Sequence dependence of the binding energy in chaperone-driven polymer translocation through a nanopore

Rouhollah Haji Abdolvahab,[1,2] Mohammad Reza Ejtehadi,[1] and Ralf Metzler[2]

[1]*Physics Department, Sharif University of Technology, P.O. Box 11155-9161, Tehran, Iran*
[2]*Physics Department, Technical University of Munich, James Franck Straße, 85747 Garching, Germany*

We study the translocation of stiff polymers through a nanopore, driven by the chemical-potential gradient exerted by binding proteins (chaperones) on the *trans* side of the pore. Bound chaperones prevent backsliding through the pore and, therefore, partially rectify the polymer passage. We show that the sequence of chain monomers with different binding affinity for the chaperones significantly affects the translocation dynamics. In particular, we investigate the effect of the nearest-neighbor adjacency probability of the two monomer types. Depending on the magnitude of the involved binding energies, the translocation speed may either increase or decrease with the adjacency probability. We determine the mean first passage time and show that, by tuning the effective binding energy, the motion changes continuously from purely diffusive to ballistic translocation.

PACS number(s): 87.15.A−, 82.35.Lr, 05.40.−a, 02.50.−r

I. INTRODUCTION

The passage of polymeric chains through a nanopore in a membrane is of fundamental importance in biology and nanotechnology [1–4]. In a biological lipid bilayer membrane, the nanopore is typically constituted by a channel protein. A prominent example are the transmembrane proteins of the haemolysin family used by bacteria to puncture or even disrupt the red blood cells of the infected host [1,5]. α-haemolysin is, in fact, commonly used in single nanopore setups for *in vitro* translocation experiments [6]. Conversely, in artificial supports such as silicon compound membranes, solid-state nanopores can be created by ion- or electron-beam techniques [7,8]. Some important examples for such *translocation* processes in living cells include the passage of proteins through the endoplasmic reticulum [1,4,9], the transport of RNA through the nucleus pore membrane, the translocation of polypeptide chains from the inner mitochondrial or chloroplast membranes through the associated matrix [1,10], or the DNA plasmid passage from cell to cell through the cell walls [1]. Similar macromolecular transport processes occur in gene transfer [11], which is fundamental to understanding the way in which DNA may be incorporated into living cells. Not least, translocation has potentially far-reaching applications in bio- and nanotechnology, including rapid DNA sequencing, gene therapy, and controlled drug delivery [6,12,13].

The principal mechanism to drive the polymer through a pore employed in single molecule setups consists of a transmembrane electric-potential difference that mainly falls off inside the pore, using the fact that biopolymers carry charges and therefore experience a drift in the presence of the electric field [6]. This type of experimental translocation has received considerable interest from a theoretical point of view; compare, for instance, Refs. [14–21]. In nature, another driving mechanism exists, namely, the building-up of a transmembrane chemical-potential gradient, visible in an effective nonzero force directed toward the *trans* side, due to the presence of binding proteins, so-called chaperones or chaperonins, that nonspecifically bind to the chain-to-be-translocated [1,22]. A chaperone binding to the chain close

to the pore on the *trans* side prevents backsliding. As we follow the details of their binding and unbinding dynamics, we find that these chaperones therefore partially rectify the translocation. In Fig. 1 we present a sketch of this process. It should be noted that, typically, a binding protein occupies more than one monomer of the chain-to-be-translocated. Thus, DNA-binding proteins usually cover between 5 and 20 base pairs [1]. This gives rise to interesting physical phenomena (for instance, the "parking-lot effect" connected with the nonoptimal packing of binding proteins on the chain discussed in the following). Chaperone-driven translocation was studied previously in terms of statistical models for a homopolymeric chain [1,14,15,23–28] or in a random energy model [29].

Here we pose the question of how a specific sequence of different monomers in the chain-to-be-translocated affects the translocation dynamics. For the chaperones, that is, the sequence of different monomers causes a local variation of the binding affinity, and therefore changes the overall translocation dynamics. We investigate the case of a relatively stiff chain, for which the persistence length is much larger than the monomer size. For moderately long chains, we therefore neglect the polymeric degrees of freedom of the translocating chain that would overlay the chaperone-driving effects. For the simplest case of a chain consisting of two different types of monomers *A* and *B*, we evaluate the mean first passage times of the translocation process under variation of the chain composition and length, chaperone binding affinity, and chaperone volume concentration. In particular, we show how different parameters lead to the transition from a purely diffusive translocation dynamics up to the case of ballistic translocation (pure ratcheting). Moreover, we study the behavior of the waiting times of individual monomers during their passage through the pore, and we determine the effective force exerted by the chaperones on the chain.

We assume that, in our test polymers, the two monomer types *A* and *B* are adjacent to each other with probability P_{AB}. We use a Monte Carlo method to simulate the translocation process of polymers with sizes from 100 to 900 monomers. Our simulations were performed for several randomly constructed polymers with the same P_{AB}, and the translocation of each of these polymers was performed at least 10^3 times. To
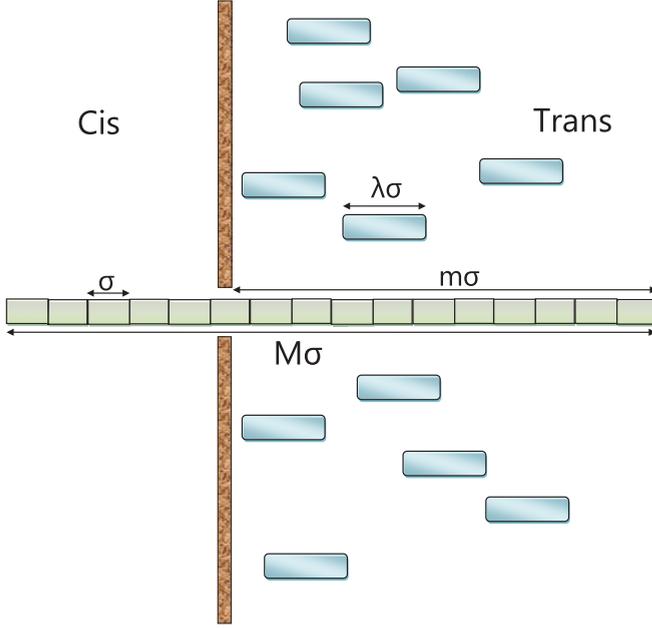
FIG. 1. (Color online) Translocating polymer consisting of monomers of size σ. On the *trans* side, binding proteins (chaperones) of size $\lambda\sigma$ bind to the translocating chain, giving rise to a chemical potential difference. The number m of already translocated monomers is a natural "reaction coordinate" of the translocation process.

compute the mean translocation time numerically, we used the Metropolis approach; cf. Refs. [28,30]. In the simulations, we attempt to move the polymer to the right and to the left with equal *a priori* probability. However, although the polymer may always move to the right, a move to the left is permitted only when the monomer next to the pore on the *trans* side is vacant [26,28,30]. The simulations results are compared to explicit calculations based on a master equation for the chaperone-driven translocation process.

We start our analysis with the master equation approach to the translocation process in Sec. II, before studying explicitly the sequence dependence of the mean translocation time in Sec. III. We draw our conclusions in Sec. IV.

