

## A Statistical Model for Translocation of Structured Polypeptide Chains through Nanopores

Alessandro Ammenti,<sup>†</sup> Fabio Cecconi,<sup>\*,‡</sup> Umberto Marini Bettolo Marconi,<sup>¶</sup> and Angelo Vulpiani<sup>§</sup>

Dipartimento di Fisica, Università di Perugia, Via A. Pascoli, 06123 Perugia, Italy, INFN-CNR, Center for Statistical Mechanics and Complexity, Istituto dei Sistemi Complessi, Via dei Taurini 19, 00185 Rome, Italy, Dipartimento di Fisica, Università di Camerino, Via Madonna delle Carceri, 68032 Camerino (MC), Italy, Dipartimento di Fisica, Università "La Sapienza", INFN and INFN-CNR, P.le Aldo Moro 2, 00185 Rome, Italy

Received: February 2, 2009; Revised Manuscript Received: May 29, 2009

The translocation process of a globular protein (ubiquitin) across a cylindrical nanopore is studied via molecular dynamics simulations. The ubiquitin is described by a native-centric model on a C<sub>α</sub> carbon backbone to investigate the influence of protein-like structural properties on the translocation mechanism. A thermodynamical and kinetic characterization of the process is obtained by studying the statistics of blockage times, the mobility, and the translocation probability as a function of the pulling force  $F$  acting in the pore. The transport dynamics occurs when the force exceeds a threshold  $F_c$  depending on a free-energy barrier that ubiquitin has to overcome in order to slide along the channel. Such a barrier results from competition of the unfolding energy and the entropy associated with the confinement effects of the pore. We implement appropriate umbrella sampling simulations to compute the free-energy profile as a function of the position of the ubiquitin center of mass inside of the channel (reaction coordinate). This free energy is then used to construct a phenomenological drift-diffusion model in the reaction coordinate which explains and reproduces the behavior of the observables during the translocation.

## 1. Introduction

Several molecular processes in living organisms involve the continuous and swift delivery of chemical compounds from production sites to target compartments where they are eventually utilized. Such a transport of biomolecules among inter- or intracellular compartments is known as translocation.<sup>1,2</sup> The understanding of this kind of molecular transfer would clarify the preferential mechanisms through which cells interact with their environment by exchanging metabolites and transmitting chemical signals or genetic information.<sup>3,4</sup> A translocation process that has been fascinating researchers for almost 40 years concerns the transport of proteins along channels (pores) opened in cellular membranes.<sup>5</sup> Its undoubted relevance stems also from the fact that often proteins cross the endoplasmatic reticulum (ER) in eukaryotes or plasma membrane in prokaryotes to complete the biosynthesis stage.<sup>4</sup>

The interest in the subject has recently increased also thanks to the possibility to insert artificial  $\alpha$ -hemolysin into a lipid bilayer, which allowed accurate voltage-driven translocation experiments of polynucleotide segments,<sup>6,7</sup> peptides,<sup>8</sup> and, recently also, proteins.<sup>9</sup> These experiments have shown that  $\alpha$ -hemolysin channels, solid-state nanopores,<sup>10</sup> and carbon nanotubes<sup>11</sup> under appropriate conditions can work as sequencing devices as they are capable of characterizing and discriminating polynucleotide molecules.<sup>12</sup>

Several studies, since the pioneering work by Blobel and collaborators,<sup>13</sup> have highlighted the crucial role of membrane pores to understand and characterize the various import/export

mechanisms of proteins. Proteins, to cross the membranes, use specialized pores which have been designed by evolution to assist and control the process.<sup>3,14,15</sup> The typical size of these pores is on the order of nanometers, and during the transport, proteins remain partially unfolded and refold often with the aid of chaperons and enzymes only when delivered to the trans side of the pore.<sup>5,16</sup>

In the past few years, thorough experimental, computational, and theoretical efforts have produced remarkable advances in understanding the principles of protein translocation. However, this subject still represents a challenge in molecular biology due to the enormous complexity of translocation machinery, the variety of systems, and the presence of several concurrent factors (e.g., chaperones).

Theoretical investigations on protein translocation comprise computational approaches and mathematical modeling. The latter, mainly based on statistical mechanical principles, attempt a phenomenological description of biopolymer translocation as diffusion over energy barriers by assuming elementary "ad hoc" polymer–pore interactions. This approach reveals to be particularly suitable to describe the properties of processes characterized by a low degree of chemical details, such as translocation of unstructured biopolymers or homopolymers.<sup>17–19</sup>

Simulations at atomic resolution<sup>20–22</sup> of both proteins and membrane pores, although very informative on molecular details, explore time scales too short with respect to real translocation times.

The present work combines molecular dynamics simulations and theoretical methods to describe the translocation of a globular protein (ubiquitin) by taking into account its structural properties and network of internal interactions. To cut down on the computational load and increase the statistics of events, we resort to a coarse-grained representation of the protein and

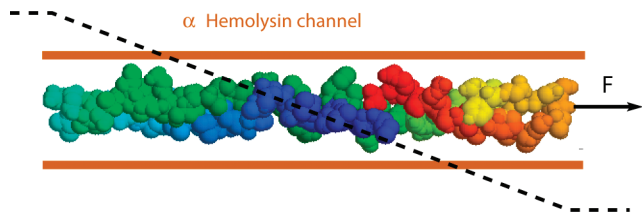
\* To whom correspondence should be addressed.

<sup>†</sup> Università di Perugia.

<sup>‡</sup> Istituto dei Sistemi Complessi.

<sup>¶</sup> Università di Camerino.

<sup>§</sup> Università "La Sapienza".



**Figure 1.** Sketch of the geometrical setup used to simulate the translocation of the ubiquitin through the pore. The arrow represents the pulling homogeneous force acting on the last bead still inside of the pore, and the dashed line is the potential across the channel.

the pore as well. Ubiquitin is modeled via a Gō-like force field on the  $C_\alpha$  carbon backbone, which offers a realistic description as it incorporates the minimal structural properties of protein-like molecules. In this phenomenological approach, the ubiquitin translocation process is simulated as the passage of a protein-like polymer through a pore with a simple geometrical shape (Figure 1). These simplifications allow the entire translocation process to be simulated in an accessible CPU time and with an approximation which retains the unaltered physical essence of the problem.

It is important to note that although ubiquitin is not a protein that is generally imported, we have chosen it for its well-defined topology of secondary structure, because its folding pathway has been well characterized by a variety of methods,<sup>23–25</sup> and also for a comparison to previous theoretical works.<sup>26</sup> Specifically, the issue we address concerns the role of entropic barriers associated with the translocation of proteins across narrow pores. The sudden confinement experienced by polymers when passing a narrow path determines a great conformational entropy reduction leading to the appearance of free-energy barriers opposing translocation. The case of proteins is further complicated with respect to unstructured polymers<sup>27,28</sup> by their natural resistance to unfolding and propensity to refold.

In our study, we followed an approach similar to that suggested by Makarov and co-workers,<sup>26</sup> but at variance with that work, we considered a Gō-like protein model of ubiquitin and finite-length pores. The finite length implies that three dynamical stages can be identified during translocation; (i) a slow process places an initial monomer at the entrance of the channel; in our simulation, this stage is neglected as the N-terminus of ubiquitin is positioned such that translocation starts with no delay; (ii) in the second stage, the protein unfolds and attempts to insert a stable nucleus of residues inside of the channel; (iii) the third stage involves the final escape of the molecule from the channel with consequent refolding. Notice that stages (ii) and (iii) cannot be ignored as they give non-negligible contribution to the full description of the biological process.

