Journal of Statistical Mechanics: Theory and Experiment
An IOP and SISSA journal

PAPER: Biological modelling and information

Mean velocity and effective diffusion constant for translocation of biopolymer chains across membrane

Xining Xu and Yunxin Zhang

Laboratory of Mathematics for Nonlinear Science, Shanghai Key Laboratory for Contemporary Applied Mathematics, Centre for Computational Systems Biology, School of Mathematical Sciences, Fudan University, Shanghai 200433, People's Republic of China
E-mail: xyz@fudan.edu.cn

Received 6 September 2019
Accepted for publication 1 December 2019
Published 13 February 2020

Online at stacks.iop.org/JSTAT/2020/023501
https://doi.org/10.1088/1742-5468/ab633d

Abstract. Chaperone-assisted translocation through a nanopore embedded in membrane holds a prominent role in the transport of biopolymers. Inspired by classical *Brownian ratchet*, we develop a theoretical framework characterizing such a translocation process through a master equation approach. In this framework, the polymer chain, provided with reversible binding of chaperones, undergoes forward/backward diffusion, which is rectified by chaperones. We drop the assumption of timescale separation and keep the length of a polymer chain finite, both of which happen to be key points in most of the previous studies. Our framework makes it accessible to derive analytical expressions for mean translocation velocity and an effective diffusion constant in a stationary state, which is the basis of a comprehensive understanding of the dynamics of such a process. Generally, the translocation of polymer chains across a membrane consist of three subprocesses: initiation, termination, and translocation of the main body part of a polymer chain, where the translocation of the main body part depends on both the binding/unbinding kinetics of chaperones and the diffusion of the biopolymer chain. This is the main concern of this study. Our results show that the increase of the forward/backward diffusion rate of a polymer chain and the binding/unbinding ratio of chaperones both raise the mean translocation velocity of a polymer chain, and the mean velocity finally reaches a saturation amount with an extremely rapid diffusion or extremely high binding/unbinding ratio. Roughly speaking, the dependence of effective

1742-5468/20/023501+22$33.00

diffusion constant on these two major processes achieves similar behavior. Besides, longer polymer chains employ higher velocity when the diffusion rate and binding/unbinding ratio are both large and similar results hold for polymer chains that are not too long in terms of the effects on the effective diffusion constant.

Keywords: biomolecules, molecular motors, systems biology

Contents

1. Introduction

Translocation across membranes are ubiquitous during or after the synthesis of biopolymers. A typical example is of proteins that translocate across the endoplasmic reticulum through an aqueous-pore protein-conducting channel [1–5]. Additional examples, including the exportation of RNA across nuclear membranes in eukaryotic cells [6, 7] and the viral injection of DNA into a host [8, 9], both involve translocation through nanoscopic pores embedded in biological membranes. New evidence also indicates that the uptake of long DNA molecules concern the linear passage through the nuclear pore complex [10]. Despite the lack of a membrane-bound nucleus or any membrane-bound organelle, the DNA of bacteria travels across the cell envelope through narrow constrictions during the process of horizontal gene transfer [11–13], such as transformation [14], which is responsible for the adaptive evolution of bacteria. Similar transport mechanisms have also been reported in the biotechnology of drug delivery [15] and rapid DNA sequencing [16].

In recent decades, wide attention has been devoted to understand the mechanism of such crucial biological processes. Pioneering work has established various mechanistic understandings of what induces the directional movement of biopolymers through narrow pores. One important driving force is the external electric field across the membrane, demonstrated by *in vitro* experiments [17, 18] and extensive theoretical analysis

J. Stat. Mech. (2020) 023501

[19–21]. Considering there is usually not a strong enough electric field in living cells, the other mechanism, *Brownian ratchet* suggests pure random thermal diffusion is rectified by chemical asymmetry between the *cis* and *trans* sides of the membrane, giving rise to directional motion [22, 23]. This is the main concern of this paper. A typical case, particularly for proteins, is that the binding of chaperones to the translocating polymer on the *trans* side of a membrane prohibits the polymer chain's backward diffusion to the *cis* side, thus leading to the directional translocation [3, 24, 25]. This mechanism appears to be confirmed on the basis of empirical data for the transport of prepro-α factor into the lumen of endoplasmatic reticulum, where prepro-α-factor is regarded as the protein and BiP serves as a chaperone molecule [26]. Recent progress on the kinetic uptake of DNA also provides evidence for analogous mechanisms [27].

Quantitative investigations devoted to chaperone-assisted transcriptions mainly focus on the dynamic properties of the system, with theoretical explorations mostly setup through a continuous space and time description [22, 23, 28]. Such continuum models can work effectively in the limit that chaperones are much larger compared with the translocating polymer. However, this may become unrealistic in living cells.

Other studies contribute to translocation process of such structure usually through discrete master equations. Early results based on one-dimensional master equations are derived on the simple assumption that both the attachment and the detachment of corresponding chaperones are much faster than the translocation rate, leaving the state of individual sites drawn from the stationary distribution [29, 30]. Dropping this approximation, a more general discrete model is then proposed [31], where the translocating polymer is represented by a one-dimensional lattice of infinite length and an analytical expression for the transcription velocity is obtained mathematically. Extensive analysis was carried out on this model to determine the diffusion constant and even higher cumulants of the translocated length [32], while the same process is discussed from the viewpoint of first-passage time [33]. These three models enjoy a common feature, they always keep track of the state of each binding site that has been translocated to the target region. This approach, although ideal, does not seems feasible, especially for long polymer chains, since the space of states shall experience an explosion as the number of binding sites grows. In this sense, such modeles makes a difference merely on the numerical simulation side, even if explicit expressions for some dynamic quantities can be obtained. To avoid this issue, a fresh study on this problem setup a modified translocation ratchet model involving the state of only a few binding sites [34]. Although a model is proposed dropping the key assumption of timescale separation between the binding/unbinding process of chaperones and the forward/backward diffusion process, analytical results become no longer accessible. Meanwhile, the length of a particular polymer chain is always assumed to be infinite on both sides of the membrane in this model, which is not quite appropriate.

In this paper, based on the classical *Brownian ratchet*, a general framework describing the chaperone-assisted translocation of a polymer chain across a membrane is provided. Since the binding/unbinding kinetics of chaperones might have a sizable effect on the dynamics of the translocation process, as addressed in previous work [31, 32], the state of binding sites located just behind the membrane, i.e. whether it is bound with a chaperone, is taken into account explicitly in our model. Moreover, the usual assumption of infinite length of polymer chain is also rejected. That is to say,

J. Stat. Mech. (2020) 023501

the number of binding sites of a particular polymer chain, or the length of a polymer chain, is assumed to be finite in our model. Through a scheme similar to that in previous investigations [35], the mean velocity and effective diffusion constant of polymer translocation are obtained explicitly.

This paper is organized as follows. On the concept of *Brownian ratchet*, a model describing the process of polymer translocation across membrane is introduced in section 2. How this model works will be explained in detail with coupled master equations. In section 3, we deduce analytical expressions for mean translocation velocity and effective diffusion constant. To show basic properties of the translocation process, numerical results are presented in section 4. Finally, we summarize our results briefly in section 5.

2. Theoretical description of polymer translocation across a membrane through a nanopore

During the translocation across a membrane, a polymer chain moves stochastically to the other side of the pore embedded in the membrane, which is the only possible way that polymers can pass through. Chaperones are assumed to exist in the target region only. Once the binding site of a polymer chain adjacent to the pore is occupied, the polymer chain is prevented from moving backward through the pore due to the large relative size of the chaperone with respect to the diameter of the narrow pore.

A polymer chain of finite length is represented as a one-dimensional lattice with $N+1$ binding sites, labeled $0, 1, 2, \cdots, N$. For convenience, the distance between each two neighboring binding sites is normalized to one. If we take the polymer chain as a reference frame, the translocation process of a polymer chain across membrane through a pore can be regarded as the membrane's stochastic motion along a one-dimensional lattice. We represent the membrane as a wall for simplicity. If the wall is standing at site k and the previous site, labeled $k-1$, remains unoccupied, the wall may hop stochastically either forward to the next site $k+1$ with rate ω_{f_0}, or backward to site $k-1$ with rate ω_b. Otherwise, if the site $k-1$ is occupied with a chaperone, the wall will have no choice but to hop forward stochastically by one step and the corresponding rate is denoted as ω_{f_1}. In summary, the stochastic motion of a membrane along a polymer chain can be divided into two categories, one is the usual diffusion process with possible forward and backward hopping, the other is the rectified diffusion with possible forward hopping alone, since the backward hopping is blocked by chaperones. In this sense, our model can be regarded as a *Brownian ratchet* translocation model.

We assume that chaperones can stochastically bind to any unoccupied site behind the wall with rate ω_a. At the same time, any attached chaperone can detach from its binding sites at random as well, where the detachment rate is denoted as ω_d. It is easy to find that the dynamics of a translocation process depend on the states of all the binding sites of the polymer chain behind the wall. An ideal model, therefore, should include the states of all these binding sites, i.e. whether it is occupied with some chaperone or not. However, keeping all these states explicitly in the model will make further

J. Stat. Mech. (2020) 023501

theoretical analysis inaccessible, especially for long polymer chains, since the state space grows exponentially with the number of binding sites.

