Exact Solutions for Kinetic Models of Macromolecular Dynamics[†]

Yann R. Chemla,[‡,#] Jeffrey R. Moffitt,[∥,#] and Carlos Bustamante*[,§,∥]

Department of Physics, University of Illinois, Urbana-Champaign, Urbana, Illinois 61801, Department of Physics, Howard Hughes Medical Institute, Department of Molecular & Cell Biology and Department of Chemistry, University of California, Berkeley, California 94720

Received: August 1, 2007; In Final Form: November 26, 2007

Dynamic biological processes such as enzyme catalysis, molecular motor translocation, and protein and nucleic acid conformational dynamics are inherently stochastic processes. However, when such processes are studied on a nonsynchronized ensemble, the inherent fluctuations are lost, and only the average rate of the process can be measured. With the recent development of methods of single-molecule manipulation and detection, it is now possible to follow the progress of an individual molecule, measuring not just the average rate but the fluctuations in this rate as well. These fluctuations can provide a great deal of detail about the underlying kinetic cycle that governs the dynamical behavior of the system. However, extracting this information from experiments requires the ability to calculate the general properties of arbitrarily complex theoretical kinetic schemes. We present here a general technique that determines the exact analytical solution for the mean velocity and for measures of the fluctuations. We adopt a formalism based on the master equation and show how the probability density for the position of a molecular motor at a given time can be solved exactly in Fourier−Laplace space. With this analytic solution, we can then calculate the mean velocity and fluctuation-related parameters, such as the randomness parameter (a dimensionless ratio of the diffusion constant and the velocity) and the dwell time distributions, which fully characterize the fluctuations of the system, both commonly used kinetic parameters in single-molecule measurements. Furthermore, we show that this formalism allows calculation of these parameters for a much wider class of general kinetic models than demonstrated with previous methods.

Introduction

In recent years, single-molecule techniques have evolved into a powerful toolset for studying the dynamical behavior of many biological processes. It is now possible to follow on the nanometer and millisecond scales the individual trajectory of a single enzyme as it catalyzes a reaction, a molecular motor as it translocates, or a single polypeptide or nucleic acid molecule as it unfolds and refolds.[1−3] In each case, this trajectory consists of a sequence of molecular events (transitions between the distinct kinetic states of the system) that are stochastic in nature. Single-molecule methods have the ability to measure not only the average trajectory but also the fluctuations around this average; thus, the statistical analysis of single-molecule trajectories can provide crucial insight into the underlying sequence of events that leads to the behavior of a biological system.

Measurements of the mean rate of a reaction as a function of experimental variables (the concentration of a reactant or product, for instance) can reveal important details about intermediate states in the reaction cycle. For example, the dependence of the velocity, v, of a molecular motor on substrate concentration [S] can reveal macroscopic constants of the enzymatic reaction such as k_{cat} and K_M from the Michaelis−Menten equation $v = dk_{cat}[S]/(K_M + [S])$,[4] where d is the step

size of the motor, k_{cat} is the maximal turnover rate, and K_M is the substrate concentration at which the reaction rate is half-maximal. Furthermore, because mechanical force is a "product" of the reaction of a molecular motor, direct application of force makes it possible to also probe how mechanical motion and catalysis are coupled and locate the force-generating step within the cycle.[5] Such studies have revealed details on the mechanochemical cycle of motors such as kinesin[6,7] and myosin V,[8,9] prokaryotic and eukaryotic RNA polymerases,[10,11] and the portal motor of bacteriophage φ29,[12,13] among others.[1−3] However, macroscopic kinetic parameters such as k_{cat} and K_M are only indirectly related to the microscopic rate constants connecting the numerous kinetic states that compose the reaction cycle.[4] As a result, measurements of the velocity alone are insufficient to establish uniquely the number and type of states that compose the kinetic cycle, locate within that cycle the physical step of the motor, or determine its size.

Measurement of fluctuations about the average rate (the diffusion constant D or the related randomness parameter $r = 2D/vd$,[14,15] a dimensionless measure of fluctuations) can lend significantly more information about the underlying kinetics of a system. The latter, specifically, has proven particularly powerful because it provides a lower limit on the number of rate-limiting kinetic transitions in a reaction cycle under given experimental conditions.[15] Measurements of the randomness parameter or related quantities[16] have been used in the study of molecular motors such as kinesin,[6,7,17] topoisomerase II,[16] and the bacterial DNA translocase FtsK.[18] In analogy to the macroscopic kinetic parameters, k_{cat} and K_M, it is also possible to define analogous parameters for r that further confine the

---

† Part of the "Attila Szabo Festschrift".
* To whom correspondence should be addressed. Phone: (510) 643-9076. Fax: (510) 643-3860. E-mail: carlos@alice.berkeley.edu.
‡ University of Illinois.
§ Howard Hughes Medical Institute.
∥ University of California, Berkeley.
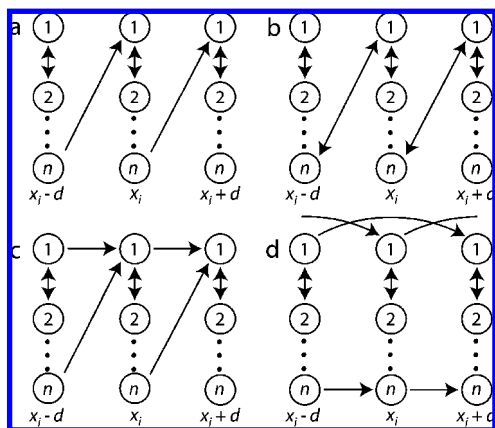# These authors contributed equally to this paper.

Figure 1. Examples of the different types of possible connectivity in different *n* state kinetic models. Circles represent distinct kinetic states; arrows represent the possible kinetic transitions between these states. Double arrowheads represent reversible transitions; the dots indicate the series of kinetic states not shown between state 2 and state *n*. The discrete position along the lattice increases from left to right. (a) A simple kinetic scheme with an irreversible step of size *d*. (b) The reversible version of the system in part a in which the motor can now take a backward step, −*d*. (c) A kinetic scheme in which the system makes physical steps during two different kinetic transitions. (d) A kinetic system with two types of steps that differ in size and that result from different kinetic transitions.

number and value of the microscopic rate constants of the underlying kinetic cycle.

An advantage of steady-state kinetic parameters, such as the velocity and the randomness parameter, is that they do not require the measurement of the exact time each reaction cycle takes to complete. In measurements of molecular motors, for example, it is possible to extract these parameters even when the individual steps of the motor are obscured by experimental noise.[16,18,19] Indeed, in cases in which certain features of the kinetic scheme are known a priori, the randomness parameter can be used to determine the step size of a motor without direct observation.[18] However, if the individual steps are resolvable, then it is possible to measure the exact time between events and compile the distribution of first passage (or dwell) times, $\psi(t)$, in addition to the distribution of step sizes $\rho(x)$. In a rigorous sense, the first two moments of the dwell time distribution, the mean and the variance, are directly related to the mean reaction rate and its fluctuations, respectively. However, since this distribution can, in principle, generate every moment, it contains the maximal amount of kinetic information that can be extracted from the data. Dwell time distributions have been measured for molecular motors such as myosin V[20,21] and kinesin,[22,23] the rotary motor F1-ATPase,[24] and enzymes such as β-galactosidase,[25] to name only a few.
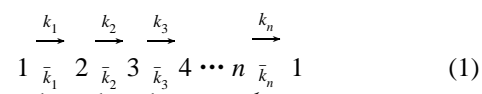
These experimentally observed quantities (the mean velocity v, randomness parameter r, dwell time distribution $\psi(t)$, and their associated macroscopic kinetic parameters) are used in single-molecule measurements to infer properties of the underlying kinetic cycle; however, their connection to the microscopic rate constants that compose this cycle is often quite complex. Moreover, the kinetic cycles that describe the biological processes studied by these techniques are also generally far more complex than those of simple textbook Michaelis−Menten enzyme reactions. They can involve numerous intermediate states that may be connected by irreversible (Figure 1a) or reversible transitions (Figure 1b), branches off the main catalytic pathway (Figure 1c), or parallel pathways with different reaction outcomes (Figure 1d). It is known, for instance, that during transcriptional pausing and arrest, RNA polymerases enter a

backtracked pathway consisting of numerous states,[11,26] whereas kinesin and dynein have been observed to take forward and backward steps or steps of varying sizes, all likely corresponding to different parallel pathways with different kinetics.[27−29] Moreover, certain metabolic enzymes are believed to fluctuate between a large number of catalytically active pathways, each with distinct rate constants.[25,30,31] Thus, to describe these myriad behaviors, it has become essential to develop a general theoretical formalism that can calculate the quantities typically measured in single-molecule experiments for cycles of arbitrary complexity. As will be discussed in more detail below, analytical and numerical methods for performing these types of calculation have already been developed and described previously,[15,30−40] but no single method has been used to calculate exact analytical expressions for all of these kinetic properties. In particular, the calculation of dwell time distributions for systems with the variety of stepping pathways demonstrated in Figure 1 has been limited.[15,36,39,40]

In this article, we present a general methodology adapted from previous work[33,34] for calculating these quantities for a wide range of kinetic systems, including all of those mentioned above. Furthermore, we also demonstrate that it is possible to use this methodology to derive exact analytical expressions for the Laplace transform of the dwell time distributions for many kinetic systems, including most branched systems for which such quantities have not been calculated previously. Finally, we end our discussion with a series of worked examples, chosen to illustrate the power of this technique and the usefulness of these kinetic parameters in the analysis of single-molecule data.

General Formalism

For illustration purposes in the development of this methodology, we will consider in this discussion a molecular motor that undergoes a cycle of chemical and conformational transitions in which it catalyzes a chemical reaction, such as ATP hydrolysis, and translocates along a linear track, such as actin or DNA. Typically, the track can be considered to be periodic, such that, upon completing each cycle, the motor returns to the same mechanochemical state. Transitions between states typically occur by thermal activation over an energy barrier well described by Kramers theory[5] and are generally exponentially distributed in time. Thus, the progress of a molecular motor along its chemical coordinates and along its track can be modeled as a random walk in continuous time along a discrete, periodic one-dimensional lattice.[37] Such assumptions allow for a wide class of kinetic models with different degrees of connectivity between kinetic states and between physical locations on the periodic lattice, as shown in the example systems in Figure 1. For now, we will consider a motor that undergoes a linear cycle of *n* reversible transitions, as denoted in Figure 1b and the following scheme:

$$1 \underset{\bar{k}_1}{\overset{k_1}{\rightleftharpoons}} 2 \underset{\bar{k}_2}{\overset{k_2}{\rightleftharpoons}} 3 \underset{\bar{k}_3}{\overset{k_3}{\rightleftharpoons}} 4 \cdots n \underset{\bar{k}_n}{\overset{k_n}{\rightleftharpoons}} 1 \qquad (1)$$

where $1 \cdots n$ are the *n* kinetic states of the motor, and k_i and $\bar{k}_i$ are the rate coefficients for the forward transition from state *i* to $i + 1$ and the backward transition from $i + 1$ to *i*, respectively. We will extend our analysis to more general cycles in the discussion below.

Because of the cyclical form of eq 1, we can choose without loss of generality to enumerate the states such that transitions 1 through $n - 1$ are catalytic and conformational steps and

Kinetic Models of Macromolecular Dynamics

J. Phys. Chem. B, Vol. 112, No. 19, 2008 **6027**

transition n corresponds to the physical step of the motor along the track. Following the formalism of Koza,[33,34] we describe the progress of the motor through its mechanochemical cycle by the probability density $P_i(x_j, t)$ of occupying state i, at position x_j along the track at time t. The time evolution of this quantity is given by the master equation

$$\frac{\partial P_i(x_j, t)}{\partial t} = -(k_i + \bar{k}_{i-1})P_i(x_j, t) + k_{i-1}P_{i-1}(x_j, t) + \bar{k}_i P_{i+1}(x_j, t) \quad (2)$$

which is modified at the start and end of the cycle to account for transitions from the previous (x_{j-1}) and subsequent (x_{j+1}) positions along the track:

$$\frac{\partial P_1(x_j, t)}{\partial t} = -(k_1 + \bar{k}_n)P_1(x_j, t) + k_n P_n(x_{j-1}, t) + \bar{k}_1 P_2(x_j, t)$$

$$\frac{\partial P_n(x_j, t)}{\partial t} = -(k_n + \bar{k}_{n-1})P_n(x_j, t) + k_{n-1}P_{n-1}(x_j, t) + \bar{k}_n P_1(x_{j+1}, t) \quad (3)$$

The use of the master equation to describe the kinetics of this system implicitly assumes that the lifetimes of each state in the cycle are exponentially distributed. Although it is possible to extend this formalism to explicitly treat arbitrary lifetime distributions, such a calculation falls outside the scope of this article.

By introducing the Fourier transform of the kinetic state probability density, $P_l(q,t) = \sum_{j=-\infty}^{\infty} P_l(x_j, t)e^{-iqx_j}$, we can express these equations concisely in matrix form,

$$\frac{\partial}{\partial t}\mathbf{P}(q, t) = \mathbf{M}(q)\mathbf{P}(q, t) \quad (4)$$

where the vector $\mathbf{P}(q, t)$ and the matrix $\mathbf{M}(q)$ are given by

$$\mathbf{P}(q, t) = \begin{pmatrix} P_1 \\ P_2 \\ P_3 \\ \vdots \\ \vdots \\ P_n \end{pmatrix}$$

$$\mathbf{M}(q) = \begin{pmatrix} -(k_1 + \bar{k}_n) & \bar{k}_1 & 0 & \cdots & k_n\rho_+(q) \\ k_1 & -(k_2 + \bar{k}_1) & \bar{k}_2 & \cdots & 0 \\ 0 & k_2 & -(k_3 + \bar{k}_2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \bar{k}_n\rho_-(q) & 0 & 0 & \cdots & -(k_n + \bar{k}_{n-1}) \end{pmatrix}$$

$$(5)$$

Diagonal elements correspond to transitions out of each state, and off-diagonal elements, to transitions into each state. The generation of a physical step during a kinetic transition is included by multiplying transitions into the state to which the motor steps by the appropriate Fourier weight $\rho_+(q) = e^{-iqd}$, for forward steps, and $\rho_-(q) = e^{iqd}$ for backward steps, where $d = x_{j+1} - x_j$ is the periodicity of the track or the step size of the motor. For a kinetic scheme with a distribution of step sizes, this matrix formalism can easily be extended to a continuous position variable with $\rho(q)$, the continuous Fourier transform of the step size distribution, $\rho(x)$, as the general weight for stepping transitions.

With the appropriate initial conditions, solution to the matrix master equation, eq 4, determines the complete time course of the motor along its mechanochemical cycle, through its multiple catalytic and conformational states, and through its discrete physical steps. In practice, single-molecule experiments do not measure the precise kinetic state in which the motor resides, only its physical position along the track; thus, we will focus on the quantity relevant to these measurements, the *position probability density*,

$$P(q, t) = \sum_{i=1}^{n} P_i(q, t) \quad (6)$$

in which these hidden kinetic states are summed over. Below, we demonstrate how to calculate this quantity and how to connect it to practical kinetic parameters, such as the mean velocity, the diffusion constant, and the full dwell time distribution or distributions.

Calculation of the Mean Velocity and Diffusion Constant. The mean velocity, v, and the diffusion constant, D, are formally defined as[32]

$$v = \lim_{t \to \infty} \frac{\partial}{\partial t}\langle x(t) \rangle$$

$$D = \lim_{t \to \infty} \frac{1}{2}\frac{\partial}{\partial t}(\langle x(t)^2 \rangle - \langle x(t) \rangle^2) \quad (7)$$

where $\langle \cdots \rangle$ denotes the ensemble average. By nature of the Fourier transform, the moments of this distribution can be related to derivatives of $P(q, t)$ by

$$\langle x(t) \rangle = i\frac{\partial P}{\partial q}\bigg|_{q=0}$$

$$\langle x(t)^2 \rangle - \langle x(t) \rangle^2 = -\frac{\partial^2 P}{\partial q^2}\bigg|_{q=0} + \left(\frac{\partial P}{\partial q}\right)^2\bigg|_{q=0} \quad (8)$$

Thus, v and D are related to the first and second moments of the position probability density $P(q, t)$, respectively. As shown by Koza,[33,34] calculation of the velocity and diffusion constant does not require a complete solution for the probability density. Instead, it is only necessary to determine properties of the smallest of the n eigenvalues of the matrix $\mathbf{M}(q)$, which dominates in the long time ($t \to \infty$) and short wavenumber ($q = 0$) limits taken in eqs 7 and 8.

We adopt here a different approach that will be useful in a later analysis. Taking the Laplace transform of the matrix master equation, eq 4, and inverting it yields $\tilde{\mathbf{P}}(q, s) = \mathbf{R}(q, s)^{-1}\mathbf{P}(0)$, where $\tilde{\mathbf{P}}(q, s) = \int_0^\infty dt\, e^{-st}\mathbf{P}(q, t)$ is the Fourier–Laplace transform of the kinetic state probability density vector, the matrix $\mathbf{R}(q, s) \equiv s\mathbf{I} - \mathbf{M}(q)$, and $\mathbf{P}(0)$ is the vector of initial conditions of the kinetic state probability densities. Applying the definition of the inverse matrix and carrying out the summation over kinetic state index i, the Fourier–Laplace transform of the position probability density, $\tilde{P}(q, s)$, takes the following form:

$$\tilde{P}(q, s) = \frac{1}{|\mathbf{R}(q, s)|}\sum_{i,j=1}^{n} Co_{ji}(q, s)P_j(0) = \frac{C(q, s)}{|\mathbf{R}(q, s)|} \quad (9)$$

where $|\mathbf{R}(q, s)|$ is the determinant of $\mathbf{R}$ (i.e., the characteristic polynomial of $\mathbf{M}(q)$), and the matrix elements $Co_{ij}(q, s)$ are the cofactors of $\mathbf{R}$.

There are several properties of this probability density that will be useful in later calculations. First, since $\mathbf{R}$ is an $n \times n$ matrix, its determinant is an nth order polynomial in s; that is,

$$|\mathbf{R}(q, s)| = s^n + \cdots + \alpha(q)s^2 + \beta(q)s + \gamma(q) \quad (10)$$

where, for reasons that will become clear below, we have only written out the last three terms of the polynomial. Furthermore, from the definition of the matrix $\mathbf{R}$, the last term in the polynomial, $\gamma(q)$, is simply the determinant $-|\mathbf{M}(q)|$. Because of conservation of probability, all of the columns of $\mathbf{M}$ sum to zero in the $q \rightarrow 0$ limit (see eq 5); thus, $|\mathbf{M}(0)|$ and $\gamma(0) = 0$. In addition, from the definition of the Fourier transform, this limit corresponds to the integral over x in real space, which by conservation of probability must be 1. Since the Laplace transform of 1 is $1/s$, this implies that $C(q = 0, s) = |\mathbf{R}(0, s)|/s$. Thus, although $C(q, s)$ depends on initial conditions, $C(0, s)$ is independent of $\mathbf{P}(0)$.

