

# REPORT DOCUMENTATION PAGE

AFRL-SR-AR-TR-05-

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, gathering data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-018 4302). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to provide information unless it is required by a notice of information collection that carries a valid OMB control number. PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

0376

<b>1. REPORT DATE (DD-MM-YYYY)</b> 08/09/2005		<b>2. REPORT TYPE</b> technical		January 2004-December 2004	
<b>4. TITLE AND SUBTITLE</b> Molecular modeling of stochastic sensing of bio-molecules				<b>5a. CONTRACT NUMBER</b>	
				<b>5b. GRANT NUMBER</b> FA 9550-04-1-0195	
				<b>5c. PROGRAM ELEMENT NUMBER</b>	
<b>6. AUTHOR(S)</b> Murugappan Muthukumar				<b>5d. PROJECT NUMBER</b>	
				<b>5e. TASK NUMBER</b>	
				<b>5f. WORK UNIT NUMBER</b>	
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b>  University of Massachusetts Polymer Science and Engineering Department, Amherst, MA 01003				<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>	
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b> Dr. Hugh DeLong AFOSR 4015 Wilson Blvd., Arlington, VA 22203-1954 <span style="float: right; font-size: 24pt;">NL</span>				<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>	
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b>  Approve for Public Release: Distribution Unlimited				<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>	
				<b>13. SUPPLEMENTARY NOTES</b>	
<b>14. ABSTRACT</b> Accurate modeling of stochastic sensing of protein analytes has been carried out using simultaneous computation of ionic current and analyte conformations, in the general platform of genetically-engineered alpha-hemolysin channels. This allows designing of optimally modified alpha-hemolysin channels for a diverse vector of analytes. Efforts were made to explore arrays of multiple pores, in terms of parallel versus series configurations. There is a strong potential to design a hybrid array of these configurations to improve the efficiency of signal integration without losing the discriminatory power. In addition, modeling was performed to successfully resolve several experimental puzzles in the electrophoretic capture of relatively large charged particles by decorated conical pores. The modeling fully complemented experimental efforts by other team members. The computer programs developed are forming the basis of ongoing efforts in the fabrication of devices under the current MOLDICE activity.					
<b>15. SUBJECT TERMS</b> DNA sequencing; detection of biohazards; single-polymer-detection					
<b>16. SECURITY CLASSIFICATION OF:</b>			<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b> M. Muthukumar
<b>a. REPORT</b>	<b>b. ABSTRACT</b>	<b>c. THIS PAGE</b>			<b>19b. TELEPHONE NUMBER (include area code)</b> (413) 577-1212

# Simulation of Polymer Translocation through Protein Channels

M. Muthukumar and C. Y. Kong

*Department of Polymer Science and Engineering*

*University of Massachusetts, Amherst, MA 01003*

## ABSTRACT

A modeling algorithm is presented to compute simultaneously polymer conformations and ionic current, as single polymer molecules undergo translocation through protein channels. The method is based on a combination of Langevin dynamics for coarse-grained models of polymers and the Poisson-Nernst-Planck formalism for ionic current. For the illustrative example of ssDNA passing through the  $\alpha$ -hemolysin, vivid details of conformational fluctuations of the polymer inside the vestibule and  $\beta$ -barrel compartments of the protein pore, and their consequent effects on the translocation time and extent of blocked ionic current are presented. In addition to shedding insights into several experimentally reported puzzles, our simulations offer experimental strategies to sequence polymers more efficiently.

# 1 Introduction

Translocation of polymers through biological channels is very complex involving many machineries, and is a fundamental step in many life processes. While several essential features of translocation are richly documented in systems such as mRNP complex through nuclear pores [1-3], only recently a simple system has been identified to follow the single-file passage of one isolated polymer through one channel [4-11]. In this system, the channel is constituted by self-assembling heptamers of the *Staphylococcus aureus*  $\alpha$ -hemolysin ( $\alpha$ HL) protein. The channel is assembled in a phospholipid bilayer, which offers a physical barrier, and the channel has an opening of diameter of about 1.4 nm at the narrowest constriction [12]. A single-stranded polynucleotide, such as poly(deoxyadenylate) and poly(deoxycytidylate), is pulled through the channel by an externally applied voltage gradient across the channel in a solution of a strong electrolyte. The idea is that the ionic current through the channel due to the passage of small ions of the electrolyte is blocked to certain extent during the event of translocation of the polymer. It has been hoped that the extent and duration of the current blockade are unique signatures of the identity of the polymer, both in terms of polymer's

chemical characteristics and physical length.

Even this simplest setup, where identical molecules undergo translocation, has generated several puzzling results. The distribution,  $P(\tau)$ , of the duration  $\tau$  of blockade of ionic current  $I_b$  is very broad and appears to exhibit at least two peaks. In addition, there are also several levels of ionic current blockade  $I_b$  for the same molecule. It is a standard practice in experimental investigations to combine the histograms of  $\tau$  and  $I_b$  [8]. The resultant scatter plots always yield two groups of data even for monodisperse homopolymers.

