Polymer translocation through a cylindrical channel

Chiu Tai Andrew Wong and M. Muthukumar^{a)}

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

(Received 31 October 2007; accepted 22 February 2008; published online 16 April 2008)

A formalism of polymer translocation through a cylindrical channel of finite diameter and length between two spherical compartments is developed. Unlike previous simplified systems, the finite diameter of the channel allows the number of polymer segments inside the channel to be adjusted during translocation according to the free energy of possible conformations. The translocation process of a Gaussian chain without excluded volume and hydrodynamic interactions is studied using exact formulas of confinement free energy under this formalism. The free energy landscape for the translocation process, the distribution of the translocation time, and the average translocation time are presented. The complex dependencies of the average translocation time on the length and diameter of the channel, the sizes of the donor and receptor compartments, and the chain length are illustrated. © 2008 American Institute of Physics. [DOI: 10.1063/1.2897932]

I. INTRODUCTION

Polymer translocation through a nanopore or nanochannel is essential for biological functions. Examples include protein and DNA transport through membrane channels and viral DNA injection into host cells.¹ Recent experimental realization of monitoring single molecule translocation through protein or synthetic nanopores leads to the possibilities of single DNA sequencing, single molecular characterization, and single molecular detection.^{2–11} Many simulations^{12–24} and theoretical works^{25–40} have attempted to elucidate the translocation process.

Many factors affect the translocation process. For example, the driving force can be thermal diffusion of the polymer molecule, an electric field, a chemical potential, or a pulling force. The shape of the pore, such as the asymmetric interior of the α -hemolysin nanopore, affects translocation dynamics. For charged polymer molecules, the low dielectric constant of the membrane may impose an electrostatic energy barrier to the polymer entry.

We focus our attention on the translocation dynamics arising from the finite diameter of the channel. Previous theories,^{29,32,34} except Ref. 37 which estimated the effect of channel width using a scaling argument, assumed that the channel is very narrow so that the polymer aligns itself in a linear fashion inside the channel. In other words, a channel of length M could accommodate at most M/b segments, where b is the segment length. In reality, a channel has a finite diameter $2a$. The number of segments allowed to reside in the channel depends on both a and M . If the translocation process is in quasiequilibrium, the chain would adjust the number of segments inside the channel at each step according to the free energy of chain conformations.

In this paper, we account for the finite diameter of the channel so that the number of segments inside the channel

can be adjusted during translocation according to the free energy. Different conformations of the chain in the translocation process are shown in Fig. 1. The corresponding free energies are obtained explicitly for a Gaussian chain. All the states except state (f) are described by a single variable, p or q , or without any variables, where p and q are the numbers of segments inside the channel and the receptor compartment, respectively. The translocation, from the donor compartment to the receptor compartment, is considered as a two-step process. The chain initially residing in the donor compartment starts the process by finding the channel entrance with a chain end. It takes time τ_1 for that chain end to travel through the channel, reaching the interface between

FIG. 1. All possible conformations of a polymer of N segments during translocation. p and q are defined as the number of segments inside the channel and the receptor compartment, respectively.

^{a)}Author to whom correspondence should be addressed. Electronic mail: muthu@polysci.umass.edu.

the channel and the receptor compartment. All the segments then transfer into the receptor compartment in time τ_2 . The total translocation time $\tau_1 + \tau_2$ is calculated using the confinement free energy of a Gaussian chain. As shown below, τ_1 and τ_2 exhibit rich dependencies on the radius of the donor (R_1), radius of the receptor (R_2), radius of the channel (a), length of the channel (M), and chain length (N). All of these parameters are expressed in units of the segment length b .

II. THEORY

Consider a Gaussian chain of N segments with Kuhn segment length b . The probability $P(\mathbf{r}, \mathbf{r}_0; N)$, relative to the free space, that the two ends are at $\mathbf{r}$ and $\mathbf{r}_0$, respectively, obeys the equation⁴¹

$$\left(\frac{\partial}{\partial N} - \frac{b^2}{6} \nabla_{\mathbf{r}}^2 \right) P(\mathbf{r}, \mathbf{r}_0; N) = 0, \quad (2.1)$$

with the boundary condition $P(\mathbf{r}, \mathbf{r}_0; 0) = \delta(\mathbf{r}, \mathbf{r}_0)$. If the allowed conformations of the chain are limited by some surfaces, Eq. (2.1) is solved with the additional boundary condition that the probability is zero for $\mathbf{r}$ or $\mathbf{r}_0$ at the surfaces. We first consider two confining geometries, a sphere and a cylinder, before constructing the free energy landscape for the translocation process.

A. Polymer confined in a sphere

For a spherical compartment of radius R , the solution of Eq. (2.1) is a function of $\mathbf{r}$ and $\mathbf{r}_0$, which consist of six independent position variables. However, from the spherical symmetry, the solution should depend only on the radial distances r , r_0 , and the angle γ between the two vectors $\mathbf{r}$ and $\mathbf{r}_0$. This can be achieved by applying the Legendre addition theorem⁴² to the solution. The probability of the chain inside the sphere is

$$P_S(\mathbf{r}, \mathbf{r}_0; N) = \frac{1}{2\pi R^2} \sum_{n=0, \alpha_{n+1/2}} P_n(\cos \gamma) (2n+1) \times \frac{J_{n+1/2}(\alpha_{n+1/2} r/R) J_{n+1/2}(\alpha_{n+1/2} r_0/R)}{(J_{n+3/2}(\alpha_{n+1/2}))^2 \sqrt{rr_0}} \times \exp\left(-\frac{b^2 \alpha_{n+1/2}^2 N}{6R^2}\right), \quad (2.2)$$

where $P_n(x)$ is the Legendre polynomial of order n , $J_n(x)$ is the Bessel function of order n , and $\alpha_{n+1/2}$ are the zeros of the Bessel function of the first kind such that $J_{n+1/2}(\alpha_{n+1/2}) = 0$. Note that for each n , there is an infinite number of $\alpha_{n+1/2}$ to be summed over.

