Analytical and numerical studies of sequence dependence of passage times for translocation of heterobiopolymers through nanopores

R. H. Abdolvahab, Farinaz Roshani, Armita Nourmohammad, Muhammad Sahimi, and M. Reza Rahimi Tabar

Citation: J. Chem. Phys. **129**, 235102 (2008); doi: 10.1063/1.3040267
View online: http://dx.doi.org/10.1063/1.3040267
View Table of Contents: http://jcp.aip.org/resource/1/JCPSA6/v129/i23
Published by the American Institute of Physics.

Analytical and numerical studies of sequence dependence of passage times for translocation of heterobiopolymers through nanopores

R. H. Abdolvahab,[1] Farinaz Roshani,[2,3] Armita Nourmohammad,[4] Muhammad Sahimi,[5,a)] and M. Reza Rahimi Tabar[1,6]
[1]*Department of Physics, Sharif University of Technology, Tehran 11155-9161, Iran*
[2]*Department of Physics, Al-Zahra University, Tehran 19834, Iran*
[3]*Institute for Studies in Theoretical Physics and Mathematics, Tehran 19395, Iran*
[4]*Institut für Theoretische Physik, Universitat zu Köln, D-50937 Köln, Germany*
[5]*Mork Family Department of Chemical Engineering and Materials Science, University of Southern California, Los Angeles, California 90089-1211, USA*
[6]*Institute of Physics, Carl von Ossietzky University, D-26111 Oldenburg, Germany*

(Received 24 June 2008; accepted 10 November 2008; published online 18 December 2008)

We consider chaperone-assisted translocation of biopolymers with two distinct monomers or bases A and B, with the size of the chaperones being $\lambda\sigma$, where σ is a monomer's size. The probability that A and B are neighbors in the biopolymer is P_{AB}. A master equation is used, together with the detailed-balanced condition, in order to derive analytical results for the statistics of the first-passage times of the biopolymer as a function of P_{AB}, λ, and the biopolymer's configuration. Monte Carlo simulations have also been carried out in order to compute the same quantities for biopolymers with 100–900 monomers and several λ. The results indicate nontrivial dependence of the variance of the translocation times on the biopolymer's composition. It is also shown that measurements of the first two moments of the biopolymer's first-passage time distribution provide information on its length and ordering. Moreover, the probability density function $Q(t)$ of the first-passage times is almost Gaussian for small chaperone size λ, but becomes non-Gaussian as λ increases. At large times, $Q(t)$ decays exponentially. © *2008 American Institute of Physics*. [DOI: 10.1063/1.3040267]

I. INTRODUCTION

Transport of biological molecules, such as DNA, RNA, and proteins, across nanoporous membranes is of fundamental importance to life processes. Such molecules are typically in the form of long biopolymers, while the membranes' pores are small enough that they do not allow the biopolymers to pass through as a single unit, hence leading to translocation, i.e., squeezing of the biopolymers through the pores.[1] Translocation is used in gene therapy,[2] drug delivery,[3] and rapid DNA sequencing.[4] Other important examples include translocation of (i) RNA through the nucleus pore membrane;[5] (ii) DNA plasmid transport from cell to cell through the walls,[1] and (iii) the polypeptide chain from inner mitochondrial and chloroplast membrane through its matrix.[6]

Due to its complexity, translocation of biopolymers was first studied experimentally.[7] In particular, in a seminal paper by Kasianowicz *et al.*[8] the passage of one single-strand DNA through a protein channel, α-hemolysin, in a planar lipid bilayer was studied. The pore studied is asymmetric[9] and stable.[10] Kasianowicz *et al.*[8] demonstrated that, using an electric field, one can drive single-stranded DNA and RNA molecules through the pore, filled by water. The passage of each molecules is indicated by a blockade in the current through the pore, the magnitude and duration of which depend on the structure of the translocating molecule. More recent experiments used solid-state nanopores that offer better control on their size, and stability over changes in the voltage, salinity, and the *p*H. Motivated by its potential applications, many more experiments on translocation of polymers in small pores have been carried out.[11–13]

An experimentally accessible quantity of prime importance in such studies is[8,12] the typical time scale that a biopolymer spends in the nanopore—the so-called *dwell time*[14]—and its dependence on the molecular weight, the pore's length, and other parameters. In general, there are two distinct dynamics: in *slow* translocation the biopolymer remains equilibrated at almost all the times on *both sides* of the pore, whereas *fast* translocation allows the biopolymer to pass through the pore.

Several mechanisms may induce the translocation.[15] One is based on an external electric field applied across the membrane (pore) which, as mentioned above, has been studied by experiments *in vitro*.[8,11–13] Theoretically, it was first studied by determining the equilibrium entropy of the polymer as a function of the position through the pore.[14,16,17] Many other analytical techniques have been used to study the problem.[15,18–22] Extensive numerical simulations have also been employed.[19,23–25] The second mechanism uses a chemical-potential difference across the membrane, as there is usually no *in vivo* strong electric field in a membrane's pore,[26,27] an example of which is experiments on chaperone-assisted translocation of proteins.[1,28] The ejection of a biopolymer, such as DNA, from a bacterial virus to its host cell is the third mechanism of translocation of a polymer along its length through a pore or membrane. While diffusion

[a)]Author to whom correspondence should be addressed. Electronic mail: moe@iran.usc.edu.

might also give rise to such a motion, the most significant factor that contributes to such translocation motion is a pressure difference, and the corresponding stress, to which the biopolymer is subjected in the viral capside.[29] Yet another important factor that contributes to the translocation is the crowding effect,[30] which may be due to macromolecular aggregates and other inclusions in cytoplasmic cells, the volume fraction of which may be as high as 50%.[31] Another example of crowding is when a biopolymer, confined within a cavity, escapes if a narrow hole is introduced on the cavity.[32] Finally, it is known that the genomes of many double-stranded DNA (dsDNA) bacteriophages, as well as certain animal viruses, are packaged into a preformed protein precursor capsid.[33] The process is driven by a molecular motor, and the mechanical work that the motor does contributes to the translocation of the dsDNA.