From MD simulations combined with the umbrella sampling method,<sup>29</sup> we computed the translocation free-energy profiles as a function of appropriate reaction coordinates, the ubiquitin center of mass along the channel. This technique, as discussed in the methods section, is particularly useful in sampling conformations of a system around energy barriers. The simulations are done in the presence of an artificial biasing potential (umbrella or window potential), forcing the system dynamics to explore, with a certain accuracy, the set of unstable conformations. The effect of the umbrella potential is removed from the Boltzmann weight by processing data through appropriate debiasing algorithms.<sup>30</sup> We employ the free-energy profile to build up a one-dimensional mathematical model based on the Smoluchowski equation. In this framework, we show

that the simulation results on ubiquitin translocation along a cylindrical channel can be interpreted and reformulated as a first-passage problem<sup>31,32</sup> for the motion of an effective particle performing driven-diffusion over a free-energy profile in one dimension.

Although both the structure and the role of membrane pores are enormously more complex, this simple idealization contains enough information to characterize the main physical phenomenology of translocation of partially unfolded proteins, in agreement with the principal experiments.<sup>12</sup>

## 2. Model of Protein Translocation and Simulation Methods

In this section, we present the model for the translocation of ubiquitin through a membrane pore and discuss its essential features.

The pore through which the protein is imported is modeled by a potential with cylindrical symmetry around the  $x$ -axis (translocation direction)

$$V_p(x, y, z) = V_0 \psi(y, z) [1 - \tanh(\alpha x - L)] \quad (1)$$

where  $V_0 = \varepsilon$ ,  $\psi(y, z) = [(y^2 + z^2)/R_p^2]^q$ ;  $L$  and  $R_p$  define the pore length and radius, respectively. The parameter  $q$  tunes the potential (soft wall) stiffness, and  $\alpha$  modulates the soft step-like profile in the  $x$ -direction; the larger the  $\alpha$ , the steeper the step. In this work, we considered  $q = 1$  and  $\alpha \approx 3 \text{ \AA}^{-2}$ . A constant force  $F$  collinear to the cylinder and acting only at the interior of the pore mimics, on average, the importation mechanism. Thus, the pore–protein interaction, apart from the homogeneous pulling force, reduces to a simple confinement effect in a cylindrical region of section  $\pi R_p^2$  and length  $L$ . Since the parameter  $R_p$  is chosen to be smaller than the folded ubiquitin gyration radius, the transversal section of the pore allows translocation to occur only in a quasi-linear conformation.

The ubiquitin is described by its backbone of  $C_\alpha$  carbons only, each residue is identified by the corresponding  $C_\alpha$  atom, and the protein is assimilated to a chain of beads centered on  $C_\alpha$  positions. The interactions among the residues are assigned by promoting the formation of those native interactions stabilizing the 76 residue long PDB structure (1UBI), according to classical Gō-like approaches. The force field, as proposed by Clementi et al.,<sup>33</sup> is defined by the potential

$$V_{\text{tot}} = \sum_{i=1}^{N-1} \frac{k_h}{2} (r_{i,i+1} - R_{i,i+1})^2 + \sum_{i=2}^{N-1} \frac{k_\theta}{2} (\theta_i - \Theta_i)^2 + \sum_{i=3}^{N-2} k_\phi^{(1)} [1 - \cos(\phi_i - \Phi_i)] + k_\phi^{(3)} [1 - \cos 3(\phi_i - \Phi_i)] + \sum_{i,j>i+2} V_{\text{nb}}(r_{ij}) \quad (2)$$

where  $r_{ij}$  is the distance between  $C_\alpha$   $i$  and  $j$ ,  $\theta_i$  is the bending angle identified by the three consecutive  $C_\alpha$ 's  $i-1$ ,  $i$ , and  $i+1$ , and  $\phi_i$  is the dihedral angle defined by the two planes formed by four consecutive  $C_\alpha$ 's  $i-2$ ,  $i-1$ ,  $i$ , and  $i+1$ . The capital symbols are the corresponding quantities in the native conformation. This force field assumes that successive beads along the chain are connected by harmonic springs  $k_h/2(r_{i,i+1} - R_{i,i+1})^2$ , mimicking the virtual peptide bonds. These springs maintain the chain connectivity and allow small fluctuations of consecutive  $C_\alpha$  distances,  $r_{i,i+1}$ , around their native values  $R_{i,i+1}$ . Angular

potentials in  $\theta$  and  $\phi$  contribute to increase the cooperativity of the model by adding a further bias toward the formation of a correct native secondary structure. Nonbonded (nb) interactions between nonconsecutive  $\alpha$ -carbons are modeled with Lennard-Jones 12–10 potentials

$$V_{\text{nb}}(r_{ij}) = \begin{cases} \varepsilon \left[ 5 \left( \frac{R_{ij}}{r_{ij}} \right)^{12} - 6 \left( \frac{R_{ij}}{r_{ij}} \right)^{10} \right] & R_{ij} \leq R_c \\ \frac{10\varepsilon}{3} \left( \frac{\sigma}{r_{ij}} \right)^{12} & R_{ij} > R_c \end{cases} \quad (3)$$

rewarding the formation of native interactions (contacts). Native contacts<sup>34</sup> are identified by a cutoff distance  $R_c = 6.5$  Å, such that two residues  $i$  and  $j$  form a native interaction if their distance  $R_{ij}$  in the native state (PDB entry 1UBI) is less than  $R_c$ . This criterion selects 124 native contacts on the 1UBI crystallographic structure. When two residues are not in native contact ( $R_{ij} > R_c$ ), they interact only through the L-J repulsive tail ( $\sigma/r_{ij}$ )<sup>12</sup>, with  $\sigma = 4.5$  Å. Finally, all of the force field coefficients are proportional to the energy scale  $\varepsilon$ , such that  $k_h = 10^3 \varepsilon / d_0^2$  ( $d_0 = 3.8$  Å),  $k_\theta = 20\varepsilon$ ,  $k_\phi^{(1)} = \varepsilon$ , and  $k_\phi^{(3)} = 0.5\varepsilon$ , and forces are measured in units of  $f_u = \varepsilon \text{ Å}^{-1}$ .

Langevin thermostats with friction  $\gamma$  have been employed to control the temperature in molecular dynamics runs<sup>35</sup>

$$M\ddot{\mathbf{r}}_i = -\gamma M\dot{\mathbf{r}}_i + \mathbf{F}_i^{\text{conf}} + \mathbf{W}_i$$

where  $\mathbf{F}_i^{\text{conf}}$  is the sum of internal and external forces acting on residue  $i$ , the random force  $\mathbf{W}_i$  has zero mean, and the variance is  $\langle W_\alpha(t)W_\alpha(t') \rangle = 2\gamma M R T \delta_{\alpha\alpha} \delta(t - t')$ . The multiple histogram technique<sup>36</sup> has been used to obtain the thermodynamic observables needed to characterize the ubiquitin folding transition.

