THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

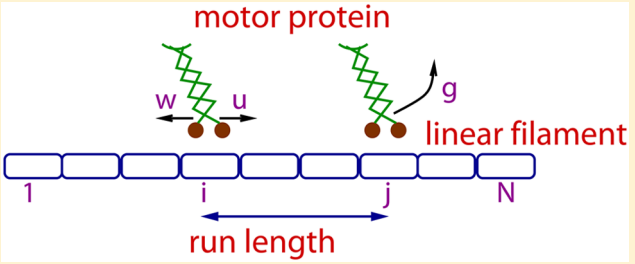
Theoretical Investigation of Distributions of Run Lengths for Biological Molecular Motors

Yunxin Zhang[†] and Anatoly B. Kolomeisky*[,‡]

[†]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

[‡]Department of Chemistry and Center for Theoretical Biological Physics, Rice University, Houston, Texas 77005, United States

ABSTRACT: Motor proteins, also known as biological molecular motors, are active enzymatic molecules that support a variety of fundamental biological processes, including cellular transport, cell division, and cell motility. They usually function by transforming chemical energy into mechanical motion, which propels them along linear structures such as protein filaments and nucleic acids. Recent single-molecule experiments measured with high-precision distributions of various dynamic properties of molecular motors. However, it is difficult to utilize these observations to obtain a better description of molecular mechanisms in motor proteins because of the lack of corresponding theoretical methods. To fill this gap, we developed a new theoretical framework to describe the distributions of dynamic properties of biological molecular motors. It is based on the method of first-passage processes. To illustrate our approach, the distributions of run lengths of motor proteins are analyzed. It is found that these distributions depend on the finite length of linear tracks along which the motors move, on the initial position of the motor proteins along the filaments, and on the intermediate chemical transitions during the enzymatic cycle. The physical mechanisms of the observed phenomena are discussed.



INTRODUCTION

Motor proteins represent important classes of enzymatic molecules that play crucial roles in supporting and maintaining major biological processes such as cellular transport, muscles functioning, cell divisions, transcription, translation, cellular motility, and many others.[1−6] There is a huge variety of different types of biological molecular motors, including cytoskeleton motor proteins (kinesins, myosins and dyneins); nucleic acid motor proteins (DNA and RNA polymerases, nucleases, gyrases and topoisomerases); and rotary motor proteins (rotary ATPases and bacterial flagellar motors).[1,2,4] However, it is now well established that all of them utilize similar mechanisms of transformation of chemical energy into mechanical work.[4] Motor proteins catalyze exotermic chemical reactions of adenosine triphosphate (ATP) hydrolysis or biopolymerization, and they are able to convert the fraction of the energy released during these processes into the work needed to support their functions. Biological molecular motors have been intensively investigated in recent decades using multiple experimental and theoretical methods.[4−9] Although some fundamental features of the mechanisms of motor proteins have been clarified,[4] many aspects of motor proteins functioning, for example, collective dynamics,[10] cooperativity,[10] and the role of stochasticity at the molecular level, remain not well understood.[4,5]

Recent advances in experimental methods, especially in single-molecule techniques, have allowed researchers to investigate properties of motor proteins with a very high temporal and spatial resolution.[4,11] More specifically, distributions of various dynamic properties of molecular motors have been measured quite precisely.[12−17] These observations include run-length distributions and distributions of translocating velocities. The stochastic nature of biochemical transitions that involve molecular motors is clearly reflected in these distributions. Thus, the measurements of distributions of dynamic properties should provide a valuable information on the mechanisms of motor proteins functioning. However, there is a fundamental problem of properly analyzing these data. Although experimentally observed distributions are usually described by Gaussian curves, it was argued that for motor proteins, especially under the external forces, this approximation is not valid anymore.[18] Furthermore, current theoretical models concentrate mostly on calculating mean dynamic properties, while the distributions are usually not addressed.[4,5,7,8]

The breakthrough in this problem came recently when a new approach, which allowed to evaluate distributions of run lengths and velocities, has been presented.[18] This elegant theoretical model utilized combinatorial arguments to obtain exact analytic expressions for various distributions of processive molecular motors. It was shown that under the external forces the distribution of velocities are non-Gaussian and bimodal, and this was explained using the discrete nature of stepping transitions in motor proteins. Although this powerful method

Special Issue: Benjamin Widom Festschrift

Received: September 20, 2017
Revised: October 20, 2017
Published: October 23, 2017

DOI: 10.1021/acs.jpcb.7b09352
J. Phys. Chem. B 2018, 122, 3272−3279

provided the analytic description of the distributions of dynamic properties made some important predictions and was able to describe experimental data for conventional kinesins under no external loads, it has some limitations. It was assumed that linear filaments, along which the motors move, are infinite in length, and there are no other proteins on the lattice that can interfere. In addition, the theory neglected the effect of the initial positions of the motor proteins. Furthermore, the intermediate biochemical transitions during the enzymatic cycle were not fully taken into account. Because real cellular conditions might strongly deviate from the simplified conditions assumed in this theoretical model,[18] it is important to have a comprehensive theoretical description that can take into account these features of cellular dynamics.

In this paper, we develop a new theoretical framework for describing the distributions of dynamic properties of generic processive molecular motors. Our method is based on the first-passage ideas, and we explained in detail the computation of the distributions of run lengths. It allows us take into account the finite length of linear filaments, the starting position of the molecular motor, and the intermediate chemical transitions during the stepping. Our calculations indicate that all these factors may strongly affect the distributions of run lengths. The physical origins of the obtained results are discussed.

