Velocity of DNA during translocation through a solid state nanopore
Calin Plesa, Nick van Loo, Philip Ketterer, Hendrik Dietz, and Cees Dekker
Nano Lett., Just Accepted Manuscript • DOI: 10.1021/nl504375c • Publication Date (Web): 11 Dec 2014
Downloaded from http://pubs.acs.org on December 15, 2014

Just Accepted

“Just Accepted” manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides “Just Accepted” as a free service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. “Just Accepted” manuscripts appear in full in PDF format accompanied by an HTML abstract. “Just Accepted” manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are accessible to all readers and citeable by the Digital Object Identifier (DOI®). “Just Accepted” is an optional service offered to authors. Therefore, the “Just Accepted” Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the “Just Accepted” Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these “Just Accepted” manuscripts.
Velocity of DNA during translocation through a solid state nanopore

Calin Plesa†, Nick van Loo†, Philip Ketterer‡, Hendrik Dietz‡, Cees Dekker†*.

† Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands.

‡ Physics Department, Walter Schottky Institute, Technische Universität München, Am Coulombwall 4a, 85748 Garching near Munich, Germany.

*Corresponding author. E-mail: c.dekker@tudelft.nl
ABSTRACT: While understanding translocation of DNA through a solid-state nanopore is vital for exploiting its potential for sensing and sequencing at the single-molecule level, surprisingly little is known about the dynamics of the propagation of DNA through the nanopore. Here we use linear double-stranded DNA molecules, assembled by the DNA origami technique, with markers at known positions in order to determine, for the first time, the local velocity of different segments along the length of the molecule. We observe large intramolecular velocity fluctuations, likely related to changes in the drag force as the DNA blob unfolds. Furthermore we observe an increase in the local translocation velocity towards the end of the translocation process, consistent with a speeding up due to unfolding of the last part of the DNA blob. We use the velocity profile to estimate the uncertainty in determining the position of a feature along the DNA given its temporal location, and demonstrate the error introduced by assuming a constant translocation velocity.

KEYWORDS: nanopore, DNA, velocity, translocation, DNA origami
The number of solid-state nanopore applications has exploded since their inception over a decade ago. Nanopores have been used to study and detect various biopolymers (particularly DNA), free protein, DNA-origami, and DNA-protein complexes. Despite this advance, some basic properties of the translocation process remain poorly understood. One such fundamental question is how the velocity of a long DNA molecule changes as the molecule translocates through the pore. Previous simulations from the Golovchenko group suggested that the wide distribution observed for the translocation velocities for molecules of equal length can be attributed to drag-induced velocity fluctuations and predicted that the DNA would strongly speed up at the end of the translocation process, as described in detail later. Currently, no experimental measurement of the velocity profile has been carried out. Since knowledge of this velocity profile is required to convert temporal signals into positional information, crucial for obtaining biological information from a translocating DNA molecule, this represents a serious gap in our current understanding of the translocation process.

In this study we introduce a new technique to determine the velocity of different segments along a DNA molecule, and we present the first experimental data on how the velocity changes during the course of a translocation event. We systematically observe a higher local velocity at the end of the translocation event, which can be attributed to the unfolding of the DNA blob outside the pore. Additionally, we observe a wide distribution in the intramolecular, intermolecular, and pore-to-pore translocation velocities. We utilize this information to estimate how well we can determine the spatial position from a measured nanopore signal.