Taking this into consideration, we assume that the binding sites of two units or more away from the wall have sufficient time to reach equilibrium, and the probability of staying unoccupied thereby is $q = \omega_d/(\omega_a + \omega_d)$. Actually, the binding/unbinding kinetics of chaperones are much faster than the forward/backward diffusion process of a polymer chain's translocation across the membrane. The ratio of their timescales lies around 1/300 [36], which therefore makes the assumption plausible.

Our model, characterizing the stochastic translocation of a polymer chain across membrane, or equivalently the stochastic motion of a wall along one-dimensional lattice of length N is depicted in figure 1. For $k = 1, \ldots, N$, we denote $P_{k,0}$ and $P_{k,1}$ as the probability that the wall stands at site k with the previous site $k-1$ unoccupied or occupied respectively, see figure 1(a) and (b). The first site of a polymer chain is labeled as 0, and the corresponding probability that the wall stands at this site, site 0, is denoted as P_0. Additionally, we denote Q as the probability that none of the sites of a polymer chain is passing through the pore, i.e. the nanopore of the membrane is vacant. In other words, Q is the probability that a polymer chain has just accomplished its translocation, and a new translocation process has not started yet.

Since chaperones only exist in the target region, only the sites behind the wall are permitted to be bound. Then, the wall, currently standing at site k, can always move forward stochastically to the next site $k+1$, while it gets quite different for backward motions. If the site $k-1$ is unoccupied, the wall at site k can also stochastically move backward to site $k-1$ with rate ω_b. However, if site $k-1$ is occupied with a chaperone, the backward motion of the wall at site k will be blocked. That is to say, the forward motion of the wall standing at site k may depend on the state of site $k-1$. To state a general case, the forward hopping rate of the wall is denoted differently, where ω_{f_0} and ω_{f_1} match the cases that site $k-1$ stays in an unoccupied and occupied state.

To picture the site-by-site translocation process, we can still imagine a wall standing at some site on the main body, labeled as k. As described in the beginning of this section, the probability that site $k-2$ stays in an unoccupied state is q, with this site reaching the equilibrium of binding/unbinding kinetics. Based on a sustained translocation process, we establish a one-dimensional system of period l, where one period indicates a complete transfer of one identical polymer chain. Then, master equations of probabilities $P_{k,0}$ and $P_{k,1}$ in the main body of a lattice, i.e. $k = 2, \cdots, N-1$, are given by

$$\begin{aligned}\frac{\partial P_{k,0}(l,t)}{\partial t} = {} & \omega_{f_0} P_{k-1,0}(l,t) + \omega_{f_1} P_{k-1,1}(l,t) + \omega_d P_{k,1}(l,t) \\ & + q\omega_b P_{k+1,0}(l,t) - (\omega_b + \omega_a + \omega_{f_0}) P_{k,0}(l,t), \end{aligned} \tag{1}$$

$$\frac{\partial P_{k,1}(l,t)}{\partial t} = \omega_a P_{k,0}(l,t) + (1-q)\omega_b P_{k+1,0}(l,t) - (\omega_d + \omega_{f_1}) P_{k,1}(l,t). \tag{2}$$

See figure 1(c) for the schematic depiction.

Since site 0 is the starting point of translocation process, see figure 1(d), there is no lattice site behind the wall if it stands at site 0, indicating that no chaperone can

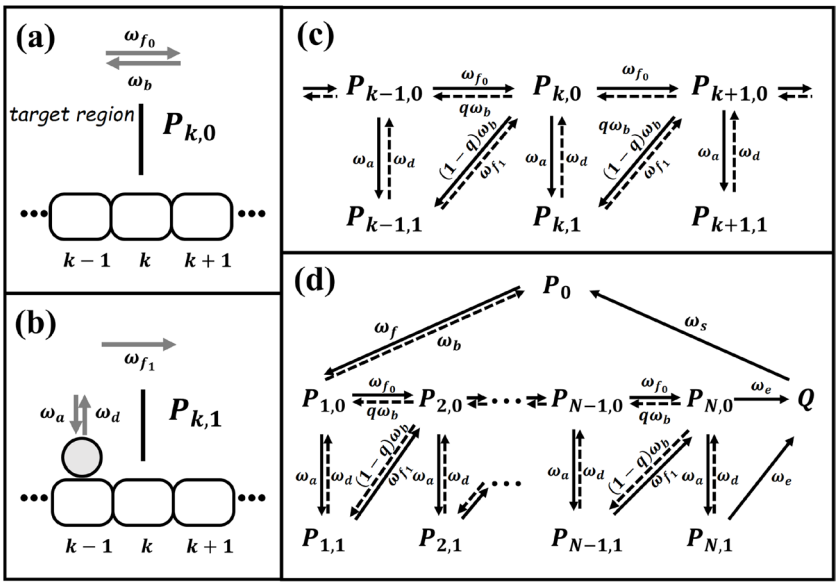

Figure 1. Schematic depiction of the chaperone-assisted translocation of a polymer chain across a membrane. (a) and (b) depict the two possible states included in our model, that is, whether the last site transferred is unoccupied or occupied with a chaperone. The corresponding probabilities are denoted as $P_{k,0}$ and $P_{k,1}$, respectively. Here, the linear polymer molecule is represented by a one-dimension lattice, the bar stands for the moving wall (membrane) and the circle is for a chaperone. The target region of polymer chain is on the left side of the wall, (c) provides a detailed description of possible transitions around site $k\,(2 \leqslant k \leqslant N-1)$, while (d) shows the transitions around boundary sites 0 and N.

be attached to the polymer chain when $k = 0$. Thus, the forward hopping rate of the wall at site 0 may be different from both ω_{f_0} and ω_{f_1}. A different notation ω_f is used to denote the forward hopping rate of the wall located at site 0. As schematically depicted in figure 1(d), equation (2) still holds for $k = 1$ while equation (1) should be modified as

$$\begin{aligned}\frac{\partial P_{1,0}(l,t)}{\partial t} =& \omega_f P_0(l,t) + \omega_d P_{1,1}(l,t) + q\omega_b P_{2,0}(l,t) \\ & - (\omega_b + \omega_a + \omega_{f_0}) P_{1,0}(l,t).\end{aligned} \tag{3}$$

When it comes to the boundary sites, things change. As discussed above, we use Q to denote the probability that the system stays in the interval between two translocation processes of separate polymer chains, and in this state, we can imagine the wall is standing at a fictitious site Q. Translocation starts with rate ω_s when the wall binds to the initial site 0 of a polymer chain, or we say the wall jumps from the fictitious site Q to site 0 with rate ω_s. Similarly, it is when the wall hops forward from the terminal site N to the fictitious site with rate ω_e that the translocation of a polymer chain comes to an end. See figure 1(d) for a schematic depiction. Based on this scenario, the master equation for probability P_0 is as follows,

$$\frac{\partial P_0(l,t)}{\partial t} = \omega_s Q(l,t) + \omega_b P_{1,0}(l,t) - \omega_f P_0(l,t). \tag{4}$$

Similarly, probabilities $P_{N,0}$ and $P_{N,1}$ are governed by

J. Stat. Mech. (2020) 023501

$$\frac{\partial P_{N,0}(l,t)}{\partial t} = \omega_{f_0} P_{N-1,0}(l,t) + \omega_{f_1} P_{N-1,1}(l,t) + \omega_d P_{N,1}(l,t) - (\omega_b + \omega_a + \omega_e) P_{N,0}(l,t), \tag{5}$$

$$\frac{\partial P_{N,1}(l,t)}{\partial t} = \omega_a P_{N,0}(l,t) - (\omega_d + \omega_e) P_{N,1}(l,t). \tag{6}$$

Once a polymer chain has been transferred to the target region completely, the translocation process enters into a pause state, where the wall will wait until it walks onto another polymer chain, that is when another polymer chain starts its translocation. With figure 1(d), the probability Q satisfies

$$\frac{\partial Q(l,t)}{\partial t} = \omega_e [P_{N,0}(l-1,t) + P_{N,1}(l-1,t)] - \omega_s Q(l,t)\,. \tag{7}$$

Hence, we have setup a system of $2N+2$ states, where the translocation of a polymer chain across a nanopore of membrane is mapped into the motion of a wall through a one-dimensional lattice in terms of relative motion.

3. Expressions for mean velocity and effective diffusion constant

In this section, we will generalize the main idea of Derrida [35, 37] to realize analytical expressions for mean velocity and effective diffusion constant of translocation of polymer chains across a membrane in the stationary-state limit, $t \to \infty$. We begin with a general expression for the mean translocation of the wall,

$$\langle x(t) \rangle = \sum_{l=-\infty}^{+\infty} \left[(Nl) Q(l,t) + \sum_{k=0}^{N} (k+Nl) P_k(l,t) \right]. \tag{8}$$

That is to say, if we assume the biopolymer chain is of length N, each translocation period will add N units to the accumulated distance that the wall travels, since hopping away from a polymer chain (translocation *termination*, from site N to pause state, or stepping into a new period (translocation *initiation*, from fictitious site Q to site 1), does not increase the total distance that the wall travels.

