Published in final edited form as:

Electrophoresis. 2012 December ; 33(23): 3466–3479. doi:10.1002/elps.201200164.

Toward detection of DNA-bound proteins using solid-state nanopores: insights from computer simulations

Jeffrey Comer, Anthony Ho, and Aleksei Aksimentiev*

Department of Physics and Beckman Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

Abstract

Through all-atom molecular dynamics simulations, we explore the use of nanopores in thin synthetic membranes for detection and identification of DNA binding proteins. Reproducing the setup of a typical experiment, we simulate electric field-driven transport of DNA-bound proteins through nanopores smaller in diameter than the proteins. As model systems, we use restriction enzymes *EcoRI* and *BamHI* specifically and nonspecifically bound to a fragment of double-stranded DNA, and streptavidin and NeutrAvidin proteins bound to double- and single-stranded DNA via a biotin linker. Our simulations elucidate the molecular mechanics of nanopore-induced rupture of a protein–DNA complex, the effective force applied to the DNA–protein bond by the electrophoretic force in a nanopore, and the role of DNA–surface interactions in the rupture process. We evaluate the ability of the nanopore ionic current and the local electrostatic potential measured by an embedded electrode to report capture of DNA, capture of a DNA-bound protein, and rupture of the DNA–protein bond. We find that changes in the strain on double-stranded DNA can reveal the rupture of a protein–DNA complex by altering both the nanopore ionic current and the potential of the embedded electrode. Based on the results of our simulations, we suggest a new method for detection of DNA binding proteins that utilizes peeling of a nicked double strand under the electrophoretic force in a nanopore.

Keywords

Nanoscale pore; nucleic acids; molecular dynamics; transmembrane transport

1 Introduction

Nanopores in thin synthetic membranes have emerged as versatile tools for manipulation and detection of nucleic acids [1, 2]. In the most common setup, a thin insulating membrane contains a single nanopore and separates an electrolyte solution into two chambers. An external electric field is applied to transport nucleic acids from one chamber to the other, one molecule at a time. The presence of a nucleic acid molecule in a nanopore modulates the ionic current flowing through it. Such modulations of ionic current can be used to determine the length of the nucleic acid polymer [3], identify the specific variety of the polymer [4, 5] and, potentially, reveal the nucleotide sequence of the polymer [6], although the latter has not been experimentally demonstrated yet [7].

The potential applications of nanopores in life sciences extend beyond DNA and RNA sequencing. At any given moment within a cell, a myriad of protein machines interact with DNA, enabling and regulating gene expression, initiating and performing DNA replication

*Corresponding author. Address: Loomis Laboratory of Physics, 1110 West Green Street, Urbana, IL 61801, U.S.A., Tel.: (217)333-6495, Fax: (217)244-6978.

and repair, and remodeling the spatial organization of the genome. Solid-state nanopores offer several plausible strategies for detection and identification of proteins bound to DNA [8–13]. The key enabling factors are the localization of the electric field in a nanopore to the volume of a single protein and the unidirectional character of DNA motion through the nanopore. The basic idea is to perform nanopore translocation experiments using DNA molecules of known sequences extracted from living cells. As DNA transits the nanopore from one compartment to the other, the proteins bound to DNA are detected and identified as they sequentially interact with the nanopores. The sequence of DNA fragments that have proteins bound to them or, ultimately, the identity of proteins bound to specific DNA sequences provide a snapshot of a cell's activity that is difficult to obtain using existing methods.

Depending on the size of the nanopore, two approaches are possible, Fig. 1. In the first approach [11, 12], DNA and the attached proteins are threaded through a pore that is wider than the largest DNA-protein assembly, Fig. 1a. In this case, one can detect the presence of one or more proteins by monitoring fluctuations of the ionic current, which provide information on the volume inside the nanopore excluded from ionic conductance. In principle, such an approach permits barcoding DNA for genotyping [14], using proteins as fiducial marks. Although the high speed of DNA transport in such free translocation experiments [15] may present a problem for protein detection, optical or magnetic tweezers can be used to displace DNA through a nanopore in a controlled manner [16]. Whereas detection of single proteins has already been demonstrated using this method [13], identifying proteins via ionic current measurement is difficult as many proteins have similar shapes and their charged surfaces modulate the current in a nontrivial way.

The second approach is to use a pore that is barely wide enough to admit a DNA molecule [8–10], Fig. 1b. In such a setup, a DNA molecule with bound proteins pauses every time a protein encounters the nanopore, providing ample time to perform DNA or protein identification measurements. Sensing proteins using ionic current alone may be difficult, as the protein is positioned at the mouth of the nanopore while the DNA already occupies a considerable fraction of the nanopore volume. However, other sensing modalities are possible. Thus, one can infer information about the protein by measuring the lifetime of the DNA-protein complex under the force applied to DNA by the electric field in the nanopore. Furthermore, as the pore constriction has dimensions comparable to the size of a DNA molecule, one may be able to extract some information about the nucleotide composition of the DNA fragment confined in the pore constriction. Such information, even if not precise enough to call the nucleotide sequence, may still be sufficient to identify the approximate location along the DNA where the translocation pauses and hence the protein bound upstream from that sequence. Finally, one can explore other sensing modalities, for example, by incorporating thin electrodes within the pore membrane [17, 18], which can report of the presence, conformation, and possibly sequence of the DNA fragment confined in the constriction. Recent work in the area of nanowire [19] and graphene nanopores [20–23] makes such sensing measurement within the grasp of present technology.

In this manuscript, we explore solid-state nanopores for detection of proteins by means of all-atom molecular dynamics simulations that link the rupture of the protein–DNA complex, nanopore ionic current and local electrostatic potential to the conformation and the effective force on DNA in a nanopore. We focus here on nanopores that have diameters smaller than the proteins they are meant to detect. As model proteins, we use restriction enzymes *EcoRI* and *BamHI* bound to double-stranded DNA and streptavidin and NeutrAvidin bound to double-stranded or single-stranded DNA via biotin a linker. Our simulations determine the feasibility of detection of proteins by measuring the voltage at which rupture of the protein–DNA bond occurs, actual detection of the rupture by measuring nanopore ionic current and

the local electrostatic potential within the membrane, and identifying the sequence and local structure of DNA downstream from the bound protein. Our simulations also reveal possible pitfalls of the approach associated with DNA sticking to the nanopore surfaces, proteins sliding along DNA, and DNA stretching during free translocation.

2 Materials and Methods

General MD methods

All MD simulations were performed using NAMD [24]. Conjugate gradient energy minimization was performed for each system. The systems were then equilibrated using a Langevin thermostat [25] with a damping constant of 5 ps^{-1} to maintain a temperature of 295 K and a Langevin piston pressure control [26] to maintain a pressure of 1 atm. Van der Waals (vdW) energies were computed using a smooth 1.0–1.2 nm cutoff for the restriction enzyme and streptavidin systems. Electrostatic energies were computed via the particle mesh Ewald method [27]. After equilibration, simulations were performed at fixed volume with temperature maintained by Langevin dynamics (using a damping constant of 1.0 ps^{-1}) applied to the membrane only. A constant electric field was applied across each system to obtain the desired biases [28]. The simulations of restriction enzyme and streptavidin systems were carried out using a 1-2-4 fs multiple time stepping scheme; rigidity of water molecules and bonds to hydrogen was not enforced. All simulations of the NeutrAvidin systems were carried out using 2-2-6 fs time stepping with SETTLE [29] and RATTLE [30] algorithms applied to covalent bonds involving hydrogens in water and DNA/protein, respectively, and 0.7–0.8 nm cutoff for computing vdW energies. We have previously shown that the two sets of parameters produce quantitatively similar results [31], the second set of parameters being factor 2 more computationally efficient. Interatomic interactions were calculated as prescribed by the CHARMM force field [32] and the Beglov and Roux parametrization of ions [33].

