Analytical theory of the nonequilibrium spatial distribution of RNA polymerase translocations

Hyung-June Woo
Department of Chemistry, University of Nevada, Reno, Nevada 89557, USA
(Received 13 April 2006; published 14 July 2006)

A continuum Fokker-Planck model is considered for the RNA polymerase in the elongation phase, where the topology of a single free energy profile as a function of the translocation variable distinguishes the Brownian ratchet and power stroke mechanisms. The model yields a simple analytical stationary solution for arbitrary functional forms of the free energy. With the translocation potential of mean force estimated by the time-series data of the recent high-resolution single-molecule experiment [Abbondanzieri *et al.* Nature (London) **438**, 460 (2005)], predictions of the model for the mechanical properties agree with experiments quantitatively with reasonable values of parameters. The evolution of the spatial distribution of translocation variable away from equilibrium with increasing nucleoside triphosphate concentration shows qualitatively different behavior in the two alternative scenarios, which could serve as an additional measurable signature of the underlying mechanism.

DOI: 10.1103/PhysRevE.74.011907 PACS number(s): 87.16.Nn, 05.40.−a, 87.14.Gg

I. INTRODUCTION

Motor proteins constitute one of the most well-characterized biological model systems operating far from equilibrium, converting chemical free energy into mechanical work [1]. One of the questions that has been raised for the mechanism of such free energy transductions is regarding the roles played by thermal fluctuations in the force production step, leading to the two scenarios, Brownian ratchet and power stroke perspectives [2–8]. In the thermal ratchet description, isotropic fluctuations inherent in nanoscale systems are biased into net movements in one preferential direction by external controls, including flashing asymmetric potentials and the binding and hydrolysis of chemical fuels such as adenosine triphosphate (ATP) [9–12]. Protein conformational changes, on the other hand, play central roles in the power stroke perspective, such as the swinging lever arm mechanism of actomyosins underlying muscle contractions [13–15]. In power stroke descriptions, the force production step is viewed as resulting from the relaxation of a certain protein conformational degrees of freedom, such as the lever arm rotation in myosins. Thermal fluctuations merely modify such movements, and therefore are deemed not essential.

One of the motor protein systems for which the question has arisen with growing interests via a series of recent experimental studies [8,16] is the RNA polymerase (RNAP). The RNAP operations, catalyzing RNA synthesis from DNA by highly processive movements along the template [17,18], differ in many aspects from other motor proteins such as myosins and kinesins: the chemical fuel molecules, the nucleoside triphosphates (NTP), get directly incorporated into the growing RNA transcript, and the natural expected step length, the base-pair rise of DNA, is subnanometer scale (Fig. 1). Angstrom-level fluctuations of local protein conformational degrees of freedom are ubiquitous, and it is natural to expect that translocations of an RNA polymerase on the track may not require any "forced" movements. Kinetic models based on the Brownian ratchet scenario have been formulated and applied successfully [19–21].

Recent high-resolution structures of the single-subunit T7 RNAP [22–24] in the elongation phase, however, showed strong evidence in support of the power stroke scenario. Up to two distinct NTP binding sites corresponding to the pre-insertion and insertion modes have been identified, with the binding events expected to be closely coupled to a conformational change allowing for the insertion of the bound substrate into the active site. The pyrophosphate (PPi) release after the catalyzed substrate incorporation into the RNA transcript would then be the triggering event for the translocation step, "pushed" by the reverse conformational change restoring the open state.

Modern single-molecule experiments on T7 [25] and multisubunit RNAPs [16,26–28], on the other hand, have provided strong support for the Brownian ratchet scenario. In particular, the subnanometer resolution has now been achieved to prove that the RNAP moves by one base-pair at a time [28]. The measured force-velocity properties have indicated that the motor complex is more susceptible to applied forces at lower [NTP], suggesting that the RNAP-hybrid complex is likely undergoing translocations back and forth via thermal fluctuations prior to the binding of an NTP substrate. The NTP binding can then only occur when the binding site has been cleared of the previous nucleoside monophosphate (NMP) monomer by RNAP translocations.

Complementary clues to the elucidation of the mechanism of motor operations could come from efforts to connect in-

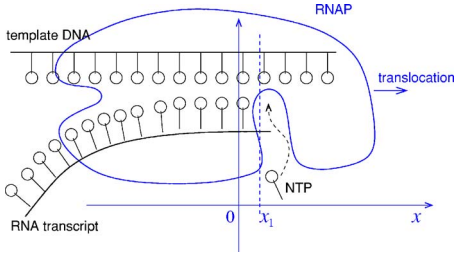

FIG. 1. (Color online) A schematic drawing of the RNA polymerase translocation in the elongation phase. The translocation reaction coordinate x is defined as the displacement of the RNAP active site relative to the position of the (chemically intact) 3′-end of the RNA transcript.

formation gathered from high-resolution structures to energetic and dynamical properties. One possibility is to utilize molecular simulation techniques including the free energy molecular dynamics methods [29,30]. Such calculations could yield the free energy landscape as a function of an appropriate set of reaction coordinates, which forms the starting point of many simple stochastic dynamical models [10,11,31], but have so far been mostly assumed or constructed. The two alternative work production mechanisms, in particular, can be differentiated in terms of the topology of the free energy profile.