Most biopolymers are not usually made of similar units. For example, DNA consists of repeated stacks of bases that are either adenine-thymine or guanine-cytosine pairs, coupled by hydrogen bonds and held together by a sugar-phosphate backbone. It has been shown that DNA sequences with different arrangements of nucleotide bases exhibit significant differences in their electronic properties[34] and breathing dynamics.[35] The question, then, is whether the composition also affects translocation. Various groups have recently begun to address this important question.

Muthukumar[16] used entropy-based arguments and a Fokker–Planck formulation to address the problem for a diblock copolymer. Slutsky *et al.*[36] used a MFPT formulation in order to study diffusion of an inhomogeneous polymer chain, while Kafri *et al.*[21] showed that anomalous dynamics arises due to sequence heterogeneity. Monte Carlo (MC) simulations were utilized by Romiszowski and Sikorski[37] to study polymer chains that were composed of hydrophilic and hydrophobic monomers. Luo *et al.*[35] used a combination of molecular dynamics simulation and the Langevin equation to study translocation of a biopolymer of the form A_mB_n $(m,n=1-3)$, and found that the heterogeneity does not affect the scaling of the average translocation time with the biopolymer's length. Furthermore, they showed that for symmetric polymers, $m=n$, the translocation times also depend on the *orientation* of the base which enters the pore first. Gauthier and Slater[38] used a MC simulation method, that they had developed earlier,[25] in order to study the translocation times of a copolymer in a pore under an external electric charge. The copolymer consisted of two types of monomers that differed only in terms of their electric charges. They found that each sequence leads to a unique value of the translocation probability and time.

In this paper, we consider the general case of a biopolymer of two distinct monomers or bases, A and B, and study its slow translocation through a nanopore, both analytically and by MC simulation. We consider chaperone-assisted translocation of biopolymers, and using a master equation together with detailed-balanced condition, as well as MC simulation, to study the problem. Analytical results for certain biopolymer's configuration, as well as numerical values for biopolymers with 100–900 monomers are presented.

The rest of this paper is organized as follows. In Sec. II the biopolymer's model that we use is described. Section III presents the analytical formulation of the problem, while Sec. IV describes the MC approach that we utilize in our study. The results are presented and discussed in Sec. V. This paper is summarized in Sec. VI, where we also describe some possible future directions in for further studying the problem.

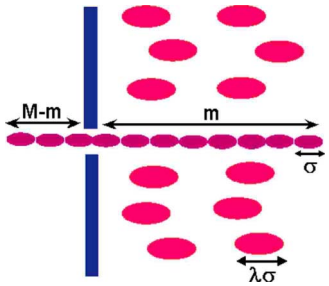

FIG. 1. (Color online) A translocating biopolymer with binding proteins of size $\lambda\sigma$.

II. THE MODEL

We consider translocation of a stiff heterobiopolymer chain with M monomers and length $L=M\sigma$ thorough a nanopore, where σ is the monomer's length. The bases, A and B, are neighbors with given probabilities. For example, the configurations AA, BB, and AB occur with the probabilities P_{AA}, P_{BB}, and P_{AB}, respectively. On the pore's right side there are small chaperones of length $\lambda\sigma$ and density c_0, the role of which is to prevent backsliding of the biopolymer, when it is in contact with the pore's walls.[15] This is shown schematically in Fig. 1. Thus, a chemical-potential difference is induced by the chaperones. The difference between the various bases is in the binding energies of a chaperone to them.

Finite-size effects are important. If the biopolymer is infinitely long, then, specifying P_{AB} should suffice for constructing the biopolymer. However, for finite polymers one must be careful about how to construct an ensemble of such polymers. For example, if $P_{AB}=0$, even a finite polymer made of only A or B may satisfy this condition. Thus, in order to generate the correct ensemble, one should first generate an infinitely long biopolymer, and then cuts it into pieces of the desired length L. In this way, an ensemble of the biopolymers with given values of P_{AB} and length L is generated. This is the model that we analyze in the present paper.

As mentioned above, Luo *et al.*[35] studied a biopolymer with a *fixed* sequence A_mB_n, and computed the mean first-passage times (MFPTs) as a function of the fraction A/B. In contrast, we consider the general case that the *probability* of A and B being neighbors is P_{AB} (biopolymers with random structure), and compute the statistics of the MFPTs as a function of P_{AB}. As we demonstrate below, the statistics are fundamental to sequencing a heterobiopolymer of *a priori* unknown structure. In addition, we consider *chaperone-assisted* translocation which, to our knowledge, has never been studied for the heteropolymers. Moreover, there has also never been any investigation of the entire statistics of the MFPT and its probability density function (PDF) in terms of λ.

These issues are all studied in the present paper for biopolymer with two distinct monomers. In particular, we show that the PDF of the FPTs becomes increasingly non-Gaussian with increasing λ.

III. ANALYTICAL ANALYSIS: FIRST-PASSAGE TIME FORMULATION

Analytical investigation of translocation—a nonequilibrium process—is very difficult. For some special ordering of the bases, we use the master equation[39] together with the detailed-balance condition[40] (DBC), in order to describe the phenomenon, and determine the statistics of the FPT in terms of the biopolymer's transition probabilities with various sequences of the monomers. Since translocation is a nonequilibrium process, the DBC is an approximation. Nevertheless, it leads to some very useful and insightful results. Analytical results are derived for $\lambda=2$ and arbitrary P_{AB}. In Sec. IV we use a MC method to compute the statistics of the FPT of a biopolymer of size M and a given λ.