Note that, in equation (8) $P_k(l,t) := P_{k,0}(l,t) + P_{k,1}(l,t)$ for $1 \leqslant k \leqslant N$, is the probability that the wall stands at site k of the $l-$th polymer at time t, see figure 1. We now define a few auxiliary functions for each state concerned, they are

$$B_Q(t) = \sum_{l=-\infty}^{+\infty} Q(l,t), \quad C_Q(t) = \sum_{l=-\infty}^{+\infty} (Nl) Q(l,t),$$

$$B_k(t) = \sum_{l=-\infty}^{+\infty} P_{k,0}(l,t), \quad C_k(t) = \sum_{l=-\infty}^{+\infty} (k+Nl) P_{k,0}(l,t), \tag{9}$$

J. Stat. Mech. (2020) 023501

$$\widetilde{B}_k(t) = \sum_{l=-\infty}^{+\infty} P_{k,1}(l,t), \quad \widetilde{C}_k(t) = \sum_{l=-\infty}^{+\infty} (k+Nl) P_{k,1}(l,t). \tag{10}$$

Notably, functions in equation (9) are defined for index k such that $0 \leqslant k \leqslant N$ whereas those in equation (10) are defined for $1 \leqslant k \leqslant N$ only. Here, for convenience, we assume $P_{k,0}(l,t) = P_k(l,t)$ for $k = 0$ in equation (9). Generalizing Derrida's original method [35, 37], we propose the following ansatz in the stationary-state limit, i.e.

$$B_Q(t) \to b_Q, \quad B_k(t) \to b_k, \quad \widetilde{B}_k(t) \to \widetilde{b}_k, \tag{11}$$

$$C_Q(t) \to a_Q t + T_Q, \; C_k(t) \to a_k t + T_k, \quad \widetilde{C}_k(t) \to \widetilde{a}_k t + \widetilde{T}_k. \tag{12}$$

Obviously, all these b_ks together with $\widetilde{b}_k$s are governed by normalization condition

$$b_Q + b_0 + \sum_{k=1}^{N} (b_k + \widetilde{b}_k) = 1, \tag{13}$$

since b_k (or $\widetilde{b}_k$) gives the probability of finding the wall at a specific state at large time.

In the stationary-state limit, (11), that is $\mathrm{d}B_k/\mathrm{d}t = 0$ and $\mathrm{d}\widetilde{B}_k/\mathrm{d}t = 0$ when time t is infinite, master equations given in section 2 are transformed into

$$\begin{aligned}
0 =& \omega_e(b_N + \widetilde{b}_N) - \omega_s b_Q, \\
0 =& \omega_s b_Q + \omega_b b_1 - \omega_f b_0, \\
0 =& \omega_f b_0 + \omega_d \widetilde{b}_1 + q\omega_b b_2 - (\omega_b + \omega_{f_0} + \omega_a) b_1, \\
0 =& \omega_{f_0} b_{k-1} + \omega_{f_1} \widetilde{b}_{k-1} + \omega_d \widetilde{b}_k + q\omega_b b_{k+1} - (\omega_a + \omega_{f_0} + \omega_b) b_k, \\
& \qquad (k = 2, \cdots, N-1), \\
0 =& \omega_{f_0} b_{N-1} + \omega_{f_1} \widetilde{b}_{N-1} + \omega_d \widetilde{b}_N - (\omega_a + \omega_e + \omega_b) b_N, \\
0 =& \omega_a b_k + (1-q)\omega_b b_{k+1} - (\omega_d + \omega_{f_1}) \widetilde{b}_k, \quad (k = 1, \cdots, N-1), \\
0 =& \omega_a b_N - (\omega_d + \omega_e) \widetilde{b}_N.
\end{aligned} \tag{14}$$

These equations lead to a recurrence

$$R b_{k+1} - S b_k = \omega_b b_1 - \omega_f b_0 = -\omega_s b_Q, \quad \text{for} \quad k = 1, \cdots, N-1, \tag{15}$$

where

$$R = \omega_b \frac{\omega_d + q\omega_{f_1}}{\omega_d + \omega_{f_1}}, \qquad S = \omega_{f_0} + \frac{\omega_a \omega_{f_1}}{\omega_d + \omega_{f_1}}. \tag{16}$$

Equation (15) means that the probability-fluxes between each two neighboring *effective* sites remain constant in steady state.

Considering the stationary state for boundary sites, i.e. the first equation and the last equation in equation (14), we have

$$b_N = \left[\omega_e + \frac{\omega_a \omega_e}{\omega_d + \omega_e} \right]^{-1} \omega_s b_0 \equiv K \omega_s b_Q. \tag{17}$$

J. Stat. Mech. (2020) 023501

The recurrence given in equation (15) along with equation (17) yields a general expression for b_k,

$$b_k = \begin{cases} \left[\widetilde{S}^{k-N}[K-(S-R)^{-1}]+(S-R)^{-1}\right]\omega_s b_Q, & (k=1,\cdots,N), \\ \omega_f^{-1}(\omega_b b_1+\omega_s b_Q), & (k=0), \end{cases} \tag{18}$$

where $\widetilde{S} \equiv S/R$.

Thus, all of b_ks and $\widetilde{b}_k$s are determined by $\omega_s b_Q$, which is the probability flux exactly. With the normalizing condition, equation (13), we obtain

$$\omega_s b_Q = \left[W_1 + W_2\widetilde{S}^{1-N} + W_3 N\right]^{-1}, \tag{19}$$

where

$$W_1 = \frac{1}{\omega_s}+\frac{1}{\omega_f}+\frac{U}{S-R}+K\left(\frac{\omega_a}{\omega_d+\omega_e}-\frac{\omega_a}{\omega_d+\omega_{f_1}}\right)+M\frac{S}{S-R}\left(K-\frac{1}{S-R}\right),$$

$$W_2 = \left(K-\frac{1}{S-R}\right)\left(U-\frac{M}{\widetilde{S}-1}\right) \quad \text{and} \quad W_3 = \frac{M}{S-R},$$

with

$$M = 1+\frac{\omega_a}{\omega_d+\omega_{f_1}}+\frac{(1-q)\omega_b}{\omega_d+\omega_{f_1}}, \qquad U = \frac{\omega_b}{\omega_f}-\frac{(1-q)\omega_b}{\omega_d+\omega_{f_1}}.$$

So far, b_Q has been obtained explicitly as a function of all constant rates. So do b_ks and $\widetilde{b}_k$s.

Let us return to the ansatz (12) to derive explicit expressions for a_ks, $\widetilde{a}_k$s, T_ks and $\widetilde{T}_k$s. With the definitions of $C_k(t)$ and $\widetilde{C}_k(t)$, i.e. (9) and (10), as well as the master equations presented in section 2, it is easy to show that $C_k(t)$ and $\widetilde{C}_k(t)$ satisfy

$$\begin{aligned}
\frac{\mathrm{d}C_Q}{\mathrm{d}t}(t) =& \omega_e(C_N+\widetilde{C}_N)-\omega_s C_Q, \\
\frac{\mathrm{d}C_0}{\mathrm{d}t}(t) =& \omega_s C_Q+\omega_b C_1-\omega_f C_0-\omega_b B_1, \\
\frac{\mathrm{d}C_1}{\mathrm{d}t}(t) =& \omega_f C_0+\omega_d\widetilde{C}_1+q\omega_b C_2-(\omega_b+\omega_a+\omega_{f_0})C_1+\omega_f B_0-q\omega_b B_2, \\
\frac{\mathrm{d}C_k}{\mathrm{d}t}(t) =& \omega_{f_0}C_{k-1}+\omega_{f_1}\widetilde{C}_{k-1}+\omega_d\widetilde{C}_k+q\omega_b C_{k+1}-(\omega_b+\omega_a+\omega_{f_0})C_k \\
&+\omega_{f_0}B_{k-1}+\omega_{f_1}\widetilde{B}_{k-1}-q\omega_b B_{k+1}, \quad (k=2,\ldots,N-1), \\
\frac{\mathrm{d}C_N}{\mathrm{d}t}(t) =& \omega_{f_0}C_{N-1}+\omega_{f_1}\widetilde{C}_{N-1}+\omega_d\widetilde{C}_N-(\omega_b+\omega_a+\omega_e)C_N \\
&+\omega_{f_0}B_{N-1}+\omega_{f_1}\widetilde{B}_{N-1}, \\
\frac{\mathrm{d}\widetilde{C}_k}{\mathrm{d}t}(t) =& \omega_a C_k+(1-q)\omega_b C_{k+1}-(\omega_d+\omega_{f_1})\widetilde{C}_k \\
&-(1-q)\omega_b B_{k+1}, \quad (k=1,\ldots,N-1), \\
\frac{\mathrm{d}\widetilde{C}_N}{\mathrm{d}t}(t) =& \omega_a C_N-(\omega_d+\omega_e)\widetilde{C}_N.
\end{aligned} \tag{20}$$

Let time $t \to \infty$ and focus on the terms proportional to time t. Then we get a set of equations analogous to equation (14). Specifically, these equations on a_ks($\widetilde{a}_k$s) are just the same as equation (14), if b_ks($\widetilde{b}_k$s) in equation (14) are all replaced by a_ks($\widetilde{a}_k$s). Thus, one can immediately conclude that

$$a_Q = Ab_Q, \quad a_k = Ab_k, \quad \widetilde{a}_k = A\widetilde{b}_k\,, \tag{21}$$

and

$$a_Q + a_0 + \sum_{k=1}^{N}(a_k + \widetilde{a}_k) = A[b_Q + b_0 + \sum_{k=1}^{N}(b_k + \widetilde{b}_k)] = A\,, \tag{22}$$

where A is a constant. Sum up all of the equations in the stationary-state limit of equation (20) and we have

$$\begin{aligned} A &= \omega_f b_0 + \omega_{f_0}\sum_{k=1}^{N-1} b_k + \omega_{f_1}\sum_{k=1}^{N-1}\widetilde{b}_k - \omega_b\sum_{k=1}^{N} b_k \\ &= \omega_b b_0 - \omega_f b_1 + \sum_{k=1}^{N-1}(Sb_k - Rb_{k+1}) \\ &= N\omega_s b_0, \end{aligned} \tag{23}$$

where the last equation follows the recurrence in equation (15) and b_Q is given in equation (19). Therefore, constant A is determined by all of the constant transition rates. Explicit expressions for a_ks and $\widetilde{a}_k$s can also be obtained via equations (18), (19), (21), (23).