If the probability of the chain depends only on the radial distance r and r_0 but not the angle γ , only the $n=0$ terms in Eq. (2.2) are retained. It reduces to

$$P_S(r, r_0; N) = \frac{1}{2\pi R r r_0} \sum_{m=1}^{\infty} \sin\left(\frac{m\pi r_0}{R}\right) \sin\left(\frac{m\pi r}{R}\right) \times \exp\left(-\frac{m^2 \pi^2 b^2 N}{6R^2}\right), \quad (2.3)$$

which was given in Ref. 29. Equation (2.2) should be used

when the probability of the chain depends on the angle γ .

If one of the chain ends is free, and the other one is anchored at $r=R-c$ near the surface of the compartment, where c is the anchoring distance and is comparable to b , Eq. (2.3) becomes²⁹

$$P_{S1}(R, N) = \frac{2c}{R} \sum_{m=1}^{\infty} \exp\left(-\frac{m^2 \pi^2 b^2 N}{6R^2}\right) \equiv \frac{2c}{R} f_{S1}(R, N). \quad (2.4)$$

The higher $O(c^2)$ terms are ignored. If both ends are free inside the sphere, the probability is

$$P_S(R, N) = \int_0^R 4\pi r^2 dr \int_0^R 4\pi r_0^2 dr_0 P_S(r, r_0; N) = \frac{8R^3}{\pi} \sum_{m=1}^{\infty} \frac{1}{m^2} \exp\left(-\frac{m^2 \pi^2 b^2 N}{6R^2}\right). \quad (2.5)$$

B. Polymer confined in a cylinder

Similarly, by solving Eq. (2.1), the probability of a chain inside a cylindrical compartment of length M and diameter $2a$ is

$$P_C(\mathbf{r}, \mathbf{r}_0; N) = \frac{8}{\pi M a^2} \sum_{n=0}^{\infty} \sum_{m=1}^{\infty} \sum_{\mu_n} (\cos n\theta \cos n\theta_0 + \sin n\theta \sin n\theta_0) \sin\left(\frac{m\pi}{M} \left(z + \frac{M}{2}\right)\right) \times \sin\left(\frac{m\pi}{M} \left(z_0 + \frac{M}{2}\right)\right) \frac{J_n(\mu_n r/a) J_n(\mu_n r_0/a)}{J_{n+1}^2(\mu_n)} \times \exp\left(-\frac{b^2}{6} \left(\frac{\mu_n^2}{a^2} + \frac{m^2 \pi^2}{M^2}\right) N\right), \quad (2.6)$$

where μ_n are the zeros of Bessel function of the first kind such that $J_n(\mu_n) = 0$. The origin of the coordinates is at the center of the cylinder. z is along the axis of the cylinder. $\mathbf{r} = (r, \theta, z)$ and $\mathbf{r}_0 = (r_0, \theta_0, z_0)$.

If the chain end $\mathbf{r}_0$ can be anywhere inside the cylinder, from Eq. (2.6), the probability that the other chain end is at $\mathbf{r}$ is

$$P_C(\mathbf{r}; N) = \int_{-\pi}^{\pi} d\theta_0 \int_0^a r_0 dr_0 \int_{-M/2}^{M/2} dz_0 P_C(\mathbf{r}, \mathbf{r}_0; N) = \frac{16}{\pi} \sum_{m\mu_0} \left(\frac{J_0(\mu_0 r/a)}{\mu_0 J_1(\mu_0)} \right) \left(\frac{1 - (-1)^m}{m} \right) \times \sin\left(\frac{m\pi}{M} \left(z + \frac{M}{2}\right)\right) \times \exp\left(-\frac{b^2}{6} \left(\frac{\mu_0^2}{a^2} + \frac{m^2 \pi^2}{M^2}\right) N\right). \quad (2.7)$$

Here μ_0 are the zeros of $J_0(x)$. If both ends are free, the probability of the confined chain is

$$\begin{aligned}
P_C(N) &= \int_{-\pi}^{\pi} d\theta \int_0^a r dr \int_{-M/2}^{M/2} dz P_C(\mathbf{r}; N) \\
&= \frac{16Ma^2}{\pi^2} \sum_{m\mu_0} \left(\frac{1 - (-1)^m}{m\mu_0} \right)^2 \\
&\quad \times \exp\left(-\frac{b^2}{6} \left(\frac{\mu_0^2}{a^2} + \frac{m^2\pi^2}{M^2} \right) N\right). \quad (2.8)
\end{aligned}$$

If one end is free while the other end is allowed to be anchored anywhere at the lateral surface at $z = -M/2 + c$, from Eq. (2.7), the probability is

$$\begin{aligned}
P_{C1}(N) &= \int_{-\pi}^{\pi} d\theta \int_0^a r dr P_C(\mathbf{r}; N)|_{z=-M/2+c} \\
&= \frac{16ca^2}{M} \sum_{m\mu_0} \frac{1 - (-1)^m}{\mu_0^2} \\
&\quad \times \exp\left(-\frac{b^2}{6} \left(\frac{\mu_0^2}{a^2} + \frac{m^2\pi^2}{M^2} \right) N\right) \\
&\equiv \frac{16ca^2}{M} f_{C1}(N), \quad (2.9)
\end{aligned}$$

where $O(c^2)$ terms are ignored. By symmetry, the probability that one chain end is anchored at $z = M/2 - c$ while the other chain end is free to move is also given by Eq. (2.9). Similarly, the probability that one end is anchored at $z = -M/2 + c_1$ while the other end is anchored at $z = M/2 - c_2$ is

$$\begin{aligned}
P_{C2}(N) &= \frac{32\pi^3 c_1 c_2}{M^3} \sum_{m\mu_0} \frac{m^2}{\mu_0^2} (-1)^{m+1} \\
&\quad \times \exp\left(-\frac{b^2}{6} \left(\frac{\mu_0^2}{a^2} + \frac{m^2\pi^2}{M^2} \right) N\right) \\
&\equiv \frac{32\pi^3 c_1 c_2}{M^3} f_{C2}(N). \quad (2.10)
\end{aligned}$$

C. Free energy of confinement

The Helmholtz free energy of a confined Gaussian polymer is given by

$$F = -TS = -k_B T \ln P \quad (2.11)$$

for the probability of the chain conformation P . We consider the spherical donor and receptor compartments of radius R_1 and R_2 , respectively, being connected with a cylindrical channel of diameter $2a$ and length M . The chain can partition among the three compartments according to the free energy of the conformation. In the present calculations, we assume that there is only one segment at the interface between the cylinder and the sphere.