To gain insight into these puzzles, we have developed the following simulation. This is complementary to a flurry of theoretical activity[13-21], based on entropic barrier dynamics[22], all of which leading to a generic  $P(\tau)$  unlike in experiments. Although it is indeed desirable to perform the computation *ab initio*, the size of the system to be simulated is forbiddingly large to enable such a computation. Therefore we restrict ourselves to generate a minimal model of polymer translocation through  $\alpha$ HL channel, and our goal is to discover the most basic concepts relevant to a molecular understanding of the experimental results. Our minimal model incorporates enough chemical details of the polymer and channel to evaluate the potential roles played by secondary structures of the polymer, orientation of the chain (3' versus

5' end), binding sites inside the lumen of the channel, etc. In addition to resolving the above puzzles, the details emerging from our simulations are so vivid that new experimental protocols can be formulated to substantially narrow down  $P(\tau)$ , and thus to enable a much faster sequencing scheme.

The key result of our simulations is that conformational entropy of the polymer plays a dominant role in dictating  $\tau$  and  $P(\tau)$ . The breadth of  $P(\tau)$  can be directly attributed to the entropic trap arising from the vestibule of  $\alpha$ HL pore.  $P(\tau)$  can be made dramatically sharper, sufficient to facilitate sequencing of polymers, by deleting the vestibule or by dragging the polymer in a direction normal to the pore axis.

## 2 Simulation Method

The system consists of essentially four parts, namely, the heptameric  $\alpha$ -hemolysin pore, a ssDNA molecule undergoing translocation, a phospholipid membrane barrier carrying the pore, and an electrolyte solution permeating through the pore. The protein pore,  $\alpha$ HL, is represented by a united atom model where the residues of the pore are replaced by beads. The coordinates of the residues are extracted from the Protein Data Bank. There can be many

choices for the description of beads in terms of their sizes and shapes. In this paper we present results for the simplest choice of the beads. In the present united atom model, each residue is replaced by a spherical bead and the diameter of each bead is  $3\text{\AA}$ . A bead is assigned -1, 0, or +1 charge according to the ionization of the residue. All aspartic acid and glutamic acid residues, and C-terminal groups are deprotonated; arginine, lysine, histidine and N-terminal groups are protonated. The rest of the residues are electrically neutral. In view of experimental contexts [23], the protein pore is not allowed to move.

Further, the protein pore is treated as a rigid structure with a dielectric constant of 2 surrounded by a solution of dielectric constant 80. The protein pore is embedded in a  $48\text{\AA}$  thick membrane of dielectric constant 2. The membrane surfaces are hard walls. The protein pore is the only space to allow passage of the polymer from the donor compartment to the recipient compartment. The symmetric axis of the pore is positioned along the X-axis, as shown in Fig. (1a) which gives the geometry of the pore. The vestibule of the pore is submerged in the donor compartment. The  $\beta$ -barrel of the protein begins at  $x = 0$  and extends up to  $x = 48\text{\AA}$ .

We have used the side-chain incorporated model to represent a single-

stranded polynucleotide [20], due to its simplicity, fast computational speed and past success. Each nucleotide is made up of three subunits (phosphate, sugar, and base) and each subunit is taken as a bead. The bead representing the phosphate moiety carries one negative charge. To distinguish the chain directionality in terms of 3' versus 5' ends, we deliberately tilt the side-chain (base) at 65° angle as shown in Fig. (1c). The chain backbone is freely jointed in view of the fact that the persistence length of single-stranded DNA has been measured in high salt concentrations to be roughly the length of 1-2 nucleotides [24-26]. Although the three subunits are of different sizes and masses, the results presented below are for the model of all subunits having equal size and mass. The diameter of these beads is taken to be 2.5Å, and the equilibrium “bond length” between these “atoms” is 2.5Å. The mass of a nucleotide is distributed equally into the three subunits. For example, for the case of poly(dC), each subunit carries 96Da.

Polymer conformations are evolved by using the velocity Verlet algorithm [27] to follow the dynamics of the  $i$ th bead,

$$m \frac{d^2 \vec{r}_i}{dt^2} = -\zeta \vec{v}_i - \vec{\nabla}_{\vec{r}_i} U + \vec{F}_i(t) \quad (2.1)$$

where  $\vec{r}_i$ ,  $m$ , and  $\zeta$  are the position vector, mass, and friction coefficient, respectively, of the  $i$ th bead.  $\vec{F}_i(t)$  is the random force from the solvent bath acting on the  $i$ th bead and is stipulated to satisfy fluctuation-dissipation theorem,

$$\langle \vec{F}_i(t) \cdot \vec{F}_j(t') \rangle = \delta_{ij} 6k_B T \zeta \delta(t - t') \quad (2.2)$$

where  $t$  is the time and  $\delta$  is the usual delta function.  $U$  in Eq.(2.1) is the total potential energy acting on the  $i$ th bead, and consists of four contributions.