Before we turn to the experimental findings, we briefly recapitulate the technique of solid-state nanopores and the DNA translocation process. Nanofabrication is used to create chips with 20 nm thick free-standing SiN membranes. A transmission electron microscope (TEM) is
subsequently used to focus an electron beam onto the membrane and create a nanopore of any
diameter desired. This membrane is then placed in a flowcell such that it separates two small
reservoirs containing 4M LiCl salt solution. Chlorinated silver electrodes are placed into each
chamber and an electric field is applied across the membrane. The ionic current passing through
the nanopore is recorded by means of a low-noise amplifier. Any biomolecule passing through
the pore causes a temporary drop in the measured ionic current for the duration of its
translocation through the pore constriction. DNA molecules, which are highly negatively
charged, are electrophoretically driven through the nanopore, as shown in Figure 1a. Long DNA
molecules typically translocate in a head-to-tail fashion\textsuperscript{20}, but can also exhibit varying degrees of
folding, when the molecule is captured along its length rather than at an end. While a piece of the
DNA traverses the nanopore, the rest of the DNA molecule remains as a large blob outside the
pore\textsuperscript{21}. A large variation is typically observed for the total translocation time, even among
molecules of the same type. These differences in the intermolecular velocity, as well as
intramolecular velocity fluctuations, have been suggested to be caused by variations in the drag
force due to the unraveling of this random blob outside the pore\textsuperscript{18}, which becomes increasingly
smaller as the translocation progresses. The translocation velocity of a molecule is dependent on
its particular initial conformation at the moment of capture, and this is believed to be the cause of
the wide distribution of intermolecular translocation times typically observed. Both simulations\textsuperscript{18}
and recapture experiments\textsuperscript{22} suggest that extended conformations lead to slower translocation
durations, presumably because the mean drag force is larger. Near the end of the translocation
process, the drag force due to the DNA blob before the pore decreases quickly, resulting in a
significant non-linear increase in the translocation velocity. This effect has been predicted by
simulations\textsuperscript{18} but never experimentally measured. A recent experimental study by Singer et al
examined the translocation time between two DNA-bound PNA probes as a function of distance, for relatively short (<3.4 kbp) DNA molecules translocating through a small (3.7 nm) pore in asymmetric salt conditions\textsuperscript{23}. They found a power law dependence of this time with distance between the probes, similar to what is observed for the total translocation time as a function of total DNA length. In contrast, we investigate long DNA molecules, with a specific focus on determining the velocity profile along the molecule and we attempt to shed some experimental light on the open questions that we outlined in this section.

We designed synthetic DNA constructs with markers at known positions in order to measure the local velocity over different segments along the molecule. DNA nanotechnology has grown significantly over the last few years, particularly due to the drop in DNA synthesis costs as well as the introduction of DNA origami\textsuperscript{24}, a simple technique to create nanoscale shapes out of DNA. In the DNA origami approach, a long single-stranded DNA scaffold is folded into any shape desired, through the addition of short oligonucleotides complementary to multiple sections of the scaffold. In this study, we used a similar strategy to create linear long DNA molecules with a protrusion at a defined position along the molecule (Fig. 1). Since the position of the protrusion is controlled by design and we can measure the time required to traverse the DNA between say the start of the molecule and the protrusion, we are able to determine the mean translocation velocity along different segments of the molecule. A 7560 base M13 ssDNA scaffold was hybridized with 42 bp oligos everywhere along its length except where the protrusion is attached (Fig. 1 and Supplementary Section 5). A protrusion is separately assembled and hybridized to the partially hybridized M13 backbone, as shown in Figure 1b. The 299 bp protrusion is created from the hybridization of six complementary oligos, ranging in length from 94 to 140 bases, as detailed in Supplementary Section 5. Three constructs were
created, i) a symmetric construct with the protrusion at the exact center of the linear molecule (Figure 2a), ii) an asymmetric construct with the protrusion positioned 1571 bp (20.8%) from the closest end (Figure 2b), iii) and a control construct with no protrusion. We chose to use a dsDNA protrusion since this will have a well-defined blockade level that can be easily distinguished from the current signals produced by knots\textsuperscript{25}, which are at least twice as high in amplitude (as measured from the single dsDNA blockade level). For the asymmetric construct we chose to attach the protrusion 1571 bp away from the end in order to be able to distinguish the protrusion from small folds which can occur at the start or the end of the translocation process. By measuring the time $\tau_p$ from the start of the translocation to the start of the protrusion, we are able to determine the mean velocity over the first part of the molecule. Similarly, we can determine the velocity for the segment in which the protrusion itself resides inside the pore, as well as the last segment of the molecule. Given the already large distribution in translocation times $\tau_{DNA}$ for DNA molecules of equal length, we use the normalized temporal position ($\tau_p/\tau_{DNA}$) for comparisons between different molecules and pores. The presence of nicks along the DNA strand can be expected to reduce the persistence length of these DNA constructs relative to dsDNA. In order to characterize this effect, we carried out gels and determined the persistence length using AFM measurements (Supplementary Section S1). We found that the persistence length of these constructs is only slightly reduced and that these DNA constructs still behave like worm-like chains as opposed to freely-jointed chains. These synthetic DNA constructs allowed us to probe the velocity of the translocation process in a way not previously possible.