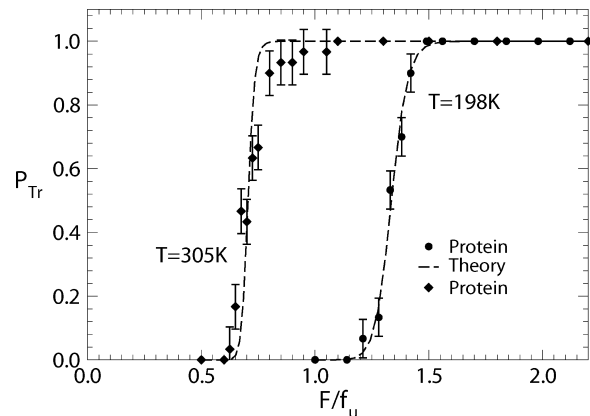
The umbrella sampling method<sup>29</sup> has been applied to extract from simulations the free-energy barriers that the ubiquitin feels across the pore during its translocation. A harmonic umbrella potential

$$V_U(X - X_k) = \frac{K_{\text{umb}}}{2} \left( \frac{1}{N} \sum_{i=1}^N x_i - X_k \right)^2 \quad (4)$$

restrained the  $x$ -coordinate of the protein center of mass,  $X = \sum_{i=1}^N x_i / N$ , to fluctuate around a chosen value  $X_k$ . The enlarged interval  $[-l, L + l]$  ( $L = 300$  Å,  $l = 50$  Å) containing the pore was partitioned into  $N_b = 2 \times 10^3$  bins of width  $\Delta X = (L + 2l)/N_b$ ; a set of  $M = 200$   $X_k$ -centers has been selected to properly sample the whole channel and the regions near its ends. For each center  $X_k$ , we thermalized the configuration and performed a run to collect the biased histogram  $h_k^{(b)}(X)$  of the coordinate  $X$  under the effect of the potential  $V_U(X - X_k)$  centered on  $X_k$ . The biased histograms need to be deweighted from the potential  $V_U(X - X_k)$  and recombined in order to provide the unbiased probability distribution  $P(X)$  along the whole channel, which is expressed as the linear superposition

$$P(X) = C \sum_{k=1}^M c_k(X) e^{\beta V_U(X - X_k)} h_k^{(b)}(X)$$

with  $C$  as a normalization constant. The weights  $c_k(X)$ , normalized such that  $\sum_k c_k(X) = 1$ , are determined by a self-consistent procedure as explained in ref 30. The free-energy  $G(X)$  as a



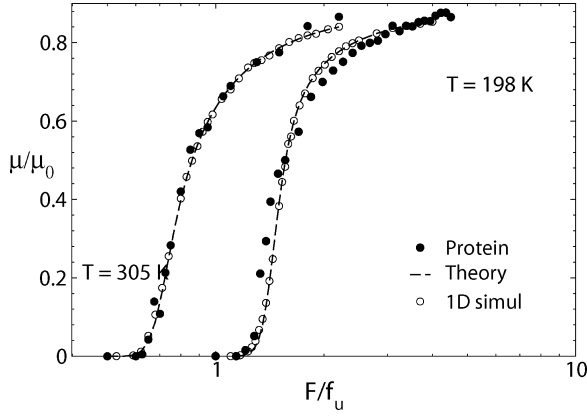
**Figure 2.** Ubiquitin translocation probability across a pore of radius of  $R_p = 4$  Å and length of  $L = 300$  Å as a function of the force (in units  $f_u = 6$  pN) at temperatures of  $T_{\text{ref}} = 198$  K and  $T_{\text{ph}} = 305$  K. Dots are MD simulation results, and dashed lines indicate the fitting via eq 13 (see below).

function of  $X$  (also termed the potential of mean force) is obtained from  $P(X)$  by inverting the Boltzmann weight  $G(X) = -RT \ln P(X)$ .

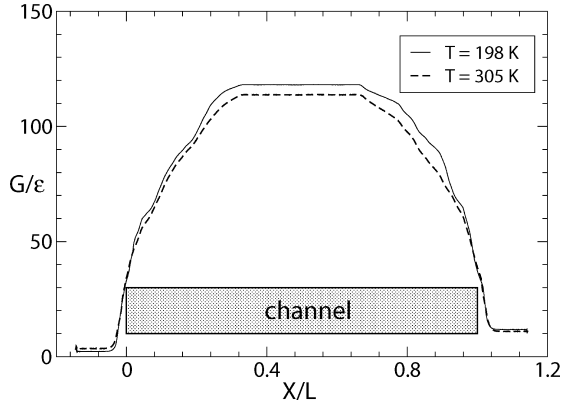
### 3. Results and Discussion

The folding thermodynamics of the 1UBI structure has been obtained through folding simulations from random open conformations at a  $T^* = 1.5$  in reduced temperature units  $R/\varepsilon$ , (equivalent to  $T = 670$  K). The folding temperature occurs at  $T^* = 0.77$ , corresponding to the experimental denaturation temperature  $T = 338$  K;<sup>23</sup> this sets the energy scale to the value  $\varepsilon \approx 0.88$  kcal mol<sup>-1</sup>, and the units of time and force are  $t_u = \sigma(M/120\varepsilon)^{1/2} \sim 0.25$  ps and  $f_u = \varepsilon \text{ Å}^{-1} \sim 6$  pN, where the average amino acid mass amounts to  $M \sim 136$  Da. The time step and friction coefficient used in the Langevin dynamics are  $h = 0.001t_u$  and  $\gamma = 0.25t_u^{-1}$ , respectively. The kinetics of the translocation process was simulated at temperatures of  $T_{\text{ref}} = 198$  K (reference) and  $T_{\text{ph}} = 305$  K (physiological) by performing different runs in which the protein was threaded into the channel by the force. At each run, the ubiquitin N-terminus was placed at the left opening of the channel  $x = 0$  to facilitate the entrance. The simulation was run until the protein was fully expelled out of the right end of the channel, and almost complete refolding occurred. However, at low forces, ubiquitin may fail to cross the pore within a large assigned waiting time  $t_w$ ; in this case, we preferred to stop the run, discard its statistics, and restart a new trajectory.

The process can be characterized by measuring the average translocation velocity and translocation time as a function of the importing force  $F$ . Averages were performed over independent runs, excluding those in which ubiquitin did not cross the pore from one edge to the other within  $t_w$ . The probability of translocation  $P_{\text{Tr}}$  can be estimated as the number of translocation successes, within the time  $t_w \approx 10^5 t_u$ , over the number of total runs; to be precise, this is the probability for the run not to be absorbed in a time  $t < t_w$  because at  $t_w$ , the run is killed. Figure 2 reports the curves of  $P_{\text{Tr}}$  versus  $F$  at physiological  $T_{\text{ph}}$  and reference  $T_{\text{ref}}$  temperatures. At both temperatures, translocation becomes probable ( $P_{\text{Tr}} > 1/2$ ) when the force exceeds a critical value  $F_c(T_{\text{ref}}) \approx 1.35f_u$ ,  $F_c(T_{\text{ph}}) \approx 0.72f_u$ , while below these thresholds, the probability goes rapidly to 0. Instead of the average translocation velocity  $V$ , it is convenient to plot the dimensionless mobility  $\mu/\mu_0 = V/F \cdot (N\gamma M)$  as a function of the force  $F$  (Figure 3), where  $\mu_0 = 1/(N\gamma M)$  is the mobility that



**Figure 3.** Dimensionless mobility  $\mu/\mu_0$  [ $\mu_0 = 1/(N\gamma M)$ ] measuring the ratio between protein speed and force applied to ubiquitin as a function of  $F$  (in units of  $f_u$ ). The curves refer to pore radius  $R_p = 4$  Å and length  $L = 300$  Å at temperatures  $T_{\text{ref}} = 198$  K and  $T_{\text{ph}} = 305$  K. Filled circles indicate the data obtained from MD simulations, and the dashed line is the corresponding theoretical fit (eq 12). Open circles are the result of the numerical integration of the one-dimensional Langevin overdamped equation with a potential  $G_{\text{fit}}(X)$  for a further check on the theory.