THEORETICAL MODEL

Let us consider a single molecular motor that moves along the linear lattice as shown in Figure 1. Because all dynamic properties, in principle, can be evaluated from the probabilities of run lengths,[18] we concentrate in this work on calculating the distribution of run lengths for generic processive molecular motors as illustrated in Figure 1. Suppose that the motor protein molecule starts from the lattice site i, and then if it dissociates from the filament at the site j, the run length will be $l = j - i$. One can observe that for the motor it will be the dissociation from the linear track *for the first time*. Then, naturally, measuring these run lengths corresponds to recording first-passage dissociation events. This is the main idea of our theoretical method. We associate the run-length distributions with the probabilities' densities of the dissociation events. The method of first-passage processes is a well-developed mathematical tool that has been successfully applied for studying various processes in chemistry, physics, and biology.[4,19]

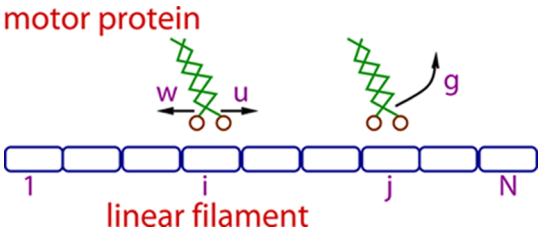


Figure 1. A schematic view of the transport of biological molecular motors along the linear filaments of length N. A single molecule starts the motion from the site i, and it might dissociate into solution from any site with a rate γ. It steps forward with a rate u, while the backward rate is equal to w.

Distribution of Run Lengths on Infinite Linear Filaments. To explain our approach in more detail, let us start with the simplest situation of infinite linear tracks on which molecular motors preferentially step in one direction. This was also considered in the theoretical work of Vu and co-workers.[18] A single motor protein starts from the site $i = 0$. It can move forward (backward) with a rate u (w), or it can irreversibly dissociate from the linear track into the surrounding solution with a rate γ: see Figure 1.

The probability to have a run length n for such molecular motor is identical to the probability of exiting the lattice exactly at the site n. We can introduce a first-passage probability density function $F_l(t)$ of detaching from the linear track at the site n at time t if at $t = 0$ the molecule started at the site l. Here the parameter l can be any integer, and the temporal evolution of first-passage probabilities is governed by a set of backward master equations

$$\frac{dF_l(t)}{dt} = uF_{l+1}(t) + wF_{l-1}(t) - (u + w + \gamma)F_l(t) \tag{1}$$

for $l \neq n$; and

$$\frac{dF_n(t)}{dt} = uF_{n+1}(t) + wF_{n-1}(t) + \gamma F_d(t) - (u + w + \gamma)F(t) \tag{2}$$

for $l = n$. In the last equation, the label d describes another state of the motor protein that dissociated into the solution exactly from the state n but not from any other state. It is convenient to postulate the existence of such state because it is motivated by experimental methods,[12−17] and it allows us to employ the backward master equations method. In our language, the dissociation from the state n is the same as the first arrival into the special state d. Consequently, we can write $F_d(t) = \delta(t)$, which means that if the motor just detached from the lattice at the site n the process is immediately accomplished.

We can solve eqs 1 and 2 by utilizing the Laplace transformations of the first-passage probabilities: $\widetilde{F}_l(s) = \int_0^\infty F_l(t)e^{-st}dt$. This modifies the original backward master equations into

$$(s + u + w + \gamma)\widetilde{F}_l(s) = u\widetilde{F_{l+1}}(s) + w\widetilde{F_{l-1}}(s) \tag{3}$$

for $l \neq n$; and

$$(s + u + w + \gamma)\widetilde{F}_n(s) = u\widetilde{F_{n+1}}(s) + w\widetilde{F_{n-1}}(s) + \gamma \tag{4}$$

for $l = n$. In the last equation we utilized the fact that $\widetilde{F}_d(s) = 1$. We are looking for the solution in the general form $\widetilde{F}_l(s) = Ax^l$, where A and x are unknown parameters that can be determined from substituting the anzats into eqs 3 and 4. This leads to a quadratic equation, which has two roots:

$$x_1 = \frac{(s + u + w + \gamma) - \sqrt{(s + u + w + \gamma)^2 - 4uw}}{2u} \tag{5}$$

and

$$x_2 = \frac{(s + u + w + \gamma) + \sqrt{(s + u + w + \gamma)^2 - 4uw}}{2u} \tag{6}$$

One can notice that the site n is a special one, and it divides the system into two uniform parts. Then the solutions of eqs 3 and 4 have a different behavior in both parts:

$$\widetilde{F}_l(s) = A_1x_1^l + A_2x_2^l \tag{7}$$

for $l < n$, and

DOI: 10.1021/acs.jpcb.7b09352
J. Phys. Chem. B 2018, 122, 3272−3279

$$\widetilde{F}_l(s) = A_3 x_1^l + A_4 x_2^l \tag{8}$$

for $l > n$. Because $x_1 < 1$ and $x_2 > 1$ and $\widetilde{F}_l(s)$ for $l \neq n$ should vanish at $\pm\infty$, we conclude from eqs 7 and (8) that $A_1 = A_4 = 0$. The coefficients A_2 and A_3 can be found from the fact that both solutions for Laplace transforms of the first-passage probability densities should be the same at $l = n$, and from direct substitution into eq 4. Then we obtain for the starting position $l = 0$

$$\widetilde{F}_0(s) = \frac{\gamma}{\sqrt{(s+u+w+\gamma)^2 - 4uw}} x_2^{-n} \tag{9}$$

for $n > 0$, and

$$\widetilde{F}_0(s) = \frac{\gamma}{\sqrt{(s+u+w+\gamma)^2 - 4uw}} x_1^{-n} \tag{10}$$

for $n < 0$.

