Origins and Consequences of Velocity Fluctuations during DNA Passage through a Nanopore

Origins and Consequences of Velocity Fluctuations during DNA Passage through a Nanopore

Lu Bo[†, ‡], Fernando Albertorio[†], David P. Hoogerheide[†], and Jene A. Golovchenko[†*]

[†]Department of Physics, Harvard University, Cambridge MA 02138
[‡]Electron Microscopy Laboratory, Department of Physics, Peking University, Beijing, China

*Correspondence: golovchenko@physics.harvard.edu

ABSTRACT

We describe experiments and modeling results that reveal and explain the distribution of times that identical double-stranded DNA (dsDNA) molecules take to pass through a voltage-biased solid-state nanopore. We show that the observed spread in this distribution is caused by viscous drag induced velocity fluctuations that are correlated with the initial conformation of nanopore-captured molecules. This contribution exceeds that due to diffusional Brownian motion during the passage. Nevertheless, and somewhat counter-intuitively, the diffusional Brownian motion determines the fundamental limitations of rapid DNA strand sequencing with a nanopore. We model both diffusional and conformational fluctuations in a Langevin description. It accounts well for passage time variations for DNA molecules of different lengths, and predicts conditions required for low error rate nanopore strand DNA sequencing with nanopores.

INTRODUCTION

Solid-state nanopores are capable of detecting and characterizing individual charged polymers in solution. Recent work in solid-state (1-4) and protein-based nanopores (5-7) has focused on detecting DNA (8-11) with the aim of developing a rapid, inexpensive single-molecule DNA strand sequencing capability. The potential advantages of a nanopore's approach to sequencing are many, as are fundamental challenges to its realization. In particular, a sequencing nanopore should have sub nanometer spatial and sufficient temporal resolution for detecting and identifying individual bases along a DNA molecule as it passes through the nanopore (8). The recent development of single-layer graphene nanopores (12), multi-layer graphene nanopores (13), and coated graphene nanopores (14), is an important step in achieving the required spatial resolution, but in connection with the time resolution, little attention has been paid to the impact of motional fluctuations of the DNA molecule in the nanopore. Thus the accuracy of a 'read,' or sequence, will depend on the uniformity of DNA motion through the pore. In particular, backward motion of the DNA in the nanopore, caused by large velocity fluctuations, will give rise to sequencing errors and pose a potential limitation to the accuracy of nanopore strand sequencing.

Virtually all of the literature on the dynamics of DNA passage, or "translocation," through biological and solid-state pores focuses on understanding the relationship between DNA length and total translocation time. Experimental results show a super-linear power-law relationship (2-3, 5, 15-16), consistent with scaling-law predictions (16-19) and simulations (20-21). Escape

time experiments in biological ion channels indicate interaction of the DNA molecule with the biological channel (6-7, 22); loss rates in single-molecule trapping experiments suggest that similar interactions may occur in solid-state nanopores (23).

Given its importance for DNA sequencing, it is surprising that there has been little discussion of the *variation* in observed translocation times of identical DNA molecules. This variation has been reported in some cases (3, 24), but quantitatively characterizing, modeling, and identifying the origins of the variations have not been achieved.

Here we present a study of the electrophoretic motions of dsDNA molecules during nanopore translocation. We present experimental results for translocation time distributions of DNA molecules of two different DNA lengths that are selected to have an unfolded, single file passage through the same nanopore. A simple model that relates the variation in translocation time to the unraveling of different (equally probable) initial conformational geometries of otherwise identical molecules is shown to agree well with experiments. The model predicts large velocity fluctuations during nanopore translocation that are produced by fluctuations in the drag force on a time-varying part of captured DNA molecules that have yet to pass through the pore. These fluctuations are correlated with the unraveling kinetics of DNA molecules from their random conformations during translocation through the nanopore. They are responsible for the spreads in translocation time distributions. Conformation-induced velocity fluctuations are to be distinguished from the diffusional Brownian motion fluctuations during the translocation event. We evaluate the consequences of Brownian motion fluctuations for strand sequencing and show that controlling their contribution is critical for strand sequencing, and derive conditions necessary for strand sequencing with a nanopore.

DNA TRANSLOCATION EXPERIMENTS

Experimental methods

Nanopore translocation experiments were designed to simultaneously study a mixture of two different length (5.3 and 10 kilobase) dsDNA molecules that pass through the same nanopore. These two length molecules exhibit well-separated translocation time distributions. A nanopore of diameter 8-10 nm was fabricated with a 200 keV JEOL 2010F transmission electron microscope in a 60 nm thick free-standing membrane of low stress silicon nitride, as previously described (25). To reduce capacitance and associated electronic noise, the nitride layer was framed by a 2-μm thick support layer of thermal silicon dioxide grown on the underlying silicon substrate. Nanopores fashioned in this way have an hourglass shape with a central cylindrical region whose effective channel length is about 20 nm (26).

Fig. 1 shows a schematic of the experiment as a captured DNA molecule, from a random conformation, passes through a nanopore. Also shown in Fig. 1 is a freely-jointed chain model of the DNA molecule that is used in the model. The free-standing membrane with the nanopore separates two reservoirs filled with 1.6 M KCl maintained at pH 8 by a 10 mM tris, 1 mM EDTA buffer. The *trans* reservoir was biased at $V_{bias} = +100$ mV relative to the *cis* reservoir by an Axopatch 200B patch-clamp amplifier coupled to the two reservoirs by Ag/AgCl electrodes. Under these conditions, a stable ionic current of 9 nA flowed through the nanopore, consistent with the nanopore geometry described above.

Double stranded DNA of length 5.3 kilobases (kb) was prepared from a $\phi \times 174RF1$ plasmid (obtained from New England Biolabs, Ipswich, MA) by cutting the plasmid with SspI restriction enzyme and purifying by gel electrophoresis. 10 kb dsDNA was purchased from New England Biolabs and purifed by gel electrophoresis to yield no detectable contaminants. A 1:1 mixture of the purified 5.3 kb and 10 kb fragments was prepared at ~1 nM each and injected into the *cis* reservoir. Single molecule DNA-nanopore translocation events were observed as transient ionic current blockades. The ionic current data for each event was filtered by an 8-pole 40 kHz low-pass Bessel filter and digitized at 250 kilosamples/second. About 4700 single molecule current blockades events were observed. Previous work (2) has confirmed that these events are the result of single dsDNA molecules translocating through the solid-state nanopore and that the structure of each event is related to the folding of captured molecules in the nanopore (Fig. 2*A*, inset). Nanopores may "clog", either showing persistent current blockages and/or extended translocation times. This clogging is likely due to the sticking of DNA molecules or other impurities on the nanopore wall. No events after such a "clogging" incident have been included in the data presented here.

All dsDNA translocation current trace events were processed using MATLAB code that fits each event to a series of sharp current steps modified by the transfer function of the experimental low-pass filter. The results comprise a data set represented by a 2-D histogram of average ionic current blockage versus event duration (Fig. 2*A*). About 1350 events corresponding to only unfolded DNA translocations (single-level events, in which only one double helix occupies the pore at all times during the translocation process) were selected. The distribution of these translocation times was then compared to simulated DNA translocation events (Fig. 2*B*) using a model described in the section on modeling below and in Appendices A and S1.

Experimental results and discussion

The populations of all 5.3 kb and 10 kb events are readily distinguishable as two crescent-shaped structures in the two-dimensional current blockage-translocation time histogram shown in Fig. 2, *A*. Only the events between the dotted lines correspond to molecules that have passed through the pore unfolded, and it is these events that allow velocity fluctuations to be modeled and applied to the analysis of translocation time distributions in a straightforward way.

Translocation time distributions of 1350 unfolded events from the experiment are plotted in Fig. 2*B*. These distributions are slightly asymmetric with the average translocation time larger than the most probable value. The most probable translocation time for 10 kb DNA (417 ± 5 μs) is slightly more than twice that of the 5.3 kb DNA (202 ± 1 μs), in agreement with previous findings (2, 16). The average translocation times for 5 kb and 10 kb DNA are 211 ± 1 μs and 443 ± 5 μs, respectively. Error bars were calculated from numerical analysis of the translocation time distributions.

The experimental distribution widths (FWHM) are seen to be ~33% of the mean unfolded event translocation times. These widths are significantly larger than would be expected from the diffusive motion of the molecule during the translocation. This can be seen by considering the velocity v of the molecule in the nanopore to be described by a one-dimensional Langevin equation:

$$m\frac{dv}{dt}+\gamma v-F=A(t). \qquad (1)$$

A DNA molecule of mass m and linear charge density λ, translocating through the nanopore experiences a driving force, $F=V_{bias}\lambda$, due to the electric field in the nanopore. With $\lambda \approx 0.2\ e^{-}$/bp (see Appendix S3 in the Supplementary Material) and a bias voltage of 100 mV, this driving force is $F \approx 9.4$ pN. The DNA also experiences a stochastic thermal (Brownian) force $A(t)$ from random molecular collisions. The driving force is opposed by a drag force $-\gamma v$ on the part of the molecule being moved through the solution. Experimentally, the mean translocation velocity $\langle \bar{v} \rangle$ for the 10 kb DNA is 7.4 mm/s. Assuming that the velocity of the DNA in the nanopore is roughly constant during the translocation, the fluctuation-dissipation theorem predicts an average diffusion constant given by $\langle D \rangle = k_B T/\gamma = \langle \bar{v} \rangle k_B T/F \approx 3.3\times10^{-12}$ m^2/s. The ratio between the translocation time FWHM, $2\sqrt{2\ln 2}\sigma_T$, and the mean translocation time, t_0, is

$$2\sqrt{2\ln 2}\frac{\sigma_T}{t_0} \approx 2.35\left(\frac{2\langle D\rangle t_0}{t_0^2\langle \bar{v}\rangle^2}\right)^{0.5} \sim 0.038, \qquad (2)$$

resulting in a translocation time width of ≈4% of the mean translocation time, whereas the experimentally observed widths in Fig. 2*B*, are ≈33% of the mean. Thus Brownian motion during the translocation is not the dominant factor responsible for the observed translocation time widths.