Figure 1 shows all possible conformations of a chain during translocation. Their corresponding free energy can be obtained from the various probabilities calculated in the previous sections. If the chain resides completely in the donor compartment as shown in Fig. 1(a), from Eqs. (2.5) and (2.11), the free energy is

$$\begin{aligned}
\frac{F_{(a)}}{k_B T} &= -\ln P_S(R_1, N) \\
&= -\ln \frac{8R_1^3}{\pi} - \ln \left(\sum_{m=1}^{\infty} \frac{1}{m^2} \exp\left(-\frac{m^2\pi^2 N}{6R_1^2}\right) \right). \quad (2.12)
\end{aligned}$$

If the polymer is confined in the donor compartment with one end anchored anywhere at the entrance of the channel at distance c , as in Fig. 1(b), its free energy is

$$\begin{aligned}
\frac{F_{(b)}}{k_B T} &\simeq -\ln(P_{S1}(R_1, N) \pi a^2) \\
&= -\ln \frac{2\pi c a^2}{R_1} - \ln \left(\sum_{m=1}^{\infty} \exp\left(-\frac{m^2\pi^2 N}{6R_1^2}\right) \right). \quad (2.13)
\end{aligned}$$

Equations (2.4) and (2.11) are used, and the factor πa^2 accounts for the area of the channel entrance that the chain end can anchor at. Suppose p segments reside in the cylindrical channel and $N-p$ segments reside in the donor compartment, as in Fig. 1(c). Now Eqs. (2.4) and (2.9) yield the probability as

$$\begin{aligned}
P_{(c)} &= \int d\mathbf{r} \int d\mathbf{r}_0 \int' d\mathbf{r}' P_C(\mathbf{r}, \mathbf{r}'; p) P_S(\mathbf{r}', \mathbf{r}_0; N-p) \\
&\simeq P_{C1}(p) P_{S1}(R_1, N-p). \quad (2.14)
\end{aligned}$$

The integrations of $\mathbf{r}$ and $\mathbf{r}_0$ are over the whole corresponding compartments. $\mathbf{r}'$ is integrated over the area connecting the two compartments. Integration of $\mathbf{r}'$ is performed by assuming that $\mathbf{r}'$ is at $z' = -M/2 + c_1$ inside the cylindrical channel while it is at $r' = R_1 - c$ inside the spherical compartment, making P_{S1} independent of the integration of $\mathbf{r}'$. The free energy is therefore

$$\frac{F_{(c)}(p)}{k_B T} \simeq -\ln(P_{C1}(p) P_{S1}(R_1, N-p)). \quad (2.15)$$

For the conformation of Fig. 1(d), one end is anchored in the cylinder surface at $z = M/2 - c_2$, and the other end can be anywhere inside the donor compartment. If there are p segments inside the channel, the free energy becomes

$$\frac{F_{(d)}(p)}{k_B T} \simeq -\ln(P_{C2}(p) P_{S1}(R_1, N-p)). \quad (2.16)$$

Likewise, the free energies of conformations in Figs. 1(e)–1(j) are

$$\frac{F_{(e)}}{k_B T} = -\ln P_{C1}(N), \quad (2.17)$$

$$\frac{F_{(f)}(p, q)}{k_B T} \simeq -\ln(P_{S1}(R_2, q) P_{C2}(p) P_{S1}(R_1, N-p-q)), \quad (2.18)$$

$$\frac{F_{(g)}(q)}{k_B T} \simeq -\ln(P_{C2}(N-q) P_{S1}(R_2, q)), \quad (2.19)$$

$$\frac{F_{(h)}(q)}{k_B T} \simeq -\ln(P_{C1}(N-q)P_{S1}(R_2, q)), \quad (2.20)$$

$$\frac{F_{(i)}}{k_B T} \simeq -\ln(P_{S1}(R_2, N)\pi a^2), \quad (2.21)$$

$$\frac{F_{(j)}}{k_B T} = -\ln P_S(R_2, N), \quad (2.22)$$

where the various probabilities P are defined in Eqs. (2.4), (2.5), (2.9), and (2.10).

III. TRANSLOCATION PROCESS

A. Free energy profile of translocation

The chain diffuses from the donor spherical compartment of radius R_1 through a cylindrical channel of diameter $2a$ and length M to the receptor compartment of radius R_2 . We construct the free energy profile as a function of q , the number of segments inside the receptor.

First, the anchoring distances c , c_1 , c_2 , and c_3 have to be determined. They are fixed by the continuity of the translocation free energy profile. The conditions are

$$F_{(b)} = F_{(c)}(1), \quad (3.1)$$

$$F_{(d)}(M) = F_{(f)}(M, 1), \quad (3.2)$$

$$F_{(f)}(M, N-M-1) = F_{(g)}(N-M), \quad (3.3)$$

$$F_{(h)}(N-1) = F_{(i)}. \quad (3.4)$$

Solving the equations gives

$$c = \frac{R_1}{2} \frac{f_{S1}(R_2, N-M)}{f_{S1}(R_2, N-M-1)f_{S1}(R_1, 1)}, \quad (3.5)$$

$$c_1 = \frac{M}{16} \frac{f_{S1}(R_1, N)}{f_{S1}(R_1, N-1)f_{C1}(1)}, \quad (3.6)$$

$$c_2 = \frac{M}{16} \frac{f_{S1}(R_2, N)}{f_{C1}(1)f_{S1}(R_2, N-1)}, \quad (3.7)$$

$$c_3 = \frac{R_2}{2} \frac{f_{S1}(R_2, N-M)}{f_{S1}(R_2, N-M-1)f_{S1}(R_1, 1)}. \quad (3.8)$$

f_{S1} , f_{C1} , and f_{C2} are defined in Eqs. (2.4), (2.9), and (2.10), respectively. Note that this is not the only way to determine the anchoring distances. The number of segments inside the channel can be larger (but not smaller) than M . However, a change from M to $2M$ led to only 0.03% change in the values of c and c_3 for $R_1=30$, $R_2=30$, $N=300$, $M=30$, and $a=3$ while c_1 and c_2 are unaffected. The results presented in this paper are based on Eqs. (3.5)–(3.8). The ground state dominance approximation⁴³ here yields unreasonably large anchoring distances as much as $R_1/2$ or $M/2$.

