

# Nanoscale

Accepted Manuscript



This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

*Accepted Manuscripts* are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.

## ARTICLE

## Effect of confinement on DNA, solvent and counterion dynamics in a model biological nanopore

Cite this: DOI: 10.1039/x0xx00000x

Received 15<sup>th</sup> May, 2014,  
Accepted 00th January 2012

DOI: 10.1039/x0xx00000x

www.rsc.org/

Suren Markosyan<sup>1\*</sup>, Pablo M De Biase<sup>1\*</sup>, Luke Czapla<sup>1,2</sup>, Olga Samoylova<sup>1</sup>, Gurpreet Singh<sup>1</sup>, Javier Cuervo<sup>1§</sup>, D. Peter Tieleman<sup>1</sup> and Sergei Yu. Noskov<sup>1</sup>

The application of recent advances in nanopore technology to high-throughput DNA sequencing requires a more detailed understanding of solvent, ion and DNA interactions occurring within these pores. Here we present a combination of atomistic and coarse-grained modeling studies of the dynamics of short single-stranded DNA (ssDNA) homopolymers within the alpha-hemolysin pore, for the two single-stranded homopolymers poly(dA)<sub>40</sub> and poly(dC)<sub>40</sub>. Analysis of atomistic simulations along with the per-residue decomposition of protein-DNA interactions in these simulations gives new insight into the very complex issues that have yet to be fully addressed with detailed MD simulations. We discuss a modification of the solvent properties and ion distribution around DNA within nanopore confinement and put it into the general framework of counterion condensation theory. There is a reasonable agreement in computed properties from our all-atom simulations and the resulting predictions from analytical theories with experimental data, and our equilibrium results here support the conclusions from our previous non-equilibrium Brownian Dynamics studies with a recently developed BROMOC protocol that cations are the primary charge carriers through alpha-hemolysin nanopores under an applied voltage in the presence of ssDNA. Clustering analysis led to an identification of distinct conformational states of captured polymer and depth of the current blockade. Therefore, our data suggest that confined polymer may act as a flickering gate, thus contributing to excess noise phenomena. We also discuss the extent of water structuring due to nanopore confinement and the relationship between the conformational dynamics of a captured polymer and the distribution of blocked current.

## Introduction

Recent advances in nanobiophysics<sup>1-3</sup> and nanobiology<sup>4-7</sup> have led to a renewed interest in the development of fast and inexpensive methods for DNA sequencing and chemical analysis based on these technologies. A nanopore-based sequencing method based on alpha-hemolysin ( $\alpha$ HL) protein pores inserted into a lipid membrane<sup>8</sup> was first proposed more than a decade ago. In a typical experiment, individual single-stranded DNA (ssDNA) molecules are electrophoretically driven through a pore and the resulting nucleobase dependent ionic-current blockade is measured. Because of differences in the nucleotide interactions with the protein and permeating ions, the ionic current is anticipated to be different for every nucleobase in the pore, therefore offering a source of rapid and efficient DNA sequencing. The transmembrane region of a protein pore can be engineered for sensitivity to specific classes of analytes (DNA, sugars, divalent cations), enabling an inherent first level of discrimination. One of the key challenges in the effective design of nanopores capable of high levels of contrast between different nucleotides lies in lack of knowledge of the key determinants of

DNA-protein interaction in the nanopore confinement that are partly responsible for the distinct kinetics of DNA escape<sup>9</sup>. Site-specific mutations of protein pores,  $\alpha$ HL in particular, have been shown to greatly increase the interaction with specific chemical groups. Furthermore, it has been suggested that  $\alpha$ HL may possess several sensing zones that might preferentially bind different nucleotides<sup>4,5</sup>.

Arguably, molecular modeling techniques play an increasingly important role in evaluation of DNA and ion dynamics within the  $\alpha$ HL pore, which has been powerfully demonstrated by previous studies<sup>3</sup>. Many aspects of mechanical properties of DNA in confinement, its interaction with counterions, and ion-specific formation of structural elements in flexible polymers were also carefully studied using non-equilibrium MD simulations<sup>10, 11</sup>. Despite an increasing number of papers on the simulation of ssDNA in nanopores, our knowledge of the equilibrium free energy of DNA association with the pore and the per-residue contributions to DNA stabilization in the pore (as well as the effect of DNA on basic properties of the pores) remains incomplete. While a non-equilibrium approach based on SMD simulations<sup>12</sup> provided

interesting insight into electromechanics of DNA in the pore, simulations met increasing difficulties in obtaining fully converged PMFs that govern translocation of the large and flexible DNA polymer<sup>13</sup>. Thus mapping a particular “sweet spot” for DNA binding inside the pore is challenging.

Even in confinement within the nanopore, a charged polymer remains partially hydrated and maintains a counterion atmosphere, and thus assessment of DNA-protein interactions requires substantial? statistical sampling. Furthermore, evaluation of ion conductance from MD simulations often requires resorting to phenomenological models that use modifications of an average area profile to match blockade to conductance. An excellent review on the state of the problem has been published by A. Aksimentiev<sup>3</sup>. Ideally, one would like to use a simple and transparent computational scheme to calculate the per-residue contribution to the DNA-protein binding from atomistic simulations and then map key residues determining stabilization of different bases in the pore. The next step would be to connect observations to fundamental properties of the pore that modulate both ion and DNA interactions, such as a position-dependent effective dielectric constant. Dielectric constants can later be used for multiscale simulations that may enable accurate assessment of DNA and ion electrostatics in the pore. An additional dimension of complexity is the fact that solvent molecules confined within a nanopore may strongly affect the ionic current. The existence of structured (ordered) water wires was proposed to contribute to stabilizing cations in carbon nanotubes, and it is believed to play an important role in proton transport across nanopores<sup>14, 15</sup>. Furthermore, a hydrophobic environment of nanopore confinement may lead to a formation of ice-like solvation shell around transported cations<sup>16</sup>. Whether or not this effect is present in biological nanopores is currently unclear.

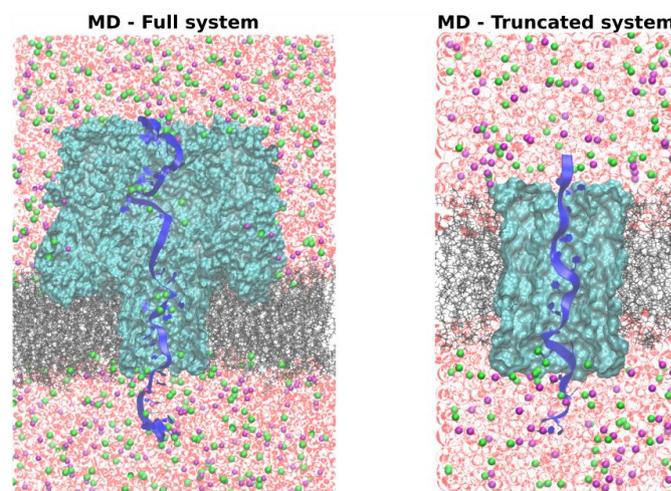
Here we present a detailed analysis of water, ion and DNA dynamics from over 0.5  $\mu\text{s}$  of all-atom MD simulations and 1000 independent 2 micro-seconds long BD simulations combined with clustering analysis and a per-residue decomposition of the interaction energy governing protein-DNA interactions. The key pore characteristics such as dielectric constant and position-dependent ion diffusion were used to study ssDNA blockade of ion currents for two model systems: ss-poly(dA)<sub>40</sub> and ss-poly(dC)<sub>40</sub> using a recently developed Grand Canonical Monte Carlo/Brownian Dynamics (GCMC/BD) protocol that explicitly accounts for ion-ion, ion-DNA and DNA-DNA interactions in the nanopore<sup>17, 18</sup>. The newly developed set of effective potentials describing ion-ion and ion-DNA interactions was used here to quantify the effect of ssDNA translocation on the predicted ion current through the pore from GCMC/BD simulations. It may be possible to improve these simulations through basic observations from more sophisticated treatments such as the all-atom explicit solvent MD simulations presented here.