$$U = U_{bond} + U_{LJ} + U_{DH} + U_V. \quad (2.3)$$

These correspond, respectively, to bond stretch, short range, screened Coulomb, and electric potential given by

$$U_{bond} = k(l - l_0)^2 \quad (2.4)$$

$$U_{LJ} = \epsilon_{LJ} \left[ \left( \frac{\sigma}{r} \right)^{12} - 2 \left( \frac{\sigma}{r} \right)^6 \right] \quad (2.5)$$

$$U_{DH}(r_{ij}) = \frac{Z_i Z_j}{4\pi \epsilon_0 \epsilon r} \exp(-\kappa r) \quad (2.6)$$

$$U_V(\vec{r}_i) = Z_i V(\vec{r}_i). \quad (2.7)$$

Here,  $l$  is the bond length,  $l_0 (= 2.5\text{\AA})$  is the equilibrium bond length,  $r$  is the distance between interacting beads,  $Z_i$  is the electric charge on the  $i$ th bead,  $\epsilon_0$  is the permittivity of the vacuum, and  $V$  is electric potential. The values of the parameters are:  $k = 171 \text{ kcal/mol} \cdot \text{\AA}^2$ ,  $\epsilon_{LJ} = 0.5 \text{ kcal/mol}$ ,  $\sigma = 2.5\text{\AA}$ ,  $\kappa^{-1} = 3\text{\AA}$  for 1 M KCl. The dielectric constant  $\epsilon$  is inhomogeneous.  $\epsilon$  is 2 inside beads and membrane and is 80 everywhere else.

The ionic current  $I(t)$  at time  $t$ , due to the passage of electrolyte ions accompanying the polymer transport, is computed by using the Poisson-Nernst-Planck (PNP) formalism [28-30]. Taking advantage of the fact that small ions relax much faster than a large polyelectrolyte molecule and that the concentration of electrolyte in the pertinent experiments is very high in comparison with monomer concentration, we assume that at every time step of the Langevin dynamics simulation of the polymer, the electrolyte ions have relaxed to the steady state so that the polymer chain is taken only as a fixed charge distribution  $\rho_p(\mathbf{r}, t)$  at this time.

With this assumption, the self-consistent coupled PNP equations for the local charge density  $C_i(\mathbf{r}, t)$  of the  $i$ th ionic species and the electric potential  $V(\mathbf{r}, t)$  for a given polymer conformation at  $t$  are

$$\nabla \cdot \left( \nabla C_i(\mathbf{r}) + \frac{Z_i C_i}{k_B T} \nabla V(\mathbf{r}) \right) = 0. \quad (2.8)$$

$$\epsilon_0 \nabla \cdot [\epsilon(\mathbf{r}) \nabla V(\mathbf{r})] = -\rho_f(\mathbf{r}) - \rho_p(\mathbf{r}, t) - \sum_i C_i(\mathbf{r}). \quad (2.9)$$

Here  $\rho_f(\mathbf{r})$  is the local charge density arising from fixed charges in the system, such as the charged beads of the protein pore. As already mentioned,  $\epsilon(\mathbf{r})$  is set to 2 for the beads of polymer and pore, and for the membrane. It is 80 everywhere else. In addition, the externally applied voltage gradient is accounted for by fixing the electrostatic potential at the boundaries in the  $x$ -direction. The box for the ionic current calculations is located at  $-100 \text{ \AA} < X < 75 \text{ \AA}$ ,  $-50 \text{ \AA} < Y < 50 \text{ \AA}$ , and  $-50 \text{ \AA} < Z < 50 \text{ \AA}$ . We set  $V = 0$  at  $X = -100 \text{ \AA}$  and  $V = V_0$  at  $X = 75 \text{ \AA}$ .

The above equations (2.8) and (2.9) are solved by successive over-relaxation method with a grid spacing of  $1 \text{ \AA}$ , by following Ref.(29). First we start with  $V(X)$  due to the externally applied electrostatic potential gradient across the membrane of dielectric constant of 2, and uniform concentrations ( $C_+$  and  $C_-$ ) of cations and anions of the dissolved electrolyte.

Then, the coordinates of beads of the pore and polymer are read in. Next,

using the successive over-relaxation method,  $V, C_+$ , and  $C_-$  are computed iteratively until they converge within the tolerance levels of  $\Delta V = 10^{-8}V$ , and  $\Delta C_{\pm} = 10^{-7}M$ . In our simulations we have considered 1 M of KCl as the electrolyte. Since  $C_{\pm}$  is very high in comparison with the counterions from the polymer, the latter are ignored. From the convergent values of  $C_{\pm}$  and  $V$  at each time unit of the Langevin simulation, the ionic current at this time  $t$  is given by

$$I(t) = A(J_+ + J_-), \quad (2.10)$$

with

$$J_{\pm}(\mathbf{r}, t) = -D_{\pm} \left( \nabla C_{\pm}(\mathbf{r}, t) + \frac{Z_{\pm} C_{\pm}}{k_B T} \nabla V(\mathbf{r}, t) \right). \quad (2.11)$$

In evaluating the ionic current, we have taken  $J_{\pm}$  at  $X = 48\text{\AA}$  and  $A$  at this exit point is  $441\text{\AA}^2$ .  $D_+$  and  $D_-$  are taken to be  $1.96 \times 10^{-5} \text{ cm}^2\text{s}^{-1}$  and  $2.03 \times 10^{-5} \text{ cm}^2\text{s}^{-1}$ , respectively, for  $\text{K}^+$  and  $\text{Cl}^-$  ions. The magnitude of the applied voltage is  $V_0 = 120mV$ .