As detailed in Appendix A, applying eqs 7 and 8 to the position probability density, $\tilde{P}(q,s)$, in eq 9, given the behavior derived above, yields the simple expression

$$v = -i \frac{\dot{\gamma}(0)}{\beta(0)} \quad (11)$$

where $\dot{\gamma}(0)$ denotes the derivative with respect to q evaluated at $q = 0$. Thus, the velocity v depends only on the coefficients, β and γ, of the lowest two orders of the polynomial $|\mathbf{R}(q, s)|$, and is independent of $C(q, s)$. The same approach (see Appendix A) can be used to calculate the diffusion coefficient D,

$$D = \frac{\ddot{\gamma}(0) - 2iv\,\dot{\beta}(0) - 2\alpha(0)\,v^2}{2\beta(0)} \quad (12)$$

which depends only on the coefficients for the three lowest order coefficients of the determinant, α, β, and γ. Equations 11 and 12 were first derived by Koza[33,34] using similar arguments.

The randomness parameter is defined as $r = 2D/vd$ which from eqs 11 and 12 yields

$$r = \frac{i}{d}\left(\frac{\ddot{\gamma}(0)}{\dot{\gamma}(0)} - 2\frac{\dot{\beta}(0)}{\beta(0)} + 2\frac{\alpha(0)\,\dot{\gamma}(0)}{\beta(0)^2}\right) \quad (13)$$

In cases for which the step size is not uniform, the definition of the randomness parameter must be generalized to account for this variability. As suggested by Shaevitz et al.,[36] d can be replaced by $\langle d \rangle$, the average step size, and thus, $r = 2D/v\langle d \rangle$. However, as discussed in more detail below, the appropriate definition of r is ambiguous in certain circumstances, exposing a potential limitation to this kinetic parameter.

To illustrate the use of eqs 11 and 13, we return to the reversible n-state linear cycle in eq 1 and Figure 1b. Due to the predominant tridiagonal symmetry of the matrix $\mathbf{R}$ in this scheme, it is possible to determine these three coefficients explicitly and obtain general expressions for the velocity and diffusion constant for an arbitrary reversible linear cycle (see Appendix B). The velocity is given by

$$v = d \frac{1 - \prod_{l=1}^{n} K_l}{\sum_{i=1}^{n} \frac{1}{k_i}\left(1 + \sum_{j=i+1}^{n+i-1} \prod_{l=i+1}^{j} K_l\right)} \quad (14)$$
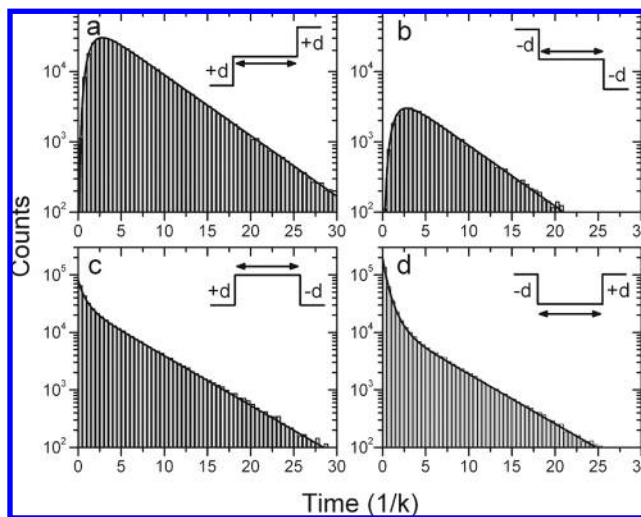

Figure 2. Dwell time distributions from a Monte Carlo simulation of a four-state fully reversible, linear, kinetic scheme, as in eq 1 and Figure 1b. Panels a–d show the different dwell time histograms for forward steps that were preceded by a forward step (a) or a backward step (d), or backward steps that were preceded by a backward step (b) or a forward step (c). Despite the apparent simplicity of the reversible linear scheme, its dynamics are described by multiple dwell time distributions that depend on the type of physical step taken both before and after the dwell. However, it is possible to calculate these dwell time distributions (solid lines) from the position probability density using the method described here. The rates used in the simulation are $k_1 = k_2 = k_3 = k_4 = k$, $\bar{k}_1 = \bar{k}_2 = \bar{k}_3 = k$, and $\bar{k}_4 = k/5$. Time is measured in units of $1/k$.

an expression first derived by Derrida[32] via a different method. Here, $K_j \equiv \bar{k}_{j-1}/k_j$ with the property that $K_{n+j} = K_j$ (note also that $K_1 = \bar{k}_n/k_1$). Due to its common usage, we derive an expression for the randomness parameter, $r = 2D/vd$, rather than the diffusion constant, D, applying eq 13,

$$r = \frac{1 + \prod_{l=1}^{n} K_l}{1 - \prod_{l=1}^{n} K_l} - 2[1 - \prod_{l=1}^{n} K_l] \times$$

$$\frac{\sum_{i=2}^{n}\sum_{j=1}^{i-1}\frac{1}{k_i}\frac{1}{k_j}(1 + \sum_{m=j+1}^{i-1}\prod_{l=j+1}^{m} K_l)(1 + \sum_{m=i+1}^{n+j-1}\prod_{l=i+1}^{m} K_l)}{\left[\sum_{i=1}^{n}\frac{1}{k_i}(1 + \sum_{j=i+1}^{n+i-1}\prod_{l=i+1}^{j} K_l)\right]^2} \quad (15)$$

which is also consistent with expressions derived by Derrida.[32] While these quantities have been calculated previously via different methods, these results illustrate that this matrix methodology can address kinetic systems of arbitrary size.

Connection to Dwell Time Distributions. We now extend the calculation of kinetic properties with this matrix formalism to the *dwell time distribution*, $\psi(t)$, the probability that the motor spends a time t at a single physical location on its track. It is important to first note that it is not always possible to describe a given kinetic scheme by a single dwell time distribution. For example, the reversible linear scheme that we have so far considered (eq 1 and Figure 1b), in which a motor can take both forward and backward steps, cannot be described by a single $\psi(t)$ (see Figure 2). To illustrate this point, consider the sequence of kinetic transitions that must occur during a dwell

Kinetic Models of Macromolecular Dynamics

J. Phys. Chem. B, Vol. 112, No. 19, 2008 **6029**

preceded by a forward step and followed by another forward step (Figure 2a). The motor starts in state 1 and makes multiple transitions to state n, where it generates a forward step, d, and returns to state 1 at the next position along the track. In contrast, in a dwell preceded by a forward step but followed by a backward step (Figure 2c), the motor starts in state 1 and may immediately take a backward step $-d$, or take excursions to any of the states 2 ... n before returning to state 1 and taking a backward step. Similar situations arise for the other two cases: dwells preceded by a backward step but followed by a forward (Figure 2d) or backward step (Figure 2b). Because the kinetic path taken in each case is different, it is clear that in general, the dwell times will be described by separate dwell time distributions. A similar argument was made recently by Tsygankov, Lindén, and Fisher[39] and Lindén and Wallin.[40]

Although we will derive a connection between the position probability and the four unique dwell time distributions required to describe the dynamics of the linear cycle considered above, we will first consider, for illustrative purposes, the situation in which the transition rate for the backward step is zero, $\bar{k}_n = 0$, and the system steps only forward (Figure 1a is an example of such a system). In this specialized case, there is only one dwell time distribution, since the motor can step only forward, and there is a simple one-to-one correspondence between the position probability density and the dwell time distribution, concisely written in Fourier−Laplace space,[41]

$$\tilde{P}(q, s) = \frac{1 - \tilde{\psi}(s)}{s(1 - \rho_+(q)\tilde{\psi}(s))} \tag{16}$$

where $\tilde{\psi}(s)$ is the Laplace transform of the dwell time distribution. Thus, if we calculate the full position probability density, $\tilde{P}(q,s)$, via eq 9, we can use this relation, eq 16, to solve for $\tilde{\psi}(s)$. $\psi(t)$ can then be generated from $\tilde{\psi}(s)$ using an inverse Laplace transform, although in general, this may not be necessary, since the moments of $\psi(t)$ can be found from derivatives of $\tilde{\psi}(s)$, as we will show below. Moreover, the calculation of the inverse transform involves finding the roots of an nth order polynomial, thus analytical solutions may not always exist.

To calculate $\tilde{P}(q, s)$ explicitly from eq 9, we must first choose the appropriate initial condition. Since the dwell time distribution describes the first passage time from one physical step to the next, the correct initial condition is $P_j(0) = \delta_{j1}$; that is, the system starts in state 1. In general, the appropriate initial condition is the state to which the motor steps (state 1 in Figure 1a and c, for instance). As derived in Appendix C, $\tilde{P}(q, s)$ for an n-state irreversible linear scheme takes a remarkably simple form,

$$\tilde{P}(q, s) = \frac{1}{s}\frac{|\mathbf{R}(0, s)|}{|\mathbf{R}(q, s)|} = \frac{s^{n-1} \cdots + \alpha s + \beta}{s^n \cdots + \alpha s^2 + \beta s + \gamma(q)} \tag{17}$$

In other words, $C(q, s) = C(0, s)$. The simple form of eq 17 is a direct consequence of the irreversibility of the stepping transition. For this cycle, the Fourier transform of the step size probability distribution is simply $\rho_+(q) = e^{-iqd}$, and application of eqs 16 and 17 yields

$$\tilde{\psi}(s) = \frac{|\mathbf{R}(q, s)| - |\mathbf{R}(0, s)|}{|\mathbf{R}(q, s)| - \rho_+(q)|\mathbf{R}(0, s)|} =$$

$$\frac{\gamma_+}{s^n \cdots + \alpha s^2 + \beta s + \gamma_+} \tag{18}$$

where $\gamma_+ = \gamma(q)/(1 - \rho_+(q)) = k_1 k_2 \cdots k_n$. In the limit $s = 0$, corresponding to the integral over all time, $\tilde{\psi}(s = 0) = 1$, and the dwell time distribution is properly normalized to unity. Equation 18 is a general, closed-form solution for the Laplace transform of the dwell time distribution for all linear cycles with an irreversible step (Figure 1a), and remarkably, it is only the characteristic polynomial of $\mathbf{M}$, and not the n eigenvalues of the system, that is needed to calculate $\tilde{\psi}(s)$. An elegant, alternate way of expressing $\tilde{\psi}(s)$ is to take the limit that $\rho_+(q) = 0$ in eq 16. It follows that $\tilde{\psi}(s) = 1 - s\tilde{P}(q,s)|_{\rho_+=0}$, which can be simplified to

$$\tilde{\psi}(s) = \frac{\gamma_+}{|\mathbf{R}'(s)|} \tag{19}$$

where $\mathbf{R}'(s)$ is the matrix $\mathbf{R}(q, s)$ in which $\rho_+(q) = 0$. Evaluating $|\mathbf{R}'(s)|$ reduces eq 19 to eq 18, and expressions for all of the coefficients in terms of kinetic rates can be found in Appendix B.

To extend this analysis to the fully reversible linear cycle, eq 1, a generalized form of eq 16 must be developed. The problem can be extended to that of any cycle involving multiple stepping mechanisms (for instance, forward or backward stepping or stepping by different distances), each attached to a different dwell time distribution (see Figure 1). The distinction must also be made between cycles in which the dwell time statistics depend on the type of step taken before and after the dwell (Figure 1b, the reversible cycle) or those in which the dwell time statistics depend on only the succeeding step (Figure 1c, in which different stepping branches lead to the same initial state). We term the former *correlated stepping*, in which the probability of taking one type of step depends on that of the preceding step, and the latter *uncorrelated stepping*, in which it does not. The treatment of correlated stepping requires the definition of a dwell time distribution matrix $\tilde{\boldsymbol{\psi}}(s)$, with elements $\tilde{\psi}_{ij}(s)$, in which the first index denotes the type of step preceding the dwell, and the second index similarly denotes the succeeding step. Using the reversible linear cycle as illustration,

$$\tilde{\boldsymbol{\psi}}(s) = \begin{pmatrix} p_{++}\tilde{\psi}_{++}(s) & p_{+-}\tilde{\psi}_{+-}(s) \\ p_{-+}\tilde{\psi}_{-+}(s) & p_{--}\tilde{\psi}_{--}(s) \end{pmatrix} \tag{20}$$

where $+$ and $-$ indicate a forward and backward step, respectively, and $p_{\pm\pm}$ represent the probability of taking a forward or backward step given the previous step, i.e., the *conditional branching probabilities*. For convenience in calculation, we define the matrix $\boldsymbol{\rho}(q)$ and the vector $\tilde{\boldsymbol{\Psi}}(s)$,

$$\boldsymbol{\rho}(q) = \begin{pmatrix} \rho_+(q) & 0 \\ 0 & \rho_-(q) \end{pmatrix}$$

$$\tilde{\boldsymbol{\Psi}}(s) = \frac{1}{s}\begin{pmatrix} 1 - p_{++}\tilde{\psi}_{++}(s) - p_{+-}\tilde{\psi}_{+-}(s) \\ 1 - p_{-+}\tilde{\psi}_{-+}(s) - p_{--}\tilde{\psi}_{--}(s) \end{pmatrix} \tag{21}$$

As derived in Appendix D, the position probability density and the multiple dwell time distributions can be related through the following matrix equation,

$$\tilde{P}(q, s) = \mathbf{p}_0^T(\mathbf{I} - \tilde{\boldsymbol{\psi}}(s)\boldsymbol{\rho}(q))^{-1}\tilde{\boldsymbol{\Psi}}(s) \tag{22}$$

where the vector $\mathbf{p}_0$ is the vector of initial conditions. In the reversible linear cycle, for instance, $\mathbf{p}_0^{\mathrm{T}} = (1\ 0)$ corresponds to the motor having taken an initial forward (+) step. This formalism is general and can be used to describe any kinetic system in which m types of steps are taken and the dwell time and the probability to take a given step depend on the type of step taken previously, such as the kinetic schemes depicted in Figure 1b and d. Furthermore, whereas the reversible linear cycle we consider as an example here satisfies the *bottleneck property* described by Lindén and Wallin[40] (in which forward and reverse steps occur during the forward and reverse transitions between the same two states), this formalism is general and can be used to address more complex systems with steps between multiple kinetic states, such as that pictured in Figure 1d.

In the limit that the motor takes only one type of step, only one element of the dwell time distribution matrix $\tilde{\psi}(s)$ survives, and eq 22 can be seen to reduce to eq 16, as expected. For kinetic cycles in which m types of steps can be taken but the dwell time statistics and the probability to take each type does not depend on the preceding step (i.e., uncorrelated systems such as the kinetic scheme shown in Figure 1c), eq 22 can also be simplified greatly (see Appendix D). Since there is no dependence on the preceding type of step, the first index i in the dwell time distribution matrix $\tilde{\psi}(s)$ can be eliminated (i.e., $p_{ij} \rightarrow p_j$ and $\psi_{ij} \rightarrow \psi_j$), and eq 22 simplifies to

$$\tilde{P}(q, s) = \frac{1 - \mathrm{Tr}(\tilde{\psi}(s))}{s(1 - \mathrm{Tr}(\tilde{\psi}(s)\boldsymbol{\rho}(q)))} \quad (23)$$

where the trace $\mathrm{Tr}(\tilde{\psi}(s)\boldsymbol{\rho}(q))$ denotes the sum $\sum_{i=1}^{m} p_i \tilde{\psi}_i(s)\rho_i(q)$, and p_i, $\psi_i(s)$, and $\rho_i(q)$ are the branching probability, the Laplace transform of the dwell time distribution, and Fourier transform of the step size distributions associated with the ith type of step, respectively.

For systems described by eqs 22 or 23, $\tilde{P}(q, s)$ is connected to m^2 or m dwell time distributions, respectively. Therefore, unlike simple kinetic schemes with only one type of step, extracting all the dwell time distributions from the position probability density involves a more complex procedure. We first treat the simpler case of the uncorrelated system described by eq 23. Manipulating this equation to obtain the quantity $(s\tilde{P}(q, s))^{-1}$ and taking the limit that all except the ith step size distribution goes to zero, $\rho_{j\neq i}(q) = 0$, yields the simple form

$$\left.\frac{1}{s\tilde{P}(q, s)}\right|_{\{\rho_{j\neq i}=0\}} = \frac{1 - p_i\tilde{\psi}_i(s)\rho_i(q)}{1 - \sum_{l=1}^{m} p_l\tilde{\psi}_l(s)} = a_0(s) + a_i(s)\,\rho_i(q) \quad (24)$$

which consists of one term independent of $\rho_i(q)$ and one linear in $\rho_i(q)$. Substituting the calculated expression for $\tilde{P}(q, s)$ for that particular cycle and applying these limits, one can solve for a_0 and a_i, matching terms by their dependence on $\rho_i(q)$, and determine the dwell time distribution by $p_i\tilde{\psi}_i(s) = -a_i/a_0$. The individual branching probabilities, p_i can be extracted using the fact that the dwell time distributions must be normalized, that is, $\tilde{\psi}_i(0) = 1$. Repeating this procedure for all $i = 1 \cdots m$ determines the other dwell time distributions and branching probabilities.

