Statistical Kinetics of Processive Enzymes

M.J. SCHNITZER[1,2] AND S.M. BLOCK[2,3]
[1]*Department of Physics,* [2]*Department of Molecular Biology and* [3]*Princeton Materials Institute, Princeton University, Princeton, New Jersey 08544*

Recent advances have enabled studies of individual biomolecules at work. One breakthrough has been the development of optical traps, which can be used to produce motility assays in vitro that function down to the level of single motor proteins (Block et al. 1990). When combined with ultrasensitive position sensors, based on either quadrant photodetectors (Finer et al. 1994; Molloy et al. 1995) or interferometers (Svoboda et al. 1994), optical traps have been used to reveal the nanometer-sized steps taken by single biological motors, such as kinesin and myosin, as well as to measure the piconewton-sized forces that these mechanoenzymes generate. Another breakthrough has been the development of low-background, total internal reflection fluorescence microscopy, which has led to the visualization of single ATP hydrolytic cycles by myosin (Funatsu et al. 1995). Video-enhanced, differential interference contrast microscopy has been used to study DNA transcription by single molecules of RNA polymerase, leading to time-resolved studies in vitro (Schafer et al. 1991; Yin et al. 1994). Using similar methods, the kinetics of loop formation and breakdown in individual DNA molecules have also been studied (Finzi and Gelles 1995). The atomic force microscope, a hybrid opto-mechanical device with nanometer spatial resolution, has led to the detection of signals that may well reflect individual enzymatic cycles in lysozyme (Radmacher et al. 1994), and to the visualization of individual protein complexes, such as RNA polymerase-DNA (Guthold et al. 1994).

Collectively, these and other experiments represent substantial progress in the study and control of biological processes. With the advent of single-molecule studies comes the possibility of performing a new type of kinetic study—one based on molecular statistics. Traditionally, biochemists have confined their attention to the average rates at which reactions occur, rather than to the fluctuations about the average. This focus was appropriate, because in studies with bulk quantities of reagents, variations in signals arise chiefly from instrument or other noise. However, in single-molecule studies, in which individual events dominate the behavior, the measured signals often reflect the underlying stochastic nature of the biochemical reaction itself.

In fact, fluctuation analysis has a well-established tradition in neuroscience and has long been used in work on transmembrane channels (Colquhoun 1971; Colquhoun and Hawkes 1977). However, analogous studies of fluctuations during the movement of processive enzymes in vitro, such as kinesin or polymerase, require a somewhat different analytical approach from that developed for single-channel work, due to certain technical differences (such as statistical non-stationarity). In this paper, we present theoretical tools for analyzing processive enzyme behavior and then use our results to compute expected fluctuations in a proposed reaction scheme for the microtubule-based motor, kinesin. Experimental studies incorporating fluctuation analysis of biological motors has only just begun, with recent studies including measurements of stochasticity in kinesin (Svoboda et al. 1994) and the bacterial rotary motor (Samuel and Berg 1995). We anticipate that the concepts developed in this paper, which review and extend the work of Svoboda et al. (1994), will help to form a theoretical basis for future studies of fluctuations in processive reactions.

THEORY

A processive enzyme may be thought of as any protein, or protein complex, that works in a cyclic fashion for a significant number of iterations before completely releasing its substrate (Stryer 1995). The mean number of such iterations, in fact, provides a quantitative measure of this *processivity* (Kornberg and Baker 1992; Gilbert et al. 1995). In most cases, processive enzymes carry out some motor (or motor-like) function, advancing relative to the substrate during each enzymatic cycle. Examples of processive protein complexes include polymerases and ribosomes, which proceed along nucleic acids (Kornberg and Baker 1992), as well as certain ATP-dependent mechanoenzymes, e.g., kinesin, which moves along microtubules (Howard et al. 1989; Block et al. 1990; Gilbert et al. 1995).

Any enzyme takes a variable time, τ, to complete a single enzymatic cycle. In this section, we address the following question: What can one learn about the kinetic behavior of such an enzyme by studying the statistical distribution of its cycle durations, $P(\tau)$, or equivalently, from the moments of this distribution? We first address the problem in broad outline, showing how moments can be determined, and then consider two specific examples of cycles.