The simulation is carried out as follows. First, a single polymer chain is equilibrated in the absence of pore and externally applied voltage gradient, by considering only the connectivity, Lennard-Jones, and Debye-Hückel forces

among various beads of the chain. Separately the electrostatic potential for the protein pore is computed for 1 M KCl solution in the steady state with a given value of  $V_0(120mV)$ , and in the absence of the polymer. The potential along the central axis of the channel is given as dotted numerical data in Fig. (2). The different gradients (slopes of the curve) at different parts of the trajectory are readily conspicuous in this figure. Since we are interested here in the details of the actual translocation event, and not in the way the polymer arrives at the channel entrance, we have extended the range of the potential gradient inside the vestibule to the donor compartment. In addition, we have ignored the finer details of the potential gradient and replaced the actual potential gradient by the approximated gradient consisting of three parts, viz., up to the  $\beta$ -barrel, inside the  $\beta$ -barrel, and outside  $\beta$ -barrel in the recipient compartment.

Next, the equilibrated polymer is inserted in the donor compartment such that the center of mass of the chain is at  $X = -100\text{\AA}$  in front of the mouth of the pore. The polymer is dragged by the electrical potential gradient across the pore. For each time unit of the Langevin dynamics simulation, the coordinates of the polymer beads are stored and the ionic current at this time is computed. When any one monomer of the polymer enters the

mouth of the pore (at  $X = -50\text{\AA}$ ) a clock for measuring the translocation time starts. As the polymer progressively invades the lumen of the pore, the ionic current is calculated at each time unit. This is how the ionic current traces given below are constructed. It is to be noted that  $\nabla V$  entering the Langevin equation for the polymer bead is the original value of the protein pore in the absence of the polymer. When the last bead of the polymer exits the outer edge of the pore (at  $X = 48\text{\AA}$ ), the clock stops with a reading of the translocation time  $\tau$ . By repeating the above sequence of simulations thousands of times we construct the histogram of  $\tau$ , and for each of the entries of the histogram, we have a trace of ionic current as the polymer traverses through the pore.

### 3 Results

First, we give results on chain conformations, ionic current, and translocation time  $\tau$  as the polymer transits through the channel in typical simulations. Then we present distribution functions of blocked current and  $\tau$ , based on thousands of simulations. Fig. (3a) represents a typical simulation result for the model chain of ss-poly(dC) with  $N = 45$  nucleotides undergoing

translocation through  $\alpha$ HL under a potential difference of 120 mV in 1 M KCl solution. One unit of simulation time is arbitrarily taken to be 0.15  $\mu$ s, by matching the peak values of  $\tau$  in our simulations and the experiments of Refs. (8) and (9), and by noting that the average translocation velocity is constant in the long polymer and high voltage gradient limits [9].

There are several key features evident from the typical trajectory of Fig. (3a). (i) The open pore current of about 135 pA is very close to the experimental value. (ii) The polymer enters the mouth of the vestibule not necessarily as a single-file. The presence of polymer monomers at the mouth of the vestibule is sufficient to provoke a response in the ionic current by providing a spatial blockade to the flow of electrolyte ions. (iii) As time progresses the chain gets into the vestibule. Due to its large volume, the vestibule acts as an entropic trap and the chain segments linger in this cavity for some definite time duration. At this stage, the ionic current is  $I_{b1}$  ( $\sim$  100 pA). The fluctuations in  $I$  around its mean value reflect the dynamics of the polymer rattling inside the vestibule. (iv) One end of the chain eventually enters the  $\beta$ -barrel. At this time,  $I(t)$  is reduced precipitously to the blocked current level  $I_{b2}$ . The average value of  $I_{b2}$  is 17 pA. This very low value reflects the reduction in the cross-sectional area of  $\beta$ -barrel for flow of

electrolyte ions by the presence of polymer segments. The fluctuations in  $I_{b2}$  is solely due to the dynamics of the polymer. The ionic current is about  $I_{b2}$  until the last monomer exits the end of the  $\beta$ -barrel.  $\tau$  is the duration of all of the above events. (v) When the chain has completed its translocation, the ionic current returns to the open pore current after leaving a distinct signature for the presence of some monomers in the back of the exit location, perturbing the flow of electrolyte ions.

It is to be noted from Fig. (3a) that the chain spends significant amount of time inside the vestibule before entering the  $\beta$ -barrel. While such trajectories are quite common, there are also more prevalent trajectories where the dwell time inside the vestibule is short as illustrated in Fig. (3b). The essential features of Fig. (3b) are the same five features discussed for Fig. (3a), except that the time spent inside the vestibule is shorter. As seen from the chain conformations presented in Fig. (3b), the shorter duration inside the vestibule is due to the initial chain orientation in line with the axis of the channel, and the chain entering the vestibule mouth essentially as a single-file.

Vivid details such as those in Figs. (3a) and (3b) are available for each of the thousands of simulations that were performed, based on which we have

constructed histograms for  $\tau$  and  $I_b$ . The normalized probability  $P(\tau)$  based on 4000 simulations is plotted against  $\tau$  in Fig. (4) for  $N = 45$ ,  $V_0 = 120mV$ , and 1 M KCl. The error bars are estimated from the square root of counts per bin before normalizing the histogram. It is evident that the distribution of  $\tau$  is very broad and two dominant peaks at  $50 \mu s$  and  $75 \mu s$  may be identified. This result is in agreement with our earlier report of two peaks based on smaller number of simulations with shorter ( $N = 15$ ) chains [20]. Now we are in a position to go back to the trajectories of the chains and study them for various events taking different translocation times. Indeed, an example of the trajectory taking  $76 \mu s$  is Fig. (3a), where the vestibule plays the role of an entropic trap, delaying the translocation process. In contrast, by stochasticity associated with chain end orientation and thermal noise from the background, the chain end can enter the  $\beta$ -barrel sooner by essentially avoiding the vestibule's entropic trap. This is the case with Fig. (3b) where  $\tau (= 23.7 \mu s)$  is much shorter, and this event belongs to the dominant peak in the histogram. We validate this mechanism for the occurrence of two peaks, by performing further simulations on nanotubes where we have cut out the vestibule part of the channel (see below). We must mention at this point that there is no significant discrimination between 3' end and 5' end entering

the  $\beta$ -barrel first in terms of translocation time.