MODELING VELOCITY FLUCTUATIONS

Basic Theory

We posit that the main contribution to spread in translocation times is from fluctuations in the drag force that can be modeled as a time dependent drag coefficient γ in Eq. 1 . Thus as each molecule unraveled from its own geometrical conformation in the solution when it passes through the nanopore only part of the molecule chain is being dragged through the solution (the "dragged part") towards the nanopore at any given time. The rest of the chain (the "undragged part") remains unaffected. The time-varying position along the molecule that separates the dragged and and undragged regions along the molecule will be called the "pivot point" (Fig. 1). Details of how the pivot point is identified in our modeling simulations are discussed below and in Appendix S1.

What happens to the "undragged" region of the molecule during the translocation event depends on the details of the relaxation kinetics. The relaxation time for the entire molecule to attain an equilibrium conformation in solution is estimated by Zimm dynamics to be

$$\tau_{Zimm} = \frac{\eta(\sqrt{N}l_0)^3}{\sqrt{3\pi}kT}, \qquad (3)$$

where N is the number of Kuhn segments of length l_0 and η is the viscosity of the solution (27). For a 10 kb dsDNA molecule in water, the Zimm time is $\tau_{Zimm} \approx 20$ ms , two orders of magnitude longer than the experimentally observed translocation times. Thus for our experimental conditions one is justified in assuming that the molecule conformation past the pivot point is frozen during the translocation process. We note that this assumption is not valid for short single-stranded DNA (ssDNA) molecules passing through a protein nanopore because the Zimm time is much shorter than the ssDNA translocation time (28). Also, the freely-jointed chain model is reasonable only for dsDNA molecules longer than a few Kuhn lengths and applies to our experimental situation.

Consider a generalized Langevin equation of motion for the time dependent "dragged" part of the translocating DNA molecule. As it unravels, the drag coefficient $\gamma(t)$ and mass $m(t)$ change with time. Thus

$$m(t)\frac{dv}{dt} - \gamma(t)v(t) - F = A(t;\gamma(t)). \qquad (4)$$

Averaging Eq. 4 over times shorter than the translocation time of a Kuhn length, but much longer than the inertial relaxation time $\tau_R = \frac{m(t)}{\gamma(t)} \sim 10^{-12}$ s (29), A and $\frac{dv}{dt}$ average to zero. One obtains a conformation-dependent time varying average velocity

$$\overline{v}(t) = F / \gamma(t). \qquad (5)$$

The fluctuations from Brownian motion are implicitly contained in equation (5) through an application of the fluctuation-dissipation theorem which relates the microscopic high frequency fluctuations in $A(t)$ to the damping constant $\gamma(t)$. The role of Brownian motion for nanopore sequencing will be discussed later in this paper,

The calculational details used to study the time-variation of the drag force in Eq. 5 are provided in Appendices A and S1. Briefly, the DNA is modeled by a freely jointed chain where each rigid segment is one Kuhn length long (100 nm or ~294 base pairs). The 5.3 kb and 10 kb DNA molecules used in the experiment are 18 and 34 Kuhn lengths, respectively. The motion of the freely jointed chain passing through the nanopore is calculated as the molecule unravels from each member of an ensemble of appropriately selected random initial conformation. For each translocating segment at the nanopore, the calculation determines the minimum unraveling motion required from the rest of the molecule that enables the segment to pass through the pore. The changing pivot point is therefore uniquely determined as each segment passes through the nanopore. The drag force, summed over each of the dragged Kuhn segments before the "pivot point", determines the molecular velocity of the translocating segment and hence the segment translocation time. This calculation is carried out for successive segments to yield the translocation time of the whole chain. The configuration of the molecule is restricted to a two-dimensional hexagonal lattice for computational accessibility. The plane of the lattice passes through the nanopore and lies perpendicular to the membrane. We find that the reduced

dimensionality and the coarse lattice structure do not have a significant influence on the results (see Appendix S2).

Modeling the translocation dynamics of chains of the same length but different initial conformations provides predictions of translocation time distributions. The obtained translocation time for each conformation is sensitive to two free parameters in the model. The first is the linear charge density λ of the DNA molecule and the second is a "center of mass offset," defined as the number of segments which are assumed unraveled and straight when the DNA translocation begins. This parameter accounts for an electric field gradient unraveling effect on the molecule before it actually begins to translocate through the nanopore, as well as any short range drag force from the portion of the molecule on the *trans* side of the nanopore. The special role of the center of mass position of a particular conformation is discussed later in the paper. Because all the equations are linear in the charge density, this parameter linearly scales the distribution with respect to the time axis. As described in Appendix A, the center of mass offset changes the initial conformation distribution. Comparison of the dashed and solid lines in Fig. 2*B* illustrates that distributions with larger center of mass offsets are narrower and shifted to longer translocation times.

Modeling results

Calculated translocation time distributions for 10^4 DNA molecules of 18 and 34 Kuhn lengths are shown as the smooth curves in Fig. 2*B*. The dashed curves come from the model translocation dynamics of random walk generated conformations with no electric field gradient molecule straightening (zero center of mass offset) before capture into the nanopore. The experimental data are however best described by a center of mass offset of 3 Kuhn lengths and an effective charge density of 0.22 e^-/bp plotted as the solid line in Fig. 2*B*. The distribution shapes, widths, and positions are in very good agreement with the experimental data, and the effective charge density is consistent with previous measurements of DNA mobility (see Appendix S3). An additional data set (not shown) of a slightly smaller pore in 1 M KCl (conductance ≈ 30 nS, compared to ≈ 90 nS in 1.6 M KCl for the data shown) was best fit with a center of mass offset of 3 Kuhn lengths and an effective charge density of 0.30 e^-/bp. The larger effective charge density in a smaller nanopore is in agreement with the literature (30).

With the help of the generalized Langevin model, we have also studied the correlation between the initial center of mass distance from the nanopore of each conformation and the translocation time for that conformation. Fig. 3, a histogram of the results for all modeled conformations of both 5.3 kb (18 Kuhn lengths) and 10 kb (34 Kuhn lengths), shows that shorter translocation times belong to molecules whose initial conformations on capture into the pore are centered near the pore, while molecules whose initial conformations are centered far from the pore take longer to translocate. The correlation is very strong and linear for both 5.3 and 10 kb DNA. For sufficiently long molecules, this means that the spread in translocation times arises primarily from the spread in this distance for different molecule conformations. This correlation likely lies behind the success of the scaling arguments used to relate most probable translocation time to molecular length (15-16).

The fluctuation of the translocation velocity for a single molecule due to its unraveling was also explored with the model. (Unfortunately no one has figured out how to measure this experimentally yet.) Fig. 4 shows the predicted velocity $\overline{v}(t)$ for three molecule conformations. The average velocities, and hence translocation times, of these three conformations are very different, and in all cases the instantaneous velocity differs significantly from the average velocity during the translocation event. The first few molecular segments that enter the pore traverse quickly because only a few segments from the molecule's initial conformation contribute to the drag. As more segments are pulled through the pore, additional segments on the *cis* side of the nanopore become correlated to the motion of the segment in the pore, increasing the drag and reducing the translocation speed; that is, as the molecule translocates, the pivot point, on average, moves away from the nanopore. Towards the end of the process, the molecule on the *cis* side is completely stretched. As it is pulled through the pore, it now becomes shorter, the drag force becomes less, and the translocation speed increases. These effects are particularly pronounced for molecules whose center of mass starts out far from the nanopore (molecule 3 in Fig. 4); for molecules whose center of mass is close to the nanopore (molecule 1), these trends are less pronounced than the velocity fluctuations due to the unraveling of a complicated initial conformation.

IMPLICATIONS FOR DNA STRAND SEQUENCING

Velocity fluctuations of the translocating molecule present a challenge to nanopore sequencing strategies by complicating the relationship between the elapsed translocation time and the "read position" along a DNA strand. The conformational velocity fluctuations discussed in the previous sections are important only on time scales corresponding to the translocation of a single Kuhn length. For DNA sequencing applications, these velocity fluctuations will be important to take into account, especially if the detector has limited time resolution or sensitivity. For an arbitrarily fast and sensitive detector operating at the much shorter time scales corresponding to successive translocations of individual bases, velocity fluctuations from Brownian motion will dominate the fluctuations over unraveling the initial conformations. This effect is illustrated by the velocity power spectra plotted in Fig. 5. The power spectra are calculated from the model time-dependent velocity profile of a 10 kb molecule with a translocation time near the most probable translocation time (molecule 2 in Fig. 4). The black curve is the power spectrum of conformational velocity fluctuations only, while the gray curve includes Brownian motion (see Appendix S4 for details). The power spectral density of the velocity profile that includes Brownian motion is frequency independent above the typical inverse translocation time of a single Kuhn length segment and dominates conformational velocity fluctuations at high frequencies. Note that under the experimental conditions in this paper, the single-base translocation rate is approximately 30 MHz (by comparison, the maximum bandwidth of the Axopatch 200B amplifier used for these measurements is 60 kHz). The contribution from the Brownian diffusional motion, $S_{thermal} = 4k_b T\overline{v}\left(t\right)/F$, also fluctuates on slow time scales where the conformational velocity fluctuations dominate. The model velocity noise spectrum extending to high frequencies is shown in Fig. 5, which is calculated for the entire duration of the translocation event. It is directly related to the average spatial diffusion constant: $S_{thermal} = 4k_B T\left\langle \overline{v} \right\rangle / F = 4\left\langle D \right\rangle$ up to the damping relaxation frequency of the correlated segments, $\gamma_i / 2\pi m_i \approx 10^{12}$ Hz (29).