The complete distribution $P(\tau)$ contains useful information about how a processive enzyme works: The challenge is to extract this information from real experimental records. At a minimum, once one determines

the average time to complete a cycle, $\langle\tau\rangle$, where angle brackets denote the expectation value, one can compute the catalytic rate of the enzyme, since trivially, $k_{cat}=1/\langle\tau\rangle$. But what, for example, can expressions such as $\langle\tau^2\rangle$ reveal about the reaction sequence? Using $\langle\tau^2\rangle$, we can form the dimensionless ratio

$$r=\left(\frac{\langle\tau^2\rangle-\langle\tau\rangle^2}{\langle\tau\rangle^2}\right) \tag{1}$$

where r, the randomness parameter, is a measure of the temporal regularity of the enzyme (Svoboda et al. 1994). If r is small, the spread in completion times is small compared to a typical completion time itself, and so the enzyme behaves in a highly regular, clock-like fashion. If the value of r is near unity, the completion cycle is highly irregular (an exponential Poisson process achieves $r = 1$, see below). The value of r therefore provides a convenient measure of the regularity of an enzyme's catalytic action. Typically, an enzyme undergoes transitions through a series of intermediate states during each cycle. Determination of the randomness parameter can shed light on the kinetics of these transitions.

A generalized reaction sequence for the cycle of a processive enzyme that passes through exactly M intermediate states is

$$1 \underset{k_{-1}}{\overset{k_1}{\leftrightarrow}} 2 \underset{k_{-2}}{\overset{k_2}{\leftrightarrow}} \cdots \underset{k_{1-M}}{\overset{k_{M-1}}{\leftrightarrow}} M \overset{k_M}{\rightarrow} 1 \tag{2}$$

where the final arrow simply indicates the return of the enzyme to its initial state, completing the cycle. For many enzymes, some of the backward reactions in the pathway may occur at sufficiently low rates as to be negligible. For a given enzyme, we seek information about the reaction pathway. How many intermediates exist, and what are the rate-limiting transitions in the forward and backward directions? For the trivial case where $M = 1$, the enzymatic cycle reduces to a Poisson process, for which the standard deviation equals the mean, and so r is unity. For the general case, such questions are more difficult to address, requiring detailed biochemical kinetic studies to determine each step in the pathway. In certain single-molecule experiments, however, it becomes possible to deduce the moments of $P(\tau)$. What can one learn about the kinetic pathway under such single-molecule conditions?

When individual molecules have been studied, it has been difficult, in most cases, to construct $P(\tau)$ directly, because it has not always been feasible to detect each and every completion of the enzymatic cycle. The difficulties are technical, since many events are lost in the noise, due to either their brevity or signal-to-noise limitations (Svoboda et al. 1993; Finer et al. 1994; Radmacher et al. 1994). Fortunately, perfect detection is not required to deduce the moments of $P(\tau)$. To determine these moments, it suffices to monitor the total number of cycles completed by a single enzyme in a given time period, $N(t)$. Since this quantity is the accumulated result of multiple turnovers, it is easier to study than the times of individual cycles, particularly for processive enzymes, because turnovers sum to produce net displacement. As shown below, one can derive the moments of the distribution $P(\tau)$ from measurements of $N(t)$. In addition, the moments of $P(\tau)$, computed from $N(t)$, constitute a statistically robust measurement, in the sense that they are relatively immune to common sources of thermal and instrumentation noise.

In the following discussion, we often need to consider a sequence of stochastic events. The time required to perform the entire sequence is the sum of times required to perform the individual events, and we further assume that the individual times are governed by independent probability distributions. For example, the time to complete a reaction cycle is the sum of the times required to progress from each intermediate state to the next. In such cases, we can make use of the well-known result that the probability distribution for such a sum is given by the convolution of the separate probability distributions for all the individual times (Feller 1968, 1971). This convolution theorem is easily demonstrated, as follows.

If one considers two sequential events, A and B, completion of both of these events within a time t requires that event A be completed at a time t' less than t, and that event B be completed in the remaining interval $t - t'$. So if $P_A(t)$, $P_B(t)$, and $P_{A+B}(t)$ are the probability densities for completion of A, B, and the pair $A + B$, respectively, then $P_{A+B}(t)$ is given by the joint probability that A finishes at t' and B finishes $t - t'$ later, at time t, integrated over all possible times $t' < t$:

$$P_{A+B}(t)=\int_0^t P_A(t')P_B(t-t')dt' \tag{3}$$

Equation 3 is the convolution of the two densities, P_A and P_B, as asserted. Similarly, if one wants to compute the probability that A but *not* B has occurred within time t, then one simply replaces $P_B(t - t')$ with the probability $1 - \int_0^{t-t'} P_B(t'')dt''$ that B does not occur within the final $t - t'$. Multiple sequential events simply produce multiple convolutions, as can easily be seen by jointly considering all but the final event to be A, and the final event to be B.