## Methods

### Molecular Dynamics Simulations

First, we briefly summarize all the details about the simulation methodology used in the current study. We considered two systems

in our studies. In the first, explicitly solvated atomic-level systems were constructed and equilibrated using the program CHARMM with inclusion of an entire alpha-hemolysin toxin protein (with Protein Data Bank<sup>19</sup> accession code 7AHL<sup>20</sup>). The ss-poly(dA)<sub>40</sub> or ss-poly(dC)<sub>40</sub> polymers and ions were placed into a pre-equilibrated system containing a membrane patch and  $\alpha\text{HL}$  solvated by 1 M of KCl solution in an orthorhombic simulation box with dimensions: 130  $\text{\AA}$  x 130  $\text{\AA}$  x 180  $\text{\AA}$ . The system comprising of protein, ssDNA, explicit DPPC membrane, counterions and solvent is shown in Figure 1 and contains about 300,000 atoms for both studied systems<sup>21</sup>.



**Figure 1.** Schematic of simulation boxes:  $\alpha\text{HL}$  protein (cyan surf) is embedded into a DPPC lipid bilayer (black sticks) and solvated in TIP3P water (pink background). ss-poly(dA/dC)<sub>40</sub> or ss-poly(dA/dC)<sub>20</sub> (blue ribbon) is threaded through the  $\alpha\text{HL}$  water-filled pore.  $\text{K}^+$  and  $\text{Cl}^-$  (green and magenta spheres) are added to the simulation box to represent 1 M of KCl solution.

The  $z$ -axis is aligned with the long axis of the orthogonal system and runs perpendicular to the surface of the lipid bilayer. The two equilibrium MD simulations were performed on the full systems (15 ns each) and trajectories were sampled every 1 ns to obtain 15 starting conformations for simulations with applied voltage. All setups for MD simulations were done in keeping with the protocol that produces best correspondence to  $\alpha\text{HL}$  experiments as described in detail by Wells and Aksimentiev<sup>22</sup>. Each of the systems was equilibrated for 5 ns and then subjected to a production run of 20 ns with a biasing electrical field corresponding to a voltage drop of 600 mV, in order to generate the current distribution from independent runs.

The temperature was maintained using a Lowe-Anderson thermostat at  $T=315$  K, as implemented in NAMD 2.9 by Wells and Aksimentiev<sup>22</sup>. The van der Waals interactions were switched at 14–16  $\text{\AA}$  by a force-switching function<sup>23</sup>. Long-range electrostatic interactions were calculated using the PME method, with a grid spacing of 0.75  $\text{\AA}$  for fast Fourier transformation,  $\kappa = 0.34 \text{\AA}^{-1}$ , and a sixth-order B-spline interpolation<sup>24</sup>. We used flexible TIP3P water and ion parameters that provide the best agreement with

experimental data as tested by Wells and Aksimentiev<sup>22</sup> with a 1 fs time step. The extended system pressure algorithm was employed with constant pressure of 1 atm and constant cross-sectional area in the XY plane. A graphical representation of both systems used for the MD simulation studies is shown in Figure 1.

### Generation of initial ssDNA structures

It has been established that ion current blockade by ssDNA in nanopores displays considerable similarity to flickering in currents across ion channels<sup>7, 9, 25</sup>. One of the key features is the presence of excess noise phenomena, which is traditionally associated with rapid conformational transitions in the gate of a channel<sup>26, 27</sup>. Accordingly, one of the common explanations for the excess noise phenomena is that ssDNA conformational dynamics acts as a fast gate in ion transport across the channel. Single-stranded DNA is very flexible and exhibits characteristic conformational transitions in the micro- to milli-seconds range. Sampling its conformational dynamics in the nanopore and linking it to the computed currents is a significant challenge especially for the full system with over 300,000 atoms. To test possible effects of the initial structure on the obtained residual current we constructed a truncated version of the system which contains only the stem region and ss-poly(dX)<sub>x</sub> described in detail in our previous publication<sup>17</sup>. The size of the system allows us to use Replica-Exchange MD simulations to generate low-energy conformations of ss-DNA. We generate ion positions and DNA conformations using Metropolis Monte Carlo and an implicit solvent/membrane model with fixed protein coordinates, first using canonical DNA structure as an initial guess for the MC routine. One million DNA translational and dihedral rotational Monte Carlo moves were applied to pre-extended and minimized ssDNA with the CHARMM27 nucleic acid force field parameters in the presence of an implicit membrane and alpha-hemolysin using CHARMM27 protein parameters<sup>28</sup>, followed by random placement of ions and 100,000 ion translational Monte Carlo moves, all using a temperature  $T = 300$  K within the MC module in CHARMM. We then clustered structures from MC runs to obtain initial guesses for the Temperature Replica Exchange MD simulations (TREM) with the REPDSTR module of CHARMM. REPDSTR was used to run REMD. The exponential distribution of temperature points between 290 and 420 K<sup>29, 30</sup> allowed better sampling. The GBSW module of CHARMM was used to represent membrane and solvent<sup>28</sup>. We used 15 replicas with an average exchange probability of ~0.3. The structures generated by REMD were clustered using RMSD clustering with ssDNA backbone atom positions, and these structures were used to generate all-atom systems for MD simulations with an applied electrical field and truncated systems. Each of the truncated systems (10 runs per ssDNA) was re-equilibrated for 5 ns and then subjected to a production run of 45 ns with a biasing electrical field corresponding to a voltage drop of 600 mV

### Analysis of all-atom simulations

The following scheme was used to evaluate the per-residue decomposition of protein-DNA interaction enthalpy using the

protocol developed previously for studies of P53-DNA complexation<sup>31, 32</sup>:

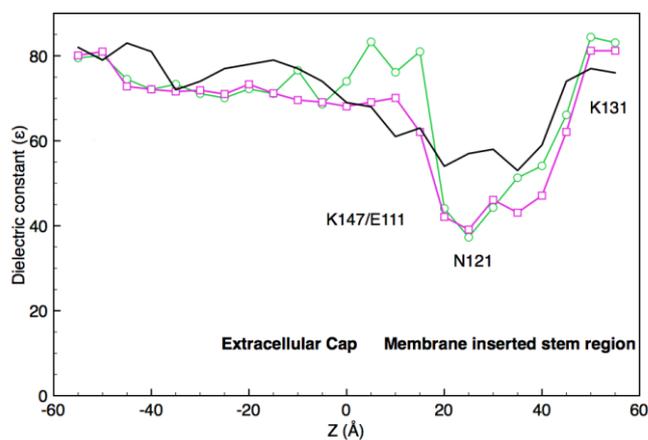
$$\Delta G_{Non-bonded} = \Delta G_{elect} + \Delta G_{vdw} \quad (1)$$

$\Delta G_{elect}$  was obtained by solving Poisson-Boltzmann (PB) equation. The protein, implicit membrane and DNA dielectric constants were set to  $\epsilon_1 = 12$ , similar to previous studies<sup>33</sup>, whereas the dielectric constant of solvent was set to  $\epsilon_2 = 80$  with an electrolyte concentration of 1 M.