In this paper, another possible route is considered for obtaining such a free energy profile directly for RNAP translocations, utilizing the unprecedented resolution of modern single-molecule experiments [28] by inverting the translocation time-series histogram. To provide a sufficiently general foundation for such analysis, we consider a simple continuum stochastic dynamical model for RNAP translocations, extending the existing kinetic models that utilize discrete chemical and mechanical states [19,20,32–35]. The formulation uses the Fokker-Planck (or Smoluchowsky) equations coupled by chemical reactions [36–39], for which a simple analytical solution for the stationary states is obtained. Properties measurable in experiments are considered as functionals of the potential of mean force of translocation $G(x)$, including the Michaelis-Menten-type force-velocity relations and the evolution of nonequilibrium stationary distribution of the translocation variable with increasing substrate concentrations.

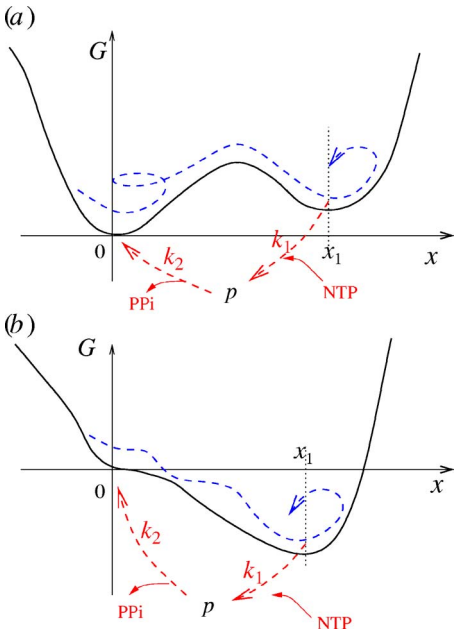

FIG. 2. (Color online) Two possible topologies of the free energy of translocation $G(x)$: (a) Brownian ratchet scenario. Thermal fluctuations drive the translocation from $x=0$ to $x=x_1$ (and back), and the chemical reactions at $x=x_1$ bring the system back to the pretranslocation state at $x=0$ with the transcript grown by one unit. (b) Power stroke scenario. The translocation is driven by the thermodynamic force due to the gradient of G, triggered by the PPi-release. The reactions indicated are reversible, with the backward rates k_{-1} and k_{-2} not shown in the figure.

II. FREE ENERGY OF TRANSLOCATION

A. General considerations

The purely kinetic description of the coupled translocation of RNAP can be summarized by the reaction scheme in the simplest form as follows:

$$R_{n,0} \rightleftharpoons R_{n,x_1} \underset{k_{-1}}{\overset{k_1[\mathrm{NTP}]}{\rightleftharpoons}} R_{n,x_1}\cdot \mathrm{PPi} \underset{k_{-2}[\mathrm{PPi}]}{\overset{k_2}{\rightleftharpoons}} R_{n+1,0}, \qquad (1)$$

where R represents the RNAP-template-transcript complex with the subscripts indicating its length and translocation state, and the states displaced by more than single base-pair rise distance $x_1=3.4$ Å as well as back-tracked states are not included. In Eq. (1), the second step combines the NTP binding into the preinsertion site, its transfer to the insertion site possibly accompanied by partial conformational changes of the RNAP [24,23], and the catalyzed incorporation of the NTP in the 3′-end of the RNA chain, with the effective forward rate $k_1[\mathrm{NTP}]$ and the backward rate k_{-1}, respectively. The third step in Eq. (1) represents the PPi-release, with the forward and backward rates k_2 and $k_{-2}[\mathrm{PPi}]$, respectively. The end-product is a complex containing the transcript elongated by one additional unit, which by definition is the pretranslocation complex for the next round of the process (Fig. 1). The distinction between the Brownian ratchet and power stroke mechanisms lies in the degree of reversibility of the first step in Eq. (1); the translocation is reversible with the former, while the latter would implicate a one-sided translocation dominated by protein conformational changes.

In the continuum Fokker-Planck description of motor proteins, the system is viewed as undergoing Brownian motions in a subset of reaction coordinate spaces, influenced by random forces, frictional drag, and the thermodynamic force arising from the free energy. Chemical reactions provide a means of nonequilibrium pumping maintaining nonzero fluxes in the reaction coordinate space. A version appropriate for the RNAP translocation is described by the free energy profiles depicted in Fig. 2. The motor protein RNAP undergoes spatial translocations in the reaction coordinate space x, coupled to NTP binding at $x=x_1$ and its catalyzed incorporation into the RNA-DNA hybrid.