Because convolutions of probability densities involve integration, it is generally more convenient to work with their Laplace transforms, $\tilde{P}(s) = \int_0^\infty e^{-st}P(t)dt$. By so doing, convolutions become multiplications in the Laplace transform domain (Carrier et al. 1983), hence $\tilde{P}_{A+B}(s) = \tilde{P}_A(s)\cdot\tilde{P}_B(s)$. At the end of a calculation, one takes the inverse Laplace transform to recover the answer in terms of the original variables. In this fashion, it is easier to compute moments of $P(\tau)$ and $N(t)$ if we consider $\tilde{P}(s)$ and $\tilde{N}(s)$, their respective

Laplace transforms. We will first find $\langle \tilde{N}(s) \rangle$ and $\langle \tilde{N}^2(s) \rangle$ in terms of $\tilde{P}(s)$. To do so, we require $P(N, t)$, the probability that after time t, exactly N completions have occurred, and $\tilde{P}(N, s)$, its Laplace transform.

Completing exactly N cycles in a time interval t may be considered to comprise $N + 1$ events: The enzyme makes N completions *before* t, and then makes *no more* completions in the remaining time. It follows from the above discussion that

$$\tilde{P}(N,s) = \tilde{P}^N(s)\left(\frac{1-\tilde{P}(s)}{s}\right) \tag{4}$$

where the factor inside parentheses is the Laplace transform of $1 - \int_0^t P_B(t')dt'$. To find $\langle \tilde{N}(s) \rangle$ and $\langle \tilde{N}^2(s) \rangle$, it is useful to introduce the dummy variable, x, and to form the probability generating function for N

$$Z(x,s) = \sum_{N=0}^{\infty} \tilde{P}(N,s)x^N \tag{5}$$

The generating function is reminiscent of the partition function from statistical mechanics, because it yields statistical averages when differentiated:

$$\left.\frac{\partial Z}{\partial x}\right|_{x=1} = \sum_{N=0}^{\infty} N\tilde{P}(N,s) = \langle \tilde{N}(s) \rangle \tag{6}$$

$$\left.\frac{\partial^2 Z}{\partial x^2}\right|_{x=1} = \sum_{N=0}^{\infty} (N^2 - N)\tilde{P}(N,s) = \langle \tilde{N}^2(s) \rangle - \langle \tilde{N}(s) \rangle \tag{7}$$

In our case, Z is a straightforward geometric sum

$$Z(x,s) = \frac{1}{s}\left(\frac{1-\tilde{P}(s)}{1-x\tilde{P}(s)}\right) \tag{8}$$

so

$$\langle \tilde{N}(s) \rangle = \frac{1}{s}\left(\frac{\tilde{P}(s)}{1-\tilde{P}(s)}\right) \tag{9}$$

and

$$\langle \tilde{N}^2(s) \rangle = \frac{\tilde{P}^2(s) + \tilde{P}(s)}{s(1-\tilde{P}(s))^2} \tag{10}$$

Equations 9 and 10 were derived in Svoboda et al. (1994). Using these results, we now show that the moments of the distribution $P(\tau)$ are related to those of $N(t)$.

Our strategy will be to re-express $\tilde{P}(s)$ in Equations 9 and 10 in terms of the moments of $P(\tau)$, using the Taylor expansion

$$\tilde{P}(s) = \sum_{n=0}^{\infty} \frac{\langle \tau^n \rangle (-s)^n}{n!} \tag{11}$$

where the moments of $P(\tau)$ are

$$\langle \tau^n \rangle = \int_0^{\infty} \tau^n P(\tau) d\tau = (-1)^n \left.\frac{d^n \tilde{P}(s)}{ds^n}\right|_{s=0} \tag{12}$$

This substitution yields

$$\langle \tilde{N}(s) \rangle = \frac{1}{s^2\langle \tau \rangle} - \left(\frac{\langle \tau \rangle^2 - \langle \tau^2 \rangle/2}{s\langle \tau \rangle^2}\right) + O(1) + \ldots \tag{13}$$

and

$$\langle \tilde{N}^2(s) \rangle = \frac{2}{s^3\langle \tau \rangle^2} - \left(\frac{3\langle \tau \rangle^2 - 2\langle \tau^2 \rangle}{s^2\langle \tau \rangle^3}\right) + O(s^{-1}) + \ldots \tag{14}$$

where we have also carried out an expansion in small s, which corresponds to large t, after taking the inverse Laplace transform, to produce:

$$\langle N(t) \rangle = \frac{t}{\langle \tau \rangle} + \frac{\langle \tau^2 \rangle/2 - \langle \tau \rangle^2}{\langle \tau \rangle^2} + O(t^{-1}) + \ldots \tag{15}$$

$$\langle N^2(t) \rangle - \langle N(t) \rangle^2 = t\left(\frac{\langle \tau^2 \rangle - \langle \tau \rangle^2}{\langle \tau \rangle^3}\right) + O(1) + \ldots \tag{16}$$

In Equations 15 and 16, the series have been truncated in the lower-order powers of t, which do not matter in the limit of large t, i.e., after multiple cycles. This limit applies in a typical experiment, in which hundreds of cycles might accumulate.