To compute the probability for the molecular motor to have the run length equal to n, $P(n)$, we notice that the first-passage probability functions calculated above contain a full dynamic description of the system. Specifically, $\widetilde{F}_0(s=0)$ describes the probability to dissociate exactly at the site n, which is equal to the probability to have the run length n, that is, $P(n) \equiv \widetilde{F}_0(s=0)$. Finally, we obtain the explicit formulas for the run-length distributions

$$P(n>0) = \frac{\gamma}{\sqrt{(u+w+\gamma)^2 - 4uw}} \left[\frac{2u}{(u+w+\gamma) + \sqrt{(u+w+\gamma)^2 - 4uw}}\right]^n \tag{11}$$

$$P(n<0) = \frac{\gamma}{\sqrt{(u+w+\gamma)^2 - 4uw}} \left[\frac{2w}{(u+w+\gamma) + \sqrt{(u+w+\gamma)^2 - 4uw}}\right]^{-n} \tag{12}$$

These expressions, as expected, are exactly the same as those obtained earlier in ref 18. Surprisingly, these results imply that that ratio of probabilities of the same run lengths in the positive and negative directions are independent of the dissociation rate,

$$\frac{P(+n)}{P(-n)} = \left(\frac{u}{w}\right)^n \tag{13}$$

This result can be also understood as a simple consequence of the fluctuation theorems for stochastic dynamics.

The results of calculations for run-length distributions of the molecular motors walking on the infinite filaments are presented in Figure 2. Our theoretical model predicts that there are two branches in the distribution. For the motor that started at the origin ($n = 0$), the probability of positive run lengths decays exponentially with a characteristic length

$$\lambda_+ = \frac{1}{\ln\left[\frac{(u+w+\gamma) + \sqrt{(u+w+\gamma)^2 - 4uw}}{2u}\right]} \tag{14}$$

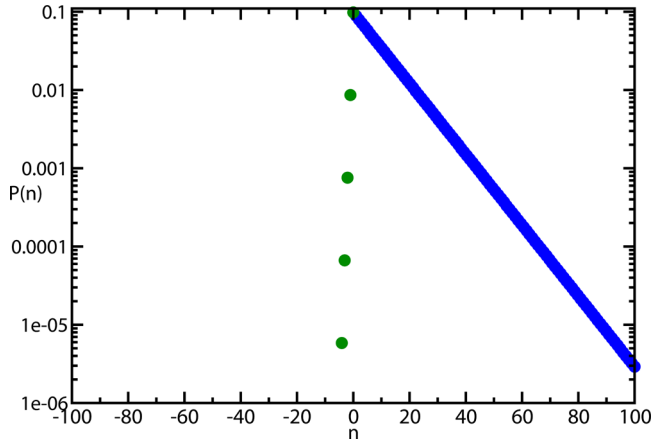


Figure 2. Run-length distribution for a generic motor protein on the infinite linear filament. The starting position of the molecular motors is at the origin. Blue symbols, which look like a line because of the high density of points, correspond to positive run lengths, while green symbols describe the run lengths in the opposite directions. The following parameters are utilized for calculations: $u = 10\ s^{-1}$ and $w = \gamma = 1\ s^{-1}$.

The probability of negative run lengths decays much faster with a characteristic length,

$$\lambda_- = \frac{1}{\ln\left[\frac{(u+w+\gamma) + \sqrt{(u+w+\gamma)^2 - 4uw}}{2w}\right]} \tag{15}$$

because we typically have $u \gg w$. We note here that although the motor protein preferentially moves forward, the stochasticity of stepping leads to the possibility of negative run lengths. Under the external load F, which decreases the forward rate u and increases the backward rate w and the dissociation rate γ, one should expect that the asymmetric run-length distribution presented in Figure 2 will become more symmetric until reaching the stall force when the positive and negative run lengths will be equally probable. This asymmetric run-length distribution is also

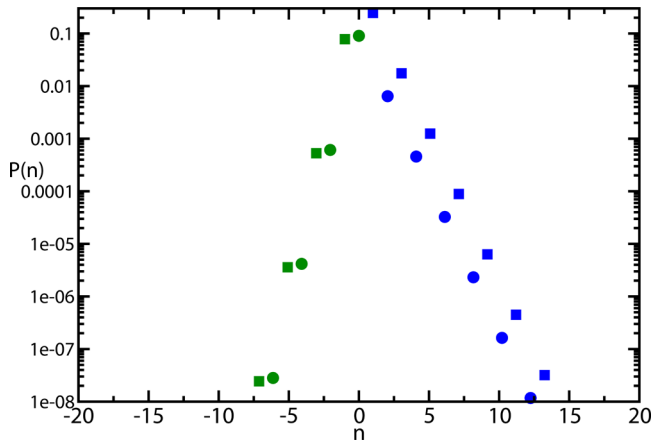


Figure 3. Run-length distribution for a generic motor protein on the infinite linear filament with one intermediate chemical state. The starting position of the molecular motors is at the origin. Blue symbols correspond to even positive run lengths, and blue squares correspond to odd positive run lengths; while green circles describe the even negative run lengths, and green squares describe the odd negative run lengths. The following parameters are utilized for calculations: $u_0 = \gamma_1 = 10\ s^{-1}$ and $u_1 = w_0 = w_1 = \gamma_0 = 1\ s^{-1}$.

DOI: 10.1021/acs.jpcb.7b09352
J. Phys. Chem. B 2018, 122, 3272−3279

closely related to the bimodal distribution of motor proteins velocities as discussed in ref 18.

Distribution of Run Lengths with Intermediate Chemical Transitions. The advantage of our first-passage approach is that it can be extended to take into account more complex features of the motor proteins functioning. For example, it is known that the enzymatic cycles of motor proteins involve several intermediate biochemical states, from which the motor can dissociate into the solution with different rates.[4] To analyze this effect, let us again consider a single motor protein that travels along the infinite linear filament (see Figure 1). We assume that there is only one intermediate state, which leads to the following dynamic rules. At even sites of the lattice, the protein steps forward or backward with rates u_0 or w_0, respectively, while at odd sites the corresponding rates are u_1 and w_1, respectively. In addition, the dissociation rates from the even sites are γ_0, and from the odd sites they are equal to γ_1. Note that here the enzymatic cycle is completed when the motor makes two steps forward. The protein molecule starts from the site $i = 0$, and our goal is to calculate the run-length distributions.