A similar procedure can be applied to calculate the dwell time distributions for the correlated stepper governed by eq 22. For illustrative purposes, we consider the fully reversible linear system. We first choose the initial condition $\mathbf{p}_0^{\mathrm{T}} = (1\ 0)$, corresponding to an initial forward step, and accordingly solve

for $\tilde{P}(q, s)$ with state 1 as the proper initial condition (see Appendix C), $\tilde{P}_+(q, s)$. Taking the limit that $\rho_-(q) = 0$, $(s\tilde{P}_+(q, s))^{-1}$ takes the form

$$\left.\frac{1}{s\tilde{P}_+(q, s)}\right|_{\rho_-=0} = \frac{1 - p_{++}\tilde{\psi}_{++}(s)\rho_+(q)}{1 - p_{++}\tilde{\psi}_{++}(s) - p_{+-}\tilde{\psi}_{+-}(s)}$$

$$= a_0(s) + a_+(s)\rho_+(q) \quad (25)$$

similar to that found above, and we can solve for $p_{++}\tilde{\psi}_{++}(s)$ in an identical fashion, using $p_{++}\tilde{\psi}_{++}(s) = -a_+(s)/a_0(s)$. Choosing the other condition, $\mathbf{p}_0^{\mathrm{T}} = (0\ 1)$, for an initial backward step, and solving for $\tilde{P}(q, s)$ with state n as the appropriate initial condition, $\tilde{P}_-(q, s)$, one can solve for $p_{--}\tilde{\psi}_{--}(s)$ in an analogous fashion. For a general cycle with m types of steps, this procedure can be used to extract all the diagonal elements of the dwell time distribution matrix $\tilde{\psi}(s)$. However, the remaining $m(m - 1)$ off-diagonal dwell time distributions cannot in general be extracted by this method. Fortunately, for the reversible cycle considered here where $m = 2$, the solution to $p_{++}\tilde{\psi}_{++}(s)$ makes it possible to also determine the off-diagonal element $p_{+-}\tilde{\psi}_{+-}(s)$, since the above expression can be solved for $p_{+-}\tilde{\psi}_{+-}(s)$,

$$p_{+-}\tilde{\psi}_{+-}(s) =$$
$$1 - s\tilde{P}_+(q, s)|_{\rho_-=0} - p_{++}\tilde{\psi}_{++}(s)(1 - s\tilde{P}_+(q, s)|_{\rho_-=0}\,\rho_+(q)) =$$
$$\frac{a_0(s) + a_+(s) - 1}{a_0(s)} \quad (26)$$

Analogously, using $\tilde{P}_-(q, s)$ and $p_{--}\tilde{\psi}_{--}(s)$ yields $p_{-+}\tilde{\psi}_{-+}(s)$.

Using this procedure, we can write exact and remarkably simple expressions for the four dwell time distributions of the reversible linear cycle, eq 1,

$$p_{++}\tilde{\psi}_{++}(s) = \frac{\gamma_+}{|\mathbf{R}'(s)|}$$

$$p_{--}\tilde{\psi}_{--}(s) = \frac{\gamma_-}{|\mathbf{R}'(s)|}$$

$$p_{+-}\tilde{\psi}_{+-}(s) = \frac{\bar{k}_n|\mathbf{R}_-(s)|}{|\mathbf{R}'(s)|}$$

$$p_{-+}\tilde{\psi}_{-+}(s) = \frac{k_n|\mathbf{R}_+(s)|}{|\mathbf{R}'(s)|} \quad (27)$$

where $\gamma_- = \bar{k}_1\bar{k}_2 \cdots \bar{k}_n$ and $\mathbf{R}_+(s)$ and $\mathbf{R}_-(s)$ are defined as $n - 1 \times n - 1$ submatrices of $\mathbf{R}(q, s)$, in which the last row and column have been removed and in which the first row and column have been removed, respectively (see Appendices B and C). By taking $s = 0$, we find that the branching probabilities are

$$p_{++} = \frac{\gamma_+}{\gamma_+ + \gamma_- + \gamma_{+-}}$$

$$p_{--} = \frac{\gamma_-}{\gamma_+ + \gamma_- + \gamma_{+-}}$$

$$p_{+-} = \frac{\gamma_- + \gamma_{+-}}{\gamma_+ + \gamma_- + \gamma_{+-}}$$

$$p_{-+} = \frac{\gamma_+ + \gamma_{+-}}{\gamma_+ + \gamma_- + \gamma_{+-}} \quad (28)$$

Kinetic Models of Macromolecular Dynamics

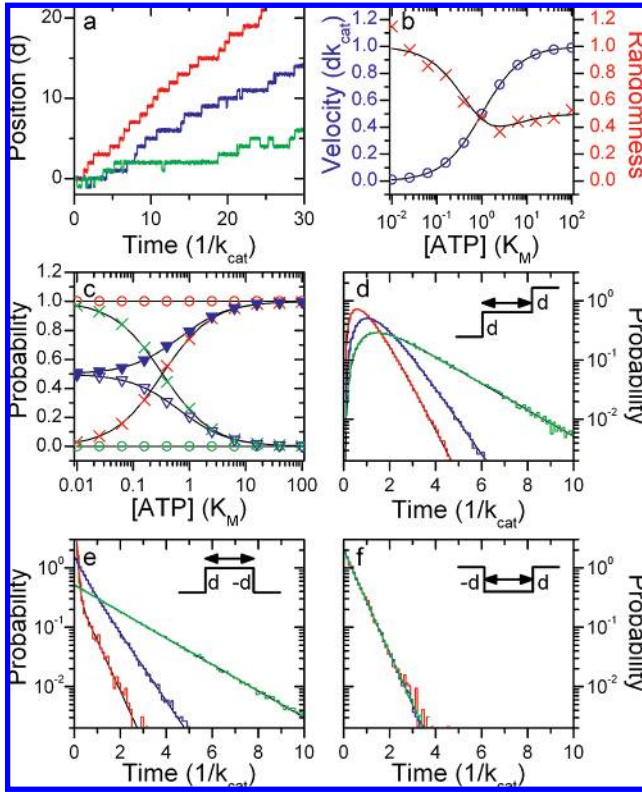*J. Phys. Chem. B, Vol. 112, No. 19, 2008* **6031**

Figure 3. Simulated experimental data for the kinetic scheme in model 3. (a) Observed motor motion measured in units of the step size, d, and $1/k_{cat}$ for three concentrations of ATP: $1/10K_M$, K_M, and $10K_M$ plotted in green, blue, and red respectively. (b) Velocity (blue circles) and randomness parameter, $r = 2D/vd$ (red crosses), measured for different concentrations of ATP. Solid lines are predictions from the theory in model 3. (c) Branching probabilities p_{++}, p_{+-}, p_{-+}, p_{--}, p_+, and p_- plotted as red crosses, green crosses, red circles, green circles, solid blue triangles, and open blue triangles, respectively, as a function of [ATP]. Solid lines are predictions from model 3. (d) $\psi_{++}(t)$, (e) $\psi_{+-}(t)$, and (f) $\psi_{-+}(t)$ plotted for [ATP] = $1/10K_M$, K_M, and $10K_M$ in green, blue, and red, respectively. The solid black lines are predictions from model 3. All data were simulated using model 3 with $k_1 = 36/25k_{cat}$, $k_2 = k_3 = 2k_{cat}$, and $\bar{k}_1 = \bar{k}_3 = 2/5k_{cat}$, and dwell time distributions were computed from $\sim 10^6$ total steps.

where $\gamma_{+-} = k_1 k_2 \cdots k_n \sum_{m=1}^{n-1} \prod_{l=1}^{m} K_l$ and $\gamma_\pm$ are defined above and in Appendix B. These expressions are consistent with those calculated previously by a different method.[39] We use eq 27 to produce the predictions for the four different dwell time distributions plotted in Figure 2.

Example Calculations and Modeling. The different kinetic parameters described above (mean velocity, randomness parameter, branching probabilities, and dwell time distributions) are commonly used parameters in analyzing molecular motor data. Their ability to discriminate between possible kinetic models to describe the measured data, however, varies greatly. In this section, we illustrate this point by considering possible kinetic models for a set of simulated data. We demonstrate not only the usefulness of this formalism in calculating the desired parameters of different kinetic models, but also the ways in which these parameters can be used to aid an experimenter in assessing possible kinetic models for a set of experimental data. We thus imagine an experiment that measures a molecular motor stepping along its track at varying ATP concentrations, as depicted in Figure 3a.

Model Pathway 1: Linear Cycle. We assume that, initially, the experimenter determines only the velocity and the randomness parameter as measures of the kinetics of the motor, plotted

TABLE 1: Coefficients of $|\mathbf{R}(q, s)|$ for Examples 1–3

example		
1	σ	$k_1[T]k_2 k_3$
	$\gamma(q)/\sigma$	$1 - \rho(q)$
	β/σ	$1/(k_1[T])(1 + \bar{k}_1/k_2) + 1/k_2 + 1/k_3$
	α/σ	$1/(k_1[T]k_2) + 1/(k_1[T]k_3)(1 + \bar{k}_1/k_2) + 1/(k_2 k_3)$
2	σ	$k_1[T](k_{2+} + k_{2-})k_{3+}k_{3-}$
	$\gamma(q)/\sigma$	$p_+(1 - \rho_+(q)) + p_-(1 - \rho_-(q))$
	$\beta(q)/\sigma$	$p_+\{(1 - \rho_+(q))/k_{3-} + 1/(k_1[T])(1 + \bar{k}_1/(2k_{2+})) +$ $1/(2k_{2+}) + 1/k_{3+}\} + p_-\{(1 - \rho_-(q))/k_{3+} +$ $1/(k_1[T])(1 + \bar{k}_1/(2k_{2-})) + 1/(2k_{2-}) + 1/k_{3-}\}$
	α/σ	$p_+\{1/(k_1[T]k_{3-}) + 1/(k_{3+}k_{3-}) + 1/(2k_1[T]k_{2+}) +$ $1/(k_1[T]k_{3+})(1 + \bar{k}_1/k_{2+}) + 1/(k_{2+}k_{3+})\} +$ $p_-\{1/(k_1[T]k_{3+}) + 1/(k_{3+}k_{3-}) + 1/(2k_1[T]k_{2-}) +$ $1/(k_1[T]k_{3-})(1 + \bar{k}_1/k_{2-}) + 1/(k_{2-}k_{3-})\}$
3	σ	$k_1[T]k_2 k_3$
	$\gamma(q)/\sigma$	$1 - \rho_+(q) + \bar{k}_3/(k_1[T])(1 + \bar{k}_1/k_2)(1 - \rho_+(q) \rho_-(q))$
	$\beta(q)/\sigma$	$\bar{k}_3/(k_1[T]k_2)(1 - \rho_+(q) \rho_-(q)) + 1/(k_1[T])(1 +$ $\bar{k}_1/k_2)(1 + \bar{k}_3/k_3) + 1/k_2 + 1/k_3$
	α/σ	$1/(k_1[T]k_2)(1 + \bar{k}_3/k_3) + 1/(k_1[T]k_3)(1 + \bar{k}_1/k_2) +$ $1/(k_2 k_3)$

in Figure 3b. To describe this data, it is sensible to first consider a simple kinetic model that contains only the minimal number of chemical states during hydrolysis,

$$E_1 \leftrightarrow T_2 \rightarrow D_3 \xrightarrow{d} E_1 \tag{29}$$

with three states corresponding to empty (E), ATP-bound (T), and post-hydrolysis (D). In this model, we assume that hydrolysis is followed by fast, irreversible phosphate release, and the two transitions are combined into one. Furthermore, we assume for simplicity that the translocation step (denoted by the right arrow superscripted with the letter d) occurs simultaneously with the release of ADP, here also assumed to be irreversible. For simplicity of analysis, we consider only the ATP binding transition to be reversible.

To use the above formalism to determine the kinetic parameters of interest, we first write the rate matrix **M**, corresponding to this cycle, by applying the master equation eq 4, ensuring that transitions corresponding to a physical step of the motor, such as that from state D_3 into E_1, are multiplied by $\rho(q) = e^{-iqd}$ for a uniform step size, d. Thus, for this model cycle

$$\mathbf{M}(q) = \begin{pmatrix} -k_1[T] & \bar{k}_1 & k_3\rho(q) \\ k_1[T] & -(k_2 + \bar{k}_1) & 0 \\ 0 & k_2 & -k_3 \end{pmatrix} \tag{30}$$

where we have expressed the rate coefficients in pseudo-first-order notation, such that the second-order ATP binding reaction in the cycle is written as a second-order rate constant multiplied by the substrate concentrations [T].

The position probability density can be solved by the procedure outlined above. The appropriate initial condition is the state E_1, since this is the first state visited after the motor steps. Carrying out the algebra, $\tilde{P}(q, s)$ takes the form denoted in eq 17 with $n = 3$,

$$\tilde{P}(q,s) = \frac{s^2 + \alpha s + \beta}{s^3 + \alpha s^2 + \beta s + \gamma(q)} \tag{31}$$

in which the coefficients α, β, and γ are solved by calculating the determinant $|\mathbf{R}(q, s)|$ and by collecting terms of common order in s. These coefficients are listed in Table 1,

TABLE 2: Macroscopic Michaelis−Menten Parameters in Terms of Kinetic Rates

example		
1	$1/k_{cat}$	$1/k_2 + 1/k_3$
	K_M/k_{cat}	$1/k_1(1 + \bar{k}_1/k_2)$
2	$1/k_{cat}$	$p_+(1/(2k_{2+}) + 1/k_{3+}) + p_-(1/(2k_{2-}) + 1/k_{3-})$
	K_M/k_{cat}	$p_+/k_1(1 + \bar{k}_1/(2k_{2+})) + p_-/k_1(1 + \bar{k}_1/(2k_{2-}))$
3	$1/k_{cat}$	$1/k_2 + 1/k_3$
	K_M/k_{cat}	$1/k_1(1 + \bar{k}_1/k_2)(1 + \bar{k}_3/k_3)$

TABLE 3: Macroscopic Randomness Parameters in Terms of Kinetic Rates

example		
1	$1/q^2_{cat}$	$1/(k_2 k_3)$
	Q_M/q^2_{cat}	$1/(k_1 k_2) + 1/(k_1 k_3)(1 + \bar{k}_1/k_2)$
2	$1/q^2_{cat}$	$p^2_+/(k_{2+} k_{3+}) + p^2_-/(k_{2-} k_{3-}) - p_+ p_-(1/k_{3+} - 1/k_{3-})^2$
	Q_M/q^2_{cat}	$p_+(1/(2k_1 k_{2+}) + 1/(k_1 k_{3+}))(1 + p_+\bar{k}_1/k_{2+})) + p_-(1/(2k_1 k_{2-}) + 1/(k_1 k_{3-}))(1 + p_-\bar{k}_1/k_{2-}))$
	$1/q'_{cat}$	$p_+ p_-(1/k_{3-} - 1/k_{3+})$
3	$1/q^2_{cat}$	$1/(k_2 k_3)$
	Q_M/q^2_{cat}	$1/(k_1 k_2)(1 + \bar{k}_3/k_3) + 1/(k_1 k_3)(1 + \bar{k}_1/k_2)$

divided for clarity by the product of the forward rates $\sigma = k_1[T]k_2 k_3$.

Mean Velocity and Randomness Parameter. The mean velocity can be calculated from these coefficients by eq 11. Given the simplicity of the kinetic cycle, the velocity has the expected Michaelis−Menten form,

$$v = d\frac{k_{cat}[T]}{K_M + [T]} \quad (32)$$

which is hyperbolic in ATP concentration, asymptotically reaching a maximum value of dk_{cat} for $[T] \gg K_M$ and a half-maximal value when $[T] = K_M$ (solid line in Figure 3b). How these macroscopic kinetic parameters relate to the microscopic rate constants in this example cycle is evident by manipulating the Michaelis−Menten equation into its Lineweaver−Burke form,

$$\frac{d}{v} = \frac{K_M}{k_{cat}[T]} + \frac{1}{k_{cat}} \quad (33)$$

which, by eq 11, is equal to $id\beta/\dot{\gamma}(0)$. Using the coefficients from Table 1 and matching terms with the above expression, it follows that the ATP-dependent term in β/σ is equal to K_M/k_{cat} and the ATP-independent term corresponds to $1/k_{cat}$. Table 2 lists these parameters explicitly.

A similar analysis can be carried out for the randomness parameter using eq 13. Unlike the case of the mean velocity, there are no traditionally defined macroscopic kinetic constants for the randomness parameter. However, to express r in a concise form, we have found it necessary to define two constants, q_{cat} and Q_M, which, in analogy to k_{cat} and K_M in the Michaelis−Menten equation, represent an effective rate constant and a binding equilibrium constant, respectively. The connection between these constants and the microscopic rates and their physical meaning will be discussed shortly; for now, their definition allows r to be expressed in a simple, Michaelis−Menten-like equation

$$r = 1 - 2\frac{k^2_{cat}}{q^2_{cat}}\frac{[T](Q_M + [T])}{(K_M + [T])^2} \quad (34)$$

With some manipulation of this equation, its relation to the coefficients α, β, and γ and the rate constants of the cycle is more transparent.

$$r = 1 - 2\frac{\dfrac{Q_M}{q^2_{cat}[T]} + \dfrac{1}{q^2_{cat}}}{\left(\dfrac{K_M}{k_{cat}[T]} + \dfrac{1}{k_{cat}}\right)^2} \quad (35)$$

Here, the denominator is clearly the expression for d/v presented above, squared. From eq 13, it follows that the numerator is

$id\alpha/\dot{\gamma}(0)$. Substituting these coefficients from Table 1 and comparing terms with those of the numerator above, it follows that Q_M/q^2_{cat} and $1/q^2_{cat}$ correspond to the ATP-dependent and -independent terms of α/σ, respectively. Both constants are listed in Table 3.

The physical meaning of these two new macroscopic kinetic parameters is revealed from the dependence of r on ATP concentration. At $[T] \gg K_M$, in which the catalytic transitions in the cycle are rate-limiting, r takes a value of $r_\infty = 1 - 2k^2_{cat}/q^2_{cat} \leq 1$. r_∞ and by extension q_{cat} are thus related to the number of rate-limiting catalytic transitions in the cycle.[15] For instance, taking the limit that hydrolysis is very fast ($k_2 \rightarrow \infty$), it follows that $q_{cat} \rightarrow \infty$ and $1/r_\infty \approx 1$, corresponding to one rate-limiting step. At the ATP concentration $[T]_{min} = K_M Q_M/(Q_M - 2K_M)$, r drops to a minimum value of

$$r_{min} = 1 - \frac{k^2_{cat}}{2q^2_{cat}}\frac{Q_M}{K_M}\frac{Q_M}{(Q_M - K_M)} \quad (36)$$

Here, $[T]_{min}$ represents the concentration at which the catalytic transitions and ATP binding are equally rate-limiting. In the case that there is only one rate-limiting catalytic transition, $Q_M \rightarrow \infty$ and $[T]_{min} = K_M$, since this also corresponds to the point of half-maximal velocity. In general, $[T]_{min} > K_M$ if there are additional catalytic rate-limiting transitions. The minimal value r_{min} places a bound on the total number of rate-limiting transitions in the cycle; in the case of a single catalytic rate-limiting transition, $1/r_{min} \approx 2$. Finally, at low $[T] \ll K_M$, $r_0 = 1$, which corresponds to ATP binding being the one rate-limiting step in the reaction. Thus, the randomness parameter can reveal additional details of the kinetic scheme that may not be apparent in the average velocity.