II. MASTER EQUATION FOR CHAPERONE-DRIVEN TRANSLOCATION

The translocating stiff polymer chain consists of M monomers of length σ. The number of already translocated monomers is denoted by m, and therefore the length of the chain on the *trans* side is $m\sigma$. As sketched in Fig. 1, the chaperones are located exclusively on the *trans* side, at a volume concentration c_0. One chaperone covers λ monomers. On binding to the polymer, the chaperones prevent backsliding through the pore when the bound chaperone is located next to the pore [26–28]. In this study, for simplicity we assume that $\lambda = 2$. Given the two different monomer types A and B, we therefore have three different binding energies for a given chaperone. Namely, depending on the nature of the two monomers to which the chaperone binds (AA, AB, BB, and BA), we only need three binding energies per chaperone, denoted by ε_{AA}, $\varepsilon_{AB} = \varepsilon_{BA}$, and ε_{BB}.

A convenient stochastic description of the chaperone-driven translocation process may be based on the probability distribution $P(m,t)$ for finding m translocated monomers at time t. Given the discrete nature of the variable m, the dynamics of $P(m,t)$ follows the master equation [27,28]

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

where $W^+(m)$ and $W^-(m)$ are, respectively, the transition probabilities per unit time for forward and backward motion, further specified later. Thus, the first term on the right-hand side of Eq. (1) denotes an increase of m coming from state $m - 1$; that is, the chain moves one monomer from the *cis* to the *trans* side. Similarly, the second term stands for the backsliding of the chain through the pore toward the *cis* side by one monomer (passage from state $m + 1$ to m). The last line of Eq. (1) stands for the passage to the *trans* side ($m \to m + 1$) and *cis* side ($m \to m - 1$), respectively. Note that in Eq. (1), we already averaged out the degrees of freedom of the chaperones. This is possible in the reversible binding regime [28] when the typical chaperone binding and unbinding times are short compared to the time it takes the polymer to diffuse a distance σ [10,31].

As no chaperones exist on the *cis* side and we assume that chaperones cannot actively pull the chain through the pore, the forward rate is given by the constant

$$W^+(m) = k \equiv \frac{1}{2\tau_0}, \qquad (2)$$

where τ_0 is the typical time it takes the bare chain to diffuse over the distance σ of a monomer. The factor $\frac{1}{2}$ is introduced here as, on average, only every other attempted step is directed toward the *trans* side of the pore. As we assume, furthermore, that the system should eventually reach thermal equilibrium, the backward rate is fixed by the detailed balance condition [32]

$$W^-(m)\mathscr{Z}_\lambda(m) = W^+(m-1)\mathscr{Z}_\lambda(m-1), \qquad (3)$$

where $\mathscr{Z}_\lambda(m)$ is the partition function of the problem that will be defined later. Thus, we obtain

$$W^-(m) = k \frac{\mathscr{Z}_\lambda(m-1)}{\mathscr{Z}_\lambda(m)}. \qquad (4)$$

Note that, for flexible translocating polymers, the close-to-equilibrium assumption behind the detailed balance condition may be violated under conditions of strong driving [21]. From the forward and backward rates, we may then deduce the mean translocation time, corresponding to the calculation of the mean first passage time (MFPT). The result is given by [33]

$$T = \sum_{m=0}^{M} \left(\Phi(m) \sum_{m'=0}^{m} \frac{1}{W^+(m')\Phi(m')} \right), \qquad (5)$$

where we made use of the abbreviation

$$\Phi(m) = \prod_{u=1}^{m} \left(\frac{W^-(u)}{W^+(u)} \right). \qquad (6)$$

In the absence of polymeric degrees of freedom of the chain-to-be-translocated, the partition function $\mathscr{Z}_\lambda$ is solely a measure for the contributions of the chaperones. It is made up of the following three contributions: (i) The first part is the Boltzmann weight

$$\exp\left(-\frac{\varepsilon}{k_B T}\right),\qquad(7)$$

where ε represents the binding energy per chain monomer for a translocating homopolymer chain; the heterogeneous case will be considered later. Moreover, $k_B T$ denotes thermal energy. Note that, in our notation, the value of ε may be positive or negative. A negative value of ε indicates an affinity between chaperone and chain, while a positive value of ε indicates that binding is unfavorable. (ii) The second contribution corresponds to the number of complexions $\Omega_\lambda(m,n)$, counting the different ways in which one can distribute n chaperones of length λ over m monomers. This quantity is given by [27,28]

$$\Omega_\lambda(m,n) = \binom{m-(\lambda-1)n}{n} = \frac{[m-(\lambda-1)n]!}{n!(m-\lambda n)!}.\qquad(8)$$

(iii) Finally, the third part accounts for the chaperone entropy in the solution [28,34]. This results in the factor

$$\left(\frac{N}{N_t}\right)^n,\qquad(9)$$

in which N is the total number of (indistinguishable) chaperones. Moreover, $N_t = V/v_0$ is the number of voxels in the total volume V of the compartment on the *trans* side of the membrane in units of the volume v_0 occupied by an individual binding particle [35]. By taking all three contributions together, we obtain the partition factor for a given combination of m and n,

$$\mathscr{Z}_\lambda(m,n) = \Omega_\lambda(m,n)\chi^{\lambda n},\qquad(10)$$

where we took

$$\chi = \left[c_0 v_0 \exp\left(-\frac{\varepsilon}{k_B T}\right)\right]^{1/\lambda}\qquad(11)$$

and the volume concentration $c_0 = N/V$. Based on expression χ we define the effective binding energy (EBE)

$$\mathcal{E}_{\mathrm{eff}} = -\log(\chi),\qquad(12)$$

which is a useful quantity to characterize the translocation process, as we will show. The partition function $\mathscr{Z}_\lambda(m)$ needed for the definition of the transfer rates follows from expression (10) by summation over all allowed values of n,

$$\mathscr{Z}_\lambda(m) = \sum_{n=0}^{n_{\max}} \mathscr{Z}_\lambda(m,n),\qquad(13)$$

where the maximum number of binding proteins on the chain is given by $n_{\max} = [m/\lambda]$, the largest integer number smaller than or equal to m/λ.

Physically, from our definition of the partition function (10), we see that the backward rate (4) is proportional to the probability $\Omega_\lambda(m-1,n)/\Omega_\lambda(m,n)$ that the binding site closest to the pore on the *trans* side is vacant. Similarly, the expression $N/N_t = Nv_0/V = c_0 v_0$ is the probability that a chaperone is next to a given binding site. In terms of the

free energy $F(m,n) = -k_B T \log \mathscr{Z}_\lambda(m,n)$ we see that a larger χ value leads to a lower free energy (i.e., a more efficient translocation process). This can be achieved by increasing the volume concentration or the absolute binding energy $|\varepsilon|$.

A. The special case $\lambda = 1$

The consideration of the simplest case $\lambda = 1$ (i.e., when the chaperone size matches the monomer size of the chain-to-be-translocated) will be instructive for the subsequent discussion. We therefore present some analytical results pertaining to this case.

