Enzyme-Modulated DNA Translocation through a Nanopore

Ajay S. Panwar and M. Muthukumar*

Department of Polymer Science and Engineering, University of Massachusetts, Amherst, Massachusetts 01003

Received May 22, 2009; E-mail: muthu@polysci.umass.edu

Abstract: We present a Langevin dynamics simulation study of enzyme-modulated translocation of a single-stranded DNA molecule through a cylindrical nanopore. The toroidal-shaped enzyme placed along the axis of the pore, threads a DNA molecule at a constant rate. As a result of this controlled release process, the length of DNA available for translocation varies with time. We examine the effect of time-dependent conformational entropy of the DNA on the translocation process. In addition, we also examine the effects of both the separation between the exonuclease and the pore, and the rate at which DNA is released by the enzyme. Our results indicate that the separation distance primarily influences the entry of the DNA into the pore. The length of the DNA released by the exonuclease that is most likely to enter the pore is nearly equal to separation distance between the pore and the exonuclease despite the flexibility of the polymer. However, the speed at which the DNA translocates through the nanopore is solely determined by the rate at which the exonuclease releases the DNA. We find that the translocation velocity is directly proportional to the rate of release.

I. Introduction

The phenomenon of DNA translocation through a nanopore involves forcing either a single-stranded (ss)[1-3] or double-stranded (ds)[4-7] DNA molecule through a transmembrane nanometer-sized pore. An externally applied electric field originating because of a potential drop across the pore imparts a driving force to the DNA and also causes an ionic current to flow through the pore.[8] When the DNA moves through the nanopore, it is accompanied by a reduction in ionic current. Consequently the nature of this blockade, involving both its duration and level, is thought to be a unique signature of the analyte passing through the pore. Hence, the phenomenon carries immense biotechnological potential promising ultrafast genome sequencing; the ultimate technological dream being a base-by-base read-out as DNA is fed into the nanopore.[8-10] A tremendous amount of scientific effort has been invested in understanding the mechanisms of polymer translocation[11-19] and the associated processes of charge transport[20-24] through a nanopore. However, experiments have been unable to resolve the composition of the translocating DNA to a single base level.[10,25,26]

The difficulty in detection arises because DNA translocation through a nanopore occurs too rapidly, approximately at the rate of a million bases per second.[10,27] This is beyond the resolution of present instrumentation in use. A possible strategy to enable single-base detection would be an attempt at slowing the translocation process down to a speed that enables detection by the instrumentation in use. To slow down a translocating DNA molecule, an additional force should act on the DNA during the translocation process which opposes the force due to the externally applied electric field. If the extent of this opposing force can be tuned, then one can, in principle, control the speed at which the DNA moves through the nanopore.

(1) Kasianowicz, J. J.; Brandin, E.; Branton, D.; Deamer, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13770–13773.
(2) Meller, A.; Nivon, L.; Branton, D. *Phys. Rev. Lett.* **2001**, *86*, 3435–3438.
(3) Fologea, D.; Gershow, M.; Ledden, B.; McNabb, D. S.; Golovchenko, J. K.; Li, J. L. *Nano Lett.* **2005**, *5*, 1905–1909.
(4) Li, J.; Gershow, M.; Stein, D.; Brandin, E.; Golovchenko, J. A. *Nat. Mater.* **2003**, *2*, 611–615.
(5) Chen, P.; Gu, J.; Brandin, E.; Kim, Y.-R.; Wang, Q.; Branton, D. *Nano Lett.* **2004**, *4*, 2293–2298.
(6) Storm, A. J.; Storm, C.; Chen, J.; Zandbergen, H.; Joanny, J.-F.; Dekker, C. *Nano Lett.* **2005**, *5*, 1193–1197.
(7) Storm, A. J.; Chen, J. H.; Zandbergen, H. W.; Dekker, C. *Phys. Rev. E* **2005**, *71*, 051903.
(8) Kasianowicz, J. J.; Henrickson, S. E.; Weetall, H. H.; Robertson, B. *Anal. Chem.* **2001**, *73*, 2268–2272.
(9) Meller, A. *J. Phys.: Condens. Matter* **2003**, *15*, R581–R607.
(10) Branton, D.; et al. *Nat. Biotechnol.* **2008**, *26*, 1146.
(11) Sung, W.; Park, P. J. *Phys. Rev. Lett.* **1996**, *77*, 783–786.
(12) Park, P. J.; Sung, W. *J. Chem. Phys.* **1999**, *111*, 5259.
(13) Muthukumar, M. *J. Chem. Phys.* **1999**, *111*, 10371–10374.
(14) Lubensky, D. K.; Nelson, D. R *Biophys. J.* **1999**, *77*, 1824–1838.
(15) Henrickson, S. E.; Misakian, M.; Robertson, B.; Kasianowicz, J. J. *Phys. Rev. Lett.* **2000**, *85*, 3057–3060.
(16) Ambjörnsson, T.; Apell, S. P.; Konkoli, Z.; Di Marzio, E. A.; Kasianowics, J. J. *J. Chem. Phys.* **2002**, *117*, 4063–4073.
(17) Meller, A.; Branton, D. *Electrophoresis* **2002**, *23*, 2583–2591.
(18) Muthukumar, M. *J. Chem. Phys.* **2003**, *118*, 5174–5184.
(19) Kantor, Y.; Kardar, M. *Phys. Rev. E* **2004**, *69*, 021806.
(20) Kong, C. Y.; Muthukumar, M. *Electrophoresis* **2002**, *23*, 2697–2703.
(21) Rabin, Y.; Tanaka, M. *Phys. Rev. Lett.* **2005**, *94*, 148103.
(22) Muthukumar, M.; Kong, C. Y. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 5273–5278.
(23) Bonthuis, D. J.; Zhang, J.; Hornblower, B.; Mathe, J.; Shklovskii, B.; Meller, A. *Phys. Rev. Lett.* **2006**, *97*, 128104-1–128104-4.
(24) Mohan, A.; Kolomeisky, A. B.; Pasquali, M. *J. Chem. Phys.* **2008**, *128*, 125104.
(25) Meller, A.; Nivon, L.; Brandin, E.; Golovchenko, J.; Branton, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 1079–1084.
(26) Howorka, S.; Cheley, S.; Bayley, H. *Nat. Biotechnol.* **1999**, *19*, 636–639.
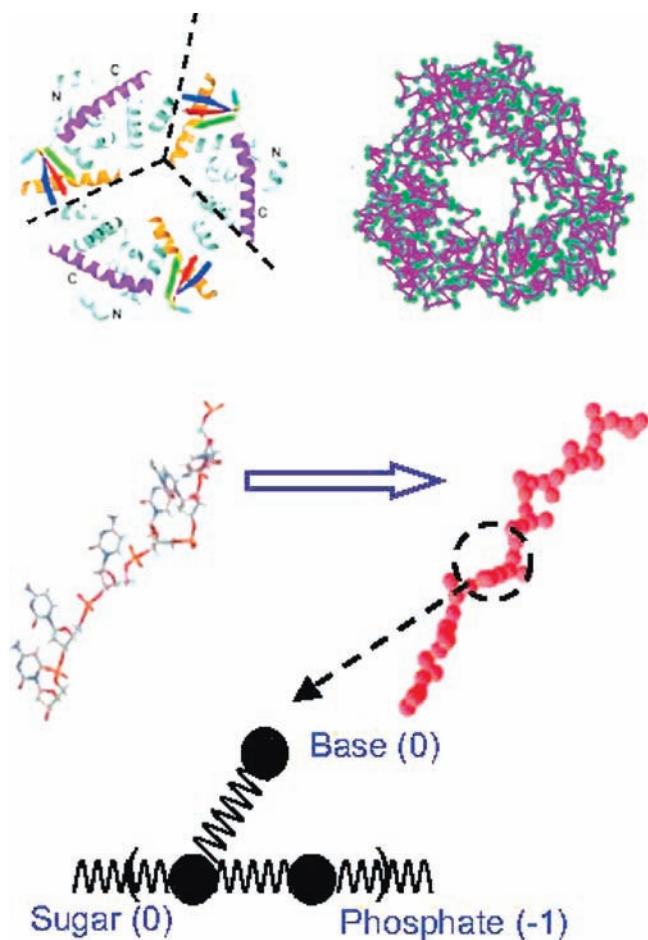(27) Chan, E. Y. *Mutat. Res.* **2005**, *573*, 13–40.