The inset to Fig. 5 demonstrates how Brownian velocity fluctuations result in sequencing errors. In the small time window presented here, the conformational velocity fluctuations are minimal (black curve), and Brownian motion dominates (gray curve). When the instantaneous velocity drops below zero, the molecule is moving backward. (Large forward velocities fluctuations can result in skipped bases if the sampling interval is comparable to the rate of successive base passage.) Note that the conformational velocity fluctuations modeled earlier in the paper do not result in backwards molecular motions.

Consider a dsDNA molecule with base pair spacing $a = 0.34$ nm traversing a nanopore with our experimental mean velocity $\overline{v} \approx \langle \overline{v} \rangle = 7.4$ mm/s . The minimum bandwidth required for sequencing is $f_s = \overline{v} / a \approx 22$ MHz . To avoid "re-reading" the same base pair from a backward fluctuation in the motion (Fig. 5, inset) the contributions to the rms thermal velocity fluctuation up to f_s, $\overline{v}_s$, must be smaller than $\overline{v}$. That is,

$\overline{v}^2 \gg \overline{v}_s^2 = S_{thermal} f_s = \left(4k_B T\overline{v} / F\right) \cdot \left(\overline{v} / a\right)$,

or

$$\frac{4k_B T}{Fa} \ll 1. \tag{6}$$

For our experimental conditions a value of ~5 is obtained for the left hand side of Eq. 6. Therefore, Brownian diffusion would not make accurately sequencing of two adjacent base pairs possible under these conditions.

A more precise calculation of the error rate of strand sequencing from Brownian motion can be obtained by using a moving spatial Gaussian diffusion kernel, which describes the probability distribution of the position of a particle undergoing Brownian motion. If x_0 is the known position of the object in question at time $t = 0$, then

$$P\left(x,t\right) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{\left(x - x_0 - \overline{v}t\right)^2}{4Dt}\right). \tag{7}$$

Assume a high spatial resolution base identification mechanism,(such as with a 0.6 nm long graphene nanopore (12)) located at $x = 0$ that reads a base at t = 0. The probability that the next base is in the detection region between $\left[-a/2, a/2\right]$ after a time t has is:

$$P_2\left(t\right) = \int_{-a/2}^{a/2} \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{\left(x + a - \overline{v}t\right)^2}{4Dt}\right) dx \tag{8}$$

The probability this next base is at the detector after a characteristic time $\tau = a / \overline{v}$ is

$$P_2\left(\tau\right) = \int_{-a/2}^{a/2} \frac{1}{\sqrt{4\pi D\tau}} \exp\left(-\frac{x^2}{4D\tau}\right) dx = \mathrm{erf}\left(\frac{1}{2}\sqrt{\frac{a^2}{4D\tau}}\right) = \mathrm{erf}\left(\frac{1}{4}\sqrt{\frac{Fa}{k_B T}}\right) \tag{9}$$

The misread or error probability is:

$$E_2 = 1 - P_2(\tau) = \mathrm{erfc}\sqrt{\frac{Fa}{16k_BT}} \qquad (10)$$

Equation 10 predicts a 75% 'read error' under the conditions in our experiments. These expressions are easily generalized to the *n*th base:

$$E_n = 1 - P_n(\tau) = \mathrm{erfc}\sqrt{\frac{Fa}{16(n-1)k_BT}}\,. \qquad (11)$$

The effective error-free "read length" for a single read of a nanopore-based strand sequencing device is the base number n for which E_n falls below a critical accuracy level, say 95%. These results are plotted in Fig. 6 for $n = 2,\ 4,$ and 6. For $n = 2$, a driving force about 50 times larger than that used in these experiments is necessary to achieve the 95% accuracy level. These results should also be applicable to ssDNA with the appropriate base spacing and driving force.

As shown in Eqs. 6 and 9, the relevant parameter for predicting strand sequencing error rates is the ratio of thermal energy to the work done to translocate the DNA from one base to the next, k_BT / Fa. One way to experimentally reduce the errors introduced by Brownian motion is to increase the driving force *F*. The resulting shorter times available for experimentally identifying bases can then in principle be offset by increasing the viscosity of the solution. It must be noted that simply changing the viscosity to slow down the molecule's motion does not reduce the diffusional contribution. Any reduction in the diffusion constant by increasing the viscosity of the solution is offset by the longer time each base has to diffuse as it passes through the nanopore, which is why there is no viscosity dependence in the sequencing parameter k_BT / Fa.

The importance of considering the Brownian motion of the strand in the nanopore is illustrated in Fig. 7, which demonstrates the limitations Brownian motion imposes on the ability to discriminate between 10 base-pairs with an ideal instrument. We do not consider other important practical issues, such as electronic readout signal to noise or the spatial resolution of the nanopore. Under ideal conditions ($k_BT / Fa \to 0$), the molecule translocates at a constant speed, and the signals from each base will be readily distinguished (Fig. 7*A*). Fig. 7*B*, shows the effect of Brownian motion under our current experimental conditions, denoted by the red circle in Fig. 6. With a 75% read-error, the Brownian fluctuations are dominant within these timescales; we obtain periodic discrimination between base pairs but not enough to generate an accurate sequence in a single read. Increasing the driving force by at least an order of magnitude is sufficient to achieve base discrimination, as shown in Fig. 7*C*. Under these conditions, the accuracy is around 95%, and although variations in the velocity of the molecule still exist, these have been sufficiently suppressed to yield distinguishable signals from each of the 10 bases in our simulated molecule.

The foregoing treatment of Brownian motion in the nanopore may require some modification to very accurately account for the relaxation dynamics of the diffusing DNA strand. During translocation of ssDNA, for example, the translocation time of a single base is similar to molecular relaxation times, and the relaxation dynamics may contribute to the random motion of the strand in the nanopore on these time scales. For dsDNA, the translocation time of a single base is much faster than the molecular relaxation times of the molecule, and we do not expect the

relaxation dynamics to affect the calculated error rate. A kind of effective diffusion, which is influenced by the conformation of the entire molecule, has been described by "fractional Brownian motion" (31-33). It will be interesting to see if further refinement of our model is necessary to account for experimental sequencing results.

Finally we note that other forces than the driving force *F* may be used to overcome the Brownian diffusion. For example, additional large clamping or ratcheting forces acting on the DNA during the read can reduce the errors as appears to have been achieved in Ref. (34) by biochemical means. Control of a single molecule's motion through a nanopore has also been achieved by optical tweezers (35-36), and could potentially be achievable by AFM/nanotube sensors (37), though the practicality of these methods for rapid sequencing is unclear. Also, errors can be reduced by performing multiple reads on identical molecules or on nanopore recaptured molecules as has been reported in Ref. (23).

CONCLUSION

We demonstrate the existence of different mechanisms that give rise to mechanical fluctuation effects for single molecules passing through solid state nanopores. Conformational differences between molecules of the same length determine the spread in their translocation time distribution. These differences in translocation time impact measurements whose goal is to determine molecular lengths. During each translocation event, a simple Langevin model shows that the translocation velocity undergoes large fluctuations as the initial molecular conformation is unraveled. Brownian fluctuations dominate at the short timescales required to achieve rapid single molecule strand sequencing and may be overcome with sufficiently high driving voltage and other applied control forces to the translocating DNA molecule. Realizing base scale sub-nanometer molecular control that overcomes Brownian motion for strand sequencing is a major challenge that may be solved by introducing large forces based on biochemical or physics-based nanotechnology methods. Improvements by error correction based on resequencing may also be anticipated.

APPENDIX A: SIMPLE TRANSLOCATION MODEL

Each initial conformation is generated on a two-dimensional hexagonal lattice as a random walk from a node corresponding to the nanopore. As shown in Fig. 8, the first *m* steps of the random walk are directly away from the nanopore, and the random walk is also excluded from a distance of *m* Kuhn lengths from the membrane. The parameter *m* is called the "center of mass offset" and is a free parameter in our model. The center of mass offset accounts for the fact that the electric field gradient just outside the nanopore will stretch the "equilibrium" conformations as a molecule approaches the nanopore; it also can account for the added drag associated with pushing a short length of the molecule into the fluid on the *trans* side of the membrane, which, although not modeled explicitly, should correspond to the drag from only one or two Kuhn lengths.

The center of mass offset $m = 3$ cited in the main text was chosen to provide the best agreement between the simulated and experimental translocation times. Note that in general the center of

mass offset may depend weakly on both the applied potential and the length of the molecule. As high field gradients distort the "equilibrium" molecular conformation, the center of mass offset will become larger. Long molecules may also contribute additional drag as the molecule is pushed away from the nanopore on the *trans* side of the membrane. As our model is intended for qualitative understanding of the effects of molecular rearrangement on the translocation velocity, we do not consider these effects further.

Because we experimentally study unfolded events, we are interested in only those initial conformations in which an end of the molecule is presented to the nanopore. To generate our initial conformations, we discard all conformations for which the molecule crosses into the excluded membrane region. This procedure is in contrast with the usual method of generating random conformations *in equilibrium*, which involves a reflection principle for random walks that cross the membrane (38).

For the translocation process of a molecule chain, each simulation step describes the motion of one Kuhn length, or *segment*, of dsDNA through the nanopore. The index of the segment passing through the nanopore is *i*. The value of the drag coefficient γ_i depends on the details of the molecular conformation resulting from the previous step and is calculated for use in evaluating the average segment velocity $\overline{v}_i = F / \gamma_i$ and the segment translocation time $\tau_i = l_0 / \overline{v}_i$. The total translocation time for each molecule is $\tau = \sum_i \tau_i$. The detailed algorithm is described in Appendix S1 of the Supplementary Materials.