In the long-time limit, Equations 15 and 16 tell us how to compute the randomness parameter, r. If we can measure $N(t)$, then the ratio

$$\lim_{t\to\infty}\left(\frac{\langle N^2(t) \rangle - \langle N(t) \rangle^2}{\langle N(t) \rangle}\right) = \left(\frac{\langle \tau^2 \rangle - \langle \tau \rangle^2}{\langle \tau \rangle^2}\right) = r \tag{17}$$

yields r. To study the moments of $P(\tau)$, one need only measure the more accessible quantity $N(t)$, because the terms in Equations 15 and 16 that are proportional to t eventually outgrow any terms that are constant or decay with time. This dominance of terms proportional to t constitutes the mathematical reason that experimental measurements of this parameter remain comparatively immune to sources of noise: Eventually, such terms outstrip any noise terms that are time-invariant.

In the foregoing discussion, we have used the integer $N(t)$ to denote the quantity obtained during an experiment. In practice, however, experimental measurements will not generate integral values for $N(t)$, but instead a continuous version of $N(t)$ corrupted by thermal and instrumentation noise. Denoting the corrupted version of $N(t)$ as $\Omega(t)$ and the com-

bined sources of noise as $\xi(t)$, we have

$$\Omega(t) = N(t) + \xi(t) \tag{18}$$

Does this additional noise affect the measurement of r? Assuming that the additional noise has zero mean, $\langle\Omega(t)\rangle = \langle N(t)\rangle$, and is also statistically stationary, so $\langle\xi^2\rangle$ is constant, then the measured value of r approaches the true value during a long experiment

$$\left(\frac{\langle\Omega^2(t)\rangle - \langle\Omega(t)\rangle^2}{\langle\Omega(t)\rangle}\right) = \left(\frac{\langle N^2(t)\rangle - \langle N(t)\rangle^2 + \langle\xi^2\rangle}{\langle N(t)\rangle}\right) \longrightarrow r \tag{19}$$

The noise $\xi(t)$ simply adds a constant term to the variance in $N(t)$; but in the limit of large t, this term will again be outstripped by terms proportional to t, just as for the other neglected terms in Equation 17. For this reason, measurements of r need not demand a tour de force of experimental technique, because if data can be collected for a sufficient time (which, in turn, places demands on the processivity of the enzyme, as defined previously), the analysis is forgiving with respect to thermal and instrumentation noise. In fact, the ratio in Equation 19, which approaches r in the long-time limit, is the definition of the randomness parameter previously adopted by Svoboda et al. (1994).

Having motivated the feasibility of measuring of r, we proceed to examine a simple reaction pathway. What is the value of r when the enzymatic pathway consists of a sequence of M forward reactions only? In this case, we have

$$1 \xrightarrow{k_1} 2 \xrightarrow{k_2} \cdots \xrightarrow{k_{M-1}} M \xrightarrow{k_M} 1 \tag{20}$$

Again, r can always be computed by solving the system of differential equations corresponding to the reaction of Equation 20, but it is possible to simply write down the answer by inspection in the Laplace transform domain. Since the pathway involves a sequence of M independent Poisson processes, the total cycle time is a sum of M independent Poisson variables. $P(t)$ is therefore the convolution of the M constituent Poisson distributions, which are equivalent to M multiplications in the Laplace transform domain. The Laplace transform of a Poisson density ke^{-kt}, is $\int_0^\infty ke^{-(s+k)t}dt = k/(k+s)$, so

$$\tilde{P}(s) = \prod_{i=1}^{M}\left(\frac{k_i}{k_i + s}\right) \tag{21}$$

Using equations 1, 12, and 21 yields

$$r = \frac{\sum_{i=1}^{M} k_i^{-2}}{\left(\sum_{i=1}^{M} k_i^{-1}\right)^2} \tag{22}$$

Equation 22 can yield insights into the number of rate-limiting steps in the pathway. If, for example, p of the k_i rates are comparable and the other $M - p$ rates are much smaller, then

$$r \approx \frac{1}{p} \tag{23}$$

In the case of the pathway in Equation 20, measurements of $N(t)$ for a single enzyme can lead to a determination of the number of rate-limiting steps, p, through Equations 17 and 23. The reciprocal of r is a lower bound to the total number of intermediate states. In some sense $1/r$ can be thought of as providing a continuous measure of the number of rate-limiting steps, even when the rate constants are all distinct and $1/r$ is non-integral.