Restriction enzyme systems

A computational model of a crystalline Si_3N_4 pore was constructed as described in Zhao et al. [8]. The resulting double-cone pore had a minimum diameter of 2.7 nm and was housed in a membrane with a thickness of 10.3 nm and a hexagonal cross section of 117.23 nm^2 . As before [8], the entire Si_3N_4 surface had a uniform charge $100 \text{ mC/m}^2 \approx 0.6 \text{ e/nm}^2$. Using this pore, we created five systems each containing a different complement of molecules. Three of the systems contained DNA and a bound enzyme. One contained DNA but no enzyme. The last contained neither DNA nor enzymes.

To construct the systems having enzyme–DNA complexes, structures of *EcoRI* bound to its cognate sequence (PDB ID: 1CKQ), *BamHI* bound to its cognate sequence (PDB ID: 1BHM) [34], or *BamHI* bound nonspecifically [35] (PDB ID: 1ESG) were placed near the pore. The molecules of dsDNA were extended with random sequences from each end, resulting in 62-basepair molecules with one end threading through the pore and an enzyme bound to the middle. For the system containing only DNA, an 80-basepair molecule of dsDNA was threaded through the pore. Each end of this molecule was bonded to its periodic image, making the molecule effectively infinite. Each system was solvated using the TIPS3P [36] water model of the CHARMM force field. Ions were added to obtain 0.1 M KCl solutions. Such low ion concentration conditions were used in nanopore experiments aimed at detection of restriction enzyme–DNA complexes [8]. The resulting systems contained between 310,000 and 320,000 atoms.

Streptavidin/Neutravidin systems

NeutrAvidin is a deglycosylated version of the egg white glycoprotein avidin with chemical modification on its lysine and arginine residues to reduce its isoelectronic point pI to ~6.3 [37]. Both avidin and NeutrAvidin consist of four identical subunits, each of which is 128 residues-long and contains one biotin binding site. The construction of the all-atom model of the tetrameric NeutrAvidin started with duplicating and transforming PDB entry 2AVI, the x-ray structure of a dimer of deglycosylated avidin with biotin bound to both of its binding sites [38]. The x-ray structure of 2AVI bears a residual *N*-acetylglucosamine moiety per subunit which was removed in our model. Since the x-ray coordinates in 2AVI contain only residues 3–123 of the subunits, we completed the N and C termini of each subunit using the standard CHARMM22 protein patches [32, 39]. To mimic the non-selective chemical modifications on lysines and arginines in the production of NeutrAvidin [40], we replaced 5 positively-charged solvent-exposed residues (4 lysines, 1 arginine) by the neutral acetylated lysines. The choices of the residues to be replaced are based on their solvent-exposed areas and ideally the candidates should not be involved in any salt bridge interactions. A disulfide bond was added between Cys4 and Cys83 residues of each of the subunit.

The coordinates of the biotin molecules in the binding pockets of each subunit were obtained from the same crystal structure. The topology and CHARMM-compatible parameters for biotin were kindly provided by Chris Chipot from an earlier study [41]. The resulting NeutrAvidin-biotin complex was solvated in a cube of water measuring 6.0 nm on a side; K⁺ and Cl[−] ions were added to produce a 1.0 M KCl solution. The system was then minimized and equilibrated using standard thermostat and barostat protocols [42], first with constraints on the biotin and all protein atoms except for the newly extended termini for 0.5 ns, then with constraints on biotin and protein backbone except for the N and C termini for 1 ns, and finally without constraints for another 1 ns. Custom topology and parameter files were made to describe the flexible linker connecting the biotin to DNA. Biotinylated DNA strands were made by combining the structures of biotin, linker and a single DNA strand using custom patches. Analogous streptavidin systems were built using the above protocol, starting from a crystal structure of the core streptavidin reported in [43] (PDB ID: 1MK5) combined with the C terminal from PDB entry 2BC3 [44].

All-atom models of 49-basepair biotinylated dsDNA (either an A•T or G•C homopolymer) bound to streptavidin or NeutrAvidin were combined with the all-atom model of a silicon nitride pore such that about 30 DNA basepairs were threaded through the constriction of the pore. The two pores used in our simulations were made in a 13.0 nm-thick silicon nitride membrane and had a 2.6 nm×2.4 nm constriction, which was located in the middle of the membrane. In the simulations performed using the streptavidin constructs, the pore had a bi-conical shape with a cone angle (the angle between the pore axis and the pore surface) of 20° (see Fig. 8b). In the simulations of the NeutrAvidin constructs, the pore had a parabolic inner surface with a cone angle of 20° near the edge (Fig. 8a). The resulting systems were submerged in 1.0 M KCl solution and simulated using protocols described previously [45]. The simulations of the streptavidin and NeutrAvidin systems were carried out with transmembrane biases of 1 V and 400 mV, respectively.

Force calibration

The effective force exerted by the DNA molecule can be determined by measuring the tension in the flexible linker that connects biotin to DNA. A truncated version of the system described above was used to calibrate the dependence of the extension of the linker on the tension across it. In this system, a biotin molecule, which was bound to the binding pocket of a truncated NeutrAvidin, was connected to a single DNA nucleotide via the linker. The extension of the linker was measured by the separation between two selected atoms on the

linker (the Results for detailed description). In each calibration simulation, a constant force, ranging from 40 to 800 pN in 40 pN intervals, was applied to the O3 atom downstream of the flexible linker on the biotinylated nucleotide, in a direction away from the binding pocket. Harmonic restraints were placed to the heavy atoms of the biotin molecule upstream of the flexible linker. The truncated NeutrAvidin was present to mimic the protein-biotin interaction. Two sets of independent calibration simulations (show in red and blue data points in the plot) were carried out and used to obtain a fitted curve. The latter was used to determine the effective force by measuring the separation between the red and blue atom of the flexible linker in the simulations of complete nanopore/dsDNA-biotin-streptavidin/NeutrAvidin systems.

3 Results and Discussion

Fig. 2 schematically illustrates a sequence of events in an idealized measurement considered in this study. The apparatus contains two chambers separated by a thin (nanometer scale) membrane. The membrane contains a nanopore that is large enough to easily permit the passage of double-stranded DNA, but small enough to prevent the passage of a folded protein. The cross section of such a nanopore is shown in Fig. 2a. Double-stranded DNA with multiple proteins bound (which could be transcription factors, nucleosomes, etc) is added to the solution on one side of the membrane. By applying a voltage bias transverse to the membrane, the DNA, which is negatively charged, can be made to transit the pore.

At the beginning of experiment, the pore is open, which is characterized by a well defined open pore current, I_0 . Eventually, the electric field near the pore opening captures the DNA (Fig. 2b). The DNA threads through the pore (Fig. 2c) until a protein reaches the membrane and halts the translocation (Fig. 2d). The presence of a protein is detected by a change in the current level. The bias applied to the system thus far has been less than that required to dissociate the protein from the DNA; however, by slowly increasing the bias (Fig. 2e), we eventually induce rupture of the DNA-protein bond (Fig. 2f). After the rupture event is detected, the bias is then reduced to its initial value (Fig. 2g). The cycle is continued until all proteins have been removed from the DNA strand. Note that Fig. 2 represents only a desired outcome of nanopore detection experiments and, as detailed below, our simulations provide only partial support for realizing the method depicted in the figure in practice.

The schematic in Fig. 2 features a feedback control system where a single input—the transmembrane bias—is adjusted based on measurements of the two outputs—the ionic current through the pore and electrostatic potential on an electrode embedded in the membrane. Fig. 2h shows the transmembrane bias input, whereas Fig. 2i,j schematically shows the ionic current and electrostatic potential outputs. The same measurement scheme can, in principle, be employed for genotyping DNA using a set of restriction enzymes of known affinities to specific DNA sequences.

For the above approach to be successful, the following four conditions must be met. One must be able to (i) detect halting of DNA due to a bound enzyme reaching the pore opening, (ii) detect rupture of the complex, (iii) distinguish the enzymes by either measuring the rupture bias or the DNA sequence downstream from the bound protein, and (iv) ensure that the system can be reset for the next protein in the detection cycle. Below we describe a set of MD simulations designed to provide insights into all four aspects of the process. For this purpose, we will use two types of DNA binding proteins: restrictions enzymes and streptavidin/NeutrAvidin constructs in the low salt (0.1 M KCl) and high salt (1 M KCl) concentration regimes, respectively. Whereas high-salt conditions are preferred in experiments and have been used for optimization of the detection, the low-salt conditions

are most relevant if one wants to eventually detect generic DNA binding proteins taken from a cell.