Following the approach described above, we again introduce the first-passage probability density function $F_l(t)$ of dissociating the linear filament at the site n if the motor started at $t = 0$ at the site l. For convenience, we assume that the starting position l is the even number, but the results are independent of this choice. The corresponding set of backward master equations, which describe the temporal evolution of the first-passage probabilities, can be written as (for $l \neq n$),

$$\frac{dF_l(t)}{dt} = u_0F_{l+1}(t) + w_0F_{l-1}(t) - (u_0 + w_0 + \gamma_0)F_l(t) \tag{16}$$

$$\frac{dF_{l+1}(t)}{dt} = u_1F_{l+2}(t) + w_1F_l(t) - (u_1 + w_1 + \gamma_1)F_{l+1}(t) \tag{17}$$

For $l = n$ these equations depend on the value of n being even or odd:

$$\frac{dF_n(t)}{dt} = u_0F_{n+1}(t) + w_0F_{n-1}(t) + \gamma_0F_d(t) - (u_0 + w_0 + \gamma_0)F_n(t) \tag{18}$$

for even n, while for odd n we have

$$\frac{dF_n(t)}{dt} = u_1F_{n+1}(t) + w_1F_{n-1}(t) + \gamma_1F_d(t) - (u_1 + w_1 + \gamma_1)F_n(t) \tag{19}$$

The initial condition requires that $F_d(t) = \delta(t)$.

Applying the Laplace transformations, modifies the backward master equations,

$$a_0\widetilde{F}_l(s) = u_0\widetilde{F_{l+1}}(s) + w_0\widetilde{F_{l-1}}(s) \tag{20}$$

$$a_1\widetilde{F_{l+1}}(s) = u_1\widetilde{F_{l+2}}(s) + w_1\widetilde{F}_l(s) \tag{21}$$

where new parameters are introduced,

$$a_j = s + u_j + w_j + \gamma_j, \quad \text{for } j = 0, 1 \tag{22}$$

At the special site n we have

$$a_0\widetilde{F}_n(s) = u_0\widetilde{F_{n+1}}(s) + w_0\widetilde{F_{n-1}}(s) + \gamma_0 \quad \text{(even sites)} \tag{23}$$

$$a_1\widetilde{F_{n+1}}(s) = u_1\widetilde{F_{n+2}}(s) + w_1\widetilde{F}_n(s) + \gamma_1 \quad \text{(odd sites)} \tag{24}$$

To solve these equations, we notice that combining eqs 20 and 21 leads to

$$A\widetilde{F}_l(s) = U\widetilde{F_{l+2}}(s) + W\widetilde{F_{l-2}}(s) \tag{25}$$

with

$$A = a_0a_1 - (u_0w_1 + u_1w_0), \quad U = u_0u_1, \quad W = w_0w_1 \tag{26}$$

At the special site n we obtain

$$A\widetilde{F}_n(s) = U\widetilde{F_{n+2}}(s) + W\widetilde{F_{n-2}}(s) + \gamma_0a_1 \quad \text{(for } n \text{ even)} \tag{27}$$

$$A\widetilde{F}_n(s) = U\widetilde{F_{n+2}}(s) + W\widetilde{F_{n-2}}(s) + \gamma_1a_0 \quad \text{(for } n \text{ odd)} \tag{28}$$

eq 25 is an important result because it shows that the inhomogeneous problem of calculating run-length distributions with the intermediate chemical state can be mapped into the homogeneous problem without intermediate states, which is very similar to what we analyzed in the previous section. Then the first-passage probability functions can be written as

$$\widetilde{F}_l(s) = C_1x_2^{l/2} \tag{29}$$

for $l < n$; and

$$\widetilde{F}_l(s) = C_2x_1^{l/2} \tag{30}$$

for $l > n$. Parameters x_1 and x_2 are given by

$$x_1 = \frac{A - \sqrt{A^2 - 4UW}}{2U}, \quad x_2 = \frac{A + \sqrt{A^2 - 4UW}}{2U} \tag{31}$$

The coefficients C_1 and C_2 can be found from the boundary conditions given in eqs 27 and 28, yielding for even n

$$C_1 = \frac{a_1\gamma_0N_nU}{Q}, \quad C_2 = \frac{a_1\gamma_0M_nW}{Q} \tag{32}$$

and

$$\widetilde{F}_n(s) = \frac{a_1\gamma_0M_nN_n}{Q} \tag{33}$$

In the above expressions, we introduced new parameters:

$$M_n = Ax_2^{n/2-1} - Wx_2^{n/2-2}, \quad N_n = Ax_1^{n/2+1} - Ux_1^{n/2+2} \tag{34}$$

and

$$Q = AM_nN_n - WUM_nx_1^{n/2+1} - WUN_nx_2^{n/2-1} \tag{35}$$

Similar calculations for odd n produce

$$C_1 = \frac{a_0\gamma_1N_nU}{Q}, \quad C_2 = \frac{a_0\gamma_1M_nW}{Q} \tag{36}$$

and

$$\widetilde{F}_n(s) = \frac{a_0\gamma_1M_nN_n}{Q} \tag{37}$$

Finally, after some simplifications and assuming again the starting position at $l = 0$, we obtain the explicit expressions for the run-length distributions in the system with one intermediate chemical state, $P(n) = \widetilde{F}_0(s = 0)$, leading to the following expressions:

DOI: 10.1021/acs.jpcb.7b09352
J. Phys. Chem. B 2018, 122, 3272−3279

$$P\left(n > 0,\ even\right) = \frac{\gamma_0(u_1 + w_1 + \gamma_1)}{\sqrt{\left[(u_0 + w_0 + \gamma_0)(u_1 + w_1 + \gamma_1) - (u_0w_1 + u_1w_0)\right]^2 - 4u_0u_1w_0w_1}} \times \left[\frac{2u_0u_1}{\left(u_0 + w_0 + \gamma_0\right)\left(u_1 + w_1 + \gamma_1\right) - \left(u_0w_1 + u_1w_0\right) + \sqrt{\left[(u_0 + w_0 + \gamma_0)(u_1 + w_1 + \gamma_1) - (u_0w_1 + u_1w_0)\right]^2 - 4u_0u_1w_0w_1}}\right]^{n/2}; \tag{38}$$

$$P\left(n < 0,\ even\right) = \frac{\gamma_0(u_1 + w_1 + \gamma_1)}{\sqrt{\left[(u_0 + w_0 + \gamma_0)(u_1 + w_1 + \gamma_1) - (u_0w_1 + u_1w_0)\right]^2 - 4u_0u_1w_0w_1}} \times \left[\frac{2w_0w_1}{\left(u_0 + w_0 + \gamma_0\right)\left(u_1 + w_1 + \gamma_1\right) - \left(u_0w_1 + u_1w_0\right) + \sqrt{\left[(u_0 + w_0 + \gamma_0)(u_1 + w_1 + \gamma_1) - (u_0w_1 + u_1w_0)\right]^2 - 4u_0u_1w_0w_1}}\right]^{n/2}; \tag{39}$$

$$P\left(n > 0,\ odd\right) = \frac{\gamma_1(u_0 + w_0 + \gamma_0)}{\sqrt{\left[(u_0 + w_0 + \gamma_0)(u_1 + w_1 + \gamma_1) - (u_0w_1 + u_1w_0)\right]^2 - 4u_0u_1w_0w_1}} \times \left[\frac{2u_0u_1}{\left(u_0 + w_0 + \gamma_0\right)\left(u_1 + w_1 + \gamma_1\right) - \left(u_0w_1 + u_1w_0\right) + \sqrt{\left[(u_0 + w_0 + \gamma_0)(u_1 + w_1 + \gamma_1) - (u_0w_1 + u_1w_0)\right]^2 - 4u_0u_1w_0w_1}}\right]^{n/2}; \tag{40}$$

$$P\left(n < 0,\ odd\right) = \frac{\gamma_1(u_0 + w_0 + \gamma_0)}{\sqrt{\left[(u_0 + w_0 + \gamma_0)(u_1 + w_1 + \gamma_1) - (u_0w_1 + u_1w_0)\right]^2 - 4u_0u_1w_0w_1}} \times \left[\frac{2w_0w_1}{\left(u_0 + w_0 + \gamma_0\right)\left(u_1 + w_1 + \gamma_1\right) - \left(u_0w_1 + u_1w_0\right) + \sqrt{\left[(u_0 + w_0 + \gamma_0)(u_1 + w_1 + \gamma_1) - (u_0w_1 + u_1w_0)\right]^2 - 4u_0u_1w_0w_1}}\right]^{n/2}. \tag{41}$$

One can easily see that for the symmetric case, $u_0 = u_1$, $w_0 = w_1$ and $\gamma_0 = \gamma_1$, these run-length distributions reduce to the results already obtained for homogeneous system: see eqs 11 and 12. In addition, the ratio of probabilities of the same positive and negative run lengths again is independent of the dissociation rates

$$\frac{P(+n)}{P(-n)} = \left(\frac{u_0u_1}{w_0w_1}\right)^{n/2} \tag{42}$$

The results for run-length distributions in the system with the intermediate chemical state are presented in Figure 3. We again observe two asymmetric branches in the run-length distributions, which can be explained by the fact that the motor protein has a higher probability to step in the positive direction. However, the presence of the intermediate state modifies the original exponentially decaying run-length distribution by adding the oscillations. This is because the molecular motor has different dissociation rates at the even and the odd sites, giving different dissociation probabilities from these sites, which look like oscillations in the run-length distributions. It is expected that such oscillations should also show up in the distribution of velocities. Thus, the intermediate chemical transitions in motor proteins should modify the run-lengths distributions in the way that reflects the periodicity of enzymatic cycles. Similar behavior