Figure 1. Coarse-grained united atom descriptions of λ-exonuclease (top) which has a toroidal structure made up of three protomers, and a ss-DNA molecule (bottom). For the case of the enzyme, each residue is replaced by a united atom, whereas in the case of the ss-DNA, the sugar, base, and phosphate moieties are replaced each by a united atom.

A potential candidate for such a "molecular braking mechanism" would be the λ-exonuclease enzyme.[10] λ-Exonuclease has a toroidal structure[28,29] and participates in DNA repair, replication, and recombination (see Figure 1). This enzyme binds to the free-end of a ds-DNA molecule and in the presence of Mg^{2+} ions digests one of the two strands in the 5' to 3' direction, releasing 5' mononucleotides, while keeping the other strand intact. Despite extensive investigations during the past four decades,[30-39] the catalytic mechanism of this enzyme is not fully understood. Depending on the particular methodology used

in following the kinetics of the hydrolysis, whether ensemble assays or single-molecule techniques, which in turn depends on the level of resolution and buffer conditions, the average digestion rates of 12 nt/s (refs 30 and 38), 15–20 nt/s (refs 33 and 36), and 32 nt/s (ref 37) have been reported. A much larger average rate of about 1000 nt/s (ref 35) has also been reported. These measurements also show that the instantaneous rate fluctuates about the average value and that in some cases the instantaneous rate of hydrolysis is dictated by the melting of the base pair at the catalytic site. Nevertheless, in spite of the elaborate rich local forces acting at the catalytic site, an average rate of hydrolysis for strings of thousands of bases has been reported in these studies.[30-39] Thus, in the present context of providing a molecular brake for a polymer undergoing translocation inside a nanopore, the hydrolysis process by the enzyme can be taken as the exonuclease feeding a single strand of DNA at an average rate, which is then forced into the nanopore due to the applied electric field. However, now the translocation speed is limited by the rate at which the exonuclease processes the ds-DNA (which defines the processivity of the enzyme). In a recent study,[40] Hornblower et al. demonstrated that the speed of DNA translocation can be retarded by binding ssDNA to Exonuclease I enzyme. Exonuclease I has a scissorlike structure, participates in DNA repair and recombination, and processively digests a ss-DNA in the 3'–5' direction in the presence of Mg^{2+} ions.[41] The idea of using exonuclease-assisted DNA sequencing was first proposed[42] by Jett et al. This involved a DNA that is attached to a fixed polymer bead at one end, whereas the other end is processively digested by an exonuclease. Once a nucleotide is cleaved by the enzyme, it is carried by a flow to be fluoroscently detected downstream. In the context of a nanopore, the translocation process enables detection via an ionic current change, whereas the exonuclease activity is proposed to control the rate at which detection takes place. Recent studies[43-45] have also examined the binding of DNA to the enzyme, DNA polymerase I, by forcing DNA–polymerase complexes into the α-hemolysin nanopore.

During the process of driven polymer translocation through a nanopore, the loss of conformational entropy upon entry into the nanopore leads to a free energy barrier associated with pore entry.[15-18] Thus, the applied potential gradient which attempts to push the DNA into the nanopore has to work against the entropic advantage enjoyed by the chain outside the pore. It is understood that the chain would make several attempts at entering the pore before a critical length of the polymer is placed inside it. When this happens, the polymer overcomes the largely entropic barrier, and from that point onward is carried downhill and through the nanopore by the applied electric field. In this case, the conformational entropy of the chain outside the pore is known and can be computed. However, when the DNA is complexed with the exonuclease, the conformational entropy