Next, we begin to derive the expressions for T_ks and $\widetilde{T}_k$s. In the stationary-state limit, equation (20) employs the following expressions,

$$\begin{aligned} a_Q &= \omega_e(T_N + \widetilde{T}_N) - \omega_s T_Q, \\ a_0 &= \omega_s T_Q + \omega_b T_1 - \omega_f T_0 - \omega_b b_1, \\ a_1 &= \omega_f T_0 + \omega_d\widetilde{T}_1 + q\omega_b T_2 - (\omega_b + \omega_a + \omega_{f_0})T_1 + \omega_f b_0 - q\omega_b b_2, \\ a_k &= \omega_{f_0}T_{k-1} + \omega_{f_1}\widetilde{T}_{k-1} + \omega_d\widetilde{T}_k + q\omega_b T_{k+1} - (\omega_b + \omega_a + \omega_{f_0})T_k \\ &\quad + \omega_{f_0}b_{k-1} + \omega_{f_1}\widetilde{b}_{k-1} - q\omega_b b_{k+1}, \qquad (k = 2, \ldots, N-1), \\ a_N &= \omega_{f_0}T_{N-1} + \omega_{f_1}\widetilde{T}_{N-1} + \omega_d\widetilde{T}_N - (\omega_b + \omega_a + \omega_e)T_N \\ &\quad + \omega_{f_0}b_{N-1} + \omega_{f_1}\widetilde{b}_{N-1}, \\ \widetilde{a}_k &= \omega_a T_k + (1-q)\omega_b T_{k+1} - (\omega_d + \omega_{f_1})\widetilde{T}_k - (1-q)\omega_b b_{k+1}, \\ &\quad (k = 1, \ldots, N-1), \\ \widetilde{a}_N &= \omega_a T_N - (\omega_d + \omega_e)\widetilde{T}_N. \end{aligned} \tag{24}$$

Here, in the stationary-state limit, attention is paid to the terms independent of time t and thus equation (20) reduces to (24).

With some new auxiliary quantities

$$y_k \equiv \begin{cases} \omega_f T_0 - \omega_b T_1, & \text{for } k = 0 \\ S T_k - R T_{k+1}, & \text{for } k = 1, \cdots, N-1. \end{cases}$$

Equation (24) can be transformed into a recurrence, that is

$$y_k - y_{k-1} = Z_k, \tag{25}$$

and the first term y_0 comes as

$$y_0 = \omega_s T_Q + Z_0. \tag{26}$$

Since we have already obtained the explicit expressions for b_ks as well as a_ks, Z_ks turn to known terms, see equation (A.1).

Through a simple summation, it produces

$$y_k = \omega_s T_Q + r_k, \tag{27}$$

where

$$r_k = \sum_{i=0}^{k} Z_i,$$

and $k = 0, \cdots, N-1$. See equation (A.6) for expressions of r_k.

It means that each y_k consists of two components: one has been determined explicitly and the other depends on the undetermined constant T_Q. In fact, the same holds for each T_k, which can be shown through backward iterations. That is

$$T_k = \begin{cases} K(a_Q + \omega_s T_Q) + (\omega_a + \omega_d + \omega_e)^{-1} \widetilde{a}_N, & \text{for } k = N, \\ S^{-1}(R T_{k+1} + y_k), & \text{for } k = 1, \cdots, N-1, \\ \omega_f^{-1}(\omega_b T_1 + y_0), & \text{for } k = 0. \end{cases} \tag{28}$$

Obviously, T_N and consequently all T_ks are made up of these two components as well, i.e. one has been determined explicitly by constant transition rates and the other depends on the undetermined constant T_Q.

We can now write $T_k = X_k + Y_k$, where X_k is proportional to T_Q and Y_k is described in some determined factors. Particularly, $T_Q = X_Q$ and $Y_Q = 0$. At the same time, $\widetilde{T}_k$s can be derived from the last two equations in equation (24) and they are shown as

$$\widetilde{T}_k = \begin{cases} (\omega_d + \omega_e)^{-1}[\omega_a T_k - \widetilde{a}_k], & \text{for} \quad k = N, \\ (\omega_b + \omega_{f_1})^{-1}[\omega_a T_k + (1-q)\omega_{f_1} T_{k+1} - \widetilde{a}_k - (1-q) b_{k+1}], \\ \qquad\qquad \text{for} \quad k = 1, \cdots, N-1. \end{cases} \tag{29}$$

That is to say, each $\widetilde{T}_k$ can also be split into those two parts and we can then write $\widetilde{T}_k = \widetilde{X}_k + \widetilde{Y}_k$ with $\widetilde{X}_k$ proportional to T_Q and $\widetilde{Y}_k$ determined.

Surprisingly, the undetermined constant T_Q will cancel out in the final expressions for the dynamic properties concerned. Keeping this in mind, we first let $T_Q = 0$ for the calculation of Y_k and show how that works later.

With $T_Q = 0$, that is when both X_k and $\widetilde{X}_k$ equal zero, we find

J. Stat. Mech. (2020) 023501

$$\begin{aligned} Y_N &= Ka_Q + (\omega_a + \omega_d + \omega_e)^{-1}\widetilde{a}_N, \\ Y_k &= \widetilde{S}^{k-N}Y_N + S^{-1}\sum_{i=0}^{N-1-k}\frac{r_{i+k}}{\widetilde{S}^i} \quad \text{for} \quad k = 1, \cdots, N-1, \\ Y_0 &= \omega_f^{-1}[\omega_b Y_1 + Z_0], \\ Y_Q &= 0, \end{aligned} \tag{30}$$

and $\widetilde{Y}_k$ is given by Y_k in the same form as equation (29), i.e.

$$\widetilde{Y}_k = \begin{cases} (\omega_d + \omega_e)^{-1}[\omega_a Y_k - \widetilde{a}_k]\,, & \text{for} \quad k = N, \\ (\omega_d + \omega_{f_1})^{-1}[\omega_a Y_k + (1-q)\omega_{f_1}Y_{k+1} - \widetilde{a}_k - (1-q)b_{k+1}], \\ \text{for} \quad k = 1, \cdots, N-1. \end{cases} \tag{31}$$

Now, we are fully prepared to derive analytical expressions for dynamic properties of interest, i.e. the mean translocation velocity V and effective diffusion constant D in the steady state, which is respectively defined as

$$V = \lim_{t\to+\infty}\frac{\mathrm{d}\langle x(t)\rangle}{\mathrm{d}t}\,, \tag{32}$$

and

$$D = \lim_{t\to+\infty}\frac{1}{2}\frac{\mathrm{d}\left[\langle x^2(t)\rangle - \langle x(t)\rangle^2\right]}{\mathrm{d}t}. \tag{33}$$

With equation (8) and the auxiliary functions given in equations (9) and (10),

$$\frac{\mathrm{d}\langle x(t)\rangle}{\mathrm{d}t} = \frac{\mathrm{d}C_Q(t)}{\mathrm{d}t} + \sum_{k=0}^{N}\frac{\mathrm{d}C_k(t)}{\mathrm{d}t} + \sum_{k=1}^{N}\frac{\mathrm{d}\widetilde{C}_k(t)}{\mathrm{d}t}, \tag{34}$$

which is exactly the summation of equation (20). Recall the derivation of constant A (see equations (21)–(23)), as well as the stationary-state assumption as is given in equation (12), and we can easily get the mean velocity in the stationary-state limit, i.e.

$$V = a_Q + \sum_{k=0}^{N}a_k + \sum_{k=1}^{N}\widetilde{a}_k = A = N\omega_s b_Q\,, \tag{35}$$

where the probability flux $\omega_s b_0$ is given by (19).