**Figure 4.** Free-energy profiles from umbrella sampling simulations, with an umbrella potential (eq 4) of  $K_{\text{umb}} = 0.5\epsilon$ , of the ubiquitin translocation across the pore  $V_p = \epsilon$ ,  $R_p = 4$  Å at temperatures  $T_{\text{ref}}$  and  $T_{\text{ph}}$ .

ubiquitin would have in the absence of the channel. The curves show a nonlinear  $V$ – $F$  characteristic, indicating the existence of a free-energy barrier which the molecule has to overcome to activate its translocation. The threshold mechanism is generated by the channel–protein interaction, and the theoretical approach explaining the results requires the determination of the free-energy profiles associated with ubiquitin translocation. For instance, the decrease of the critical force with temperature is a clear consequence of the lowering of free-energy translocation barriers with temperature increase. This can be achieved via umbrella sampling simulations with the umbrella potential (eq 4) as discussed in section 2.

Typical free-energy profiles  $G(X)$  in the absence of force experienced by ubiquitin in different positions inside of the pore are reported in Figure 4 as a function of the “natural” reaction coordinate, the center of mass position  $X$ . The pore occupies the region  $[0, L]$ , and the plateau in the middle indicates that once the protein is unfolded and well inside of the pore, it can slide along the axis without free-energy increase. It is worth remarking that the presence of the plateau is a consequence of the channel length, which is long enough to fully accommodate all of the residues of the elongated conformations of ubiquitin. By reducing the length  $L$ , this plateau is expected to shorten

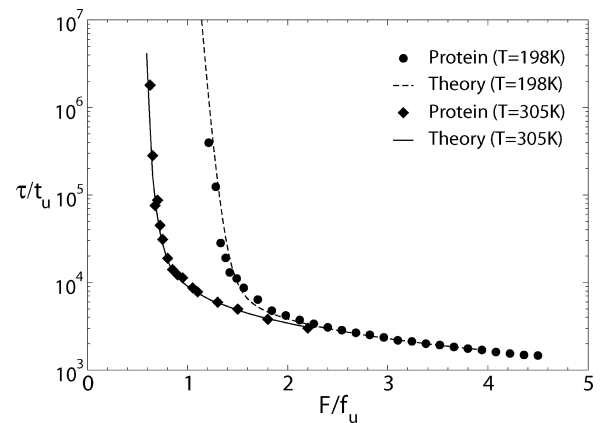
**TABLE 1: Parameter Values of the Fitting Function  $G_{\text{fit}}(X)$  at Temperatures  $T_{\text{ref}}$  and  $T_{\text{ph}}$**

$s$	$G_s$ [kcal/mol]	$\mu_s$ [Å <sup>-2</sup> ]	$l_s$ [Å]
$T_{\text{ref}} = 198$ K			
1	33.4ε	$2.79 \times 10^{-4}$	150.8
2	10.5ε	$4.90 \times 10^{-4}$	122.3
3	15.4ε	$2.90 \times 10^{-4}$	88.2
$T_{\text{ph}} = 305$ K			
1	13.4ε	$5.90 \times 10^{-4}$	152.9
2	34.1ε	$1.20 \times 10^{-4}$	130.0
3	8.5ε	$3.30 \times 10^{-4}$	81.6

until disappearing below a certain value of  $L$ . The major variation of  $G(X)$  occurs at the boundaries of the channel, and its shape can be fit by a one-dimensional potential

$$G_{\text{fit}}(X) = \sum_{s=1}^3 G_s g\{\mu_s[(X - L/2)^2 - l_s^2]\}$$

which is a combination of step-like functions  $g(u) = [1 - \tanh(u)]$ , suggested by the potential in eq 1, depending on the tuning parameters  $\{G_s\}$ ,  $\{\mu_s\}$ , and  $\{l_s\}$ , whose values are listed in Table 1. One may wonder how the free-energy profiles are influenced by the parameters  $R_p$ ,  $\mu$ , and  $q$  in the pore potential definition (eq 1). Of course, if  $R_p$  is increased, the pore becomes larger, and the translocation is facilitated; then, the free energy barrier decreases. The role of  $\alpha$  is very marginal; in fact, we checked that results do not change by replacing  $\tanh[\alpha x(L - x)]$  with the step function  $\Theta[x(L - x)]$ . The role of  $q$  is less trivial than  $\alpha$ , although we do not see appreciable variations on the translocation dynamics passing from  $q = 1$  (present simulations) to 2. However, from eq 1, it is clear that changing  $q$  corresponds to making the wall less soft. To some extent, this is equivalent to a small change in the radius of the channel. For example, in the limit of  $q \rightarrow \infty$ , the channel is a perfect hard cylindrical wall with radius  $R_p$ . In this case, there is no energetic contribution of the channel, and the free-energy barrier is only due to ubiquitin properties. Even if we had not explored the range of  $q$ , we are confident that the scenario that we report is robust with respect to variation of  $q$ . We preferred to work with low  $q$  to avoid as many instabilities as possible in the MD code. Once the ubiquitin had placed enough residues in the pore (i.e., CM coordinate  $X \approx 100$  Å in Figure 4), it could slide



**Figure 5.** Average translocation time  $\tau$  (in units of  $t_u = 0.25$  ps) of ubiquitin as a function of the force  $F$  at temperatures  $T = 198$  and  $305$  K and  $R_p = 4$  Å. Symbols indicate simulation results and lines the corresponding fit by eq 15.



without further free-energy cost. In the proximity of the channel right end, the molecule started to be almost spontaneously expelled, and the results of Figure 3 and Figure 5 suggest a description based on an effective Smoluchowski equation for the probability density  $P(X,t)$  of the reaction coordinate  $X$

$$\frac{\partial P}{\partial t} = D_0 \frac{\partial}{\partial X} \left\{ e^{-\beta U(X)} \frac{\partial}{\partial X} e^{\beta U(X)} P \right\} \quad (5)$$

where  $U(X) = G(X) - FX$ ,  $\beta = (RT)^{-1}$ , and  $D_0$  and  $F$  are the effective diffusion constant of the protein and pulling force, respectively. The theory requires the assignment of proper boundary conditions at the edges of the channel  $X = 0$  and  $L$ . A meaningful choice is the radiative boundary conditions<sup>37,38</sup> for the current

$$J(X,t) = -D_0 e^{-\beta U(X)} \partial_X [e^{\beta U(X)} P(X,t)]$$

$J(0,t) = -R_0 P(0,t)$  and  $J(L,t) = R_L P(L,t)$ , which explicitly read as

$$\begin{cases} \left. \partial_X e^{\beta U(X)} P(X,t) \right|_0 = R_0 / D_0 e^{\beta U(0)} P(0,t) \\ \left. \partial_X e^{\beta U(X)} P(X,t) \right|_L = -R_L / D_0 e^{\beta U(L)} P(L,t) \end{cases} \quad (6)$$

taking into account the possibility for the protein to be spontaneously expelled from the channel at rates  $R_0$  and  $R_L$  from the left and right, respectively.