To begin with, we note that, compared to translocation, the binding process is fast. The ratio of the time scales for the two processes is[41] about 1/300. It is, therefore, reasonable to assume that the fast variable approaches rapidly its equilibrium distribution. Therefore, as Zandi *et al.*[27] pointed out, the entire process is described by the evolution of the slow variable, which is the number m of the bases or monomers that pass through the nanopore. The master equation for $P(m,t)$, the probability that m bases have passed through the nanopore at time t, is given by

$$\frac{\partial P(m,t)}{\partial t} = W^{+}(m-1)P(m-1,t) + W^{-}(m+1)P(m+1,t) - [W^{+}(m) + W^{-}(m)]P(m,t), \tag{1}$$

where $W^{+}(m)$ and $W^{-}(m)$ are, respectively, the transition rates for forward and backward motions, which are evaluated using the DBC, according to which

$$W^{-}(m)Z_{\lambda}(m) = W^{+}(m-1)Z_{\lambda}(m-1). \tag{2}$$

The stationary solution of the problem is then given by

$$P(m) = \frac{Z_{\lambda}(m)}{\sum_{m=1}^{M} Z_{\lambda}(m)}, \tag{3}$$

Here, Z_{λ} is the partition function.[42]

Since, as described above, the chaperones only induce a chemical-potential difference along the nanopore, they do not affect the forward transition rate. Therefore, $W^{+}(m)=k$, where k is a constant. One also has

$$W^{-}(m) = k\frac{Z_{\lambda}(m-1)}{Z_{\lambda}(m)}. \tag{4}$$

The MFPT T is then given by[42]

FIG. 2. (Color online) The structure of the biopolymer used in the study.

$$T = \sum_{m=0}^{M}\left[\Phi(m)\sum_{m'=0}^{m}\frac{1}{W^{+}(m')\Phi(m')}\right], \qquad \Phi(m) = \prod_{u=1}^{m}\left[\frac{W^{-}(u)}{W^{+}(u)}\right]. \tag{5}$$

Thus, if the partition function is computed, Eqs. (4) and (5) enable one to calculate the MFPT T. The partition function Z_{λ} should account for three important factors.

(i) The first factor is

$$\Omega_{\lambda}(m,n) = \frac{[m-(\lambda-1)n]!}{n!\,(m-\lambda n)!}, \tag{6}$$

which is the total number of ways that n chaperones of length λ are distributed on m sites. The chaperons' length is measured in units of the monomers.

(ii) The equilibrium Boltzmann weight, $\exp(-\varepsilon/k_BT)$, where ε is the chaperones' binding energy with the monomers, T the temperature, and k_B the Boltzmann's constant, is the second factor.

(iii) The third factor, $Z_{\text{space}}(n)$, accounts for the effect of decreasing the number of chaperones through a cell,[6,43] with

$$Z_{\text{space}}(n) = \frac{N!\,[(N_t-m)-N]!}{(N-n)!\,[(N_t-m)-(N-n)]!}, \tag{7}$$

where N is the total number of the chaperones and, $N_t=V/v_0$. Here, V is the volume of the cell on the right side of the nanopore, and v_0 is the volume of a chaperone. The factor m in Eq. (7) accounts for the space that the biopolymer occupies. This effect is negligible when N_t is very large, which is also the limit that we consider. In the limit, $N_t \gg N$, $Z_{\text{space}}(n) \to (N/N_t)^n$.[44] Moreover, we also have $c_0=Nv_0/V$.

The translocation times depend not only on the biopolymer's length but also on the monomers' ordering. If the biopolymer consists of two bases, A and B, with the corresponding configuration probabilities P_{AA}, P_{BB}, and P_{AB}, the binding energies are ε_{AA}, ε_{BB}, and ε_{AB}. Consider, first, the case, $\lambda=2$, so that the size of the chaperones is $\lambda\sigma=2\sigma$ and assume, for simplicity, equal numbers of A and B monomers, $M_A=M_B$. Analytical determination of the partition function Z_{λ} for a general ordering of A and B is not possible. One can, however, determine its exact form for some special cases. Consider, for example, a biopolymer with M_{AA} units

of $AAA\cdots$. Then, there exist $\frac{1}{2}(M_{AB}+1)$ of $ABAB\cdots$, and $M_{BB}=M-(M_{AA}+M_{AB})$ of $BBB\cdots$ in the biopolymer; see Fig. 2. Here,

$$M_{AA}=M_{BB}\simeq\tfrac{1}{2}(1-P_{AB})(M-1), \tag{8}$$

$$M_{AB}\simeq P_{AB}(M-1). \tag{9}$$

If, $p=P_{AB}$, and $r=M_{AB}+M_{BB}$, then

$$Z_\lambda(m;p)=\begin{cases}\sum_n Z_\lambda^{(1)}(m,n;p), & \text{if } M_{BB}>m,\\ \sum_n Z_\lambda^{(2)}(m,n;p), & \text{if } M_{BB}<m<M_{BB}+M_{AB},\\ \sum_n Z_\lambda^{(3)}(m,n;p), & \text{if } M_{BB}+M_{AB}<m,\end{cases} \tag{10}$$

where

$$Z_\lambda^{(1)}(m,n;p)=\Omega_\lambda(m,n)(\chi_B)^{2n}, \tag{11}$$

$$\begin{aligned}Z_\lambda^{(2)}(m,n;)=&\left\{\sum_{n_1}[\Omega_\lambda(M_{BB},n_1)(\chi_B)^{2n_1}]\right.\\&\left.\times[\Omega_\lambda(m-M_{BB}-2,n-1-n_1)(\chi_A\chi_B)^{n-n_1}]\right\}\\&\times\left\{\sum_{n_1}[\Omega_\lambda(M_{BB}+1,n_1)(\chi_B)^{2n_1}]\right.\\&\left.\times[\Omega_\lambda(m-M_{BB}-1,n-n_1)(\chi_A\chi_B)^{n-n_1}]\right\},\end{aligned} \tag{12}$$

$$\begin{aligned}Z_\lambda^{(3)}(m,n;p)=&\sum_{n_1}\{[Z_\lambda^{(1)}(r+1,n_1;p)+Z_\lambda^{(2)}(r+1,n_1;p)]\\&\times\{\Omega_\lambda[m-(r+1),n-n_1](\chi_A)^{2(n-n_1)}\}\}\\&+\sum_{n_1}\{[Z_\lambda^{(1)}(r,n_1,p)+Z_\lambda^{(2)}(r,n_1;p)]\\&\times\{\Omega_\lambda[m-(r+1)-1,n-n_1-1]\\&\times(\chi_A)^{2(n-n_1)}\}\}.\end{aligned} \tag{13}$$