ACKNOWLEDGMENTS

We thank Dr. Slaven Garaj, Dr. Mike Burns and Professor Dan Branton for helpful discussions and E. Brandin for technical assistance.
L.B. acknowledges support from the State Scholarship Fund of China. D.P.H. acknowledges support from the National Science Foundation. F.A. acknowledges support from a Ruth L. Kirschstein National Research Service Award from the National Institutes of Health (grant no. 1F32HG004692). This work was supported by National Institutes of Health award HG003703.

REFERENCES

1. Li, J., D. Stein, C. McMullan, D. Branton, M. J. Aziz, and J. A. Golovchenko. 2001. Ion-beam sculpting at nanometre length scales. Nature 412:166-169.
2. Li, J., M. Gershow, D. Stein, E. Brandin, and J. A. Golovchenko. 2003. DNA molecules and configurations in a solid-state nanopore microscope. Nature Mater. 2:611-615.
3. Storm, A. J., J. H. Chen, H. W. Zandbergen, and C. Dekker. 2005. Translocation of double-strand DNA through a silicon oxide nanopore. Phys. Rev. E 71:051903.
4. Fologea, D., M. Gershow, B. Ledden, D. S. McNabb, J. A. Golovchenko, and J. Li. 2005. Detecting single stranded DNA with a solid state nanopore. Nano Lett. 5:1905-1909.
5. Meller, A., L. Nivon, and D. Branton. 2001. Voltage-driven DNA translocations through a nanopore. Phys. Rev. Lett. 86:3435-3438.

6. Bates, M., M. Burns, and A. Meller. 2003. Dynamics of DNA molecules in a membrane channel probed by active control techniques. Biophys. J. 84:2366-2372.
7. Wiggin, M., C. Tropini, V. Tabard-Cossa, N. N. Jetha, and A. Marziali. 2008. Nonexponential Kinetics of DNA Escape from alpha-Hemolysin Nanopores. Biophys. J. 95:5317-5323.
8. Branton, D., D. W. Deamer, A. Marziali, H. Bayley, S. A. Benner, T. Butler, M. Di Ventra, S. Garaj, A. Hibbs, X. H. Huang, S. B. Jovanovich, P. S. Krstic, S. Lindsay, X. S. S. Ling, C. H. Mastrangelo, A. Meller, J. S. Oliver, Y. V. Pershin, J. M. Ramsey, R. Riehn, G. V. Soni, V. Tabard-Cossa, M. Wanunu, M. Wiggin, and J. A. Schloss. 2008. The potential and challenges of nanopore sequencing. Nat. Biotechnol. 26:1146-1153.
9. DeGuzman, V. S., C. C. Lee, D. W. Deamer, and W. A. Vercoutere. 2006. Sequence-dependent gating of an ion channel by DNA hairpin molecules. Nucleic Acids Res. 34:6425-6437.
10. Butler, T. Z., M. Pavlenok, I. M. Derrington, M. Niederweis, and J. H. Gundlach. 2008. Single-molecule DNA detection with an engineered MspA protein nanopore. Proc. Natl. Acad. Sci. U. S. A. 105:20647-20652.
11. Stoddart, D., A. J. Heron, E. Mikhailova, G. Maglia, and H. Bayley. 2009. Single-nucleotide discrimination in immobilized DNA oligonucleotides with a biological nanopore. Proc. Natl. Acad. Sci. U. S. A. 106:7702-7707.
12. Garaj, S., W. Hubbard, A. Reina, J. Kong, D. Branton, and J. A. Golovchenko. 2010. Graphene as a subnanometre trans-electrode membrane. Nature 467:190-U173.
13. Merchant, C. A., K. Healy, M. Wanunu, V. Ray, N. Peterman, J. Bartel, M. D. Fischbein, K. Venta, Z. T. Luo, A. T. C. Johnson, and M. Drndic. 2010. DNA Translocation through Graphene Nanopores. Nano Lett. 10:2915-2921.
14. Schneider, G. F., S. W. Kowalczyk, V. E. Calado, G. Pandraud, H. W. Zandbergen, L. M. K. Vandersypen, and C. Dekker. 2010. DNA Translocation through Graphene Nanopores. Nano Lett. 10:3163-3167.
15. Fologea, D., E. Brandin, J. Uplinger, D. Branton, and J. Li. 2007. DNA conformation and base number simultaneously determined in a nanopore. Electrophoresis 28:3186-3192.
16. Storm, A. J., C. Storm, J. Chen, H. Zandbergen, J. F. Joanny, and C. Dekker. 2005. Fast DNA translocation through a solid-state nanopore. Nano Lett. 5:1193-1197.
17. Grosberg, A. Y., S. Nechaev, M. Tamm, and O. Vasilyev. 2006. How long does it take to pull an ideal polymer into a small hole? Phys. Rev. Lett. 96:228105.
18. Zoia, A., A. Rosso, and S. N. Majumdar. 2009. Asymptotic Behavior of Self-Affine Processes in Semi-Infinite Domains. Phys. Rev. Lett. 102:120602.
19. Sakaue, T. 2010. Sucking genes into pores: Insight into driven translocation. Phys. Rev. E 81:041808.
20. Forrey, C., and M. Muthukumar. 2007. Langevin dynamics simulations of ds-DNA translocation through synthetic nanopores. J. Chem. Phys. 127:015102.
21. Kong, C. Y., and M. Muthukumar. 2002. Modeling of polynucleotide translocation through protein pores and nanotubes. Electrophoresis 23:2697-2703.
22. Lathrop, D. K., E. N. Ervin, G. A. Barrall, M. G. Keehan, R. Kawano, M. A. Krupka, H. S. White, and A. H. Hibbs. 2010. Monitoring the Escape of DNA from a Nanopore Using an Alternating Current Signal. J. Am. Chem. Soc. 132:1878-1885.
23. Gershow, M., and J. A. Golovchenko. 2007. Recapturing and trapping single molecules with a solid-state nanopore. Nature Nanotechnol. 2:775-779.

24. 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.
25. Storm, A. J., J. H. Chen, X. S. Ling, H. W. Zandbergen, and C. Dekker. 2003. Fabrication of solid-state nanopores with single-nanometre precision. Nature Mater. 2:537-540.
26. Kim, M. J., B. McNally, K. Murata, and A. Meller. 2007. Characteristics of solid-state nanometre pores fabricated using a transmission electron microscope. Nanotechnology 18:205302.
27. Grosberg, A. I. U., and A. R. Khokhlov. 1994. Statistical physics of macromolecules. AIP Press, New York.
28. Meller, A., L. Nivon, E. Brandin, J. Golovchenko, and D. Branton. 2000. Rapid nanopore discrimination between single polynucleotide molecules. Proc. Natl. Acad. Sci. U. S. A. 97:1079-1084.
29. Wannier, G. H. 1966. Statistical physics. Wiley, New York.
30. van Dorp, S., U. F. Keyser, N. H. Dekker, C. Dekker, and S. G. Lemay. 2009. Origin of the electrophoretic force on DNA in solid-state nanopores. Nature Physics 5:347-351.
31. Kantor, Y., and M. Kardar. 2007. Anomalous diffusion with absorbing boundary. Phys. Rev. E 76.
32. Panja, D. 2010. Anomalous polymer dynamics is non-Markovian: memory effects and the generalized Langevin equation formulation. Journal of Statistical Mechanics-Theory and Experiment.
33. Dubbeldam, J. L. A., V. G. Rostiashvili, A. Milchev, and T. A. Vilgis. 2011. Fractional Brownian motion approach to polymer translocation: The governing equation of motion. Phys. Rev. E 83.
34. Leiberman, K. R., G. M. Cherf, M. J. Doody, F. Olasagasti, Y. Kolodji, and M. Akeson. 2010. Processive Replication of Single DNA Molecules in a Nanopore Catalyzed by phi29 DNA Polymerase. J. Am. Chem. Soc. 132:17961–17972.
35. Keyser, U. F., B. N. Koeleman, S. Van Dorp, D. Krapf, R. M. M. Smeets, S. G. Lemay, N. H. Dekker, and C. Dekker. 2006. Direct force measurements on DNA in a solid-state nanopore. Nature Physics 2:473-477.
36. Trepagnier, E. H., A. Radenovic, D. Sivak, P. Geissler, and J. Liphardt. 2007. Controlling DNA capture and propagation through artificial nanopores. Nano Lett. 7:2824-2830.
37. King, G. M., and J. A. Golovchenko. 2005. Probing nanotube-nanopore interactions. Phys. Rev. Lett. 95:216103.
38. Wu, D. C., and J. A. Kang. 1996. Conformational statistics of polymer chain terminally attached to wall .1. NRW model tail chain. Science in China Series B-Chemistry 39:608-617.

FIGURE LEGENDS

FIGURE 1. Geometry of DNA nanopore translocation experiment (schematic) and model. The DNA is rendered as a molecule with smooth conformational changes (gray) and as a freely jointed chain used in the model (solid and dashed lines). In the model, a "pivot point" is defined as the (changing) point on the molecule nearest to the nanopore and beyond which its conformation is unchanged as each Kuhn length is driven through the nanopore.

FIGURE 2. Translocation time distributions. (*A*) Density histogram of translocation events for 5.3 kb and 10 kb DNA. Inset shows typical time traces for folded and unfolded events. (*B*) Translocation time histogram of 1350 unfolded events; error bars represent the counting error in each histogram bin. The bin size is 4.9 μs for the 5.3 kb data and 7.8 μs for the 10 kb data. Solid and dashed curves are the predictions of the 2D models with and without a 300 nm center of mass offset, respectively.

FIGURE 3. Distribution of translocation times as a function of initial center of mass position from the nanopore. Parameters are those used in the solid curves in Fig. 2.