DOI: 10.1021/acs.jpcb.7b09352
J. Phys. Chem. B 2018, 122, 3272−3279

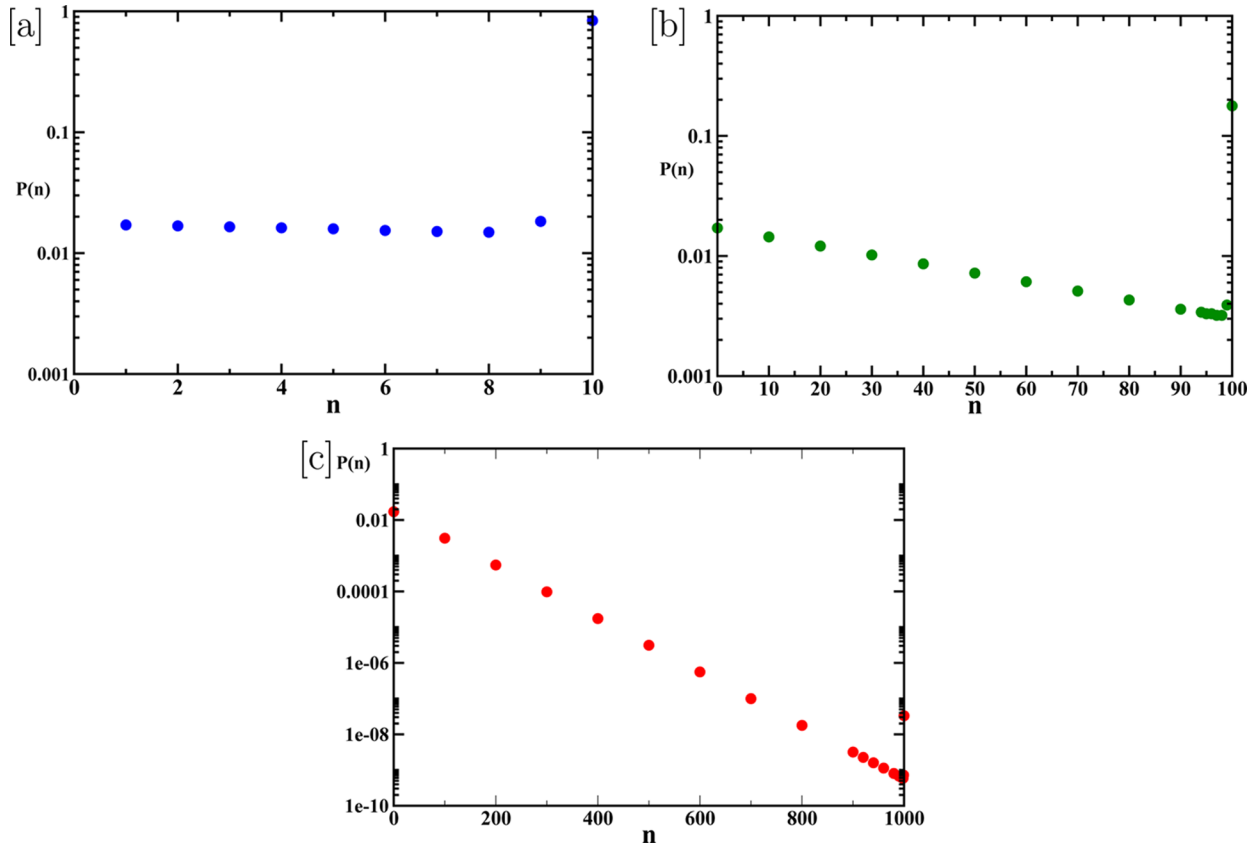


Figure 4. Run-length distribution for a motor protein on the finite length linear filaments. The starting point for the molecular motor is at the middle of the linear filament at $i = 0$. Only positive run-lengths probabilities are shown because for utilized parameters essentially no backward motion is observed. The following parameters are utilized for calculations: $u = 133\ s^{-1}$ $w = 0.6\ s^{-1}$, $\gamma = 2.3\ s^{-1}$; and (a) $N = 10$; (b) $N = 100$; and (c) $N = 1000$.

is expected for more complex chemical kinetic schemes of the enzymatic action for biological molecular motors.

Run-Length Distributions on Finite Linear Filaments. In real cellular systems, the linear tracks on which the motor proteins move have finite lengths. Our theoretical framework can be extended to take this into account. Let us consider a linear filament of the length $2N+1$, and we label its sites as running from $-N$ to N. For convenience, we assume that there is no intermediate chemical states, and the forward, backward, and dissociation rates are equal to u, w, and γ, respectively: see Figure 1. It is assumed also that the linear filaments ends are reflecting; that is, the motor protein at the last site can only go backward or it can dissociate into the solution. We define $F_{i,n}(t)$ as the first-passage probability function of exiting the linear track at the site n if started at $t = 0$ at the site i. The following set of backward master equations describe the evolution of these probability functions,

$$\frac{dF_{-N,n}(t)}{dt} = uF_{-N+1,n}(t) - (u+\gamma)F_{-N,n}(t) \tag{43}$$

$$\frac{dF_{i,n}(t)}{dt} = uF_{i+1,n}(t) + wF_{i-1,n}(t) - (u+w+\gamma)F_{i,n}(t), \quad \text{for } i \neq n \tag{44}$$

$$\frac{dF_{n,n}(t)}{dt} = uF_{n+1,n}(t) + wF_{n-1,n}(t) + \gamma F_d(t) - (u+w+\gamma)F_{n,n}(t) \tag{45}$$

$$\frac{dF_{N,n}(t)}{dt} = wF_{N-1,n}(t) - (w+\gamma)F_{i,n}(t) \tag{46}$$

Using Laplace transformations, these equations are modified into

$$(s+u+\gamma)\widetilde{F_{-N,n}}(s) = u\widetilde{F_{-N+1,n}}(s) \tag{47}$$

$$(s+u+w+\gamma)\widetilde{F_{i,n}}(s) = u\widetilde{F_{i+1,n}}(s) + w\widetilde{F_{i-1,n}}(s), \quad i \neq n \tag{48}$$

$$(s+u+w+\gamma)\widetilde{F_{n,n}}(s) = u\widetilde{F_{i+1,n}}(s) + w\widetilde{F_{i-1,n}}(s) + \gamma \tag{49}$$

$$(s+u+\gamma)\widetilde{F_{N,n}}(s) = w\widetilde{F_{N-1,n}}(s) \tag{50}$$

These equations can be solved using the same approach as was utilized in the previous cases by dividing the system into two parts depending on the position of the dissociation site. We obtain explicit expressions for Laplace transforms of the first-passage probabilities. Specifically, we derive that

$$\widetilde{F_{-N,n}}(s) = \frac{\gamma}{(s+u+w+\gamma)A_{N+n-1} - uA_{N+n-1}(D_{N-n-2}/D_{N-n-1}) - wA_{N+n-2}} \tag{51}$$

where

$$A_k = (a_1 - b_1)a_k - b_k, \qquad D_k = (c_1 - d_1)c_k - d_k \tag{52}$$

The new parameters are given by

$$a_1 = \frac{s+u+w+\gamma}{u}, \quad c_1 = \frac{s+u+w+\gamma}{w}, \quad b_1 = w/u, \quad d_1 = u/w \tag{53}$$

and

DOI: 10.1021/acs.jpcb.7b09352
J. Phys. Chem. B 2018, 122, 3272−3279

$$a_k = \left(1 + \frac{2b_1}{\Delta + a_1\sqrt{\Delta}}\right)\left(\frac{a_1 + \sqrt{\Delta}}{2}\right)^k - \frac{2b_1}{\Delta + a_1\sqrt{\Delta}}\left(\frac{a_1 - \sqrt{\Delta}}{2}\right)^k \quad (54)$$

$$b_k = \left(1 + \frac{2b_1}{\Delta + a_1\sqrt{\Delta}}\right)b_1\left(\frac{a_1 + \sqrt{\Delta}}{2}\right)^{k-1} - \frac{2b_1^2}{\Delta + a_1\sqrt{\Delta}}\left(\frac{a_1 - \sqrt{\Delta}}{2}\right)^{k-1} \quad (55)$$