(28) Kovall, R. A.; Matthews, B. W. *Science* **1997**, *277*, 1824–1827.
(29) Kovall, R. A.; Matthews, B. W. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 7893–7897.
(30) Radding, C. M. *J. Mol. Biol.* **1966**, *18*, 235–250.
(31) Little, J. W. *J. Biol. Chem.* **1967**, *242*, 679–686.
(32) Carter, D. M.; Radding, C. M. *J. Biol. Chem.* **1971**, *246*, 2502–2512.
(33) Dapprich, J. *Cytometry* **1999**, *36*, 163–168.
(34) Mitsis, P. G.; Kwagh, J. G. *Nucleic Acids Res.* **1999**, *27*, 3057–3063.
(35) Matsuura, S.; Komatsu, J.; Hirano, K.; Yasuda, H.; Takashima, K.; Katsura, S.; Mizuno, A. *Nucleic Acids Res.* **2001**, *29*, e79.
(36) Subramanian, K.; Rutvisuttinunt, W.; Scott, W.; Myers, R. S. *Nucleic Acids Res.* **2003**, *31*, 1585–1596.
(37) van Oijen, A. M.; Blainey, P. C.; Crampton, D. J.; Richardson, C. C.; Ellenberger, T.; Xie, X. S. *Science* **2003**, *301*, 1235–1238.
(38) Perkins, T. T.; Dalal, R. V.; Mitsis, P. G.; Block, S. M. *Science* **2003**, *301*, 1914–1918.
(39) Conroy, R. S.; Koretsky, A. P.; Moreland, J. Eur. Biophys. J., published online June 27, 2009, http://dx.doi.org/10.1007/s00249-009-0502-5.

(40) Hornblower, B.; Coombs, A.; Whitaker, R. D.; Kolomeisky, A.; Picone, S. J.; Meller, A.; Akeson, M. *Nat. Methods* **2007**, *4*, 315–317.
(41) Breyer, W. A.; Matthews, B. W. *Nat. Struct. Biol.* **2000**, *7*, 1125–1128.
(42) Jett, J. H.; Keller, R. A.; Martin, J. C.; Marrone, B. L.; Moyzis, R. K.; Ratliff, R. L.; Seitzinger, N. K.; Shera, E. B.; Stewart, C. C. *J. Biomol. Struct. Dyn.* **1989**, *7*, 301–309.
(43) Benner, S.; Chen, R. J. A.; Wilson, N. A.; Abu-Shumays, R.; Hurt, N.; Lieberman, Dunbar, W. B.; K, R; Akeson, M. *Nat. Nanotechnol.* **2007**, *2*, 718–724.
(44) Cockroft, S. L.; Chu, J.; Amorin, M.; Ghadiri, M. R. *J. Am. Chem. Soc.* **2008**, *130*, 818–820.
(45) Hurt, N.; Wang, H.; Akeson, M.; Lieberman, K. R. *J. Am. Chem. Soc.* **2009**, *131*, 3772–3778.

of the part of the chain attempting to enter the nanopore, varies in a time-dependent manner. The manner in which the number of conformations sampled by the DNA changes with time is determined by the processivity of the enzyme-DNA complex. Since the entropic resistance to entry varies with time, it is not immediately clear how the competition between the entropic penalty and electric driving force influences the translocation process. Thus, in addition to providing insight into developing new strategies for sequencing DNA through nanopores, studying modulated polymer translocation examines previously unexplored aspects of time-varying conformational entropy.

While introducing some sort of "molecular braking mechanism" is one alternative, we recognize that improvements in instrument resolution and methods of data analysis would also constitute efforts in the right direction. In this work, we choose to focus on modulating the speed of translocation because it provides insight both into polymer dynamics in nanoporous media and single-molecule manipulation techniques. An investigation motivated by these two factors carries relevance to understanding macromolecular dynamical processes at the physiological level and designing biomimetic processes.

We may intuitively expect a nonmonotonic time-dependence of the probability, P_c, of capture of the DNA free end by the pore. For short times, given a fixed distance, L, between the location of the enzyme generating the chain and the pore mouth, the DNA length is so short that P_c is zero. In the other limit of very long times, the DNA chain generated is so long that it is entropically unfavorable to find its end at the pore mouth. Hence, there is an intermediate time at which the chain entry into the pore is maximized. Furthermore, the chain conformation might not be adequately equilibrated for faster generation of the chain. Thus, the chain entropy and the entropic barrier for translocation become effectively time-dependent. The time at which P_c is a maximum depends on L and the length, N, of the DNA that is generated, which in turn is controlled by the processing speed of the enzyme. The basic questions associated with this process are the following: (a) Is there an optimum distance L between the enzyme and the pore? (b) What is the most probable length, N^*, of the DNA that needs to be generated for successful capture, given a fixed L? How does N scale with L? Does the DNA capture process follow Gaussian chain statistics, where N^* is expected to scale as L^2? (c) What is the relation between the speed of translocation and the release rate of DNA by the enzyme ?

We use Langevin dynamics simulations to investigate these issues because of their inherent capacity to study nonequilibrium phenomena. In addition, using coarse-grained simulations we can resolve detail at the length scale of a DNA nucleotide ($\approx$ 0.5 nm), and also span longer time scales which are relevant to the translocation process (>1 μs). In the next section, we will set up the problem and list the details of both the model and the simulation technique. This is followed by a description of results in the following section. We find that the configuration of the DNA which is most likely to enter the nanopore corresponds to a stretched segment that is nearly the same length as the separation distance. The value of this most probable entry length, surprisingly, depends only on the distance between the pore and the enzyme, and is found to be independent of the rate at which the exonuclease releases the DNA. However, the speed of translocation is found to depend only on the rate at which the DNA is released by the exonuclease, and not the separation distance. In conclusion, we summarize our investigation in the final section of the paper.
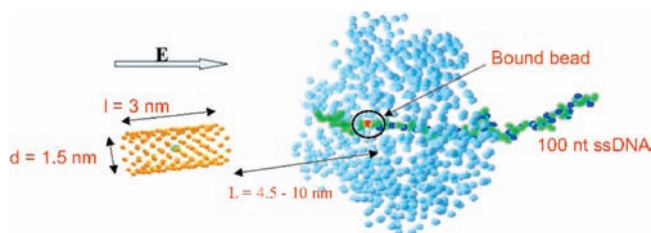

Figure 2. Schematic of enzyme-modulated translocation process with the exonuclease placed along the axis of the nanopore. The pore-exonuclease separation, L, is varied, as is also the binding time, τ_{bind}.