Here, $\chi_i=\sqrt{c_0\upsilon_0}\exp(-\varepsilon_i/k_BT)$, and, $i=A$ and B. Having determined the partition function, the MFPT T is computed *numerically* through Eq. (5).

IV. MONTE CARLO SIMULATIONS

To compute T, the MFPT, by numerical simulations, we used the Metropolis algorithm, using the typical values, $\chi_A=0.42$ and $\chi_B=1.15$.[6,40] The biopolymer is moved to the right and left with equal probabilities. When, however, there is a vacant monomer next to the nanopore, the biopolymer moves to the left side (in contrast with the case in which it can go to the right without any restriction). In other words, if a chaperone is bound near the pore, it rectifies the polymer's motion to the right.

To obtain reliable statistics, we generated 10^4 ensembles for the Metropolis move (see Sec. II about how the ensemble was constructed). As already mentioned above, it takes the biopolymer about $\tau_D\sim 5$ s to translocate, which is much longer than the diffusion time for the chaperones, $\tau_c\sim 1$ ms.[15] This implies that the chaperones have plenty of time to relax. In order to obtain better statistics with higher efficiency, we also built biopolymers with the *same* probability P_{AB}, but with various sequences (configurations), and computed the maximum and minimum translocation times.

V. RESULTS AND DISCUSSIONS

We carried out extensive MC simulations, in order to understand the statistics of the FPTs, and compare them to the analytical results. In what follows we present the results and describe their implications.

A. The statistics of the first-passage times

We compare in Fig. 3 the theoretical MFPTs for the ordering shown in Fig. 2 with the results of the MC simulations for biopolymers with $M=100$. Figure 3 indicates that the computed MFPT T is always smaller than those obtained from the analytical analysis. The difference is due to the fact that, under the DBC, one finds that T scales with the size of the biopolymer as

$$T=C_1M, \tag{14}$$

whereas from the simulations we find that

$$T=C_2M^\nu, \tag{15}$$

with $\nu\simeq 1.05\pm 0.01$, and $C_1/C_2\simeq 1.2\pm 0.1$.

We also carried out MC simulation of translocation of biopolymers with chaperon size (in units of the monomers' size σ) $\lambda=2$, 5, and 10, with $M=100-900$ monomers, and a given P_{AB}, making no assumption about the ordering of the monomers A and B. In Figs. 4 and 5 we show the dependence of the MFPT T and its standard deviations S on M, for $\lambda=2$ and various P_{AB}. One has

$$T\sim M^{1.05}. \tag{16}$$

Therefore, although the exponent ν that relates T to M is essentially independent of the probability P_{AB}, the coefficients C_1 and C_2 in Eqs. (14) and (15), as well as the implied amplitude in Eq. (16), are not and depend, in general, on the probability P_{AB}. Equation (16) indicates that the presence of the chaperones increases the speed of the biopolymer inside the cell, as ν decreases from 2 for random walk-type motion[15] to nearly 1. This is attributed to the fact that due to the chaperons that generate a net nonzero chemical-potential difference along the pore, diffusion of the biopolymer is biased toward the right side of the pore. Moreover, the almost linear dependence of T on M also implies that there are no strong fluctuations in T when measured as a function of M; see Fig. 4. In other words, the series T versus M is essentially continuous. As for the standard deviations S of the FPTs, we find that

$$S\sim M^\mu, \tag{17}$$

with $\mu\simeq 0.65\pm 0.02$, which is a nontrivial result.

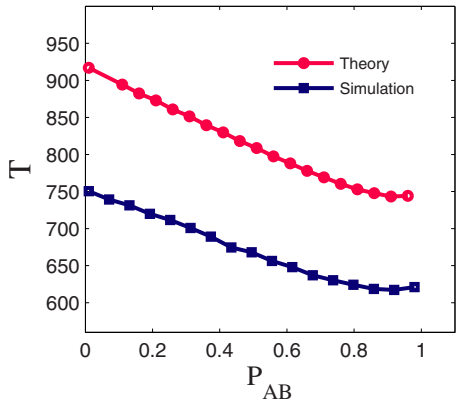

FIG. 3. (Color online) Comparison of the analytical predictions with the results of numerical simulations for a polymer with $M=100$ monomers and $\lambda=2$.