Dwell Time Distribution. Although not measured directly by the experimenter in this imagined scenario, it is still useful to calculate the dwell time distribution for this system because it is an example of the simplest type of stepping mechanism. As detailed above, the Laplace transform of the dwell time distribution, $\tilde{\psi}(s)$, is determined from the position probability density by applying eq 16 to eq 31, yielding

$$\tilde{\psi}(s) = \frac{\gamma_+}{s^3 + \alpha s^2 + \beta s + \gamma_+} \quad (37)$$

where the coefficient γ_+ is defined from the relation $\gamma(q) = \gamma_+(1 - e^{-iqd})$.

Due to the nature of the Laplace transform, it is possible to calculate moments of the dwell time distribution, $\psi(t)$, from derivatives of $\tilde{\psi}(s)$ evaluated at $s = 0$. The first two moments are found via

Kinetic Models of Macromolecular Dynamics

J. Phys. Chem. B, Vol. 112, No. 19, 2008 **6033**

$$\langle t \rangle = -\frac{d\tilde{\psi}}{ds}\Big|_{s=0}$$

$$\langle t^2 \rangle - \langle t \rangle^2 = \frac{d^2\tilde{\psi}}{ds^2}\Big|_{s=0} - \left(\frac{d\tilde{\psi}}{ds}\right)^2\Big|_{s=0} \tag{38}$$

Using eqs 37 and 38, we find that $\langle t \rangle = \beta/\gamma_+$ and $\langle t^2 \rangle - \langle t \rangle^2 = \beta^2/\gamma_+^2 - 2\alpha/\gamma_+$. Because this system has only one step size and can be described by a single dwell time distribution, the corresponding mean velocity is

$$v = \frac{d}{\langle t \rangle} \tag{39}$$

As demonstrated previously,[15] and detailed in Appendix E, in the case of a uniform step size, the randomness parameter r can also be expressed in terms of the variance and mean of the dwell times by the relation

$$r = \frac{\langle t^2 \rangle - \langle t \rangle^2}{\langle t \rangle^2} \tag{40}$$

Using the values in Tables 1 and 2, it is simple to show that these expressions agree with the values derived above.

Because of the simple form of this model kinetic cycle, the inverse Laplace transform can be determined analytically, yielding an exact solution to the dwell time distribution in the time domain. Substituting the expressions for the coefficients α, β, and γ_+ into eq 36, $\tilde{\psi}(s)$ can be written explicitly as

$$\tilde{\psi}(s) = \frac{k_1[T]k_2 k_3}{(s^2 + s(k_1[T] + \bar{k}_1 + k_2) + k_1[T]k_2)(s + k_3)} \tag{41}$$

Applying the inverse Laplace transform to this expression reveals a triexponential dwell time distribution with decay rates $\lambda_{1,2}$ and k_3,

$$\psi(t) = \frac{k_1[T]k_2 k_3}{(\lambda_2 - \lambda_1)(k_3 - \lambda_1)}e^{-\lambda_1 t} + \frac{k_1[T]k_2 k_3}{(\lambda_1 - \lambda_2)(k_3 - \lambda_2)}e^{-\lambda_2 t} +$$
$$\frac{k_1[T]k_2 k_3}{(\lambda_1 - k_3)(\lambda_2 - k_3)}e^{-k_3 t} \tag{42}$$
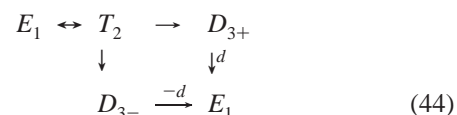
where $\lambda_{1,2}$ are given by

$$\lambda_{1,2} = \frac{1}{2}(k_1[T] + \bar{k}_1 + k_2) \pm$$
$$\frac{1}{2}\sqrt{(k_1[T] + \bar{k}_1 + k_2)^2 - 4k_1[T]k_2} \tag{43}$$

It should again be emphasized that it will not always be possible to perform the inverse Laplace transform analytically, especially with an increasing number of kinetic states.

Comparison to Data. Having calculated the velocity and randomness parameter, the experimenter can evaluate this model against the observed data. Since many different kinetic schemes with different numbers of states can produce a Michaelis−Menten-like velocity,[4] this velocity measurement contains little information on the number or connectivity of the states in the kinetic scheme. The randomness parameter, on the other hand, provides a lower limit on the number of kinetic states,[15] both binding and catalytic. Figure 3b illustrates that the measured randomness parameter as a function of [ATP] is well described by the macroscopic randomness parameter derived above, eq 33. The asymptotic value of the randomness parameter is $r_\infty \approx$

1/2, which suggests that at saturating [ATP], only two states are rate-limiting and that they have comparable transition rates; in this case, $k_2 = k_3 \approx 2k_{cat}$. The minimum value of the randomness parameter, on the other hand, is $r_{min} \approx 1/2.5$. A value higher than 1/3 suggests that binding is reversible, since reversible transitions tend to increase r,[15] a result that is again consistent with the model assumed here. The randomness parameter, however, does not dictate the number of kinetic states, only a lower limit, so at this point, the experimenter can only conclude that the data is consistent with this simple model, but cannot rule out more complicated models.

Model Pathway 2: Branched Cycle with Two Stepping Pathways. Closer inspection of the motor stepping traces reveals that the motor not only steps forward but also occasionally takes backward steps (see blue and green traces, Figure 3a), a result that is clearly inconsistent with the previous model. To account for the backward steps, our experimenter next considers a model cycle with two parallel branches, one to account for each direction of stepping. Upon binding ATP, the molecular motor may select one of two branches, denoted $+$ and $-$, in which translocation occurs in forward steps, d, or backward steps, $-d$:

$$
\begin{array}{ccc}
E_1 & \leftrightarrow\ T_2 & \rightarrow\ D_{3+} \\
& \downarrow & \downarrow d \\
& D_{3-} & \xrightarrow{-d}\ E_1
\end{array}
\tag{44}
$$

As before, translocation in each branch occurs during an irreversible ADP release transition. This scheme could describe a motor with two different "gears", allowing it to step forward by d or backward by $-d$, or more generally, by steps of different size d_+ and d_-. In the limit that one step size is set to zero, for instance, the corresponding branch could represent a pathway for futile hydrolysis, in which the energy from ATP hydrolysis is not utilized to generate mechanical motion. Note that this model allows the rates of each branch to be distinct.

We calculate the kinetic parameters for this cycle by writing the rate matrix, as before, enumerating the matrix elements in the order: 1, 2, 3+, 3−:

$$\mathbf{M}(q) = \begin{pmatrix} -k_1[T] & \bar{k}_1 & k_{3+}\rho_+(q) & k_{3-}\rho_-(q) \\ k_1[T] & -(k_{2-} + k_{2+} + \bar{k}_1) & 0 & 0 \\ 0 & k_{2+} & -k_{3+} & 0 \\ 0 & k_{2-} & 0 & -k_{3-} \end{pmatrix} \tag{45}$$

where $\rho_\pm(q) = e^{\mp iqd}$ are the Fourier transforms of uniform forward $(+)$ and backward $(-)$ steps of size d.

Although there are now two stepping transitions, the position probability density $\tilde{P}(q, s)$ can be shown by the same procedure outlined in Appendix C to take the form

$$\tilde{P}(q, s) = \frac{s^3 \cdots + \alpha s + \beta(0)}{s^4 \cdots + \alpha s^2 + \beta(q)s + \gamma(q)} \tag{46}$$

in which the proper initial condition is state E_1. In contrast to the linear cycle discussed in model 1, a feature of this branched pathway is that both coefficients β and γ depend on q. For purposes that will be evident further below, we express in Table 1 the coefficients α, $\beta(q)$, and $\gamma(q)$ divided by the factor $\sigma = k_1[T](k_{2+} + k_{2-})k_{3+}k_{3-}$ and define the quantities $p_\pm$ as

$$p_\pm = \frac{k_{2\pm}}{k_{2+} + k_{2-}} \tag{47}$$

6034 *J. Phys. Chem. B, Vol. 112, No. 19, 2008*

Chemla et al.

such that $p_+ + p_- = 1$. As we will demonstrate below, these represent the probabilities of taking the + or − branch. It can be verified that taking the limits $k_{2-} = k_{3-} = 0$, equivalent to removing the two transitions between states T_2 and D_{3-} and D_{3-} and E_1 in the cycle, $p_+ = 1$ and $p_- = 0$, and the coefficients reduce to those for the linear cycle in model 1, as expected.

Mean Velocity and Randomness Parameter. Calculating the mean velocity from the position probability density by eq 11 again yields a simple Michaelis−Menten expression,

$$v = \langle d \rangle \frac{k_{cat}[T]}{K_M + [T]} \tag{48}$$

where $\langle d \rangle = (p_+ - p_-)d$. The parameters k_{cat} and K_M can be related to the rate constants of the cycle by following the same procedure as in model 1 and are listed in Table 2.

For this cycle, the appropriate definition for the randomness parameter is ambiguous. Since the motor can step forward and backward, r may be defined either in terms of the magnitude of the step size, d, or its average, $\langle d \rangle$, at the discretion of the experimenter. In a more general case where the step sizes in each branch, $d_\pm$, are arbitrary, the definition using $\langle d \rangle$ would appear necessary. Here, however, since the magnitude of the step is the same, only its direction is different, it is possible to use $r = 2D/vd$, which is the definition used in Figure 3b. In this fashion, the randomness parameter for the branched system has a general form very different from that found in the previous model,

$$r = \frac{1}{p_+ - p_-} - 2\frac{k_{cat}^2}{q_{cat}^2}\frac{[T](Q_M + [T])}{(K_M + [T])^2}(p_+ - p_-)$$
$$+ 4\frac{k_{cat}}{q'_{cat}}\frac{[T]}{K_M + [T]} \tag{49}$$

with Q_M/q_{cat}^2, $1/q_{cat}^2$, and $1/q'_{cat}$ listed in Table 3. Note that in contrast to the randomness parameter in the previous model, $r_0 > 1$ in the limit that $[T] \ll K_M$, since $p_+ - p_- < 1$.

Dwell Time Distributions. In this example, it is clear that there is no unique dwell time distribution to describe the stepping time statistics because there are two separate branches from which the motor can translocate, each with a unique lifetime. It is thus necessary to utilize a generalized form of eq 16 to connect the position probability density, $\tilde{P}(q, s)$, to multiple dwell time distributions. Inspection of this cycle reveals that the choice of dwell time distribution depends only on the step succeeding the dwell; thus, the steps in this system are uncorrelated, and only two distributions $\tilde{\psi}_\pm(s)$ are necessary, each corresponding to one of the branches. (As an interesting aside, if the stepping transitions occur between the kinetic states T_2 and $D_{3\pm}$, the statistics of the steps are correlated, and four dwell time distributions are needed to describe the behavior of the system.) The appropriate generalization of eq 16 to use for uncorrelated systems is eq 23, which in this case becomes

$$\tilde{P}(q, s) = \frac{1 - p_+\tilde{\psi}_+(s) - p_-\tilde{\psi}_-(s)}{s(1 - \rho_+(q)p_+\tilde{\psi}_+(s) - \rho_-(q)p_-\tilde{\psi}_-(s))} \tag{50}$$

As we will demonstrate shortly, the quantities $p_\pm$ are identical in form to those defined above in eq 47. Using the method described above to determine dwell time distributions from the position probability density for a system with multiple stepping branches, we find that

$$p_\pm\tilde{\psi}_\pm(s) = \frac{\beta_\pm s + \gamma_\pm}{s^4 \cdots + \alpha s^2 + (\beta_+ + \beta_- + \beta(0))s + (\gamma_+ + \gamma_-)} \tag{51}$$

where the coefficients $\gamma_\pm$, $\beta_\pm$, and $\beta(0)$ are related to $\gamma(q)$ and $\beta(q)$ by the relations $\gamma(q) = \gamma_+(1 - e^{-iqd}) + \gamma_-(1 - e^{+iqd})$ and $\beta(q) = \beta_+(1 - e^{-iqd}) + \beta_-(1 - e^{+iqd}) + \beta(0)$. Since $\tilde{\psi}_\pm(0) = 1$, taking the limit $s = 0$ reveals the form of the branching probabilities $p_\pm$,

$$p_\pm = \frac{\gamma_\pm}{\gamma_+ + \gamma_-} \tag{52}$$

which reproduces the values in eq 47. Thus, the dwell time distribution for each branch and the probabilities of taking each branch easily derive from the position probability density, illustrating the power and usefulness of this formalism.

Because there are two dwell time distributions, the connection between their moments and the velocity as derived from the position probability density is more complex than in model 1. Nevertheless, it is easy to demonstrate that v can also be calculated from the first moments $\langle t_\pm \rangle$ of the distributions $\tilde{\psi}_\pm(s)$. As detailed in Appendix E, the velocity is given by

$$v = \frac{\langle d \rangle}{\langle t \rangle} \tag{53}$$

as expected, since the average time to complete the cycle is simply $\langle t \rangle = p_+\langle t_+ \rangle + p_-\langle t_- \rangle$.

The relation between the randomness parameter of the entire cycle and the moments of the individual dwell time distributions is also more complex than in the previous examples (see Appendix E). It can be shown that

$$r = \frac{1}{p_+ - p_-} + \frac{\langle t^2 \rangle}{\langle t \rangle^2}(p_+ - p_-) - 2\frac{p_+\langle t_+ \rangle - p_-\langle t_- \rangle}{\langle t \rangle} \tag{54}$$

where $\langle t^2 \rangle = p_+\langle t_+^2 \rangle + p_-\langle t_-^2 \rangle$. A similar relation containing the first two terms of eq 54 was derived previously;[6] in contrast to these published results, eq 54 also contains a third term with a more complex dependence on branching probabilities and dwell times, which directly results from the fact that steps occur in different branches with different kinetics. In the limit that the mean time in each branch $\langle t_\pm \rangle$ is identical, the step size is then uncoupled from the dwell time: only one dwell time distribution is necessary to describe the stepping statistics of the motor, and the expressions for r simplify to those derived previously.[6]

Again, due to the simplicity of the model kinetic cycle, it is possible to calculate the inverse Laplace transform of $p_\pm\tilde{\psi}_\pm(s)$ directly. Substituting the coefficients into eq 51, it is possible to write

$$p_\pm\tilde{\psi}_\pm(s) =$$
$$\frac{k_1[T]k_{2\pm}k_{3\pm}}{(s^2 + s(k_1[T] + \bar{k}_1 + k_{2+} + k_{2-}) + k_1[T](k_{2+} + k_{2-}))(s + k_{3\pm})} \tag{55}$$

which, upon taking the inverse Laplace transform, yields the following triexponential,

Kinetic Models of Macromolecular Dynamics

J. Phys. Chem. B, Vol. 112, No. 19, 2008 **6035**

$$p_\pm \psi_\pm(t) = \frac{k_1[T]k_{2\pm}k_{3\pm}}{(\lambda_2 - \lambda_1)(k_{3\pm} - \lambda_1)}e^{-\lambda_1 t} +$$

$$\frac{k_1[T]k_{2\pm}k_{3\pm}}{(\lambda_1 - \lambda_2)(k_{3\pm} - \lambda_2)}e^{-\lambda_2 t} +$$

$$\frac{k_1[T]k_{2\pm}k_{3\pm}}{(\lambda_1 - k_{3\pm})(\lambda_2 - k_{3\pm})}e^{-k_{3\pm}t} \quad (56)$$

where

$$\lambda_{1,2} = \frac{1}{2}(k_1[T] + \bar{k}_1 + k_{2+} + k_{2-}) \pm$$

$$\frac{1}{2}\sqrt{(k_1[T] + \bar{k}_1 + k_{2+} + k_{2-})^2 - 4k_1[T](k_{2+} + k_{2-})} \quad (57)$$

Comparison to Data. With the calculated kinetic parameters, the experimenter can now determine if such a model is consistent with the observed data. Although the velocity has the same Michaelis−Menten form as in the previous model, the randomness parameter has a more complex behavior. Notably, the asymptotic value of r at low [ATP] exceeds 1 for this model, which is not consistent with the observed data (see Figure 3b). However, it should be noted that measurements of the randomness parameter can be noisy, and it may not always be simple to determine if the asymptotic value of r is consistent with a value equal to or larger than 1, as seen in Figure 3b.

A more unequivocal test of this model is to consider the probabilities of forward and backward stepping. The ability to detect the discrete steps of the motor allows the experimenter to measure directly the probabilities of taking a forward or a backward step, $p_\pm$, and the probabilities of observing pairs of steps $p_{\pm\pm}$ as a function of [ATP] (see Figure 3c). If the steps are uncorrelated, as in the model considered here, then the probability of taking two steps in succession should simply be the product of the probability of taking each step, that is, $p_{\pm\pm} = p_\pm p_\pm$. This is clearly not the case for the observed branching probabilities, since p_- is finite, whereas p_{--} is zero for all [ATP]. Thus, without measuring a single dwell time distribution, the experimenter can rule out not only this model but also any kinetic model that does not produce correlated stepping.

Model Pathway 3: Cycle with Reversible Stepping Transition. With the observation that the steps of the molecular motor are correlated, the experimenter now considers a simple modification to the initial model that introduces the possibility of backward steps in a fashion that produces correlated stepping statistics. Here, the kinetic cycle is nearly identical to that in model 1, with the exception that the product release transition is now reversible.

$$E_1 \leftrightarrow T_2 \rightarrow D_3 \xrightarrow{\pm d} E_1 \quad (58)$$
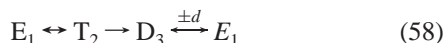
Note, however, that the entire cycle is not reversible due to the irreversible hydrolysis transition, a feature of this model motivated experimentally by the fact that the experimenter never observes two successive backward steps. The corresponding rate matrix is only minimally different from that found in model 1, incorporating a rate, $\bar{k}_3$, and weight, $\rho_-(q) = e^{iqd}$, for backward steps.

$$\mathbf{M}(q) = \begin{pmatrix} -(k_1[T] + \bar{k}_3) & \bar{k}_1 & k_3\rho_+(q) \\ k_1[T] & -(k_2 + \bar{k}_1) & 0 \\ \bar{k}_3\rho_-(q) & k_2 & -k_3 \end{pmatrix} \quad (59)$$

Following the procedure outlined above and in Appendix C, we calculate the position probability density. In contrast to the previous two models, two initial conditions are appropriate here: state E_1, which would follow a forward step, and state D_3, which would follow a backward step. Although the choice of initial conditions does not affect the mean velocity and randomness parameter as demonstrated above, it is relevant for calculating the dwell time distribution. The position probability densities are thus given by

$$\tilde{P}_+(q, s) = \frac{s^2 + \alpha s + \beta(0)}{s^3 + \alpha s^2 + \beta(q)s + \gamma(q)} -$$
$$\frac{\bar{k}_3(1 - \rho_-(q))(s + k_2 + \bar{k}_1)}{s^3 + \alpha s^2 + \beta(q)s + \gamma(q)} \quad (60)$$

$$\tilde{P}_-(q, s) = \frac{s^2 + \alpha s + \beta(0)}{s^3 + \alpha s^2 + \beta(q)s + \gamma(q)} -$$
$$\frac{k_3(1 - \rho_+(q))(s + k_1[T] + k_2 + \bar{k}_1)}{s^3 + \alpha s^2 + \beta s + \gamma(q)} \quad (61)$$

where the subscript $+$ and $-$ denote that the initial condition corresponds to a forward or backward step, respectively. The coefficients α, β, and γ are listed in Table 1, normalized by the product of the forward rates $\sigma = k_1[T]k_2k_3$.