The solution-by-inspection method outlined above becomes even handier when reversible reactions are present, i.e., some of the k_{-i} are non-zero. Consider the following pathway:

$$1 \underset{k_{-1}}{\overset{k_1}{\leftrightarrow}} 2 \xrightarrow{k_2} 3 \xrightarrow{k_3} \cdots \xrightarrow{k_{M-1}} M \xrightarrow{k_M} 1 \tag{24}$$

where a single reversible reaction takes place at the first step. The extension to the general pathway of Equation 2 is straightforward. The key insight is that any particular realization of the cycle (Eq. 24) can be viewed as a sequence of Poisson processes. In contrast to the sequence of Equation 20, the sequence in Equation 24 is of variable length, because the enzyme may make any number of trips through the reverse reaction. Each time the enzyme reaches intermediate state 2, there are two possible routes: forward, to state 3, or backward, to state 1. The lifetime of state 2 is a Poisson variable with distribution $(k_{-1}+k_2)e^{-(k_{-1}+k_2)t}$ and the probability of leaving by the backward route is $k_{-1}/(k_{-1}+k_2)$. One can characterize any arbitrary trip through the cycle by the exact number of times, q, that the backward route is chosen from state 2. For fixed q, we apply Equation 21 to find the cycle duration distribution

$$\tilde{P}_q(s) = \left(\frac{k_1}{k_1 + s}\right)^{q+1}\left(\frac{k_{-1} + k_2}{k_{-1} + k_2 + s}\right)^{q+1}\prod_{i=3}^{M}\left(\frac{k_i}{k_i + s}\right) \tag{25}$$

Now, q itself is a random variable with the distribution

$$P(q) = \left(\frac{k_{-1}}{k_{-1} + k_2}\right)^q\left(\frac{k_2}{k_{-1} + k_2}\right) \tag{26}$$

because the probability of making exactly q reversals is the product of the probability of choosing the backward route q times followed by the forward route precisely once. The total cycle time will therefore have the distribution

$$\tilde{P}(s) = \sum_{q=0}^{\infty}\tilde{P}_q(s)P(q) \tag{27}$$

because each $\tilde{P}_q(s)$ is weighted by the probability of its particular value of q. Substituting Equations 25 and 26 into Equation 27 gives

$$\tilde{P}(s)=\sum_{q=0}^{\infty}\left\{\left(\frac{k_{-1}}{k_{-1}+k_2}\right)^q\left(\frac{k_2}{k_2+k_{-1}}\right)\cdot\left(\frac{k_1}{k_1+s}\right)^{q+1}\left(\frac{k_{-1}+k_2}{k_{-1}+k_2+s}\right)^{q+1}\right\}\prod_{i=3}^{M}\left(\frac{k_i}{k_i+s}\right) \quad (28)$$

The continued product in Equation 28 comes from the factor in $\tilde{P}(s)$ arising from the final $M - 2$ reaction states, whereas the sum is the factor arising from the first two states. This sum is geometric and yields

$$\tilde{P}(s)=\left(\frac{k_2k_1}{s^2+s(k_{-1}+k_2+k_1)+k_2k_1}\right)\prod_{i=3}^{M}\left(\frac{k_i}{k_i+s}\right) \quad (29)$$

It should be noted how much simpler it is to write down Equation 28 than to solve the corresponding system of differential equations. For those familiar with probability theory, the sum in Equation 28 is basically the composition of the generating function for q with the Laplace transform of the probability distribution for the time required to leave and then return to state 2. This form of composition is typical for sums of a random number of continuous stochastic variables.