Consider a chain that is about to enter the receptor compartment so that one of the chain ends is fixed at the end of the channel. Two situations are possible. (1) There are p

segments in the channel and $N-p$ segments in the donor [Fig. 1(d)], and (2) the whole chain resides within the channel [Fig. 1(e)]. For situation (1), there are $N-M$ states for partitioning into the channel and the donor compartment because $M \leq p \leq N-1$. The probability of each conformation is calculated in the previous sections. Since the free energy $F/k_B T = -\ln P$ and P itself is the probability of the state, the average free energy of a chain that anchors at the entrance of the receptor compartment is

$$\bar{F}_{an} = \frac{-\sum_{p=M}^{N-1} P_{(d)}(p) \ln P_{(d)}(p) - P_{(e)} \ln P_{(e)}}{\sum_{p=M}^{N-1} P_{(d)}(p) + P_{(e)}}. \quad (3.9)$$

Next, we calculate the free energy of translocation when there are $q < N-M$ segments in the receptor compartment. Two conformations as depicted in Figs. 1(f) and 1(h) are possible. To map the translocation into a single variable process which is completely specified by q , we assume that the process is in quasiequilibrium so that when the chain takes the conformation in Fig. 1(f), all possible values of p are accessible at each step q . The average free energy is then

$$\bar{F}_1(q) = \frac{-\sum_{p=M}^{N-q-1} P_{(f)}(p, q) \ln P_{(f)}(p, q) - P_{(h)}(q) \ln P_{(h)}(q)}{\sum_{p=M}^{N-q-1} P_{(f)}(p, q) + P_{(h)}(q)}. \quad (3.10)$$

For $q \geq N-M$, the free energy of translocation is given by $F_{(h)}(q)$ because the chain is no longer long enough to partition into the donor compartment.

Figure 2 shows the free energy profiles of translocation. The free energies labeled in the figure are defined in Eqs. (2.12), (2.13), (3.9), (3.10), and (2.20)–(2.22). A typical translocation process goes as follows. In the beginning, the chain resides in the donor compartment with both chain ends free and the free energy is $F_{(a)}$ [Eq. (2.12)]. Then one of the chain ends finds the entrance of the channel and is fixed at the interface between the donor compartment and the channel. The free energy increases to $F_{(b)}$ [Eq. (2.13)] due to the loss in the translational entropy of the chain end. The translocation into the receptor compartment starts when a chain end is fixed at the interface between the channel and the receptor compartment. At the present state, the chain can either partition between the channel and the donor [Fig. 1(d)] or reside completely in the channel [Fig. 1(e)], with average free energy $\bar{F}_{an}$ [Eq. (3.9)]. For $0 \leq q < N-M$, where q is the number of segments inside the receptor compartment, the chain can either have p segments inside the channel and $N-p-q$ segments inside the donor compartment [Fig. 1(f)], or $N-q$ segments completely inside the channel [Fig. 1(h)]. The average free energy of translocation $\bar{F}_1(q)$ is given by Eq. (3.10). For $N-M \leq q \leq N$, the chain end is not long enough to partition into the donor compartment, and the free energy of translocation is $F_{(h)}(q)$, given by Eq. (2.20). The translocation process ends when all segments translocate into the receptor compartment, with one chain end fixed at the interface between the channel and the receptor compartment, with free energy $F_{(i)}$ [Eq. (2.21)]. Finally, the chain resides completely in the receptor compartment with both ends free, with free energy $F_{(j)}$ [Eq. (2.22)]. In Figs. 2(a) and 2(b), the shape of

FIG. 2. The free energy profiles of translocation of a chain with length $N=300$ and donor compartment radius $R_1=30$. The translocation coordinate q denotes the number of segments inside the receptor compartment. The various free energy labels F in (a) are defined in Eqs. (2.12), (2.13), (3.9), (3.10), and (2.20)–(2.22). (a) $M=30$, $R_2=30$, and $a=3$. (b) $M=30$, $R_2=60$, and $a=3$. (c) $M=30$, $R_2=60$, and $a=5$. (d) $M=45$, $R_2=60$, and $a=3$.

the free energy profile shows the transition from state (f) to state (h) as depicted in Fig. 1. State (h) is more favorable whenever $N-q$ is small or the channel (M and a) is large. Note that the free energy curve is not symmetric even when $R_1=R_2$ because of the definition of q .

When the translocation starts, the chain takes the conformation in Fig. 1(b) so that it is at the entrance of the channel. It takes τ_1 to translocate to the conformation in Fig. 1(d) or 1(e) so that the leading chain end is at the end of the channel. The free energy change is

$$F^+ = \bar{F}_{an} - F_{(b)}. \quad (3.11)$$

F^+ is the free energy barrier for the chain to pass through the channel and reach the receptor compartment. Its dependence on different variables is shown in Fig. 3. Crossover behaviors can be seen in Figs. 3(a) and 3(b). For small values of the cylinder radius a and the channel length M , state (d) is preferred. State (e) dominates for larger values of a and M .

B. Average translocation time

The translocation time is defined as the time needed for the polymer to go from the state that only one segment enters the channel [Fig. 1(b)] to the state that only one segment is left in the channel [Fig. 1(i)]. The translocation can be considered as a two-step process: (1) The chain overcomes the free energy barrier F^+ of channel filling and (2) the chain fills up the receptor with the number of segments from $q=1$ to $q=N-1$. The average translocation time of step (1) is

$$\langle \tau_1 \rangle \approx \frac{1}{k_0} \exp(F^+/k_B T). \quad (3.12)$$

where k_0 is a phenomenological rate constant, representing the friction between one monomer and the channel.