The two different mechanisms are represented by the different topologies of the potential of mean force $G(x)$ [$G(0)=0$ is set as the reference value of the free energy]. In the Brownian ratchet case, the pre- and post-translocational states at $x=0$ and $x=x_1$ are both local minima similar in free energy, separated by a barrier of height $\sim k_BT$ (the Boltzmann constant times temperature) or smaller. In the power stroke scenario, the translocation step would be a monotonic downhill drift powered by the thermodynamic force, which would commence immediately after the PPi-release that completes the synthesis of $R_{n+1,0}$. In both cases, the combined reaction can only be initiated near the immediate neighborhood of $x=x_1$, reflecting the fact that the NTP binding is possible only after the translocation due to steric hin-

drances within the protein, and the catalytic activity of the protein requires sharply pinpointed conformational states. The PPi-bound state with probability p is to be regarded as spatially localized at $x=x_1$. We also note that in general depending on the particular motor systems being described, the distinction between the two possible mechanisms is often not entirely clear-cut, with the reality possibly possessing partial characteristics of the two extremes at the same time [4,5]. The use of the free energy profile $G(x)$ for RNAP translocation offers a well-defined description connecting the two limiting cases.

The adequacy of using only a single free energy profile $G(x)$ for the minimal description of the stochastic dynamics is peculiar to RNAP, compared to other conventional motor proteins such as myosins or kinesins, where at least two such curves are necessary [40]. For the RNAP, as soon as the catalyzed reactions are complete and PPi dissociates for a given transcript length n, the RNAP complex becomes structurally identical to the starting configuration for the addition of $n+1$'th NTP monomer without any changes in the reaction coordinate. A more explicit version could employ a series of curves vertically displaced by the amount of reaction free energy change for given concentrations [33,39,41]. It is then possible to derive approximately from such extended models the reduced description of the type shown in Fig. 2 by reduction (Appendix). The simpler scheme of Fig. 2 nevertheless ignores the dependence of the free energy on the total length of the transcript, as well as the effects of the sequence heterogeneity, which have been considered in Refs. [20,34,35]. The latter, however, could be included with minor extensions by adopting a disordered free energy.

B. Free energy profile from experiments

In modern single-molecule experiments on RNAP using optical tweezers [16,25–28], a bead attached to a DNA molecule is confined by an optical trap, supplying controlled external loads to an RNAP transcribing the DNA while attached to another bead. Abbondanzieri *et al.* [28] recently achieved the subnanometer resolution of the RNAP translocation measurements. The increased resolution in fact allows us to consider inverting the single-molecule translocation time-series data to obtain the free energy profile $G(x)$ directly.

In equilibrium conditions with concentrations $[\mathrm{NTP}]_{\mathrm{eq}}$ and $[\mathrm{PPi}]_{\mathrm{eq}}$, the detailed balance implies that the forward and backward rates of the combined (as well as the individual) chemical reaction of the second and third steps in Eq. (1) should be equal to each other and to the forward and backward rates of translocations:

$$k_1 k_2 [\mathrm{NTP}]_{\mathrm{eq}} P_{\mathrm{eq}}(x_1) = k_{-1} k_{-2} [\mathrm{PPi}]_{\mathrm{eq}} P_{\mathrm{eq}}(0), \tag{2}$$

or using the equilibrium probability density

$$P_{\mathrm{eq}}(x) = \frac{1}{y} e^{\beta[Fx - G(x)]}, \tag{3}$$

where F is the external force defined as positive for assisting loads, $1/\beta = k_B T$, and y is the normalization constant,

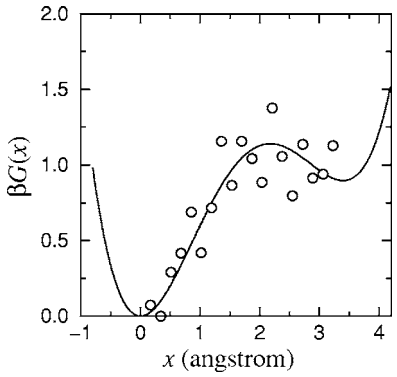

FIG. 3. Free energy profile $G(x)$ in units of $k_B T$, obtained from the histogram of x from the time-series data reported in the single-molecule optical tweezer experiment by Abbondanzieri *et al.* [28]. The circles represent the negative logarithm of the histogram, which is also shown in Fig. 6(a). The solid line is a polynomial fit $\beta G = x^2(a_1 x^2 + a_2 x - 2a_1 x_1^2 - 3a_2 x_1/2)$ where $a_1 = 0.0717$ and $a_2 = -0.5332$, such that $x=0$ and $x=x_1$ are both minima.

$$\frac{[\mathrm{NTP}]_{\mathrm{eq}}}{[\mathrm{PPi}]_{\mathrm{eq}}} = \frac{k_{-1} k_{-2}}{g k_1 k_2}, \tag{4}$$

where $g = e^{\beta F x_1 - \beta G(x_1)}$. Away from equilibrium where [NTP] and [PPi] are maintained at constant nonequilibrium values with their ratio larger than Eq. (4), the stationary distribution $P(x)$ deviates from the equilibrium profile (3), resulting in nonzero net cyclic flux and translocation rates.