Mean Velocity and Randomness Parameter. The mean velocity and randomness parameter are calculated from the position probability density by the same procedure used in models 1 and 2. Despite the possibility of backward steps in this model, the mean velocity is identical in form to that found in the previous models with k_{cat} and K_M given in Table 2. Surprisingly, the randomness parameter also has the same form as that in model 1, eq 34, provided the definition $r = 2D/vd$ is used. The parameters q_{cat} and Q_M are listed in Table 3.

Dwell Time Distributions. As in model 2, multiple dwell time distributions are necessary to describe the kinetics of this model. However, unlike the case of the previous branched cycle in which the choice of dwell time distribution depends only on the step succeeding the dwell, here, this choice depends on both the preceding and succeeding steps and, thus, involves four dwell time distributions. As a result, it is necessary to use the matrix formalism developed to address systems with correlated stepping statistics (eq 22) to extract these quantities from the position probability density. As described above, the distributions for dwells occurring between two forward steps, $\tilde{\psi}_{++}(s)$, and between a forward and backward step, $\tilde{\psi}_{+-}(s)$, are calculated from the position probability density $\tilde{P}_+(q, s)$ (eq 60), whose initial condition corresponds to an initial forward step. Following the procedure outlined in the text, we evaluate $\tilde{P}_+(q, s)$ in the limit $\rho_-(q) = 0$ and extract the dwell time distributions,

$$p_{++}\tilde{\psi}_{++}(s) = \frac{\gamma_+}{s^3 + \alpha s^2 + (\beta_{+-} + \beta(0))s + (\gamma_+ + \gamma_{+-})} \quad (62)$$

$$p_{+-}\tilde{\psi}_{+-}(s) = \frac{\beta_{+-}s + \gamma_{+-} + s\bar{k}_3(s + k_2 + \bar{k}_1)}{s^3 + \alpha s^2 + (\beta_{+-} + \beta(0))s + (\gamma_+ + \gamma_{+-})} \quad (63)$$

where the coefficients γ_+, γ_{+-}, β_{+-}, and $\beta(0)$ are related to $\gamma(q)$ and $\beta(q)$ by the relations $\gamma(q) = \gamma_+(1 - \rho_+(q)) + \gamma_{+-}(1 - \rho_+(q)\rho_-(q))$ and $\beta(q) = \beta_{+-}(1 - \rho_+(q)\rho_-(q)) + \beta(0)$. The quantities $p_{\pm\pm}$ correspond to the probabilities of following a forward step with a forward or backward step, respectively, and are determined by taking the limit $s = 0$.

$$p_{++} = \frac{\gamma_+}{\gamma_+ + \gamma_{+-}}$$

$$p_{+-} = \frac{\gamma_{+-}}{\gamma_+ + \gamma_{+-}} \qquad (64)$$

From the position probability density $\tilde{P}_-(q, s)$ (eq 61), whose initial condition corresponds to a backward step, the distributions for dwells between a backward and forward step, $\tilde{\psi}_{-+}(s)$, and between two backward steps, $\tilde{\psi}_{--}(s)$, can be determined. Evaluating $\tilde{P}_-(q, s)$ in the limit $\rho_+(q) = 0$,

$$p_{-+}\tilde{\psi}_{-+}(s) =$$
$$\frac{\beta_{+-}s + \gamma_+ + \gamma_{+-} + sk_3(s + k_1[T] + k_2 + \bar{k}_1)}{s^3 + \alpha s^2 + (\beta_{+-} + \beta(0))s + (\gamma_+ + \gamma_{+-})}$$

$$p_{--}\tilde{\psi}_{--}(s) = 0 \qquad (65)$$

and the associated probabilities $p_{-+} = 1$ and $p_{--} = 0$. The latter stems from the fact that the hydrolysis step in the model is irreversible, eliminating the possibility of two consecutive backward steps. As a result, only three dwell time distributions are required to characterize the stepping kinetics in this model.

As in model 2, the connection between the moments of these dwell time distributions and the mean velocity and randomness parameter is more complex. As shown in Appendix E, the mean velocity has the same simple form, $v = \langle d \rangle / \langle t \rangle$, where the average step size is now given by $\langle d \rangle = d(p_+ - p_-)$ and the average dwell time by $\langle t \rangle = p_+(p_{++}\langle t_{++} \rangle + p_{+-}\langle t_{+-} \rangle) + p_-p_{-+}\langle t_{-+} \rangle$. $p_\pm$ is the probability of taking a forward or backward step, independent of the previous step, defined in Appendix E in terms of the conditional probabilities $p_{\pm\pm}$.

Calculation of the randomness parameter $r = 2D/vd$ yields a more complex expression,

$$r = \frac{1}{p_+ - p_-} + \frac{\langle t^2 \rangle}{\langle t \rangle^2}(p_+ - p_-) +$$
$$2\frac{p_+-p_{-+}\langle t_{+-} \rangle \langle t_{-+} \rangle}{(p_{+-} + p_{-+})\langle t \rangle^2}(p_+ - p_-) +$$
$$2\frac{p_+p_{+-}\langle t_{+-} \rangle + p_-(p_{++}\langle t_{++} \rangle + p_-\langle t_{-+} \rangle) - (p_{+-} + p_{-+})\langle t \rangle}{(p_{-+} - p_{+-})\langle t \rangle}$$

$$(66)$$

where the mean square of the dwell time is given by $\langle t^2 \rangle = p_+(p_{++}\langle t_{++}^2 \rangle + p_{+-}\langle t_{+-}^2 \rangle) + p_-p_{-+}\langle t_{-+}^2 \rangle$.

As with models 1 and 2, it is possible to calculate the inverse Laplace transform of the three dwell time distributions eqs 62, 63, and 65 analytically. Substituting the coefficients from Table 1 into these expressions and simplifying, we obtain

$$p_{++}\tilde{\psi}_{++}(s) =$$
$$\frac{k_1[T]k_2k_3}{(s^2 + s(k_1[T] + \bar{k}_1 + k_2 + \bar{k}_3) + k_1[T]k_2 + (k_2 + \bar{k}_1)\bar{k}_3)(s + k_3)}$$

$$p_{+-}\tilde{\psi}_{+-}(s) =$$
$$\frac{(s + k_2 + \bar{k}_1)\bar{k}_3}{s^2 + s(k_1[T] + \bar{k}_1 + k_2 + \bar{k}_3) + k_1[T]k_2 + (k_2 + \bar{k}_1)\bar{k}_3}$$

$$p_{-+}\tilde{\psi}_{-+}(s) = \frac{k_3}{s + k_3} \qquad (67)$$

The corresponding stepping probabilities are given by

$$p_{++} = \frac{k_1[T]k_2}{k_1[T]k_2 + (k_2 + \bar{k}_1)\bar{k}_3}$$

$$p_{+-} = \frac{(k_2 + \bar{k}_1)\bar{k}_3}{k_1[T]k_2 + (k_2 + \bar{k}_1)\bar{k}_3}$$

$$p_{-+} = 1 \qquad (68)$$

In the time domain, the three dwell time distributions take the form

$$p_{++}\psi_{++}(t) = \frac{k_1[T]k_2k_3}{(\lambda_2 - \lambda_1)(k_3 - \lambda_1)}e^{-\lambda_1 t} +$$
$$\frac{k_1[T]k_2k_3}{(\lambda_1 - \lambda_2)(k_3 - \lambda_2)}e^{-\lambda_2 t} + \frac{k_1[T]k_2k_3}{(\lambda_1 - k_3)(\lambda_2 - k_3)}e^{-k_3 t}$$

$$p_{+-}\psi_{+-}(t) = \frac{(k_2 + \bar{k}_1 - \lambda_1)\bar{k}_3}{(\lambda_2 - \lambda_1)}e^{-\lambda_1 t} + \frac{(k_2 + \bar{k}_1 - \lambda_2)\bar{k}_3}{(\lambda_1 - \lambda_2)}e^{-\lambda_2 t}$$

$$p_{-+}\psi_{-+}(t) = k_3 e^{-k_3 t} \qquad (69)$$

with

$$\lambda_{1,2} = \frac{1}{2}(k_1[T] + \bar{k}_1 + k_2 + \bar{k}_3) \pm$$
$$\frac{1}{2}\sqrt{(k_1[T] + \bar{k}_1 + k_2 + \bar{k}_3)^2 - 4(k_1[T]k_2 + \bar{k}_3(\bar{k}_1 + k_2))} \quad (70)$$

Comparison to Data. As seen from these calculations, although very different in the stepping behavior from model 1, this model produces exactly the same velocity and randomness parameter, illustrating the limitations of these kinetic measures. However, with the ability to observe steps of the motor directly, the experimenter can now calculate histograms of the dwell times. Since the system clearly displays correlated stepping behaviors from the branching probabilities, it is necessary to keep track of the preceding and succeeding step when collecting dwell times. These distributions are plotted in Figure 3d−f for steps at three different ATP concentrations. If the quality of these distributions is high enough, then the experimenter can in principle fit them to determine how many exponentials sum to create the observed dwell distribution. This number is the exact number of kinetic states in the system. Moreover, by investigating how these distributions change with [ATP], the experimenter can learn something about the connectivity of these states. For example, the observed dwell time distributions $\psi_{++}(t)$ (Figure 3d), $\psi_{+-}(t)$ (Figure 3e), and $\psi_{-+}(t)$ (Figure 3f) are well-described by the sums of three, two, and one exponential, respectively. The fact that $\psi_{-+}(t)$ is well-described by a single exponential implies that the only kinetic transition that occurs during the dwell preceded by a backward step and followed by a forward step is the stepping transition itself. In addition, $\psi_{-+}(t)$ is independent of [ATP], which implies that the stepping transition cannot be the binding of ATP. These two facts confirm the position of the irreversible transition and that of the stepping transition in the kinetic cycle considered here. Furthermore, the fact that $\psi_{+-}(t)$ is described by two additional exponentials with unique decay rates and that $\psi_{++}(t)$ is described by all three of the observed exponential decay rates confirms that the number of states in the underlying kinetic model is three. Given this

Kinetic Models of Macromolecular Dynamics

J. Phys. Chem. B, Vol. 112, No. 19, 2008 **6037**

data, the experimenter can conclude that model 3 is an excellent model for the underlying behavior of the molecular motor.

It is in this fashion that the dwell time distributions provide the experimenter with the most information on the behavior of a molecular motor. However, it should be noted that fitting these distributions will require a large numbers of steps to accurately measure the distribution and distinguish exponential decays that have similar decay rates. Furthermore, kinetic transitions that are faster than the time resolution of the measurement will not be present in the measured distributions. All of these situations, however, are inherent limitations to any measure of the fluctuations of a kinetic system.

Conclusions

In this article, we have extended a matrix methodology first proposed by Koza[33] that is based on the master equation to solve the kinetics of a molecular motor undergoing cycles of arbitrary complexity. By formulating the problem in Fourier–Laplace space, we have derived a simple matrix representation for the system and have shown how it can be used to solve for the position probability density $\tilde{P}(q, s)$. From this quantity, we can derive not only all the kinetic parameters typically associated with bulk and single-molecule experiments, such as the velocity v, but also quantities uniquely derivable from single-molecule experiments, such as the randomness parameter r, the branching probabilities, and the dwell time distributions $\tilde{\psi}(s)$. Furthermore, through this formalism, we can connect the microscopic rates of these cycles to the macroscopic parameters, k_{cat} and K_M, derivable from traditional bulk or single-molecule measurements and propose analogous parameters for the randomness parameter, q_{cat} and Q_M, which should allow classification of kinetic schemes on the basis of their fluctuation characteristics. Although we have explicitly treated the case of a motor for illustrative purposes, it is important to note that one need not associate the discrete step d in our formalism with a necessarily physical motion. This step can just as easily represent product formation or a conformational change at the completion of a chemical reaction that may or may not be coupled to a mechanical task. Thus, the methodology described here can be extended to a large variety of kinetic processes.

Several alternative analytical methods for treating complex kinetics schemes have been developed and described previously.[15,32,33,36,37] One approach, using a methodology first developed by Derrida for treating one-dimensional diffusion in random media,[32] has been used to calculate general expressions for the velocity and diffusion constant for a wide range of kinetic schemes (systems with parallel paths,[42,43] arbitrary off-pathway branches, or deaths[35]) and has been modified to include states with arbitrary lifetimes;[44] however, this technique has not been reported to yield dwell time distributions. Another type of analytical method calculates the dwell time distribution directly using a recursive algorithm and derives the velocity and randomness parameter from its moments.[36] However, as we have shown here, a wide class of kinetic schemes—systems with multiple parallel kinetic pathways with different step sizes (model 2) or, more generally, systems in which the dwell time statistics depend on both the preceding and succeeding step (model 3)—cannot be described by a single dwell time distribution, a situation not addressed with this method. Recently, a third method for calculating dwell time distributions in a reversible system, such as the one considered above, was developed in which the dwell time distributions were derived by finding the eigenvalues of the rate matrix for a reduced master equation with absorbing boundary conditions.[40] This method is similar to the one proposed here; however, because it works in the time domain, it requires the calculation of the eigenvalues of the rate matrix in order to calculate analytical expressions for the dwell time distributions, which may not be possible for systems with large numbers of internal kinetic states.

Our methodology, in contrast, can accommodate all these types of complex kinetic cycles and yield, in a straightforward fashion, analytical expressions for all the kinetic parameters typically accessible through bulk and single-molecule experiments. Because the methodology introduced here is based on matrix algebra, it is particularly amenable to implementation with computer algebra programs. As we have shown, in cycles in which the system steps irreversibly, it is only necessary to calculate the characteristic polynomial of the rate matrix $\mathbf{M}$ to determine the position probability density and dwell time distribution, and only its three lowest-order coefficients are necessary to calculate the velocity and randomness parameter, simple operations that can be carried out analytically by many computer programs. The latter point is significant because the amount of algebra involved in calculating these kinetic parameters increases dramatically with the number of kinetic states. Moreover, because the calculation of the position probability density requires at most the calculation of a matrix inverse, as opposed to the n eigenvalues of the system, this method can be easily extended to kinetic systems and, thus, matrices of arbitrary size. Finally, by working in Laplace space, our method provides exact, closed-form solutions for the Laplace transform of the dwell time distributions for a much wider class of kinetic schemes than previously reported, and although the calculation of the inverse Laplace transform requires solution of essentially the same eigenvalue problem as a direct solution in the time domain, this may not be necessary, since the Laplace transform of the dwell time distributions can yield exact expressions for all of the moments of that distribution.

As single-molecule techniques continue to unravel the details of molecular motion, it has become clear that many systems display kinetic behaviors that cannot be captured by simple kinetic schemes, and the need arises for the development of theoretical tools to calculate the properties of very general kinetic models. We have shown that our methodology can be applied to a large class of kinetic systems, including branched kinetic schemes that take multiple types of steps. Recent technical innovations[3,45] have improved the resolution of single-molecule techniques to the point that in the future, it will be possible to observe the steps of nearly all molecular motors and determine dwell time distributions directly. The promise of precision fits to dwell time distributions, with the ability to determine the intricate details of the kinetic cycle of the system, is the exciting future of single-molecule measurements.

Appendix A: Derivation of Velocity and Randomness Parameter

The general form of the Fourier–Laplace transform of the position probability density, eq 9, allows us to derive general relations for the velocity and diffusion coefficient as defined by eq 7. Applying eq 8 to eq 9, we calculate the average position in Laplace space,

$$\langle \tilde{x}(s) \rangle = i \left. \frac{\partial \tilde{P}(q, s)}{\partial q} \right|_{q=0} = \frac{i\dot{C}(0, s)}{|\mathbf{R}(0, s)|} - \frac{i|\dot{\mathbf{R}}(0, s)|}{s|\mathbf{R}(0, s)|} \quad \text{(A1)}$$

6038 *J. Phys. Chem. B, Vol. 112, No. 19, 2008*

Chemla et al.

and the average square of the position in Laplace space,

$$\langle \tilde{x}^2(s) \rangle = -\frac{\partial^2 \tilde{P}(q, s)}{\partial q^2}\bigg|_{q=0} = -\frac{\ddot{C}(0, s)}{|\mathbf{R}(0, s)|} +$$
$$\frac{2\dot{C}(0, s)|\dot{\mathbf{R}}(0, s)|}{|\mathbf{R}(0, s)|^2} + \frac{|\ddot{\mathbf{R}}(0, s)|}{s|\mathbf{R}(0, s)|} - 2\frac{|\dot{\mathbf{R}}(0, s)|^2}{s|\mathbf{R}(0, s)|^2} \quad \text{(A2)}$$

where we have used the property that $C(0, s) = |\mathbf{R}(0, s)|/s$, as discussed in the text.

Calculating v and D requires taking the inverse Laplace transform of these expressions using the Bromwich integral, which is calculated by evaluating the residues at the poles of each term. For instance,

$$\langle x(t) \rangle = \frac{1}{2\pi i}\int_{-i\infty}^{i\infty} \mathrm{d}s\, e^{st}\langle \tilde{x}(s) \rangle = \sum_j \mathrm{Res}(e^{st}\langle \tilde{x}(s) \rangle; -p_j) =$$
$$\sum_j \lim_{s \to -p_j} \frac{1}{(m_j - 1)!}\frac{\mathrm{d}^{m_j-1}}{\mathrm{d}s^{m_j-1}}((s + p_j)^{m_j}\, e^{st}\langle \tilde{x}(s) \rangle) \quad \text{(A3)}$$

for the mean position $\langle \tilde{x}(s) \rangle$, where $-p_j$ are its poles of order m_j. In general, this procedure is quite difficult because of the complex dependence of both $C(q, s)$ and $|\mathbf{R}(q, s)|$ on s. However, as stated in eq 7, we are only concerned with the long time limit, $t \to \infty$, where most terms vanish as $e^{-p_j t} \to 0$, and only the pole at $s = 0$ contributes to the sum. Thus, we can write for a general function $f(t)$,

$$\lim_{t \to \infty} f(t) = \mathrm{Res}(e^{st}\tilde{f}(s); s = 0) =$$
$$\lim_{s \to 0} \frac{1}{(m - 1)!}\frac{\mathrm{d}^{m-1}}{\mathrm{d}s^{m-1}}(s^m e^{st}\tilde{f}(s)) \quad \text{(A4)}$$

where m is the order of a pole at $s = 0$, and a sum over all such poles is implicit.