3.1 Rupture voltage as a measure of strength of a DNA-protein bond

Dissociation of a protein–DNA bond under an external force is a stochastic process that can be modeled according to the Kramers theory of diffusive crossing of a free energy barrier [47–49]. The bound state is seen as a free energy well separated from the dissociated state by a barrier, while an external force reduces the effective height of the barrier, increasing the probability of escape from the bound state by diffusion. Sufficiently large external forces may remove the barrier completely, in which case the system can freely diffuse out of the bound state.

Typically, the mean time to dissociation after the application of the external force increases very sharply with magnitude of the external force [49]. Although it is also not clear whether the experiments discussed below can be modeled by assuming a high barrier for dissociation or operate in the regime where this barrier is low compared to the thermal energy, experiments in which nanopores are used to apply external forces to a protein–DNA bond show a strong dependence of dissociation rate on the external force (transmembrane bias) [10]. The dramatic change of dissociation rate with the transmembrane bias observed in experiments has been characterized as a “threshold bias”. One should note, however, that rupture of the protein–DNA bond remains a stochastic process and that rupture may occur spontaneously at bias values below the “threshold”. Moreover, the exact value of the threshold bias depends on the time scale over which the system is observed, since longer time scales will allow the observation of more dissociation events.

Here, we discuss experimental and computational studies that indicate that the proteins bound to DNA can indeed be distinguished from one another by measuring the transmembrane bias at rupture. We have chosen to focus on restriction enzymes *EcoRI* and *BamHI*, which have very similar binding sites although their amino acid sequences show little similarity [34]. Experiments showed that the threshold bias for dissociating *EcoRI* from its cognate sequence (GAATTC) was much larger than that for dissociating *EcoRI* from sequences that differ by one basepair [8, 10]. No threshold bias existed for a sequence that differed by two basepairs [8], similar to the case of non-specific binding. Furthermore, it was demonstrated that dissociation events involving *EcoRI* were distinguishable from those involving *BamHI* due a significant difference in the bias at which the events occur [8].

Although MD simulations were limited to nanosecond time scales and did not obtain a statistical sampling of rupture times, they nevertheless revealed rupture kinetics that agree qualitatively with experimental results. Fig. 3a illustrates a typical simulation system, where a fragment of double-stranded DNA with a single enzyme bound to it is initially placed halfway through the nanopore. Under the influence of an external electric field, DNA translocates through the nanopore until the protein encounters the membrane (Fig. 3b), which halts DNA transport. Eventually, the DNA-enzyme bound breaks (Fig. 3c) and the DNA resumes translocation through the nanopore (Fig. 3d,e). An animation illustrating this simulation trajectory is available in the Supporting Information.

We have performed several similar simulations at different constant values of the transmembrane bias for three systems containing the same nanopore but different enzyme–DNA complexes. The first enzyme was *EcoRI*, which was previously used in experimental and computation work on nanopore force spectroscopy [8, 10]. The second system contained *BamHI* bound to cognate fragment of dsDNA. Experimentally, *BamHI* was found to have a lower threshold bias than *EcoRI*. The third system contained *BamHI* nonspecifically bound

to DNA, which is expected to have a very low rupture threshold, as it is thought to be able to diffuse along the DNA. Fig. 4a–c illustrates the protein–DNA contacts in all three systems.

The results of these simulations are summarized in Fig. 4d–f, where we plot, as a function of time, the distance between a particular part of the enzyme and a portion of the DNA sequence to which the enzyme is bound. To characterize the enzyme–DNA separation, we computed the distance between the center of mass of the glutamic acid (E113) in enzyme's active site [34] and the phosphate of the third nucleotide in the cognate sequence (GAATTC for *EcoRI* and GGATCC for *BamHI*). These distances were chosen because they have comparable values for the two enzymes due to the strikingly similar structure of their active sites [34]. The equilibrium separation was ~ 0.8 nm for the cognate systems and ~ 2.3 nm for the nonspecific *BamHI* system, which reflects a significant difference in the enzyme's conformation upon binding a cognate substrate. The rupture of an enzyme–DNA bond can be determined from the plots as a rapid increase of the enzyme–DNA separation, typically exceeding 0.3 nm (one DNA base pair) in less than 5 ns.

Fig. 5 clearly shows the difference in the rupture kinetics among the three systems. As expected from the experiments, the nonspecifically bound *BamHI* complex ruptures quickly for even the smallest biases. With a bias of 1.0 V, the complex ruptured in under 30 ns. In contrast, the *BamHI* cognate complex did not rupture in 70 ns of simulation at this bias. Supposing that the bias required to cause rupture within a given time is indicative of the strength of the bond, we find that the *EcoRI* cognate complex is most strongly bound, followed by the *BamHI* cognate complex, and that the *BamHI* nonspecific complex is the most weakly bound. Thus, the ranking of the bond strengths found in the simulation agreed with the ranking of the threshold biases found in experiment. Working with a $3.4 \text{ nm} \times 4.5 \text{ nm}$ pore in Si_3N_4 membrane having a thickness of 12 ± 2 nm, Zhao et al. [8] identified threshold biases of 1.80 V and 1.58 V, respectively, for dissociating *EcoRI* and *BamHI* from their respective cognate sequences. Furthermore, they found that single-basepair substitutions to the cognate sequence of *EcoRI* reduced the threshold bias from 2.1 V to 1.1 V in a $3.4 \text{ nm} \times 4.7 \text{ nm}$ pore [10]. A similar reduction is expected for *BamHI* non-specifically bound to DNA.

Although the results of our simulations are rather encouraging, they also reveal potential pitfalls of the approach. First of all, rupture of a bond subject to an external force is a stochastic process [50, 51]. Thus, variations in the rupture voltage are expected from one measurement to another even for the same system. Our two independent simulations of the same cognate *BamHI* systems at 2 V showed rather similar times to rupture, Fig. 4e. Greater variations in the kinetics of rupture can be expected for systems that permit the protein–DNA complex to encounter the nanopore in multiple conformations. Most likely, a large number of repeat measurements will be required to pinpoint rupture voltage characteristics of a specific protein–DNA complex in experiment. Second, interactions between DNA and the pore wall may obscure measurements of the strength of a protein–DNA bond, in particular, at low transmembrane biases. For example, our MD simulations of dsDNA translocation in the absence of bound proteins revealed a stick-slip character of dsDNA motion for voltages of 1 V and below [52–55], which is associated with dsDNA interaction with the pore wall (see also Fig. 6f). Treatment of the pore surface with a coating that eliminates DNA sticking within the nanopore (and protein sticking to the membrane) will be required to perform reliable measurements of the rupture voltage.

3.2 Ionic current reports on protein loading, not rupture

Next, we examine the effect of DNA capture, protein lodging at the nanopore mouth and rupture of the protein–DNA bond on the nanopore ionic current. Following the protocol described earlier [45, 56], ionic

$$I(t+\Delta t/2) = \frac{1}{\Delta t \ell} \sum_{i=1}^N q_i (\zeta_i(t+\Delta t) - \zeta_i(t)), \quad (1)$$

current was computed as

where

$$\zeta_i(t) = \begin{cases} z_i(t), & |z_i(t)| \leq \ell/2 \\ -\ell/2, & z_i(t) < -\ell/2 \\ \ell/2, & z_i(t) > \ell/2 \end{cases} \quad (2)$$

and z_i and q_i were respectively the z coordinate and charge of ion i , N was the total number of ions, and $\Delta t = 5$ ps was the time between trajectory frames. In our setup, the z axis is normal to the membrane and parallel to the applied electric field. To minimize thermal noise, the current was computed over a region within the pore ($-\ell/2 \leq z \leq \ell/2$ where $\ell = 9$ nm)

Table 1 shows the mean current values for fragments of the MD trajectories along with the applied bias and the associated conductance. The time shown in the table is the amount of time over which the mean current was computed. This is the time excluding the initial current transient until rupture, for Simulations EC-2V, BC-1.5V, BC-2Va, BC-2Vb, BN-1V and BN-1.5V where rupture occurred. For Simulations O-2V, O-4V, D-1V, D-2V, BC-1V and BN-0.5V, it is the time from the end of the initial current transient to the end of the simulation. Some simulations have been omitted from the table because there was insufficient time to establish a steady current.