Since we are interested in those events for which the molecule occupies the channel for a given time (blockage time), we consider the survival probability at time  $t$ , that is, the probability that the molecule has not yet escaped the channel

$$S(t) = \int_0^L dX P(X,t) \quad (7)$$

The probability  $S(t)$  is related to the distribution of blockage times  $\psi(t)$  via the expression

$$S(t) = 1 - \int_0^t d\tau \psi(\tau)$$

thus, by a simple differentiation, we obtain

$$\psi(t) = -\frac{dS(t)}{dt} \quad (8)$$

The use of eqs 5 and 7 and of the boundary conditions leads to

$$\psi(t) = R_0 P(0,t) + R_L P(L,t) = f_0(t) + f_L(t) \quad (9)$$

where  $f_0(t)$  and  $f_L(t)$  are the fluxes at the boundaries whose explicit expressions are obtained by solving the boundary value problem with the initial condition  $P(X,0) = \delta(X - X_0)$ . This initial condition assumes that the translocating molecule is released exactly at the place  $X = X_0$  inside of the channel. By definition, the integration of the flux  $f_L(t)$  over all time provides the probability of translocation

$$P_{\text{Tr}} = R_L \int_0^\infty dt P(L,t) \quad (10)$$

The average time spent in the channel by the molecule is given by

$$\tau = \langle t \rangle = \int_0^\infty dt t \psi(t) = \int_0^\infty dt S(t) \quad (11)$$

The last identity stems from eq 8 and an integration by parts. Time  $\tau$  would correspond in the experiments and simulations to blockage time; then, the average velocity can be estimated via the formula

$$V \simeq \frac{L}{\tau} \quad (12)$$

A complete solution of eq 5 is needed to make the analytical expression of  $V$ ,  $\tau$ , and  $\psi(t)$  explicit. Equation 5 cannot be solved in general; however, as we show in the Appendix, key quantities such as  $P_{\text{Tr}}$  and  $\tau$  can be obtained. These are given by the formulas

$$P_{\text{Tr}}(F) = \frac{D_0 K_L}{D_0 (K_0 e^{-\beta FL} + K_L) + K_0 K_L M_+(F)} \quad (13)$$

and

$$\tau(F) = \frac{D_0 M_-(F) + K_L e^{\beta FL} M_0(F)}{D_0 (K_0 + K_L e^{\beta FL}) + K_0 K_L e^{\beta FL} M_+(F)} \quad (14)$$

respectively. To simplify the notation, we have introduced  $K_0 = R_0 \exp\{-\beta G(0)\}$  and  $K_L = R_L \exp\{-\beta G(L)\}$

$$M_\pm(F) = \int_0^L dx e^{\pm \beta [G(x) - Fx]}$$

and

$$M_0(F) = \int_0^L dx e^{-\beta [G(x) - Fx]} \int_x^L dy e^{\beta [G(y) - Fy]}$$

Since, in our simulations, the average time is computed only over the runs for which translocations occurred, that is,  $P_{\text{Tr}}(F) = 1$ , we need to take the limits  $K_0 \rightarrow 0$  and  $K_L \rightarrow \infty$  in the above expression to get the quantity directly comparable to the data

$$\tau(F) \simeq \frac{M_0(F)}{D_0} \quad (15)$$

For each value of  $F$ , the integration is carried out numerically, replacing  $G(x)$  by its fitting function  $G_{\text{fit}}(x)$ . Equations 13 and 15 allow the observed translocation phenomenology to be qualitatively explained at different forces. Moreover, once parameters  $D_0$ ,  $K_0$ , and  $K_L$  have been properly adjusted to fit the data, they reproduce also quantitatively the behavior of the corresponding quantities measured in MD simulations. The agreement between numerical data and theory can be appreciated

in Figures 2 and 3, where the behavior of the simulated translocation probability and the velocities of ubiquitin as functions of  $F$  are well fitted by eqs 13 and 12, respectively, with  $\tau$  from eq 15. To obtain satisfactory fitting curves, the free energy (Figure 4) has been rescaled by a factor of 0.72 at  $T_{\text{ref}} = 198$  K and a factor of 0.85 at  $T_{\text{ph}} = 305$  K. Such a renormalization is presumably ascribable to the crude approximation of considering the ubiquitin translocation as a one-dimensional process described by a single reaction coordinate, which, although reasonable, might not be fully appropriate. Other reaction coordinates are needed to fully characterize the kinetics.

The analytical expression of  $P_{\text{Tr}}(F)$  can be used to determine the nontrivial dependence of the critical force on the free-energy barrier. In fact  $F_c$  can be defined as the  $F$  value such that the probability to translocate and that not to translocate are equal,  $P_{\text{Tr}}(F_c) = 1/2$ . This leads to the nonlinear equation

$$1 = \frac{K_0}{K_L} e^{-\beta F_c L} + \frac{K_0}{D_0} M_+(F_c)$$

where the dependence on the free-energy barrier  $\Delta = G_{\text{max}} - G(0)$  is hidden in the function  $M_+(F)$  previously defined. We can differentiate the above expression

$$0 = - \left\{ \beta L \frac{K_0}{K_L} e^{-\beta F_c L} + \frac{K_0}{D_0} \left| \frac{\partial M_+}{\partial F_c} \right| \right\} \frac{dF_c}{d\Delta} + \frac{K_0}{D_0} \frac{\partial M_+}{\partial \Delta}$$

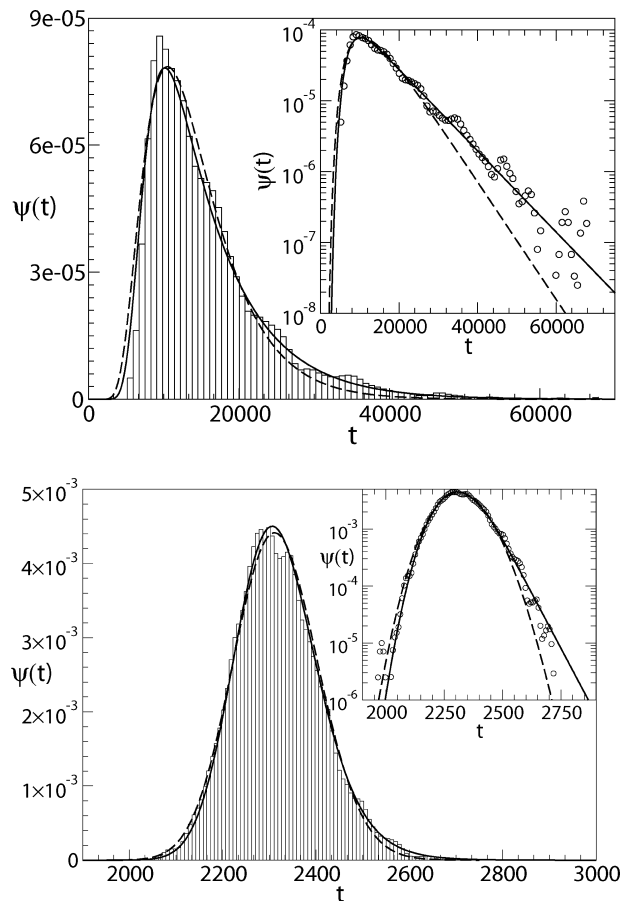
where, we used the existence of a relationship  $F_c = F_c(\Delta)$  between the critical force and the barrier  $\Delta$  and the fact that, by definition,  $M_+$  is a function decreasing with  $F$  and increasing with  $\Delta$ ; thus, its derivatives with respect to  $F_c$  and  $\Delta$  are negative and positive, respectively. Then, solving for the derivative

$$\frac{dF_c}{d\Delta} = \frac{\frac{\partial M_+}{\partial \Delta}}{\frac{\beta L D_0}{K_L} e^{-\beta F_c L} + \left| \frac{\partial M_+}{\partial F_c} \right|} > 0$$

we see that  $dF_c/d\Delta$  is positive as the second member is positive. Thus, even if we cannot make explicit the dependence of  $F_c$  on  $\Delta$  without performing trivial approximations, we can say that the positiveness of  $dF_c/d\Delta$  clearly justifies the statement that an increase in the barrier corresponds to an increase of  $F_c$ , in agreement with the numerical results of Figure 2.