An explicit expression for effective diffusion constant D requires more detailed analysis. Utilizing the master equations in section 2, the differentiation of $\langle x^2(t)\rangle$ turns to

$$\frac{1}{2}\frac{\mathrm{d}\langle x^2(t)\rangle}{\mathrm{d}t} = D_1 + D_2\,, \tag{36}$$

where

$$D_1 = \left[\omega_f C_0 + \omega_{f_0}\sum_{k=1}^{N-1}C_k + \omega_{f_1}\sum_{k=1}^{N-1}\widetilde{C}_k - \omega_b\sum_{k=1}^{N}C_k\right], \tag{37}$$

J. Stat. Mech. (2020) 023501

$$D_2 = \frac{1}{2}\left[\omega_f B_0 + \omega_{f_0}\sum_{k=1}^{N-1} B_k + \omega_{f_1}\sum_{k=1}^{N-1}\widetilde{B}_k + \omega_b\sum_{k=1}^{N} B_k\right]. \tag{38}$$

At the same time, the second part of D as given in (33) can be described as

$$D_3 \equiv \frac{1}{2}\frac{\mathrm{d}\langle x(t)\rangle^2}{\mathrm{d}t} = V\langle x(t)\rangle = A\left[C_Q + \sum_{k=0}^{N} C_k + \sum_{k=1}^{N}\widetilde{C}_k\right]. \tag{39}$$

In the stationary-state limit, we first claim that terms proportional to time t in the expression for D cancel out. Since $B_k(\widetilde{B}_k)$ tends to be a constant $b_k(\widetilde{b}_k)$ as time t goes infinity, we just concentrate on equations (37) and (39) to verify it. Replacing $C_k(\widetilde{C}_k)$ with $a_k(\widetilde{a}_k)$ in equation (37), one can easily find that the coefficient of t in D_1 comes as $A\cdot N\omega_s b_Q = A^2$ in comparison with equation (23). Similarly, the coefficient of t in D_3 is

$$A\left[a_Q + \sum_{k=0}^{N} a_k + \sum_{k=1}^{N}\widetilde{a}_k\right] = A^2.$$

Then, we focus on the terms proportional to T_Q in order to show that the effective diffusion constant D is independent of them. As mentioned before, each T_k and $\widetilde{T}_k$ can be written as $T_k = X_k + Y_k$ and $\widetilde{T}_k = \widetilde{X}_k + \widetilde{Y}_k$, respectively. With a glance at equations (27) and (28), we find X_k follows the same recurrence as b_k does. In other words, every b_k in equation (15) can be replaced with X_k. So does $\widetilde{X}_k$ with respect to $\widetilde{b}_k$. Thus, we have

$$X_k = \frac{T_Q}{b_Q}b_k, \qquad \widetilde{X}_k = \frac{T_Q}{b_Q}\widetilde{b}_k. \tag{40}$$

This can be better understood if the procedure of seeking b_k and $\widetilde{b}_k$ is regarded as solving linear equations $\boldsymbol{\Omega x} = \mathbf{0}$ corresponding to equation (14), where $\boldsymbol{\Omega}$ is a matrix of $2(N+1)\times 2(N+1)$ in size. Since the rank of $\boldsymbol{\Omega}$ is $2N+1$, $\boldsymbol{b} = (b_Q, b_0, b_1, \widetilde{b}_1, \cdots, b_N, \widetilde{b}_N)^{\intercal}$, what we obtained in the foregoing, is determined on the normalizing condition. If we denote $\boldsymbol{T}$ as the vector $(T_Q, T_0, T_1, \widetilde{T}_1, \cdots, T_N, \widetilde{T}_N)^{\intercal}$, one can find the derivation of $\boldsymbol{T}$ is just solving the linear equation $\boldsymbol{\Omega x} = \boldsymbol{\beta}$ due to equation (24), where each component in $\boldsymbol{\beta}$ is determined by $b_k(\widetilde{b}_k)$ and $a_k(\widetilde{a}_k)$. Actually, we have already obtained a solution for this equation, which is $\boldsymbol{Y} = (Y_Q, Y_0, Y_1, \widetilde{Y}_1, \cdots, Y_N, \widetilde{Y}_N)^{\intercal}$ with each component given in equations (30) and (31). Then, a general solution $\boldsymbol{T}$ can be written as $\boldsymbol{T} = \alpha\boldsymbol{b} + \boldsymbol{Y}$. Particularly, the first component in $\boldsymbol{Y}$ is zero, i.e. $Y_Q = 0$. If we regard T_Q as a undetermined constant, the undetermined constant should be restricted to $\alpha = T_Q/b_Q$ in the consideration of the first component in $\boldsymbol{T}$, $\boldsymbol{Y}$ and $\boldsymbol{b}$. Equation (40) holds therefore. That is exactly what we have done.

In this way, considering whether the terms proportional to T_Q in equations (37) and (39) can cancel out is to examine exactly the performance of X_ks. It is easy to see such terms in the former equation sum up to $N\omega_s T_Q$, if we take X_k as a substitution for C_k in equation (37) and compare the coefficients with those in equation (23). The latter

J. Stat. Mech. (2020) 023501

equation (39), leads to a summation of X_ks and $\widetilde{X}_k$s, which is $N\omega_s T_Q$ as well. As a result, terms proportional to T_Q cancel out.

Due to these two cancellations, we finally derive the explicit expression for effective diffusion constant in the stationary-state limit,

$$\begin{aligned} D &= D_1 + D_2 - D_3 \\ &= -\sum_{k=0}^{N-1} s_k - \sum_{k=1}^{N-1} \widetilde{s}_k + \frac{V}{2}N - V\left[Y_0 + \sum_{k=1}^{N}(Y_k + \widetilde{Y}_k)\right], \end{aligned} \tag{41}$$

where Y_k has already been obtained in equation (30), $\widetilde{Y}_k$ follows equation (31) and

$$s_k = \sum_{i=1}^{k} a_i\,, \quad \widetilde{s}_k = \sum_{i=2}^{k} \widetilde{a}_i.$$

See appendix A for more details.

4. Properties of polymer translocation across membrane

To further understand the analytical results and get an intuition about the translocation process of polymer chains across a membrane, extensive numerical explorations are carried out according to explicit results given in equations (35) and (41). Instead of general analysis with respect to each transition rate, particular attention is paid to a special case where $\omega_{f_0} = \omega_{f_1} = \omega_f = \omega_b \equiv w$, due to the fact that the forward/backward diffusion of a polymer chain essentially comes from Brownian fluctuations. This means that it is the binding of chaperones that rectifies the diffusion of polymer chain to a directional motion on average. In fact, results remain valid even if the polymer chain undergoes a biased random walk when crossing the membrane. Note that both the binding and unbinding of chaperones are extremely fast compared to the forward/backward diffusion of a polymer chain. In all calculations below, we always keep the binding/unbinding rate of a chaperone at least 2 orders of magnitude higher than the forward/backward diffusion rate.

Numerical results on mean translocation velocity V are displayed in figure 2. When trailing binding sites (i.e. binding sites of a polymer chain lying in the target region) are occupied with chaperones in large probability, the polymer chain with a larger forward/backward diffusion rate will pass through the pore more rapidly, since its backward motion is hindered by chaperones. So translocation velocity increases with the diffusion rate w. This rise, however, will not last indefinitely, since for sufficiently large diffusion rate w, the translocation velocity will be limited by the binding/unbinding process of chaperones as well as the finite starting rate ω_s and ending rate ω_e. See figure 2(a).

It means that there is an upper limit for mean velocity V as diffusion rate w tends to infinity, see figure 2(a). Besides, we notice that a longer polymer chain employs greater mean velocity, provided large diffusion rate w and high ratio ω_a/ω_d. This matches the fact that the pretty slow initiation process and termination process, acting as restrictions, produce greater impact on the overall mean velocity of a shorter polymer chain.

J. Stat. Mech. (2020) 023501

From another perspective, we consider the mean dwell time t for some translocation, which is

$$t = \frac{N}{V} = W_1 + W_2 \widetilde{S}^{1-N} + W_3 N\,. \tag{42}$$

Here, the numerator N is the length of a polymer chain and the denominator V is the mean velocity of translocation. It reflects the effect of binding sites' number on polymers of fixed length. According to the definition of constant S, if the diffusion rate w is low, S increases almost linearly with w, while W_1, W_2, W_3 are all insensitive to the low diffusion rate w. Therefore, equation (42) indicates that the relation between dwell time t and diffusion rate w follows power law approximately, when the polymer chain diffuses at a low diffusion rate w. See figure 2(b).

The dependence of mean velocity V on the ratio of binding/unbinding rate of a chaperone is presented in figure 2(c). With high ratio ω_a/ω_d, the binding site lying in the target region will be more likely to be occupied with a chaperone, which will, then, help rectify the forward/backward diffusion. That is, high binding/unbinding ratio speeds up the translocation of a polymer chain. The dependence of mean velocity on ratio ω_a/ω_d displays similar performance to that on diffusion rate w, as is shown in figures 2(a) and (c). If the ratio ω_a/ω_d is large enough, mean translocation velocity V increases with the length N of a polymer chain and vice versa. See figure 2(c) and the inset. This is because the longer a polymer chain is, the less influence is likely to be exerted through the relatively low starting rate ω_s and ending rate ω_e. Conversely, low ratio ω_a/ω_d leads to the slow translocation of the main body part of a polymer chain, and the relatively large starting rate ω_s as well as ending rate ω_e plays a more active role in promoting the overall mean translocation velocity V of a shorter polymer chain. It means that the mean velocity will decrease with the growth of the polymer length N.