An important question is whether the deviation from unity of the exponent ν in Eq. (15) is physical, or that it is merely numerical and larger-scale simulations should yield $\nu=1$ exactly. We believe that it is the former case. To see this, consider the following. If we assume that the two transition rates $W^+(m)$ and $W^-(m)$ are independent of m, Eq. (5) yields

$$\alpha^m \equiv \Phi(m) = \prod_{u=1}^{m} \left[\frac{W^-(u)}{W^+(u)} \right] = \left(\frac{W^-}{W^+} \right)^m, \tag{18}$$

and, therefore, the MFPT T is given by

$$T = \sum_{m=0}^{M} \left[\Phi(m) \sum_{m'=0}^{m} \frac{1}{W^+(m')\Phi(m')} \right] = \frac{1}{W^+} \sum_{m=0}^{M} \left(\frac{1-\alpha^m}{1-\alpha} \right)$$

$$= \left(\frac{1}{W^+ - W^-} \right) \left[(M+1) - \alpha \left(\frac{1-\alpha^{M+1}}{1-\alpha} \right) \right]. \tag{19}$$

But, since $\alpha<1$, we obtain

$$T = \left(\frac{1}{W^+ - W^-} \right) \left(M + 1 - \frac{\alpha}{1-\alpha} \right), \tag{20}$$

so that in the limit, $M \gg 1$, we obtain $T \sim M$, i.e., one obtains

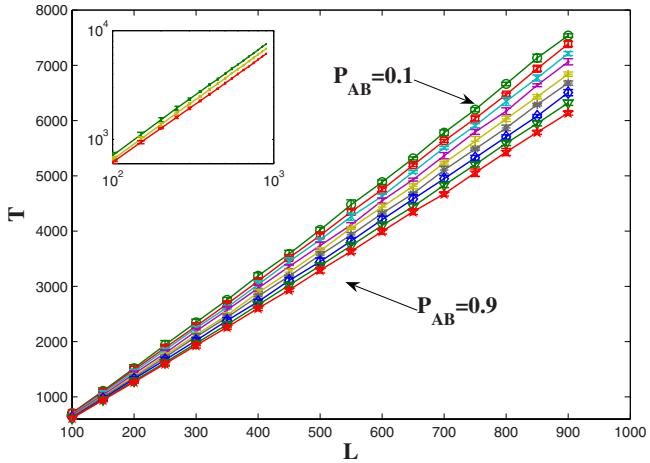

FIG. 4. (Color online) Dependence of the MFPT T on the biopolymer length and the ordering probability P_{AB}. The inset displays the same data for $P_{AB}=0.9$, 0.5 and 0.1 in a logarithmic plot.

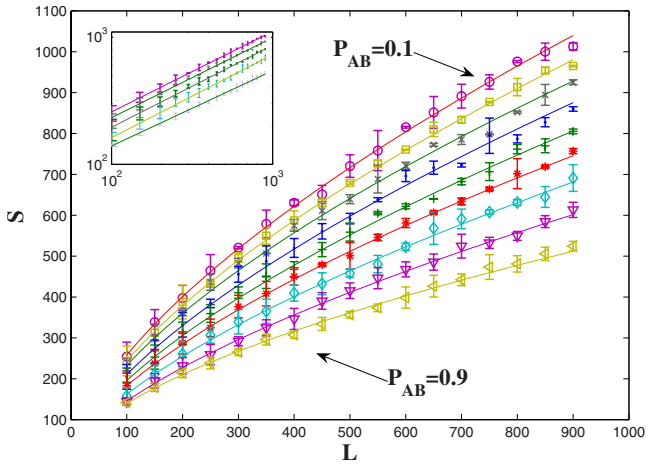

FIG. 5. (Color online) Dependence of the standard deviations S on the biopolymer length and the ordering probability P_{AB}. The inset displays the same data for $P_{AB}=0.9$, 0.7, 0.5, 0.3, and 0.1 in a logarithmic plot.

$\nu=1$ *only* when the transition rates are constant and independent of m. However, both in the analytical treatment and the numerical simulations, the transition rates are *not* independent of m and are, therefore, correlated. Such correlations are the reason for $\nu>1$.

While the above results cannot be directly compared to the previous ones reported for translocation of biopolymers made of only one type of monomers, it might be instructive to summarize the previous results. Specifically, for *forced* translocation, one has[23] a crossover in the scaling from

$$T \sim M^{2\nu_F} \sim M^{1.18}, \tag{21}$$

to

$$T \sim M^{1+\nu_F} \sim M^{1.59}, \tag{22}$$

where ν_F the Flory exponent of the polymer which, as is well known, arises due to the excluded-volume effect. Thus, the linear polymer or the self-avoiding walk used in these works is different from the stiff biopolymer that we consider in this paper. For *free* translocation, on the other hand, one finds that[23]

$$T \sim M^{1+2\nu_F} \sim M^{2.18}. \tag{23}$$

Note that the Flory exponent is given by, $\nu_F \simeq 0.588$. In the problem that we study, the presence of the chaperones gives rise to a net chemical-potential difference, hence a net force. Therefore, the question is notwithstanding the differences between our model and the previous ones with one type of monomers, which of the above scaling regimes might correspond to what we study in this paper?

In the presence of a *small* external force F, the biopolymer can move to the right *and* left, although with some drift to the right. This would then be similar to a random walk with $p \neq q$, where p and q are the transition probabilities to move to the right and left.[45] In some sense, this is similar to what we have studied in that the biopolymer just *waits* (instead of moving to the left) and then moves to the right because, if the chaperons allows the polymer, it moves inside the pore; otherwise, the biopolymer waits up to a waiting time t_w until the possibility of moving into the pore is favor-

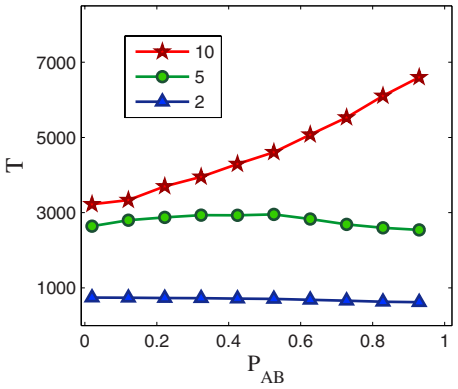

FIG. 6. (Color online) Dependence of the MFPT T on the biopolymer length and the ordering probability P_{AB}. The error bars represent one standard deviation.