If the histogram of translocation trajectories are measured in single-molecule experiments with subnanometer resolution, Eq. (3) allows us to obtain $G(x)$ by inversion. In the experiments in Ref. [28], the translocation time-series data were reported for the condition where $[\mathrm{PPi}] \simeq 0$ and $F = 18$ pN, and [NTP] is small enough (6.875 μM when averaged over the four types of monomers) such that the translocation velocity $v \simeq 1$ base-pair (bp) s^{-1}. In this paper, we approximate the conditions as *near*-equilibrium [justified *a posteriori* in Fig. 6(a) below], and use the data to obtain $G(x)$. In Fig. 3, the histogram of x-distribution was obtained by collecting statistics over the trajectories with the assumed periodicity of x_1, whose negative logarithm gives the potential $\beta G(x)$ with the term due to the force subtracted. The data were then fit into a quartic polynomial with two minima at $x=0$ and $x=x_1$. The free energy profile thus obtained is consistent with the assumptions of the ratchet models formulated in Refs. [19,20], showing a translocation barrier of the order of $k_B T$ with the posttranslocated state higher in free energy.

Thermal noises inherent in single-molecule experiments will affect the quality of $G(x)$ data obtained by inversion. The establishment of (near) equilibrium conditions for single-molecule properties generally takes long time-averages, while at the same time noises are themselves requirements for fulfilling the fluctuation-dissipation relation. The efficiency and reliability of the inversion data, therefore, will depend sensitively on the experimental setup, such as the characteristics of feedback loops maintaining constant forces.

III. STATIONARY STATES

A. Model predictions

We write the Smoluchowsky equation of the probability distribution $P(x)$ [36,38] for the RNAP complex without the NTP substrate in the active site in terms of the translocation variable x, together with the discrete-state probability p for the complex with the bound-PPi, as (Fig. 2 and Appendix)

$$\frac{\partial P}{\partial t} = -\frac{\partial J}{\partial x} + \delta(x)(k_2 p - k'_{-2}P) + \delta(x - x_1)(k_{-1}p - k'_1 P), \tag{5a}$$

$$\frac{dp}{dt} = -(k_{-1} + k_2)p + k'_{-2}P(0) + k'_1 P(x_1), \tag{5b}$$

where the flux J is given by

$$J = \gamma^{-1}(F - G')P - D\frac{\partial P}{\partial x}, \tag{6}$$

where D is the diffusion coefficient, γ is the friction coefficient, and $G' \equiv dG/dx$. In Eq. (5), the effective rate constants are given by

$$k'_1 = \epsilon k_1[\mathrm{NTP}],$$

$$k'_{-2} = \epsilon k_{-2}[\mathrm{PPi}], \tag{7}$$

with the characteristic length scale ϵ of the localization of rates near the posttranslocated minima, and the delta functions approximate the x-dependence of the rates. The local form of rates is not only physically reasonable, but also leads to a great simplification of the analysis, as has been used for a similar study for nonprocessive motor proteins [40].

From Eqs. (5) and (6), for $x \neq 0$ and $x \neq x_1$, the spatial part of the stationary solution $\{\bar{P}(x), \bar{p}\}$ to $\partial P/\partial t = 0$ or $J = \mathrm{const}$ has the general functional form of

$$\bar{P}(x) = e^{f(x)}\left[a + b\int_0^x dz e^{-f(z)}\right], \tag{8}$$

where $f(x) = \beta[Fx - G(x)]$, $\beta = 1/D\gamma$, and the coefficients a and b are to be determined by appropriate boundary conditions in each of the piecewise-analytic regions [40]. Dividing the x-range into three domains $x < 0$, $0 \leq x < x_1$, and $x \geq x_1$, the boundary conditions read

$$\bar{P}(0^+) = \bar{P}(0^-), \tag{9a}$$

$$\bar{P}(x_1^+) = \bar{P}(x_1^-), \tag{9b}$$

$$\bar{P}'(0^+) - \bar{P}'(0^-) = -\frac{k_2}{D}\bar{p} + \frac{k'_{-2}}{D}\bar{P}(0), \tag{9c}$$

$$\bar{P}'(x_1^+) - \bar{P}'(x_1^-) = -\frac{k_{-1}}{D}\bar{p} + \frac{k'_1}{D}\bar{P}(x_1), \tag{9d}$$

where $\bar{P}'(x) \equiv d\bar{P}/dx$ and Eqs. (9c) and (9d) follow from integrating the right-hand side of Eq. (5a) across the immediate neighborhoods of the singular points $x = 0$ and $x = x_1$.