For $\lambda = 1$, the partition function $\mathscr{Z}_1(m)$ reduces to the exact form

$$\mathscr{Z}_1(m) = \sum_{n=0}^{m} \frac{m!}{n!(m-n)!}(\chi)^n = (1+\chi)^m,\qquad(14)$$

and we find the simple result for the quantity $\Phi(m)$,

$$\Phi(m) = (\chi+1)^{-m}.\qquad(15)$$

The mean translocation time therefore becomes

$$T(M) = 2\tau_0 \sum_{m=0}^{M}\left((\chi+1)^{-m}\sum_{m'=0}^{m}(\chi+1)^{m'}\right).\qquad(16)$$

By carrying out the sum, we obtain the result

$$T(M) = 2\frac{\tau_0}{\chi^2}(\chi+1)[(\chi-1)+\chi M + (\chi+1)^{-(M+1)}].\qquad(17)$$

For large values of the constant $\chi > 1$ (i.e., negative EBE), we expect an efficient translocation. Indeed, we find in the limit $M \gg 1$ that the mean translocation time

$$T(M) \sim 2\tau_0 \frac{1+\chi}{\chi}M \equiv \frac{M}{v}\qquad(18)$$

scales as the chain length M. This corresponds to diffusion under a constant bias. In Eq. (18) we defined the velocity

$$v \equiv W^+ - W^- = \frac{1}{2\tau_0}\frac{\chi}{1+\chi},\qquad(19)$$

in which the expression $\chi/(1+\chi)$ is the binding probability P. For very strong chaperone binding ($\chi \gg 1$), the binding probability tends to 1, and we recover the ballistic limit $T(M) = 2M\tau_0$ of completely ratcheted motion (i.e., on average at every other time step τ_0, the chain successfully moves one monomer to the *trans* side, while all attempts to move to the *cis* side are prohibited due to full chaperone coverage).

Conversely, for poor binding we have $\chi < 1$ and obtain the expansion

$$T(M) = 2\frac{\tau_0}{\chi^2}(\chi+1)\left(\frac{(M+1)(M+2)}{2}\chi^2 - \frac{(M+1)(M+2)(M+3)}{6}\chi^3 + \cdots\right).\qquad(20)$$

In the limit $\chi M \ll 1$, the result

$$T(M) \sim 2\tau_0 \frac{\chi+1}{2}M^2 \equiv \frac{\chi+1}{2K}M^2\qquad(21)$$
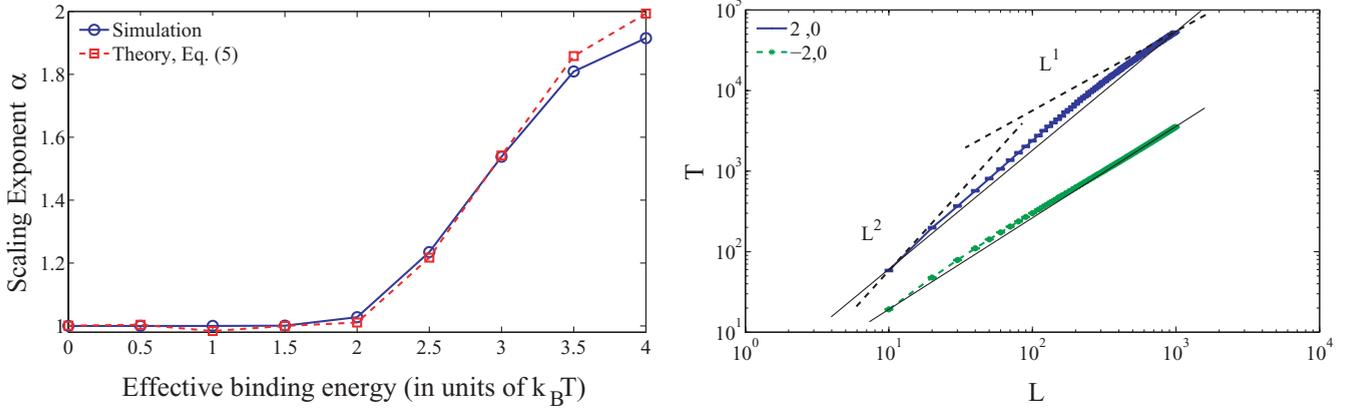
FIG. 2. (Color online) Left: Scaling exponent α of the mean translocation time $T(M) \simeq M^\alpha$ for $\lambda = 2$ and with identical effective binding energy for all monomers, as function of the EBE. We see the turnover from drift-dominated motion ($\alpha = 1$) to the diffusive scaling with $\alpha = 2$. The theoretically predicted behavior (5) agrees well with the simulation results. Note that α does not cross the threshold value 2. For $\mathcal{E}_{\text{eff}} = 4.0$ we find $\alpha = 1.95$, while for $\mathcal{E}_{\text{eff}} = 4.5$ we find $\alpha = 1.97$. Right: Examples for the scaling behavior of $T(M) \simeq M^\alpha$ at $P_{AB} = 0.10$, for $\lambda = 2$ with fitted slopes ($L = M\sigma$). The two dashed lines indicate the limiting behaviors $T(M) \simeq M^1$ and $T(M) \simeq M^2$. The pairs $\mathcal{E}_{\text{eff}}(A), \mathcal{E}_{\text{eff}}(B)$ are indicated in the plots.

yields, corresponding to the diffusive scaling, $T(M) \simeq M^2$. In Eq. (21) we defined the one-dimensional diffusion constant $K = 1/(2\tau_0)$. In the limit of very small χ (i.e., highly positive EBE), essentially no binding of chaperones occurs, and we obtain the result $T(M) = M^2/(2K)$ of unbiased diffusion. Depending on the order parameter χM, the mean translocation time $T(M)$ reaches from linear scaling in the case of complete ratcheting ($\chi M \gg 1$) to the purely diffusive M^2 dependence when no binding occurs ($\chi M \ll 1$). An analogous result pertains to the case of general λ for a homopolymer chain in the long-chain limit [28,30].

Figure 2 shows the scaling exponent α of the mean translocation time with chain length $T(M) \simeq M^\alpha$ as a function of the EBE, for $\lambda = 2$. In this case the EBE shown on the axis is assumed homogeneous along the translocating chain. In the figure we include the results from Monte Carlo simulations and the theoretical prediction based on numerical evaluation of Eq. (5) (for chain lengths 100–900). By sweeping the EBE, the exponent changes continuously from the diffusive scaling $\alpha = 2$ to the linear behavior corresponding to biased diffusion. The figure is universal in the sense that the EBE is the only free parameter determining the scaling exponent α, for finite M. For fixed values of $c_0 v_0$, the variation of α is solely due to a change in the binding energy ε; conversely, at fixed ε, the value of the EBE reflects changes of the concentration c_0. For instance, if the volume concentration of chaperones is fixed to $c_0 = 10\ \mu$M and their eigenvolume is fixed to $v_0 = 2\text{nm}^3$, then the binding energies $\varepsilon = -6.66 k_B T$ and $\varepsilon = -4.66 k_B T$ correspond to EBE values of -1 and 1, respectively. It is interesting to see that, at around $\mathcal{E}_{\text{eff}} = 4 k_B T$, the slope $\alpha = 2$ is reached, while up to $\mathcal{E}_{\text{eff}} = 1 k_B T$, essentially no influence of the EBE on α is observed, and thus the motion is completely drift dominated. Note, however, that the proportionality factor in the scaling law $T(M) \simeq M^\alpha$ does depend on the EBE; as we will see later. In the right panel of Fig. 2 we show examples for the scaling of the mean translocation time with chain length.