As already noted, there are two levels of blocked ionic current during the translocation event. For the event corresponding to Fig. (3a), the normalized histogram  $P(I_b)$  of blocked current  $I_b$  within the duration of  $\tau$ , given in Fig. (5a), exhibits two separate distributions. The average values of these two peaks represent whether the blockade corresponds to the vestibule (weak blockade) or the  $\beta$ -barrel (strong blockade). The widths of the peaks correspond to the accompanying polymer dynamics. It must be remarked that the occurrence of two populations of blocked ionic current has nothing to do with the actual value of  $\tau$ . The histogram of Fig. (5a) corresponds to  $\tau = 76\mu s$ . For the case of  $\tau = 23.7\mu s$  corresponding to Fig. (3b), the histogram of  $I_b$  is given in Fig. (5b). Here again, there are two distributions representing polymer dynamics inside the vestibule and the  $\beta$ -barrel. However the weight of  $I_b$  population relating to the vestibule is weaker for the faster translocation events. These results emphasize the need to interpret the scatter plots of Ref. [8] differently by having to go into individual events instead of clumping them all together.

The intervention by the vestibule of  $\alpha$ HL as an entropic trap for the translocation of the polymer leads to the breadth of  $P(\tau)$ . To further validate

this observation, we have repeated our simulations by replacing the  $\alpha$ HL channel by a cylindrical tube of diameter  $18 \text{ \AA}$  and length  $98 \text{ \AA}$ , as illustrated in Fig. (1b). The tube is inserted inside the membrane of dielectric constant 2, and the membrane thickness is the same as the tube length.

A typical trajectory of polymer translocation through the cylindrical tube is given in Fig. (6) for  $N = 45$ ,  $V_0 = 120mV$ , and 1 M KCl. This also exhibits all of the features described for Figs. (3a) and (3b), except the part corresponding to the role of the vestibule, where the weaker blockade is prominently absent. Based on 1000 simulations, the histogram  $P(\tau)$  is given in Fig. (7), where the results of Fig. (4) are included for comparison. It is clear that  $P(\tau)$  is very sharply peaked and there is only one peak for the nanotube. Our simulations thus suggest that it is preferable to have a nanotube or just the  $\beta$ -barrel for sequencing purposes instead of  $\alpha$ HL.

## 4 Discussion and conclusions

Our simulations using coarse-grained models of the polymer and the channel with appropriate accounting of dielectric heterogeneity reproduce all of the essential features observed experimentally with  $\alpha$ HL. The vestibule is found

to play a significant role by providing an entropic trap and thereby slowing down the process. Another observation is that there is no direct correlation between translocation time and blocked current. For each value of  $\tau$ , there are two dominant values of blocked current.

As already mentioned, there is a perturbation in the ionic current through the channel, as soon as a monomer is eclipsing the pathway for small ions in front of the channel (see Figs. 3 and 6). This suggests a novel experimental setup for sequencing polymers using measurements of ionic current through channels. The strategy is illustrated in Fig. (8a), where the circle represents the mouth of the pore through which electrolyte ions pass and create the measured ionic current under an applied voltage gradient. The pore can be either  $\alpha$ HL or a nanotube with prescribed diameters. Then a polymer is dragged across the front of the pore in the direction of the arrow in Fig. (8a). The physical size of the monomer at the pore-front excludes small ions to pass by and consequently reduces the magnitude of the ionic current. When monomers with different sizes are dragged normal to the ion flow, different levels of ionic current are registered. When a polymer with eight dC units is dragged at the speed of  $1.0\text{\AA}/\mu s$  near the pore entrance, in the normal direction to the current flow, the time-dependence of ionic current is given

in Figs. (8b) and (8c). The saw-tooth nature of the trace is due to the periodicity arising from the monomer entering the aperture, then generating maximum eclipse, and then exiting the aperture. If an octomer of dA is simulated under identical conditions (where the diameter of the side-chain bead is  $5\text{\AA}$  instead of  $2.5$  for dC and the bond length connecting the side-chain to the backbone is  $3.75\text{\AA}$ ), the ionic trace is as given in Figs. (8b) and (8c). In these particular simulations, the pore diameter and length are taken as  $14\text{\AA}$  and  $50\text{\AA}$ , respectively, to detect the identity of only one monomer. Depending on the size of the aperture, signals corresponding to dimers, trimers, etc., can be generated. In the present illustration of the concept, we now drag octomers with different sequences containing dA and dC. Depending on which monomer sequence is being dragged, the ionic current will toggle back and forth between the values corresponding to pure dA and dC values. There is only one unique ionic current trace for a given sequence, as illustrated by two distinctly different traces for two sequences. Although the difference in the ionic current from our simulations are rather small, the current experimental status [31] is able to discern such small difference. The strategy emerging from our simulations will enable the precise sequencing of polymers by several orders of magnitude faster than the currently available techniques.