Further information on the translocation process can be gained by studying the statistics of translocation times. The time statistics of translocation events is accessible to experiments<sup>6,12</sup> in which current drops signal the occupation of the channel by the passing molecule; for this reason, these times are also referred to as blockage times. In our simulations, translocation times can be measured as first arrival times  $t$  at the channel end  $x = L$  of the protein center of mass. Figure 5 shows the average translocation time  $\tau$  at temperatures  $T_{\text{ref}}$  and  $T_{\text{ph}}$  as a function of the applied force  $F$ . The reduction of  $\tau$  is not simply proportional to  $F$  but exhibits the behavior predicted by eq 15. Only in the large force regime does it follow an Arrhenius-like law

$$\tau(F) \sim \tau_0 \exp(-\beta FL)$$



**Figure 6.** Distributions of translocation times across a channel of length  $L = 300$ , radius  $R_p = 4$  Å, and temperature  $T = 198$  K. The top panel refers to field  $F = 1.38f_u$ , near the critical region. The bottom panel corresponds to field  $F = 3.0f_u$ , well outside of the critical force region. The dashed line represents the fit via eq 16, with parameters  $D_0$  and  $\mu_0$ . The solid line indicates the semianalytical result obtained from eq 5 in the approximation  $G(x) \approx G_0 - Fx$  and the radiative boundary conditions (eq 6). The inset shows the same data on a semilog scale.

with  $\tau_0$  being the Kramers's time in the absence of the field  $F$ . Distributions of times shown in Figure 6 are obtained by collecting the histograms from MD simulations in about  $10^3$  numerical experiments of ubiquitin translocation on regimes of small and large forces. The distributions are not Gaussian as they present a moderate degree of skewness; however, the tail asymmetry becomes less marked as  $F$  increases, and in the limit of very large forces, the Gaussian character is recovered. This is in agreement with experimental results.<sup>10,12</sup> Unfortunately, the analytical determination of the time distribution requires the complete Smoluchowski problem to be solved. Since this is not possible, one has to resort to reasonable approximations. A first attempt to fit data can be done via the function

$$\psi(t) = \frac{L}{\sqrt{4\pi D_0 t^3}} \exp \left\{ -\frac{(\mu_0 F t - L)^2}{4 D_0 t} \right\} \quad (16)$$

which is the distribution of first arrival times to the position  $L$  of biased random walkers starting from the origin in the assumptions of constant drift  $v_0 = \mu_0 F$ , semi-infinite channel  $[-\infty, L]$ , and absorbing boundaries at  $x = -\infty$  and  $L$ . Although eq 16 provides a reasonable fit to histograms (see Figure 6), the simple biased random walk results in being a naive

approximation to explain the complexity of a protein translocation process strongly characterized by free-energy barriers. Translocation events could be assimilated to the paths of biased random walks only as long as the process is so fast that the protein has no time to explore the boundary barriers. The semilog scale (inset of Figure 6) reveals that at low forces, the large-time tail predicted by eq 16 deviates more and more from simulation data.

A better description of  $\psi(t)$  is gained by considering eq 5 in the approximation  $G(x) \simeq G_0 = \text{const}$ , an assumption rigorous only in the region of  $100 < X < 200$  Å, well inside of the pore (see Figure 4). In this approximation, the effects of the potential barriers at the two channel ends can be partially recast in a proper redefinition of parameters  $R_0$  and  $R_L$  defining the boundary conditions.

The Smoluchowski problem, eq 5 in a constant potential and boundaries in eq 6, can now be fully solved via the technique of Laplace transformation (see the Appendix and Berezhkovskii et al.<sup>38</sup>), and we obtain an analytical expression for the Laplace transform

$$\hat{f}_L(s) = \frac{RZ(s)e^{v_0L/2D_0}}{RZ(s) \cosh(\lambda L) + (v_0R + 2D_0s) \sinh(\lambda L)}$$

of the translocation time distribution  $\psi(t)$ , with  $\lambda = Z(s)/2D_0$ ,  $R = R_L$  (radiation coefficient),  $v_0 = \mu_0F$ , and  $Z(s) = (v_0^2 + 4D_0s)^{1/2}$ . The antitransformation of the above expression must be done numerically by standard packages<sup>39</sup> and provides a time distribution that formally reads as

$$\psi(t) = a(t)e^{-\Gamma t}$$

where  $a(t)$  is a time-dependent amplitude and the exponential tail at large times is characterized by the rate  $\Gamma$ , determined by the first pole of the  $\hat{f}_L(s)$ . Such a pole is obtained by retaining the negative root  $s_0$  of the equation

$$\tanh\left[\frac{L}{2D_0}\sqrt{v_0^2 + 4D_0s}\right] = -\frac{v_0R + 2D_0s}{R\sqrt{v_0^2 + 4D_0s}}$$

with the minimal modulus  $|s_0|$ . The fine-tuning of parameters  $D_0$ ,  $v_0$ , and  $R$  allows a satisfactory fitting of the bulk of time distributions and of the tails as well. The comparison between the simulated distribution and theory is shown in Figure 6 for a force near the critical value  $F_c$  and well beyond it. The agreement is remarkable also in a semilog plot, showing that the tails are better reproduced with respect to the simple decay  $\exp\{-t\mu_0^2F^2/(4D_0)\}$  predicted by eq 16. Of course, the presence of a supplementary parameter  $R$  offers a greater flexibility in fitting data.

#### 4. Conclusions

We proposed a simplified model of translocation where a small globular protein, represented by the ubiquitin, is imported across a finite size channel by a uniform pulling force acting inside of the channel only. This paper is an attempt to compare MD simulations at the coarse-grain level of protein translocation to its theoretical interpretation by the drift-diffusion Smoluchowski equation with radiation boundary conditions. With respect to previous studies,<sup>26</sup> we considered finite length channels, a simulation setup closer to real translocation processes

occurring in cellular membrane pores and also similar to voltage experiments on synthetic  $\alpha$ -hemolysin channels.<sup>6,12</sup>

We simulated translocation kinetics via Langevin molecular dynamics to control the temperature. The analysis of trajectories has shown that the translocation of structured polypeptide chains (like proteins) leads to very different results than the translocation of small peptides and unstructured polymers (e.g., homopolymers). For proteins, indeed, the process occurs via three stages, unfolding  $\rightarrow$  translocation  $\rightarrow$  refolding. This introduces two time scales that interplay to generate the total blockade time  $\tau_B$  of the pore. One is related to delay due to unfolding/refolding processes occurring at the ends of the channels, and the other one corresponds to the net translocation motion. This scale separation can be resolved only if working with sufficiently long pores. The central point of this work is to have shown that the phenomenology of a computational (coarse-grained) model of protein translocation can be described by the formalism of the one-dimensional driven-diffusion Smoluchowski equation in the collective variable  $X$  (ubiquitin center of mass) and by employing the free-energy profile  $G(X)$  as an effective potential. The shape of  $G(X)$ , computed from umbrella sampling simulations, looks like a bump of large width, characterized by a plateau in the region well inside of the pore, while its maximal variation occurs at the boundaries. The plateau occurs as we choose the pore length of  $L = 300$  Å, comparable to the length of the fully stretched ubiquitin conformation,  $Nd_0 \simeq 76 \times 3.8$  Å.