III. SEQUENCE DEPENDENCE OF THE MEAN TRANSLOCATION TIME

We now include the effect of the sequence of monomers with different binding affinities for chaperone binding. As stated earlier, we consider two sorts of monomers, A and B, and we consider chaperones of size $\lambda = 2$. The adjacency probabilities of finding two different monomers next to each other are P_{AA}, P_{BB}, and P_{AB}, with the corresponding binding energies ε_{AA}, ε_{BB}, and $\varepsilon_{AB} = \varepsilon_{BA}$. For simplicity, we assume equal numbers of the two monomers, $N_A = N_B$. Our simulation was performed for five to ten randomly constructed polymers with the same P_{AB}, and each polymer was translocated for at least 10^3 times.

A given sequence of monomers A and B along the chain-to-be-translocated corresponds to a form of quenched disorder. That means that the contribution from the binding energy enters the partition function through the fixed sequence ε_m, where $m = 1, 2, \ldots, M$ denotes the monomers along the translocating chain, and assumes the two values ε_A and ε_B. For a random sequence of monomers, only specified by the probability P_{AB} that monomer A is adjacent to monomer B, we can investigate the effect of P_{AB} in a mean-field approach. To this end, we write the backward rate $W^-(m)$ in the form

$$W^- = \frac{1}{2\tau_0}(1 - P), \tag{22}$$

where P is the mean-field binding probability. This quantity is based on the explicit probabilities for finding a chaperone bound to the monomer pair AA, AB, or BB, for which the equilibrium expressions are, respectively,

$$P_{AA}^{\text{bind}} = \frac{\chi(A)^2}{1 + \chi(A)^2}, \quad P_{BB}^{\text{bind}} = \frac{\chi(B)^2}{1 + \chi(B)^2},$$

$$P_{AB}^{\text{bind}} = \frac{\chi(A)\chi(B)}{1 + \chi(A)\chi(B)}. \tag{23}$$
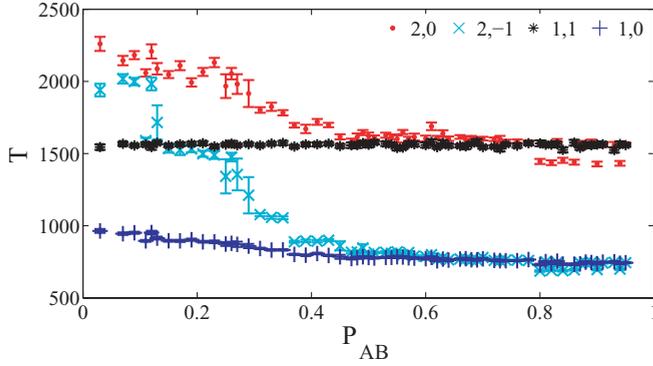
FIG. 3. (Color online) Mean translocation time T as a function of the adjacency probability P_{AB} for different values of EBE for monomer types A and B, varying between -2 and 2. The chain length is $M = 100$ monomers, and $\lambda = 2$. In all cases the EBE is comparatively small, and the translocation speed improves for increasing adjacency probability P_{AB}. In the plots, the pairs of numbers a, b denote the effective binding energies $\mathcal{E}_{\mathrm{eff}}(A), \mathcal{E}_{\mathrm{eff}}(B)$.

The binding probability per chaperone in this mean-field approach is therefore

$$P = \tfrac{1}{2}(1 - P_{AB})\big(P_{AA}^{\mathrm{bind}} + P_{BB}^{\mathrm{bind}}\big) + P_{AB} P_{AB}^{\mathrm{bind}}, \quad (24)$$

in terms of the adjacency probability P_{AB}. Thus, when we have a fully alternating sequence $(AB)_{M/2}$ corresponding to $P_{AB} = 1$, then $P = P_{AB}^{\mathrm{bind}}$, and the translocation is dominated exclusively by the sum of the two binding energies (Fig. 3). In the opposite limit of a block copolymer $(A)_{M/2}(B)_{M/2}$ or $(B)_{M/2}(A)_{M/2}$, we observe $P_{AB} = 0$ and $P \sim \tfrac{1}{2}(P_{AA}^{\mathrm{bind}} + P_{BB}^{\mathrm{bind}})$ [36].

The mean-field translocation time, according to Eq. (5), becomes

$$T = \frac{2\tau_0}{P}\left(M + 1 - \frac{1 - P}{P}[1 - (1 - P)^{M+1}]\right). \quad (25)$$

For sufficiently long chains $M \gg 1$, any value of P different from 0 leads to the approximate drift-diffusion behavior $T \sim 2\tau_0 M / P$, reaching its minimum for maximal binding, $P = 1$. Conversely, when binding becomes very poor, $P \to 0$, the leading-order scaling of T with the chain length follows the diffusive scaling $T \sim \tau_0 M^2$. In the Appendix, we introduce the diffusion advection equation as the continuum limit of the translocation process. As shown there, the Péclet number $\mathrm{Pe} = PM/2$ distinguishes between drift-dominated and diffusion-dominated regimes.