II. Methods: Simulation Algorithm and Parameters

Figure 2 shows a schematic of the enzyme-DNA complex lying along the axis of the nanopore. The length L represents the separation between the complex and the pore. Specifically, it is the distance between the entrance of the nanopore and the center-of-mass of the exonuclease, which can be assumed to be the location of the DNA binding site. The nanopore lies across an impenetrable membrane, and is characterized by a length, l, and a diameter, d. The translocating polymer is a 100 nucleotide (nt) ss-DNA molecule. In reality, a ds-DNA binds to the enzyme on one side, and as one of its strands is hydrolyzed, a ss-DNA strand exits on the other side of the enzyme. However, we model this process as the threading of a ss-DNA through the exonuclease cavity at a constant rate which in turn depends on the experimental conditions.[30−39] We define the binding time, τ_{bind}, as the time taken by the exonuclease to digest one base. An externally applied electric field, E, arising due to an applied potential gradient, drives the translocation process.

The ss-DNA is modeled[20,22] as a side-chain incorporated united atom model, where a collection of three sites represent a nucleotide (see Figure 1). Two of these beads lie along the chain backbone and represent the sugar and phosphate moieties, respectively, whereas the bead on the side chain represents the base. The bond angle of the polymer backbone is fixed at a value of 120°. This bond angle yields a persistence length of approximately 7 Å, which corresponds to experimentally estimated persistence length for ssDNA under conditions of high salt concentrations. However, the bond angle between a backbone bond and the side-chain is fixed at 60° to model the 3′ or 5′ end of the ssDNA. All beads in the polymer model are ascribed an excluded volume diameter and are connected to each other by harmonic springs. Both the sugar and base units are assumed to be charge neutral, whereas the phosphate unit is ascribed a charge of $-1e$. The charges in our system interact electrostatically via screened Coulombic interactions, which implicitly account for the screening effect due to the presence of salt ions. In most nanopore translocation experiments, as well as under physiological conditions, salt concentrations are usually on the higher side, and electrostatic interactions are heavily screened.

The motion of every bead constituting the ss-DNA model is described by a Langevin equation, since every bead represents a Brownian particle in Langevin dynamics simulations. The Langevin equation, which accounts for the different forces acting on a bead, is given by

$$m\frac{d^2\mathbf{r}_i}{dt^2} = \zeta\mathbf{v}_i - \nabla_i U + \mathbf{F}_i(t) \tag{1}$$

where m, $\mathbf{r}_i$, $\mathbf{v}_i$, and ζ define the mass, position, velocity, and friction coefficient for the ith bead, respectively. Here, $\mathbf{F}_i$ is the average random force due to collisions from the solvent molecules. The form of this force is given by the fluctuation−dissipation theorem

$$\mathbf{F}_i = \left(\frac{6\zeta k_B T}{\Delta t}\right)^{1/2}\mathbf{n}_i \tag{2}$$

where $\mathbf{n}_i$ is a random vector chosen such that it is uniformly distributed in the interval $[-1,1]$, and Δt is the size of the time

step in the simulation. The net interaction potential between different beads in the system is given by U. This potential is composed of the following terms,

$$U = U_{bond} + U_\theta + U_{LJ} + U_{DH} + U_V \qquad (3)$$

where the different contributions correspond to potentials due to bond stretching, bending, excluded volume interactions, screened Coulombic interactions, and the applied external potential, respectively. The bond stretching is defined by a harmonic potential,

$$U_{bond} = K_{bond}(l - l_0)^2 \qquad (4)$$

characterized by a spring constant, $K_{bond} = 300$ KJ/(mol-Å^2), and an equilibrium bond length, $l_0 = 2.5$ Å. The extension of the bond is defined by $(l - l_0)$ where l is the instantaneous bond length. Similarly, the bending potential between two adjacent bonds is also given by a harmonic potential,

$$U_\theta = K_\theta(\cos\theta - \cos\theta_0)^2 \qquad (5)$$

where $K_\theta = 312$ KJ/mol and θ and θ_0 define the instantaneous and the equilibrium bond angles, respectively. The values of the equilibrium bond angles used in this study have been defined in the preceding text. To account for stearic effects, we model excluded volume effects via a Lennard-Jones potential given by

$$U_{LJ} = \begin{cases} \varepsilon_{LJ}\left[\left(\dfrac{\sigma}{r}\right)^{12} - 2\left(\dfrac{\sigma}{r}\right)^6\right], & r \le \sigma \\ 0, & r > \sigma \end{cases} \qquad (6)$$

where r is the separation between any two beads, $\varepsilon_{LJ} = k_B T$, and $\sigma = 2.5$ Å are, respectively, the location and value of the minimum of the energy well-defined by the potential U_{LJ}. All electrostatic interactions in our system are defined by a screened Coulombic interaction,

$$U_{DH} = \frac{Z_i Z_j}{4\pi\varepsilon\varepsilon_0}\frac{\exp(-\kappa r)}{r} \qquad (7)$$

Here, Z_i and Z_j are the charges on the two beads i and j, respectively, ε is the relative permittivity, ε_0 is permittivity of free space, and κ is the inverse Debye length. Assuming an electrolyte corresponding to a 1 M KCl solution, the Debye length κ^{-1} is equal to 3 Å. The potential due to the externally applied voltage V_{ext}, that drives the DNA through the nanopore, is given by

$$U_v = Z_i V_{ext} \qquad (8)$$

Upon integration, eq 1 yields bead positions as functions of time. This in turn gives us polymer conformations as they evolve in time. The role of dielectric mismatch[22] between the membrane and the solution is not explicitly taken into account. Here, we use the constant electric field scenario. A velocity Verlet algorithm is used to integrate this equation.

λ-Exonuclease is a toroidal-shaped enzyme which is composed of three protomeric units. For the purpose of coarse-grained simulations, we model the exonuclease by a united atom representation. Here, every amino acid unit is represented by a united atom by substituting a finite diameter bead at the center-of-mass of the amino acid. In this study, we are concerned with effects of the processivity rate and separation of the exonuclease from the nanopore. As a result, the internal motion of the exonuclease need not be resolved. The only properties of the exonuclease which matter are its occupied physical volume, the charge distribution, and the binding interaction with the DNA. Consequently, we assign no connectivity between the beads representing the exonuclease. As mentioned in the previous paragraph, all electrostatic interactions, including those between the DNA and the enzyme, are treated via a screened Coulombic potential.