For step (2), the average translocation time is calculated from the mean first passage time of the corresponding Fokker-Planck equation^{25,27,44}

$$\frac{\partial W_m(t)}{\partial t} = \frac{\partial}{\partial m} \left(\frac{k_0}{k_B T} \frac{\partial F(m)}{\partial m} W_m(t) + k_0 \frac{\partial W_m(t)}{\partial m} \right), \quad (3.13)$$

where $W_m(t)$ is the probability that the polymer has m segments in the receptor compartment at time t . It can be shown that the mean first passage time τ satisfies the equation⁴⁴

$$k_0 \tau''(z) - \frac{F'(z)}{k_B T} k_0 \tau'(z) = -1, \quad (3.14)$$

where the prime denotes the derivative with respect to z , the initial position in free energy profile $F(q)$. The solution is

$$\tau(z) = \frac{1}{k_0} \int_z^s dx \int_r^x dy \exp(F(x) - F(y))/k_B T, \quad (3.15)$$

with reflecting and absorbing boundary conditions at r and s , respectively. Therefore, the average translocation time of step (2) is²⁷

FIG. 3. The free energy barrier F^+ of channel filling as a function of different parameters. (a) $M=30$, $R_1=R_2=30$, and $N=300$. (b) $a=5$, $R_1=R_2=30$, and $N=300$. (c) $a=5$, $R_2=30$, $M=30$, and $N=300$. (d) $a=3$, $R_1=10$, $R_2=15$, and $M=6$.

$$\langle \tau_2 \rangle = \frac{1}{k_0} \int_1^{N-1} dx \int_1^x dy \exp(F(x) - F(y))/k_B T, \quad (3.16)$$

where $F(q)$ is given by $\bar{F}_1(q)$ [Eq. (3.10)] for $q < N-M$ and by $F_{(h)}(q)$ [Eq. (2.20)] for $q \geq N-M$. The translocation time is calculated numerically.

The average total translocation time $\langle \tau_1 \rangle + \langle \tau_2 \rangle$ varies nonmonotonically with the channel length M . Figure 4 shows some typical results for the average total translocation time. The presence of the channel increases the free energy barrier F^+ of entering the receptor compartment by making the conformations in Figs. 1(d) and 1(e) less entropically favorable. The net effect is that $\langle \tau_1 \rangle$ increases with M , as shown in Fig. 4(a). The chain has to partition into the channel before entering the receptor, taking a more confined state. Nevertheless, for a Gaussian chain, the driving force of translocation increases with the increase in entropy per step of movement. The channel increases the entropy difference, and hence the driving force, to enter the more spacious receptor compartment, resulting in the decrease of $\langle \tau_2 \rangle$ with M . The interplay between these two effects results in the existence of a minimum in the average total translocation time. The location of this minimum depends on the parameters R_1 , R_2 , a , and N , as demonstrated in Figs. 4(a)–4(c).

The finite size of the channel leads to two different driving forces. First, if the chain is long enough, the conformations of the chain changes from (f) to (h) as translocation proceeds. For conformation (f), the driving force is given by the entropy difference between the donor and the receptor if

the number of segments inside the channel remains constant, as it is likely to happen for a long chain. For conformation (h), the tail of the chain resides completely in the channel. The driving force is given by the favorable entropy difference between a segment in the channel and that in the receptor. In general, this driving force is larger than the force involved in transferring segments between compartments. These two driving forces are evident in the free energy landscape in terms of the change in the slope with respect to q . Note that the change occurs well before $m=N-M$ if the diameter of the channel is large.

C. Translocation time distribution

The probability distribution of the translocation time τ_2 is

$$P(\tau_2) = - \left(\frac{k}{k_B T} \frac{\partial F(m)}{\partial m} W_m(\tau_2) + \frac{\partial W_m(\tau_2)}{\partial m} \right)_{m=N}, \quad (3.17)$$

where the probability distribution of the polymer $W_m(t)$ is obtained by solving Eq. (3.13) numerically using the Crank–Nicolson method.⁴⁵ The result is shown in Fig. 5. The distribution of the translocation time is a non-Gaussian peaked distribution with a long tail, with the character of drift-diffusion stochastic processes. τ_2 increases with the channel radius a because the entropy gain of escaping the channel into the receptor decreases. Note that the total translocation time $\tau_1 + \tau_2$ decreases as a increases, see Figs. 4(c) and 4(d). The reason is that τ_1 decreases as a increases, and this effect dominates over τ_2 .

FIG. 4. Total translocation time $\langle \tau_1 \rangle + \langle \tau_2 \rangle$. (a) $a=3$, $R_1=30$, $N=300$. Solid lines: $R_2=30, 40, 50$, and 60 . Dotted line: $\langle \tau_1 \rangle$ for $R_2=60$. Dashed line: $\langle \tau_2 \rangle$ for $R_2=60$. (b) $N=300$, $a=5$, $R_1=30$, $R_2=30, 40, 50$, and 60 . (c) $N=300$, $R_1=30$, $R_2=50$, $a=3, 4, 5$, and 6 . (d) $R_1=10$, $R_2=15$, $M=6$, $a=3, 4$, and 5 .

IV. DISCUSSION AND CONCLUSION

A formalism of polymer translocation between two spherical compartments through a cylindrical channel is developed. The average translocation time and its distribution are obtained. As a test case, the translocation of a Gaussian chain, which is purely entropy driven, is studied. The merit of the system is that exact formulas of the confinement free energies can be obtained. The translocation is composed of two steps, (1) worming through the channel to reach the entrance of the receptor compartment and (2) translocating all segments into the receptor.

FIG. 5. Translocation time distribution of τ_2 for $a=3, 4, 5$, respectively. $R_1=30$, $R_2=50$, $M=10$, and $N=300$.

The driving force of translocation may change during step (2). If the number of segments inside the channel remains the same, from Eq. (2.4), the driving force of step (2) is

$$f_2 \sim \frac{\pi^2}{6} \left(\frac{1}{R_1^2} - \frac{1}{R_2^2} \right) \quad (4.1)$$

under the ground state dominance approximation, which is qualitatively correct. If the space of the channel is large compared to the radius of gyration of the remaining $N-q$ segments, the segments would prefer to reside in the channel completely. Conformation (h) rather than (f), see Fig. 1, would dominate the translocation. Using Eqs. (2.4) and (2.9), the driving force of step (2) becomes

$$f'_2 \sim \frac{1}{6} \left(\frac{\mu_0^2}{a^2} + \frac{\pi^2}{M^2} - \frac{\pi^2}{R_2^2} \right), \quad (4.2)$$

where $\mu_0 \approx 2.4048$ is the first zero of Bessel function $J_0(x)$. f'_2 is larger than f_2 in the dimension scale we are interested in. Therefore, the limited space of the channel results in a larger driving force. Likewise, the longer the channel is, the more dominant conformation (h) in step (2) is and the shorter τ_2 is. Of course τ_1 , the time for filling up the channel, increases with the channel length. As a result, there is a minimum in the translocation time as a function of M , solely caused by entropy. Recently, this effect was observed in a computer simulation,¹⁹ as was also predicted for an infinitely thin cylinder.^{29,34}