To see how the translocation dynamics depends on the sequence of monomer types A and B we perform simulations, in which we vary the respective EBE between $-2k_BT$ and $2k_BT$ and study the dependence of the mean translocation time T on the adjacency probability P_{AB}. Figure 3, for a polymer length of $M = 100$, confirms that higher values of the EBE, implying weaker binding affinity, indeed lead to longer translocation times. We observe that the mean translocation time is increasingly sensitive to the adjacency probability for the growing difference between the EBE of monomers A and B: T is highest when A and B separate into a block copolymer. In this case, there is one part of the chain that shows high affinity to the chaperones, and therefore translocates faster, while the second part of the chain has low chaperone affinity, and experiences a smaller translocation bias. For instance,
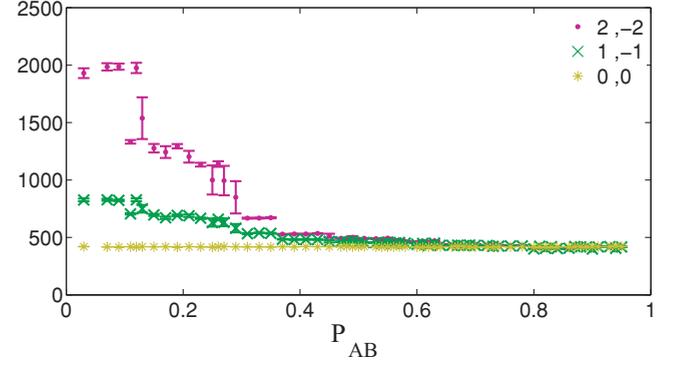
for the combination $2, -2$, the effect for the relatively short chain $M = 100$ already amounts to a factor of 4 between vanishing and full adjacency. At around $P_{AB} \approx 0.5$, the decay of T with growing P_{AB} turns over to a plateau for larger P_{AB}. This is an interesting observation; it indicates that for the best translocation performance, there is approximately no difference between a completely random sequence of monomers A and B ($P_{AB} = 1/2$), or a completely adjacent sequence $(AB)_{M/2}$. For small difference of the EBE for monomers A and B, an almost constant dependence of T with P_{AB} is observed. As it should, no sensitivity of T with P_{AB} shows when both monomer types have the same EBE. We also see a clear difference of the T values for identical EBE values of 1 and 0 [e.g., $(1,1)$ vs $(0,0)$]. Thus, despite the fact that the scaling exponent α is insensitive to values of EBE below ≈ 1 (cf. Fig. 2), the corresponding prefactor still significantly varies when the EBE values range in this interval. We also observe that several curves converge to each other. As expected from our mean-field analysis, this occurs for all cases in which the sum of the two EBEs coincides, as the combined binding energy dominates the translocation time at high adjacency.
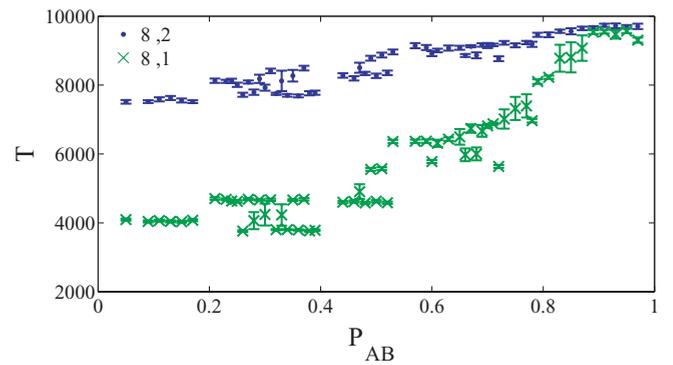
FIG. 4. (Color online) Mean translocation time T as a function of the adjacency probability P_{AB} for different values of EBE for monomer types A and B, varying between 1 and 8. The chain length is $M = 100$ monomers. Here the sum of the two EBEs is large, effecting an increase of T with P_{AB}, in contrast to the trend for smaller EBE values in Fig. 3.
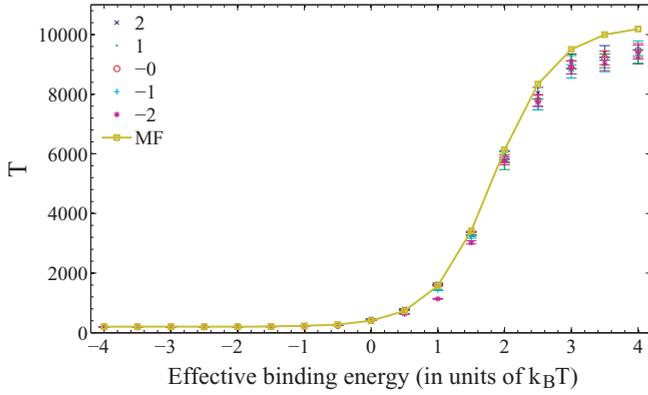
FIG. 5. (Color online) MFPT vs EBE [$\mathcal{E}_{\text{eff}}(A) = \text{EBE}(B)$] for $M = 100$ and $\lambda = 2$. Hardly any variation is observed up until $\mathcal{E}_{\text{eff}} \approx 0$, while above $\mathcal{E}_{\text{eff}} \approx 4$ a plateau is reached. An inflection occurs at $\mathcal{E}_{\text{eff}} \approx 2$. In this plot we also consider cases with chaperone-chaperone cooperativity. These are indicated by the numbers from -2 to 2: For each bound chaperone that binds right next to another chaperone, this EBE offset is added to the overall binding affinity. No significant dependence on the cooperativity is observed. The solid line is the result of the mean-field approximation, Eq. (25).

Will the mean translocation time always decrease with P_{AB}? As shown in Fig. 4, there may be situations when T actually increases with P_{AB}. To understand this phenomenon, consider the results for the mean translocation time of a homopolymer as function of EBE plotted in Fig. 5. We see that T is essentially insensitive to EBE for values below ≈ 0. At larger values of EBE, the curve $T(\text{EBE})$ exhibits an inflection, changing from positive to negative curvature. With the results from our mean-field approach this implies that, also in a heteropolymer, for a highly adjacent sequence of A and B monomers the sum of the two binding energies reaches high values when only one of the two EBE values is high. Thus, in a random or completely adjacent A-B sequence, the mean binding affinity is still low, and chaperones do not bind to any part of the chain. By contrast, when the sequence is less mixed, chaperones may still bind to the domains rich in the monomer with the higher chaperone affinity. In this case a portion of the

chain induces a more directed translocation dynamics, whereas the rest of the chain has almost freely diffusive character. As a consequence, for such EBE values, the mean translocation time T indeed increases with P_{AB}. Thus, the absolute values of EBE strongly influence the translocation dynamics. Note that, for the parameters considered here, cooperativity between adjacently binding chaperones does not significantly influence the magnitude of T (Fig. 5).

A. Probability density function of translocation times

A measure for the spread of the translocation times is the corresponding probability density function (PDF), which is expected to decay exponentially [37]. In Fig. 6 we show the results for the translocation time PDF for polymers with effective binding energies $\mathcal{E}_{\text{eff}}(A) = 2$ and $\mathcal{E}_{\text{eff}}(B) = -2$, for different adjacency probabilities P_{AB}. The chains used in these simulations were constructed such that, for a given value of P_{AB}, the center of the chain consisted exclusively of B (A) monomers, around these we put monomer pairs AB, and toward the extremities we locate exclusively A (B) monomers. In the left part of Fig. 6 we collect the results for B monomers in the center, while in the right part the central monomers are of type A. It can be seen that the translocation becomes less efficient with decreasing adjacency. Moreover, those chains with higher chaperone affinity in their center (central B monomers) translocate faster. For (statistically) symmetrical distributions of chain monomers around the middle of the chain, the passage of the central part of the chain therefore corresponds to a limiting regime of the translocation process. For block copolymers this behavior changes, as we will demonstrate for the monomer waiting times.