Previous investigations of the λ-exonuclease binding with ds-DNA have revealed four binding sites on each protomer of the enzyme.[28] In our model, beads corresponding to binding sites on the exonuclease interact with the bound nucleotide of the ss-DNA through a Lennard-Jones potential. Thus, in our coarse-grained description, the center-point of the exonuclease cavity is the approximate location of DNA-exonuclease binding. Then, the nonbonded potential attempts to localize the bound nucleotide at the center of the exonuclease cavity and is given by

$$U_{exo} = \begin{cases} \varepsilon_{exo}\left[\left(\dfrac{\sigma_{exo}}{r}\right)^{12} - 2\left(\dfrac{\sigma_{exo}}{r}\right)^6\right], & r \le r_c \\ 0, & r > r_c \end{cases} \qquad (9)$$

Here, $\varepsilon_{exo} = 20k_B T$, $\sigma_{exo} = 20$ Å, and $r_c = 40$ Å is the cutoff distance applied to this interaction. The value of σ_{exo} chosen in the simulation is approximately equal to the distance between the binding sites on the exonuclease and the center-point of the exonuclease cavity. The parameter that characterizes the exonuclease processivity is the binding time, τ_{bind}. With the passing of every τ_{bind}, the exonuclease-bound ssDNA translates a distance equivalent to the size of one nucleotide closer to the nanopore.[30−39] The size of every nucleotide, or equivalently, the distance between two adjacent phosphate beads on the ssDNA is 0.5 nm. Hence, after every τ_{bind}, the exonuclease feeds out 0.5 nm of the bound DNA strand.

The dimensions of the nanopore are chosen to approximately match the dimensions of the β-barrel of α-hemolysin nanopore. The diameter, d, of the pore is taken to be 1.5 nm, whereas the length, l, of the pore is fixed at 3 nm. We vary the separation length, L, between 4.5 and 10 nm. An electric potential of 75 mV, which drives the DNA, is applied across the length of the nanopore. In the cis (between the enzyme and the pore) and trans (pore exit) regions, the strength of the electric field is chosen to be a third of the strength inside the pore. The starting configuration for the ssDNA corresponds to the seventh nucleotide (or the seventh phosphate bead), from the DNA end (end closer to the nanopore), bound to the exonuclease. Thus, 3 nm of the DNA is assumed to be "processed" by the exonuclease at the start of the simulation. This particular configuration is chosen to maintain the initial stability of the ssDNA-exonuclease complex, and prevent backward slippage of DNA from the exonuclease during the simulation.

In the present first study of modeling polymer behavior in the tug between the pore and enzyme, we have taken only the most basic starting point as detailed above. There are several important questions appertaining to the activity of the enzyme which are beyond the scope of the present paper. In particular, the roles of DNA sequence, load, and the electric field on the hydrolysis rate are of interest. On the basis of the experimental results of refs 30−39, the hydrolysis site is localized. The fate of a phosphate ion in the hydrated environment of the catalysis cleft under an externally imposed electric field is beyond the scope of the coarse-grained model presented here, which is designed to explore the conformational entropy of the polymer chain in relation to the feeding by the enzyme and capturing by the pore. It is also known experimentally that the local rate of hydrolysis can vary depending on the particular base pair being melted (ref 37). In the present study, we model homopolymers and we have taken uniform hydrolysis rate.

III. Results and Discussion

Previous studies[15−18] have widely recognized that the process of DNA translocation is composed of two major steps; the first being the chain end finding the pore entrance, which is followed by the second step involving motion of DNA through the nanopore. During the first step, the chain faces an uphill task associated with overcoming an entropic barrier at the pore entrance. Once this barrier is overcome, the chain moves down
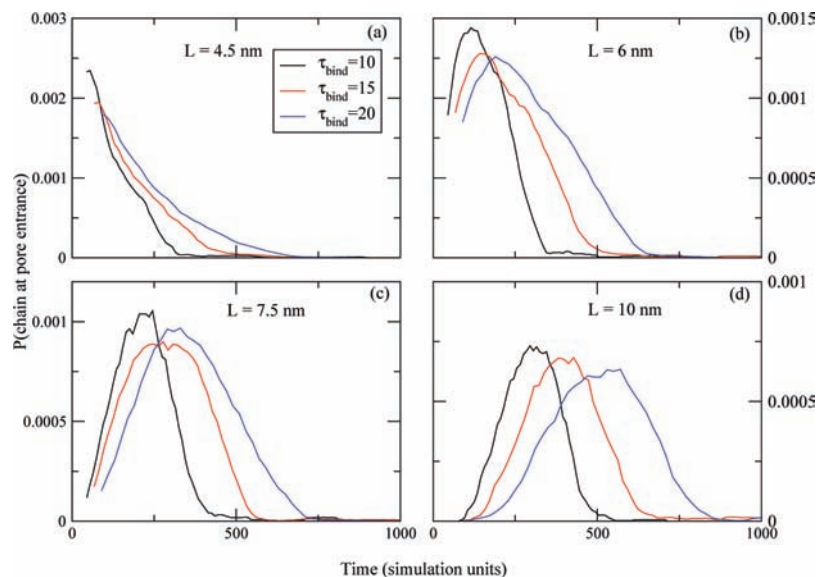
Figure 3. Plots showing the variation in time-dependent probability for the free end of the ss-DNA to enter the nanopore with the separation L and the binding time, τ_{bind}.

a free energy gradient. We break down the analysis of our simulation results along these lines. The length of DNA that is available for translocation varies with time, as a result of the processive binding with the exonuclease. Thus, the probability of the chain end finding the pore entrance varies with time in a way that is not immediately clear, but may depend on the processivity of the enzyme. The second step in translocation is characterized by a translocation velocity, which also varies with time, and should be influenced by the processivity of the exonuclease. In our study, we vary both the processivity (binding time, τ_{bind}) and the nanopore-enzyme separation, L. Corresponding to every data point (L and τ_{bind}), we simulated an ensemble of 500 ssDNA molecules over which the statistics were evaluated. Of these 500, not every realization led to a successful end-first translocation of the ssDNA through the nanopore. However, for every parameter set the size of the subensemble, corresponding to successful translocations, was easily over a 100.