Simulations O-2V and O-4V, in which no DNA or enzyme was present, show that the current through the empty pore does not scale linearly with the voltage, giving conductances of 0.76 ± 0.01 nS and 0.63 ± 0.01 nS for biases of 2.0 and 4.0 V, respectively. The presence of DNA in the pore caused an increase in the conductance of about 60% as can be seen by comparison of Simulations O-2V and D-2V. Such current enhancements have been observed before [46, 52, 57] and are due to the increased ion concentration in the vicinity of the DNA.

In Simulations EC, BC, and BN, an enzyme blocks the entrance to the pore, reducing the conductance. Apart from the clear reduction of the current from its value in the absence of the enzymes, there is no clear relation between the conductance and the type of enzyme–DNA complex. For example, although Simulations BC-2Va and BC-2Vb were performed for the same complex (*Bam*HI–DNA) and under the same bias, there was a significant difference in the conductances due to differences in the conformations of the complex.

Thus, our simulations have demonstrated that capture of a dsDNA molecule by the pore is associated with increase in the nanopore conductance, which we attribute to a low (0.1 M) bulk ion concentration and the negative charge of the nanopore. When a protein bound to DNA lodges at the nanopore mouth, the current is reduced, which is a signal that can be used to trigger the voltage ramp. Within the limited sampling afforded by the MD method, we find the current blockades produced by different enzymes to be indistinguishable from one another amid variations in the current blockade produced by the same enzyme in different conformations.

Note that the precise shape of the ionic current recording will depend on several factors, most importantly, the ion concentration, the size of the nanopore, the size of the protein and

the charge of the nanopore [58]. Thus, in a higher molarity buffer (e.g. 1 M KCl), capture of dsDNA will decrease the ionic current relative to the open pore level [46], whereas protein lodging will manifest itself in an additional current decrease due to reduced pore cross section [58]. For the conditions considered in our simulations, the protein/DNA level is similar to the open pore level. Using slightly different pore diameter, surface treatment or salt conditions should make the open pore current distinguishable from the DNA/protein current, based on the current knowledge of the mechanism of the ionic current blockades and enhancements [46, 58]. However, even in the worst-case scenario (which is actually realized in our simulations), an ion current signal from a DNA strand containing many proteins can be distinguished from a signal originating from two or more independent translocation events. In the first case, the ionic current will alternate between the levels above and close to the open pore current with the rate determined by the DNA translocation speed and not by the concentration of the DNA-protein molecules in the sample.

Based on our previous simulations of a dsDNA system [53], we expect a higher sensitivity of the measurement at higher salt concentrations. Thus, in the case of dsDNA, the current enhancement relative to the open pore level at 0.1 M KCl was 36%, whereas at 1 M, the current blockade (reduction) was 62% of the open pore. In terms of absolute magnitudes, the current change with the addition of dsDNA threaded through the pore was from 1.15 ± 0.04 to 1.57 ± 0.05 nA at 0.1 M KCl, while at 1.0 M KCl it was from 7.5 ± 0.2 and 2.8 ± 0.3 nA.

Our simulations, however, suggest that detecting protein rupture by measuring the ionic current may be difficult, as protein does not dissociate from the DNA after the DNA–protein bonds has ruptured, certainly not within the time scale of our simulations, see Fig. d,e. To get more insight into this process, we carried out a set of additional MD simulations using a frictionless, featureless surface—the phantom pore [59]—which minimized the effect of DNA and protein adsorption to the nanopore surface. In all phantom pore simulations (data not shown) the protein was observed to remain loosely attached to DNA after the specific protein–DNA bond was broken. Recent experiments using DNA curtains [60] revealed similar behavior when a translocase encountered a restriction enzymes bound on the same DNA track: the enzyme slid along the DNA together with the translocase.

3.3 Effective force on the protein–DNA bound

To gather insights into the effective force applied to the DNA-protein bond, we simulated DNA-biotin-streptavidin and DNA-biotin-NeutrAvidin constructs in a solid-state nanopore subject to electric field of different strength. Such systems have been extensively explored in conventional [41, 51, 61–63] and nanopore force spectroscopy measurements [64–66].

Figures 6a–e illustrate a typical chain of microscopic events following the capture of dsDNA with a NeutrAvidin protein attached to one of its strands via a biotin linker. Subject to a 400mV bias, dsDNA permeates through the pore until the NeutrAvidin reaches the pore surface, Fig. 6a. Subsequently, dsDNA stretches, Fig. 6b,c, and adheres to the pore surface, Fig. 6c,d. Fig. 6f quantitatively characterizes the process of dsDNA permeation by plotting the number of DNA bps remaining above the pore constriction versus the simulation time. For the MD trajectory shown in Figures 6a–e, this number reaches a constant, non-zero value indicating that DNA had reached a steady-state stretching. Similar behavior was observed in the simulation of the poly(GC)-biotin-NeutrAvidin construct at 400mV and in the simulation of poly(AT)-biotin-streptavidin construct at 1V. One could clearly identify steps in the traces shown in Fig. 6f, which reflects the stick-slip character of dsDNA motion through pores smaller in diameter than B-form DNA. Such steps have also been observed in MD simulations DNA hairpin transport through nanopores [52].

However, not all MD simulation resulted in trapping dsDNA. Thus, the biotin-streptavidin bond broke in the simulation of the poly(GC)-biotin-streptavidin at 1V and dsDNA was observed to escape the pore constriction.

To determine which factors influence the outcome of the trapping experiments, we determined the force exerted by dsDNA on biotin in the binding pocket of streptavidin and NeutrAvidin. In principle, the force acting on every atom at every step of an MD simulation is known. However, extracting the effective force between groups of atoms, in our case, between the biotin and its binding pocket, is difficult as it requires precise estimation of the force between these groups of atoms and the solvent [8]. In our case, we could determine the effective force by measuring tension in the linker that connects the biotin to one strand of dsDNA from the extension of the linker. The atomic structure of the linker is shown in Fig. 7a. To make quantitative estimates possible, we performed an independent set of calibration simulations that determined the relationship between the equilibrium extension of the linker subject to a mechanic force and the value of that force. Fig. 7b details the results of the calibration simulations; the inset illustrates the system used for the calibration simulations. The tension was found to non-linearly increase with extension of the linker, reflecting molecular-scale elasticity of the latter.

Fig. 6g plots the force acting on the biotin along the direction of streptavidin or NeutrAvidin binding pocket for the four simulations characterized in Fig. 6f. The plot reveals that the breakup of the streptavidin-biotin bond in the simulation of poly(GC)-biotin-streptavidin construct under a 1V bias is caused by the force in excess of 400pN, sustained for at least 10 ns. Similar by magnitude but much shorter (< 2 ns) by duration fluctuations of the force were observed earlier in the same trajectory and in the 1V trajectory of the poly(AT)-biotin-streptavidin system, which did not lead to the breakup of the construct. This observation is in agreement with the theoretically derived dependence that predicts exponential decrease of the lifetime of a bond under external force with the magnitude of that force [50, 51].