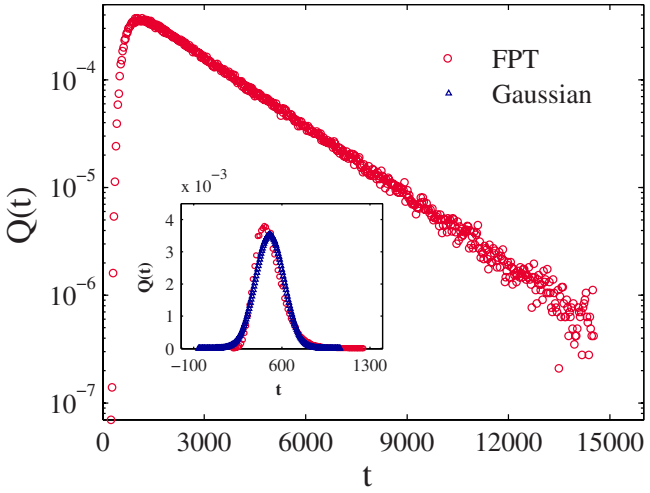

FIG. 7. (Color online) The PDF $P(t)$ of the FPTs for a biopolymer with ordering probability $P_{AB}=1/(M-1)$, shaperon size (in units of monomers) $\lambda=5$ ($\lambda=2$ in the inset), and $M=100$ monomers, and its comparison with a Gaussian distribution which was computed based on the mean value T and the standard deviations S.

able. Waiting for the time t_w is exactly equivalent to moving to the left and going back to where the motion to the left had commenced, over a period of time exactly equal to the waiting time t_w. Thus, the problem that we study is somewhat similar to the case, studied previously by others,[23] in which a *weak* external force F was applied to the linear polymer made of only one type of monomers. Indeed, the scaling law that we obtain, Eq. (16), is close to Eq. (21).

The two problems are not, of course, completely similar, because in the problem that we study the biopolymers is stiff and heterogeneous, which influences the scaling of the MFPT with the polymer size M. In addition, the chaperone-assisted translocation is not exactly identical with one under the influence of an external force. Due to such differences, the apparent closeness of Eqs. (15), (16), and (21) is intriguing and, therefore, this is a problem that must be studied further.

For a given number of monomers M, large values of the ordering probability P_{AB} result in small MFPTs. Moreover, as Fig. 5 indicates, the standard deviations S depends strongly on the ordering probability P_{AB}, implying that, to estimate P_{AB} experimentally, it is the standard deviations S that yields insight into the ordering of the monomers in the biopolymer, rather than the MFPT itself. In addition, we find that the third and higher moments of the FPTs do not yield any sharper bounds on the value of the ordering probability P_{AB} (see also below).

Figure 6 presents the average of the MFPTs for biopolymers with $M=100$ monomers and $\lambda=2$, 5, and 10. The results suggest that the mean translocation time T is basically insensitive to the biopolymer's structure (i.e., to the value of P_{AB}) for small chaperones, $\lambda \leq 5$.

We can estimate the ordering probability P_{AB} using the statistics of T and the standard deviations S. Suppose, for example, that measurements yield, $T=2000$ and $S=300$ (with proper units). Starting with Fig. 4, we estimate the biopolymer's length (in units of the monomers' size σ) to be between 260 and 320. Using the estimated length and Fig. 5, one finds that, $0.6 \leq P_{AB} \leq 0.8$. Then, utilizing the estimate of P_{AB} and Fig. 4 again, one obtains the tighter bounds, $0.7 < P_{AB} < 0.8$, and the length of the biopolymer to be between 290 and 320. After a few iterations between Figs. 4 and 5, one finds that the length of the biopolymer is between 300 and 315 with an ordering probability $0.7 \leq P_{AB} \leq 0.8$.

B. The probability density function of the first-passage times

Figure 7 presents the computed $Q(t)$, the PDF of the FPTs, for the chaperone sizes $\lambda=2$ and 5 (in units of the monomers), computed by the MC simulations. The results indicate that, while for small values of λ the PDF $Q(t)$ is almost Guassian and symmetric, the skewness of $Q(t)$ increases with λ, and becomes non-Gaussian. Long tails develop that are the result of the biopolymer waiting for a favorable condition to move to the right side. In fact, as Fig. 7 indicates, the PDF $Q(t)$ for $\lambda=5$ decays exponentially fast with t.

In a recent paper, Chatelain *et al.*[46] presented numerical results for $Q(t)$ in another model of translocation, in which a self-avoiding polymer (made of the same monomers) translocates through a membrane pore. Theirs is a model in which M monomers are restricted to the sites of a square lattice with the excluded-volume interactions taken into account. A randomly selected monomer is moved by one lattice spacing in an arbitrarily chosen direction. If the new configuration of the polymer is allowed, the step is taken; otherwise, the polymer's configuration remains intact. The thickness of the membrane is two lattice spacing, with the size of its pore being three lattice spacing. Chatelain *et al.*[46] found that for large t,

$$Q(t) \sim \exp(-t/t_0), \tag{24}$$

where t_0 is a decay constant. Equation (21) is in agreement with the results shown in Fig. 7 for $\lambda=5$, although our model and that of Chatelain *et al.*[46] are very different. The similarity indicates that the exponential dependence of $Q(t)$ on t may be a generic feature of first-passage time distributions for such problems. Indeed, exponential decay of the PDF of the first-passage times is a general feature of many random

processes,[45] although power-law decay of such PDFs is also quite common.[47]

VI. SUMMARY

The effect of the ordering of distinct monomers or bases of a biopolymer on the statistics of its translocation time was studied. For certain types of the ordering, a master equation approach together with the DBC was used in order to derive analytical results for the statistics of the first-passage times. Monte Carlo simulations were also utilized for more general orderings of the biopolymer. Knowledge of the first two moments of the first-passage times enables one to obtain information about the ordering probability—the probability that the two distinct monomers of the biopolymer are neighbors—and the biopolymer's length. It was also shown that the PDF $Q(t)$ of the first-passage times is almost Gaussian for small chaperones, but that the non-Gaussianity increases rapidly with the size of the binding chaperones. In fact, $Q(t)$ decays exponentially for large t.