In summary, our combination of Langevin dynamics simulations of polymer dynamics with the PNP formalism for ionic current, in the general premise of united atom description, offers a computational platform to discover generic physical principles behind polymer translocation through biological channels.

## Acknowledgment

Acknowledgment is made to NIH Grant No. 1R01HG002776-01, NSF Grant No. DMR-0209256, and to the donors of The American Chemical Society Petroleum Research Fund for partial support of this research.

## References

- [1] Macara, I.G., *Microbiology and Molecular Biology Reviews*, 2001 *65*, 570-594.
- [2] Dreyfuss, G., Kim, V.N., Kataoka, N., *Nature Reviews Molecular Cell biology*, 2002 *3*, 195-205.
- [3] Cullen, B.R., *Journal of Cell Science*, 2003 *116*, 587-597.
- [4] Kasianowicz, J. J., Brandin, E., Branton, D., Deamer, D. W., *Proc. Natl. Acad. Sci. USA*. 1996, *93*, 13770-13773.
- [5] Bezrukov, S. M., Vodyanoy, I., Brutyan, R. A., Kasianowicz, J. J., *Macromolecules*, 1996, *29*, 8517-8522.
- [6] Akeson M., Branton, D., Kasianowicz, J. J., Brandin, E., Deamer, D. W., *Biophys. J.*, 1999, *77*, 3227-3233.

- [7] Henrickson, S. E., Misakian, M., Robertson, B., Kasianowicz, J. J.,  
Phys. Rev. Lett. 2000, *85*, 3057-3060.
- [8] Meller, A., Nivon, L., Brandin, E., Golovchenko, J., Branton, D., Proc.  
Natl. Acad. Sci. USA. 2000, *97*, 1079-1084.
- [9] Meller, A., Nivon, L., Branton, D., Phys. Rev. Lett. 2001, *86*, 3435-  
3438.
- [10] Meller, A., Branton, D., Electrophoresis, 2002, *23*, 2583-2591.
- [11] Vercoutere, W. A., Winters-Hilt, S., DeGuzman, V. S., Deamer, D.,  
Ridino, S. E., Rodgers, J. T., Olsen, H. E., Marziali, A., Akeson, M.,  
Nucleic Acids Res. 2003, *31*, 1311-1318.
- [12] Song, L., Hobaugh, M. R., Shustak, C., Cheley, S., Bayley, H., Gouaux,  
J. E., Science, 1996, *274*, 1859-1866.
- [13] Sung, W., Park, P. J., Phys. Rev. Lett. 1996, *77*, 783-786.
- [14] Lubensky, D. K., Nelson, D. R., Biophys. J. 1999, *77*, 1824-1838.
- [15] Muthukumar, M., J. Chem. Phys., 1999, *111*, 10371-10374.
- [16] Muthukumar, M., Phys. Rev. Lett. 2001, *86*, 3188-3191.

- [17] Chern, S. S., Cardenas, A. E., Coalson, R. D., J. Chem. Phys. 2001, *115*, 7772-7782.
- [18] Muthukumar, M., Electrophoresis, 2002, *23*, 1417-1420.
- [19] Ambjornsson, T., Apell, S.P., Konkoli, Z., Di Marzio, E.A., Kasianowicz, J.J., J. Chem. Phys. 2002, *117*, 4063-4073.
- [20] Kong, C., Muthukumar, M., Electrophoresis, 2002, *23* 2697-2703.
- [21] Slonkina E., Kolomeisky A.B., J. Chem. Phys., 2003, *118*, 7112-7118.
- [22] Muthukumar M. and Baumgärtner A., Macromolecules, 1989, *22*, 1937-1941.
- [23] Korchev, Y. E., Bashford. C. L., Alder, G. M., Kasianowicz, J. J., Pasternak. C. A., J. Membr. Biol. 1995, *147*,233-239.
- [24] Achter, E. K., Felsenfeld, G., Biopolymers. 1971, *10*, 1625-1634.
- [25] Smith, S. B., Cui, Y., Bustamante, C., Science. 1996, *271*, 795-799.
- [26] Tinland, B., Pluen, A., Sturm, J., Weill, G., Macromolecules. 1997, *30*, 5763-5765.

- [27] Allen, M. P., Tildesley, D. J., Computer Simulation of Liquids. Oxford: Clarendon, 1987.
- [28] Eisenberg, R. S., J. Membr. Biol. 1996, *150*, 1-25.
- [29] Kurnikova M.G., Coalson R.D., Graf P., Nitzan A., Biophys. J. 1999, *76*, 642-656.
- [30] Corry, B., Kuyucak, S., Chung S. H., Biophys. J. 2000, *78*, 2364-2381.
- [31] Xie, H.Z., Braha, O., Gu, L.Q., Cheley, S., Bayley, H., Chemistry and Biology, 2005, *12* 109-120.

## 5 Figure Captions

Figure 1. (a) The end and side views of  $\alpha$ HL. (b) The internal wall of the nanotube is made of spherical neutral beads on a curved hexagonal lattice. The interaction between polymer bead and tube bead is taken as the Lennard-Jones potential of Eq. 2.5 with  $\epsilon = 0.2 \text{ kcal/mol}$  and  $\sigma = 2.75 \text{ \AA}$ . (c) United atom representation of ss-poly(dC).