From Equations 1, 12, and 29, r for this system is

$$r=\frac{\sum_{i=3}^{M}\frac{1}{k_i^2}+\frac{(k_1+k_2+k_{-1})^2-2k_1k_2}{(k_1k_2)^2}}{\left(\sum_{i=3}^{M}\frac{1}{k_i}+\frac{k_1+k_2+k_{-1}}{k_1k_2}\right)^2} \quad (30)$$

As required, Equation 30 reduces to Equation 22 when $k_{-1}\to 0$. When all the reaction rates are comparable,

$$r\approx\frac{M+5}{(M+1)^2} \quad (31)$$

As reflected in Equation 31, adding a single reverse process to the reaction pathway in Equation 20, to generate Equation 24, raises the value of r. Even though the mean time, $\langle\tau\rangle$, of the cycle is increased by k_{-1}/k_1k_2, the variance is sufficiently increased that r is always raised. As before, $1/r$ may be interpreted as a lower bound on the total number of intermediate states in the reaction pathway. Figure 1 shows the monotonic rise in r as a function of k_{-1} for $M = 3$, 5, and 7 steps. In the next two sections, we apply Equation 31 to the study of specific enzyme processes.

RESULTS

Enzyme Catalysis with Reversible Substrate Binding

The pathway of Equation 24 describes the action of any enzyme for which there exists just one non-negligible reverse reaction rate (reversible pathways had been independently considered by P. Mitra and K. Svoboda [pers. comm.]). Because convolutions in time become multiplications in the Laplace transform domain, associativity implies that the result in Equation 31 remains unchanged regardless of which particular step in the pathway is reversible. A corollary of this is that values for r do not depend on the particular sequence of the reaction steps. For many enzymes, the primary nonnegligible reversible reaction is that of substrate binding before catalysis. This class of enzymes appears to include several of the molecular motors, among others.

We assume now that substrate binding is first order in the substrate concentration, i.e., $k_1 = k_S[S]$, and that there are no cooperative effects. In this case, we may consider r to be a function of substrate concentration, $[S]$. Figure 2 shows how the value of r changes as a function of $k_S[S]$, for several pathways, with all other forward kinetic rates taken to be equal. The effect of adding reversible substrate binding is demonstrated by the solid curves, as compared with the dashed curves computed for an irreversible binding process. At high substrate concentrations, the unbinding rate is negligible compared to the binding rate, and the corresponding curves approach the same asymptote. At low substrate concentrations, substrate binding becomes the sole rate-limiting step, and all curves approach $r = 1$. At intermediate substrate concentrations, substrate-binding rate becomes comparable to the other rates, hence the randomness parameter dips, but the curves for pathways with reversible binding always lie above those without, as expected from Figure 1. In principle, curves such as those shown in Figure 2 should be observable experimentally, using single-enzyme assays.

How Many ATP Molecules Are Hydrolyzed per 8-nm Kinesin Step?

Unlike myosin, which appears to release from its actin substrate during every enzymatic cycle (Spudich 1993), the kinesin motor is processive, typically undergoing multiple ATP hydrolysis cycles and mechanical steps subsequent to an encounter with a microtubule (Howard et al. 1989; Block et al. 1990; Svoboda et al. 1993; Gilbert et al. 1995). Single kinesin molecules have been observed to propel microtubules or cargo (vesicles or small microspheres) for distances of up to micrometers in vitro (Howard et al. 1989; Block et al. 1990), and the size of their elementary steps has been measured at 8 nm (Svoboda et al. 1994). The kinesin ATPase reaction pathway appears to be distinct from that of myosin, and recent progress by several groups has defined many of the rate constants in the cycle (Hackney 1988; Sadhu and Taylor 1992; Gilbert et al. 1995). Denoting kinesin by K and microtubule by M, two alternative reaction schemes for kinesin have been recently proposed (Gilbert et al. 1995; Johnson and Gilbert 1995):