The steady-state force in the simulation where the rupture of the constructs was not observed does not scale with the magnitude of the transmembrane bias and strongly depends on the conformation of dsDNA in the pore. Thus, for the simulation illustrated in Figs. 6a–e, the force increases roughly twofold (from 60pN to 120pN) when the NeutrAvidin reaches the top of the membrane, Fig. 6a. The force remains roughly at the same level as the dsDNA stretches until pairing of the bases in the pore constriction breaks and the force transiently increases to 250pN. Following that, the force decreases to 50 pN as dsDNA adopts a new conformation that allows it to interact more strongly with the nanopore surface, Fig. 6d,e. Thus, our simulations have demonstrated that interactions of dsDNA with the nanopore surface can substantially reduce the effective force applied to the biotin-streptavidin (or biotin-NeutrAvidin) bond, thereby increasing the lifetime of the bond by many orders of magnitude. We speculate this to be the mechanism explaining the existence of a sharp voltage threshold for rupture of DNA-protein bond reported earlier [8, 64]. Note that such direct surface-DNA interactions reduce the effective force in addition to the electro-osmotic effect [67] that can be particularly large when the nanopore surface is negatively charged [68] but is negligible in very narrow pores as long as the no-slip boundary condition holds for the solvent.

We have used the conformation of the dsDNA-biotin-NeutrAvidin systems to simulate the force and the ionic current blockades produced by a single DNA strand attached to NeutrAvidin. The NeutrAvidin-bound poly(dA) and poly(dG) ssDNA systems are built by removing the non-biotinylated strands of the NeutrAvidin-bound AT and GC dsDNAs followed by a short equilibration to allow the resolution of the ssDNA. Following that, the systems were simulated for 50 ns under a 400 mV bias, see Fig. 8c. In both systems, we

found that removing the non-biotinylated strand did not change the effective force, which indicates that the conformation of the biotinylated DNA strand is the major determinant of the effective force applied to biotin in experiments performed using such narrow pores. This surprising result appears to be due to adhesion of the non-biotinylated strand to the pore surface. Prior to removal of this strand, the effective force applied to the non-biotinylated was balanced by the interaction of the strand with the nanopore surface. Therefore, removal of this strand had a minor effect on the force experienced by the biotinylated strand and hence the DNA-protein bond.

3.4 Detection of local DNA structure downstream from protein binding

Transient arrest of DNA motion through a nanopore by a bound protein exposes a short DNA fragment downstream from the protein binding site to the narrowest part of the pore. If one could determine the sequence of that fragment via ionic current or other measurement, one would immediately learn the sequence of the upstream binding site, which, in many cases, would be sufficient to identify the bound protein. While sequencing DNA by measuring ionic current in solid-state nanopores remains elusive, here we demonstrate that the nanopore ionic current and the local electrostatic potential can report the local structure of DNA in the pore constriction, which can facilitate protein detection.

In all our MD simulations, dsDNA was observed to stretch shortly after the motion of dsDNA was arrested by a bound protein. Although the bias used in the majority of our simulations was 1 V and above, we also observed DNA stretching at 400 mV in our simulations of NeutrAvidin-dsDNA systems, Fig. 6a–e. Based on the results of direct force measurement in a nanopore [16], one can expect the effective force on dsDNA to exceed the overstretching threshold of ~ 60 pN at transmembrane voltages over ~ 300 mV.

Fig. 8 summarizes the dependence of the ionic current blockade on the sequence and conformation of DNA in our simulations of streptavidin/NeutrAvidin systems. As these simulations were performed at 1 M KCl, the presence of DNA in a nanopore was observed to reduce the nanopore current. For this analysis, we defined the pore constriction as the volume within 1.5 nm of the narrowest part of the pore. For dsDNA, the nanopore current was observed to vary within each simulation trajectory. However, the ratio of blocked and open-pore currents (I/I_0) appear to correlate with the average separation of the DNA nucleotides in the pore constriction for both pores used in these simulations, Fig. 8a,b. For B-form DNA, Fig. 8a, data cluster around $I/I_0 = 0.2$. For stretched DNA, 8b, there is a linear dependence of the ionic current blockade and the average separation of the nucleotides. In the case of ssDNA, Fig. 8c, the blockades are shallow ($I/I_0 > 0.75$) and do not appear to depend on DNA stretching. Within the limited sampling of our MD runs, the simulations reveal no sequence specificity of the ionic current.

Thus, although the local sequence of dsDNA in the pore constriction does not appear to alter the ionic current, stretching of dsDNA does, and may be used for sequence detection. Indeed, as the overstretching threshold for short fragments of dsDNA sensitively depends on their nucleotide sequence [69, 70], measurements of the voltage threshold for overstretching dsDNA in a nanopore can provide information about the DNA sequence confined between pore constriction and protein binding site.

Stretching of dsDNA may also help to identify rupture of a protein–DNA bound in a nanopore force spectroscopy measurement. Indeed, upon rupture of a DNA-protein bond, one can expect tension in DNA to drop substantially. As DNA resumes translocation, B-form DNA fragments will occupy pore constriction, and the blockade current will return to the level observed before the DNA-bound protein entered the nanopore. This scenario is depicted in our scheme of an idealized experiment, Fig. 2. Our simulations give only limited

support for the above hypothesis: the mean distance between base pairs in the pore constriction was observed to decrease after dissociation of a restriction enzyme, Fig. 9, but did not return to the B-form value. Future experiments and simulations will determine if DNA can be locally stretched in free-translocations experiments due to hydrodynamic drag apposing DNA motions, stick-slip motions or due to friction with non-specifically bound protein retaining loosely bound to DNA after specific bond rupture.

Whereas the difference between stretched and B-form DNA is considerable (20% of the open pore current), the difference between B-form DNA and ssDNA is even bigger (60% of the open pore current). This difference can be exploited for protein detection. In our simulations of dsDNA-streptavidin/NeutrAvidin systems, the free (non-biotinylated) strand of dsDNA was observed to begin to peel off the biotinylated strand, see Fig. 8b. Limited by the time scale of our simulations, complete removal of the non-biotinylated strand was not observed, but, given the overstretched conformation of dsDNA and the electrophoretic force in a nanopore, one can expect the free strand to peel off completely and exit the nanopore constriction, leaving a single DNA strand attached to the protein, Fig. 8c. Such a transition from dsDNA to ssDNA can be exploited for detection of proteins that nick DNA at the binding site and in assays that employ nicked dsDNA or use endonuclease to nick DNA downstream from the expected protein binding sequence.

3.5 Embedded electrodes may assist local structure detection

Finally, we estimate the effect of the local DNA conformation on the electrostatic potential on the embedded electrode. The potential at the electrodes of the metallic membrane gives an approximate measure of the charge density of the DNA along the pore axis. This charge density can then be used to infer the extent to which the DNA is stretched. Here we develop a simple model validated by MD simulations that approximately relates the potential at the electrodes to the extension of the DNA. We approximate the dsDNA as a uniformly charged cylinder of infinite length. The diameter of the cylinder is chosen to be 2.0 nm, near the diameter of the dsDNA, although any reasonable choice of this diameter (between 0.5 and 3.0 nm) gives essentially the same results at this level of approximation. The charge per basepair of the dsDNA is $-e$, but the axial length per basepair changes as tension is applied along the DNA strand. The potential at an electrode with a distance R from the center of the pore is given by

$$\phi(\mathbf{R}, t) = \frac{1}{4\pi\epsilon_0} \frac{Q}{\pi D^2 L(t)/4} \int_{\text{cylinder}} d\mathbf{r} \frac{\exp(-\kappa|\mathbf{r} - \mathbf{R}|)}{|\mathbf{r} - \mathbf{R}|} \quad (3)$$

where $\epsilon_0 = 80$ is the dielectric constant of the solvent, $Q = -2e$ is the charge of a DNA basepair, D is the diameter of a DNA basepair, $L(t)$ the length of a DNA basepair along the pore axis at time t , and $\kappa = 0.3163 \text{ \AA}^{-1} [C/(\text{mol/L})]^{1/2}$ is the Debye screening length at salt concentration C . Our electrodes are located 2.19 nm from the pore axis. Solving the integral numerically with $R = 2.19 \text{ nm}$, we obtain

$$\phi(t) = \frac{-7.4 \text{ mVnm}}{L(t)}, \quad (4)$$

We use this expression below for a proof-of-principle demonstration, which is valid in the limit of infinitely small electrodes, neglecting the effect of leads and the method to read out the voltage. The actual measured signal is expected to be much smaller, but still should reflect the change in the DNA charge density. More accurate estimates of the potential

require a self-consistent approach [17, 71], which can also be performed at the device level [72].