In order to test our results by experiments, one must prepare biopolymers of different sequences, made of the same bases or monomers, and measure their translocation times through a nanopore, similar to the early experiments on translocation of homopolymers.[7,8] If our results are confirmed, as we expect to, then, experimental studies of translocation of biopolymers and their application must take into account important effect of the biopolymer's configuration and its ordering structure.

[1] B. Alberts, D. Bray, J. Lewis *et al.*, *Molecular Biology of the Cell* (Garland, New York, 2002).
[2] J. P. Henry, J. F. Chich, D. Goldschmidt, and M. Thieffry, J. Membr. Biol. **112**, 139 (1989); J. Akimaru, S. Matsuyama, H. Tokuda, and S. Mizushima, Proc. Natl. Acad. Sci. U.S.A. **88**, 6545 (1991); E. Di Marzio and J. J. Kasianowicz, J. Chem. Phys. **119**, 6378 (2003).
[3] Y.-L. Tseng, J.-J. Liu, and R.-L. Hong, Mol. Pharmacol. **62**, 864 (2002); J. M. Tsutsui, F. Xie, and R. T. Porter, Cardiovasc. Ultrasound **2**, 23 (2004).
[4] J. Han, S. W. Turner, and G. H. Craighead, Phys. Rev. Lett. **83**, 1688 (1999); S. W. P. Turner, M. Cabodi, and H. G. Craighead, *ibid.* **88**, 128103 (2002).
[5] M. Bukrinsky, Mol. Med. **10**, 1 (2004).
[6] W. Liebermeister, T. A. Rapoport, and R. Heinrich, J. Mol. Biol. **305**, 643 (2001).
[7] W. T. Wickner and H. F. Lodish, Science **230**, 400 (1985); K. Verner and G. Schatz, *ibid.* **241**, 1307 (1988); D. Georlich and I. W. Mattaj, *ibid.* **271**, 1513 (1991).
[8] J. J. Kasianowicz, E. Brandin, D. Branton, and D. W. Deamer, Proc. Natl. Acad. Sci. U.S.A. **93**, 13770 (1996).
[9] L. Song, M. R. Hobaugh, C. Shustak, S. Cheley, H. Barley, and J. E. Gouaux, Science **274**, 1859 (1996).
[10] Y. Jung, H. Bayley, and L. Movileanu, J. Am. Chem. Soc. **128**, 15332 (2006); S. M. Bezrukov and J. J. Kasianowicz, Eur. Biophys. J. **26**, 471 (1997); M. Pastoriza-Gallego, G. Oukhaled, J. Mathé, B. Thiébot, J. M. Bretton, L. Auvray, and J. Pelta, FEBS Lett. **581**, 3371 (2007).
[11] M. Akeson, D. Branton, J. J. Kasianowicz, E. Brandin, and D. W. Deamer, Biophys. J. **77**, 3227 (1999); A. Aksimentiev, J. B. Heng, G. Timp, and K. Schulten, *ibid.* **87**, 2086 (2004); S. E. Henrickson, M. Misakian, B. Robertson, and J. J. Kasianowicz, Phys. Rev. Lett. **85**, 3057 (2000); A. Meller, L. Nivon, and D. Branton, *ibid.* **86**, 3435 (2001).
[12] A. Meller, L. Nivon, E. Brandin, J. A. Golovchenko, and D. Branton, Proc. Natl. Acad. Sci. U.S.A. **97**, 1079 (2000).
[13] A. Meller and D. Branton, Electrophoresis **23**, 2583 (2002); A. F. Sauer-Budge, J. A. Nyamwanda, D. K. Lubensky, and D. Branton, Phys. Rev. Lett. **90**, 238101 (2003); J. Mathé, H. Visram, V. Visnoff, Y. Rubin, and A. Meller, Biophys. J. **87**, 3205 (2004).
[14] W. Sung and P. J. Park, Phys. Rev. Lett. **77**, 783 (1996); M. Muthukumar, J. Chem. Phys. **111**, 10371 (1999); M. Muthukumar, *ibid.* **118**, 5174 (2003).
[15] J.-F. Chauwin, G. Oster, and B. S. Glick, Biophys. J. **74**, 1732 (1998); T. C. Elston, *ibid.* **82**, 1239 (2002); T. Ambjörnsson and R. Metzler, Phys. Biol. **1**, 77 (2004); T. Ambjörnsson, M. A. Lomholt, and R. Metzler, J. Phys.: Condens. Matter **17**, S3945 (2005).
[16] M. Muthukumar, Electrophoresis **23**, 1417 (2002).
[17] B. Slonkina and A. B. Kolomeisky, J. Chem. Phys. **118**, 7112 (2003).
[18] A. J. Storm, C. Storm, J. Chen, H. Zandbergen, J.-F. Joanny, and C. Dekker, Nano Lett. **5**, 1193 (2005); P. J. Park and W. Sung, J. Chem. Phys. **108**, 3013 (1998); D. K. Lubensky and D. R. Nelson, Biophys. J. **77**, 1824 (1999).
[19] J. Chuang, Y. Kantor, and M. Kardar, Phys. Rev. E **65**, 011802 (2001); Y. Kantor and M. Kardar, *ibid.* **69**, 021806 (2004); A. Milchev, K. Binder, and A. Bhattacharya, J. Chem. Phys. **121**, 6042 (2004); A. Corsi, A. Milchev, V. G. Rostiashvili, and T. A. Vilgis, Macromolecules **39**, 7115 (2006).