The value  $\Delta G_{vdw}$  is the scaled vdW interaction from the CHARMM27 force field, as suggested by double mutant cycle studies of Eriksson and Roux<sup>33, 34</sup>. A full evaluation of the protein-DNA binding free energy would require computations of entropic terms, which is a formidable task due to the flexibility of ssDNA molecules, and therefore we limited our analysis to enthalpy computations. This MM/PB analysis is intended to illuminate differences in DNA-pore interactions (if any) between the two homopolymers. The position-dependent effective dielectric constants were evaluated using average fluctuations in the dipole moment of the volume slice (5 Å) using the following equation:

$$\epsilon = \epsilon_{\infty} + 4\pi / (3 \langle V \rangle k_b T) (\langle M^2 \rangle - \langle M \rangle^2) \quad (2)$$

where  $\langle V \rangle$  is the average volume occupied by solvent molecules in the slab, estimated by a grid-search algorithm and  $\langle M \rangle$  is the net dipole moment of the water molecules in the slab.



**Figure 2.** The one-dimensional profile of the dielectric constant in an open aHL pore (solid black), the aHL-poly(dA)40 pore (green) and the aHL-poly(dC)40 pore (magenta). The profile was computed from equilibrium all-atom MD simulations for the full system. The standard errors for position-dependent dielectric constants do not exceed 5 %. The uncertainties were estimated from block-averaging with 5 blocks.

The high-frequency correction ( $\epsilon_{\infty}$ ) has been set to a constant value of 1.4 to simplify analysis. The hydrogen bond lifetimes were computed from MD simulations with a truncated aHL system from

the 200 ns equilibration runs. The donor-acceptor definitions were taken from DeLoof *et al.*<sup>35</sup> and the resultant number of hydrogen bonds per water molecule were binned along the  $z$ -axis of the system every 1 Å. A similar strategy was used to estimate the orientation of the dipole moment for water molecules along  $z$ -axis<sup>36</sup>. The rate of ssDNA transport under applied electrical field was assessed simply from the time-displacement relation. To test whether or not ssDNA moves in a synchronous manner along its entire length or if each of the segments exhibit different dynamical regimes, we use a position-specific measurement similar to the position-dependent diffusion approach of Im and Roux<sup>37</sup>.

### Brownian Dynamics Simulations

All of the Brownian Dynamics simulations were performed using our BROMOC package<sup>13</sup>. This program was launched recently, and interested readers are advised to refer to the original paper for a description of the conceptual principles of the program. In all of the BROMOC simulations reported here, the biological nanopore (wild type  $\alpha$ HL) was treated as a rigid structure with a dielectric constant of 2 surrounded by a high dielectric solvent ( $\epsilon_w=80$ ) and embedded in an implicit membrane ( $\epsilon_m=2$ ) with a 38 Å thickness. The cylindrical region with  $\epsilon_c=80$  and radius of 14.0 Å was introduced to mimic a water-filled pore environment. The simulation box covered the entire length of the protein-DNA system and was surrounded by two symmetric buffer regions (of length 3.5 Å). The ssDNA molecule was represented as a coarse-grained homopolymer single-stranded chain (poly(dA)<sub>40</sub> or poly(dC)<sub>40</sub>). One thousand initial DNA structures were obtained from short BROMOC equilibrium runs. The 5' end was immobilized at the wide entrance of the pore and simulations were performed at a potential of +250 mV (negative in the wide pore entrance and positive in the narrow pore entrance). KCl salt concentration of 1 M and at a temperature of 315.38 K was used. A uniform diffusion coefficient of 0.001 Å<sup>2</sup>/ps was assigned to all nucleotides. One thousand separate simulations were run to obtain accurate statistics and relatively realistic ion distributions. The duration of each simulation was 2  $\mu$ s.

### Analysis of BROMOC simulations

The enhanced sampling and data availability for ion currents across the nanopore from BROMOC simulations allows a direct correlation of the ssDNA conformational dynamics with current levels in a blocked nanopore. We performed a cluster analysis using the clustering method of Daura *et al.*<sup>38</sup>. All BROMOC trajectories of ss-poly(dC)<sub>40</sub> were combined together and conformations were sampled every 5 ns. The choice of the poly(dC)<sub>40</sub> system was dictated by a) higher residual currents and b) better resolution of shallow and deep blockade levels. Next, Root Mean Square Deviation (RMSd) between every structure was computed and clusters were generated with a cut-off of 5 Å similar to techniques implemented in clustering studies of protein folding and conformational dynamics<sup>39, 40</sup>. The numbering of the nucleotides is from the cap to the stem. The nucleotides 1 to 28 spanning across the nanopore were considered for structural analysis. The alignment prior to RMSd computation was achieved by a discrete rotation along  $x$ - $y$  plane by an angle of

360°/7 due to the spatial equivalency given by the heptameric symmetry of the protein.

### Structural analysis of ssDNA

All-atom homopolymer structures for poly(dC) in the pore for the high and low conducting states observed during the coarse-grained (GCMC) Brownian dynamics simulations were obtained from RMSD fitting of atomistic structure from the base, sugar, and phosphate coordinates of simulated CG beads in two representative structures using CHARMM; an averaged strand structure from this simulation was also analyzed for the high conducting state. The conformational properties of ssDNA were analyzed and extracted in the software package 3DNA (X3DNA)<sup>41-43</sup>.

## Results and Discussions

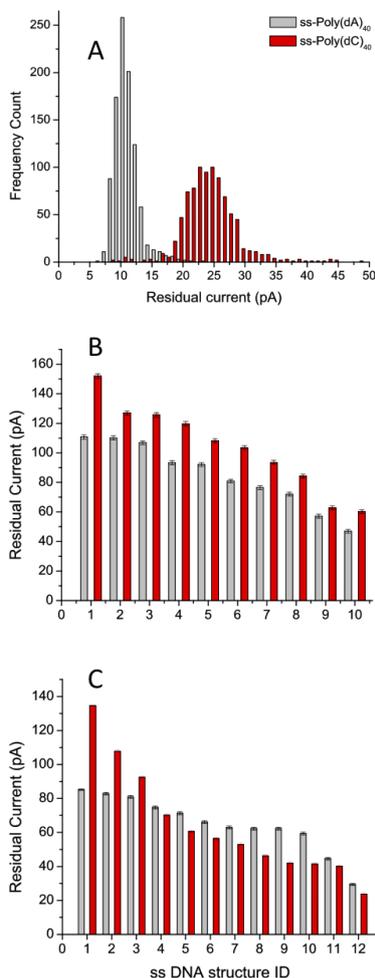
### Dielectric environment inside the pore

One of the most important characteristics of the nanopore required for evaluation of counterion dynamics is the effective dielectric constant with and without ssDNA blocking the pore. Accordingly, one-dimensional profiles of the dielectric constant in simulations of  $\alpha$ HL with single-stranded DNA and in simulations of an unblocked  $\alpha$ HL without DNA are shown in Figure 2. The block-averaging analysis of the trajectories suggests that standard errors in the position-dependent dielectric constant do not exceed 5%. An unblocked protein displays a mild drop in the average dielectric constant from 80 to 60-70 in the stem region, whereas the presence of a DNA molecule has a notable impact on the dielectric environment inside the pore by both excluded volume effects and the local electric field exerted on the solvent, as illustrated by an increase in the local dielectric constant around K147 ( $z=0$ ) and a well-pronounced drop around the middle of the bilayer (N121). The dielectric constant estimated from MD simulations drops from 80 to about 30-40.