We tested the model by comparing its predictions to those of simulation. First, the potential on the electrodes of the capacitor were computed as in Sigalov et al. [73] from the positions and charges of all atoms of the DNA. As suggested by Eq. 4, we calculated the inverse basepair length near the center of the pore. The end-to-end vector of the ten basepairs nearest the center of the pore was calculated. The component of this vector along the z axis was divided by the number of basepairs (10) to yield the average basepair length along the z axis, $L(t)$.

The electrode potential was nearly inversely proportional to the extension of the DNA in all simulations. Fig. 9 shows the inverse length of a basepair ($1/L(t)$) overlaid by the potential on the top capacitor electrode ($\phi(t)$) scaled by -7.4 mV nm for two exemplary systems. For DNA fragment of a known force-extension dependence [70, 74, 75], one may be able to determine the force on the molecule by measuring the electrode potential. One can also monitor the change in the membrane potential concurrently with the change in the ionic current to detect sequence-dependent DNA overstretching. An even bigger effect can be expected when one of the strands of dsDNA peels off, reducing the charge density more than twofold. Note that in the case of asymmetric ion concentration conditions and asymmetric electrode placement, the change in the nanopore ionic current may have an additional effect on the electrode potential [19].

4 Conclusions

We have used all-atom molecular dynamics simulations to assess the feasibility of employing solid-state nanopores for detection and identification of proteins bound to DNA. We considered a situation schematically depicted in Fig. 2, where multiple proteins bound to the same DNA molecule are sequentially removed from that DNA molecule as it transits a nanopore.

We have determined that measurements of nanopore ionic current can report on presence of a DNA-bound protein in the proximity of a nanopore entrance. Furthermore, our simulations suggest that specific protein–DNA bonds may be identified by measuring a typical voltage required to rupture the bond using a nanopore. At the same time, our MD simulations suggest that detecting rupture by measuring ionic current is difficult, as the protein remains near pore constriction after the rupture of the specific DNA-protein bound. Indirectly, rupture of a bond can be inferred by measuring ionic current and electrostatic potential signals associated with stretching of dsDNA fragment downstream from the protein binding site. Nicking of the DNA strand at or immediately downstream from the binding site greatly facilitates protein detections as, under the electrophoretic force in a nanopore, the nicked strand peels off from the intact strand, giving ample signal for detections.

Among possible difficulties with the approach depicted in Fig. 2, our MD simulations identified nonspecific interactions between DNA and the pore surface, sliding of proteins along DNA after dissociation of their cognate bonds, and possible stretching of dsDNA in free-translocation experiments.

Despite these shortcomings, the ability of solid-state nanopores to apply a local force of controlled magnitude to a DNA-protein bond, the unidirectional character of DNA motion through the nanopore, and inherently high-throughput nature of the measurements make solid state nanopores an attractive platform for future development of protein detection strategies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors gladly acknowledge stimulating discussions with Greg Timp and members of his research group. This work was supported by the grants from the National Science Foundation (DMR-0955959 and PHY-0822613) and the National Institutes of Health (R01-HG005115). The supercomputer time provided through TeraGrid Allocation grant MCA05S028. A.A. would like thank the Department of Bionanosciences at Delft University for hospitality and support from the Netherlands Organisation for Scientific Research (NWO).

References

1. Li J, Stein D, McMullan C, Branton D, Aziz MJ, Golovchenko JA. *Nature*. 2001; 412:166–169. [PubMed: 11449268]
2. Dekker C. *Nature Nanotech*. 2007; 2:209–215.
3. Storm AJ, Chen JH, Zandbergen HW, Dekker C. *Phys Rev E*. 2005; 71:051903–051913.
4. Skinner GM, van den Hout M, Broekmans O, Dekker C, Dekker NH. *Nano Lett*. 2009; 9:2953–2960. [PubMed: 19537802]
5. Wanunu M, Dadosh T, Ray V, Jin J, McReynolds L, Drndic M. *Nature Nanotech*. 2010; 5:807–814.
6. Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, et al. *Nature Biotech*. 2008; 26:1146–1153.
7. Venkatesan BM, Bashir R. *Nature Nanotech*. 2011; 6:615–624.
8. Zhao Q, Sigalov G, Dimitrov V, Dorvel B, Mirsaidov U, Sligar S, et al. *Nano Lett*. 2007; 7:1680–1685. [PubMed: 17500578]
9. Hornblower B, Coombs A, Whitaker RD, Kolomeisky A, Picone SJ, Meller A, et al. *Nature Mater*. 2007; 4:315–317.
10. Dorvel B, Sigalov G, Zhao Q, Comer J, Dimitrov V, Mirsaidov U, et al. *Nucl Acids Res*. 2009; 37:4170–4179. [PubMed: 19433506]
11. Smeets RMM, Kowalczyk SW, Hall AR, Dekker NH, Dekker C. *Nano Lett*. 2009; 9:3089–3095. [PubMed: 19053490]
12. Kowalczyk SW, Hall AR, Dekker C. *Nano Lett*. 2010; 10:324–328. [PubMed: 19902919]
13. Spiering A, Getfert S, Sischka A, Reimann P, Anselmetti D. *Nano Lett*. 2011; 11:2978–2982. [PubMed: 21667921]
14. Singer A, Rapireddy S, Ly DH, Meller A. *Nano Lett*. 2012; 12:1722–1728. [PubMed: 22352964]
15. Storm AJ, Storm C, Chen J, Zandbergen H, Joanny JF, Dekker C. *Nano Lett*. 2005; 7:1193–1197. [PubMed: 16178209]
16. Keyser U, Koeleman B, Dorp S, Krapf D, Smeets R, Lemay S, et al. *Nature Phys*. 2006; 2:473–477.
17. Gracheva ME, Xiong A, Aksimentiev A, Schulten K, Timp G, Leburton JP. *Nanotechnology*. 2006; 17:622–633.
18. Luan B, Peng H, Polonsky S, Rossnagel S, Stolovitzky G, Martyna G. *Phys Rev Lett*. 2010; 104:238103. [PubMed: 20867275]
19. Xie P, Xiong Q, Fang Y, Qing Q, Lieber CM. *Nature Nanotech*. 2012; 116:119–125.
20. Garaj S, Hubbard W, Reina A, Kong J, Branton D, Golovchenko JA. *Nature*. 2010; 467:190–193. [PubMed: 20720538]
21. Merchant CA, Healy K, Wanunu M, Ray V, Peterman N, Bartel J, et al. *Nano Lett*. 2010; 10:2915–2921. [PubMed: 20698604]
22. Schneider GF, Kowalczyk SW, Calado VE, Pandraud G, Zandbergen HW, Vandersypen LMK, et al. *Nano Lett*. 2010; 10:3163–3167. [PubMed: 20608744]
23. Venkatesan BM, Estrada D, Banerjee S, Jin X, Dorgan VE, Bae MH, et al. *ACS Nano*. 2012; 6:441–450. [PubMed: 22165962]

24. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. J Comput Chem. 2005; 26:1781–1802. [PubMed: 16222654]
25. Brünger, AT. X-PLOR, Version 3.1: A System for X-ray Crystallography and NMR. The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University; 1992.
26. Martyna GJ, Tobias DJ, Klein ML. J Chem Phys. 1994; 101:4177–4189.
27. Batcho PF, Case DA, Schlick T. J Chem Phys. 2001; 115:4003–4018.
28. Gumbart J, Khalili-Araghi F, Sotomayor M, Roux B. Biochim Biophys Acta Biomembr. 2011
29. Miyamoto S, Kollman PA. J Comput Chem. 1992; 13:952–962.
30. Andersen H. J Comp Phys. 1983; 52:24–34.
31. Wells, DB.; Bhattacharya, S.; Carr, R.; Maffeo, C.; Ho, A.; Comer, J., et al. Nanopore-based technology: single molecule characterization and DNA sequencing. Vol. 870, chapter 10. Humana Press; 2012. p. 165-186.
32. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck J, Field MJ, et al. J Phys Chem B. 1998; 102:3586–3616.
33. Beglov D, Roux B. J Chem Phys. 1994; 100:9050–9063.
34. Newman M, Strzelecka T, Dorner L, Schildkraut I, Aggarwal A. Science. 1995; 269:656–663. [PubMed: 7624794]
35. Viadiu H, Aggarwal A. Molecular cell. 2000; 5:889–895. [PubMed: 10882125]
36. Neria E, Fischer S, Karplus M. J Chem Phys. 1996; 105:1902.
37. Marttila AT, Laitinen OH, Airenne KJ, Kulik T, Bayer EA, Wilchek M, et al. FEBS Lett. 2000; 467:31–36. [PubMed: 10664451]
38. Livnah O, Bayer EA, Wilchek M, Sussman JL. Proc Natl Acad Sci U S A. 1993; 90:5076–5080. [PubMed: 8506353]
39. MacKerell AD, Feig M, Brooks CL III. J Comput Chem. 2004; 25:1400–1415. [PubMed: 15185334]
40. Bayer E, De Meester F. personal communication.
41. Dixit S, Chipot C. J Phys Chem A. 2001; 105:9795–9799.
42. Comer, J.; Wells, DB.; Aksimentiev, A. Protocols in DNA nanotechnology. Vol. 749, chapter 22. Humana Press; 2011. p. 317-358.
43. Hyre DE, Le Trong I, Merritt EA, Eccleston JF, Green NM, Stenkamp RE, et al. Prot Sci. 2006; 15:459–467.
44. Le Trong I, Humbert N, Ward T, Stenkamp R. J Mol Biol. 2006; 356:738–745. [PubMed: 16384581]
45. Aksimentiev A, Heng JB, Timp G, Schulten K. Biophys J. 2004; 87:2086–2097. [PubMed: 15345583]
46. Smeets RMM, Keyser UF, Krapf D, Wu MY, Dekker NH, Dekker C. Nano Lett. 2006; 6:89–95. [PubMed: 16402793]
47. Kramers H. Physica. 1940; 7:284–304.
48. Evans E, Ritchie K. Biophys J. 1997; 72:1541–1555. [PubMed: 9083660]
49. Dudko O, Hummer G, Szabo A. Phys Rev Lett. 2006; 96:108101. [PubMed: 16605793]
50. Bell GI. Science. 1978; 200:618–627. [PubMed: 347575]
51. Evans E, Ritchie K. Biophys J. 1997; 72:1541–1555. [PubMed: 9083660]
52. Comer J, Dimitrov V, Zhao Q, Timp G, Aksimentiev A. Biophys J. 2009; 96:593–608. [PubMed: 19167307]
53. Mirsaidov U, Comer J, Dimitrov V, Aksimentiev A, Timp G. Nanotechnology. 2010; 21:395501. [PubMed: 20808032]
54. Luan B, Afzali A, Harrer S, Peng H, Waggoner P, Polonsky S, et al. J Phys Chem B. 2010; 114:17172–17176. [PubMed: 21128651]
55. Wanunu M, Sutin J, B BM, Chow A, Meller A. Biophys J. 2008; 95:4716–4725. [PubMed: 18708467]
56. Aksimentiev A, Schulten K. Biophys J. 2005; 88:3745–3761. [PubMed: 15764651]

57. Chang H, Kosari F, Andreadakis G, Alam MA, Vasmatazis G, Bashir R. *Nano Lett.* 2004; 4:1551–1556.
58. Aksimentiev A. *Nanoscale.* 2010; 2:468–483. [PubMed: 20644747]
59. Heng JB, Aksimentiev A, Ho C, Marks P, Grinkova YV, Sligar S, et al. *Biophys J.* 2006; 90:1098–1106. [PubMed: 16284270]
60. Finkelstein IJ, Visnapuu ML, Greene EC. *Nature.* 2010; 468:983–987. [PubMed: 21107319]
61. Miyamoto S, Kollman PA. *Proteins: Struct, Func, Gen.* 1993; 16:226–245.
62. Izrailev S, Stepaniants S, Balsera M, Oono Y, Schulten K. *Biophys J.* 1997; 72:1568–1581. [PubMed: 9083662]
63. Grubmüller H, Heymann B, Tavan P. *Science.* 1996; 271:997–999. [PubMed: 8584939]
64. Tabard-Cossa V, Wiggin M, Trivedi D, Jetha NN, Dwyer JR, Marziali A. *ACS Nano.* 2009; 3:3009–3014. [PubMed: 19751064]
65. Wiggin M, Tropini C, Tabard-Cossa V, Jetha N, Marziali A. *Biophys J.* 2008; 95:5317–5323. [PubMed: 18775965]
66. Tropini C, Marziali A. *Biophys J.* 2007; 95:1632–1637. [PubMed: 17158571]
67. Luan B, Aksimentiev A. *Phys Rev E.* 2008; 78:021912.
68. Luan B, Aksimentiev A. *J Phys: Condens Matter.* 2010; 22:454123. [PubMed: 21339610]
69. Luan B, Aksimentiev A. *Phys Rev Lett.* 2008; 101:118101. [PubMed: 18851334]
70. Clausen-Schaumann H, Rief M, Tolksdorf C, Gaub HE. *Biophys J.* 2000; 78:1997–2007. [PubMed: 10733978]
71. Gracheva M, Aksimentiev A, Leburton JP. *Nanotechnology.* 2006; 17:3160–3165.
72. Lerou A, Destine J, Vanderheyden B, Gracheva M, Leburton JP. *IEEE TN.* 2010; 9:322–329.
73. Sigalov G, Comer J, Timp G, Aksimentiev A. *Nano Lett.* 2008; 8:56–63. [PubMed: 18069865]
74. Bustamante C, Bryant Z, Smith SB. *Nature.* 2003; 421:423–427. [PubMed: 12540915]
75. Cluzel P, Lebrun A, Heller C, Lavery R, Viovy JL, Chatenay D, et al. *Science.* 1996; 271:792–794. [PubMed: 8628993]

Figure 1.

Detection of a DNA-bound protein using a nanopore. (a) A protein–DNA complex is threaded through a nanopore wider than the complex. The protein is detected as a modulation of the nanopore ionic current. (b) A protein–DNA complex halts DNA transport when a protein encounters a nanopore smaller than the protein. The protein can be detected and, possibly, identified by measuring the nanopore ionic current, a lifetime of the protein–DNA complex near the nanopore, and signals associated with the local structure of the DNA in the pore constriction (see also Fig. 2).

Figure 2.

Idealized schematics of protein detection using nanopore force spectroscopy. (a) The transmembrane bias V_0 induces ionic current I_0 through an open pore. (b) A molecule of DNA is captured by the pore opening and the DNA begins to thread through the pore. The nanopore ionic current rises to I_1 due to the increased concentration of ions near the DNA. The potential on the membrane electrode drops due to the presence of the DNA. The DNA is expected to produce enhancements of the ionic current at physiological salt concentrations (~ 0.1 M KCl), and blockades at KCl concentrations exceeding 0.5 M [46]. (c) DNA is threaded through the pore. (d) A protein bound to the DNA reaches the pore and halts the translocation. Ionic current decreases to I_2 due to blockage of the pore by the protein. The device detects halting of translocation by the change in the current. (e) The bias is increased slowly beyond V_0 , producing an increasing stress on the protein–DNA bond. The strain on the DNA reduces its linear charge density and can be recognized from the change in the potential on the membrane electrode. The ionic current rises with the bias; however, the rise is not linear as explained in the text. (f) The bias reaches a sufficiently large value that the protein dissociates from the DNA. The value of the bias at which rupture occurs can be used to identify the bound protein. The ionic current rises because the protein no longer obstructs the pore opening. The relaxation of the DNA can be recognized in the potential on the membrane electrode. (g) The bias is reduced once more to V_0 to prepare for the arrival of the next protein. The ionic current returns to I_1 . Steps d–g are repeated until the DNA exits the pore. (i) Diagram of the bias applied transverse to the membrane by the Ag/AgCl electrodes immersed in the solution during each step in the process. (j) Diagram of the ionic current through the pore. (h) Diagram of the electric potential of the electrode embedded in the membrane. ★ The ionic current levels depend on the conditions of experiment (see text).