This observation makes the evaluation of the inverse Laplace transform much more tractable. Writing out the terms of the determinant $|\mathbf{R}(q, s)|$ in the mean position explicitly,

$$\langle \tilde{x}(s) \rangle = \frac{i\dot{C}(0, s)}{s(\cdots + \alpha(0)s + \beta(0))} - \frac{i(\cdots + \dot{\beta}(0)s + \dot{\gamma}(0))}{s^2(\cdots + \alpha(0)s + \beta(0))} \quad \text{(A5)}$$

it is evident that $\langle \tilde{x}(s) \rangle$ exhibits a pole of order 1 and a pole of order 2 at $s = 0$. Applying eq A4 to each term, it follows that

$$\lim_{t \to \infty} \langle x(t) \rangle = \frac{i\dot{C}(0, 0)}{\beta(0)} - \frac{i\dot{\beta}(0)}{\beta(0)} - \frac{i\dot{\gamma}(0)\, t}{\beta(0)} + \frac{i\alpha(0)\,\dot{\gamma}(0)}{\beta(0)^2} \quad \text{(A6)}$$

and the velocity is simply obtained from the time derivative,

$$v = \lim_{t \to \infty} \frac{\partial}{\partial t}\langle x(t) \rangle = -\frac{i\dot{\gamma}(0)}{\beta(0)} \quad \text{(A7)}$$

which is eq 11 in the text. Note that the time derivative could also have been applied directly to eq A4 before evaluating the residue, leading to the following relation:

$$\lim_{t \to \infty} \frac{\partial}{\partial t} f(t) = \mathrm{Res}(s e^{st}\tilde{f}(s); s = 0) =$$
$$\lim_{s \to 0} \frac{1}{(m - 1)!}\frac{\mathrm{d}^{m-1}}{\mathrm{d}s^{m-1}}(s^{m+1} e^{st}\tilde{f}(s)) \quad \text{(A8)}$$

It is easy to show that this equation also reproduces eq 11 for the velocity.

We will use this latter relation to evaluate the time derivative of the inverse Laplace transform of $\langle \tilde{x}(s)^2 \rangle$. As above, we expand the determinant $|\mathbf{R}(q, s)|$ to lowest orders in s

$$\langle \tilde{x}^2(s) \rangle = \frac{-\ddot{C}(0, s)}{s(\cdots + \alpha(0)s + \beta(0))} +$$
$$\frac{2\dot{C}(0, s)(\cdots + \dot{\beta}(0)s + \dot{\gamma}(0))}{s^2(\cdots + \alpha(0)s + \beta(0))^2} + \frac{(\cdots + \ddot{\beta}(0)s + \ddot{\gamma}(0))}{s^2(\cdots + \alpha(0)s + \beta(0))} -$$
$$\frac{2(\cdots + \dot{\beta}(0)s + \dot{\gamma}(0))^2}{s^3(\cdots + \alpha(0)s + \beta(0))^2} \quad \text{(A9)}$$

Applying eq A8, the poles of first order do not contribute to the time derivative, and

$$\lim_{t \to \infty} \frac{\partial}{\partial t}\langle x(t)^2 \rangle = \frac{2\dot{C}(0, 0)\,\dot{\gamma}(0)}{\beta(0)^2} + \frac{\ddot{\gamma}(0)}{\beta(0)} - \frac{4\dot{\gamma}(0)\,\dot{\beta}(0)}{\beta(0)^2} +$$
$$\frac{4\alpha(0)\,\dot{\gamma}(0)^2}{\beta(0)^3} - \frac{2\dot{\gamma}(0)^2 t}{\beta(0)^2} \quad \text{(A10)}$$

The diffusion coefficient can be calculated by making the following algebraic manipulation to eq 7:

$$2D = \lim_{t \to \infty}\left(\frac{\partial}{\partial t}\langle x(t)^2 \rangle - 2\langle x(t) \rangle\frac{\partial}{\partial t}\langle x(t) \rangle\right) \quad \text{(A11)}$$

Substituting eqs A6, A7, and A10, it follows that

$$2D = \frac{\ddot{\gamma}(0)}{\beta(0)} - \frac{2\dot{\gamma}(0)\,\dot{\beta}(0)}{\beta(0)^2} + \frac{2\alpha(0)\,\dot{\gamma}(0)^2}{\beta(0)^3} \quad \text{(A12)}$$

which is equivalent to eq 12 listed in the text.

Appendix B: Calculation of $|\mathbf{R}(q, s)|$ for an n-State Linear Cycle

Let d_j be the determinant of a $j \times j$ submatrix of $\mathbf{R}$ starting at the upper left element. The first such determinant is simply $d_1 = s + k_1 + \bar{k}_n$. The second is related to d_1 by the relation $d_2 = (s + k_2 + \bar{k}_1)d_1 - k_1\bar{k}_1$. In general, each determinant d_j follows from the previous two determinants in the series, d_{j-1} and d_{j-2}, by the following recursion relation

$$d_j = (s + k_j + \bar{k}_{j-1})d_{j-1} - k_{j-1}\bar{k}_{j-1}d_{j-2} \quad \text{(B1)}$$

where $j = 1, 2, ..., n$. Note that the nth determinant in the series, d_n, is not equal to $|\mathbf{R}(q, s)|$ because the elements at the upper right and lower left of the matrix $\mathbf{R}$ break its tridiagonal symmetry. To account for these elements, we define a second determinant, d_j^*, starting one element below and one to the right of the upper left corner of the matrix; that is, $d_1^* = s + k_2 + \bar{k}_1$, $d_2^* = (s + k_3 + \bar{k}_2)d_1^* - k_2\bar{k}_2$, etc. Its recursion relations are given by

$$d_j^* = (s + k_{j+1} + \bar{k}_j)d_{j-1}^* - k_j\bar{k}_j d_{j-2}^* \quad \text{(B2)}$$

for $j = 1, 2, ..., n - 1$. The full $n \times n$ determinant, $|\mathbf{R}(q, s)|$, is thus given by

$$|\mathbf{R}(q, s)| = d_n - d_{n-2}^* k_n\bar{k}_n\rho_+(q)\rho_-(q) - k_1 k_2 \cdots k_n\rho_+(q) - \bar{k}_1\bar{k}_2 \cdots \bar{k}_n\rho_-(q) \quad \text{(B3)}$$

Closer inspection of eqs B1−B3 reveals some important properties of the determinants: d_j and d_j^* are jth order polynomials in s and are independent of q. As a result, the q

Kinetic Models of Macromolecular Dynamics

J. Phys. Chem. B, Vol. 112, No. 19, 2008 **6039**

dependence of $|\mathbf{R}(q, s)|$ results explicitly from the last three terms containing $\rho_+(q)$ and $\rho_-(q)$ in eq B3. Since $\rho_\pm(q = 0) = 1$, it follows that

$$|\mathbf{R}(q, s)| = |\mathbf{R}(0, s)| + d^*_{n-2}k_n\bar{k}_n(1 - \rho_+(q)\rho_-(q)) +$$
$$k_1 k_2 \cdots k_n(1 - \rho_+(q)) + \bar{k}_1\bar{k}_2 \cdots \bar{k}_n(1 - \rho_-(q)) \quad \text{(B4)}$$

It is further convenient to introduce the matrix $\mathbf{R}'(s)$, equal to $\mathbf{R}(q, s)$ in which both $\rho_+(q)$ and $\rho_-(q)$ are set to zero. From eqs B3 and B4, it follows that

$$|\mathbf{R}'(s)| = d_n = |\mathbf{R}(0, s)| + d^*_{n-2}k_n\bar{k}_n + k_1 k_2 \cdots k_n +$$
$$\bar{k}_1\bar{k}_2 \cdots \bar{k}_n \quad \text{(B5)}$$

These matrices will be useful in a later calculation.

A convenient way of rewriting eqs B1−B3 is to normalize the determinants d_j by $\sigma_j = k_1 k_2 \cdots k_j$ and d^*_j by $\sigma^*_j = k_2 k_3 \cdots k_{j+1}$, leading to the modified recursion relations

$$\tilde{d}_j = \tilde{d}_{j-1}(1 + K_j + s/k_j) - \tilde{d}_{j-2}K_j$$
$$\tilde{d}^*_j = \tilde{d}^*_{j-1}(1 + K_{j+1} + s/k_{j+1}) - \tilde{d}^*_{j-2}K_{j+1} \quad \text{(B6)}$$

and for the full determinant,

$$|\tilde{\mathbf{R}}(q, s)| = \tilde{d}_n - \tilde{d}^*_{n-2}K_1\rho_+(q)\,\rho_-(q) - \rho_+(q) -$$
$$K_1 K_2 \cdots K_n\rho_-(q) \quad \text{(B7)}$$

where $\tilde{d}_j$ and $\tilde{d}^*_j$ are the scaled determinants and $K_j \equiv \bar{k}_{j-1}/k_j$ (with $K_1 \equiv \bar{k}_n/k_1$). The recursion relations eqs B4 and B5 provide a prescription for calculating the characteristic polynomial. However, to calculate v and D, we only need to determine the three lowest-order terms in the characteristic polynomial $|\mathbf{R}(q, s)| = \cdots \alpha s^2 + \beta s + \gamma$. These three coefficients are calculated by similarly expanding the submatrix determinants to lowest orders in s:

$$\tilde{d}_j = \sum_{l=0}^{j} a^{(l)}_j s^l = a^{(j)}_j s^j \cdots + a^{(2)}_j s^2 + a^{(1)}_j s + a^{(0)}_j$$

$$\tilde{d}^*_j = \sum_{l=0}^{j} a^{*(l)}_j s^l \quad \text{(B8)}$$

Substituting into the recursion relations and collecting like orders, we obtain the following relations for $a^{(l)}_j$,

$$\Delta a^{(l)}_j = K_j \Delta a^{(l)}_{j-1} + a^{(l-1)}_{j-1}/k_j \quad \text{(B9)}$$

and the following for $a^{*(l)}_j$,

$$\Delta a^{*(l)}_j = K_{j+1} \Delta a^{*(l)}_{j-1} + a^{*(l-1)}_{j-1}/k_{j+1} \quad \text{(B10)}$$

where $\Delta a^{(l)}_j \equiv a^{(l)}_j - a^{(l)}_{j-1}$ and similarly for the other coefficients and where it is understood that $a^{(l)}_j$ and $a^{*(l)}_j$ are zero for $l < 0$. Using the determinant $\tilde{d}_1 = s/k_1 + 1 + K_1$ and $\tilde{d}^*_1 = s/k_2 + 1 + K_2$ to provide the initial values $a^{(0)}_1 = 1 + K_1$ and $a^{*(0)}_1 = 1 + K_2$ (we also define $a^{(0)}_0 = a^{*(0)}_0 = 1$), the solutions to the recursion relations eqs B9 and B10 for $l = 0$ are calculated,

$$a^{(0)}_j = \Gamma(1, j)$$
$$a^{*(0)}_j = \Gamma(2, j + 1) \quad \text{(B11)}$$

where we have introducing the function

$$\Gamma(n_1, n_2) = 1 + \sum_{m=n_1}^{n_2} \prod_{l=n_1}^{m} K_l \quad \text{(B12)}$$

Similarly, repeated use of the recursion relations eqs B9 and B10 yields the coefficients for all orders $l > 0$. The calculation is simplified by the fact that $a^{(l)}_j = a^{*(l)}_j = 0$ for $l > j$. It thus follows that $\Delta a^{(l)}_l = a^{(l-1)}_{l-1}/k_l$ and $\Delta a^{*(l)}_l = a^{*(l-1)}_{l-1}/k_{l+1}$, terminating the recursion relations. Carrying out the algebra, we find that

$$a^{(l)}_j = \sum_{i=l}^{j} \frac{a^{(l-1)}_{i-1}}{k_i}\Gamma(i + 1, j)$$

$$a^{*(l)}_j = \sum_{i=l}^{j} \frac{a^{*(l-1)}_{i-1}}{k_{i+1}}\Gamma(i + 2, j + 1) \quad \text{(B13)}$$

for orders $l > 0$.

Equation B13 is another recursion relation that connects each coefficient to those with lower orders in l. By repeated application of this equation, we can derive exact solutions for each of these coefficients in terms of sums of products of Γs. We find that

$$a^{(l)}_j = \sum_{i_1=l}^{j} \sum_{i_2=1}^{i_1-1} \cdots \sum_{i_l=1}^{i_{l-1}-1} \frac{1}{k_{i_1}} \cdots \frac{1}{k_{i_l}}\Gamma(1, i_l - 1) \times$$
$$\Gamma(i_l + 1, i_{l-1} - 1) \cdots \Gamma(i_2 + 1, i_1 - 1)\,\Gamma(i_1 + 1, j) \quad \text{(B14)}$$

and

$$a^{*(l)}_j = \sum_{i_1=l}^{j} \sum_{i_2=1}^{i_1-1} \cdots \sum_{i_l=1}^{i_{l-1}-1} \frac{1}{k_{i_1+1}} \cdots \frac{1}{k_{i_l+1}}\Gamma(2, i_l) \times$$
$$\Gamma(i_l + 2, i_{l-1}) \cdots \Gamma(i_2 + 2, i_1)\,\Gamma(i_1 + 2, j + 1) \quad \text{(B15)}$$

which hold for all coefficients of order $l > 0$. Note that these expressions determine the coefficients of the determinant $|\mathbf{R}'(s)|$, since $|\mathbf{R}'(s)| = d_n = (a^{(n)}_n s^n \cdots + a^{(1)}_n s + a^{(0)}_n)\sigma_n$. Moreover, all of these coefficients are independent of q.

These recursion relations can be used to solve for the coefficients of the full determinant $|\mathbf{R}(q, s)|$, where $|\tilde{\mathbf{R}}(q, s)| = \sum_{l=0}^{n} a^{(l)}(q)\, s^l$, using eqs B7 and B8 to calculate the coefficients

$$a^{(0)}(q) =$$
$$a^{(0)}_n - a^{*(0)}_{n-2}K_1\rho_+(q)\,\rho_-(q) - \rho_+(q) - K_1 K_2 \cdots K_n\rho_-(q)$$

$$a^{(l)}(q) = a^{(l)}_n - a^{*(l)}_{n-2}K_1\rho_+(q)\,\rho_-(q) \quad \text{(B16)}$$

Here, the q dependence is explicit in $\rho_+(q)$ and $\rho_-(q)$, and the second expression holds for all $l > 0$. Equation B16 can thus be used to determine every coefficient of the determinant $|\mathbf{R}(q, s)|$. For our purposes here, we consider explicitly the three coefficients required to calculate the mean velocity and randomness parameter. Substituting eqs B11, B14, and B15 and

6040 *J. Phys. Chem. B, Vol. 112, No. 19, 2008*

Chemla et al.

simplifying, we obtain the following general expressions for $\alpha = a^{(2)}\sigma_n$, $\beta = a^{(1)}\sigma_n$, and $\gamma = a^{(0)}\sigma_n$:

$$\gamma(q) = \gamma_+(1 - \rho_+(q)) + \gamma_-(1 - \rho_-(q)) +$$
$$\gamma_{+-}(1 - \rho_+(q)\,\rho_-(q))$$

$$\beta(q) = \beta(0) + \beta_{+-}(1 - \rho_+(q)\,\rho_-(q))$$

$$\alpha(q) = \alpha(0) + \alpha_{+-}(1 - \rho_+(q)\,\rho_-(q)) \qquad \text{(B17)}$$

where these coefficients can be expressed compactly as

$$\tilde{\gamma}_+ = \sigma_n \quad \tilde{\gamma}_- = \sigma_n \prod_{l=1}^{n} K_l$$

$$\gamma_{+-} = \sigma_n K_1 \, \Gamma(2, n-1)$$

$$\tilde{\beta}(0) = \sigma_n \sum_{i=1}^{n} \frac{1}{k_i} \Gamma(i+1, n-1+i)$$

$$\beta_{+-} = \sigma_n K_1 \sum_{i=1}^{n-2} \frac{1}{k_{i+1}} \Gamma(2, i)\,\Gamma(i+2, n-1)$$

$$\tilde{\alpha}(0) = \sigma_n \sum_{i=2}^{n} \sum_{j=1}^{i-1} \frac{1}{k_i}\frac{1}{k_j}\Gamma(j+1, i-1)\,\Gamma(i+1, n+j-1)$$

$$\tilde{\alpha}_{+-} = \sigma_n K_1 \sum_{i=2}^{n-2} \sum_{j=1}^{i-1} \frac{1}{k_{i+1}}\frac{1}{k_{j+1}}\Gamma(2, j)\,\Gamma(j+2, i) \times$$
$$\Gamma(i+2, n-1) \quad \text{(B18)}$$

Note that eqs B17 and B18 correctly demonstrate that $\gamma(q = 0) = 0$ and that the determinant $|\mathbf{R}(0, s)| = \cdots \alpha(0)s^2 + \beta(0)s$. For the irreversible cycle considered in the text where $\bar{k}_n = 0$, note that the terms proportional to $1 - \rho_+(q)\rho_-(q)$ in eq B17 vanish since $K_1 = 0$. Furthermore, if the step size is uniform such that $\rho_\pm(q) = e^{\mp iqd}$, a case also treated in the text, these terms also vanish. Using eqs B17 and B18, the mean velocity and diffusion constant can be calculated for an n-state reversible linear cycle, yielding eqs 11 and 12.

Appendix C: Calculation of $\tilde{\psi}_{\pm\pm}(s)$ for an n-State Linear Cycle

We start by considering the Fourier−Laplace transform of the individual components of the master equation eq 4,

$$s\tilde{P}_i(q, s) - P_i(0) = \sum_{j=1}^{n} M_{ij}(q)\,\tilde{P}_j(q, s) \qquad \text{(C1)}$$

and sum over the kinetic state index i. It is apparent from the structure of the matrix M_{ij} for a linear cycle, eq 5, that all the columns sum to zero *except* for the first and the last. Thus,

$$s\tilde{P}(q, s) - 1 = -\bar{k}_n(1 - \rho_-(q))\,\tilde{P}_1(q, s) -$$
$$k_n(1 - \rho_+(q))\,\tilde{P}_n(q, s) \quad \text{(C2)}$$

Using the fact that $\tilde{P}_i(q, s) = \sum_{j=1}^{n} \text{Co}_{ji}(q, s)P_j(0)/|\mathbf{R}(0, s)|$ from eqs 6 and 9, we find that

$$s\tilde{P}(q, s) - 1 = -\frac{\bar{k}_n(1 - \rho_-(q))}{|\mathbf{R}(q, s)|} \sum_{j=1}^{n} \text{Co}_{j1}\,P_j(0) -$$
$$\frac{k_n(1 - \rho_+(q))}{|\mathbf{R}(q, s)|} \sum_{j=1}^{n} \text{Co}_{jn}\,P_j(0) \quad \text{(C3)}$$

Equation C3 is a convenient way to express $\tilde{P}(q, s)$, since for a suitable choice of initial conditions in which only one kinetic state is occupied at $t = 0$, only two matrix cofactors need to be determined (as opposed to all n^2 for the full inverse matrix).