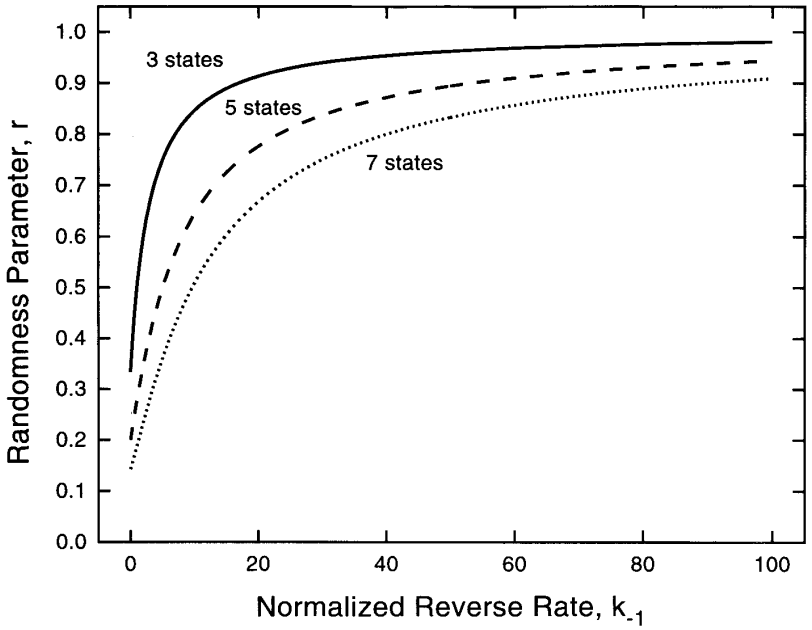

Figure 1. The randomness parameter, r, as a function of the reverse reaction rate, k_{-1}, in the pathway of Eq. 24, computed for different numbers of intermediate states, assuming that all forward rate constants have the same value. The magnitude of k_{-1} is normalized and displayed as a ratio to the value of the forward rate constant. Representative curves are shown for $M = 3$, 5, and 7 states in the pathway. At $k_{-1} = 0$, r equals the reciprocal of the number of states in the pathway (Eq. 23). When k_{-1} grows large, reversals of the first reaction dominate to such an extent that the pathway has effectively one rate-limiting step, and so $r \rightarrow 1$.

Scheme 1:

$$M \cdot K + ATP \underset{200\ s^{-1}}{\overset{2\ \mu M^{-1}s^{-1}}{\leftrightarrow}} M \cdot K \cdot ATP \overset{100\ s^{-1}}{\rightarrow} M \cdot K \cdot ADP \cdot P \overset{20\ s^{-1}}{\rightarrow}$$

$$K \cdot ADP \cdot P \overset{\text{fast, } >100\ s^{-1}}{\rightarrow} K \cdot ADP + P_i \overset{20\ \mu M^{-1}s^{-1}}{\rightarrow}$$

$$M \cdot K \cdot ADP \overset{300\ s^{-1}}{\rightarrow} M \cdot K + ADP \qquad (32)$$

Scheme 2:

$$M \cdot K + ATP \underset{200\ s^{-1}}{\overset{2\ \mu M^{-1}s^{-1}}{\leftrightarrow}} M \cdot K \cdot ATP \overset{100\ s^{-1}}{\rightarrow} M \cdot K \cdot ADP \cdot P \overset{20\ s^{-1}}{\rightarrow}$$

$$M \cdot K^* \cdot ADP + P_i \overset{\text{fast, } >100\ s^{-1}}{\rightarrow} K^* \cdot ADP \overset{\text{fast, } >100 s^{-1}}{\rightarrow}$$

$$K \cdot ADP \overset{20\ \mu M^{-1}s^{-1}}{\rightarrow} M \cdot K \cdot ADP \overset{300\ s^{-1}}{\rightarrow} M \cdot K + ADP \qquad (33)$$

For simplicity, the main catalytic routes are shown (for complete pathways, see Fig. 1 in Gilbert et al. 1995). In both schemes, the only reversible step of any consequence is ATP unbinding: Note that the kinesin ATPase pathway does not involve a reversible hydrolysis step, in contrast to the myosin pathway (Hackney 1988; Gilbert et al. 1995). The large fraction of time spent attached to the microtubule in either scheme would account for the processivity observed in motility assays for single kinesin molecules. An essential difference between these two schemes is that, in Scheme 1, phosphate release occurs subsequent to kinesin unbinding from the microtubule, whereas in Scheme 2, the order of these events is reversed. Both schemes appear consistent with the available data, but Scheme 2 has the drawback of invoking an intermediate kinesin state, K^*, of unproved existence (for discussion, see Johnson and Gilbert 1995). Scheme 1, on the other hand, has microtubule release associated with the rate-limiting step, whereas the rate-limiting step has traditionally been taken to be the point in the cycle at which the putative power stroke occurs, and which would therefore necessitate enzyme-substrate contact in order to bear load. Due to experimental limitations, the two schemes could not be distinguished, and a lower bound of 100 s^{-1} was placed on the indicated fast reactions (Johnson and Gilbert 1995). Note that in Scheme 2, both of the indicated reactions must occur at a combined rate of over 100 s^{-1}. In the vicinity of a microtubule, the rebinding rate $K \cdot ADP \rightarrow M \cdot K \cdot ADP$ is probably very fast, within microseconds.