$$c_k = \left(1 + \frac{2d_1}{\nabla + c_1\sqrt{\nabla}}\right)\left(\frac{c_1 + \sqrt{\nabla}}{2}\right)^k - \frac{2d_1}{\nabla + c_1\sqrt{\nabla}}\left(\frac{c_1 - \sqrt{\nabla}}{2}\right)^k \quad (56)$$

$$d_k = \left(1 + \frac{2d_1}{\nabla + c_1\sqrt{\nabla}}\right)d_1\left(\frac{c_1 + \sqrt{\nabla}}{2}\right)^{k-1} - \frac{2d_1^2}{\nabla + c_1\sqrt{\nabla}}\left(\frac{c_1 - \sqrt{\nabla}}{2}\right)^{k-1} \quad (57)$$

with $\Delta = a_1^2 - 4b_1$ and $\nabla = c_1^2 - 4d_1$. For $(-N+2) \le i \le n$ we have

$$\widetilde{F_{i,n}}(s) = [(a_1 - b_1)a_{N+i-1} - b_{N+i-1}]\widetilde{F_{-N,n}}(s) \quad (58)$$

while for $n \le i \le N-2$ the expressions are,

$$\widetilde{F_{i,n}}(s) = [(c_1 - d_1)c_{N-i-1} - d_{N-i-1}]\widetilde{F_{N,n}}(s) \quad (59)$$

with

$$\widetilde{F_{N,n}}(s) = \frac{\gamma}{(s+u+w+\gamma)D_{N-n-1} - wD_{N-n-1}(A_{N+n-2}/A_{N+n-1}) - uD_{N-n-2}}. \quad (60)$$

We also can show that

$$\widetilde{F_{-N+1,n}}(s) = \frac{s+u+\gamma}{u}\widetilde{F_{-N,n}}(s), \quad \widetilde{F_{N-1,n}}(s) = \frac{s+w+\gamma}{w}\widetilde{F_{N,n}}(s). \quad (61)$$

Finally, the run-length distribution $P(n|i)$, defined as the probability of dissociating at the site n if the molecular motor started at the site$i$, is calculated as $P(n|i) = \widetilde{F_{i,n}}(s=0)$.

Results of theoretical analysis of run-length distributions on finite linear tracks are illustrated in Figures 4 and 5. Interestingly, the parameters used for calculations are the same that have been utilized to describe the dynamics kinesin motor proteins in ref 18.

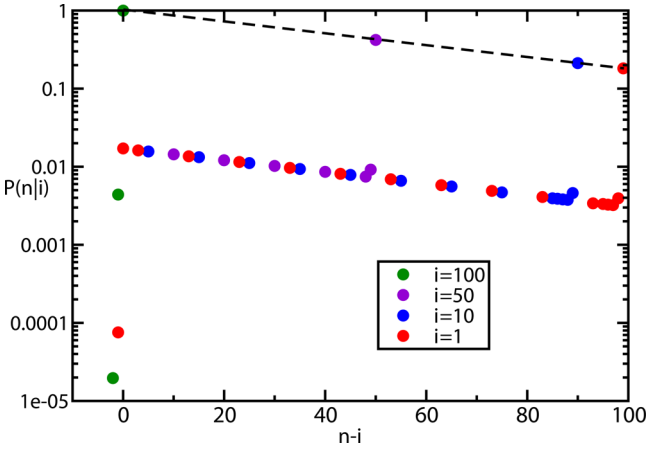


Figure 5. Run-length distributions for a motor protein on the linear filament of length $2N + 1 = 201$. The starting point for the molecular motor varies from $i = 1$ to $i = 100$. The dashed line connects the points corresponding to the motor being at the last site of the filament. The following parameters are utilized for calculations: $u = 133\ s^{-1}$, $w = 0.6\ s^{-1}$, and $\gamma = 2.3\ s^{-1}$.

One can see that the finite length influences the run-length distributions but very locally (see Figure 4). For the motor protein starting at the site i, the probability of the run length equal to $N - i$ is significantly increased and deviates strongly from the exponentially decaying trend. This is easy to understand because the last site on the filament (in the positive direction) serves as an effective trap. The molecular motor that reached the last site can stay there for a long time because it has a lower probability to move backward ($w < u$), and this leads to the enhanced probability to dissociate from the last site and to increase in the run-length probability $P(N|i)$. The effect is stronger for short filaments (Figure 4a), and it decreases with N (Figure 4b,c) since the motor protein lowers its probability of visiting any site on the lattice for large N.

Figure 5 shows the effect of the initial position on the run-length distributions of molecular motors stepping on filaments of finite length. For the motor proteins moving in the positive direction, starting the close to the right end severely limits possible run lengths: for $i = 100$ zero run length dominates, and there is a very low probability of negative run lengths (yellow circles in Figure 5). Moving the initial position further away from the right end increases the possibility of different run lengths. However, the trapping effect of the last site on the filament can still be clearly observed. The probabilities of run lengths equal to $N - i$ are significantly larger than other probabilities. Theoretical results presented in Figures 4 and 5 suggest that the starting position of the motor and the overall length of the linear track might strongly influence the run-length distributions of molecular motors.

SUMMARY AND CONCLUSIONS

We developed a new theoretical framework to calculate and analyze the distributions of dynamic properties of biological molecular motors. It employs the first-passage processes method that allows us to compute distributions analytically for various sets of parameters. The framework is specifically applied for describing the run-length distributions of motor proteins moving along linear tracks. We easily recover the results obtained in the earlier theoretical work,[18] but the advantage of our method is that it can be generalized to include more complex realistic features of the motor proteins dynamics. Our analysis shows that the run-length distributions might be affected by the presence of intermediate chemical states in the enzymatic cycles. Although the overall exponentially decaying behavior is preserved, different dissociation probabilities from intermediate states modify the run-length distributions by introducing effective oscillations. Much stronger effects are observed if the finite length of the linear filaments and the variation of the starting position of the motor are taken into account. It is argued that the last site on the lattice behaves as an effective trap, significantly increasing the probability of run lengths that end up at this site. For finite linear tracks, varying the initial position of the molecular motor limits the possibility of some run lengths, affecting the overall distribution. It is also argued that these features of the motor proteins transport should also influence distributions of other dynamic properties such velocities and diffusivities.