Figure 3 panels a–d plot the probability of finding the chain end at the pore entrance as a function of time as both L and τ_{bind} are varied. The time-dependent probability distributions for different values of L display the same generic behavior. At time $t = 0$, there is no polymer segment in the vicinity of the pore entrance, as a result of which the probability is zero. As the enzyme releases the DNA, over time the length of the chain in the separation gap increases. With this, the probability of entry into the pore also increases until it reaches a maximum at a time which depends on both L and τ_{bind}. However, with further release of DNA, we find that the probability of finding the chain end at the pore entrance decreases with time. One should note that in 3a, the distributions indeed have a zero value at $t = 0$. This is not immediately clear from the plots which were obtained as running averages of binned distributions. For smaller values of L, it takes less time for the DNA end to traverse the separation gap before it attempts an entry into the nanopore. Consequently, the maximum in probability occurs at later times as the value of L increases. Intuitively, one would expect that the probability of entry into the nanopore would be higher when the enzyme-DNA complex is placed closer to the pore. Indeed, our simulations confirm this: the height of the maximum in probability decreases for larger values of L. For a given
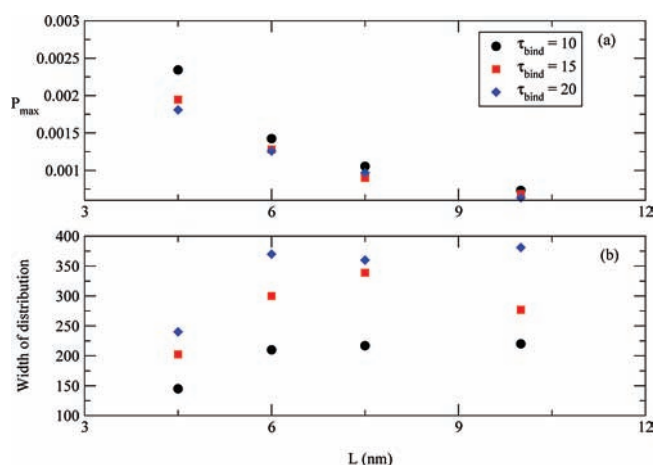

Figure 4. Plots showing the variations of (a) the most probable length at entry and (b) the width of distribution from 3 on the pore-enzyme separation, L.

separation L, the position of the maximum shifts to later times with increasing value of the binding time, τ_{bind}. In addition, the tail of the distribution also shifts to later times. This is a direct consequence of the slower processivity of the enzyme, which results in a delayed approach to the pore entrance. In Figure 4a, the variation of the maximum in probability is plotted as a function of L. We find that the value of the maximum decreases as the exonuclease is moved away from the nanopore. However, the height of the maximum is not significantly altered with a change in the binding time. Figure 4b plots the width of the distribution, defined by the full width at half-maximum, as a function of L and for different values of τ_{bind}. We find that the width of the distribution does not seem to change with L, but depends only on τ_{bind}. The only exceptions seem to be data sets corresponding to $L = 4.5$ nm. For any particular value of L, the width increases with increasing τ_{bind}, which is a consequence of the slower processivity of the enzyme. Thus, the plots in Figure 4, suggest that, while the height of the maximum is determined by the separation, L, the width of the distributions is determined by τ_{bind}.
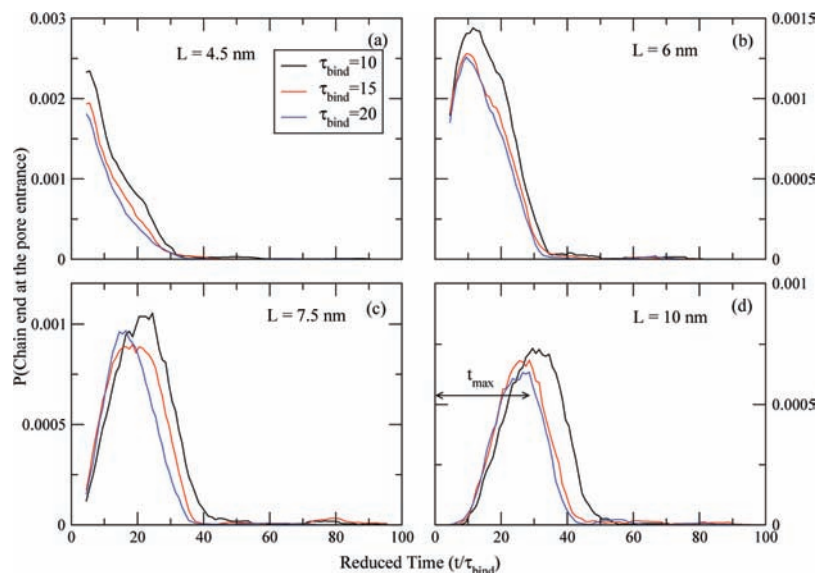
Figure 5. Plots showing the time-dependent probability for the free end of the ss-DNA to enter the nanopore plotted against a reduced time defined as t/τ_{bind}.