Translocation of constructs containing a protrusion results in events with a characteristic current spike within the DNA blockade. As expected, when the constructs are translocated through a nanopore we observe that a significant fraction (Supplementary Section S2) of
unfolded events contain current spikes caused by the presence of the protrusions. Typical examples are shown in Figure 2, where we see (i) the DNA entering the pore, causing a current drop $I_1$, then (ii) during the translocation, we observe a sharp spike with an additional amplitude $I_1$ due to the protrusion, and (iii) eventually the current level returns to the open-pore level when the translocation is complete. These types of events are not observed in the control measurements on the construct without any protrusion. The fraction of events containing protrusions is set by a number of factors as discussed in Supplementary Section S2. The presence of the protrusion in these DNA constructs is clearly visible as a current spike present within a large fraction of the DNA events observed and allows us to determine the mean translocation velocity for any segment desired.

We first discuss the results based on the symmetric construct with the protrusion at the center of the DNA molecule. Figure 3 shows the distribution of the center protrusion as a function of the normalized temporal position ($\tau_p/\tau_{DNA}$). A Gaussian fit of this 296 event distribution has a mean of 0.528 and a standard deviation of 0.137. This distribution is extremely wide considering the well-defined position of the protrusion at the center of the molecule. If the molecules were travelling at a constant velocity we would expect a peak centered at 0.500. The slight shift observed in the mean of the distribution suggests that the molecule, on average, travelled slower in the first half relative to the second half. We also plot the difference between the average velocity over the last segment and the first segment in the DNA construct to examine the typical velocity fluctuations, as shown in Figure 5a. Positive differences occur if the molecule sped up in the latter half while negative differences occurred if the molecule slowed down. We can see that the Gaussian distribution (with several positive outliers cut off in the figure), has a mean value of 0.015 bp/µs and STD of 0.79. Nearly all values fall within the $\pm 2$ bp/µs range, with some outliers
as far out as +17 bp/µs. These velocity fluctuations are quite large in comparison with the mean velocity of 2.0 bp/µs for these molecules, calculated using the most probable translocation time. This very sizable variation indicates that the intramolecular speed distribution is very broad, similar to the inter-molecule variation.

Analysis of the translocations for the asymmetric DNA construct reveal that the translocation velocity of the molecule increases as it approaches the end. Figure 4a shows the normalized temporal positions observed for the position of the protrusion. The ‘protrusion first’ orientation (I) has a 114 event distribution with a Gaussian mean of 0.238 and a standard deviation of 0.103, while the ‘protrusion last’ orientation (II) has a more narrow distribution of 98 events with a mean at 0.830 and standard deviation of 0.054. If these constructs would translocate with a constant velocity, we would expect the distributions to be centered at 0.208 and 0.792 respectively, which are not within the standard error for the measured means. The small shift observed in the mean values again indicates that the molecules move slower at the start compared to the end. Figure 4b shows the average translocation velocity obtained for different segments based on both orientations. We observe that the last 21% of the molecule translocates 19% faster, on average, than the first 21%. The slightly higher mean velocity observed for the first 21% of the molecule (short red segment) compared to the first 79% of the molecule (long blue segment) may suggest that there is a slightly higher velocity right at the start of the translocation process, which was also observed in the simulations of Lu et al\textsuperscript{18} and could be due to the low drag at the very start of the translocation process. Supplementary Figure S7 zooms in on the data for each orientation and shows the average translocation velocity determined for the three different segments of the molecule for each orientation. These results further demonstrate that significant fluctuations in the translocation velocity are present and show that the DNA
speed increases during translocation with a strong increase in the local velocity at the end of the translocation process.