Due to the overall conservation law of the total probability, only three of the Eqs. (9) lead to independent relations between undetermined coefficients of the stationary solution. Using Eqs. (8), (9), and the condition that $b = 0$ for $x > x_1$ and $x < 0$ since $\bar{P}(x)$ is finite, one can obtain the complete stationary solution as

$$\bar{P}(x) = \frac{1}{y}\begin{cases} e^{f(x)} & (x < 0), \\ e^{f(x)}\left[1 + \dfrac{1}{Q}(k_{-1}k'_{-2} - gk'_1 k_2)\displaystyle\int_0^x dz e^{-f(z)}\right] & (0 \leq x < x_1), \\ \dfrac{1}{Q}[D(k_{-1} + k_2) + k_{-1}k'_{-2}I]e^{f(x)} & (x \geq x_1), \end{cases} \tag{10a}$$

$$\bar{p} = \frac{1}{yQ}[D(gk'_1 + k'_{-2}) + gk'_1 k'_{-2} I], \tag{10b}$$

where $Q = D(k_{-1} + k_2) + gk'_1 k_2 I$, $I = \int_0^{x_1} dz e^{-f(z)}$, and y is the normalization constant with the dimension of length determined by

$$\bar{p} + \int dx \bar{P}(x) = 1, \tag{11}$$

and can be regarded as a characteristic length scale of the localization of $\bar{P}(x)$.

It is easily verified that $\bar{P}(x)$ reduces to the Boltzmann distribution (3) when the detailed balance condition (4) is satisfied. Away from equilibrium, a stationary nonequilibrium flux $\bar{J} = v$ equivalent to the translocation velocity [in units of bp s^{-1}]

$$v = \frac{D}{yQ}(gk'_1 k_2 - k_{-1}k'_{-2}) \tag{12}$$

is established for $0 \leq x < x_1$, which again vanishes unless the detailed balance is broken. Equation (12) can be rewritten in a Michaelis-Menten-type form

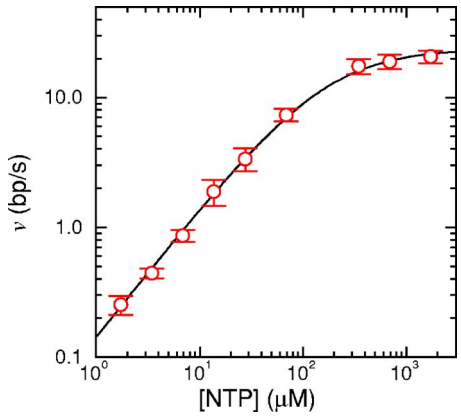

FIG. 4. (Color online) Translocation velocity with increasing [NTP]. The symbols are the experimental data from Ref. [28], plotted by taking $[\mathrm{NTP}]_{\mathrm{eq}}=6.875\ \mu\mathrm{M}$, the average of the four reference concentration values and [PPi]=0. The solid line is Eq. (13) with parameters in Table I with the experimental value of the force, $F=27$ pN.

$$v=\frac{v_{\max}\left\{[\mathrm{NTP}]-\frac{k_{-1}k_{-2}}{gk_1k_2}[\mathrm{PPi}]\right\}}{K_M+[\mathrm{NTP}]}, \tag{13}$$

where

$$v_{\max}=\frac{D}{Iy}, \tag{14a}$$

$$K_M=\frac{D(k_{-1}+k_2)}{g\epsilon k_1k_2I}. \tag{14b}$$

The velocity $v_{\max}$ is a function of [NTP] due to y, and Eq. (13) differs from the standard Michaelis-Menten rate law derived from purely kinetic models [28]. It, however, approaches the maximum velocity in the high-[NTP] limit $v_{\max}\to D/I\bar{y}$, where $\bar{y}$ is the corresponding limit of y.

B. Comparison with experiments

To examine predictions of the model in the context of single-molecule experiments, the translocation velocity expression (13) was fit (Fig. 4) into the experimental data from Ref. [28] using the Brownian ratchet free energy profile $G(x)$ determined in Fig. 3, which yielded the parameter values shown in Table I. The rate constant values show a reasonable agreement with the results of the sequence-dependent kinetic model of Bai *et al.* [20], where the values reported were (in our notation) $K_d=k_{-1}/k_1=15.6\ \mu$M, $k_2=24.7\ \mathrm{s}^{-1}$, and $k_{-2}=0$. The quantitative deviations of Eq. (13) from the kinetic

TABLE I. Parameter values chosen for the comparison with experiments in this work.