Our calculation of τ_1 has to be used with caution. In our prescription, the free energy increase associated with the filling of the channel F^+ [Eq. (3.11)] depends on M nonmonotonically because we do not consider the intermediate states of Fig. 1(c) in obtaining $\bar{F}_{an}$. As M increases, the conformation of the final state of step (1) changes from (d) to (e) in Fig. 1. F^+ increases linearly with M at first, comes to a maximum, and then converges to a constant value at large M . Our estimation of τ_1 , using Eq. (3.12), is valid only for the linear region of F^+ . For large M , F^+ becomes constant and Eq. (3.12) predicts that τ_1 is independent of M , which is clearly wrong. Instead of estimating τ_1 by assuming chain anchoring with free energy $\bar{F}_{an}$, a translocation coordinate of the chain entering the channel [Fig. 1(c)] is necessary. For very long channels, the whole chain may translocate into the channel without reaching the end of the channel. The chain then diffuses along the channel until it reaches the end or returns to the donor compartment. The translocation times shown in Fig. 4 are all calculated in the linear region of F^+ . Our calculation of τ_2 is valid for any M .

The average total translocation time changes nonlinearly with N for short chains and linearly for long chains, as shown in Fig. 4(d). The nonlinearity in the limit of short chains mainly arises from $\langle\tau_1\rangle$. However, for long chains, $\langle\tau_1\rangle$ is essentially independent of N as the change in free energy of the conformations (d) is not significant. On the other hand, for short chains, the conformations change from Fig. 1(e) to Fig. 1(d) as N increases, and the free energy increases rapidly with N . The transition can be seen in Fig. 3(d). It is to be emphasized that in our formalism, the friction coefficient of every monomer is assumed to be the same, independent of N , whether it is within the spheres or the cylinder. For longer chains, $\langle\tau_2\rangle$ dominates over $\langle\tau_1\rangle$ and now $\langle\tau_2\rangle$ increases linearly with N , as expected for a process dominated by drift.

The formalism is applied to only a Gaussian chain in the interest of getting exact results. These results need to be appended by a proper account of excluded volume effect in order to compare with experimental systems. Even for an isolated spherical cavity, there is a continuing debate^{46–48} for the form of the confinement free energy. However, when the excluded volume effect is removed, our exact results must be recovered.

Our estimation of the translocation time by calculating $\langle\tau_1\rangle$ and $\langle\tau_2\rangle$ separately and adding them is valid only in certain situations. Under this approximation, the chain is not allowed to return from step (2) to step (1). As shown in the Appendix, the approximation is good for a free energy profile of step (2) that is significantly downhill without any initial barrier, such as those in Figs. 2(c) and 2(d). In this case, the upper bound of the correction factor to $\langle\tau_1\rangle$ is related to the inverse of the gradient of the free energy profile. If the free energy profile is initially flat up to x_m and then significantly downhill, $\langle\tau_1\rangle$ would need to be corrected by a factor up to $\sim x_m$. Finally, if there is an initial free energy barrier of height Δf , then the upper bound of the correction factor to $\langle\tau_1\rangle$ is $\sim \exp(\Delta f)$. Physically, the polymer is unlikely to retreat from step (2) to step (1) in the first case, so that the

approximation is valid. For the second and third cases, returning to step (1) cannot be ignored. The details of arriving at these conclusions are given in the Appendix.

The accuracy of the free energy expressions given in this paper depends on the geometry of the system. Explicitly, we require that $c/R \ll 1$ and $c/(M/2) \ll 1$, where c is the anchoring distance, R is the radius of the corresponding spherical compartment, and M is the length of the channel. These conditions are satisfied when $(a/R)^2 \ll 1$ and $a^2/RM \ll 1$, where a is the radius of the channel. It should be noted that all of these dimensions are in units of the Kuhn length b .

It must be noted that the present formalism relies on the quasiequilibrium assumption so that the free energy during each step of translocation is given by the equilibrium conformational entropy of the chain. There exists another regime where the translocation time is shorter than or comparable to the relaxation time of the chain. Simulations and scaling arguments addressing this regime show conflicting conclusions and the results are under active discussion.^{18,19,33,40,49–51} In the present paper, we have focused our attention on situations satisfying the quasiequilibrium assumption.

Further, it must be noted that we ignored the possibility that the chain may enter or leave the channel by forming hairpins. The wider the channel is, the more probable the formation of hairpins is. The occupancy of chain entrance into the channel as hairpins is influenced by the extent of excluded volume effect and the bending elasticity of the polymer. However, in the present paper for Gaussian chains, we have assumed that the entry of the polymer into the channel and its exit occur only through chain ends.

APPENDIX: ERROR ESTIMATION OF TWO-STEP TRANSLOCATION TIME

In this paper, we separate the translocation time into two parts, $\tau = \tau_1 + \tau_2$. In fact, the average translocation time $\langle\tau\rangle = \langle\tau_1 + \tau_2\rangle \neq \langle\tau_1\rangle + \langle\tau_2\rangle$ so it is necessary to estimate the error introduced by calculating $\langle\tau_1\rangle$ and $\langle\tau_2\rangle$ separately. Physically, this approximation does not allow the translocating polymer to go back from step (2) to step (1). Intuitively, this approximation is valid when the free energy profile of step (2) is monotonically downhill with a large gradient and becomes worse when the free energy profile is flat, is uphill, or has a free energy barrier. We estimate the correction quantitatively as follows.