Next to the intra- and intermolecular fluctuations in the velocity, we also observe significant variations in the mean translocation velocity between different pores. Although the experiments all concern 10 nm pores that probe the same 7560 bp construct, we observe most-probable translocation times varying from 1.9 ms to 4.8 ms corresponding to velocities of 3.98 to 1.58 bp/µs respectively, as shown in Figure 5c. This significant variation can be attributed to DNA-pore interactions, which lead to a characteristic long tail in the translocation time distribution. We also note that pores fabricated in older membranes typically show slower translocation velocities. This effect could be due to more oxides present in the membrane material and requires further investigation. As we have shown, it is however still well possible to determine and directly compare the positions of local protrusions between different experiments by using the normalized temporal position from ensemble distributions. This approach is supported by our finding that the time required to reach the protrusion $\tau_p$ scales linearly with the total translocation time of the event $\tau_{DNA}$, as shown in Fig. S9.

How accurately can we determine the spatial position of a feature of a local structure (bound protein, side group, …) if we know its temporal profile? This question is central to many nanopore applications. The measurements carried out in our study allow us to estimate how accurately position can be determined from either a single measurement or an ensemble of measurements. Figure 6 shows the measured normalized temporal position as a function of the spatial position, from the data points of Fig. 3 and 4a (cf. Table 1). We use a cubic spline fit in order to interpolate values and the shaded area shows the range of the standard deviation observed. The solid blue line is the velocity profile that would apply in the case of a constant
velocity. We observe that the actual position, is always less than the position estimated assuming a constant velocity, with a statistically significant difference. This approach can be used to quantify the uncertainties in determining spatial position for various measurements, as discussed next.

Determining the spatial position from a single temporal measurement leads to very large uncertainties, which can be reduced significantly using ensemble measurements. We expect the spatial position corresponding to a single normalized temporal point to fall anywhere within the horizontal intercept of that normalized time point and the shaded red area. For example, if a normalized temporal position of 0.2 is measured, one would estimate that this is caused by a feature at a position along the molecule somewhere between 800 bp to 2200 bp. This result emphasizes the fact that single measurements produce very inaccurate results because the stochastic fluctuations are large. The solution to improving the accuracy thus is to carry out ensemble measurements and fit the resulting distribution of normalized time points, which results in much smaller uncertainties. Using the standard error of the mean for the Gaussian fits (Figure 3 and 4a), we estimate that the uncertainty in the spatial position (from an ensemble-measured normalized temporal position) to be about 200 bp over the first part of the molecule, which reduces to about 90 bp after the midpoint. The larger uncertainty in the first part of the molecule is due to the wider distribution observed in this region. If the velocity profile is not known or cannot be measured, one could assume a constant velocity (blue line Figure 6). What would be the error associated with estimating the spatial position from ensemble-measured temporal data, assuming a constant velocity? Since the value of the mean velocity is slightly higher than the actual local velocity for most of the molecule, the estimated spatial position determined assuming a constant velocity will be between 130 and 330 bp further along the
molecule than the true position, at all positions except at the start and the very end of the translocation process. The results demonstrate the importance of knowing the actual velocity profile and provide the first numbers for estimating the uncertainty in the spatial position of a feature along a molecule.