[20] A. Baumgärtner and J. Skolnick, Phys. Rev. Lett. **74**, 2142 (1995); R. Bundschuh and U. Gerland, *ibid.* **95**, 208104 (2006); S. Matysiak, A. Montesi, M. Pasquali, A. B. Kolomeisky, and C. Clementi, *ibid.* **96**, 118103 (2006).
[21] S. F. Simon, C. S. Peskin, and G. F. Oster, Proc. Natl. Acad. Sci. U.S.A. **89**, 3770 (1992); A. Meller, J. Phys.: Condens. Matter **15**, R581 (2003); Y. Kafri, D. K. Lubensky, and D. R. Nelson, Biophys. J. **86**, 3373 (2004).
[22] T. Ambjörnsson, S. P. Apell, Z. Konkoli, E. A. DiMarzio, and J. J. Kasainowicz, J. Chem. Phys. **117**, 4063 (2002); C. T. A. Wong and M. Muthukumar, *ibid.* **128**, 154903 (2008); R. Metzler and J. Klafter, Biophys. J. **85**, 2776 (2003); U. Gerland, R. Bundschuh, and T. Hwa, Phys. Biol. **1**, 19 (2004).
[23] K. F. Luo, T. Ala-Nissila, and S.-Ch. Ying, J. Chem. Phys. **124**, 034714 (2006); K. F. Luo, I. Huopaniemi, T. Ala-Nissila, and S.-Ch. Ying, *ibid.* **124**, 114704 (2006); I. Huopaniemi, K. F. Luo, T. Ala-Nissila, and S.-Ch. Ying, *ibid.* **125**, 124901 (2006); I. Huopaniemi, K. F. Luo, T. Ala-Nissila, and S.-Ch. Ying, Phys. Rev. E **75**, 061912 (2007).
[24] C. Y. Kong and M. Muthukumar, Electrophoresis **23**, 2697 (2002); C. Y. Kong and M. Muthukumar, J. Am. Chem. Soc. **127**, 18252 (2005); M. Muthukumar and C. Y. Kong, Proc. Natl. Acad. Sci. U.S.A. **103**, 5273 (2006).
[25] S.-S. Chern, A. E. Cardenas, and R. D. Coalson, J. Chem. Phys. **115**, 7772 (2001); M. G. Gauthier and G. W. Slater, *ibid.* **128**, 065103 (2008); H. C. Loebl, R. Randel, S. P. Goodwin, and C. C. Mathai, Phys. Rev. E **67**, 041913 (2003).
[26] H. Salman, D. Zbaida, Y. Rabin, D. Chatenay, and M. Elbaum, Proc. Natl. Acad. Sci. U.S.A. **98**, 7247 (2001); Z. Farkas, I. Derenyi, and T. Vicsek, J. Phys.: Condens. Matter **15**, S1767 (2003); J. K. Wolterink, G. T. Barkema, and D. Panja, Phys. Rev. Lett. **96**, 208301 (2006).
[27] R. Zandi, D. Reguera, J. Rudnick, and W. M. Gelbart, Proc. Natl. Acad. Sci. U.S.A. **100**, 8649 (2003).
[28] D. Klotsa, R. A. Romer, and M. S. Turner, Biophys. J. **89**, 2187 (2005); A. Rodriguez, R. A. Römer, and M. S. Turner, Phys. Status Solidi B **243**, 373 (2006).
[29] M. M. Inamdar, W. M. Gelbart, and R. Phillips, Biophys. J. **91**, 411 (2006).
[30] A. Gopinathan and Y. W. Kim, Phys. Rev. Lett. **99**, 228106 (2007).
[31] A. B. Fulton, Cell **30**, 345 (1982).
[32] M. Muthukumar, Phys. Rev. Lett. **86**, 3188 (2001).
[33] Y. R. Chemla, K. Aathavan, J. Michaelis, S. Grimes, P. J. Jardine, D. W. Anderson, and C. Bustamante, Cell **122**, 683 (2005); J. P. Rickgauer, D. N. Fuller, S. Grimes, P. J. Jardine, D. L. Anderson, and D. K. Smith, Biophys. J. **94**, 159 (2008).
[34] T. Ambjörnsson, S. K. Banik, O. Krichevsky, and R. Metzler, Phys. Rev. Lett. **97**, 128105 (2006).
[35] K. Luo, T. Ala-Nissila, S.-Ch. Ying, and A. Bhattacharya, J. Chem. Phys. **126**, 145101 (2007); K. Luo, T. Ala-Nissila, S.-Ch. Ying, and A. Bhattacharya, Phys. Rev. Lett. **100**, 058101 (2008).
[36] M. Slutsky, M. Kardar, and L. A. Mirny, Phys. Rev. E **69**, 061903 (2004).
[37] P. Romiszowski and A. Sikorski, Comput. Mater. Sci. **38**, 533 (2007).
[38] M. G. Gauthier and G. W. Slater, J. Chem. Phys. **128**, 175103 (2008).
[39] O. Flomenbom and J. Klafter, Phys. Rev. E **68**, 041910 (2003); O. Flomenbom and J. Klafter, Biophys. J. **86**, 3576 (2004).
[40] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry*

(North-Holland, Amsterdam, 1992); T. Ambjörnsson and R. Metzler, J. Phys.: Condens. Matter **17**, S1841 (2005).
[41] U. Gerland, J. D. Moroz, and T. Hwa, Proc. Natl. Acad. Sci. U.S.A. **99**, 12015 (2002).
[42] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences* (Springer, New York, 2002).
[43] I. R. Epstein, Biophys. Chem. **8**, 327 (1978).
[44] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).
[45] B. D. Hughes, *Random Walks and Random Environments* (Oxford University Press, London, 1995), Vol. 1.
[46] C. Chatelain, Y. Kantor, and M. Kardar, Phys. Rev. E **78**, 021129 (2008).
[47] S. Redner, *A Guide to First-Passage Time Processes* (Cambridge University Press, London, 2001).