D	k_1	k_{-1}	k_2	k_{-2}	ϵ
400 Å^2 s^{-1}	1 μM^{-1} s^{-1}	20 s^{-1}	36 s^{-1}	0	0.5 Å

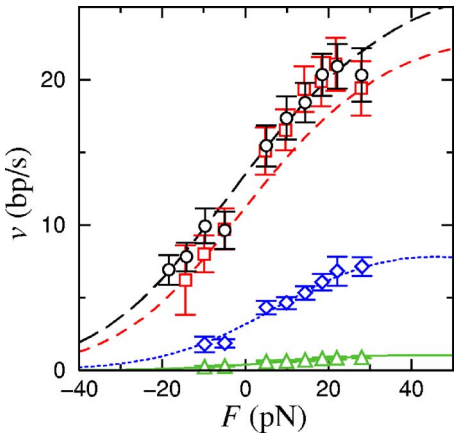

FIG. 5. (Color online) Force-velocity curves from Eq. (13) with experimental data from Ref. [28]. The triangles/solid line, diamonds/dotted line, squares/dashed line, and circles/long-dashed line are for $[\mathrm{NTP}]/[\mathrm{NTP}]_{\mathrm{eq}}=1$, 10, 100, and 250 with $[\mathrm{NTP}]_{\mathrm{eq}}=6.875\ \mu$M, respectively. All parameter values were fixed as those in Table I.

Michaelis-Menten law ($v_{\max}$ independent of [NTP]) were found to be small for the [NTP]-v relation, while the F-v relation in Fig. 5 appeared to generally give a slightly better fit to experimental data than the kinetic model fit used in Ref. [28]. One additional difference of Eqs. (13) and (14) from the kinetic model expression worth noting is that there exists a maximum of v as a function of F, and the translocation velocity eventually reaches zero as $F\to\infty$. The physical origin of the feature is the extreme predominance of the external-load biasing the effective free energy $G(x)-Fx$ for large F. It results in the distribution $P(x)$ dominated by x-ranges increasingly larger than the neighborhood of x_1, eventually suppressing the reactions and translocations. In reality, it is also expected that a sufficiently large assisting load would cause RNAP to translocate forward by more than one base-pair units without NTP incorporations ("hypertranslocation," which nevertheless may not be distinguishable from conventional translocations in experiments), a feature that is not easily captured in purely kinetic descriptions.

The major advantage of the continuum stochastic dynamical theory over the kinetic models is the ability to predict the spatial distribution of the reaction coordinate. Figure 6 shows the evolution of the probability distribution of the translocation variable x away from equilibrium for the two different cases of the topology of $G(x)$, each corresponding to the Brownian ratchet and power stroke mechanisms. The near-equilibrium assumption for the experimental condition of Ref. [28] used to construct the Brownian ratchet $G(x)$ is seen to be justified within the theory, where [NTP]=0 and [NTP]=6.875 μM cases are essentially identical. For the Brownian ratchet case, the two-local-maxima structure arising from the near-equilibrium free energy profile, Eq. (3), gradually evolves into those biased toward a single maximum near $x=0$ far from equilibrium. In contrast, for the power stroke mechanism with a free energy profile possessing a monotonic downhill feature as in Fig. 3(b), near-equilibrium distributions show a single-maximum near the potential minimum. The location of the maximum is

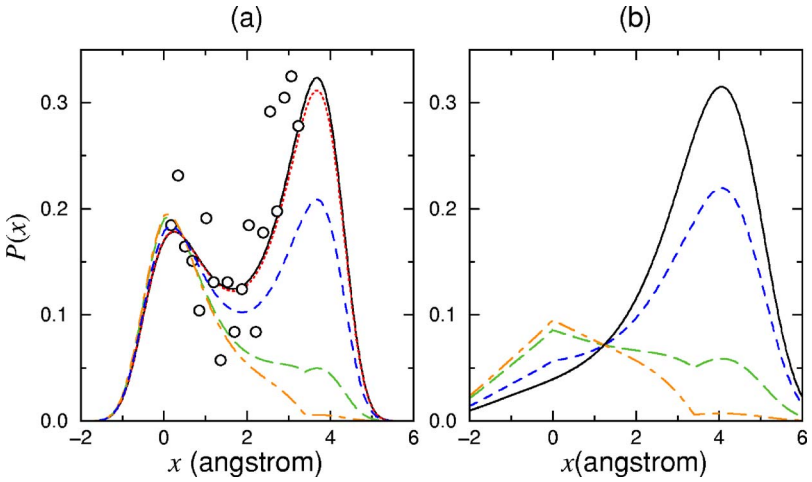

FIG. 6. (Color online) Stationary distribution $\bar{P}(x)$ for different [NTP] values: (a) Brownian ratchet case with the free energy profile shown in Fig. 2; and (b) power stroke case with the model potential $\beta G(x)=0.01x^3(x-4x_1/3)$. The solid, dotted, dashed, long-dashed, and dot-dashed lines are for [NTP]=0, 6.875 μM, 0.1 mM, 1 mM, and 10 mM, respectively. The circles in (a) are the histogram data (shifted by an arbitrary constant and the force contribution subtracted) from Ref. [28] at [NTP]=6.875 μM, which shows that the experimental condition is near-equilibrium. Parameter values are as shown in Table I with $F=18$ pN for all cases.