FIGURE 4. Examples of velocity and translocation time fluctuations from modeling results of 10 kb DNA molecules. (*A*) Initial conformations of three otherwise identical molecules with short, average, and long translocation times. (*B*) Modeled distribution of 10 kb molecules showing the translocation times of the conformations in *A*. (*C*) Velocity profiles of these three molecules during the translocation event.

FIGURE 5. Power spectral density $S_v(f)$ of the modeled translocation velocity profile of a 10 kb DNA molecule. The gray power spectrum includes Brownian motion, while the black curve represents velocity fluctuations from unraveling the initial conformation only. Power spectra were calculated with a Blackman window function and slightly smoothed for clarity. The upper limit of frequency, 25 MHz, is similar to the single base translocation rate in our experiments. (Inset) 4 μs of a simulated translocation sampled at 50 MHz. The solid line represents the average velocity in this time window, while the dashed line shows zero velocity.

FIGURE 6. The sequencing probability success rate depends on the ratio of thermal motion to the work done by the driving force to drive a single base through the nanopore. The solid circle indicates current experimental conditions. The dashed circle represents the force necessary to achieve a 95% single-read success rate for two bases.

FIGURE 7. A 10-base DNA strand passes through a detector with single base resolution under various experimental conditions. Time units are normalized to the average base translocation time $\tau = a/\overline{v}$. (*A*) If the molecule traverses the detector with a constant speed, the signal recorded will be clear. (*B*) With Brownian motion with the current experimental driving force,

the recorded signal is full of errors. (*C*) By applying a driving force 50 times larger to suppress the relative motion fluctuations (and increasing the solution viscosity by a factor of 50 to maintain the same velocity), the read is much improved.

FIGURE 8. Illustration of the center of mass offset m used to generate the initial molecule conformations. The first m segments are straight; the remaining $N - m$ segments undergo a random walk but are excluded from a distance m from the membrane.

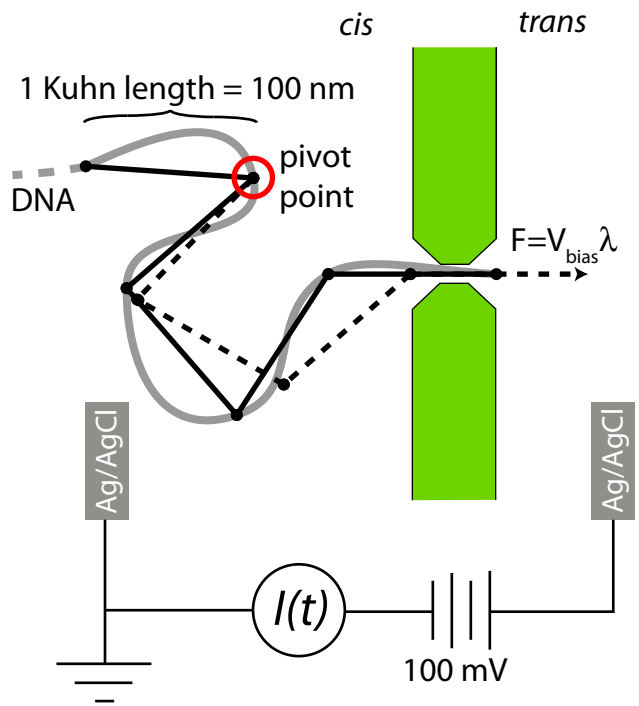

cis
trans
1 Kuhn length = 100 nm
pivot
point
DNA
F=V_bias λ
Ag/AgCl
Ag/AgCl
I(t)
100 mV

Figure 1

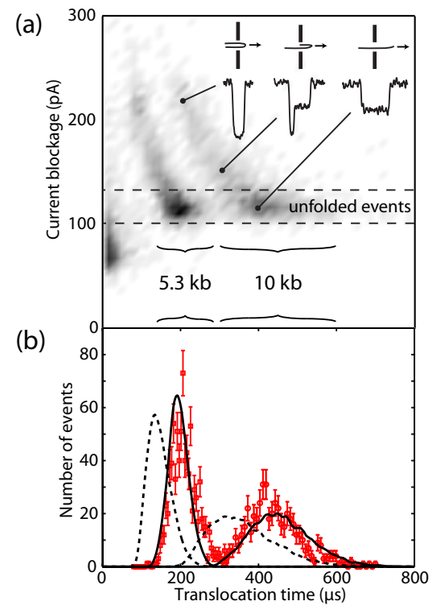
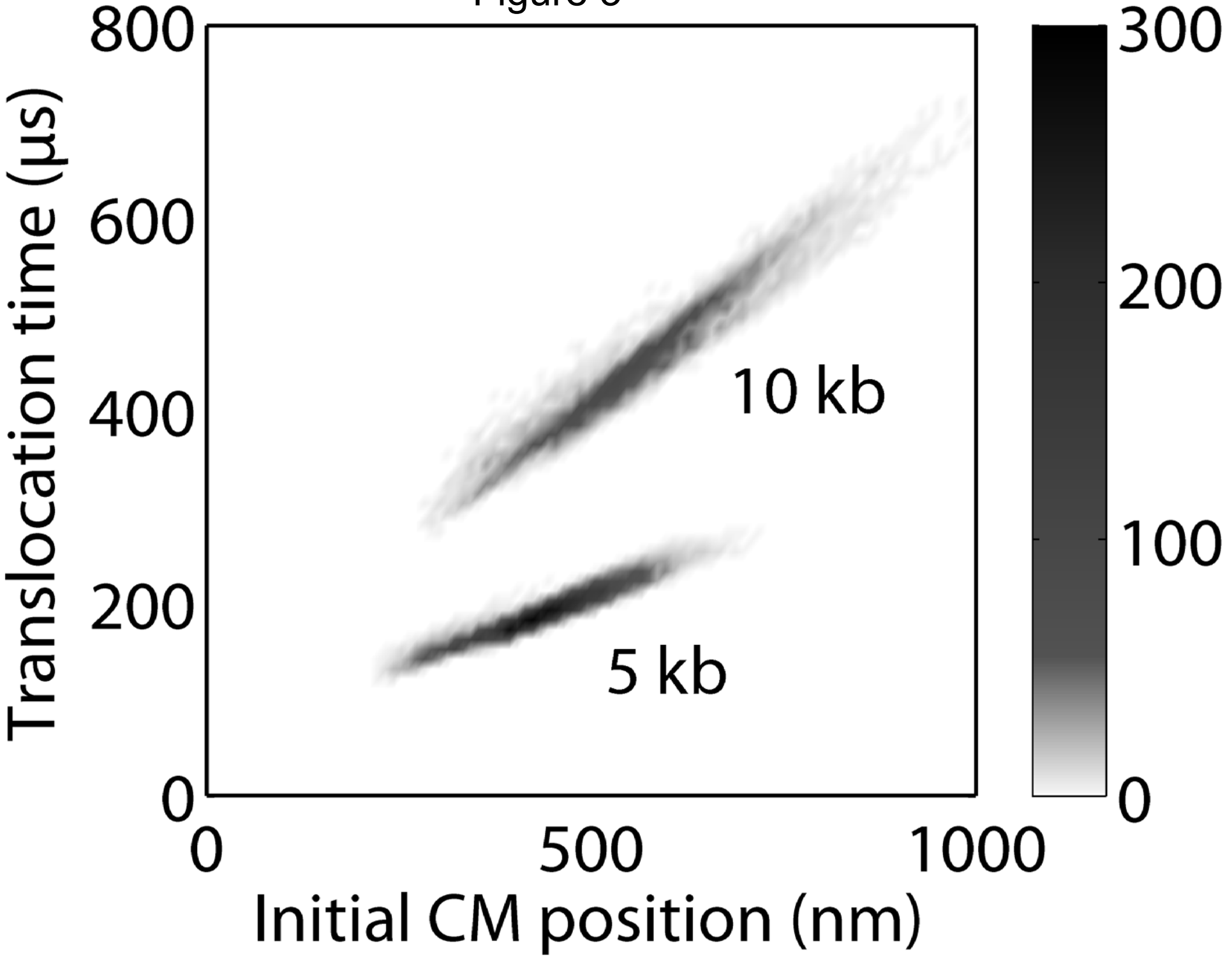

(a)
300
200
100
0
Current blockage (pA)
unfolded events
5.3 kb
10 kb
(b)
80
60
40
20
0
Number of events
0
200
400
600
800
Translocation time (μs)

Figure 2

Figure 3


800
600
400
200
0
Translocation time (μs)
0
500
1000
Initial CM position (nm)
10 kb
5 kb
300
200
100
0

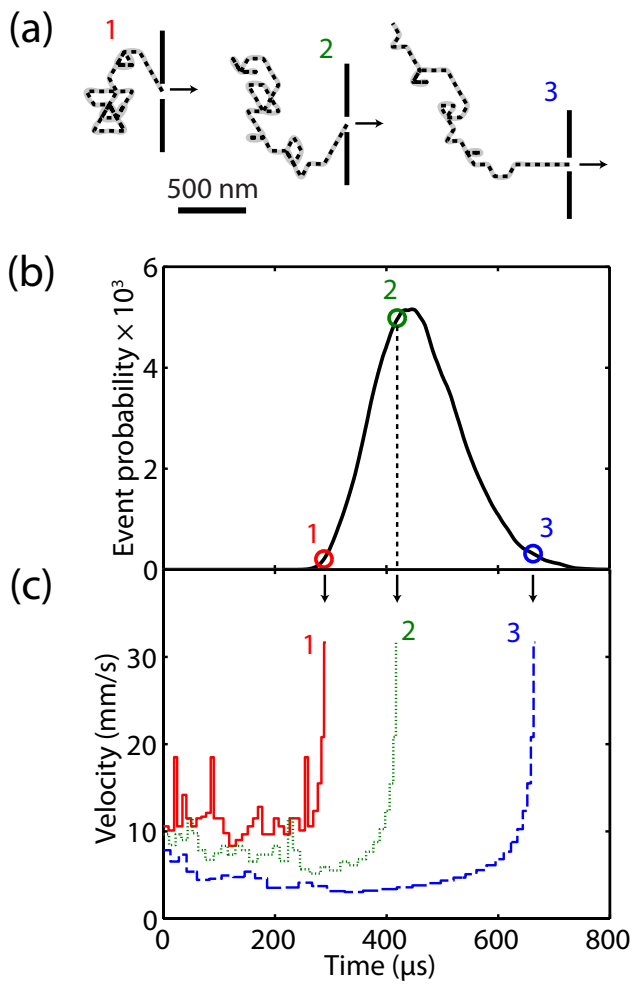

(a)
1
2
3
500 nm
(b)
Event probability × 10³
6
4
2
0
1
2
3
(c)
Velocity (mm/s)
30
20
10
0
1
2
3
0
200
400
600
800
Time (μs)