We consider a hypothetical free energy profile shown in Fig. 6(a),

$$F(x) = \begin{cases} 0, & -\delta \leq x < 0 \\ \Delta F + G(x), & 0 \leq x \leq N, \end{cases} \quad (\text{A1})$$

where ΔF is the free energy barrier in step (1), $G(x)$ is an arbitrary free energy profile satisfying $G(0)=0$, and δ is a small arbitrary positive constant representing the length of the trapped state. In our calculations, we treat the translocation as a two-step process,

$$\langle\tau_1\rangle = \frac{1}{k_0} \exp(\Delta F), \quad (\text{A2})$$

FIG. 6. (a) Free energy profiles of translocation involve crossing a free energy barrier ΔF [step (1)] and an arbitrary profile $G(x)$ [step (2)]. The corresponding correction factors to $\langle\tau_1\rangle$ (in the limits explained in the text) are (b) $C=N$, (c) $C=1/|\mu|$, (d) $C=\exp(\mu N)/\mu$, (e) $C=\exp(\mu_1 x_m)(1/\mu_1 + 1/|\mu_2|)$, and (f) $C=x_m + 1/|\mu|$.

$$\langle\tau_2\rangle = \frac{1}{k_0} \int_0^N dx \int_0^x dy \exp(G(x) - G(y))/k_B T. \quad (\text{A3})$$

On the other hand, the average total translocation time $\langle\tau\rangle = \langle\tau_1 + \tau_2\rangle$ is given by

$$\langle\tau\rangle = \frac{1}{k_0} \int_{-\delta}^N dx \int_{-\delta}^x dy \exp(F(x) - F(y))/k_B T, \quad (\text{A4})$$

with reflecting and absorbing boundary conditions at $x=-\delta$ and $x=N$, respectively. Equation (A4) can be rewritten as

$$\langle\tau\rangle = \frac{\delta^2}{2k_0} + \left[\delta \int_0^N \exp \frac{G(x)}{k_B T} dx \right] \langle\tau_1\rangle + \langle\tau_2\rangle \quad (\text{A5})$$

by using Eqs. (A2) and (A3). The first term with $O(\delta^2)$ is negligible since δ is small. Immediately, we identify that $\langle\tau_1\rangle$ has to be corrected by a factor $C\delta$, where $C = \int_0^N \exp(G(x)/k_B T) dx$. The correction factor accounts for the fact that the polymer can retreat from step (2) to the trapped state step (1), which is missing in our calculations. Note that $\langle\tau_2\rangle$ is unchanged and still given by the reflecting boundary condition at $x=0$. This is because after falling back to the trapped state step (1), the polymer has to start step (2) again from $x=0$. The net effect is that the time spent in step (2) is exactly the same as if there is a reflecting boundary at $x=0$, and only $\langle\tau_1\rangle$ has to be modified because of multiple visits.

It is instructive to estimate the correction factor C for some typical free energy profiles. If step (2) is purely diffusive, such that $G(x)=0$ as shown in Fig. 6(b), $C=N$. Physically, this means that the probability of the polymer going from step (2) to step (1) is proportional to the length of step (2), N . Before touching $x=N$, the polymer on average retracts back to step (1) for N times. If the free energy profile of step (2) is a straight line with slope μ as shown in Figs. 6(c) and 6(d), $C=1/|\mu|$ for $\mu<0$ and $|\mu|N \gg 1$ while $C=\exp(\mu N)/\mu$ for $\mu>0$ and $\mu N \gg 1$. The results have clear physical meanings. For $\mu<0$ and $|\mu|N \gg 1$, the free energy profile of step (2) is straightly downhill so the polymer is unlikely to go from step (2) back to step (1). For $\mu>0$ and $\mu N \gg 1$, the free energy profile of step (2) is straightly uphill so that step (2) also acts like a free energy barrier with height μN . The total translocation time is then $\langle\tau\rangle \sim \exp(\Delta F + \mu N)$. Finally, we estimate C when the free energy profile of step (2) has a maximum at x_m [Fig. 6(e)], or it is diffusive until x_m and goes

down with a constant slope μ [Fig. 6(f)]. For the profile shown in Fig. 6(e) with $\mu_1 x_m \gg 1$ and $|\mu_2|(N-x_m) \gg 1$, $C = \exp(\mu_1 x_m)(1/\mu_1 + 1/|\mu_2|)$. For the profile shown in Fig. 6(f) with $|\mu|(N-x_m) \gg 1$, $C = x_m + 1/|\mu|$. Thus, when there is a significant free energy barrier in step (2) with height $\mu_1 x_m$, $\langle\tau_1\rangle$ is modified as if the total barrier height is $\Delta F + \mu_1 x_m$, so that $\langle\tau_1\rangle \sim \exp(\Delta F + \mu_1 x_m)$ while $\langle\tau_2\rangle$ is unchanged. For a profile that is initially flat until x_m and then significantly downhill, $\langle\tau_1\rangle$ is corrected by a factor $\sim x_m$.

Finally, we note that our estimation of the correction factor C to $\langle\tau_1\rangle$ cannot be applied to our problem in a straightforward manner because we have defined the starting point of step (1) as the conformation that a chain end has just entered the channel [Fig. 1(b)]. There is no reason to assume that the chain cannot retreat back to the donor compartment [Fig. 1(a)]. In other words, instead of reflecting boundary at $x=-\delta$, we should use the absorbing boundary condition. Then the average translocation time $\langle\tau\rangle$ is given by the conditional mean first passage time, starting from somewhere near $x=-\delta$ and ending at $x=N$. However, the time calculated this way is always shorter than that given by a reflecting boundary condition at $x=-\delta$ as assumed above. Therefore, the above estimation serves as an *upper bound* of the correction to the total translocation time. Moreover, $\langle\tau_2\rangle$ will also be modified due to the fact that some reflections at $x=0$ belong to unsuccessful translocations where the chain goes from step (2) to step (1) and then back to the donor compartment. Nevertheless, the complications can be avoided if we define the beginning of step (1) as the trapped conformation shown in Fig. 1(a) instead of Fig. 1(b). Then the correction is exact and can be calculated numerically for an arbitrary free energy profile.

ACKNOWLEDGMENTS

It is a pleasure to thank Dr. C. Y. Kong for helpful discussions. Acknowledgment is made to NIH Grant No. 2R01HG002776-04A1 and the Materials Research Science and Engineering Center at the University of Massachusetts at Amherst.

¹B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular Biology of the Cell*, 4th ed. (Garland Science, New York, 2002).