In Figure 5a−d, we plot the same probability distributions as in Figure 3, but as functions of a reduced time defined as, t/τ_{bind}. Here, we find that for a given value of L, the distributions for different values of τ_{bind} collapse asymptotically onto a single plot. We find that for a given value of L, both the location and height of the maximum do not change. By definition, the reduced time is associated with the processivity of the enzyme and is a measure of the number of releases that take place over time t. Thus, for a given value of L, the time-dependent probability of the chain end entering the nanopore depends on the length of the chain released by the exonuclease at that time. Given the setup of this modulated translocation process, one end of the part of the chain available for translocation is always held by the exonuclease. The length of the chain in the separation gap varies with time, while the free end seeks the pore opening. For short lengths of chain released, the free end is initially far from the pore opening to make an attempt at entry. As longer lengths of the chain are spilled, the probability of the end entering the pore would increase. However, one would not expect the probability to be very high for short lengths, since one end of the segment is always constrained. The effect of this constraint is a force exerted by the exonuclease on the DNA which prevents it from entering the nanopore. For very long segments released by the exonuclease, the role of the fixed end in preventing entry into the nanopore is not expected to be large. This is because DNA is expected to adopt a coil-like conformation in the separation gap, and thus the barrier to entry of the free end arises mainly from the entropic freedom the chain enjoys outside the pore. Hence, for longer segments, the number of conformations available for the free end increases with increasing length of the chain spilled by the enzyme. Consequently, it becomes harder for the free end to enter the pore for longer chain lengths. This is seen in decreasing probabilities in the tail of the probability distributions in Figure 5. At an optimal intermediate length of the polymer, the probability for chain entry is maximum. This would correspond to a segment length which is long enough to reach the pore entrance, yet at the same time is short enough not to suffer from significant configurational entropy which would prevent it from entering.

From Figure 5, we find that such an optimal chain length is observed for all four values of L. This optimal length increases
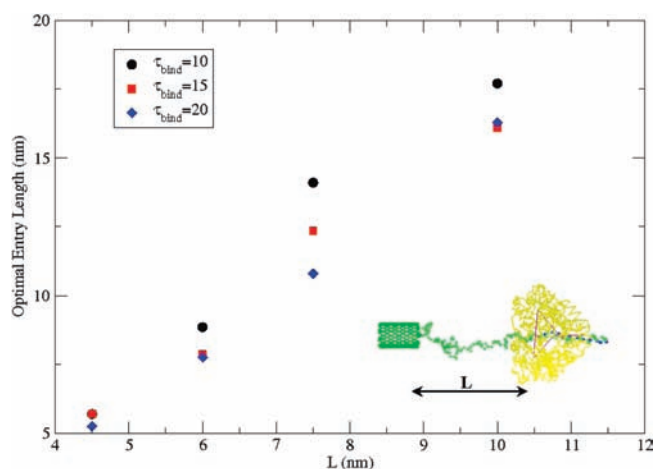

Figure 6. The most optimal length of the DNA at the time of entry into the nanopore displays a linear dependence on L. The inset shows a simulation snapshot for the case $L = 10$ nm and $\tau_{bind} = 20$, corresponding to a stretched conformation of the DNA just before it enters the nanopore.

with increasing value of L. Another interpretation of the maximum in probability distribution is that it corresponds to the length of the DNA segment that is most likely to enter the nanopore. Through the binding time, τ_{bind}, the length of the segment spilled by the enzyme can be related to the time elapsed during the process. Thus, the probability of the chain end entering the pore varies with time and reaches a maximum for an optimal length of DNA spilled out. The next obvious question examines the dependence of an optimal entry length on the two simulation parameters, L and τ_{bind}. We use Figure 5d, corresponding to $L = 10$ nm, to define a particular value of the reduced time, t_{max}, when the entry probability displays a maximum. For this case, $t_{max} \approx 30$, which corresponds to 30 releases over the time $\tau_{bind}t_{max}$. Then, the contour length of the segment released over this time corresponds to 18 nm (30 × 0.5 + 3 nm). This length is comparable to the separation gap, $L = 10$ nm. Figure 6 plots the dependence of the optimal entry length on L. We find that the optimal length increases linearly with L, and its value is close to the separation distance L. This
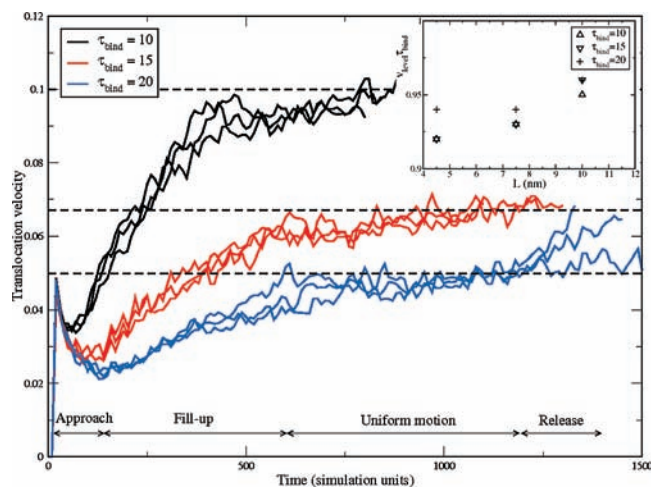
Figure 7. Plot showing the variation of the time-dependent translocation velocity with L and τ_{bind}. Data sets corresponding to the same value of L collapse on top of each other, implying that the velocity is solely determined by the binding time, τ_{bind}. The time evolution of the translocation velocity is composed of four regimes: (a) approach to the pore, (b) filling-up of the pore, (c) uniform motion through the pore, and (d) release of the DNA by the exonuclease. The dashed horizontal lines represent the plateau value of the release velocity, $1/\tau_{\text{bind}}$, for different values of τ_{bind}. In the inset we observe that the quantity $v_{\text{level}}\tau_{\text{bind}}$ is close to 1 for all values of L.

is in contrast to the quadratic dependence on L, expected from the Gaussian statistics of the chain. In addition, as anticipated the optimal entry length displays no dependence on τ_{bind}. The inset shows a simulation snapshot of the DNA just before it enters the nanopore for the case of $L = 10$ nm. The chain displays a stretched conformation and clearly suggests a contour length that is comparable to the gap separation, L. Hence, this result implies that the most probable length at the time of entry into the nanopore corresponds to one that is approximately equal to L and does not depend on τ_{bind}.