### MD and BD Simulations: Distribution of Currents

One thousand separate Brownian Dynamics (BROMOC) simulations were carried out to obtain statistics on the distribution of residual currents. The Molecular Dynamics simulations are considerably more expensive, and therefore we ran only 15 full-system MD and 10 truncated-systems MD simulations with different starting structures at a higher voltage. The truncated system is known to produce reliable (relative) estimates of blocked currents<sup>22, 44</sup>. The summary of results from the three methods is shown in Figures 3A to 3C. An exhaustive sampling from BD simulations suggests that the blocked currents for both ss-poly(dA)<sub>40</sub> and ss-poly(dC)<sub>40</sub> display distributions with highly populated “mid” level residual current state and long tails corresponding to the “high” and “low” conductance levels of the nanopore. This is in good agreement with experimental measurements<sup>1, 5, 6, 9</sup> displaying both shallow and deep blockade of  $\alpha$ HL nanopore by ssDNA. This may also suggest that ssDNA may act as a “gate” in ion conductance across the nanopore leading to multiple conductance levels. The current distributions computed from BD simulations display a low conducting level of 10 pA and a high conducting level of 15-20 pA for ss-poly(dA)<sub>40</sub>. The

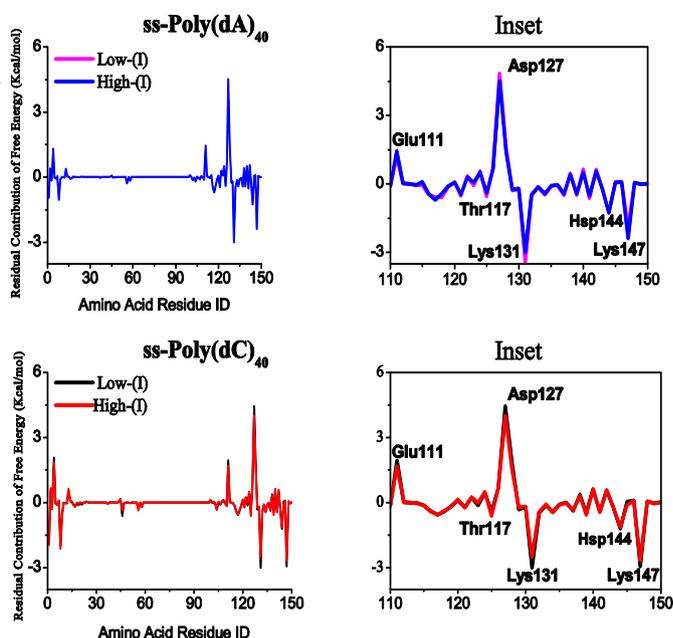
distribution of currents from MD simulations is considerably noisier due to shorter simulation times and large bias voltages, but is also in good agreement with previous experimental and theoretical studies<sup>3, 5, 6, 45</sup>. Importantly, even equilibrium MD simulations show that a flexible molecule such as ssDNA displays considerable conformational dynamics. The plateau in the computed relative RMSD (with respect to the starting conformation) is reached at around 9.5-12 Å for either of the polymers.



**Figure 3.** A) The frequency count distribution of residual ion currents from 1000 independent BD simulations ( $V = 250$  mV). B) and C) The residual currents from 12 independent full-MD (full protein) and 10 truncated-MD (truncated protein) simulations ( $V = 600$  mV) for poly(dA)<sub>x</sub> and poly(dC)<sub>x</sub>, respectively.

The  $\alpha$ HL protein, on the other hand, is very stable throughout all MD simulations with relative RMSD of only about 1.4-2.0 Å, indicating that thermal fluctuations, while important, may be coarse-grained by using an average protein structure. In spite of these apparent challenges, both methods were capable of reproducing a clear contrast between adenine and cytosine nucleotides blocking the

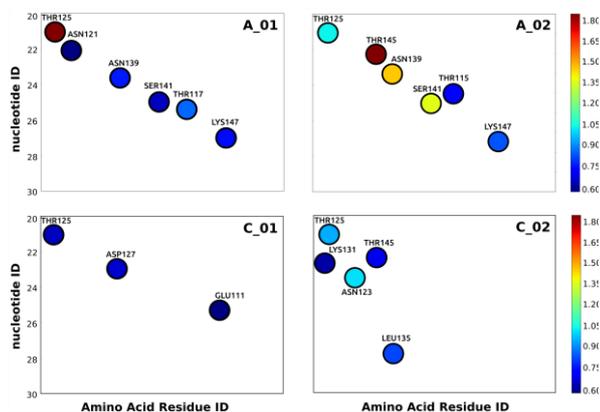
pore. As expected, the major charge carriers (>90 %) of ion current in the blocked state are cations.



**Figure 4.** The free energy decomposition per amino acid residues interacting with the DNA molecule in the channel. (Top Panel) Results for single-stranded poly(dA)<sub>40</sub>. (Bottom panel) Results for single-stranded poly(dC)<sub>40</sub>. The Inset plots on the right show the constriction zone between -20 and 20 Å. The legends (low and high) refer to 2 different levels of conductivity observed in simulations.

### MD Simulations and Protein-DNA Contact Surface

To better understand the origins of “high” and “low” current states for the blocked pore we analyzed corresponding trajectories and ensemble averages for all production runs. Long-lived protein-DNA contacts and energetics, combining both spatial analysis and MM/PB analysis, are shown in Figures 4 and 5. MM/PB decomposition based on the per-residue electrostatic potential average allows for identification of persistent protein-DNA contacts and strong thermodynamic factors for DNA binding, respectively. Data from Figures 4 and 5 suggest that even in the confinement zone, DNA is likely to be partially solvated and protein-DNA interactions remain shielded by electrolyte solution. Interestingly, poly(dA)<sub>40</sub> tends to interact more favorably with the primary constriction zone formed by K147, E111 and M113, whereas poly(dC)<sub>40</sub> displays a more favorable interaction with the secondary confinement marked by N121, N123, D127 and L135.



**Figure 5.** Key amino acid residues in protein-DNA interactions Map of close contacts observed between protein residues of  $\alpha$ HL and the nucleotides of ssDNA (top panel). This map is zoomed to show average close contact occupancies in and around the constriction (the narrowest region of the pore), since this is the region where all the long-living contacts were observed. Two different systems: polyA and polyC were subject to detailed mapping in the figure (Right panel). A\_01 (polyA) and C\_01 (polyC) have low, A\_02 (polyA) and C\_02 (polyC) have high conductance levels. The bottom panel shows the location of the residues in different sensing zones.

The data collected in Figs. 4 and 5 indicate that the ability for hydrogen bonding is an essential factor for preferential interactions between cytosine and two asparagine residues, N121 and N139. These interactions are supplemented by stabilizing contacts with the polar side chain of T145. While average per-residue contributions are relatively small, with absolute contributions of between 1 to 5 kcal/mol to the base stabilization, they might be very important to collectively contribute to ssDNA stabilization in the pore. While the

most confined region of the alpha-hemolysin pore (near K147 residues at the junction of the protein cap and stem) has strong ionic interactions with the phosphate backbone charge of DNA, it does not display a large degree of discrimination between adenine and cytosine nucleotides. The second zone, coined as a secondary constriction (centered at N121-N123-N139), shows some of the strongest combined interactions located near the *trans* side of the pore within the stem. The average protein-DNA interaction energies display considerable fluctuations between different runs, especially in the region of the secondary constriction. Protein-DNA contact analysis indicates that the DNA strand in the pore displays significant flexibility, allowing for up to 3-4 different nucleotides to interact with the same protein residue (the  $\alpha$ HL protein is a symmetric heptamer) in both sensing regions within the beta barrel. Experimentally, this region has been recently identified as a putative recognition site for nucleotide sequence, consistent with the differences we observe in this network of protein-DNA interactions for the two different homopolymers<sup>5, 6</sup>. Significant interactions near the *cis* entrance of the pore observed in the simulations may be useful for engineering optimal alpha-hemolysin mutants for DNA sequencing applications. Although the DNA bases closest to the confinement region of the pore appear to have the greatest effect on current blockade, this site may also significantly influence translocation dynamics and create significant biological noise in ion current readouts.