Note that, experimental binding/unbinding ratio is approximate to 833.3 [27]. Figure 2(c) shows that the velocity V is almost independent of ratio ω_a/ω_d when $\omega_a/\omega_d \geqslant 833.3$, implying that the translocation process is limited by the forward/backward diffusion rate w as well as starting rate ω_s and ending rate ω_e. Moreover, the plots in figures 2(a) and (b) show that, for diffusion rate w larger than 1000 s^{-1}, mean velocity V is also insensitive to the change of rate w. Therefore, the translocation of polymer chain is limited merely by the starting and the ending process when $\omega_a/\omega_d \geqslant 833.3$ and diffusion rate $w \geqslant 1000\ \mathrm{s}^{-1}$. Obviously, velocity V increases with both rate ω_s and rate ω_e, see figure 1(d) or equations (19), (35). Note that, in figures 2(a) and (b), the ratio ω_a/ω_d is always kept 833.3.

The dependence of dwell time t on length N of a polymer chain, with different values of ratio ω_a/ω_d, is plotted in figure 2(d). One can easily find that t increases almost linearly with N, especially for large binding/unbinding ratio ω_a/ω_d. It can been seen in equation (42) as well. Meanwhile, figure 2(d) also shows that dwell time t decreases with the ratio ω_a/ω_d, since a polymer chain requires high mean velocity V provided with a great binding/unbinding ratio ω_a/ω_d. See figure 2(c) and the definition of t given in equation (42).

The discussion above on mean velocity V provides a chief description of the dynamics of polymer translocation across a membrane, but it is not sufficient for many biologically relevant cases, especially when the polymer chain is not long. In these situations,

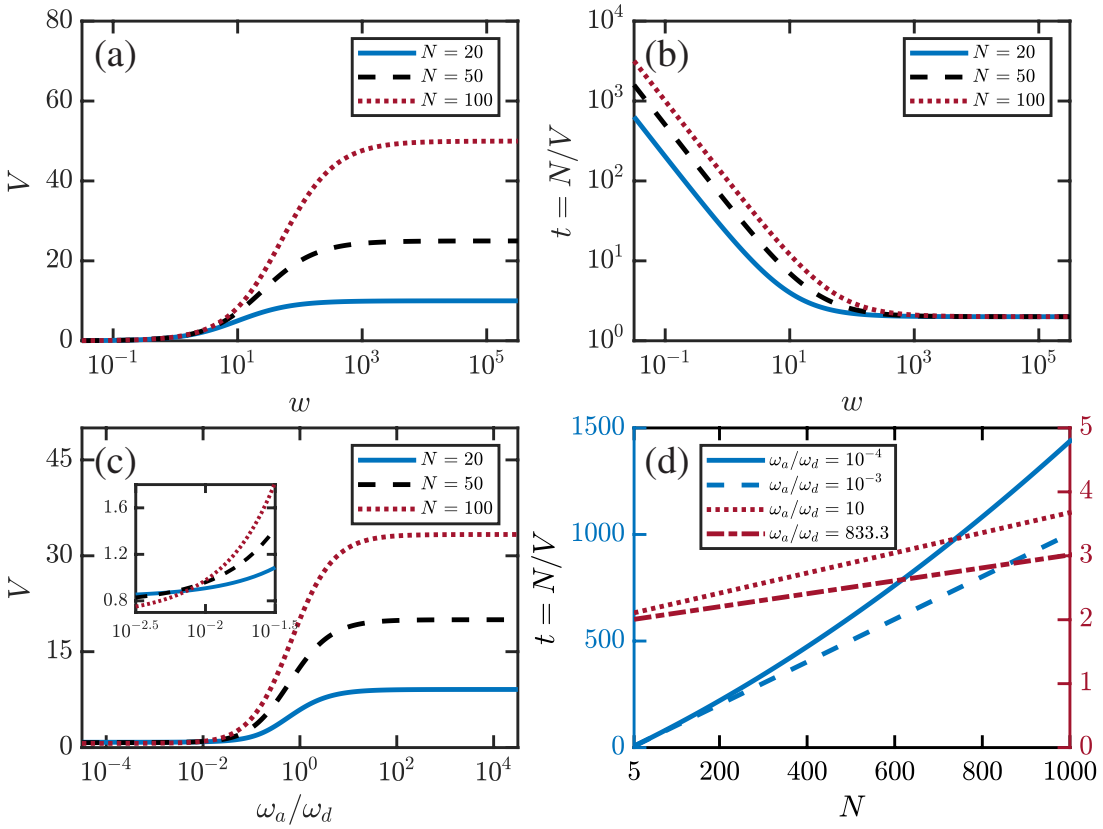

Figure 2. Dependence of (a) mean velocity V and (b) mean dwell time t on forward/backward diffusion rate w, where the fast binding rate $\omega_a = 1000 \times 10^8$ s^{-1} and the unbinding rate $\omega_d = 1.2 \times 10^8$ s^{-1}, keeping their ratio $\omega_d/\omega_a = 0.0012$ [27]. (c) Mean velocity V as a function of binding/unbinding ratio ω_a/ω_d with the inset displaying the detail in enlarged scale. (d) Mean dwell time t of translocation as a function of the length of polymer chain with different ratio of the binding/unbinding rate. Left axis: solid line and dashed line. Right axis: dotted line and dash dotted line. In (c) and (d), the diffusion rate is kept $w = 10^2$ s^{-1} and the unbinding rate ω_d is kept of order 10^8. Thus, the binding/unbinding process is more than 2 orders of magnitude faster than the diffusion process. Other parameters used in calculations are $\omega_s = \omega_e = 1$ s^{-1}.

effective diffusion constant D, or say dispersion, plays an important role. The plots in figure 3(a) show that, for high diffusion rate w, diffusion constant D increases with both the ratio ω_a/ω_d and the polymer length N, when the polymer is not long. At the same time, diffusion constant D tends to its lower limit or upper limit monotonically as the ratio ω_a/ω_d tends to 0 or infinity, respectively. However, something interesting happens as diffusion rate w goes down approaching the starting rate ω_s and ending rate ω_e. The plots in figure 3(b) show that, for such cases, D will no longer change monotonically with rate ω_a/ω_d. Neither will it increase with the polymer length N. With a particular value of ratio $\omega_a/\omega_d = 833.3$ [27], the dependence of D on diffusion rate w is plotted in figure 3(c), where the effective diffusion constant D increases with w as expected. Besides, the saturation level for effective diffusion constant D rises with the polymer length N, as long as the polymer chain is not too long.