switched over into the pretranslocated region for high [NTP]. The qualitative difference between the two alternative scenarios in their features of the stationary distribution is therefore largest near equilibrium, where [NTP] is relatively smaller.

IV. CONCLUSIONS

A simple and general formulation of the spatially extended stationary distribution of translocation in RNAP has been presented. The model extends previously formulated kinetic descriptions by employing a free energy profile $G(x)$, of which the results of the theory have been expressed as functionals, along with physically transparent parameters of rate constants and the diffusion coefficient. The free energy $G(x)$ could either be calculated using all-atom simulations based on high-resolution crystal structures [23,24], or could directly be measured in single-molecule experiments [28] as argued in this paper. Further high-resolution experiments could allow the testing of the detailed predictions for the evolution of nonequilibrium stationary distributions illustrated in Fig. 6. Such analysis could provide a more direct means of probing the mechanism of translocations than the force-velocity relations, especially in cases where the true mechanism might lie somewhere in between the two alternatives, as, for example, has been suggested in Ref. [25] for the T7 RNAP. With a number of possible extensions such as inclusions of disorder in $G(x)$ arising from the sequence-heterogeneity of nucleic acids, Eqs. (10) and (13) could provide a simple and flexible means of bridging structural features of protein constituents and dynamical properties measured in single-molecule experiments.

ACKNOWLEDGMENT

This work was supported by the University of Nevada, Reno.

APPENDIX: MODEL REDUCTION

In this appendix, we show the explicit reduction of the extended model with a series of free energy curves (Fig. 7)

$$G_n(x) = G(x - x_{n-1}) + (n-1)\Delta\mu, \tag{A1}$$

where $n=1,2,\ldots$ is the index for the free energy profile with the total length of transcript n, and

$$\Delta\mu = \Delta\mu^{(0)} + k_B T \ln \frac{[\mathrm{PPi}]}{[\mathrm{NTP}]} \tag{A2}$$

is the change in chemical potential per NTP addition reaction. In Eq. (A1), $G(x)=G_1(x)$ approximates the functional form of the general $G_n(x)$, ignoring its dependence on the transcript length. It is expected to be a reasonable approximation for the RNAP in the elongation phase where $n \gg 1$. The set of singular points are defined as

$$x_n = n x_1, \tag{A3}$$

where $n=0,1,\ldots$ with $x=0$ as the origin. The full Fokker-Planck equation for the probability density $P_n(x,t)$ of observing the translocation variable x with the length of transcript n and the discrete PPi-bound state probability $p_n(t)$ can be written as

$$\frac{\partial P_n}{\partial t} = -\frac{\partial J_n}{\partial x} + k_2\delta(x - x_{n-1})p_{n-1} - k'_{-2}\delta(x - x_{n-1})P_n + k_{-1}\delta(x - x_n)p_n - k'_1\delta(x - x_n)P_n, \tag{A4a}$$

$$\frac{dp_n}{dt} = -k_{-1}p_n + k'_1 P_n(x_n) - k_2 p_n + k'_{-2}P_{n+1}(x_n) \tag{A4b}$$

for $n=1,2,\ldots$, where

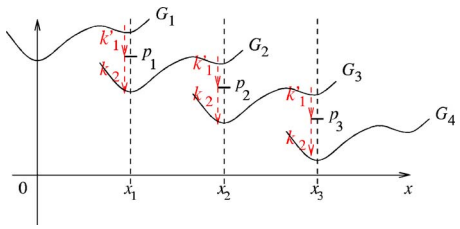

FIG. 7. (Color online) Schematic illustration of the free energy landscape of extended model represented by Eqs. (A4), where G_n is the free energy for the transcript of length n.

$$J_n = \gamma^{-1}(F - G_n')P_n - D\frac{\partial P_n}{\partial x}. \tag{A5}$$

We define the reduced distribution as [10]

$$P(x) = \sum_{n=1}^{\infty} P_n(x + x_{n-1}), \tag{A6a}$$

$$p = \sum_{n=1}^{\infty} p_n. \tag{A6b}$$

Taking time-derivatives of Eqs. (A6), using Eqs. (A4), $G_n'(x+x_{n-1}) = G'(x)$ that follows from Eq. (A1), and assuming $P_1(0,t) \simeq 0$ for $t \gg 0$ (elongation phase), we obtain Eqs. (5) for the reduced distribution.