1. Direction invariance of translocation time probability density

It was demonstrated quite generally that the probability densities for the first passage in a finite geometry are direction independent, despite the presence of an external bias [38]. For the translocation problem this means that, with respect to the translocation force, the downhill and uphill translocation time distributions are identical. The reason for this seeming contradiction is that, for these distributions, only successful
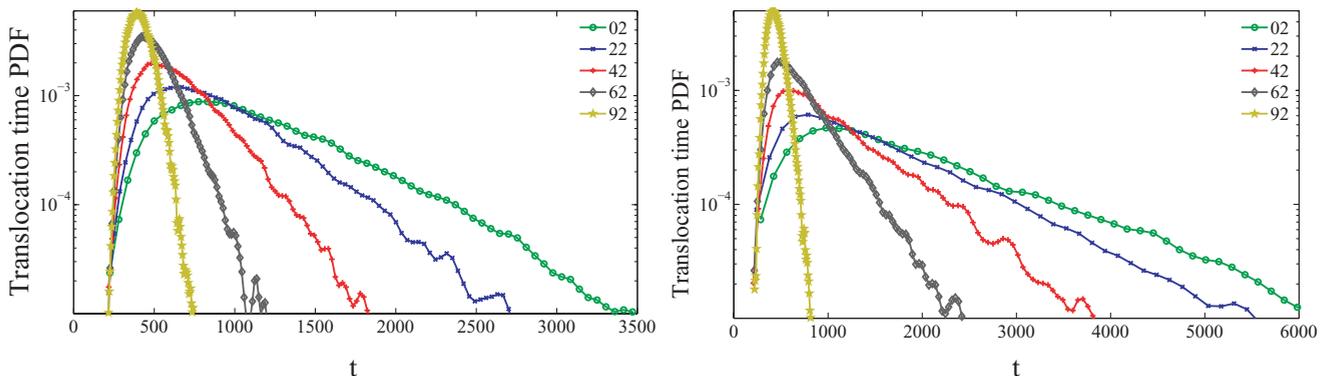
FIG. 6. (Color online) Transloaction time PDFs for polymers with effective binding energies $\mathcal{E}_{\text{eff}}(A) = 2$, and $\mathcal{E}_{\text{eff}}(B) = -2$, and with varying adjacency probabilities P_{AB} (indicated in % for the different curves in the plots). Left: Translocation time PDF for chains with central B monomers. Right: The same, but with central A monomers. In the right graph both the maxima are shifted to longer times and the exponential tails are wider.
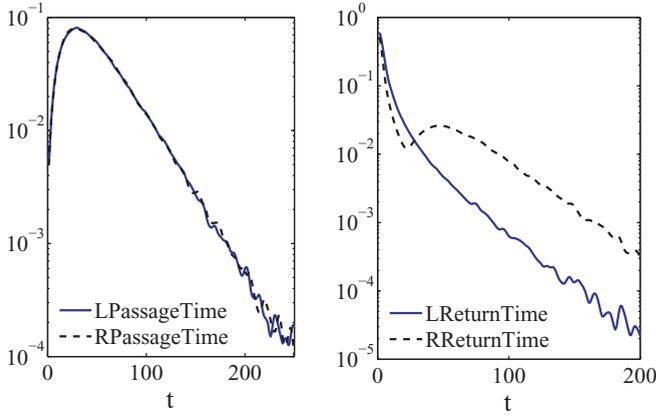
FIG. 7. (Color online) Left: PDF of the first passage times for a polymer with EBE = 0 and $M = 12$. The curve "LPassage" stands for chains that are released on the *trans* side of the pore (at $m = M$) and then translocate uphill to the *cis* side ($m = 0$). The curve "RPassage" stems from the opposite, downhill translocation from $m = 0$ to $m = M$. Both curves show excellent agreement, as predicted theoretically. Right: Plot of the return time distributions. "LReturn" denotes chains that are released at $m = 0$ and eventually return uphill, rather than translocating to $m = M$. "RReturn" corresponds to chains that are released at $m = M$ and return to that point, and do not translocate uphill. In this plot the effect of the chaperones leading to a bias toward the *trans* side is visible at longer times.

events are considered: The net flux of the translocation process remains directed.

We analyzed this prediction for chaperone-driven translocation; the results are displayed in Fig. 7. As can be seen, the agreement between left-to-right and right-to-left translocation time probability densities is perfect, thus corroborating the invariance for reversible chaperone driving. The chains used for this analysis are relatively short ($M = 12$), to allow sufficient repeats (10^5 times) of the uphill translocation. In this setup, the previously reflecting boundary at $m = 0$ is changed to an absorbing boundary in the simulations. For the chosen parameters, the probability to move from *trans* to *cis* side versus moving from *cis* to *trans* side is about 0.7. Note the logarithmic ordinates in both plots of Fig. 7, in which the linear asymptotic corresponds to the expected exponential decay of first passage distributions in a finite domain. In Fig. 7 we also plot the return time distribution to the left ($m = 0$) and right ($m = M$) boundaries. Here the bias toward the right (*trans*) side is reflected in the obtained statistics.

B. Waiting times and effective translocation force

We now proceed to analyze the waiting times per monomer (i.e., the time elapsing between the passage of monomer m and $m + 1$ through the pore). In Fig. 8 we show the mean waiting time, averaged over 10^3 translocation runs, for three cases with different adjacency and chain orientations versus translocation direction. For negligible adjacency, corresponding to the block copolymers $A_{M/2}B_{M/2}$ or $B_{M/2}A_{M/2}$, depending on the chain direction, the waiting times clearly differ for the two polymer blocks. For the completely mixed chain, the mean waiting time is approximately constant throughout the translocation process, as expected.
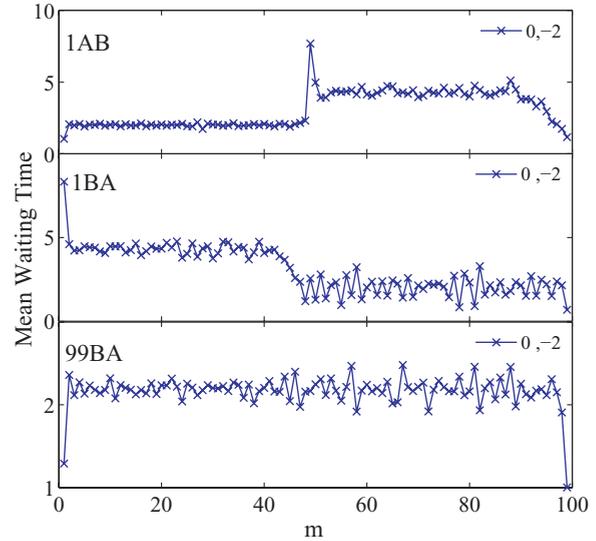
FIG. 8. (Color online) Mean waiting time for the following cases. Top panel: Low adjacency $P_{AB} \simeq 0.01$, and EBEs $\mathcal{E}_{\text{eff}}(A) = 0$ and $\mathcal{E}_{\text{eff}}(B) = -2$. The chain part rich in monomer type B enters the pore first, and translocation is faster for the first half of the chain. Middle panel: Same as before, but now the chain part rich in monomer type A enters the pore first. Bottom panel: Same EBE values, but almost complete adjacency ($P_{AB} = 0.99$); the waiting times are approximately constant throughout the chain. Note the oscillations due to the size of the chaperones, $\lambda = 2$.

Figure 9 features the cumulative waiting time corresponding to Fig. 8 (i.e., we plot the sum of the waiting times up until position m). In this plot, for the block copolymer case the difference between higher and lower binding affinity and the resulting effective force exerted on the chain is distinct, and the summation smoothens the behavior. In contrast to the block copolymer case, the completely adjacent chain does not lead to a change in slope.