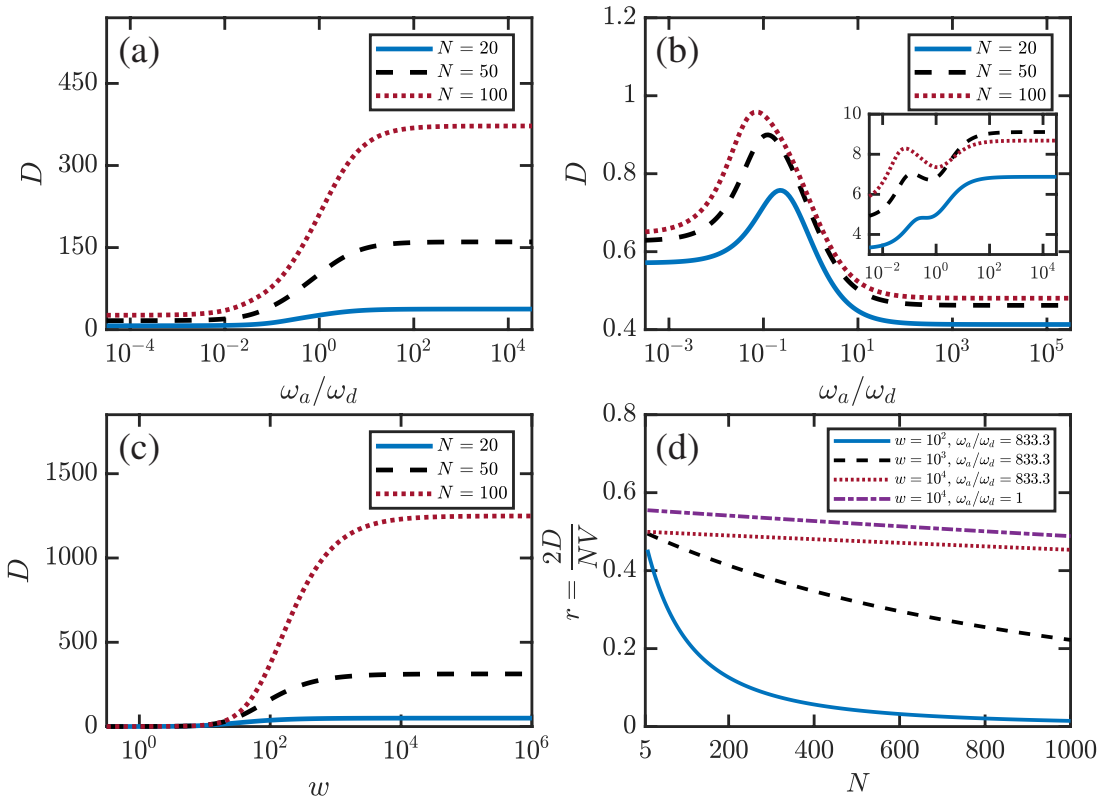

Figure 3. Effective diffusion constant D as a function of the ratio ω_a/ω_d is displayed in (a) and (b). The diffusion rate used in (a) is $w = 10^2$ s^{-1}, while calculations in (b) are carried out with lower diffusion rates. $w = 1$ s^{-1} in (b) and $w = 10$ s^{-1} in inset. (c) Dependence of D on the diffusion rate w, with ratio $\omega_a/\omega_d = 833.3$ as is given in [27]. (d) The randomness, $r \equiv 2D/NV$, as a function of the length N of a polymer chain. Several cases with different values of diffusion rate w and ratio ω_a/ω_d are considered. Detailed instructions on parameters are listed in the legend. Similar to figure 2, ω_d is held of order 10^8 s^{-1} to make the binding/unbinding process of a chaperone much faster than the forward/backward diffusion process of a polymer chain. Other parameters used in calculations are $\omega_s = 1$ s^{-1} and $\omega_e = 1$ s^{-1}.

Compare figures 2(a), (c) with 3(a), (c), and we can find that mean velocity V and effective diffusion constant D show almost the same behavior. Both of them increase significantly when the parameter, either forward/backward diffusion rate w or binding/unbinding ratio ω_a/ω_d, varies within some range. With the parameter varying beyond this range, however, mean velocity V and effective diffusion constant D both stay almost constant. The reason is that the dynamics of translocation process is regulated jointly by several factors. Besides diffusion rate w along with the ratio ω_a/ω_d, these factors include starting rate ω_s and ending rate ω_e. See figure 1. When the corresponding parameter, rate w or ω_a/ω_d, varies within an appropriate range, the translocation process of a polymer is dominated by its diffusion process and the binding/unbinding process of a chaperone, and the mean velocity or diffusion constant changes significantly with the parameter, rate w or ratio ω_a/ω_d. Otherwise, the dynamics of translocation

J. Stat. Mech. (2020) 023501

will be limited by rate ω_s and ω_e, and therefore is almost independent of the rate w and ratio ω_a/ω_d. In other words, the translocation of a polymer chain across a membrane consists of several kinetics, namely initiation, termination, binding/unbinding of chaperones and forward/backward diffusion of a polymer chain. For a given set of parameters, the overall translocation process may be limited by only one or some of them, with others fast enough to be neglected. But the limit process may be switched from one to another, with the change of one or some parameter values.

It is also noteworthy to compare the degree of fluctuations of this stochastic process. An important dimensionless function to evaluate this quantity is randomness, which is defined as [37, 38]

$$r = \frac{2D/N^2}{V/N} = \frac{2D}{NV}.$$

The plots in figure 3(d) show that the randomness r always decreases with the polymer length N. With the ratio $\omega_a/\omega_d = 833.3$ [27], r increases with the forward/backward diffusion rate w. Our results also suggest a decrease of r with the number of binding sites N. A rapid decay in randomness r can be observed on the condition of a low diffusion rate w, whereas the length of a polymer reduces r slightly, but almost linearly, as long as the diffusion rate w is high.

5. Conclusions and remarks

The *Brownian ratchet* builds a general framework describing how diffusive motion is rectified by chemical energy. Particularly, the chaperone-assisted translocation of polymer chains across a membrane constitutes a fine example of how directed motion can emerge from random diffusion, where the binding and unbinding of chaperones plays a prominent role.

In this paper, a theoretical model is presented mapping the process of translocation across membrane into the discrete motions of a membrane on the polymer chain, where the probability of finding membrane at some site of polymer chain is governed by usual master equations. With this model, the mean velocity and effective diffusion constant of translocation process can be derived explicitly, allowing us to discuss the dynamics of translocation across membrane much more efficiently.

Based on the exact expressions, detailed discussions on particular cases are presented, where the polymer chain is assumed to diffuse freely around the nanopore embedded in a membrane if there is no chaperone molecule. Our results show that both the increase of forward/backward diffusion rate w of polymer chains and the rise in binding/unbinding rate ratio ω_a/ω_d of chaperones raise the mean translocation velocity monotonically to a polymer length dependent limit. With large diffusion rate w or great ratio ω_a/ω_d, the mean velocity increases with the length of polymer chain, while opposite performance can be observed when w or ω_a/ω_d stays at a low level. The effective diffusion constant also increases with the diffusion rate w, but it will exhibit complicated properties when regarded as a function of ratio ω_a/ω_d. Meanwhile, the randomness decreases monotonically with the length of a polymer chain.

J. Stat. Mech. (2020) 023501

For most of biological processes in living cells, mean velocity together with effective diffusion constant is usually sufficient to describe the basic dynamic properties. Nevertheless, higher order moments, or higher cumulants, can be obtained as well, simply via the same methods in this paper. Meanwhile, all of the transition rates are assumed to be independent of the site of a polymer chain in our discussion. But in a promoted model with site-dependent transition rates, explicit expressions for mean velocity and effective diffusion constant can still be obtained sketching the key idea of this study.

Appendix. Derivation of effective diffusion constant equation (41)

Since the final expression for diffusion constant D does not depend on the undetermined constant T_Q, we let $T_Q = 0$, which indicates that both X_ks and $\widetilde{X}_k$s are equal to zero. At the same time, Y_k follows equation (30) and $\widetilde{Y}_k$ follows (31).

With equation (24),

$$\begin{aligned} Z_0 &= -a_0 - \omega_b b_1, \\ Z_1 &= -a_1 - \frac{\omega_d}{\omega_d + \omega_{f_1}}\widetilde{a}_1 - Rb_2 + \omega_f b_0, \\ Z_k &= -a_k - \frac{\omega_{f_1}}{\omega_d + \omega_{f_1}}\widetilde{a}_{k-1} - \frac{\omega_d}{\omega_d + \omega_{f_1}}\widetilde{a}_k - Rb_{k+1} + Sb_{k-1}, \\ &\quad (k = 2, \cdots, N-1). \end{aligned} \tag{A.1}$$

In the stationary-state limit, equations (37)–(39) turn to

$$D_1 = \omega_f Y_0 + \omega_{f_0} \sum_{k=1}^{N-1} Y_k + \omega_{f_1} \sum_{k=1}^{N-1} \widetilde{Y}_k - \omega_b \sum_{k=1}^{N} Y_k, \tag{A.2}$$

$$D_2 = \frac{1}{2}\left[\omega_f b_0 + \omega_{f_0} \sum_{k=1}^{N-1} b_k + \omega_{f_1} \sum_{k=1}^{N-1} \widetilde{b}_k + \omega_b \sum_{k=1}^{N} b_k\right], \tag{A.3}$$

and

$$-D_3 = -A\left[Y_0 + \sum_{k=1}^{N}(Y_k + \widetilde{Y}_k)\right]. \tag{A.4}$$

Take equation (29) into equation (A.2),

$$\begin{aligned} D_1 &= (\omega_f Y_0 - \omega_b Y_1) + \sum_{k=1}^{N-1}(SY_k - RY_{k+1}) - \frac{\omega_{f_1}}{\omega_d + \omega_{f_1}} \sum_{k=1}^{N-1} \widetilde{a}_k \\ &\quad - (1-q)\omega_b \frac{\omega_{f_1}}{\omega_d + \omega_{f_1}} \sum_{k=1}^{N-1} b_{k+1}, \\ &= -\frac{\omega_{f_1}}{\omega_d + \omega_{f_1}} \sum_{k=1}^{N-1} \widetilde{a}_k - (1-q)\omega_b \frac{\omega_{f_1}}{\omega_d + \omega_{f_1}} \sum_{k=1}^{N-1} b_{k+1} + \sum_{k=0}^{N-1} r_k. \end{aligned} \tag{A.5}$$

J. Stat. Mech. (2020) 023501

With the definition of r_k, we have

$$\begin{aligned} r_0 =& Z_0 = -a_0 - \omega_b b_1, \\ r_1 =& \sum_{i=0}^{1} Z_i = -\sum_{i=0}^{1} a_i - \frac{\omega_{f_1}}{\omega_d + \omega_{f_1}} \widetilde{a}_1 - R b_2 + \omega_s b_Q \,, \\ r_k =& \sum_{i=0}^{k} Z_i = -\sum_{i=0}^{k} a_i - \sum_{i=1}^{k-1} \widetilde{a}_i - \frac{\omega_{f_1}}{\omega_d + \omega_{f_1}} \widetilde{a}_k - R b_{k+1} + k\omega_s b_Q, \\ & (k = 2, \cdots, N-1). \end{aligned} \tag{A.6}$$