To calculate the position probability density explicitly and then the dwell time distributions, it is necessary to select the appropriate initial conditions. If the initial step is forward, the proper choice is that the kinetic state 1 is occupied at $t = 0$, $P_i(0) = \delta_{i1}$. Thus, eq C3 simplifies to

$$s\tilde{P}_+(q, s) - 1 =$$
$$-\frac{\bar{k}_n(1 - \rho_-(q))\,\text{Co}_{11} + k_n(1 - \rho_+(q))\,\text{Co}_{1n}}{|\mathbf{R}(q, s)|} \quad \text{(C4)}$$

where $\tilde{P}_+(q, s)$ is the position probability density calculated with this initial condition. These cofactors are related to determinants that have already been calculated in Appendix B: $\text{Co}_{11} = d^*_{n-1}$ and $\text{Co}_{1n} = \bar{k}_n\rho_-(q)d^*_{n-2} + k_1k_2 \cdots k_{n-1}$. Substituting these expressions, eq C4 becomes

$$s\tilde{P}_+(q, s) = \{|\mathbf{R}(q, s)| - \bar{k}_n(d^*_{n-1} - k_nd^*_{n-2})(1 - \rho_-(q)) -$$
$$\bar{k}_nk_nd^*_{n-2}(1 - \rho_+(q)\rho_-(q)) - k_1k_2 \cdots k_n(1 - \rho_+(q))\}/|\mathbf{R}(q, s)|$$
$$\text{(C5)}$$

In the limit that the stepping transition is irreversible ($\bar{k}_n = 0$), the numerator simplifies, $|\mathbf{R}(q, s)| - k_1k_2 \cdots k_n(1 - \rho_+(q)) = |\mathbf{R}(0, s)|$ by eq B4, and eq C5 reduces to eq 17.

From this expression, we can calculate the dwell time distributions $p_{++}\tilde{\psi}_{++}(s)$ and $p_{+-}\tilde{\psi}_{+-}(s)$ following the procedure outlined in the text. Applying eq 25 yields

$$\frac{1}{s\tilde{P}_+(q, s)}\bigg|_{\rho_-=0} = \frac{d_n - k_1k_2 \cdots k_n\rho_+(q)}{d_n - k_1k_2 \cdots k_n - \bar{k}_nd^*_{n-1}} \quad \text{(C6)}$$

where we have used the fact that $|\mathbf{R}(q, s)||_{\rho_-=0} = d_n - k_1k_2 \cdots k_n\rho_+(q)$ from eq B7 to greatly simplify eq C6.

Collecting terms dependent and independent of q allows the coefficients a_0 and a_+ from eq 25 to be determined and the dwell time distribution $p_{++}\tilde{\psi}_{++}(s) = -a_+/a_0$ to be evaluated,

$$p_{++}\tilde{\psi}_{++}(s) = \frac{k_1k_2 \cdots k_n}{d_n} = \frac{\gamma_+}{|\mathbf{R}'(s)|} \quad \text{(C7)}$$

where we have used eqs B5 and B18 to further simplify this expression. From eq 26, we can calculate the dwell time distribution $p_{+-}\tilde{\psi}_{+-}(s) = (a_0 + a_+ - 1)/a_0$:

$$p_{+-}\tilde{\psi}_{+-}(s) =$$
$$\frac{d_n - k_1k_2 \cdots k_n - d_n + k_1k_2 \cdots k_n + \bar{k}_nd^*_{n-1}}{d_n} =$$
$$\frac{\bar{k}_nd^*_{n-1}}{|\mathbf{R}'(s)|} \quad \text{(C8)}$$

We now choose initial conditions such that the last kinetic state is occupied at $t = 0$, $P_i(0) = \delta_{in}$, and eq C3 simplifies to

Kinetic Models of Macromolecular Dynamics

J. Phys. Chem. B, Vol. 112, No. 19, 2008 **6041**

$$s\tilde{P}_-(q, s) - 1 = -\frac{\bar{k}_n(1 - \rho_-(q))\,\mathrm{Co}_{n1} + k_n(1 - \rho_+(q))\,\mathrm{Co}_{nn}}{|\mathbf{R}(q, s)|} \tag{C9}$$

where $\tilde{P}_-(q, s)$ is the position probability density calculated with an initial backward step.

These cofactors are also related to determinants derived in Appendix B: $\mathrm{Co}_{nn} = d_{n-1}$ and $\mathrm{Co}_{n1} = k_n\rho_+(q)d^*_{n-2} + \bar{k}_1\bar{k}_2 \cdots \bar{k}_{n-1}$. Substituting these expressions,

$$s\tilde{P}_-(q, s) = \{|\mathbf{R}(q, s)| - k_n(d_{n-1} - \bar{k}_n d^*_{n-2})(1 - \rho_+(q)) - \bar{k}_n k_n d^*_{n-2}(1 - \rho_+(q)\rho_-(q)) - \bar{k}_1\bar{k}_2 \cdots \bar{k}_n(1 - \rho_-(q))\}/|\mathbf{R}(q, s)| \tag{C10}$$

Using the fact that $|\mathbf{R}(q, s)||_{\rho_+=0} = d_n - \bar{k}_1\bar{k}_2 \cdots \bar{k}_n\rho_-(q)$, this simplifies to

$$\left.\frac{1}{s\tilde{P}_-(q, s)}\right|_{\rho_+=0} = \frac{d_n - \bar{k}_1\bar{k}_2 \cdots \bar{k}_n\rho_-(q)}{d_n - k_n d_{n-1} - \bar{k}_1\bar{k}_2 \cdots \bar{k}_n} \tag{C11}$$

From this expression, we can calculate the dwell time distributions $p_{--}\tilde{\psi}_{--}(s)$ and $p_{-+}\tilde{\psi}_{-+}(s)$ by applying eqs 25 and 26, from which it follows that

$$p_{--}\tilde{\psi}_{--}(s) = \frac{\gamma_-}{|\mathbf{R}'(s)|}$$

$$p_{-+}\tilde{\psi}_{-+}(s) = \frac{k_n d_{n-1}}{|\mathbf{R}'(s)|} \tag{C12}$$

where we have again simplified these expressions using eqs B5 and B18.

Appendix D: General Relation between $\tilde{\psi}(s)$, $\rho(q)$, and $\tilde{P}(q, s)$

Consider a kinetic system in which there is the possibility of taking m different types of steps, each of which is preceded by a distinct kinetic pathway. We allow the dwell time for each step to depend not only on the type of step that succeeds the dwell but also the type of step that precedes it, as well. Thus, there are m^2 distinct dwell time distributions, which in Laplace space we denote by $\tilde{\psi}_{ij}(s)$, where i is the index of the type of step preceding the dwell and j is the type of step following it. There will also be m different step-size distributions, denoted in Fourier space by $\rho_i(q)$.

The probability to be at a position, x, at a given time, t, is the sum of the probability to arrive at x by taking exactly N steps in the duration, t, for all N. In Fourier–Laplace space, this can be written as

$$\tilde{P}(q, s) = \sum_{N=0}^{\infty} \tilde{P}(q, s, N) \tag{D1}$$

where $\tilde{P}(q, s, N)$ is the Fourier–Laplace transform of the probability of arriving at position x in a duration t by exactly N steps. We will determine the general form for $\tilde{P}(q, s, N)$ by extrapolating a pattern from explicit values for low N.

Consider first the case that the motor has not taken a step in time t; that is, $N = 0$. Physically, the motor starts in the kinetic state corresponding to a preceding step of type i with some initial probability p_{0i}, chooses the path corresponding to a step of type j with some probability p_{ij}, and then does not complete it. The final probability is the sum over all initial conditions and possible choices for the different kinetic paths,

$$\tilde{P}(q, s, 0) = \sum_{i,j=1}^{m} p_{0i}p_{ij}\frac{1 - \tilde{\psi}_{ij}(s)}{s} \tag{D2}$$

where we have used the fact that if the dwell time distribution is given by $\tilde{\psi}(s)$, the probability of not taking a step within a given time t is given by $(1 - \tilde{\psi}(s))/s$. For the $N = 1$ case, the motor starts in its initial state i with probability p_{0i}; chooses a path out of that state j with probability p_{ij}; completes this path generating a step of type j, the size and duration of which are governed in Fourier–Laplace space by $\rho_j(q)\psi_{ij}(s)$; chooses another path k with a probability p_{jk}; and then does not complete this path, with probability $(1 - \tilde{\psi}_{jk}(s))/s$. Again, the total probability is the sum over all possible combinations of paths.

$$\tilde{P}(q, s, 1) = \sum_{i,j,k=1}^{m} p_{0i}p_{ij}\,\tilde{\psi}_{ij}(s)\rho_j(q)\,p_{jk}\frac{1 - \tilde{\psi}_{jk}(s)}{s} \tag{D3}$$

The $N = 2$ case adds one more completed step; thus, the probability is

$$\tilde{P}(q, s, 2) = \sum_{i,j,k,l=1}^{m} p_{0i}(p_{ij}\tilde{\psi}_{ij}(s)\rho_j(q)) \times (p_{jk}\tilde{\psi}_{jk}(s)\rho_k(q))p_{kl}\frac{1 - \tilde{\psi}_{kl}(s)}{s} \tag{D4}$$

Equations D2 – D4 can be simplified greatly by defining two matrices, $\tilde{\psi}(s)$ and $\rho(q)$, whose ijth elements equal $p_{ij}\tilde{\psi}_{ij}(s)$ and $\rho_i(q)\delta_{ij}$, respectively. In addition, we define the vectors $\mathbf{p}_0$ and $\tilde{\mathbf{\Psi}}(s)$, whose ith elements equal the initial probability p_{0i} and the sum $\sum_{j=1}^{m} p_{ij}(1 - \tilde{\psi}_{ij}(s))/s$. Because the sum of the probability of taking all steps j given the previous step i must be 1, the latter simplifies to $\tilde{\Psi}_i(q, s) = (1 - \sum_{j=1}^{m} p_{ij}\tilde{\psi}_{ij}(s))/s$.

With these definitions, eq D2 – D4 can be written compactly as

$$\tilde{P}(q, s, 0) = \mathbf{p}_0^{\mathrm{T}}(\tilde{\psi}(s)\rho(q))^0\,\tilde{\mathbf{\Psi}}(s)$$

$$\tilde{P}(q, s, 1) = \mathbf{p}_0^{\mathrm{T}}(\tilde{\psi}(s)\rho(q))^1\,\tilde{\mathbf{\Psi}}(s)$$

$$\tilde{P}(q, s, 2) = \mathbf{p}_0^{\mathrm{T}}(\tilde{\psi}(s)\rho(q))^2\,\tilde{\mathbf{\Psi}}(s) \tag{D5}$$

Thus, for all N,

$$\tilde{P}(q, s, N) = \mathbf{p}_0^{\mathrm{T}}(\tilde{\psi}(s)\rho(q))^N\,\tilde{\mathbf{\Psi}}(s) \tag{D6}$$

and the position probability density is

$$\tilde{P}(q, s) = \sum_{N=0}^{\infty} \mathbf{p}_0^{\mathrm{T}}(\tilde{\psi}(s)\rho(q))^N\,\tilde{\mathbf{\Psi}}(s) \tag{D7}$$

Evaluating the infinite sum produces the final relation, eq 22 in the text,

$$\tilde{P}(q, s) = \mathbf{p}_0^{\mathrm{T}}(\mathbf{I} - \tilde{\psi}(s)\rho(q))^{-1}\,\tilde{\mathbf{\Psi}}(s) \tag{D8}$$

where $\mathbf{I}$ is the identity matrix.

For a system in which there is no memory of the previous step, all distributions and probabilities have only one index, that is, $p_{ij} \rightarrow p_j$ and $\psi_{ij} \rightarrow \psi_j$, reflecting the fact that the choice of a given step and the preceding dwell time do not depend on the previous step. The above relation for the position probability density can be simplified for such a system by noting that the

6042 *J. Phys. Chem. B, Vol. 112, No. 19, 2008*

Chemla et al.

elements of each column of the matrix product $\tilde{\psi}(s)\rho(q)$ are now identical; that is, the elements of the jth column are $p_j\tilde{\psi}_j(s)\rho_j(q)$. Since the elements of $\mathbf{p}_0^T$ must sum to 1, the product $\mathbf{p}_0^T\tilde{\psi}(s)\rho(q)$ is thus a vector with elements $p_j\tilde{\psi}_j(s)\rho_j(q)$. Moreover, the vector $\boldsymbol{\Psi}(s)$ consists of identical elements $(1 - \sum_{j=1}^m p_j\tilde{\psi}_j(s))/s$, which can be compactly written as $(1 - \mathrm{Tr}(\tilde{\psi}(s)))/s$, where $\mathrm{Tr}(\tilde{\psi}(s))$ is the trace of $\tilde{\psi}(s)$. It thus follows that

$$\mathbf{p}_0^T(\tilde{\psi}(s)\rho(q))\boldsymbol{\Psi}(s) = \mathrm{Tr}(\tilde{\psi}(s)\rho(q))\frac{1}{s}(1 - \mathrm{Tr}(\tilde{\psi}(s))) \qquad \text{(D9)}$$

It can be shown by direct matrix multiplication that powers of the matrix product $\tilde{\psi}(s)\rho(q)$ are given by the relation $(\tilde{\psi}(s)\rho(q))^N = (\mathrm{Tr}(\tilde{\psi}(s)\rho(q)))^{N-1}\tilde{\psi}(s)\rho(q)$. Thus,

$$\tilde{P}(q, s) = \sum_{N=0}^{\infty} \mathbf{p}_0^T(\tilde{\psi}(s)\rho(q))^N \boldsymbol{\Psi}(s) =$$

$$\sum_{N=0}^{\infty} (\mathrm{Tr}(\tilde{\psi}(s)\rho(q)))^N \frac{1}{s}(1 - \mathrm{Tr}(\tilde{\psi}(s))) \quad \text{(D10)}$$

which, summing over N, simplifies to

$$\tilde{P}(q, s) = \frac{1}{s}\frac{1 - \mathrm{Tr}(\tilde{\psi}(s))}{1 - \mathrm{Tr}(\tilde{\psi}(s)\rho(q))} \qquad \text{(D11)}$$

which is eq 23 in the main text.

Appendix E: Connecting the Moments of $\psi(t)$ and $\rho(x)$ to v and r

Equations 16, 22, and 23 in the text connect the Fourier−Laplace transform of the position probability density, $\tilde{P}(q, s)$, to the Laplace transform of the dwell time distributions, $\tilde{\psi}(s)$, and the Fourier transform of the step size distributions, $\rho(q)$. These expressions allow us to express the moments of $\tilde{P}(q, s)$ (more specifically, the related velocity v and randomness parameter r) in terms of the moments of the dwell time distributions and step-size distributions. We adopt here the definition $r = 2D/v\langle d\rangle$ so that the final expressions are completely general. The expressions derived below can be related to the different definition used in the text by multiplying by $\langle d\rangle/d$.

We first consider a system that displays correlated stepping statistics in which Laplace transforms of the dwell time distributions, $\tilde{\psi}_{\pm\pm}(s)$, and the Fourier transforms of the step size distributions, $\rho_{\pm}(q)$, are related to the position probability density, $\tilde{P}(q, s)$, via eq 22. We consider a system with only two types of steps, labeled with $+$ or $-$, but this calculation can be extended to systems with m step distributions. The mean velocity v is determined from the average position in Laplace space, obtained by calculating the derivative of eq 22 at $q = 0$:

$$\langle \tilde{x}(s)\rangle = i\left.\frac{\partial \tilde{P}(q, s)}{\partial q}\right|_{q=0} \qquad \text{(E1)}$$

With no loss in generality, we can assume that the initial step is a forward step, and evaluating the derivative in eq E1, we find

$$\langle \tilde{x}(s)\rangle =$$
$$\{(p_{++}\tilde{\psi}_{++} + p_{+-}\tilde{\psi}_{+-}p_{-+}\tilde{\psi}_{-+} - p_{++}\tilde{\psi}_{++}p_{--}\tilde{\psi}_{--}) \times$$
$$\langle d_+\rangle + p_{+-}\tilde{\psi}_{+-}\langle d_-\rangle\}/\{s(1 - p_{++}\tilde{\psi}_{++} - p_{--}\tilde{\psi}_{--} +$$
$$p_{++}\tilde{\psi}_{++}p_{--}\tilde{\psi}_{--} - p_{+-}\tilde{\psi}_{+-}p_{-+}\tilde{\psi}_{-+})\} \quad \text{(E2)}$$

Expanding the Laplace transform of the dwell time distributions around $s = 0$ (as in Appendix A) and relating the derivatives to the moments of these distributions yields

$$\tilde{\psi}_{\pm\pm}(s) = \tilde{\psi}_{\pm\pm}(0) + s\left.\frac{d\tilde{\psi}_{\pm\pm}}{ds}\right|_{s=0} + \frac{1}{2}s^2\left.\frac{d^2\tilde{\psi}_{\pm\pm}}{ds^2}\right|_{s=0} + \cdots =$$
$$1 - s\langle t_{\pm\pm}\rangle + \frac{1}{2}s^2\langle t_{\pm\pm}^2\rangle + \cdots \quad \text{(E3)}$$

Inserting these expressions into eq E2 and expanding in powers of $1/s$ yields

$$\langle \tilde{x}(s)\rangle = \frac{\langle d\rangle}{s^2\langle t\rangle} - \langle d\rangle\frac{p_{++}p_{--}\langle t_{++}\rangle\langle t_{--}\rangle - p_{+-}p_{-+}\langle t_{+-}\rangle\langle t_{-+}\rangle}{s\langle t\rangle^2(p_{+-} + p_{-+})} +$$
$$\frac{\langle d\rangle\langle t^2\rangle}{2s\langle t\rangle^2} - \frac{p_{+-}\langle t_{+-}\rangle\langle d_-\rangle - p_{--}\langle t_{--}\rangle\langle d_+\rangle}{s\langle t\rangle(p_{+-} + p_{-+})} -$$
$$\frac{\langle d_+\rangle}{s} + O(s^0) \quad \text{(E4)}$$

where we have defined composite step-size and dwell-time moments via $\langle d^n\rangle = p_+\langle d_+^n\rangle + p_-\langle d_-^n\rangle$ and $\langle t^n\rangle = p_+\langle t_+^n\rangle + p_-\langle t_-^n\rangle$, where $p_\pm$ is the probability of taking a given step, and $\langle t_\pm^n\rangle$ is the nth moment of the distribution of all dwells preceding a forward or backward step. These expressions can be related to the individual branching probabilities and dwell time moments via $p_\pm = p_{\mp\pm}/(p_{+-} + p_{-+})$ and $\langle t_\pm^n\rangle = p_{\pm\pm}\langle t_{\pm\pm}^n\rangle + p_{\pm\mp}\langle t_{\pm\mp}^n\rangle$, expressions derived by Tsygankov et al.[39] Only terms with negative powers of s are needed, since they are the only terms that will survive the inverse Laplace transform.