Finally we evaluate the effective force exerted on each monomer during the translocation process; cf. Refs. [28,39]. Figure 10 shows the results for two block copolymers, and a case of maximal adjacency. In the block copolymer case, an appreciable force is only exerted for those monomers with significant binding energy. In that case strong oscillations occur, stemming from the parking-lot effect: A single chaperone
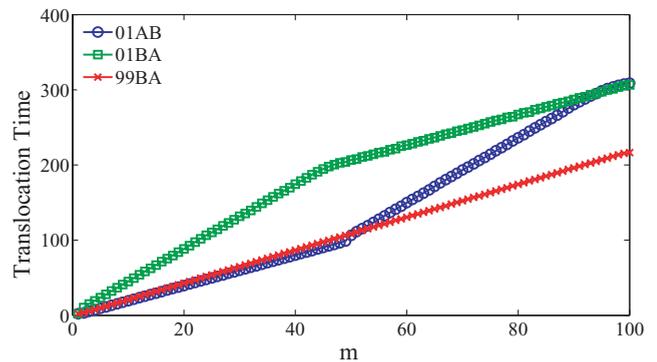
FIG. 9. (Color online) Cumulative waiting times corresponding to Fig. 8.
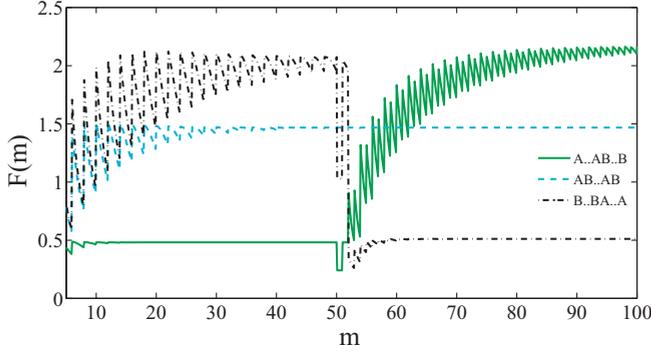
FIG. 10. (Color online) Effective force $F(m)$ exerted by the chaperones on the translocating chain. The EBE values are $\mathcal{E}_{\text{eff}}(A) = 0$ and $\mathcal{E}_{\text{eff}}(B) = -2$, and we have $\lambda = 2$. Solid (green) line: Block copolymer $(A)_{50}(B)_{50}$. Dash-dotted (black) line: Block copolymer $(B)_{50}(A)_{50}$. Dashed (blue) line: Completely adjacent chain $(AB)_{50}$.

needs two free binding sites to bind. The chain therefore needs to randomly make two steps to the *trans* side before an additional chaperone can bind. Indeed, the sawtooth pattern has a repeat of two monomers. The fact that the value of the force does not drop down to zero stems from the finite binding energy for the second polymer block. In the opposite case of complete adjacency, $(AB)_{M/2}$, initial oscillations due to the parking-lot effect occur; however, they die out after some 40 monomers arrive on the *cis* side.

IV. CONCLUSIONS

We investigated the effect of the sequence dependence of the binding energies of binding proteins (chaperones), which partially rectify the passage of a stiff polymer chain through a small pore in a membrane. In particular, we analyzed the dependence of the mean translocation time T on the adjacency probability P_{AB} of two kinds of monomers, A and B. Depending on the value of the effective binding energies for chaperone binding to A and B, we observe three behaviors of T as a function of P_{AB} that can be understood in terms of a mean-field approach: (i) When the difference between the binding affinity to monomers A and B is small, T is approximately independent of the adjacency probability P_{AB}. (ii) When the values of the EBE are sufficiently small, an increase of P_{AB} leads to appreciable chaperone binding throughout the translocating chain, and therefore T decreases with growing P_{AB}. (iii) The opposite case is observed when the EBEs are large. This effect is related to the curvature change in the $T(\mathcal{E}_{\text{eff}})$ dependence.

As a function of the adjacency probability P_{AB} at given EBE values, there exist two regimes: one, in which the value of T varies with P_{AB}, and another, in which T reaches a plateau. This effect is due to the mixing of A and B; we demonstrated that already at values of P_{AB} around $1/2$ the dynamic response approaches that of a completely mixed chain. The key parameters of sequence-dependent chaperone-driven translocation are therefore the effective binding energy and the degree of mixing of monomers A and B along the chain.

Additional details were shown in the analysis of the probability density function for the first passage of the chain across the pore. The decay of this PDF is exponential, and the maximum becomes increasingly sharp when the translocation approaches the completely rectified case (ratchet). Moreover, it turns out that a limiting step is connected to the translocation of the central part of the chain. We also showed that, as predicted, the translocation time PDF for chaperone driving is symmetric for uphill and downhill motion, while the existing bias is preserved in the left-right return time distributions. The waiting times per monomer and the effective force exerted on the translocating chain mirror the size effect of the chaperones, covering more than one chain monomer: This effect produces pronounced oscillations as a function of the translocation coordinate m.

The detailed distribution of binding affinities along chains-to-be-translocated may therefore be used to tune the translocation dynamics in chaperone-driven translocation processes. It would be interesting to extend this study to naturally occurring chaperones and typical chain sequences, including possible chaperone-chaperone cooperativity effects. In particular, it would be interesting to see to what extent the chaperone driving is sequence optimized in living cells.

ACKNOWLEDGMENTS

APPENDIX: DRIFT-DIFFUSION MODEL FOR CHAPERONE-INDUCED TRANSLOCATION

For the case of a homogeneous chain and chaperone size $\lambda\sigma$ we can introduce the continuum limit of the master equation (1). With the diffusion constant $D = \sigma^2/[2\tau_0]$ and the velocity

$$v = W^+ - W^- = \frac{1}{2\tau_0}\frac{\chi^\lambda}{1 + \chi^\lambda}, \tag{A1}$$

we can pass from the difference quotient to the differentials, and reach the diffusion advection equation

$$\frac{\partial \mathcal{P}(x,t)}{\partial t} + v\frac{\partial \mathcal{P}(x,t)}{\partial x} = D\frac{\partial^2 \mathcal{P}(x,t)}{\partial x^2}, \tag{A2}$$

where x is the dimensional translocation coordinate, obtained from $m\sigma$ in the simultaneous limit $\sigma, \tau_0 \to 0$. Finally, $\mathcal{P}(x,t)$ is the probability density to find the chain at coordinate x at time t. The initial condition is to start fully on the *cis* side of the pore (i.e., at $x = 0$). The boundary conditions are reflecting at $x = 0$ and absorbing at $x = L = M\sigma$. Note that, due to the absorbing boundary condition, the distribution $\mathcal{P}(x,t)$ is not normalized but decays to zero over time.

In this so-called transmission mode, the problem can be solved exactly in Laplace space. The result for the corresponding probability flux across the absorbing boundary can be written in the form [40]

$$j(L,s) = \frac{P_s e^{\text{Pe}}}{\text{Pe}\sinh(P_s) + P_s\cosh(P_s)}, \tag{A3}$$

where

$$\text{Pe} = \frac{vL}{2D} \tag{A4}$$

is the Péclet number defined as the dimensionless ratio of the flow versus the diffusion strength. For large Pe, the problem is dominated by the (directed) drift; in the opposite case, by (random) diffusion. With the binding probability $P = \chi^\lambda/(1 + \chi^\lambda)$ we therefore see that

$$\text{Pe} = \tfrac{1}{2} P M. \tag{A5}$$

Moreover, in Eq. (A3) we defined the quantity

$$P_s = \frac{1}{2D}\sqrt{v^2 + 4Ds}\,L. \tag{A6}$$

In Eq. (A3) the Laplace transform of the probability flux is given by

$$j(L,s) = \int_0^\infty j(x,t)e^{-st}\,dt\Big|_{x=L}. \tag{A7}$$

Taylor expansion of this relation in terms of the Laplace variable s then produces the moments of the first passage time. In particular, for the mean first passage time, corresponding to the mean translocation time, one obtains [40]

$$T = \frac{L^2}{D}\left(\frac{1}{2\text{Pe}} - \frac{1}{4\text{Pe}^2}(1 - e^{-2\text{Pe}})\right). \tag{A8}$$

At low driving, corresponding to high EBE values, the translocation is purely diffusive, $T \sim L^2/[2D]$, while at high Péclet numbers, the mean translocation time decreases as $T \sim L/v$.