#### The relationship between current distributions and protein-DNA contact surface

Analysis of conductivity of alpha-hemolysin as a function of the single-stranded DNA conformation inside the nanopore reveals some interesting features. Earlier in this paper, we showed that the distribution of residual currents in  $\alpha$ HL demonstrates different conduction levels both in MD and BD simulations. The conducting levels may be directly related to the formation of distinct ssDNA conformational states. Figure 5 displays averages for long-lived contacts observed between the nucleotides of the ssDNA and amino acid residues within the  $\alpha$ HL nanopore, based on hydrogen bonding and salt-bridging criteria described in the Methods section. Interestingly, different conducting levels display notable differences in the lifetime and distribution of protein-DNA contacts. For example, in a “high” conductance simulation interactions between ssDNA and some asparagine residues (N139 and N121) were reduced or absent. The contact analysis done for the high conducting level (Figure 4A) observed in MD (as well as BD) simulations shows minima in the number and stability of protein-DNA contacts, leading to a more elongated form of ssDNA allowing for across-the-wall transport of ions. This in turn may facilitate permeation and promote the high flux of cations through the  $\alpha$ HL nanopore. Hydrogen bonding to pockets of threonines in positions 115; 117; 125 and 145 also plays a specific role in ion transport through the  $\alpha$ HL nanopore. T145 appears to be particularly important for the formation of deep-blockade states in the studied systems. The low-conductance state displays virtually no interactions between ssDNA

and T145 as compared to the other state (Fig. 5). These residues ( $\alpha$ HL is a heptamer) are located in the sensing zone of  $\alpha$ HL

### The relationship between ssDNA conformational dynamics and current blockade levels

To analyse further the relationship between particular conformations of ss-DNA and depth of the current blockade we focus first on two extremes from all-atom MD simulations, the deep- and shallow-blocking currents. Two characteristic features of the captured polymer were chosen for analysis – end-to-end distance and radius of gyration. The time-series of these two properties are shown in the Supplementary Materials. There is a modest correlation between the radius of gyration and ion current allowed by the pore. As expected, the more “stretched” states of ssDNA allow higher conductance. The major limitation for this analysis is the relatively poor sampling of ssDNA dynamics in nano-seconds long simulations. To provide additional insights we performed clustering analysis of BROMOC trajectories to identify particular structural features of ss-DNA responsible for deep- and shallow-blockade of the pore. The coarse-grained nature of BROMOC allows us to include over 1.5 milliseconds of dynamics and to correlate this with the observed residual currents. Figure 6 shows the composition of cluster populations of the trajectories associated with low (8.3-14.4 pA, 12 trajectories representing ~ 24 micro-seconds of dynamics), average (20.5-26.6 pA, 500 trajectories) and high (38.8-44.9 pA, 5 trajectories) current. The structural composition for the high current runs is clearly similar to the average one. However, the cluster composition associated with low current has a very particular pattern. Cluster index reflects the occupancy of the cluster from highest (0) to the lowest (2123) in an entire set of trajectories (1.5 milli-seconds) used for the analysis. To test the current dispersion dependence on the structure of each cluster we ran current simulations for 200 structures in cluster 0, 1, 15, 16, 17, 23 and 32. The currents obtained were within ~3 pA of the average values, indicating the structural criteria for the clustering produce clusters with structures that yield the same current.

To test the stability of the structures and its relation to current reproducibility we selected two initial structures associated with the lowest (8.3 pA) and highest (44.9 pA) current for all of the trajectories. Each DNA structure with ions was re-run 500 times with different random number generator seeds. The average current obtained for the low-current structure was  $10.5 \pm 1.9$  (11+/-2) pA and for high-current structure  $27.8 \pm 2.4$  pA (28+/-2). This indicates that the low-current structure is a stable and long-lived one but the structures associated with the shallow-block are not kinetically stable and rapidly morph back to the most probable conductance states.

Next we performed structural analysis of these clusters. Figures 6 and S1 shows the structure of the center of the cluster for low- and high-current states, respectively. We observed only minor differences in structures associated with average or high current levels. However, low-current structures (deep block state) consistently display a coiled structure in the alpha-hemolysin cap that reduces the pore's capacity for ion transport.

A series of structural descriptors is collected in the Supplementary information. In a representative conformation for the low conducting state, ssDNA is compacted within the wide cap region of the pore and a single 360-degree loop is observed in the simulated coarse-grained chain in this region, while in the high conducting state the chain is more extended through the pore with a displacement of ~4.5 Å per nucleotide along the pore axis. The constructed atomistic model of poly(dC) in the low residual current state displays significant distortions in the backbone structure near the 11<sup>th</sup> and 16<sup>th</sup> nucleotide from the 1<sup>st</sup> nucleotide entering into the cap at the 5'-end. The observed coiling allows for a significantly more nucleotides to be present within the cap region. The key characteristics for low- and high-conducting conformation of ssDNA are collected in the supplementary information (high.dat and low.dat). Briefly, cytosine bases in this looped region within the  $\alpha$ HL cap are inclined upward towards the stem of the pore, while adjacent bases are nearly perpendicular to the axis of the pore or slightly bent downward towards the entrance of the cap. The deoxyribose sugars of the polynucleotide backbone adopt C2'-endo and a few C3'-endo/exo rotational states at the majority of nucleotides in both high and low conducting state ssDNA chain conformations. However, for the low conducting state conformation, at the 11<sup>th</sup> nucleotide near the start of the loop there is a ~70 degree rotation of the backbone  $\delta$  torsion (C5'-C4'-C3'-O3') relative to adjacent nucleotides of the chain and the deoxyribose sugar is observed in a lone C4'-exo conformation; there is also a complementary ~70 degree rotation of the backbone  $\epsilon$  (C4'-C3'-O3'-P) torsion with the deoxyribose sugar in the lone C1'-exo conformation for the 16<sup>th</sup> nucleotide near the end of the loop. The looped nucleotides in between have more gradual bending and all adopt the C2'-endo deoxyribose sugar conformation. It is important to emphasize that similar looping patterns were observed outside of the pore in modeled chains in both high- and low- conducting states (Figure S1), consistent with the low persistence length of ssDNA in corresponding solution conditions. Therefore, we conclude that formation of the coiled structure in the cap is a structural feature that is associated with observed deeply blocking state of the ssDNA.

### Counterion screening in the pore

Next, we evaluated counterion densities inside the pore. As expected, Cl- density is negligibly small both for BD and MD runs in keeping with prior studies<sup>17, 22, 45</sup>. The ion densities within the pore are in apparent agreement with a very general theory (Counterion Condensation of Onsager-Manning-Oosawa) and its application to DNA in the pore<sup>46, 47</sup> and to observations from other experimental work<sup>48 49 50</sup>, when the appropriate dielectric properties are accounted for in the associated analytical expression. The effective strengthening of electrostatic interactions by the decrease in solvent shielding creates a dense counterion atmosphere and nearly complete protein screening from DNA anionic charge within some regions of the pore. The predictions of counterion condensation theory are in keeping with the observed DNA-counterion association in the simulations, when considering the observed reduced dielectric screening of charge-charge interactions.