Figure 4

Figure 5

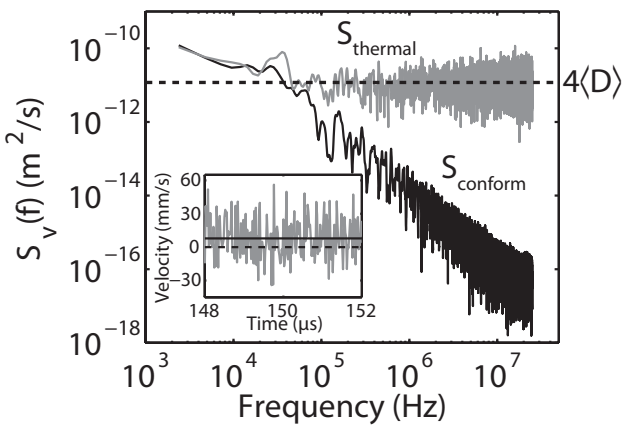

$S_{thermal}$
$4\langle D\rangle$
$S_{conform}$
$S_v(f)$ (m^2/s)
10^{-10}
10^{-12}
10^{-14}
10^{-16}
10^{-18}
10^3
10^4
10^5
10^6
10^7
Frequency (Hz)
Velocity (mm/s)
60
30
0
−30
148
150
152
Time (μs)

Figure 6

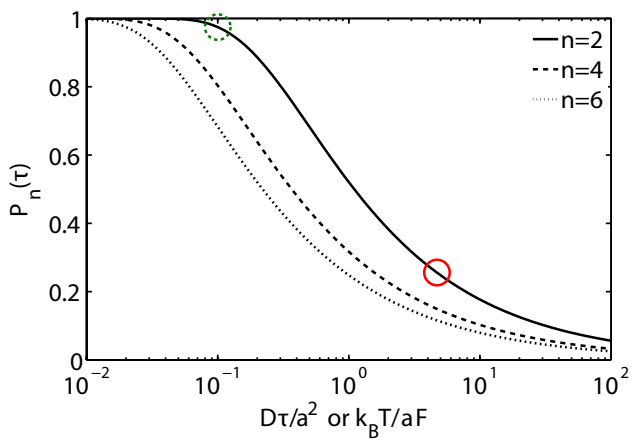
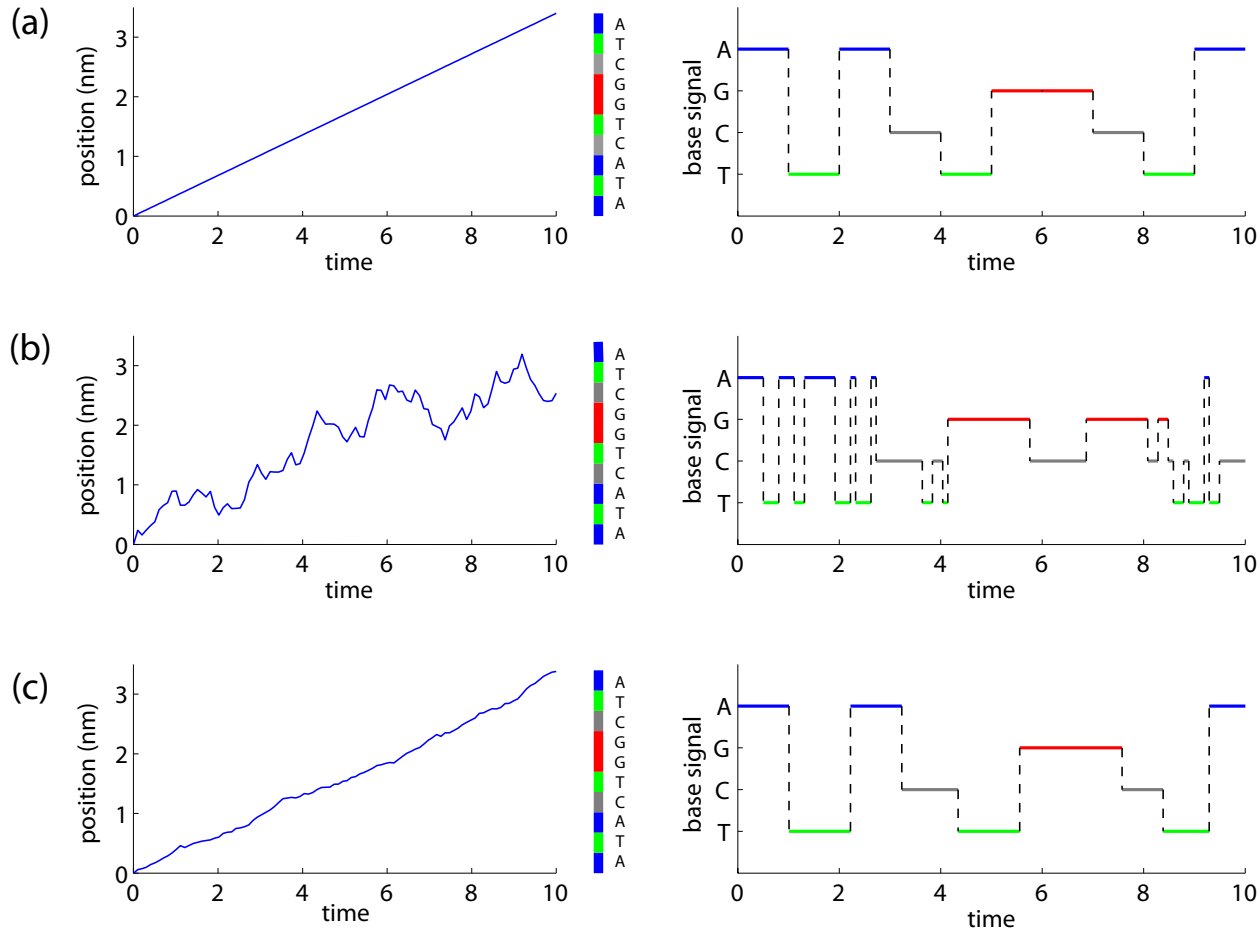


Figure 7


(a)
position (nm)
time
A T C G G T C A T A
base signal
A
G
C
T
(b)
position (nm)
time
A T C G G T C A T A
base signal
A
G
C
T
(c)
position (nm)
time
A T C G G T C A T A
base signal
A
G
C
T

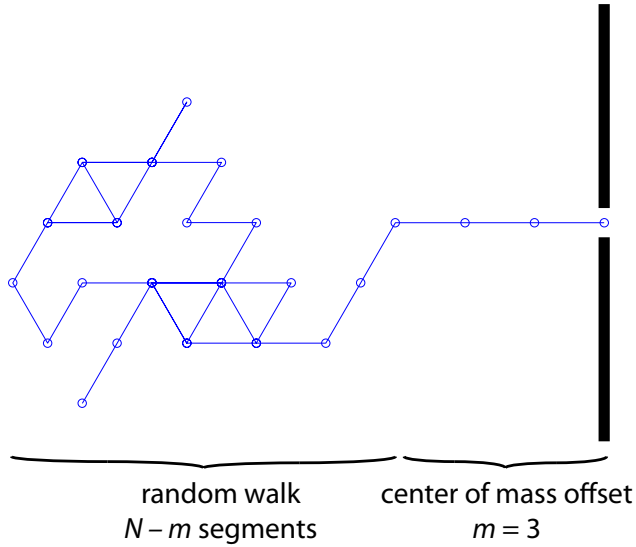


Figure 8

Supplementary Material for:

Origins and Consequences of Velocity Fluctuations during DNA Passage through a Nanopore

Lu Bo[†, ‡], Fernando Albertorio[†], David P. Hoogerheide[†], and Jene A. Golovchenko[†*]

[†]Department of Physics, Harvard University, Cambridge MA 02138

[‡]Electron Microscopy Laboratory, Department of Physics, Peking University, Beijing, China

*Correspondence: golovchenko@physics.harvard.edu

The following additional appendices deal with various aspects of the modeling results presented in this paper. Appendix S1 contains all the details of the model of molecule unraveling. Appendix S2 deals with the question of dimensionality and concludes that changes in the dimensionality of the model will not be significant enough to alter the conclusions of our paper. Appendix S3 contains an extensive discussion of the magnitude of the driving force and the estimated magnitude of fluctuations in this force during translocation. Finally, Appendix S4 details how Brownian motion is added to our translocation model. All references refer to the references at the end of the supplementary material.

APPENDIX S1: DETAILS OF THE TRANSLOCATION MODEL

For the translocation process of a molecule chain, each simulation step describes the motion of one Kuhn length, or *segment*, of dsDNA through the nanopore. The index of the segment passing through the nanopore is *i*. The value of the drag coefficient γ_i depends on the details of the molecular conformation resulting from the previous step and is calculated for use in evaluating the average segment velocity $\overline{v}_i = F / \gamma_i$ and the segment translocation time $\tau_i = l_0 / \overline{v}_i$. The total translocation time for each molecule is $\tau = \sum_i \tau_i$.