Although these estimates are specific for a 7560 bp long molecule in 4M LiCl, the results apply more generally. In the case of shorter molecules, the speed up at the end of the translocation will concern a larger fraction of the total translocation time, and accordingly there will be a larger difference between the mean velocity and the actual velocity over most of the molecule, and hence larger overestimates of the position if a constant velocity is assumed. In the case of longer molecules the effect will be opposite since the increase in velocity at the end will be an increasingly smaller fraction of the total translocation time and the mean velocity will converge towards the actual velocity over most of the molecule. The effect that the local velocity is smaller than the mean velocity may be even larger than our experiments reveal. In most nanopore measurements on dsDNA, the translocation time is much smaller than the polymer’s Zimm relaxation time. Due to the presence of nicks every 42 bp along our DNA-origami construct and the use of very high-salt solutions, the relaxation time of the DNA constructs in these experiments, however, is likely smaller than or similar to their translocation time, which may reduce the magnitude of some of the effects observed.

Conclusion

This study has introduced a novel method for probing the local velocities of DNA molecules translocating through solid state nanopores using synthetic DNA constructs. This was used to measure the mean velocity over several segments of a 7560 bp DNA molecule. Significant
fluctuations are observed in both the intramolecular and intermolecular translocation velocity, as well as between different nanopores of the same diameter. The size of the intramolecular velocity fluctuations is surprisingly large, and they are apparent even when averaged over length scales corresponding to half the total length of the molecule. We also systematically observe an increase in the velocity at the end of the translocation process, an effect attributed to the reduced drag force as the last of the DNA translocates through. We have used the measured velocity profile to estimate the error in determining the spatial position both if the velocity profile is known or if a constant velocity profile is assumed. The results demonstrate the utility of this approach for illuminating the biophysics of the translocation process.

Methods DNA construct self-assembly: The DNA construct was assembled using a circular strand of 7560 bases length derived from the genome of bacteriophage M13. Staple oligonucleotide strands were prepared by solid-phase chemical synthesis (Eurofins MWG, Ebersberg, Germany, HPSF grade). Production of the full DNA construct was accomplished by two separate reactions for assembling the long line backbone and the protrusion. The backbone was assembled in a reaction mixtures containing M13 phage DNA at a concentration of 50 nM, 42 base complementary DNA oligonucleotides at 200 nM each, and 5 mM TRIS, 1 mM EDTA, 20 mM MgCl₂ and 5 mM NaCl (pH 8). The reaction mixtures were subjected to a thermal annealing protocol using TETRAD (Biorad) thermal cycling devices. The mixtures were first incubated at 65°C for 15 min and then annealed from 60 to 40°C in steps of 1°C per hour. The protrusions were separately pre-annealed according to the same protocol above with 10mM MgCl₂. After separate assembly, the protrusion and backbone were incubated at RT for 12 hours in a ratio of 1.2:1. After the assembly all objects were purified using 100kDa Amicon filters to separate the construct from the excess staple strands. The filter purification was carried out 4
times with a buffer containing 5 mM TRIS, 1 mM EDTA, 5 mM MgCl\(_2\) and 5 mM NaCl (pH 8) centrifuging at 2000 rcf for 30 min. The constructs were then linearized with 10U HincII in a total reaction volume of 56.4 µL (for 3 hr at 37°C) and either used as is or purified with an phenol/chloroform extraction and ethanol precipitation.