[1] R. D. Vale and R. A. Milligan, Science **288**, 88 (2000).
[2] N. J. Córdova, B. Ermentrout, and G. F. Oster, Proc. Natl. Acad. Sci. U.S.A. **89**, 339 (1992).
[3] T. C. Elston, Biophys. J. **82**, 1239 (2002).
[4] H. Qian, Phys. Rev. Lett. **81**, 3063 (1998).
[5] H. Y. Wang and G. Oster, Appl. Phys. A: Mater. Sci. Process. **75**, 315 (2002).
[6] R. D. Astumian, J. Phys.: Condens. Matter **17**, S3753 (2005).
[7] W. Zheng and S. Doniach, Proc. Natl. Acad. Sci. U.S.A. **100**, 13253 (2003).
[8] D. G. Vassylyev and I. Artsimovitch, Cell **123**, 977 (2005).
[9] R. D. Astumian, Science **276**, 917 (1997).
[10] P. Reimann, Phys. Rep. **361**, 57 (2002).
[11] F. Jülicher, A. Ajdari, and J. Prost, Rev. Mod. Phys. **69**, 1269 (1997).
[12] M. T. Downton, M. J. Zuckermann, E. M. Craig, M. Plischke, and H. Linke, Phys. Rev. E **73**, 011909 (2006).
[13] R. W. Lymn and E. W. Taylor, Biochemistry **10**, 4617 (1971).
[14] M. A. Geeves and K. C. Holmes, Annu. Rev. Biochem. **68**, 687 (1999).
[15] R. Cooke, J. Gen. Physiol. **123**, 643 (2004).
[16] L. Bai, T. J. Santagelo, and M. D. Wang, Annu. Rev. Biophys. Biomol. Struct. **35**, 343 (2006).
[17] P. H. von Hippel, Science **281**, 660 (1998).
[18] J. Gelles and R. Landick, Cell **93**, 13 (1998).
[19] R. Guajardo and R. Sousa, J. Mol. Biol. **265**, 8 (1997).
[20] L. Bai, A. Shundrovsky, and M. D. Wang, J. Mol. Biol. **344**, 335 (2004).
[21] Q. Guo and R. Sousa, J. Mol. Biol. **358**, 241 (2006).
[22] Y. W. Yin and T. A. Steitz, Science **298**, 1387 (2002).
[23] D. Temiakov, V. Patlan, M. Anikin, W. T. McAllister, S. Yokoyama, and D. G. Vassylyev, Cell **116**, 381 (2004).
[24] Y. W. Yin and T. A. Steitz, Cell **116**, 393 (2004).
[25] P. Thomen, P. J. Lopez, and F. Heslot, Phys. Rev. Lett. **94**, 128102 (2005).
[26] M. D. Wang, M. J. Schnitzer, H. Yin, R. Landick, J. Gelles, and S. M. Block, Science **282**, 902 (1998).
[27] R. J. Davenport, G. J. L. Wuite, R. Landick, and C. Bustamante, Science **287**, 2497 (2000).
[28] E. A. Abbondanzieri, W. J. Greenleaf, J. W. Shaevitz, R. Landick, and S. M. Block, Nature (London) **438**, 460 (2005).
[29] P. Kollman, Chem. Rev. (Washington, D.C.) **93**, 2395 (1993).
[30] T. Simonson, G. Archontis, and M. Karplus, Acc. Chem. Res. **35**, 430 (2002).
[31] C. Bustamante, D. Keller, and G. Oster, Acc. Chem. Res. **34**, 412 (2001).
[32] F. Jülicher and R. Bruinsma, Biophys. J. **74**, 1169 (1998).
[33] H. Wang, T. Elston, A. Mogilner, and G. Oster, Biophys. J. **74**, 1186 (1998).
[34] Y. Kafri, D. K. Lubensky, and D. R. Nelson, Phys. Rev. E **71**, 041906 (2005).
[35] Y. Kafri, D. K. Lubensky, and D. R. Nelson, Biophys. J. **86**, 3373 (2004).
[36] J. Prost, J. F. Chauwin, L. Peliti, and A. Ajdari, Phys. Rev. Lett. **72**, 2652 (1994).
[37] H. Wang and G. Oster, Nature (London) **396**, 279 (1998).
[38] H. Qian, J. Phys. Chem. B **106**, 2065 (2002).
[39] J. Xing, H. Wang, and G. Oster, Biophys. J. **89**, 1551 (2005).
[40] H.-J. Woo and C. L. Moss, Phys. Rev. E **72**, 051924 (2005).
[41] J. Xing, J.-C. Liao, and G. Oster, Proc. Natl. Acad. Sci. U.S.A. **102**, 16539 (2005).