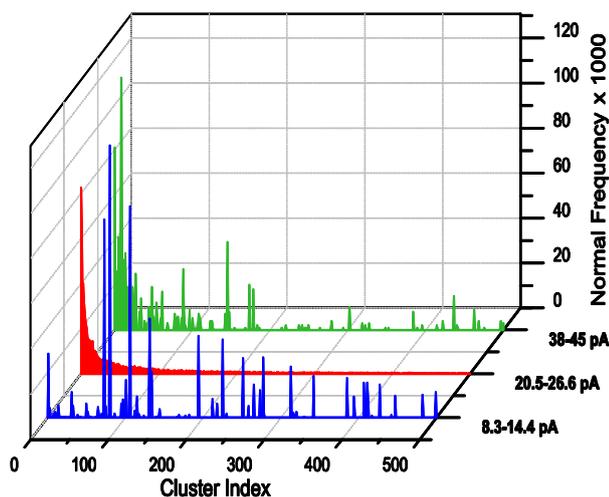
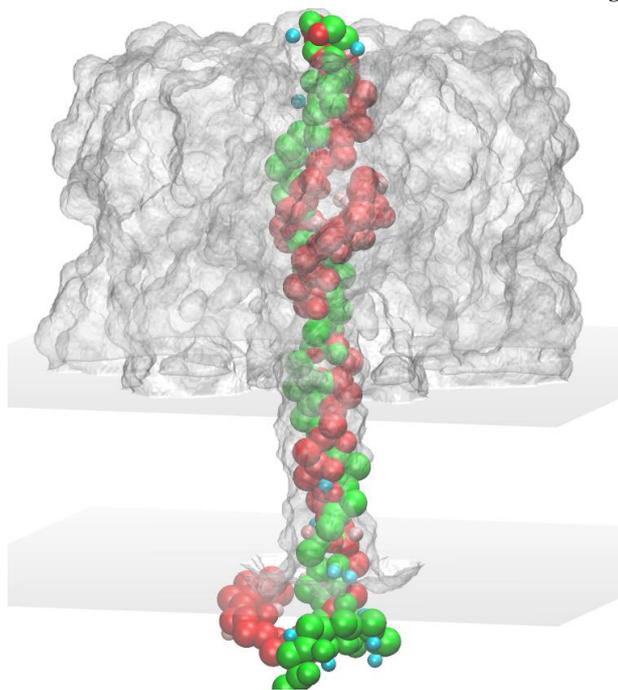


Figure 6.



Top panel: The RMSD-cluster composition of trajectories associated with high-, mid- and low-current levels. Bottom panel: Graphical illustration of ssDNA structure blocking the pore in high- and low-conductance states shown in green and red, respectively. The K<sup>+</sup> ions within 5 Å are shown in blue

An estimated fractional neutralization of ~80% of the DNA charge is observed in our simulations for DNA of the narrowest stem region, compared to the overall neutralization of ~45% of the charge in the wide cap of the pore, and a neutralization of ~30% of the charge outside of the pore.

This is consistent with the predictions from counterion condensation theory from the work of Manning<sup>51</sup>, using an expression for the fractional DNA charge neutralization and an effective dielectric constant of  $\epsilon=40$ :

$$f = 1 - \frac{b}{b_B} = 1 - \frac{N \epsilon \epsilon_0 k T}{L e^2} \quad (3)$$

where  $N$  is the magnitude of the charge per monomer (equal to 1 for ssDNA) and  $L$  is the distance between monomer units of approximately 5 Å observed in the simulation. The ratio of these values multiplied by the inverse Bjerrum length  $b_B$ , a value modified by the dielectric constant value in each individual region of the system, predicts a very similar pattern of neutralization with dielectric constant. On the other hand, the confinement effect also leads to a decrease in the ion desolvation penalty and thus tighter non-electrostatic interactions for the non-polar moieties of DNA around the middle of the pore.

### Single-stranded DNA translocation and its position dependence in the sensing areas

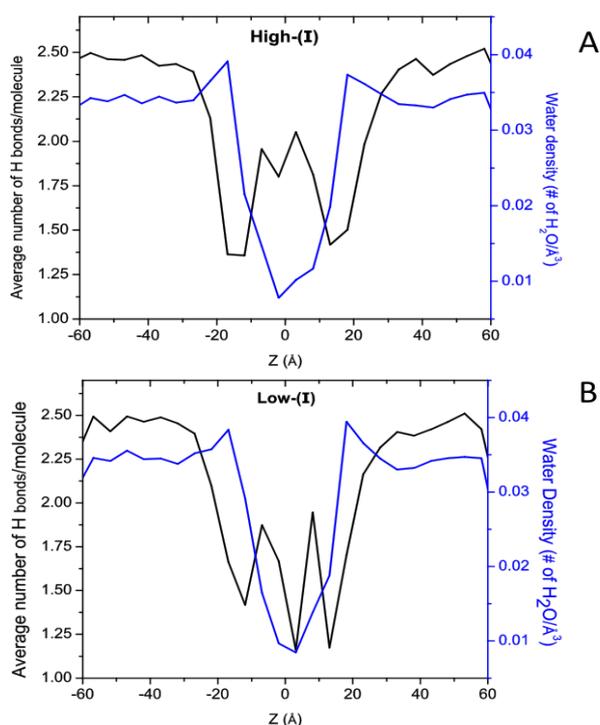
To illustrate one of the possible reasons for such a broad distribution of currents and to test the time-displacement relationship in the studied nanopores, we computed the position-dependent mean square displacement (MSD) for five all-atom MD trajectories representing different conducting states of the system (Figure S2). We used the same method as described by Mathe *et al.*<sup>45</sup>. The system was separated into three regions (the  $\alpha$ HL cap and two sensing zones in the stem) and the MSD for a subset of nucleotides was measured in the presence of a biasing potential of  $V = 600$  mV. The nucleotide displacements for the part confined in the cap region display Rouse-Zimm behavior, with two well-defined time dependencies as described previously for single-stranded DNA dynamics by Shusterman *et al.*<sup>52, 53</sup> with a relatively small distribution of MSDs between different simulations. Although the calculations are noisy in MD simulations, it is clear that the single nucleotide dynamics in the first and second constriction regions are distinctly different. The captured polymers exhibit a broad distribution of dynamical regimes from almost immobile to strongly extending form of captured ssDNA. This is expected to affect ion currents and result in a broad distribution of computed currents, as shown in Figure 3. Interestingly, several groups reported long-lived captured DNA states for alpha-hemolysin nanopores with non-exponential escape kinetics<sup>9</sup>. The lack of mobility in the captured DNA correlates well with a maximum in protein-DNA contacts and maximization of protein-DNA interaction energies. Protein residues compete with counterions for interaction with polar groups in DNA, but it is hard to assess survival probabilities of “immobile” states due to limitations in sampling.

### Water Ordering in the Model Nanopore

A number of papers suggest that nanopore confinement may lead to a significant stabilization of hydrogen-bonded water wires over

macroscopical lengths<sup>54</sup>. The lifetime of stable wires can extend to microseconds and even minutes, and insertion of additional charge breaking stable wires or emptying/filling transitions will result in significant fluctuations in ion transport rates. To test hydrogen-bonding patterns of water in DNA-blocked nanopores, we performed a position dependent connectivity analysis. The resulting hydrogen bonding per molecule profile along with particle density profiles normalized by the solvent accessible volume (excluding volume taken up by the ssDNA molecule) is shown in Figure 7.