Although the presented theoretical method is capable of describing the run-lengths distribution for processive biological molecular motors, it is only the beginning of the comprehensive investigation on role of stochastic fluctuations in dynamic properties of motor proteins. It will be interesting to extend this study in several directions. One will need to look into more realistic biochemical schemes, into the effect of external loads on

DOI: 10.1021/acs.jpcb.7b09352
J. Phys. Chem. B 2018, 122, 3272−3279

distributions and how changes in the run-lengths distributions affect the distribution of velocities and diffusivities. Another interesting question is what are the distributions of dynamic properties for interacting molecular motors, since motor proteins in cells typically function collectively and typically move in groups. Finally, it will be important to test these theoretical ideas in extensive single-molecule experiments, and this should lead to a better understanding of the mechanisms of motor proteins.

AUTHOR INFORMATION

Corresponding Author

*E-mail: tolya@rice.edu. Phone: +1 713 3485672.

ORCID

Anatoly B. Kolomeisky: 0000-0001-5677-6690

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

A.B.K. acknowledges the support from Center for Theoretical Biological Physics, NSF Grant PHY-1427654, from NSF (Grants CHE-1360979 and CHE-1664218) and from the Welch Foundation (Grant C-1559).

REFERENCES

(1) Alberts, B.; Johnson, A.; Lewis, J.; Morgan, D.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*, 6th ed.; Garland Science: New York, 2014.

(2) Lodish, H.; Berk, A.; Kaiser, C.; Krieger, M.; Scott, M.; Bretscher, A.; Ploegh, H.; Matsudaira, P. *Molecular Cell Biology*, 6th ed.; W.H. Freeman: New York, 2007.

(3) Howard, J. *Mechanics of Motor Proteins and the Cytoskeleton*, 2nd ed.; Sinauer Associates, Inc.: Sunderland, MA, 2001.

(4) Kolomeisky, A. B. *Motor Proteins and Molecular Motors*; CRC Press, Taylor and Francis: New York, 2015.

(5) Chowdhury, D. Stochastic Mechano-Chemical Kinetics of Molecular Motors: A Multidisciplinary Enterprise from a Physicist's Perspective. *Phys. Rep.* **2013**, *529*, 1−197.

(6) Cross, R. A.; McAinsh, A. Prime Movers: The Mechanochemistry of Mitotic Kinesins. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 257−271.

(7) Kolomeisky, A. B.; Fisher, M. E. Molecular Motors: A Theorists Perspective. *Annu. Rev. Phys. Chem.* **2007**, *58*, 675−695.

(8) Kolomeisky, A. B. Motor Proteins and Molecular Motors: How to Operate Machines at the Nanoscale. *J. Phys.: Condens. Matter* **2013**, *25*, 463101.

(9) Greenleaf, W. J.; Woodside, M. T.; Block, S. M. High-Resolution Single-Molecule Measurements of Biomolecular Motion. *Annu. Rev. Biophys. Biomol. Struct.* **2007**, *36*, 171−190.

(10) McLaughlin, R. T.; Diehl, M. R.; Kolomeisky, A. B. Collective Dynamics of Processive Cytoskeletal Motors. *Soft Matter* **2016**, *12*, 14−21.

(11) Veigel, C.; Schmidt, C. F. Moving Into the Cell: Single-Molecule Studies of Molecular Motors in Complex Environment. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 163−176.

(12) Walter, W. J.; Beranek, V.; Fischermeier, E.; Diez, S. Tubulin Acetylation Alone Does Not Affect Kinesin-1 Velocity and Run Length In Vitro. *PLoS One* **2012**, *7*, e42218.

(13) Courty, S.; Luccardini, C.; Bellaiche, Y.; Cappello, G.; Dahan, M. Tracking Individual Kinesin Motors in Living Cells Using Single Quantum-Dot Imaging. *Nano Lett.* **2006**, *6*, 1491−1495.

(14) Ali, M. Y.; Lu, H.; Bookwalter, C. S.; Warshaw, D. M.; Trybus, K. M. Myosin V and Kinesin Act as Tethers to Enhance Each Others' Processivity. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 4691−4696.

(15) Soppina, V.; Norris, S. R.; Dizaji, A. S.; Kortus, M.; Veatch, S.; Peckham, M.; Verhey, K. J. Dimerization of Mammalian Kinesin-3 Motors Results in Superprocessive Motion. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 5562−5567.

(16) Xu, J.; King, S. J.; Lapierre-Landry, M.; Nemec, B. Interplay between Velocity and Travel Distance of Kinesin-based Transport in the Presence of Tau. *Biophys. J.* **2013**, *105*, L23−L26.

(17) Hammond, J. W.; Cai, D.; Blasius, T. L.; Li, Z.; Jiang, Y.; Jih, G. T.; Meyhofer, E.; Verhey, K. J. Mammalian Kinesin-3 Motors Are Dimeric In Vivo and Move by Processive Motility upon Release of Auto-inhibition. *PLoS Biol.* **2009**, *7*, e1000072.

(18) Vu, H. T.; Chakrabarti, S.; Hinczewski, M.; Thirumalai, D. Discrete Step Sizes of Molecular Motors Lead to Bimodal Non-Gaussian Velocity Distributions under Force. *Phys. Rev. Lett.* **2016**, *117*, 078101.

(19) Redner, S. A. *A Guide to First-Passage Processes*; Cambridge University Press: Cambridge, 2001.

DOI: 10.1021/acs.jpcb.7b09352
J. Phys. Chem. B 2018, 122, 3272−3279