The driven-diffusion Smoluchowski model for the probability  $P(x,t)$  that ubiquitin has its center of mass  $X$  on the position  $x$  at time  $t$  is amenable to analytic treatment. We have shown that upon tuning four parameters plus the free energy barrier at the entrance of the channel, it accounts for the general behavior of the observables characterizing the translocation as a function of the importing force  $F$ . This theory reproduces the simulation results rather well, such as the translocation probability, mobility of the protein inside of the channel, and average blockade times and their distribution upon changing  $F$ . However, despite the success, two basic assumptions: (i) the center of mass of a protein is a reliable collective coordinate to describe translocation pathways and (ii) quasi one-dimensional approximations have to be carefully verified in future works. This approach, indeed, is expected to fail when several factors affect the translocation and more collective coordinates are necessary. In the case studied, the 1-D theory works pretty well, as we considered the pore as an ideal object whose only effect is to confine the protein and prevent its partial refolding inside of the channel. If the pore, for instance, would be large (or short) enough to allow the folding of non-negligible regions of native secondary structure, then a second coordinate  $Q$  (protein nativeness) at least would be necessary to discriminate different translocation pathways.

Previous interesting computational works (see, e.g., ref 26), which inspired the present one, limited the use of free-energy computation to draw qualitative characterization of the translocation; we pushed the analysis further. Indeed, the free-energy landscape in one or a few reaction coordinates turns to be the unifying element between the low-dimensional statistical description of translocation and computations or experiments. It is the result of the complex microscopic interaction of the passing protein with itself and with the pore, but, in principle, it contains almost all of the information to describe protein translocation as a first-passage time problem of a driven-diffusion stochastic process.

Mathematical approaches,<sup>18,19,37,40</sup> instead, which focus on the statistical properties (e.g., scaling laws) of translocation, obviously tend, for the sake of generality, to overlook the real shape of free-energy profiles by introducing “ad hoc” approximations. All of these models, although having the advantage of being analytically or semianalytically solvable, underestimate the contribution of chain conformational entropy as long as they limit to consider its simple approximations borrowed from polymer statistical physics. The approach results arbitrary in protein-like polymers whose resistance to unfolding introduces further slowdown to the process and uncertainty on entropy contributions. In conclusion, our analysis supports driven-diffusion phenomenological modeling of translocation, providing the possibility to replace arbitrary phenomenological potentials with the true free-energy landscapes in few appropriate reaction coordinates. On the experimental side, our results suggest increasing the efforts focused on the determination of spatial distribution of molecules inside of the channel. This, besides improving our understanding of translocation phenomenology, would provide a solid ground and benchmarks to both computational modeling and theoretical explanations.

A more realistic computational model of translocation could be realized by introducing more microscopic details in the pore–protein interaction and improving the microscopic description of the pore structure.

As a conclusive remark, it is interesting to note that both the mobility and translocation probability feature a clear nonlinear behavior typical of activated transport phenomena across a barrier observed in other physical systems (e.g., granulars) in which interactions and fluctuations requires some amount of collective behavior for transport to occur.<sup>41</sup>

**Acknowledgment.** The CPU time was provided by the PC-farm from Istituto Nazionale di Fisica Nucleare and Dipartimento di Fisica Università degli Studi di Perugia. We thank Massimo Cencini for a careful reading of the manuscript. This work has been supported by the Grant No. 827 “Dinamiche cooperative in strutture quasi unidimensionali” within the CNR programme Ricerca Spontanea a Tema Libero.

## Appendix

In this appendix, we show how to derive the translocation quantities  $P_{Tr}$  and  $\tau$  as a function of  $F$ . There is no need to solve eq 5 in the potential  $U(x) = G(x) - Fx$ . In this respect, it is useful to introduce the Laplace transformation

$$\hat{P}(x, s) = \int_0^\infty dt e^{-st} P(x, t) \quad (17)$$

that changes the partial differential eq 5 into the ordinary differential equation

$$D_0 \frac{d}{dx} \left\{ e^{-\beta U(x)} \frac{d}{dx} e^{\beta U(x)} \hat{P} \right\} - s \hat{P} = -\delta(x - x_0) \quad (18)$$

with the boundary in eq 6 and initial condition  $P(x, t = 0) = \delta(x - x_0)$ . Notice that the initial condition  $\delta(x - x_0)$  implies that we are looking for a fundamental solution  $P(x, t|x_0, 0)$  whose physical interpretation is the probability that particles released in  $x_0 \in [0, L]$  at time  $t = 0$  reach the point  $x$  at time  $t$ . For sake of notation simplicity, we drop out the explicit dependence on  $x_0$ ; thus, we write  $P(x, t)$  instead of  $P(x, t|x_0, 0)$  and  $\hat{P}(x, s)$  for its Laplace transform.

The translocation probability and translocation time are obtained by the solution of eq 18 for  $s = 0$  and  $\hat{P}(x, s = 0)$  (denoted by  $\hat{P}(x)$ ). Indeed, from eqs 11 and 10, one can write

$$\tau = \int_0^\infty dt \int_0^L dx P(x, t) = \int_0^L dx \hat{P}(x) \quad (19)$$

Moreover, eq 13, in terms of  $\hat{P}(x)$ , becomes

$$P_{Tr} = R_L \hat{P}(L) \quad (20)$$

It is convenient to introduce the new function  $Y(x) = \exp\{\beta U(x)\} \hat{P}(x)$  and to solve the equation

$$D_0 \frac{d}{dx} \{ e^{-\beta U(x)} Y'(x) \} = -\delta(x - x_0) \quad (21)$$

The theory of Green's Functions prescribes that a general solution  $Y(x)$  of eq 21, satisfying the boundary conditions  $Y_1'(0) = R_L/D_0 Y_1(0)$  and  $Y_2'(L) = R_L/D_0 Y_2(L)$ , respectively, is built up by two independent solutions  $Y_1(x)$  and  $Y_2(x)$  of the homogeneous equation. The two solutions are

$$Y_1(x) = \frac{D_0}{R_0} e^{\beta U(0)} + \int_0^x d\xi e^{\beta U(\xi)}$$

and

$$Y_2(x) = -\frac{D_0}{R_L} e^{\beta U(L)} - \int_x^L d\xi e^{\beta U(\xi)}$$

which provides the solution of eq 18 for  $s = 0$ , in the form

$$\hat{P}(x) = \begin{cases} -\hat{P}_1(x) \hat{P}_2(x_0) / \text{Wr}(x_0) & 0 < x < x_0 \\ -\hat{P}_2(x) \hat{P}_1(x_0) / \text{Wr}(x_0) & x_0 < x < L \end{cases}$$

where  $\text{Wr}(x) = [\hat{P}_1(x) \hat{P}_2'(x) - \hat{P}_2(x) \hat{P}_1'(x)]$  is the Wronskian. By taking the limit of  $x_0 \rightarrow 0$  and applying simple algebraic manipulation on eqs 19 and 20, we derive from  $\hat{P}(x)$  eqs 13 and 15, respectively.