**Methods pore measurements:** The synthetic DNA constructs were diluted into solutions of 4M LiCl TE pH8, which facilitates high-resolution nanopore measurements\(^{27}\). (Attempts to measure in 1M KCl were hampered by the limited resolution.) In this study, the DNA constructs were translocated through 10 nm nanopores at 100 mV applied voltage. Current traces were digitized at 500 kHz, low-pass Gaussian filtered at 40 kHz, and analyzed with the Transalyzer Matlab package\(^{28}\). We selected only non-folded DNA translocation events with a single spike of amplitude \(I_1\) for further analysis.
**Figure 1** – **a)** Schematic illustration of a synthetic DNA construct containing a protrusion translocating through a solid-state nanopore. **b)** Close up schematic of the protrusion showing it is assembled from multiple individual DNA oligomers.
Figure 2 – a) Example current trace of the symmetric DNA construct. b) Typical current traces produced by the asymmetric DNA construct. Two orientations are possible as depicted in the molecular configurations shown above each current trace.
**Figure 3** – Histogram of the normalized temporal position observed for the symmetric construct with a protrusion. The red line is the expected position of the protrusion assuming a constant translocation velocity. The solid black line is a Gaussian fit to the distribution. The distribution has a mean of 0.528 and a standard deviation of 0.137.
Figure 4  – **a)** Histogram of the normalized temporal position observed for the asymmetric construct. We took a cutoff point at 0.6 to separate the two orientations. The distribution of orientation I has a mean of 0.238 and STD of 0.103, while orientation II has a mean of 0.830 and STD of 0.054. Both peaks occur later than the position expected based on a constant velocity (0.208 and 0.792, respectively), which implies the first part of the molecule goes slower while the last part goes faster. **b)** The mean local translocation velocity over various segments of the asymmetric DNA construct. The horizontal line indicates the length of the segment that was used to elucidate its average speed while the vertical line indicates the standard error. The dashed green line is the mean translocation velocity of the dataset. Red points correspond to orientation I while blue points correspond to orientation II.
Figure 5 – Significant velocity fluctuations occur both within molecules (a), among molecules (b), and among different nanopores (c). a) Histogram of the difference in translocation velocity between the last half of the molecule and the first half. Positive values indicate the velocity was faster in the second half of the molecule. The distribution has a mean of 0.015 bp/µs and STD of 0.79. The solid green line shows the FWHM. b) Histogram of the differences observed in the
mean velocity for equal length molecules during the same experiment with the control construct containing no protrusion.  

c) Histograms showing the variation in the mean velocity between different 10 nm pores, *ceteris paribus*. We observe significant differences in the mean velocities for different pores.

**Figure 6** – The relationship between the normalized temporal position and the spatial position determined for our 7560 bp construct translocating through a 10 nm pore in 4M LiCl. Red points were experimentally measured, while the red line is a cubic spline interpolation, and the shaded area is the interpolated standard deviation. The blue line represents the case of constant velocity. The actual spatial position typically lags behind the position estimated assuming a constant velocity.
Table 1 – The value of the measured normalized temporal position, its standard deviation, and its standard error of the mean for three spatial positions on a 7560 bp DNA molecule.

<table>
<thead>
<tr>
<th>Spatial Position (bp)</th>
<th>Normalized Temporal Position</th>
<th>STD</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1571</td>
<td>0.238</td>
<td>0.137</td>
<td>0.01</td>
</tr>
<tr>
<td>3780</td>
<td>0.528</td>
<td>0.103</td>
<td>0.006</td>
</tr>
<tr>
<td>5989</td>
<td>0.830</td>
<td>0.054</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Supporting Information. Construct characterization, yield, velocity data, scaling of $\tau_p$ of with $\tau_{DNA}$, protrusion oligo sequences, M13 scaffold oligo sequences. This material is available free of charge via the Internet at http://pubs.acs.org.

ACKNOWLEDGMENT

The authors would like to acknowledge Meng-Yue Wu for TEM nanopore drilling, Jaco van der Torre for advice on purification protocols, and Katharina Häußermann for helpful discussions about the assembly strategy. This work was supported by the Netherlands Organisation for Scientific Research (NWO/OCW), as part of the Frontiers of Nanoscience program, as well as the European Research Council under research grant NanoforBio (no. 247072).
REFERENCES