The density profile displays a sharp drop in particle density around the first constriction zone ( $0 \text{ \AA} < z < 5 \text{ \AA}$ ). The average number of hydrogen bonds per water molecule also considerably decreases in the pore from 2.5 to 1.5, which does not support the idea of having structured water clusters both locally and across the entire length of the beta barrel.



**Figure 7.** The hydrogen bond per water molecule profile along the pore axis (solid black line) and the numerical water density profile (solid blue line) for a truncated version of  $\alpha$ HL pore from MD simulations with applied voltage for high- (A) and low-conductance states (B) of  $\alpha$ HL-polyd(C)<sub>20</sub>.

It seems that water ordering in the nanopore displays a very modest dependence on the blocking state of ssDNA. To investigate the ordering of water molecules in the confinement of the nanopore we computed the average dipole orientation from 200 ns of equilibrium (no voltage) MD simulation of a truncated pore. Figure S3 shows the water dipole orientation along the  $z$ -axis. The computed water profile is reminiscent of that published by Klauda *et al.* for pure lipid bilayers in the absence of ion channels<sup>36</sup>. The maximum of the water dipole moment is located in the choline region, while water

molecules in the bulk or inside the pore display two regions where water is apparently oriented, e.g., between  $0 \text{ \AA} < z < 5 \text{ \AA}$  and to a lesser degree between  $18 \text{ \AA} < z < 22 \text{ \AA}$ , overlapping with the proposed location of the first and second sensing zones. Although water molecules inside the wide nanopore, even partially blocked by ssDNA, appear to form short-lived pockets with oriented water molecules similar perhaps to those found in a narrow hydrophobic nanopores with a notable capacity for water transport<sup>55</sup>, it appears to lack the hydrogen bond connectivity associated with long-lived “ice-like” structures found in carbon nanotubes<sup>16</sup>, and is therefore unlikely to contribute significantly to observed excess noise in nanopore recordings.

### Comparison between experimental and computational results

Direct comparison of the simulation results (MD and BD simulations) to electrophysiological recordings is challenging because of several approximations<sup>3, 44</sup>. All-atom MD simulations will arguably provide the most detailed description of ion, DNA and solvent dynamics in the confinement of a nanopore. However, the significant computational costs involved in all-atom simulations, the resulting limited sampling time and therefore the need to use artificially high applied voltages (up to 2V), and the difficulties in describing slow conformational dynamics of captured polymers and its relation to fast ion transport events, make application of atomistic simulations to studies of DNA transport a true challenge<sup>22</sup>. The use of effective potentials derived from all-atom MD simulations in applications to DNA transport offers a remedy to some of these challenges, although this procedure itself introduces several approximations<sup>17, 18</sup>. The assignment of uniform transport coefficients to all nucleotides in the strand, the rigid structure of the nanopore and the resulting apparent over-estimation of the electrostatic barriers associated with ion entrances to the nanopore are probably the most important limitations. The results of the BD simulations reported here display residual currents that are somewhat lower than those reported from patch-clamp measurements of the  $\alpha$ HL pore in lipid bilayers<sup>1, 2</sup>. However, they are providing correct estimates for relative current blockade between different strands<sup>18</sup>. Interestingly, BD data is in surprisingly good agreement with the reported residual currents for the  $\alpha$ HL pore captured by a solid-state nanopore<sup>56</sup>. The solid-state environment should result in decreased conformational dynamics of the biological pore (hemolysin) itself and would lead to more significant DNA blockade compared to previous studies performed in lipid bilayers.

### Conclusions

A combination of all-atom Molecular Dynamics and coarse-grained Brownian Dynamics simulations was used to study ion, DNA and water dynamics in the confinement of a model nanopore,  $\alpha$ -hemolysin. It was found that both methods are capable of accurate separation between blocked currents produced by different single-stranded polynucleotides, but the complexity of single-stranded DNA dynamics makes it difficult to obtain accurate estimates of ion current from just a single MD simulation, and thus the approach of multiple MD simulations was used. The captured single-stranded DNA displays several conducting levels and its conformational dynamics may act as a fast gate, consistent with reported excess

noise in electrophysiological recordings. Based on analysis of water connectivity and hydrogen bonding we conclude that water clustering observed in other systems is unlikely to be a critical factor in the observed excess noise in aHL. The captured nucleotides exhibit several dynamical regimes depending on position in the confinement region. In the wide mushroom-like cap region of the protein, nucleotide dynamics are similar to those expected for a free polymer in solution. However, nucleotides in the stem region display different regimes ranging from an immobilized strand to a voltage-driven extended state. Based on cluster analysis of the trajectories from BROMOC simulations we conclude that ssDNA conformational dynamics in the stem and cap regions of aHL contributes to the deep and shallow blocks observed experimentally. A long-lived state of coiled structures formed in the cap of aHL contributes significantly to the blockade of ion transport. The conformational dynamics of DNA clearly modulates ion dynamics. Therefore complete sampling of DNA conformations from all-atom MD simulations is likely required for accurate results but this presents a significant challenge due to slow DNA dynamics on the MD time scale. Sampling could be improved in future studies, but it is not possible to ascertain conclusively whether structures predicted in any of the DNA-nanopore simulation literature or in this work are accurate, absent of crystallographic data or other direct experimental evidence. It is likely that the timescale for ssDNA relaxation even from carefully prepared starting structures is far too long for the structure to fully evolve over the course of a 0.2  $\mu$ s or slightly longer MD simulation trajectory. One viable strategy would build on the combination of enhanced sampling techniques and MD simulations for future studies into this very challenging problem. We hope that future studies based on the approaches presented here will add additional insight into the role of DNA translocation dynamics on ion conductance and other important properties of these biological nanopores for DNA sequencing applications.

### Acknowledgements

Useful discussions with G. Barall, A. Aksimentiev, B. Roux and W. Im are acknowledged. This work was supported by an NHGRI Grant R01 HG005095 and the National Sciences and Engineering Research Council (NSERC) (Discovery Grant RGPIN-315019 to S.Y.N.). S.Y.N. is an Alberta Innovates Technology Futures New Faculty, Canadian Institute for Health Research New Investigator, and an Alberta Innovates Health Solutions (AIHS) Scholar. Work in DPTs group is supported by NSERC (RGPIN/238357). DPT is an AIHS Scientist and Alberta Innovates Technology Futures Strategic Chair in (Bio)Molecular Simulation. Computations were performed on the West-Grid/Compute Canada facilities, and the University of Calgary TNK cluster acquired with direct support by the Canada Foundation for Innovation.