The drag coefficient is calculated for the motions of the untranslocated dragged segments as they are advanced towards the nanopore. Three steps of the algorithm for updating a 14-segment molecule's geometry are depicted in Fig. S1, *A-D*. In Fig. S1*A*, we have chosen $i = 1$ for clarity. In the following discussion, a *vertex* is a joint between two segments of the molecule. The vertex number is the same as the number of the segment immediately preceding the vertex. For example, in Fig. S1*A*, the vertex to the right of segment 1 is vertex 1.

In general, the molecule will have three distinct regions of motion:

1. Part of the molecule, between the nanopore and the region of the molecule that is unraveling, is fully stretched; that is, each segment follows its preceding segment toward the nanopore without lateral motion. The vertex after the last segment corresponding to the fully stretched region we call the *unraveling point.* The unraveling points are labeled by the broken green circles in Fig. S1, *A-D*. When moving to a subsequent step, the molecular conformation between the nanopore and the current step's unraveling point remains unchanged, even though the identity of the segments in each position of this region has changed. The unraveling point changes with each step and can even move backward along the molecule.

2. Part of the molecule will be undragged. We have already introduced the *pivot point*, which separates the dragged and undragged regions of the molecule. In Fig. S1, *A-D*, the pivot points are denoted by a solid red circle. When moving to a subsequent step, the molecular conformation between the current step's pivot point and the end of the molecule remains unchanged. The pivot point, however, does change with each step but never moves backward along the molecule.
3. In the domain between the unraveling point and the pivot point, the molecule will undergo a complicated rearrangement described next.

Our simulation algorithm begins each step by identifying this third domain, which is between the unraveling and pivot points. The requirement for a domain between two vertices j and k to be rearranged during step i is that the end-to-end distance between j and k *along the molecule* (D_{jk}) must differ from the end-to-end separation of j and k *along the lattice* (L_{jk}) by either 1 or 2 Kuhn lengths. In practice, the first domain satisfying this requirement is identified by sequentially searching the vertices $k \geq i+2$ until a vertex j between i and k can be found such that $D_{jk} - L_{jk} = 1 \text{ or } 2$. The point k is then the pivot point and j is the unraveling point. Because j and k bracket the domain closest to the nanopore satisfying the rearrangement requirement, this algorithm selects the region to be updated such that the net motion of all molecule segments is minimized while allowing segment i to pass through the nanopore.

Once the unraveling and pivot points have been found, the motions of each segment can be determined. For region 1, between the nanopore and the unraveling point, each segment before the unraveling point occupies its leading segment's previous position. For region 2, after the pivot point, the segments do not move. The only complicated motion occurs between the unraveling point *j* and the pivot point *k*. The condition $D_{jk} - L_{jk} = 1$ corresponds to a geometry in which the segments between the unraveling and pivot points are not overlapped, as in Fig. S1*A*. The new molecule conformation between the unraveling point and the pivot point will describe a straight line from the unraveling point to the pivot point, as shown by comparing Fig. S1*A*, to Fig.

S1*B*. Another motion of this type occurs between Fig. S1*C*, and Fig. S1*D*. The condition $D_{jk} - L_{jk} = 2$, on the other hand, corresponds to the overlapped case in Fig. S1*B*. In this case, both segments must rotate away from their initial position, and the segment closer to the nanopore also undergoes a translation, as seen by comparing Fig. S1*B* and Fig. S1*C*. When ambiguous, as between Fig. S1*B* and Fig. S1*C*, the direction the segments rotate does not make a significant difference in the translocation time results and is chosen randomly.

The drag coefficient is calculated from the motion of each segment. It can be expressed as

$$\gamma_i = 2\pi\eta l_0 \sum_{n \geq i} C_n \,, \qquad \text{(S1)}$$

where η is the solution viscosity and C_n is a dimensionless geometric factor representing the contribution of the hydrodynamic drag from the motion of the *n*th segment as the *i*th segment is pulled through the nanopore. The C_n are determined from the hydrodynamics of a rigid cylinder of diameter 2 nm and length 100 nm (one Kuhn length) at low Reynolds number (1) and can take different values depending on the motion of the segment. Possible motions with their drag contributions are: axial translation, or $n \leq j$ ($C_n = 0.27$); perpendicular translation (0.44); 60° rotation (0.22); 120° rotation (0.44); no motion (0). The effect of the nanopore walls on the segment in the nanopore is calculated from a finite element integration of the viscous force on a 2 nm cylinder in an hourglass-shaped pore of minimum radius 10 nm that approximates the geometry of our system (2); in this case $C_i = 0.52$. If C_i is chosen to be simply the drag coefficient for axial translation (0.27), the effective charge density required to fit the experimental data changes by about 10%, from 0.22 e^-/bp to 0.20 e^-/bp. For segments that experience both rotation and translation, the order of these operations is chosen to minimize the drag coefficient, though this choice has little effect on the translocation time distributions. We assume γ_i does not change and the velocity is constant during the translocation of a single segment through the nanopore. The translocation time of one molecule chain is calculated by summing the translocation times of the segments, $\tau_i = l_0 / \overline{v}_i = l_0 \gamma_i / F$.

The entire algorithm was implemented with custom MATLAB code.

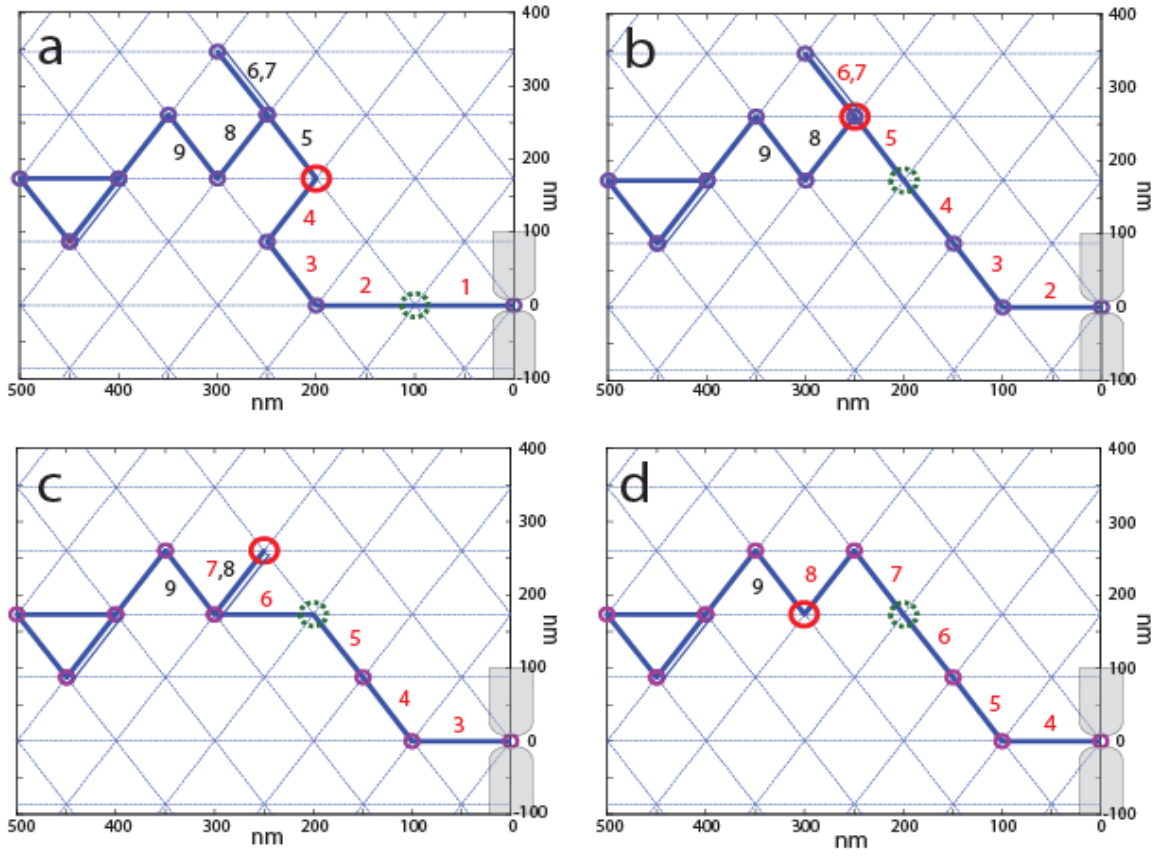


FIGURE S1. Three steps of the simulated translocation process. The DNA molecule is rendered as a freely jointed chain (blue lines; double blue lines are overlapping segments) on a two-dimensional hexagonal lattice. Each small purple circle indicates a freely rotating joint, or vertex, between two adjacent segments. Beyond the pivot point (red circle, black labels) the molecule's conformation will remain stationary on the hexagonal grid as one Kuhn length is pulled into the nanopore, while the segments with colored labels contribute to the drag force. Green, dashed circles denote the unraveling points.

APPENDIX S2. EFFECT OF DIMENSIONALITY

We expect to see some difference between a simple two-dimensional (2D) model and a full three-dimensional (3D) simulation due to the different distributions of center-of-mass distance to the nanopore (see Fig. 3 in the main text). We explored the effect of dimensionality by modifying a one-dimensional (1D) random walk model, in which the initial conformation was generated as a random walk with forward, backward, and stationary (transverse) directions. This model can be considered as quasi-2D projected to 1D if it includes one transverse dimension (with two additional directions for the random walk) and as quasi-3D if there are two transverse dimensions (each with two possible random walk directions). The translocation time distributions from the quasi-2D and quasi-3D models are compared to the results from the hexagonal 2D model in Fig. S2. As expected, because of the additional spatial degrees of freedom in the quasi-3D model, the center-of-mass distribution (and hence the translocation time) is concentrated closer to the nanopore. The differences among the three methods are not great enough to alter our conclusion that the unraveling model is sufficient to demonstrate the physics underlying the experimentally observed translocation time distributions.

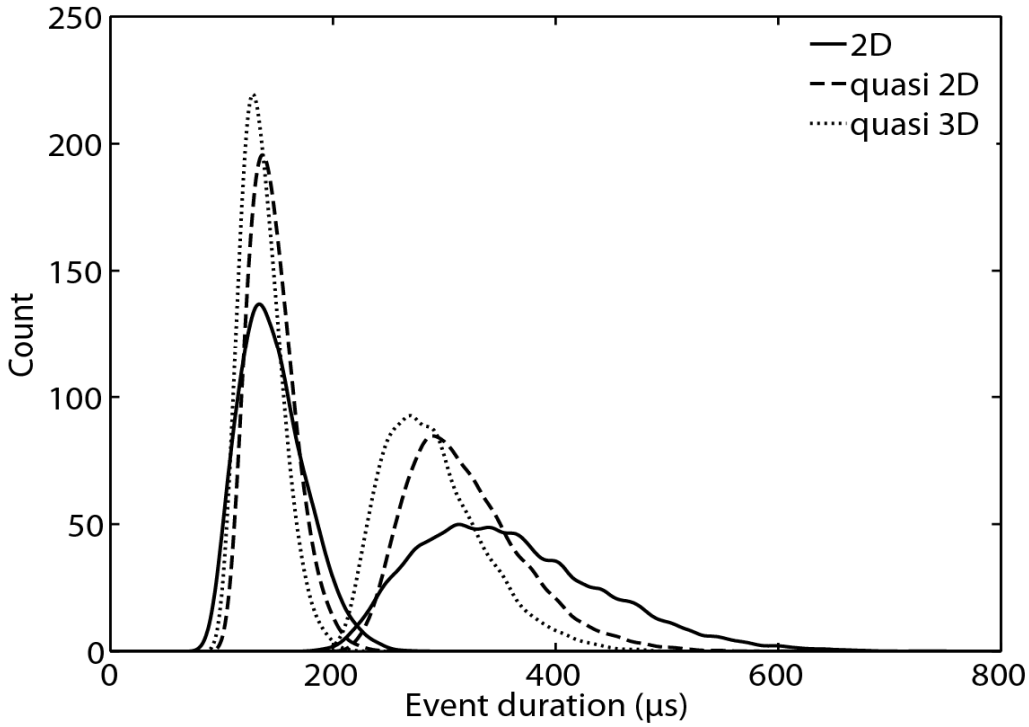


FIGURE S2. Predicted translocation time distributions from the hexagonal 2D, quasi-2D, and quasi-3D models without a center of mass offset. Curves are normalized to the same integrated area.

APPENDIX S3. THE DRIVING FORCE AND ITS FLUCTUATIONS

In the model, we assume that the driving force is constant. Here we discuss this assumption in more detail and estimate the driving force fluctuations if the molecule is not axially centered in the nanopore. In the absence of radial forces in the nanopore (a reasonable assumption more than a Debye length (0.3 nm in 1 M KCl) away from the side walls), one expects such fluctuations to occur on time scales corresponding to the radial diffusion time of a single Kuhn length in the nanopore. The relevant diffusion constant is 13×10^{-12} m^2/s (3); a typical correlation time is $R^2/D \sim 2\ \mu\mathrm{s}$, if $R \approx 5$ nm is the radius of the nanopore. This time scale is fast enough so that if the magnitude of the difference in driving force at different radial positions in the channel is large enough, the fluctuation of the driving force could in principle account for some of the observed variance in the translocation time.

The magnitude of the driving force depends primarily on the effective charge density of the DNA backbone. The bare charge of the DNA backbone is 2 electron charges per basepair (e^-/bp), and this is the charge on which the applied potential acts. The applied potential also acts on counterions that are attracted to the negative charge surface of the SiN nanopore and to the negative charge of the double strand DNA backbone. The DNA is pulled against the resulting electroosmotic flow (EOF) and moves slower than it would in the absence of the counterion flow (4). The resulting retarded motion of the DNA can be described by an effective charge density that is less than $2\ e^-/\mathrm{bp}$. Literature values for this charge density vary widely and have been measured to be anywhere from $0.10\ e^-/\mathrm{bp}$ to $0.9\ e^-/\mathrm{bp}$ (5-7). In nanopores, the EOF, and hence the effective charge, depend strongly on the surface charge state of the nanopore and therefore can be different for nanopores with different fabrication, cleaning, or wetting histories (4, 8).

The value of the charge density can be estimated from the free electrophoretic mobility measured by Nkodo et al. (9), $\mu_0 = 4.1\times10^{-8}$ m^2 / V·s . This mobility is simply the ratio of the total charge of the molecule to the total drag of the molecule, and it is found to be independent of length. If we consider a polymer as a freely jointed chain in thermal equilibrium, the total charge on the molecule is simply $Nl_0\lambda$, where N is the number of Kuhn length segments. The drag can be estimated from the fact that on average 1/3 of the molecule segments are aligned with an electric field, while 2/3 of them are transverse to the field. The total molecule drag can then be estimated from the axial and transverse drag coefficients, $\gamma_{\|}$ and $\gamma_{\perp}$, to be $\gamma = N\left(\gamma_{\|} + 2\gamma_{\perp}\right)/3$. We then obtain $\lambda = \mu_0\gamma / l_0 \approx 0.21\ e^{-}$ / bp , in good agreement with both the literature and with the charge densities found by fitting our model to the experimental result.

Different radial positions of the molecule give rise to different EOF fields in the nanopore and could in principle significantly affect the effective charge density. A simple estimate, however, shows that such fluctuations should be minimal. The EOF is generated by the electrical field acting on the region of positive charge density close to the negatively charged wall surface, so the radial gradient of the axial velocity exists only within a distance of the Debye length from the wall. With an electrolyte concentration of 1 M, the Debye length is less than 1 nm, so the enhanced drag from the EOF should be nearly constant over the width of the nanopore. We therefore expect that the effective charge density will be stable if the molecule undergoes radial position fluctuations.

APPENDIX S4: INCLUSION OF BROWNIAN MOTION IN THE COMPUTER MODEL.

The probability distribution of the one-dimensional diffusional displacement Δx of a Brownian particle with diffusion constant D after a time Δt, given an initial position $\Delta x = 0$ and uniform drift velocity $\bar{v}$, is the one-dimensional moving Gaussian kernel

$$P(\Delta x; \Delta t) = \frac{1}{\sqrt{4\pi D \Delta t}} \exp\left(-\frac{(\Delta x - \bar{v}\Delta t)^2}{4D\Delta t} \right). \quad \text{(S2)}$$

If we define the apparent diffusional velocity of the Brownian object to be $\tilde{v} = \Delta x / \Delta t$, we can write the probability distribution of the apparent velocity as

$$P(\tilde{v}; \Delta t) = \frac{\Delta t}{\sqrt{4\pi D \Delta t}} \exp\left(-\frac{(\tilde{v} - \bar{v})^2}{4D / \Delta t} \right), \quad \text{(S3)}$$

which is a Gaussian kernel centered on the uniform velocity. We use the probability distribution given in Eq. S3 to include the effects of Brownian motion in our model results. In this case, the time interval in question is the sampling interval of a hypothetical instrument, which we take to be $\Delta t = 20$ ns, the minimum sampling rate required for detecting single bases at the average molecular speeds determined by our experiments. The actual velocity of the DNA strand in the nanopore is then randomly generated from the distribution equation, Eq. S3. The velocity $\bar{v}$ and the diffusion constant $D = k_B T \bar{v} / F$ are simply those obtained from the previously described model. They vary as the molecule assumes different conformations while being pulled through the nanopore, but are assumed constant over the much smaller time scale Δt.

REFERENCES

S1. Batchelor, G. K. 1970. Slender-Body Theory for Particles of Arbitrary Cross-Section in Stokes Flow. *J. Fluid Mech.* 44:419.

S2. Kim, M. J., B. McNally, K. Murata, and A. Meller. 2007. Characteristics of solid-state nanometre pores fabricated using a transmission electron microscope. *Nanotechnology* 18:205302.

S3. Doi, M., and S. F. Edwards. 1986. The Theory of Polymer Dynamics. Oxford University Press, New York.

S4. van Dorp, S., U. F. Keyser, N. H. Dekker, C. Dekker, and S. G. Lemay. 2009. Origin of the electrophoretic force on DNA in solid-state nanopores. *Nature Physics* 5:347-351.

S5. Keyser, U. F., B. N. Koeleman, S. Van Dorp, D. Krapf, R. M. M. Smeets, S. G. Lemay, N. H. Dekker, and C. Dekker. 2006. Direct force measurements on DNA in a solid-state nanopore. *Nature Physics* 2:473-477.

S6. Schellman, J. A., and D. Stigter. 1977. Electrical Double-Layer, Zeta Potential, and Electrophoretic Charge of Double-Stranded DNA. *Biopolymers* 16:1415-1434.

S7. Smith, S. B., and A. J. Bendich. 1990. Electrophoretic Charge-Density and Persistence Length of DNA as Measured by Fluorescence Microscopy. *Biopolymers* 29:1167-1173.

S8. Ghosal, S. 2007. Electrokinetic-flow-induced viscous drag on a tethered DNA inside a nanopore. *Phys. Rev. E* 76:061916.

S9. Nkodo, A. E., J. M. Garnier, B. Tinland, H. J. Ren, C. Desruisseaux, L. C. McCormick, G. Drouin, and G. W. Slater. 2001. Diffusion coefficient of DNA molecules during free solution electrophoresis. *Electrophoresis* 22:2424-2432.