To obtain the time distribution, we need to solve eq 5 for arbitrary free-energy profiles  $U(x) = G(x) - Fx$  and not vanishing  $s$ . This is impossible; thus, we have to resort to the simplest approximation  $G(x) = G_0 = \text{const}$  along the channel and try to transfer the effect of free-energy barriers on the boundary conditions. Equation 5 simplifies to

$$D_0 \hat{P}''(x) + \mu_0 F \hat{P}'(x) - s \hat{P}(x) = -\delta(x - x_0)$$

With the radiation boundary conditions in eq 6 that take on the form  $\hat{P}'(0) = (\nu_0 + R_0)/D_0 \hat{P}(0)$  and  $\hat{P}'(L) = (\nu_0 - R_L)/D_0 \hat{P}(L)$ , we introduce  $\nu_0 = \mu_0 F$ ,  $K_0 = \nu_0 + R_0$ , and  $K_L = \nu_0 - R_L$ . As above, two fundamental solutions are needed to obtain the Green function of this simplified problem; they are

$$\hat{P}_1(x) = e^{\nu_0 x / 2D_0} \{ Z(s) \cosh(\lambda L) + A \sinh(\lambda x) \}$$



$$A = 2K_0 - v_0$$

for the  $x = 0$  boundary and

$$\hat{P}_2(x) = e^{v_0 x/2D_0} \{Z(s) \cosh[\lambda(L - x)] + B \sinh[\lambda(L - x)]\}$$

$$B = v_0 - 2K_L$$

for  $x = L$ , where  $\lambda = Z(s)/2D_0$  and  $Z(s) = (v_0^2 + 4D_0s)^{1/2}$ . The Laplace transform of the Green function at the boundary  $x = L$  provides the flux  $\hat{f}_L(s) = R_L \hat{P}(x = L)$

$$\hat{f}_L(s) = R_L Z(s) e^{v_0 L/2D_0} / \{ (K_0 - K_L) Z(s) \cosh(\lambda L) + [v_0(K_0 + K_L) - 2K_0 K_L + 2D_0 s] \sinh(\lambda L) \}$$

Actually, we are interested in situations very close to our simulation setup that corresponds to a condition where  $R_0 = 0 \rightarrow K_0 = v_0$ , reflecting boundary (zero current) at  $x = 0$  and radiative ( $R_L = R > 0 \rightarrow K_L = v_0 - R$ ) at  $x = L$ . After algebraic manipulations, we obtain the analytical expression of the flux corresponding to translocation; then, we get

$$\hat{f}_L(s) = \frac{RZ(s)e^{v_0 L/2D_0}}{RZ(s) \cosh(\lambda L) + (v_0 R + 2D_0 s) \sinh(\lambda L)} \quad (22)$$

We need, now, to invert the Laplace Transform of  $\hat{f}_L(s)$  to determine the distribution  $\psi(t)$  according to eq 9. This can be done numerically through the algorithm.<sup>39</sup> The parameters  $D_0$  and  $R$  have to be considered as tunable quantities to fit simulation or experimental data.

## References and Notes

- (1) Singer, S. J. *Annu. Rev. Cell Biol.* **1990**, *6*, 247–296.
- (2) Schatz, G.; Dobberstein, B. *Science* **1996**, *271*, 1519–1526.
- (3) Wickner, W.; Schekman, R. *Science* **2005**, *310*, 247–296.
- (4) Lodish, H.; Baltimore, D.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Durnell, J. *Molecular cell biology*; Freeman: New York, 1996.
- (5) Agarraberes, F.; Dice, J. *Biochem. Biophys. Acta* **2001**, *1513*, 1–24.
- (6) Kasianowicz, J.; Brandin, E.; Branton, D.; Deamer, D. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13770–13773.
- (7) Meller, A. *J. Phys.: Condens. Matter* **2003**, *15*, R581–R607.
- (8) Sutherland, T.; Long, Y.-T.; Stefureac, R.; Bediako-Amoa, I.; Kraatz, H.-B.; Lee, J. *Nano Lett.* **2004**, *4*, 1273–1277.

- (9) Oukhaled, G.; Mathé, J.; Biance, A. L.; Bacri, L.; Betton, J. M.; Lairez, D.; Pelta, J.; Auvray, L. *Phys. Rev. Lett.* **2007**, *98*, 158101–158104.
- (10) Jiali, L.; Gershow, M.; Stein, D.; Brandin, E.; Golovchenko, J. *Nat. Mater.* **2003**, *2*, 611–615.
- (11) Yeh, I.-C.; Hammer, G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 12177–12182.
- (12) Meller, A.; Nivon, L.; Brandin, E.; Golovchenko, J.; Branton, D. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 1079–1084.
- (13) Lingappa, V. R.; Katz, F. N.; Lodish, H. F.; Blobel, G. *J. Biol. Chem.* **1978**, *253*, 8667–8670.
- (14) Römisch, K.; Collie, N.; Soto, N.; Logue, J.; Lindsay, M.; Scheper, W.; Cheng, C.-H. C. *J. Cell Sci.* **2003**, *116*, 2875–2883.
- (15) Rapaport, T. *Nature* **2007**, *450*, 663–669.
- (16) Goloubinoff, P.; De-Los-Rios, P. *Trends Biochem. Sci.* **2007**, *32*, 372.
- (17) Sung, W.; Park, P. J. *Phys. Rev. Lett.* **1996**, *77*, 783–786.
- (18) Lubensky, D.; Nelson, D. *Biophys. J.* **1999**, *77*, 1824–1838.
- (19) Muthukumar, M. *Phys. Rev. Lett.* **2001**, *86*, 3188–3191.
- (20) Tian, P.; Andricioaei, I. *J. Mol. Biol.* **2005**, *350*, 1017–1034.
- (21) Tian, P.; Andricioaei, I. *Biophys. J.* **2006**, *90*, 2718–2730.
- (22) Wells, D.; Abramkina, V.; Aksimentiev, A. *J. Chem. Phys.* **2007**, *127*, 125101.
- (23) Wintrobe, P.; Makhatadze, G.; Privalov, P. *Proteins: Struct., Funct., Genet* **1994**, *18*, 246–253.
- (24) Briggs, M.; Roder, H. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2017–2021.
- (25) Marianayagam, N.; Jackson, S. *Biophys. Chem.* **2004**, *111*, 159–171.
- (26) Huang, L.; Kirmizialtin, S.; Makarov, D. *J. Chem. Phys.* **2005**, *123*, 124903.
- (27) Matysiak, S.; Montesi, A.; Pasquali, M.; Kolomeisky, A.; Clementi, C. *Phys. Rev. Lett.* **2006**, *96*, 118103/1–118103/4.
- (28) Fyta, M.; Melchionna, S.; Succi, S.; Kaxiras, E. *Phys. Rev. E* **2008**, *78*, 036704/1–036704/7.
- (29) Kumar, S.; Rosenberg, J.; Bouzida, D.; Swendsen, R.; Kollman, P. *J. Comput. Chem.* **1992**, *13*, 1011–1021.
- (30) Souaille, M.; Roux, B. *Comput. Phys. Commun.* **2001**, *135*, 40–57.
- (31) Risken, H. *The Fokker-Planck equation. Methods of solution and applications*, 2nd ed.; Springer Series in Synergetics; Springer: Berlin, New York, 1989.
- (32) Redner, S. *A guide to first-passage processes*; Cambridge University Press: Cambridge, U.K., 2001.
- (33) Clementi, C.; Nymeyer, H.; Onuchic, J. *J. Mol. Biol.* **2000**, *298*, 937–953.
- (34) Miyazawa, S.; Jernigan, R. *J. Mol. Biol.* **1996**, *256*, 623–644.
- (35) Schlick, T. *Molecular Modeling and Simulation - An Interdisciplinary Guide*, 2nd ed.; Interdisciplinary applied mathematics mathematical biology; Springer: New York, 2002.
- (36) Ferrenberg, A.; Swendsen, R. *Phys. Rev. Lett.* **1989**, *63*, 1195–1198.
- (37) Berezhkovskii, A.; Pustovoit, M.; Bezrukov, S. *J. Chem. Phys.* **2002**, *116*, 9952–9956.
- (38) Berezhkovskii, A.; Gopich, I. *Biophys. J.* **2003**, *84*, 787–793.
- (39) Murl, A.; Rizzardi, M. *ACM Trans. Math. Software* **1990**, *16*, 158–168.
- (40) Metzler, R.; Klafter, J. *Biophys. J.* **2003**, *85*, 2776–2779.
- (41) Costantini, G.; Cecconi, F.; Marini-Bettolo-Marconi, U. *J. Chem. Phys.* **2006**, *125*, 204711.

JP900947F