### Notes and references

<sup>1</sup> Centre for Molecular Simulation, Department of Biological Sciences, 2500 University Drive, Calgary, AB, CANADA, T2N 2N4  
<sup>2</sup> Present Address: Department of Chemistry and Engineering Science and Physics, College of Staten Island, The City University of New York (CUNY), 2800 Victory Blvd, Staten Island, NY, USA,

10314. <sup>§</sup> Present address: Enbridge (Polymath), Edmonton, AB, Canada

\* The first two authors contributed equally to this work

Author for correspondence: S.Y. Noskov,

Phone: 403-210-7971, e-mail: [snoskov@ucalgary.ca](mailto:snoskov@ucalgary.ca)

1. D. Branton, 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, *Nat Biotechnol*, 2008, 26, 1146-1153.
2. D. W. Deamer and M. Akeson, *Trends Biotechnol*, 2000, 18, 147-151.
3. A. Aksimentiev, *Nanoscale*, 2010, 2, 468-483.
4. D. Stoddart, A. J. Heron, J. Klingelhofer, E. Mikhailova, G. Maglia and H. Bayley, *Nano Letters*, 2010, 10, 3633-3637.
5. D. Stoddart, A. J. Heron, E. Mikhailova, G. Maglia and H. Bayley, *P Natl Acad Sci USA*, 2009, 106, 7702-7707.
6. D. Stoddart, G. Maglia, E. Mikhailova, A. J. Heron and H. Bayley, *Angew Chem Intl Ed*, 2010, 49, 556-559.
7. C. Tropini and A. Marziali, *Biophys J*, 2007, 92, 1632-1637.
8. J. J. Kasianowicz, E. Brandin, D. Branton and D. W. Deamer, *P Natl Acad Sci USA*, 1996, 93, 13770-13773.
9. M. Wiggin, C. Tropini, V. Tabard-Cossa, N. N. Jetha and A. Marziali, *Biophys J*, 2008, 95, 5317-5323.
10. J. B. Heng, A. Aksimentiev, C. Ho, P. Marks, Y. V. Grinkova, S. Sligar, K. Schulten and G. Timp, *Nano Letters*, 2005, 5, 1883-1888.
11. J. B. Heng, A. Aksimentiev, C. Ho, P. Marks, Y. V. Grinkova, S. Sligar, K. Schulten and G. Timp, *Biophys J*, 2006, 90, 1098-1106.
12. H. S. C. Martin, S. Jha, S. Howorka and P. V. Coveney, *J Chem Theory Comput*, 2009, 5, 2135-2148.
13. X. Daura, R. Affentranger and A. E. Mark, *Chemphyschem*, 2010, 11, 3734-3743.
14. J. Kofinger, G. Hummer and C. Dellago, *P Natl Acad Sci USA*, 2008, 105, 13218-13222.
15. B. Corry, *J Phys Chem B*, 2008, 112, 1427-1434.
16. Z. J. He, J. Zhou, X. H. Lu and B. Corry, *J Phys Chem C*, 2013, 117, 11412-11420.
17. P. M. De Biase, C. J. F. Solano, S. Markosyan, L. Czapla and S. Y. Noskov, *J Chem Theor Comp*, 2012, 8, 2540-2551.
18. P. M. De Biase, S. Markosyan and S. Y. Noskov, *J Comp Chem*, 2014, 9, 711-721.
19. H. M. Berman, T. N. Bhat, P. E. Bourne, Z. Feng, G. Gilliland, H. Weissig and J. Westbrook, *Nature structural biology*, 2000, 7, 957-959.
20. L. Z. Song, M. R. Hobaugh, C. Shustak, S. Cheley, H. Bayley and J. E. Gouaux, *Science*, 1996, 274, 1859-1866.
21. J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale and K. Schulten, *J Comp Chem*, 2005, 26, 1781-1802.
22. D. B. Wells, Bhattacharya, S, Carr, R, Maffeo, C, Ho, A., Comer, J and Aksimentiev, A., in *Nanopore-based technology: single molecular characterization and DNA sequencing*, Humana Press, 2011.
23. P. J. Steinbach and B. R. Brooks, *J Comput Chem*, 1994, 15, 667-683.
24. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J Chem Phys*, 1995, 103, 8577-8593.
25. V. Tabard-Cossa, D. Trivedi, M. Wiggin, N. N. Jetha and A. Marziali, *Nanotechnology*, 2007, 18.
26. S. M. Bezrukov, L. K. J. Vandamme and L. B. Kish, *Fluct Noise Lett*, 2011, 10, 417-418.
27. S. M. Bezrukov and M. Winterhalter, *Phys Rev Lett*, 2000, 85, 202-205.
28. A. D. Mackerell, Jr., M. Feig and C. L. Brooks, 3rd, *J Comput Chem*, 2004, 25, 1400-1415.

29. R. B. Best and G. Hummer, *P Natl Acad Sci USA*, 2005, 102, 6732-6737.
30. R. B. Best and G. Hummer, *J Am Chem Soc*, 2008, 130, 3706-3712.
31. S. Y. Noskov, J. D. Wright and C. Lim, *J Phys Chem B*, 2002, 106, 13047-13057.
32. J. D. Wright, S. Y. Noskov and C. Lim, *Nucleic Acids Res*, 2002, 30, 1563-1574.
33. M. A. L. Eriksson and B. Roux, *Biophys J*, 2002, 83, 2595-2609.
34. S. Jo, M. Vargyas, J. Vasko-Szedlar, B. Roux and W. Im, *Nucleic Acids Res*, 2008, 36, 270-275.
35. H. DeLoof, L. Nilsson and R. Rigler, *J Am Chem Soc*, 1992, 114, 4028-4035.
36. J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell and R. W. Pastor, *J Phys Chem B*, 2010, 114, 7830-7843.
37. W. Im and B. Roux, *J Mol Biol*, 2002, 322, 851-869.
38. X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, *Angew Chem Intl Ed*, 1999, 38, 236-240.
39. G. Singh and D. P. Tieleman, *J Chem Theory Comput*, 2013, 9, 1657-1666.
40. Y. Yao, J. Sun, X. H. Huang, G. R. Bowman, G. Singh, M. Lesnick, L. J. Guibas, V. S. Pande and G. Carlsson, *J Chem Phys*, 2009, 130.
41. X. J. Lu and W. K. Olson, *Nucleic Acids Res*, 2003, 31, 5108-5121.
42. X. J. Lu and W. K. Olson, *Nature protocols*, 2008, 3, 1213-1227.
43. G. Zheng, X. J. Lu and W. K. Olson, *Nucleic Acids Res*, 2009, 37, W240-246.
44. C. Maffeo, S. Bhattacharya, J. Yoo, D. Wells and A. Aksimentiev, *Chem Rev*, 2012, 112, 6250-6284.
45. J. Mathe, A. Aksimentiev, D. R. Nelson, K. Schulten and A. Meller, *P Natl Acad Sci USA*, 2005, 102, 12377-12382.
46. Y. Rabin and M. Tanaka, *Phys Rev Lett*, 2005, 94, Artn 148103.
47. T. Hu and B. I. Shklovskii, *Phys Rev E*, 2008, 78, Artn 032901.
48. G. S. Manning, *Berichte der Bunsengesellschaft für physikalische Chemie*, 1996, 100, 909-922.
49. R. W. Wilson and V. A. Bloomfield, *Biochemistry*, 1979, 18, 2192-2196.
50. K. Huber and U. Scheler, *Cur Opin Coll Interface Sci*, 2012, 17, 64-73.
51. G. S. Manning, *Q Rev Biophys*, 1978, 11, 179-246.
52. R. Shusterman, S. Alon, T. Gavrinov and O. Krichevsky, *Phys Rev Lett*, 2004, 92.
53. M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*, Clarendon Press, Oxford, 1986.
54. K. Koga, G. T. Gao, H. Tanaka and X. C. Zeng, *Nature*, 2001, 412, 802-805.
55. X. B. Zhou, G. D. Liu, K. Yamato, Y. Shen, R. X. Cheng, X. X. Wei, W. L. Bai, Y. Gao, H. Li, Y. Liu, F. T. Liu, D. M. Czajkowsky, J. F. Wang, M. J. Dabney, Z. H. Cai, J. Hu, F. V. Bright, L. He, X. C. Zeng, Z. F. Shao and B. Gong, *Nat Commun*, 2012, 3, A949.
56. S. Cabello-Aguilar, S. Balme, A. Abou Chaaya, M. Bechelany, E. Balanzat, J. M. Janot, C. Pochat-Bohatier, P. Miele and P. Dejardin, *Nanoscale*, 2013, 5, 9582-9586.