Evaluating the inverse transform and limit of eq E4 using eq A4,

$$\lim_{t\to\infty}\langle x(t)\rangle = \frac{\langle d\rangle}{\langle t\rangle}t - \langle d\rangle\frac{p_{++}p_{--}\langle t_{++}\rangle\langle t_{--}\rangle - p_{+-}p_{-+}\langle t_{+-}\rangle\langle t_{-+}\rangle}{\langle t\rangle^2(p_{+-} + p_{-+})} +$$
$$\frac{\langle d\rangle\langle t^2\rangle}{2\langle t\rangle^2} - \frac{p_{+-}\langle t_{+-}\rangle\langle d_-\rangle - p_{--}\langle t_{--}\rangle\langle d_+\rangle}{\langle t\rangle(p_{+-} + p_{-+})} - \langle d_+\rangle \quad \text{(E5)}$$

Taking the time derivative required by eq 7 yields the expected relation between the velocity and the composite moments,

$$v = \frac{\langle d\rangle}{\langle t\rangle} \qquad \text{(E6)}$$

To derive the diffusion constant, D, and thus the randomness parameter, r, we need the second moment of the position, which is derived via

$$\langle \tilde{x}^2(s)\rangle = -\left.\frac{\partial^2 \tilde{P}(q, s)}{\partial q^2}\right|_{q=0} \qquad \text{(E7)}$$

Applying this to the position probability density, eq 22; expanding the dwell time distributions as above; and expanding around $s = 0$ yields

Kinetic Models of Macromolecular Dynamics

J. Phys. Chem. B, Vol. 112, No. 19, 2008 **6043**

$$\langle \tilde{x}(s)^2 \rangle = \frac{2\langle d \rangle^2}{s^3 \langle t \rangle^2} + \frac{\langle d^2 \rangle - 2p_+\langle d_+ \rangle^2 - 2p_-\langle d_- \rangle^2 - 2\langle d_+ \rangle\langle d \rangle}{s^2 \langle t \rangle} +$$

$$\frac{2\langle d \rangle^2 \langle t^2 \rangle}{s^2 \langle t \rangle^3} + 2(\langle d \rangle + p_+\langle d_+ \rangle)\frac{p_{--}\langle t_{--} \rangle\langle d_+ \rangle - p_{+-}\langle t_{+-} \rangle\langle d_- \rangle}{s^2 \langle t \rangle^2 (p_{+-} + p_{-+})} +$$

$$2p_-\langle d_- \rangle\frac{p_{++}\langle t_{++} \rangle\langle d_- \rangle - p_{-+}\langle t_{-+} \rangle\langle d_+ \rangle}{s^2 \langle t \rangle^2 (p_{+-} + p_{-+})} -$$

$$4\langle d \rangle^2 \frac{p_{++}p_{--}\langle t_{++} \rangle\langle t_{--} \rangle - p_{+-}p_{-+}\langle t_{+-} \rangle\langle t_{-+} \rangle}{s^2 \langle t \rangle^3 (p_{-+} + p_{+-})} + O(s^{-1}) \quad \text{(E8)}$$

Transforming to the time domain and taking the appropriate time derivative and limit,

$$\lim_{t \to \infty} \frac{\partial}{\partial t}\langle x(t)^2 \rangle = \frac{\langle d^2 \rangle - 2p_+\langle d_+ \rangle^2 - 2p_-\langle d_- \rangle^2 - 2\langle d_+ \rangle\langle d \rangle}{\langle t \rangle} +$$

$$\frac{2\langle d \rangle^2}{\langle t \rangle^2}t + \frac{2\langle d \rangle^2 \langle t^2 \rangle}{\langle t \rangle^3} +$$

$$2(\langle d \rangle + p_+\langle d_+ \rangle)\frac{p_{--}\langle t_{--} \rangle\langle d_+ \rangle - p_{+-}\langle t_{+-} \rangle\langle d_- \rangle}{\langle t \rangle^2 (p_{+-} + p_{-+})} +$$

$$2p_-\langle d_- \rangle\frac{p_{++}\langle t_{++} \rangle\langle d_- \rangle - p_{-+}\langle t_{-+} \rangle\langle d_+ \rangle}{\langle t \rangle^2 (p_{+-} + p_{-+})} -$$

$$4\langle d \rangle^2 \frac{p_{++}p_{--}\langle t_{++} \rangle\langle t_{--} \rangle - p_{+-}p_{-+}\langle t_{+-} \rangle\langle t_{-+} \rangle}{\langle t \rangle^3 (p_{-+} + p_{+-})} \quad \text{(E9)}$$

Equation E9 can now be used with eqs 7 and E5 to solve for the diffusion constant and for the randomness parameter, $r = 2D/v\langle d \rangle$, yielding

$$r = \frac{\langle d^2 \rangle - \langle d \rangle^2}{\langle d \rangle^2} + \frac{\langle t^2 \rangle - \langle t \rangle^2}{\langle t \rangle^2} + 2\frac{\langle d \rangle^2 - p_+\langle d_+ \rangle^2 - p_-\langle d_- \rangle^2}{\langle d \rangle^2} +$$

$$2p_+\langle d_+ \rangle\frac{p_{--}\langle t_{--} \rangle\langle d_+ \rangle - p_{+-}\langle t_{+-} \rangle\langle d_- \rangle}{(p_{+-} + p_{-+})\langle d \rangle^2 \langle t \rangle} +$$

$$2p_-\langle d_- \rangle\frac{p_{++}\langle t_{++} \rangle\langle d_- \rangle - p_{-+}\langle t_{-+} \rangle\langle d_+ \rangle}{(p_{+-} + p_{-+})\langle d \rangle^2 \langle t \rangle} -$$

$$2\frac{p_{++}p_{--}\langle t_{++} \rangle\langle t_{--} \rangle - p_{+-}p_{-+}\langle t_{+-} \rangle\langle t_{-+} \rangle}{(p_{+-} + p_{-+})\langle t \rangle^2} \quad \text{(E10)}$$

As detailed above, although the particular choice of initial conditions affects the specific values of $\langle \tilde{x}(s) \rangle$ and $\langle \tilde{x}(s)^2 \rangle$, it does not affect either the mean velocity (eq E6) or randomness parameter (eq E10). This can be verified directly by repeating this calculation with a backward step as the initial condition. The randomness parameter for model 3 can be derived from eq E10 with $p_{--} = 0$, $\langle d_\pm \rangle = \pm d$, and $\langle d_\pm^2 \rangle = d^2$ and by accounting for the different definition used, $r = 2D/vd$.

We can now derive the same relations for a system with uncorrelated stepping pathways by dropping the first index of every term in eqs E6 and E10. With some algebra, we find that the velocity obeys eq E6 and that the randomness reduces to

$$r = \frac{\langle d^2 \rangle - \langle d \rangle^2}{\langle d \rangle^2} + \frac{\langle t^2 \rangle - \langle t \rangle^2}{\langle t \rangle^2} +$$

$$2\frac{\langle d \rangle\langle t \rangle - p_+\langle d_+ \rangle\langle t_+ \rangle - p_-\langle d_- \rangle\langle t_- \rangle}{\langle d \rangle\langle t \rangle} \quad \text{(E11)}$$

where it can be verified that the above expressions for composite moments of the step size and dwell times reduce to the expected $\langle d^n \rangle = p_+\langle d_+^n \rangle + p_-\langle d_-^n \rangle$ and $\langle t^n \rangle = p_+\langle t_+^n \rangle + p_-\langle t_-^n \rangle$. This expression reduces to the randomness parameter derived for model 2, eq 49 with $\langle d_\pm \rangle = \pm d$ and $\langle d_\pm^2 \rangle = d^2$ and when scaled by $\langle d \rangle/d$ to account for the difference in definition.

Finally, we can derive the velocity and randomness for a system with only one stepping pathway by setting one of the branching probabilities to zero. We find that the velocity continues to obey eq E6 as expected and that the randomness parameter becomes

$$r = \frac{\langle d^2 \rangle - \langle d \rangle^2}{\langle d \rangle^2} + \frac{\langle t^2 \rangle - \langle t \rangle^2}{\langle t \rangle^2} \quad \text{(E12)}$$

a result derived previously.[36] This expression produces the randomness parameter for model 1 with $\langle d \rangle = d$ and $\langle d^2 \rangle = d^2$, in which case the first term in eq E12 is zero.

References and Notes

(1) Ritort, F. Single-molecule experiments in biological physics: methods and applications. *J. Phys.: Condens. Matter* **2006**, *18*, R531−R583.

(2) Cornish, P. V.; Ha, T. A survey of single-molecule techniques in chemical biology. *ACS Chem. Biol.* **2007**, *2*, 53−61.

(3) 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.

(4) Segel, I. H. *Enzyme Kinetics*; John Wiley & Sons, Inc.: New York, 1975.

(5) Bustamante, C.; Chemla, Y. R.; Forde, N. R.; Izhaky, D. Mechanical processes in biochemistry. *Annu. Rev. Biochem.* **2004**, *73*, 705−748.

(6) Schnitzer, M. J.; Block, S. M. Kinesin hydrolyses one ATP per 8-nm step. *Nature* **1997**, *388*, 386−390.

(7) Fisher, M. E.; Kolomeisky, A. B. Simple mechanochemistry describes the dynamics of kinesin molecules. *Proc. Natl. Acad. Sci.* **2001**, *98*, 7748−7753.

(8) Mehta, A. D.; Rock, R. S.; Rief, M.; Spudich, J. A.; Mooseker, M. S.; Cheney, R. E. Myosin-V is a processive actin-based motor. *Nature* **1999**, *400*, 590.

(9) Kolomeisky, A. B.; Fisher, M. E. A Simple Kinetic Model Describes the Processivity of Myosin-V. *Biophys. J.* **2003**, *84*, 1642−1650.

(10) Abbondanzieri, E. A.; Greenleaf, W. J.; Shaevitz, J. W.; Landick, R.; Block, S. M. Direct observation of base-pair stepping by RNA polymerase. *Nature* **2005**, *438*, 460−465.

(11) Galburt, E. A.; Grill, S. W.; Wiedmann, A.; Lubkowska, L.; Choy, J.; Nogales, E.; Kashlev, M.; Bustamante, C. Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner. *Nature* **2007**, *446*, 820−823.

(12) Chemla, Y. R.; Aathavan, K.; Michaelis, J.; Grimes, S.; Jardine, P. J.; Anderson, D. L.; Bustamante, C. Mechanism of force generation of a viral DNA packaging motor. *Cell* **2005**, *122*, 683−692.

(13) Smith, D. E.; Tans, S. J.; Smith, S. B.; Grimes, S.; Anderson, D. L.; Bustamante, C. The bacteriophage straight phi29 portal motor can package DNA against a large internal force. *Nature* **2001**, *413*, 748−752.

(14) Svoboda, K.; Mitra, P. P.; Block, S. M. Fluctuation analysis of Motor Protein Movement and Single Enzyme Kinetics. *Proc. Natl. Acad. Sci.* **1994**, *91*, 11782−11786.

(15) Schnitzer, M. J.; Block, S. M. Statistical kinetics of processive enzymes. *Cold Spring Harbor Symp. Quant. Biol.* **1995**, *60*, 793−802.

(16) Charvin, G.; Bensimon, D.; Croquette, V. On the Relation Between Noise Spectra and the Distribution of Time Between Steps for Single Molecular Motors. *Single Mol.* **2002**, *3*, 43−48.

(17) Valentine, M. T.; Fordyce, P. M.; Krzysiak, T. C.; Gilbert, S. P.; Block, S. M. Individual dimers of the mitotic kinesin motor Eg5 step processively and support substantial loads in vitro. *Nat. Cell. Biol.* **2006**, *8*, 470−476.

(18) Saleh, O. A.; Perals, C.; Barre, F. X.; Allemand, J. F. Fast, DNA-sequence independent translocation by FtsK in a single-molecule experiment. *Embo J.* **2004**, *23*, 2430−2439.

(19) Neuman, K. C.; Saleh, O. A.; Lionnet, T.; Lia, G.; Allemand, J. F.; Bensimon, D.; Croquette, V. Statistical determination of the step size of molecular motors. *J. Phys.: Condens. Matter* **2005**, *17*, S3811.

(20) Rief, M.; Rock, R. S.; Mehta, A. D.; Mooseker, M. S.; Cheney, R. E.; Spudich, J. A. Myosin-V stepping kinetics: a molecular model for processivity. *Proc. Natl. Acad. Sci.* U.S.A. **2000**, *97*, 9482–9486.

(21) Yildiz, A.; Forkey, J. N.; McKinney, S. A.; Ha, T.; Goldman, Y. E.; Selvin, P. R. Myosin V walks hand-over-hand: Single fluorophore imaging with 1.5-nm localization. *Science* **2003**, *300*, 2061–2065.

(22) Asbury, C. L.; Fehr, A. N.; Block, S. M. Kinesin moves by an asymmetric hand-over-hand mechanism. *Science* **2003**, *302*, 2130–2134.

(23) Yildiz, A.; Tomishige, M.; Vale, R. D.; Selvin, P. R. Kinesin walks hand-over-hand. *Science* **2004**, *303*, 676–678.

(24) Yasuda, R.; Noji, H.; Kinosita, K.; Yoshida, M. F-1-ATPase is a highly efficient molecular motor that rotates with discrete 120 degrees steps. *Cell* **1998**, *93*, 1117–1124.

(25) English, B. P.; Min, W.; van Oijen, A. M.; Lee, K. T.; Luo, G. B.; Sun, H. Y.; Cherayil, B. J.; Kou, S. C.; Xie, S. N. Ever-fluctuating single enzyme molecules: Michaelis-Menten equation revisited (vol 2, pg 87, 2006). *Nat. Chem. Biol.* **2006**, *2*, 168–168.

(26) Neuman, K. C.; Abbondanzieri, E. A.; Landick, R.; Gelles, J.; Block, S. M. Ubiquitous Transcriptional Pausing Is Independent of RNA Polymerase Backtracking. *Cell* **2003**, *115*, 437.

(27) Carter, N. J.; Cross, R. A. Mechanics of the kinesin step. *Nature* **2005**, *435*, 308.

(28) Mallik, R.; Carter, B. C.; Lex, S. A.; King, S. J.; Gross, S. P. Cytoplasmic dynein functions as a gear in response to load. *Nature* **2004**, *427*, 649–652.

(29) Reck-Peterson, S. L.; Yildiz, A.; Carter, A. P.; Gennerich, A.; Zhang, N.; Vale, R. D. Single-molecule analysis of dynein processivity and stepping behavior. *Cell* **2006**, *126*, 335–348.

(30) Min, W.; Gopich, I. V.; English, B. P.; Kou, S. C.; Xie, X. S.; Szabo, A. When does the Michaelis–Menten equation hold for fluctuating enzymes? *J. Phys. Chem. B* **2006**, *110*, 20093–20097.

(31) Kou, S. C.; Cherayil, B. J.; Min, W.; English, B. P.; Xie, X. S. Single-molecule Michaelis–Menten equations. *J. Phys. Chem. B* **2005**, *109*, 19068–19081.

(32) Derrida, B. Velocity and Diffusion Constant of a Periodic One-Dimensional Hopping Model. *J. Stat. Phys.* **1983**, *31*, (3), 433–450.

(33) Koza, Z. General technique of calculating the drift velocity and diffusion coefficient in arbitrary periodic systems. *J. Phys. A: Math. Gen.* **1999**, *32*, 7637–7651.

(34) Koza, Z. Diffusion coefficient and drift velocity in periodic media. *Physica A* **2000**, *285*, 176–186.

(35) Kolomeisky, A. B.; Fisher, M. E. Periodic sequential kinetic models with jumping, branching and deaths. *Physica A* **2000**, *279*, 1–20.

(36) Shaevitz, J. W.; Block, S. M.; Schnitzer, M. J. Statistical kinetics of macromolecular dynamics. *Biophys. J.* **2005**, *89*, 2277–2285.

(37) Kolomeisky, A. B.; Fisher, M. E. Molecular Motors: A Theorist's Perspective. *Annu. Rev. Phys. Chem.* **2007**, *58*, 675–695.

(38) Liao, J.-C.; Spudich, J. A.; Parker, D.; Delp, S. L. Extending the absorbing boundary method to fit dwell-time distributions of molecular motors with complex kinetic pathways. *Proc. Natl. Acad. Sci.* **2007**, *104*, 3171–3176.

(39) Tsygankov, D.; Linden, M.; Fisher, M. E. Back-stepping, hidden substeps, and conditional dwell times in molecular motors. *Phys. Rev. E: Stat. Nonlinear Soft Matter Phys.* **2007**, *75*, 021909.

(40) Linden, M.; Wallin, M. Dwell Time Symmetry in Random Walks and Molecular Motors. *Biophys. J.* **2007**, *92*, 3804–3816.

(41) Weiss, G. H. *Aspects and Applications of the Random Walk.* Elsevier Science B. V.: North-Holland, 1994.

(42) Kolomeisky, A. B. Exact results for parallel-chain kinetic models of biological transport. *J. Chem. Phys.* **2001**, *115*, 7253–7259.

(43) Stukalin, E. B.; Kolomeisky, A. B. Transport of single molecules along the periodic parallel lattices with coupling. *J. Chem. Phys.* **2006**, *124*, 204901.

(44) Kolomeisky, A. B.; Fisher, M. E. Extended kinetic models with waiting-time distributions: Exact results. *J. Chem. Phys.* **2000**, *113*, 10867.

(45) Moffitt, J. R.; Chemla, Y. R.; Smith, S. B.; Bustamante, C. Recent Advances in Optical Tweezers. *Annu. Rev. Biochem.* **2008**, *77*, 19.1–19.24.