²J. J. Kasianowicz, E. Brandin, D. Branton, and D. W. Deamer, *Proc. Natl.*

- Acad. Sci. U.S.A.* **93**, 13770 (1996).
- ³H. Bayley and P. S. Cremer, *Nature (London)* **413**, 226 (2001).
 - ⁴A. Meller, L. Nivon, and D. Branton, *Phys. Rev. Lett.* **86**, 3435 (2001).
 - ⁵P. Chen, J. J. Gu, E. Brandin, Y. R. Kim, Q. Wang, and D. Branton, *Nano Lett.* **4**, 2293 (2004).
 - ⁶J. L. Li, M. Gershow, D. Stein, E. Brandin, and J. A. Golovchenko, *Nat. Mater.* **2**, 611 (2003).
 - ⁷A. J. Storm, J. H. Chen, H. W. Zandbergen, and C. Dekker, *Phys. Rev. E* **71**, 051903 (2005).
 - ⁸H. Wang and D. Branton, *Nat. Biotechnol.* **19**, 622 (2001).
 - ⁹J. Mathe, A. Arinstein, Y. Rabin, and A. Meller, *Europhys. Lett.* **73**, 128 (2006).
 - ¹⁰D. Fologea, B. Ledden, D. S. McNabb, and J. L. Li, *Appl. Phys. Lett.* **91**, 053901 (2007).
 - ¹¹M. Akeson, D. Branton, J. J. Kasianowicz, E. Brandin, and D. W. Deamer, *Biophys. J.* **77**, 3227 (1999).
 - ¹²S. S. Chern, A. E. Cardenas, and R. D. Coalson, *J. Chem. Phys.* **115**, 7772 (2001).
 - ¹³C. Y. Kong and M. Muthukumar, *Electrophoresis* **23**, 2697 (2002).
 - ¹⁴P. Tian and G. D. Smith, *J. Chem. Phys.* **119**, 11475 (2003).
 - ¹⁵A. Aksimentiev, J. B. Heng, G. Timp, and K. Schulten, *Biophys. J.* **87**, 2086 (2004).
 - ¹⁶P. Cifra, *J. Chem. Phys.* **124**, 024706 (2006).
 - ¹⁷H. C. Loeb, R. Randel, S. P. Goodwin, and C. C. Matthai, *Phys. Rev. E* **67**, 041913 (2003).
 - ¹⁸K. F. Luo, I. Huopaniemi, T. Ala-Nissila, and S. C. Ying, *J. Chem. Phys.* **124**, 114704 (2006).
 - ¹⁹K. Luo, T. Ala-Nissila, and S. C. Ying, *J. Chem. Phys.* **124**, 034714 (2006).
 - ²⁰S. Matysiak, A. Montesi, M. Pasquali, A. B. Kolomeisky, and C. Clementi, *Phys. Rev. Lett.* **96**, 118103 (2006).
 - ²¹P. Romiszowski and A. Sikorski, *Comput. Mater. Sci.* **38**, 533 (2007).
 - ²²Y. S. Tsai and C. M. Chen, *J. Chem. Phys.* **126**, 144910 (2007).
 - ²³S. Tsuchiya and A. Matsuyama, *Phys. Rev. E* **76**, 011801 (2007).
 - ²⁴D. S. Wei, W. Yang, X. G. Jin, and Q. Liao, *J. Chem. Phys.* **126**, 204901 (2007).
 - ²⁵W. Sung and P. J. Park, *Phys. Rev. Lett.* **77**, 783 (1996).
 - ²⁶D. K. Lubensky and D. R. Nelson, *Biophys. J.* **77**, 1824 (1999).
 - ²⁷M. Muthukumar, *J. Chem. Phys.* **111**, 10371 (1999).
 - ²⁸M. Muthukumar, *Phys. Rev. Lett.* **86**, 3188 (2001).
 - ²⁹M. Muthukumar, *J. Chem. Phys.* **118**, 5174 (2003).
 - ³⁰T. Ambjornsson, S. P. Apell, Z. Konkoli, E. A. Di Marzio, and J. J. Kasianowicz, *J. Chem. Phys.* **117**, 4063 (2002).
 - ³¹A. M. Berezhkovskii and I. V. Gopich, *Biophys. J.* **84**, 787 (2003).
 - ³²O. Flomenbom and J. Klafter, *Biophys. J.* **86**, 3576 (2004).
 - ³³Y. Kantor and M. Kardar, *Phys. Rev. E* **69**, 021806 (2004).
 - ³⁴C. Y. Kong and M. Muthukumar, *J. Chem. Phys.* **120**, 3460 (2004).
 - ³⁵R. Metzler and J. Klafter, *Biophys. J.* **85**, 2776 (2003).
 - ³⁶K. L. Sebastian and A. K. R. Paul, *Phys. Rev. E* **62**, 927 (2000).
 - ³⁷E. Slonkina and A. B. Kolomeisky, *J. Chem. Phys.* **118**, 7112 (2003).
 - ³⁸J. L. A. Dubbeldam, A. Milchev, V. G. Rostiashvili, and T. A. Vilgis, *EPL* **79**, 18002 (2007).
 - ³⁹J. L. A. Dubbeldam, A. Milchev, V. G. Rostiashvili, and T. A. Vilgis, *Phys. Rev. E* **76**, 010801 (2007).
 - ⁴⁰T. Sakaue, *Phys. Rev. E* **76**, 021803 (2007).
 - ⁴¹M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics* (Oxford University Press, New York, 1988).
 - ⁴²G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 5th ed. (Academic, New York, 2000).
 - ⁴³P. G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell University Press, Ithaca, 1979).
 - ⁴⁴C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences* (Springer-Verlag, Berlin, 1983).
 - ⁴⁵W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1992).
 - ⁴⁶A. Cacciuto and E. Luijten, *Nano Lett.* **6**, 901 (2006).
 - ⁴⁷S. Jun, A. Arnold, and B. Y. Ha, *Phys. Rev. Lett.* **98**, 128303 (2007).
 - ⁴⁸T. Sakaue and E. Raphael, *Macromolecules* **39**, 2621 (2006).
 - ⁴⁹J. Chuang, Y. Kantor, and M. Kardar, *Phys. Rev. E* **65**, 011802 (2002).
 - ⁵⁰D. Panja, G. T. Barkema, and R. C. Ball, *J. Phys.: Condens. Matter* **19**, 432202 (2007).
 - ⁵¹J. K. Wolterink, G. T. Barkema, and D. Panja, *Phys. Rev. Lett.* **96**, 208301 (2006).