Figure 2. PNP and approximated  $V(X)$  across the protein pore.

Figure 3. Typical ionic current traces: (a) long  $\tau$ , (b) short  $\tau$ .

Figure 4. Histogram of  $\tau$  for  $\alpha$ HL ( $N=45$ ,  $V_0=120 \text{ mV}$ ).

Figure 5. Histogram of  $I_b$ . (a) and (b) correspond, respectively, to Figs. 3(a) and 3(b).

Figure 6. Typical ionic current trace for a nanotube.

Figure 7. Comparison of  $P(\tau)$  between  $\alpha$ HL and nanotube.

Figure 8. (a) Proposed experimental strategy. The ionic current traces for (b) two homopolymers and (c) two heteropolymers (sequence given in the insert). Each polymer has 8 bases.

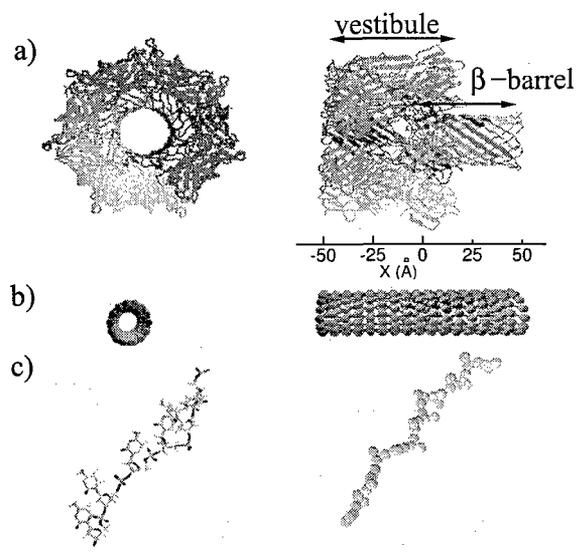


Figure 1:

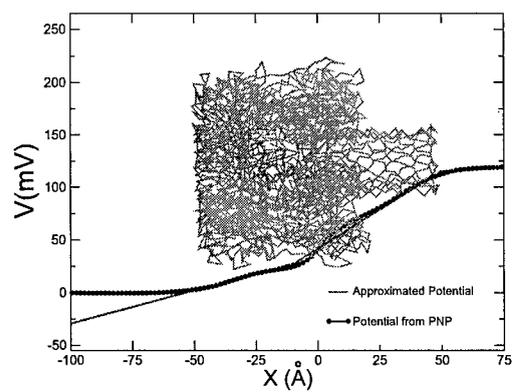


Figure 2:

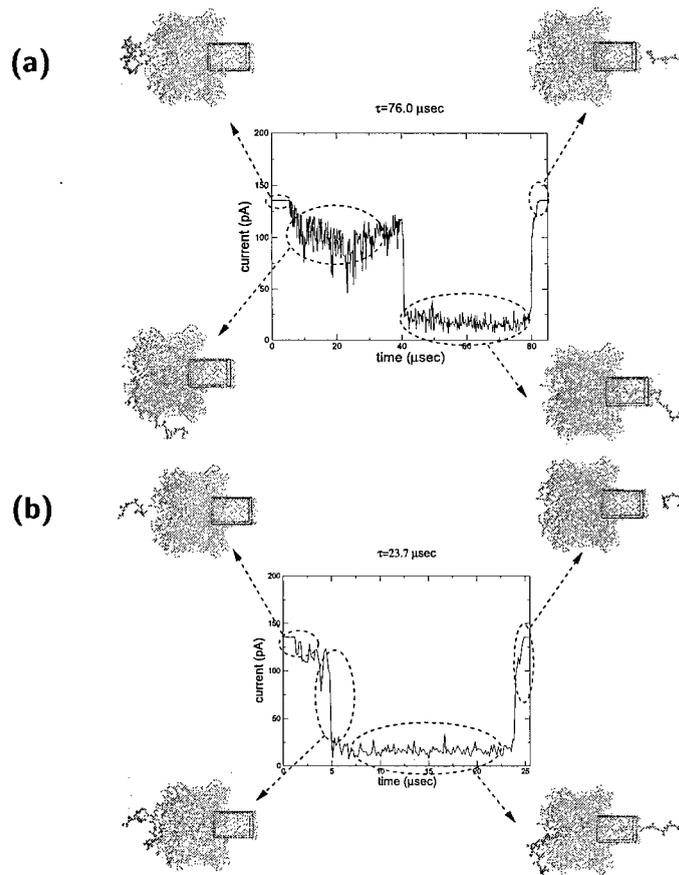


Figure 3:

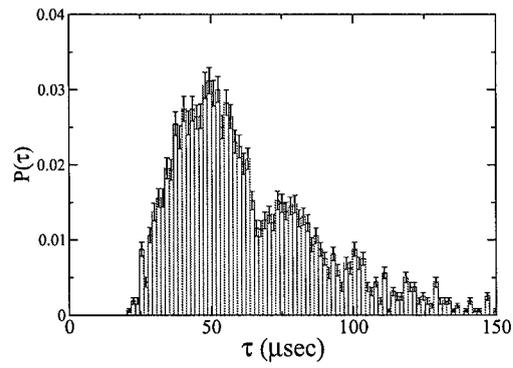


Figure 4:

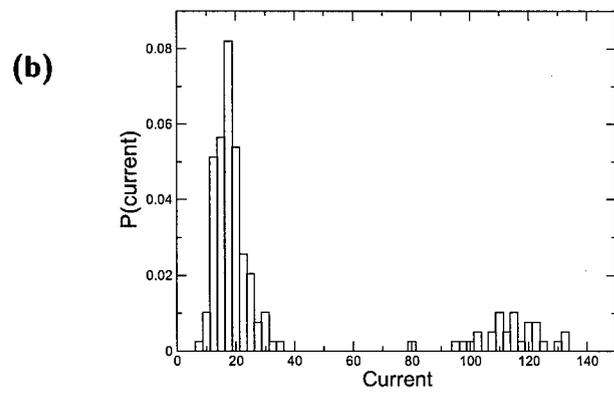
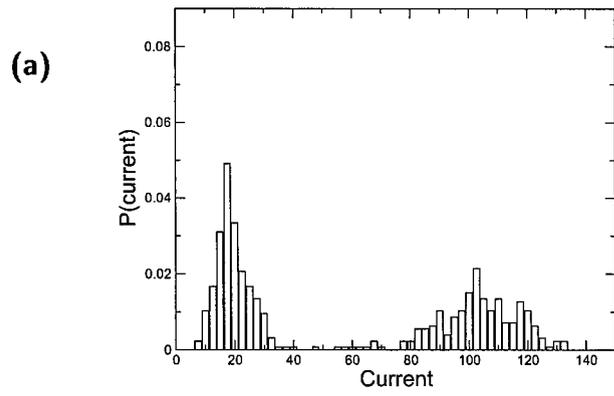


Figure 5:

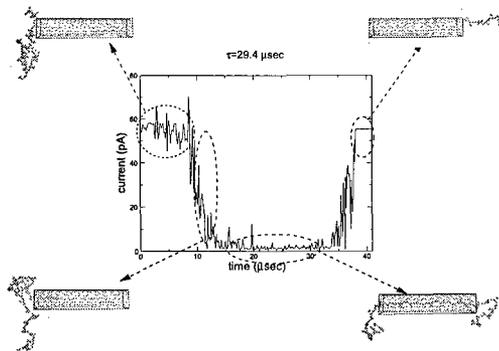


Figure 6:

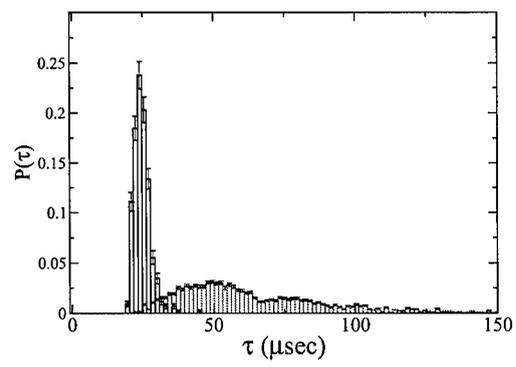


Figure 7:

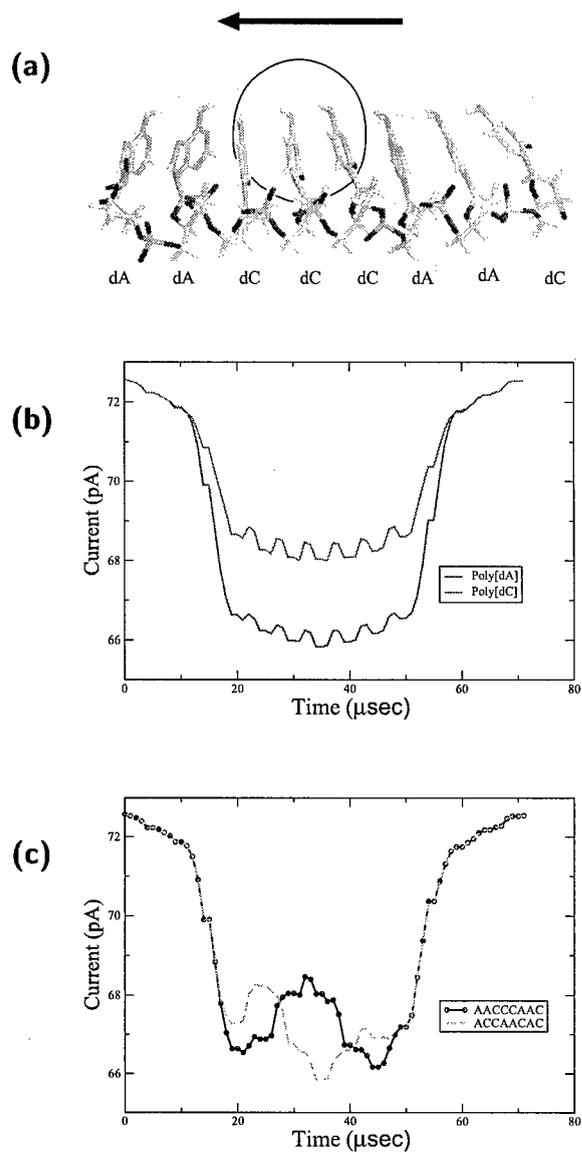


Figure 8: