different folding factors for the partner bursts. This was most likely the result of increased nanochannel surface roughness between the two focal volumes (29). The longest DNA molecules also had a lower than expected photon count and burst height, which was accounted for by choosing a smaller fit amplitude for molecules with counts over 20,000 photons. This was attributed to the lower mobility of longer DNA strands in the nanoslit, with a resulting increase in the observed loss of dye molecules to the nanoslit surfaces and photobleaching by scattered laser light. Although this mobility effect was observed previously only in shallower nanoslits (45), a similar outcome may have occurred here because of the rougher nanoslit channel surfaces.

The sample also contained a small fraction of λ-bacteriophage DNA multimers which were observed at higher photon counts and with longer length than visualized in Fig. 7. Fig. 7 *D* shows a color-coded intensity histogram of photon counts per molecule versus real length l_R. The distributions of real length and photon count per molecule are plotted in Fig. 7, *A* and *E*, respectively. The distributions in Fig. 7, *A* and *E*, were fitted with superpositions of 9 and 8 logarithmically modified Gaussians, respectively, as described in Eq. 14. In both plots, peak 3 was interpreted as a combined peak of the 2,027-bp and 2,322-bp fragments, per the analysis of Foquet et al. (12). Peak 8 in Fig. 7 *A* results from a fraction of the 4,361-bp and 23,130-bp fragments annealing at the complementary ends and creating an additional 27,491-bp fragment (12), although the sample was heated to avoid this effect. This peak is not observed in Fig. 7 *E* due to the lower resolution of inferring DNA size through photon count at longer DNA lengths.

The fit results for these peak positions and relative standard deviations are shown in Table 1. The relative standard deviations decrease with increasing strand size for both fits, with the exception of peak 7 in Fig. 7 *E*, as it contains unresolved 27,491-bp DNA strands, which are fitted as a separate peak in Fig. 7 *A*. The resolution for both measures increases for longer molecules, as opposed to gel electrophoresis. This was also observed in previous photon count measurements and is related to the fact that fluctuations in dye fluorescence, detection efficiency, and the number of dye molecules attached to strands of equal size obey Poisson statistics (12,13). These fluctuations influence the quality of the burst fits in a similar fashion and therefore lead to a similar resolution increase for longer strands. Additionally, dye binding is sequence dependent, and the number of dye molecules bound to a strand did not reach equilibrium because of the low off-rate of YOYO-1 from DNA. Both phenomena negatively affected the resolution. The measured real length of 13.1 μm for λ-bacteriophage DNA was found to be longer than in the experiments shown in Figs. 4 and 6. These experiments were performed in different devices which may have different channel widths, particularly at the critical channel entrance. In Fig. 7, *B* and *C*, the mean length and the mean photon counts per molecule for the fitted peaks were plotted against the assumed number of basepairs for the fragments. The resulting dependence was well described by a linear fit, underscoring the fact that each peak corresponded to the expected fragment. This also indicates that the stretching factor $s = l_R/L$ is independent of strand length. The slopes of the fits were $m_{lR} = 0.24 \pm 0.01$ nm/bp and $m_C = 0.63 \pm 0.02$ cnts/bp. The latter is a factor of four

TABLE 1 Fitting results for Fig. 7, *A* and *E*

Peak in 7 *A*	Peak in 7 *E*	DNA strand size (bp)	Percentage of molecules in 7 *A*	Real length l_R (μm)	Relative standard deviation of l_R (%)	Counts per molecule (cnts)	Relative standard deviation of counts (%)
1	1	125	9.5	0.0182	56	57.1	44
2	2	564	12	0.0927	24	257	24
3	3	(2027 + 2322)/2	25	0.454	15	1250	15
4	4	4361	9.4	1.10	11	2880	8.7
5	5	6557	11	1.70	7.1	4280	7.0
6	6	9416	11	2.51	6.6	6310	6.3
7	7	23,130	6.0	4.77	4.2	13,200	11
8	/	23,130 + 4361	2.7	5.71	6.6	–	–
9	8	48,502	3.3	13.1	6.1	24,800	6.9

Multiple peaks were observed and matched with the known DNA strand sizes. Fitting the real length distribution in Fig. 7 *A* gives the real length l_R and the standard deviation of the different fragment sizes. Fitting the counts per molecule distribution in Fig. 7 *E* gives the counts per molecule and associated standard deviations. The percentage of molecules was derived from Fig. 7 *A*.

smaller than the 2.5 cnts/bp that was achieved by Foquet et al. (12), which is due to the use of different optical components including smaller optical fibers and increased adsorption of the dye molecules to the nanoslit surfaces. Despite this lower detection efficiency, the smallest fragment of the λ-*Hind*III digest was resolved due to decreased noise from a smaller nanochannel and the use of two focal volumes. The second photon count signal had a filtering effect, as fluorescent bursts without partners in the time range predicted by cross correlation analysis were discarded. This allowed time range was about an order of magnitude larger than the observed time difference variation and did not influence the distributions. The photon count for the longest DNA molecules in Fig. 7 *C* lies significantly below the linear fit, as a result of increased dye loss and photobleaching for this molecule in the device used here.

The percentage of the total molecules detected in each peak is also reported in Table 1. Considering annealing effects (12), peaks 1 –8 comprise the expected percentages of the total count. Peak 9 is reduced due to the retardation of long molecules in the nanoslit. Neglecting λ-bacteriophage molecules, the total percentage of intact digest fragments in these peaks was calculated to be 82%, based on the results shown in Fig. 7 *A*. The same value was determined to be 81% using agarose gel electrophoresis. This higher percentage of intact molecules as compared to λ-bacteriophage DNA results from random fragments coinciding with digested fragment peaks.

Fig. 7 *F* is a color-coded intensity histogram of the number of molecules against the molecule speed v_S and the real length l_R. This measurement was repeated for device biases of 25 V and 100 V (results not shown) and the same trend was observed, which was also similar to that observed for DNA molecules of comparable lengths in free solution (46). This indicates that the deformation from the random coil shape of unconfined DNA into the elongated and stretched conformation has a negligible influence on the speed of strands with length longer than the channel width. The smallest fragment (125 bp) was observed to have a slightly lower speed; a similar trend was observed around 100 bp in free solution (47) and attributed to electrolyte friction. It could also be related here to DNA lengths which are smaller than the width and depth of the nanochannel, permitting molecular orientations more transverse to the channel length which result in increased hydrodynamic friction. These orientations are denied to longer molecules, as the known stiffness of double-stranded DNA characterized by its persistence length results in an orientation more aligned to the channel axis. Further investigation is needed to clarify this effect.

This measurement shows the current limits of this method. To correctly interpret shorter DNA fragments with apparent length smaller than the focal volume length $d/m = 50\,\mu\text{m}/40 = 1.25\,\mu\text{m}$, the prefactors $p_{0,1}$ and $p_{0,2}$ of the fitting formulas must be accurately predetermined. This was possible only if the burst height for single unfolded DNA strands remained constant throughout the experiment. The time limit for this was found to be ~ 2 min as a result of stage or laser drift, which limited the number of molecules detected at lower voltages. Another limit was related to the device loading time. The fluorescence intensity of the molecules was found to decrease with the time the DNA molecules spent in the nanoslit region of the device, even before being directly exposed to the focused lasers. Explanations for this effect include photobleaching by scattered or reflected laser light or loss of positively charged YOYO-1 dye molecules attracted to the negatively charged channel walls. Increasing channel background fluorescence was indeed observed after the nanochannel outlets, although the solution contained PVP to dynamically coat the channel surfaces and reduce interactions. This phenomenon was particularly evident with the device used in the experiment shown in Fig. 7.

CONCLUSION

A method is presented to rapidly analyze fluorescently labeled DNA molecules, and several experiments investigating the physics of DNA in nanochannels were performed. DNA molecules were driven electrophoretically from a nanoslit into a nanochannel to confine and dynamically elongate the strands beyond their equilibrium conformation. Two lasers were focused sequentially along the length of the nano-

channel to induce bursts of fluorescence from the molecules. These bursts showed multiple intensity levels indicative of folding, and the collected photon count signals were analyzed with various analytical fitting models which yielded molecular conformation, length, and speed of single DNA strands.

Molecular length was determined both by fitting each fluorescent burst shape to analytical models and by photon counting. Although counting photons was simpler, modeling and fitting the bursts yielded better resolution for longer strands for the optical setup used here and has several other advantages for analyzing unknown samples: no calibration sample is needed for burst shape fitting, and it is less affected by excitation intensity and variations in fluorescent labeling. The resolution for λ-bacteriophage DNA with one fold at the front end was 114 nm for individual molecules with a mean real length of 11 μm, and the analysis time was $\sim$20 ms per molecule. The speed of the molecules was found to increase only slightly with folding. A possible cause of this trend is increased hydrodynamic interactions between strand segments in proximity. Deeper insight into the influence of hydrodynamic friction and surface interactions on the molecule speed could be achieved by molecular dynamics simulations. DNA stretching was found to increase with applied device bias and electric field, which was estimated from device dimensions to be 56 times higher in the nanochannels than in the nanoslit.

The analysis of a mixture of $\lambda-$bacteriophage with its own *Hin*dIII digest demonstrated that our method can identify, by length, 9 of 10 fragments ranging over 2.5 orders of magnitude in size. When compared to standard agarose gel electrophoresis, longer strands ($>$20 kbp) can be analyzed, the measurement time is at least 30 times faster, and much less sample is consumed. Although resolution decreases with size for agarose gels, the opposite is found for this method. DNA speed was found to be almost constant for all fragments investigated, showing only a slight decrease for short fragments. Confinement, elongation, and interaction with channel surfaces did not lead to a significant size-dependent influence on speed.

Future work will focus on extracting information pertaining to the length and conformation of molecules shorter than the focal volume length that is present in the signal start and end slopes but is currently obscured by signal fluctuations. This may be achieved by further elongating the DNA molecules by using smaller nanochannels, which would lead to a higher uniformity of basepair density along the channel. This will yield fluorescence signals with smaller fluctuations, leading to improved length determination using the method here. The accuracy of other physical information such as molecular conformation and speed will also increase, leading to more insight into the behavior of DNA molecules in highly constrained environments. As burst fitting can provide spatial information beyond the diffraction limit, the method here has potential for extremely rapid and ultrahigh spatial resolution measurements. Other pertinent applications of interest are the position of a site-specific label and biophysical effects such as knot formation. As this label or knot may fall in a folded over segment of DNA, with a resulting ambiguous position, models must be developed to extract positional information. The analysis of samples containing longer DNA strands will have to use photon burst fitting models that account for the increased probability of more complicated folding conformations and knot formation.

The authors gratefully acknowledge New England Biolabs for assistance with gel electrophoresis. Device fabrication was performed in part at the Cornell NanoScale Science and Technology Facility (CNF) with the assistance of Rob Ilic as well as in the Cornell Center for Materials Research.

This work was supported by the National Human Genome Research Institute as well as the Nanobiotechnology Center, which is funded by the STC Program of the National Science Foundation under agreement No. ECS-9876771 and the New York State Office of Science, Technology, and Academic Research (NYSTAR). CNF is a member of the National Nanotechnology Infrastructure Network, which is supported by the National Science Foundation.

REFERENCES

1. Li, J., D. Stein, C. McMullan, D. Branton, M. J. Aziz, and J. A. Golovchenko. 2001. Ion-beam sculpting at nanometer length scales. *Nature*. 412:166–169.
2. Li, J. L., M. Gershow, D. Stein, E. Brandin, and J. A. Golovchenko. 2003. DNA molecules and configurations in a solid-state nanopore microscope. *Nat. Mater.* 2:611–615.
3. Han, J., and H. G. Craighead. 1999. Entropic trapping and sieving of long DNA molecules in a nanofluidic channel. *J. Vac. Sci. Technol. A.* 17:2142–2147.
4. Han, J., and H. G. Craighead. 2000. Separation of long DNA molecules in a microfabricated entropic trap array. *Science.* 288:1026–1029.
5. Han, J. Y., and H. G. Craighead. 2002. Characterization and optimization of an entropic trap for DNA separation. *Anal. Chem.* 74:394–401.
6. Samiee, K. T., M. Foquet, L. Guo, E. C. Cox, and H. G. Craighead. 2005. λ-repressor oligomerization kinetics at high concentrations using fluorescence correlation spectroscopy in zero-mode waveguides. *Biophys. J.* 88:2145–2153.
7. Levene, M. J., J. Korlach, S. W. Turner, M. Foquet, H. G. Craighead, and W. W. Webb. 2003. Zero-mode waveguides for single-molecule analysis at high concentrations. *Science.* 299:682–686.
8. Tegenfeldt, J. O., O. Bakajin, C. F. Chou, S. S. Chan, R. Austin, W. Fann, L. Liou, E. Chan, T. Duke, and E. C. Cox. 2001. Near-field scanner for moving molecules. *Phys. Rev. Lett.* 86:1378–1381.
9. Turner, S. W. P., M. Cabodi, and H. G. Craighead. 2002. Confinement-induced entropic recoil of single DNA molecules in a nanofluidic structure. *Phys. Rev. Lett.* 88:128103.
10. Kaji, N., Y. Tezuka, Y. Takamura, M. Ueda, T. Nishimoto, H. Nakanishi, Y. Horiike, and Y. Baba. 2004. Separation of long DNA molecules by quartz nanopillar chips under a direct current electric field. *Anal. Chem.* 76:15–22.
11. Huang, L. R., J. O. Tegenfeldt, J. J. Kraeft, J. C. Sturm, R. H. Austin, and E. C. Cox. 2002. A DNA prism for high-speed continuous fractionation of large DNA molecules. *Nat. Biotechnol.* 20:1048–1051.
12. Foquet, M., J. Korlach, W. Zipfel, W. W. Webb, and H. G. Craighead. 2002. DNA fragment sizing by single molecule detection in submicrometer-sized closed fluidic channels. *Anal. Chem.* 74:1415–1422.
13. Chou, H. P., C. Spence, A. Scherer, and S. Quake. 1999. A microfabricated device for sizing and sorting DNA molecules. *Proc. Natl. Acad. Sci. USA.* 96:11–13.
14. Foquet, M., J. Korlach, W. R. Zipfel, W. W. Webb, and H. G. Craighead. 2004. Focal volume confinement by submicrometer-sized fluidic channels. *Anal. Chem.* 76:1618–1626.

15. Stavis, S. M., J. B. Edel, K. T. Samiee, and H. G. Craighead. 2005. Single molecule studies of quantum dot conjugates in a submicrometer fluidic channel. *Lab Chip*. 5:337–343.

16. Stavis, S. M., J. B. Edel, Y. G. Li, K. T. Samiee, D. Luo, and H. G. Craighead. 2005. Detection and identification of nucleic acid engineered fluorescent labels in submicrometre fluidic channels. *Nanotechnology*. 16:S314–S323.

17. Stavis, S. M., J. B. Edel, Y. G. Li, K. T. Samiee, D. Luo, and H. G. Craighead. 2005. Single-molecule mobility and spectral measurements in submicrometer fluidic channels. *J. Appl. Phys*. 98:044903.

18. Stavis, S. M., S. C. Corgié, B. R. Cipriany, H. G. Craighead, and L. P. Walker. 2007. Single molecule analysis of bacterial polymerase chain reaction products in submicrometer fluidic channels. *Biomicrofluidics*. 1:034105.

19. Mannion, J. T., and H. G. Craighead. 2007. Nanofluidic structures for single biomolecule fluorescent detection. *Biopolymers*. 85:131–143.

20. Craighead, H. 2006. Future lab-on-a-chip technologies for interrogating individual molecules. *Nature*. 442:387–393.

21. Perkins, T. T., D. E. Smith, and S. Chu. 1997. Single polymer dynamics in an elongational flow. *Science*. 276:2016–2021.

22. Phillips, K. M., J. W. Larson, G. R. Yantz, C. M. D'Antoni, M. V. Gallo, K. A. Gillis, N. M. Goncalves, L. A. Neely, S. R. Gullans, and R. Gilmanshin. 2005. Application of single molecule technology to rapidly map long DNA and study the conformation of stretched DNA. *Nucleic Acids Res*. 33:5829–5837.

23. Chan, E. Y., N. M. Goncalves, R. A. Haeusler, A. J. Hatch, J. W. Larson, A. M. Maletta, G. R. Yantz, E. D. Carstea, M. Fuchs, G. G. Wong, S. R. Gullans, and R. Gilmanshin. 2004. DNA mapping using microfluidic stretching and single-molecule detection of fluorescent site-specific tags. *Genome Res*. 14:1137–1146.

24. Bakajin, O. B., T. A. J. Duke, C. F. Chou, S. S. Chan, R. H. Austin, and E. C. Cox. 1998. Electrohydrodynamic stretching of DNA in confined environments. *Phys. Rev. Lett*. 80:2737–2740.

25. Larson, J. W., G. R. Yantz, Q. Zhong, R. Charnas, C. M. D'Antoni, M. V. Gallo, K. A. Gillis, L. A. Neely, K. M. Phillips, G. G. Wong, S. R. Gullans, and R. Gilmanshin. 2006. Single DNA molecule stretching in sudden mixed shear and elongational microflows. *Lab Chip*. 6:1187–1199.

26. Tegenfeldt, J. O., C. Prinz, H. Cao, S. Chou, W. W. Reisner, R. Riehn, Y. M. Wang, E. C. Cox, J. C. Sturm, P. Silberzan, and R. H. Austin. 2004. The dynamics of genomic-length DNA molecules in 100-nm channels. *Proc. Natl. Acad. Sci. USA*. 101:10979–10983.

27. Riehn, R., M. C. Lu, Y. M. Wang, S. F. Lim, E. C. Cox, and R. H. Austin. 2005. Restriction mapping in nanofluidic devices. *Proc. Natl. Acad. Sci. USA*. 102:10012–10016.

28. Reisner, W., K. J. Morton, R. Riehn, Y. M. Wang, Z. N. Yu, M. Rosen, J. C. Sturm, S. Y. Chou, E. Frey, and R. H. Austin. 2005. Statics and dynamics of single DNA molecules confined in nanochannels. *Phys. Rev. Lett*. 94:196101.

29. Reccius, C. H., J. T. Mannion, J. D. Cross, and H. G. Craighead. 2005. Compression and free expansion of single DNA molecules in nanochannels. *Phys. Rev. Lett*. 95:268101.

30. Mannion, J. T., C. H. Reccius, J. D. Cross, and H. G. Craighead. 2006. Conformational analysis of single DNA molecules undergoing entropically induced motion in nanochannels. *Biophys. J*. 90:4538–4545.

31. Gordon, M. P., T. Ha, and P. R. Selvin. 2004. Single-molecule high-resolution imaging with photobleaching. *Proc. Natl. Acad. Sci. USA*. 101:6462–6465.

32. Thompson, R. E., D. R. Larson, and W. W. Webb. 2002. Precise nanometer localization analysis for individual fluorescent probes. *Biophys. J*. 82:2775–2783.

33. Elson, E. L., and D. Magde. 1974. Fluorescence correlation spectroscopy. 1. Conceptual basis and theory. *Biopolymers*. 13:1–27.

34. Magde, D., W. W. Webb, and E. Elson. 1972. Thermodynamic fluctuations in a reacting system—measurement by fluorescence correlation spectroscopy. *Phys. Rev. Lett*. 29:705–708.

35. Magde, D., E. L. Elson, and W. W. Webb. 1974. Fluorescence correlation spectroscopy. 2. Experimental realization. *Biopolymers*. 13: 29–61.

36. Brinkmeier, M., K. Dorre, J. Stephan, and M. Eigen. 1999. Two beam cross correlation: a method to characterize transport phenomena in micrometer-sized structures. *Anal. Chem*. 71:609–616.

37. Magde, D., W. W. Webb, and E. L. Elson. 1978. Fluorescence correlation spectroscopy. 3. Uniform translation and laminar-flow. *Biopolymers*. 17:361–376.

38. Gao, Q. F., and E. S. Yeung. 1998. A matrix for DNA separation: genotyping and sequencing using poly(vinylpyrrolidone) solution in uncoated capillaries. *Anal. Chem*. 70:1382–1388.

39. Kang, S. H., M. R. Shortreed, and E. S. Yeung. 2001. Real-time dynamics of single-DNA molecules undergoing adsorption and desorption at liquid-solid interfaces. *Anal. Chem*. 73:1091–1099.

40. Ueda, M., T. Hayama, Y. Takamura, Y. Horiike, T. Dotera, and Y. Baba. 2004. Electrophoresis of long deoxyribonucleic acid in curved channels: the effect of channel width on migration dynamics. *J. Appl. Phys*. 96:2937–2944.

41. Perkins, T. T., D. E. Smith, R. G. Larson, and S. Chu. 1995. Stretching of a single tethered polymer in a uniform flow. *Science*. 268:83–87.

42. Chen, P., J. J. Gu, E. Brandin, Y. R. Kim, Q. Wang, and D. Branton. 2004. Probing single DNA molecule transport using fabricated nanopores. *Nano Lett*. 4:2293–2298.

43. Austin, R. 2003. Nanopores—the art of sucking spaghetti. *Nat. Mater*. 2:567–568.

44. Cao, H., Z. N. Yu, J. Wang, J. O. Tegenfeldt, R. H. Austin, E. Chen, W. Wu, and S. Y. Chou. 2002. Fabrication of 10 nm enclosed nanofluidic channels. *Appl. Phys. Lett*. 81:174–176.

45. Cross, J. D., E. A. Strychalski, and H. G. Craighead. 2007. Size-dependent DNA mobility in nanochannels. *J. Appl. Phys*. 102:024701.

46. Stellwagen, N. C., C. Gelfi, and P. G. Righetti. 1997. The free solution mobility of DNA. *Biopolymers*. 42:687–703.

47. Stigter, D. 2002. Wall effects on DNA stretch and relaxation. *Biophys. Chem*. 101:447–459.