[1] B. Alberts *et al.*, *Molecular Biology of the Cell* (Garland, New York, 2002).

[2] M. Akeson, D. Branton, J. J. Kasianowicz, E. Brandin, and D. W. Deamer, Biophys. J. **77**, 3227 (1999).

[3] M. Muthukumar, Annu. Rev. Biophys. Biomol. Struct. **36**, 435 (2007).

[4] T. A. Rapoport, Nature (London) **450**, 663 (2007).

[5] See, for instance, B. B. Griffiths and O. McClain, J. Basic Microbiol. **28**, 427 (1988).

[6] A. Meller, J. Phys.: Condens. Matter **15**, R581 (2003).

[7] M. Wanunu, J. Suntin, B. McNally, A. Chow, and A. Meller, Biophys. J. **95**, 4716 (2008); A. J. Storm, C. Storm, J. H. Chen, H. Zandbergen, J. F. Joanny, and C. Dekker, Nano Lett. **5**, 1193 (2005); U. F. Keyser, B. N. Koeleman, S. van Dorp, D. Krapf, R. M. M. Smeets, S. G. Lemay, N. H. Dekker, and C. Dekker, Nat. Phys. **2**, 473 (2006); C. Dekker, Nat. Nanotech. **2**, 209 (2007).

[8] J. Li, D. Stein, C. McMullan, D. Branton, M. J. Aziz, and J. A. Golovchenko, Nature (London) **412**, 166 (2001); A. J. Storm, J. H. Chen, X. S. Ling, H. W. Zandbergen, and C. Dekker, Nat. Mater. **2**, 537 (2003).

[9] W. Liebermeister, T. A. Rapoport, and R. Heinrich, J. Mol. Biol. **305**, 643 (2001).

[10] J.-F. Chauwin, G. Oster, and S. Glick, Biophys. J. **74**, 1732 (1998).

[11] B. Dreiseikelmann, Microbiol. Rev. **58**, 293 (1994).

[12] E. Di Marzio and J. J. Kasianowicz, J. Chem. Phys. **119**, 6378 (2003).

[13] J. J. Kasianowicz, E. Brandin, D. Branton, and D. W. Deamer, Proc. Natl. Acad. Sci. USA **93**, 13770 (1996).

[14] S. F. Simon, C. S. Peskin, and G. F. Oster, Proc. Natl. Acad. Sci. USA **89**, 3770 (1992).

[15] W. Sung and P. J. Park, Phys. Rev. Lett. **77**, 783 (1996).

[16] M. Muthukumar, J. Chem. Phys. **111**, 10371 (1999); Phys. Rev. Lett. **86**, 3188 (2001); J. Chem. Phys. **118**, 5174 (2003).

[17] D. K. Lubensky and D. R. Nelson, Biophys. J. **77**, 1824 (1999).

[18] J. Chuang, Y. Kantor, and M. Kardar, Phys. Rev. E **65**, 011802 (2001).

[19] Y. Kantor and M. Kardar, Phys. Rev. E **69**, 021806 (2004).

[20] S. Matysiak, A. Montesi, M. Pasquali, A. B. Kolomeisky, and C. Clementi, Phys. Rev. Lett. **96**, 118103 (2006).

[21] K. Luo, T. Ala-Nissilä, S.-Ch. Ying, and R. Metzler, Europhys. Lett. **88**, 68006 (2009).

[22] R. A. Stuart, D. M. Cyr, E. A. Craig, and W. Neupert, Trends Biochem. Sci. **19**, 87 (1994).

[23] P. L. Krapivsky and K. Mallick, J. Stat. Mech. (**2010**) P07007.

[24] A. Depperschmidt and P. Pfaffelhuber, Stoch. Proc. Applic. **120**, 901 (2010).

[25] R. Zandi, D. Reguera, J. Rudnick, and W. M. Gelbart, Proc. Natl. Acad. Sci. USA **100**, 8649 (2003).

[26] R. H. Abdolvahab, F. Roshani, A. Nourmohammad, M. Sahimi, and M. R. R. Tabar, J. Chem. Phys. **129**, 235102 (2008).

[27] T. Ambjörnsson, M. A. Lomholt, and R. Metzler, J. Phys.: Condens. Matter **17**, S3945 (2005).

[28] T. Ambjörnsson and R. Metzler, Phys. Biol. **1**, 77 (2004).

[29] Y. Kafri, D. K. Lubensky, and D. R. Nelson, Biophys. J. **86**, 3373 (2004).

[30] T. Ambjörnsson and R. Metzler, J. Phys.: Condens. Matter **17**, S1841 (2005).

[31] Typically, in membrane channels the diffusion of the polymer chain is much slower than in the free solution. This is particularly true for situations *in vivo*, in which additional effects such as binding of the chain-to-be-translocated to the membrane itself come into play. From that point of view we expect that even for typical chaperone sizes of 60–90 kDA (in comparison, a typical DNA binding protein is twice the size), the assumption of fast kinetics is in order.

[32] N. G. Van Kampen, *Stochastic Processes in Physics and Chemistry* (North-Holland, Amsterdam, 1992).

[33] C. W. Gardiner, *Handbook of Stochastic Methods* (Springer, New York, 2002).

[34] I. R. Epstein, Biophys. Chem. **8**, 327 (1978).

[35] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1972).

[36] Due to the fact that one pair AB or BA occurs in the block copolymer, the exact expression for the adjacency probability is $P_{AB} = (M - 1)^{-1}$.

[37] C. Chatelain, Y. Kantor, and M. Kardar, Phys. Rev. E **78**, 021129 (2008).

[38] A. M. Berezhkovskii, G. Hummer, and S. M. Bezrukov, Phys. Rev. Lett. **97**, 020601 (2006); A. M. Berezhkovskii and S. M. Bezrukov, J. Phys.: Condens. Matter **19**, 065148 (2007); J. Alvarez and B. Hajek, Phys. Rev. E **73**, 046126 (2006).

[39] The chemical potential difference between *cis* and *trans* sides equals the magnitude of the force $F(m)$ in units of $k_B T$.
For instance, for $\lambda = 1$ the chemical potential difference is $\ln(1 + \chi)$; for general λ and sufficiently large m it can be written as $\ln(r_{\max})$, where $r_{\max}$ is the maximum of the roots of the following equation: $r^{\lambda+1} - r^{\lambda} - (\chi^{\lambda})r = 0$.

[40] S. Redner, *A Guide to First-Passage Processes* (Cambridge University Press, Cambridge, England, 2001).