The second step in this modulated translocation process is the translocation of the chain through the nanopore after the end has gained entry. In Figure 7, we examine the effect of both L and τ_{bind} on the time dependence of the translocation velocity of the DNA. The translocation velocity of the DNA is defined as the velocity of the center-of-mass of the entire DNA chain, which would include DNA segments on either side of the exonuclease. We compute the center-of-mass velocity for the DNA for each Langevin dynamics realization that involves successful translocation of the DNA through the nanopore. Following this, an average value of the translocation velocity, corresponding to particular values of L and τ_{bind}, is obtained by ensemble averaging over all the successful realizations. The translocation velocity displays similar qualitative behavior for different values of L and τ_{bind}, though there are quantitative differences. At $t = 0$, the velocity is zero, but jumps to a finite value over a very short period of time. This is a start-up effect arising from the way the simulation is carried out, because we assume the DNA to be stationary for $t < 0$. From this finite value, the velocity first decreases to a minimum value, then rises steeply over a short period of time to attain an almost uniform value. At the end of the translocation process, a slight upturn in velocity is also observed. Interestingly, we find that the translocation velocity does not display any dependence on L. Curves corresponding to the same value of τ_{bind} but different values of L collapse together. However, there emerges a clear dependence of the velocity on τ_{bind}. As labeled in Figure 7, the time-dependent behavior of the translocation velocity can be

broken down into four regimes; (a) approach to the entrance of the pore, (b) filling up of the pore by the DNA, (c) uniform motion through the pore, and finally (d) release of DNA by the exonuclease. After the initial start-up from $t = 0$, the translocation velocity decreases as the chain end attempts to enter the pore. As the exonuclease releases the DNA, it begins to spill in the separation gap, but has to wait for the DNA end to enter the pore before moving ahead. This results in the velocity decreasing to a minimum, the minimum velocity being smaller for a larger value τ_{bind}. Once the DNA has entered the pore, it quickly fills up the pore, resulting in a rapidly increasing translocation velocity. As one would expect, the rate of increase of the velocity is higher for a smaller value of τ_{bind}. The velocity increases until it levels off to a uniform value. The center of mass moves at a uniform velocity for a duration that increases with increasing values of τ_{bind}. This period of uniform motion corresponds to a state when the nanopore is completely filled by a segment of the DNA, while the DNA is still bound to the exonuclease. If we define this uniform velocity as v_{level}, then the ratio of v_{level} corresponding to different values of τ_{bind}, is inversely proportional to the ratio of the respective binding times. This inverse dependence is seen in the inset of Figure 7, where the product $v_{\text{level}}\tau_{\text{bind}}$ is plotted as a function of L. The value of this product is nearly one for all values of L and τ_{bind}. The size of the symbols is the error bar for these data points. Thus, the value of v_{level} is completely determined by the binding time, τ_{bind}. However, the braking of polymer translocation does not occur until the chain end is caught inside the pore. Once the DNA is completely released by the exonuclease, the chain is free and the translocation velocity shows an increase.

Finally we note that the above results are for a given electric field strength. Although the chain conformation is expected to be affected only slightly for the electric field strengths of experimental relevance,[46] the electrophoretic capture rate will depend on it. This work is being addressed at present. In addition, the electric field is expected to influence the rate of hydrolysis of DNA at the catalytic site of the enzyme in a significant manner. However, this issue is beyond the scope of the present model.

IV. Conclusions

We have performed Langevin dynamics simulations of ss-DNA translocation through a nanopore, where the motion of the DNA is modulated by an enzyme bound to the DNA. We identified the enzyme-nanopore separation, L, and the binding time, τ_{bind}, as two parameters which should most significantly influence the speed with which the DNA moves through the nanopore. The analysis of the translocation process was divided into two parts.

The first deals with the entry of the free end of the DNA into the nanopore, and involves overcoming an entropic barrier at the entrance of the pore. Since the number of conformations that the segment of DNA, in the separation gap, explores changes with time, the resistance to the entry of the chain is also expected to change with time. This results in a probability of chain entry into the pore which rises from zero at time $t = 0$ and peaks for an extended conformation for the DNA segment released in the pore-enzyme gap. We derive two important conclusions from the analysis of the time dependent probability distributions for the entry of the DNA into the pore. The first

(46) Muthukumar, M. *J. Noncryst. Solids.* **1991**, *131−133*, 654–666.

result is that the maximum value of the time dependent probability of pore entry depends only on L, the separation between the pore and the exonuclease. It displays no dependence on the rate at which the exonuclease processes the DNA, the rate being characterized by the binding time, τ_{bind}. However, the width of the distribution is independent of L, and increases with increasing value of τ_{bind}. Obviously the time at which the maximum value occurs corresponds to the length of the DNA which is most likely to enter the nanopore. Another interesting result is that the segment length, corresponding to the most probable conformation at time of entry, varies linearly with the gap separation, L. Moreover, the value of the most probable length is approximately the same as the value of L. This implies that most of the conformations entering the nanopore are stretched and not coil-like in the region between the pore and the enzyme. Our results suggest that, given the experimental conditions corresponding to nanopore translocation experiments, the electric field values are strong enough to stretch the DNA before it enters the nanopore.

The second part of the analysis is concerned with the subsequent motion of the DNA through the nanopore after it has surmounted the free energy barrier. Again, we analyzed the effect of both L and τ_{bind} on the time dependent behavior of the translocation velocity. We found that the translocation velocity does not seem to change when L is varied but τ_{bind} is kept the same. The time-dependent behavior of the translocation velocity displays four regimes, corresponding to approach to the pore entrance, filling up of the pore, uniform motion through the pore, and finally the release of the DNA by the enzyme. The third important result of our study is that the translocation velocity depends only on τ_{bind}, and is independent of L. This assertion is further supported by the result that the product of ν_{level}, the velocity corresponding to uniform motion through the pore, and τ_{bind} has a value of nearly 1 for all values of L. Hence, an inverse relation between ν_{level} and τ_{bind} implies what one would intuitively expect, that the translocation velocity is faster for a smaller value of the binding time.

Our results elucidate the mechanisms underlying modulated transport of single polymer chains through nanoscopic pores and should serve as a guide to the design of better sequencing strategies. Specifically, our investigation sheds light on the roles of enzyme processivity and its separation distance on the translocation process. Only the distance of the enzyme from the pore, and not its processing speed, controls the entry of the DNA into the nanopore. However, modulation of the release rate does have an affect on the speed of translocation, which is found to have an inverse dependence on the binding time. The translocation velocity depends only on the processivity of the bound enzyme, which means that a tunable binding time would provide an effective control on the speed of DNA translocation through a nanopore.

Supporting Information Available: Complete ref 10. This material is available free of charge via the Internet at http://pubs.acs.org.

JA904047Q