Figure 3.

MD simulation of protein–DNA rupture. The silicon nitride membrane is shown as a molecular surface, the *Bam*HI protein is drawn using vdW spheres; dsDNA is schematically shown using a ball-and-stick model. The snapshots illustrate the initial state of the system (a), the moment the protein reaches the surface of the membrane (b), rupture of the protein–DNA bond (c), and subsequent transport of dsDNA (d and e). Note that the protein does not dissociate from DNA. To the contrary, the DNA is observed to slide along the protein, rotating clockwise as it moves through the nanopore. This MD trajectory was obtained under a 2.0 V transmembrane bias.

Figure 4. Restriction enzyme systems used for MD simulations of protein detection. Shown are (a) *EcoRI* bound to its cognate sequence, (b) *BamHI* bound to its cognate sequence, and (c) *BamHI* bound to a sequence that differs from the cognate sequence by a single basepair. (d-f) Dependence of enzyme-DNA rupture on the bias and type of enzyme. The separation plotted is the distance between the center of mass of the glutamic acid (E113) in the enzyme's active site [34] and the phosphate of the third nucleotide in the cognate sequence (GAATTC for *EcoRI* and GGATCC for *BamHI*). Sharp increases in this distance indicate rupture of the enzyme-DNA complex. Two trajectories are shown for *BamHI* cognate at 2 V, panel e. Note that the scale of the time axis differs between the three plots.

Figure 5.

Time to rupture versus transmembrane bias for three enzyme–DNA systems: *Eco*RI bound to its cognate sequence (black circles), *Bam*HI bound to its cognate sequence (red squares), and *Bam*HI bound nonspecifically (blue diamonds). In the case of the cognate *Bam*HI system, two independent runs were performed at 2 V. At 1 V for this system, rupture was not observed within 70 ns. Lines are guides to the eye. Note that the rupture time is on a logarithmic scale.

Figure 6.

MD simulations of the effective force on the protein–DNA bond. (a–e) Five representative snapshots illustrating an MD trajectory of NeutrAvidin-bound AT homopolymer at a 400 mV bias. The NeutrAvidin protein is shown as vdW spheres, while the DNA strands carrying adenine and thymine nucleotides are shown in blue and orange respectively. (a) Base pairs near the top of the dsDNA start to disassociate at $t = 32$ ns. (b) Base pairing of the dsDNA at the constriction starts to break down at $t = 112$ ns. (c) Conformation of the dsDNA when the force on biotin reaches its maximum at $t = 130$ ns. (d) More base-pair disassociations at $t = 144$ ns. (e) Conformation of the dsDNA after reaching its steady state at $t = 174$ ns. (f) A plot showing the time dependence of the number of base pairs remaining above the constriction. See panel (e) for color legend. (g) A plot illustrating the effective force on biotin along the NeutrAvidin/streptavidin binding pocket versus time. The arrows in *f* and *g* correspond to the time when the snapshots (a–e) are taken.

Figure 7.

(a) Chemical structure of a biotin-linker-ssDNA fragment. (b) Tension versus extension plot for a linker connecting biotin to ssDNA. Each data point corresponds to an MD simulation of the system shown in the inset, where a constant force was applied to one atom of the linker while streptavidin was spatially restrained. The extension was measured between the two atoms shown in the inset as red and blue spheres. The inset schematically shows the systems used for calibration simulations. The streptavidin pocket is shown as a molecular surface, biotin, linker and the first nucleotide of a DNA strand are shown as bonds colored according to the atom type: nitrogen (blue), oxygen (red), hydrogen (white), cyan (carbon), phosphorous (gold). A fit to the data (solid line) was used to relate the extension of the linker in nanopore simulations to the effective force applied to the DNA-protein bond, Fig. 6f.

Figure 8.

Local structure of DNA in the pore constriction. (a-c) Three representative conformations of streptavidin/NeutrAvidin-bound DNA in a silicon nitride nanopore. The streptavidin/NeutrAvidin protein is shown as vdW spheres, while the A, C, T and G DNA bases are shown in the cartoon representation and colored in blue, yellow, orange and red respectively. (a) Typical conformation of B-DNA. Shown in the figure is a NeutrAvidin-bound AT homopolymer trapped in a parabolic 2.6×2.4 nm silicon nitride pore in its canonical form at 400 mV. (b) Typical conformation of stretched dsDNA conformation. Shown in the figure is a streptavidin-bound GC homopolymer trapped in a bi-conical 2.6×2.4 nm silicon nitride pore stretched by 60% at 1.0 V. (c) Typical conformation of stretched ssDNA. Shown in the figure is a NeutrAvidin-bound polyA ssDNA trapped in a parabolic 2.6×2.4 nm silicon nitride pore at 400 mV. (d) A plot showing the relationship between the ionic current blockage (I/I_0) and the stretching of DNA. Three groups of conformation can be categorized: B-DNA (orange circle), stretched dsDNA (red circle), and ssDNA (blue circle). Symbols: AT/A and GC/G homopolymers are shown solid and hollow symbol respectively. Circles represent streptavidin-bound dsDNA at 1.0 V, squares represent NeutrAvidin-bound dsDNA at 400 mV, while triangles represent NeutrAvidin-bound ssDNA at 400 mV.

Figure 9.

The electric potential on electrodes embedded in the pore is approximately proportional to the inverse length of the DNA segment within the pore. The black line shows the inverse of the average basepair length for the 10 basepairs nearest to the pore center (see the text for more details) during the simulation. The red circles show the electrode potential at the corresponding times, estimated as in Sigalov et al. [73] and scaled according to Eq. 4. The dashed blue line shows the approximate time at which the DNA and enzyme dissociated. (a) Simulation of *EcoRI* and cognate DNA at 4 V. (b) Simulation of *BamHI* and cognate DNA at 2 V.

Table 1

Ion current through a model nanopore for simulations with various system configurations and transmembrane biases.

sim.	enzyme	DNA	time (ns)	bias (V)	current (nA)	conductance (nS)
O-2V	none	none	23.2	2.01.52 ± 0.03		0.76 ± 0.01
O-4V	none	none	16.8	4.02.51 ± 0.03		0.63 ± 0.01
D-1V	none	fixed	18.1	1.01.40 ± 0.05		1.40 ± 0.05
D-2V	none	free	13.2	2.02.45 ± 0.06		1.23 ± 0.03
EC-2V	<i>EcoRI</i>	cognate	31.4	2.01.43 ± 0.03		0.71 ± 0.01
BC-1V	<i>BamHI</i>	cognate	59.4	1.00.87 ± 0.02		0.87 ± 0.02
BC-1.5V	<i>BamHI</i>	cognate	10.0	1.51.09 ± 0.05		0.72 ± 0.03
BC-2Va	<i>BamHI</i>	cognate	7.0	2.01.48 ± 0.06		0.74 ± 0.03
BC-2Vb	<i>BamHI</i>	cognate	12.0	2.01.30 ± 0.04		0.65 ± 0.02
BN-0.5V	<i>BamHI</i>	nonspec.	20.2	0.50.37 ± 0.04		0.75 ± 0.08
BN-1V	<i>BamHI</i>	nonspec.	29.2	1.00.85 ± 0.03		0.85 ± 0.03
BN-1.5V	<i>BamHI</i>	nonspec.	11.1	1.51.16 ± 0.05		0.78 ± 0.03

Simulations O-2V and O-4V contained only the model nanopore, water, and ions —no DNA or proteins were included. In Simulations D-1V and D-2V, an effectively infinite molecule of DNA was threaded through the pore. The remaining simulations contained a 62-basepair DNA molecule with a bound *EcoRI* or *BamHI* enzyme, as indicated. The portion of the simulations after the rupture of the enzyme–DNA complex were not included in the current calculations.