Then,

$$\begin{aligned} \sum_{k=0}^{N-1} r_k =& -\frac{\omega_d}{\omega_d + \omega_{f_1}} \sum_{k=1}^{N-1} \widetilde{a}_k - R \sum_{k=1}^{N} b_k - \omega_b b_1 \\ & + \omega_s b_Q \sum_{k=1}^{N-1} k - \sum_{k=0}^{N-1} \sum_{i=0}^{k} a_i - \sum_{k=1}^{N-2} \sum_{i=1}^{k} \widetilde{a}_i \,. \end{aligned} \tag{A.7}$$

Note that the summation of the first term in equation (A.5) and that in equation (A.7) is

$$-\frac{\omega_{f_1}}{\omega_d + \omega_{f_1}} \sum_{k=1}^{N-1} \widetilde{a}_k - \frac{\omega_d}{\omega_d + \omega_{f_1}} \sum_{k=1}^{N-1} \widetilde{a}_k = -\sum_{i=1}^{N-1} \widetilde{a}_i \,,$$

and the summation of terms related to b_k in these two equations equals

$$-\omega_b \sum_{k=1}^{N} b_k + \frac{1}{2} N(N-1)\omega_s b_Q.$$

Then

$$D_1 = -\sum_{k=0}^{N-1} s_k - \sum_{k=1}^{N-1} \widetilde{s}_k - \omega_b \sum_{k=1}^{N} b_k + \frac{1}{2} N(N-1)\omega_s b_Q \,, \tag{A.8}$$

where

$$s_k = \sum_{i=0}^{k} a_i, \quad \widetilde{s}_k = \sum_{i=1}^{k} \widetilde{a}_i.$$

For equation (A.3),

$$\begin{aligned} D_2 =& \frac{1}{2} \left[\omega_f b_0 + \omega_{f_0} \sum_{k=1}^{N-1} b_k + \omega_{f_1} \sum_{k=1}^{N-1} \widetilde{b}_k - \omega_b \sum_{k=1}^{N} b_k + 2\omega_b \sum_{k=1}^{N} b_k \right] \\ =& \omega_b \sum_{k=1}^{N} b_k + \frac{1}{2} V. \end{aligned} \tag{A.9}$$

With (35), the summation of equations (A.8), (A.9) and (A.4) can be written as

J. Stat. Mech. (2020) 023501

$$D = D_1 + D_2 - D_3$$
$$= -\sum_{k=0}^{N-1} s_k - \sum_{k=1}^{N-1} \widetilde{s}_k + N\frac{V}{2} - V\left[Y_0 + \sum_{k=1}^{N}(Y_k + \widetilde{Y}_k)\right], \tag{A.10}$$

which is exactly equation (41).

References

[1] Simon S M and Blobel G 1991 A protein-conducting channel in the endoplasmic reticulum *Cell* **65** 371–80
[2] Simon S M and Blobel G 1992 Signal peptides open protein-conducting channels in *E. coli* *Cell* **69** 677–84
[3] Neupert W and Brunner M 2002 The protein import motor of mitochondria *Nat. Rev. Mol. Cell Biol.* **3** 555–65
[4] Rapoport T A 2007 Protein translocation across the eukaryotic endoplasmic reticulum and bacterial plasma membranes *Nature* **450** 663–9
[5] Neupert W 2015 A perspective on transport of proteins into mitochondria: a myriad of open questions *J. Mol. Biol.* **427** 1135–58
[6] Schatz G and Dobberstein B 1996 Common principles of protein translocation across membranes *Science* **271** 1519–26
[7] Santos-Rosa H, Moreno H, Simos G, Segref A, Fahrenkrog B, Panté N and Hurt E 1998 Nuclear mRNA export requires complex formation between Mex67p and Mtr2p at the nuclear pores *Mol. Cell Biol.* **18** 6826–38
[8] Albert B, Johnson A, Lewis J, Raff M, Roberts K and Walter P 2008 *Molecular Biology of the Cell* 5th edn (New York: Garland Science)
[9] Liu S, Chistol G, Hetherington C L, Tafoya S, Aathavan K, Schnitzbauer J, Grimes S, Jardine P J and Bustamante C 2014 A viral packaging motor varies its DNA rotation and step size to preserve subunit coordination as the capsid fills *Cell* **157** 702–13
[10] Salman H, Zbaida D, Rabin Y, Chatenay D and Elbaum M 2001 Kinetics and mechanism of DNA uptake into the cell nucleus *Proc. Natl Acad. Sci. USA* **98** 7247–52
[11] Chen I, Christie P J and Dubnau D 2005 The ins and outs of DNA transfer in bacteria *Science* **310** 1456–60
[12] Allemand J F and Maier B 2009 Bacterial translocation motors investigated by single molecule techniques *FEMS Microbiol. Rev.* **33** 593–610
[13] Burton B and Dubnau D 2010 Membrane-associated DNA transport machines *Cold Spring Harbin Perspect. Biol.* **2** a000406
[14] Chen I and Dubnau D 2004 DNA uptake during bacterial transformation *Nat. Rev. Microbiol.* **2** 241–9
[15] Holowka E P, Sun V Z, Kamei D T and Deming T J 2007 Polyarginine segments in block copolypeptides drive both vesicular assembly and intracellular delivery *Nat. Mater.* **6** 52–7
[16] Turner S W P, Cabodi M and Craighead H G 2002 Confinement-induced entropic recoil of single DNA molecules in a nanofluidic structure *Phys. Rev. Lett.* **88** 128103
[17] Kasianowicz J J, Brandin E, Branton D and Deamer D W 1996 Characterization of individual polynucleotide molecules using a membrane channel *Proc. Natl Acad. Sci. USA* **93** 13770–3
[18] Meller A, Nivon L, Brandin E, Golovchenko J and Branton D 2000 Rapid nanopore discrimination between single polynucleotide molecules *Proc. Natl Acad. Sci. USA* **97** 1079–84
[19] Sung W and Park P J 1996 Polymer translocation through a pore in a membrane *Phys. Rev. Lett.* **77** 783–6
[20] Meller A, Nivon L and Branton D 2001 Voltage-driven DNA translocations through a nanopore *Phys. Rev. Lett.* **86** 3435–8
[21] Corsi A, Milchev A, Rostiashvili V G and Vilgis T A 2006 Field-driven translocation of regular block copolymers through a selective liquid–liquid interface *Macromolecules* **39** 7115–24
[22] Simon S M, Peskin C S and Oster G F 1992 What drives the translocation of proteins? *Proc. Natl Acad. Sci. USA* **89** 3770–4
[23] Peskin C S, Odell G M and Oster G F 1993 Cellular motions and thermal fluctuations: the Brownian ratchet *Biophys. J.* **65** 316–24
[24] Matlack K E S, Mothes W and Rapoport T A 1998 Protein translocation: tunnel vision *Cell* **92** 381–90
[25] Elston T C 2000 Models of post-translational protein translocation *Biophys. J.* **79** 2235–51
[26] Matlack K E S, Misselwitz B, Plath K and Rapoport T A 1999 BiP acts as a molecular ratchet during post-translational transport of prepro-α factor across the ER membrane *Cell* **97** 553–64

J. Stat. Mech. (2020) 023501

[27] Hepp C and Maier B 2016 Kinetics of DNA uptake during transformation provide evidence for a translocation ratchet mechanism *Proc. Natl Acad. Sci. USA* **113** 12467–72
[28] Zandi R, Reguera D, Rudnick J and Gelbart W M 2003 What drives the translocation of stiff chains? *Proc. Natl Acad. Sci. USA* **100** 8649–53
[29] Ambjörnsson T and Metzler R 2004 Chaperone-assisted translocation *Phys. Biol.* **1** 77–88
[30] Ambjörnsson T, Lomholt M A and Metzler R 2005 Directed motion emerging from two coupled random processes: translocation of a chain through a membrane nanopore driven by binding proteins *J. Phys.: Condens. Matter* **17** S3945–64
[31] D'Orsogna M R, Chou T and Antal T 2007 Exact steady-state velocity of ratchets driven by random sequential adsorption *J. Phys. A: Math. Theor.* **40** 5575–84
[32] Krapivsky P L and Mallick K 2010 Fluctuations in polymer translocation *J. Stat. Mech.* P07007
[33] Abdolvahab R H, Metzler R and Ejtehadi M R 2011 First passage time distribution of chaperone driven polymer translocation through a nanopore: homopolymer and heteropolymer cases *J. Chem. Phys.* **135** 245102
[34] Uhl M and Seifert U 2018 Force-dependent diffusion coefficient of molecular Brownian ratchets *Phys. Rev.* E **98** 022402
[35] Derrida B 1983 Velocity and diffusion constant of a periodic one-dimensional hopping model *J. Stat. Phys.* **31** 433–50
[36] Gerland U, Moroz J D and Hwa T 2002 Physical constraints and functional characteristics of transcription factor-DNA interaction *Proc. Natl Acad. Sci. USA* **99** 12015–20
[37] Kolomeisky A B and Phillips H 2005 Dynamic properties of motor proteins with two subunits *J. Phys.: Condens. Matter* **17** S3887–99
[38] Svoboda K, Mitra P P and Block S M 1994 Fluctuation analysis of motor protein movement and single enzyme kinetics *Proc. Natl Acad. Sci. USA* **91** 11782–6

J. Stat. Mech. (2020) 023501