To account for processivity, several authors have speculated about some form of cooperativity between the two heads of the kinesin molecule, in such a way

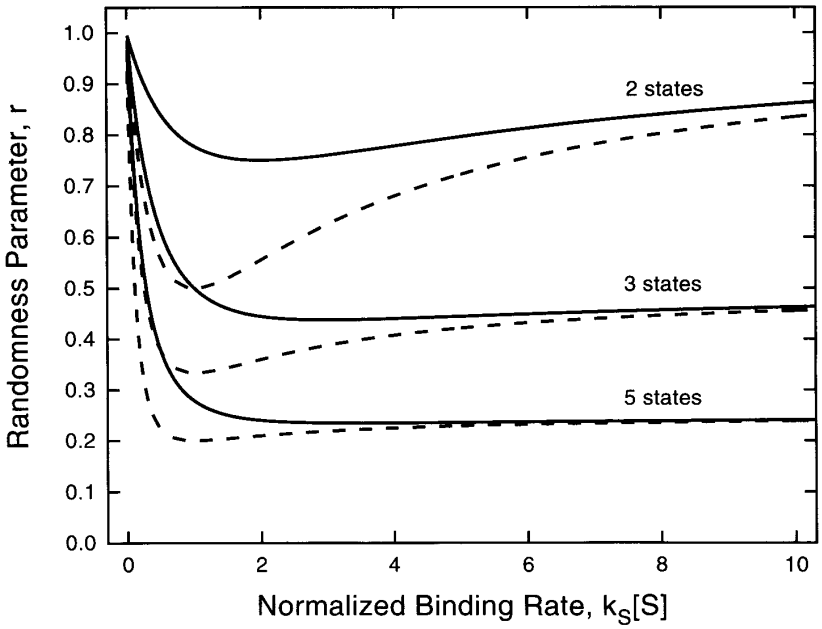

Figure 2. The randomness parameter, r, as a function of the substrate-dependent forward rate, $k = k_S[S]$, in the pathway of Eq. 24, computed for various numbers of intermediate states, both with and without reversible substrate binding. It is assumed that all other rates are constant and identical. The magnitude of the substrate-dependent forward rate is normalized and displayed as a ratio to the value of the constant rate. Representative curves are shown for $M = 2$, 3, and 5 states in the pathway. The solid curves are computed for reversible binding, and the dashed curves are computed for irreversible binding.

that it could move "hand-over-hand" (for discussion, see Block 1995). In such models, the two kinesin heads move in alternating succession (Hackney 1994a; Gilbert et al. 1995; Peskin and Oster 1995). However, biochemical measurements of the overall ATPase rate for kinesin molecules, estimated in a number of laboratories at about 20 s^{-1}•$head^{-1}$, seem hard to reconcile with such enzymes moving in 8-nm mechanical steps. Since single kinesin molecules have been observed to move at speeds of 800–1000 nm•s^{-1} in vitro (and possibly in excess of 2000 nm•s^{-1} in vivo), ATPase rates closer to about 100–200 s^{-1}•$molecule^{-1}$, corresponding to about 50–100 s^{-1}•$head^{-1}$, are needed, subject to the important working assumption that each advance of 8 nm results from the hydrolysis of a single ATP molecule. However, higher rates for truncated, single kinesin head constructs have been measured, and it remains possible that the kinesin ATPase is inhibited when it is not moving on microtubule, thereby skewing many biochemical estimates (Hackney 1994b). Nevertheless, the fundamental question remains as to whether the hydrolysis of a single ATP corresponds to an event that carries a kinesin molecule through 4 nm, 8 nm, or possibly some greater distance. Loosely speaking, this is the mechanochemical coupling problem in a nutshell (for reviews of this issue on myosin and kinesin, see Burton 1992; Block 1995). Since it is not currently feasible to study ATPase rates in the same in vitro motility assays in which stepping is measured (but for recent progress, see Funatsu et al. 1995), we hope to take an indirect approach based on statistics.

If enzymatic cycles and mechanical stepping are tightly coupled, does an advance of 8 nm correspond to ATP hydrolysis by one, or two, kinesin heads? One approach to this important question is suggested by Equation 31. Since several possible physical schemes for producing either 4- or 8-nm steps on a microtubule B-type lattice are under current consideration (Block and Svoboda 1995), statistical analysis may help to narrow down the current range of possibilities.

For the kinetic schemes of Equations 32 and 33, we may compute r as a function of ATP concentration, as shown in Figure 3. The two curves represent estimated bounds, the upper curve being calculated under the assumption that the reaction steps indicated as "fast" are essentially instantaneous, and the lower curve calculated under the assumption that such fast processes occur at a rate of 100 s^{-1}. Note that under the latter assumption, the predicted value for r at saturating ATP from the Gilbert et al. (1995) data is consistent with the earlier measurement of $r \approx 1/2$ (Svoboda et al. 1994). The values of r in Figure 3 were computed subject to the assumption that one 8-nm step corresponds to one

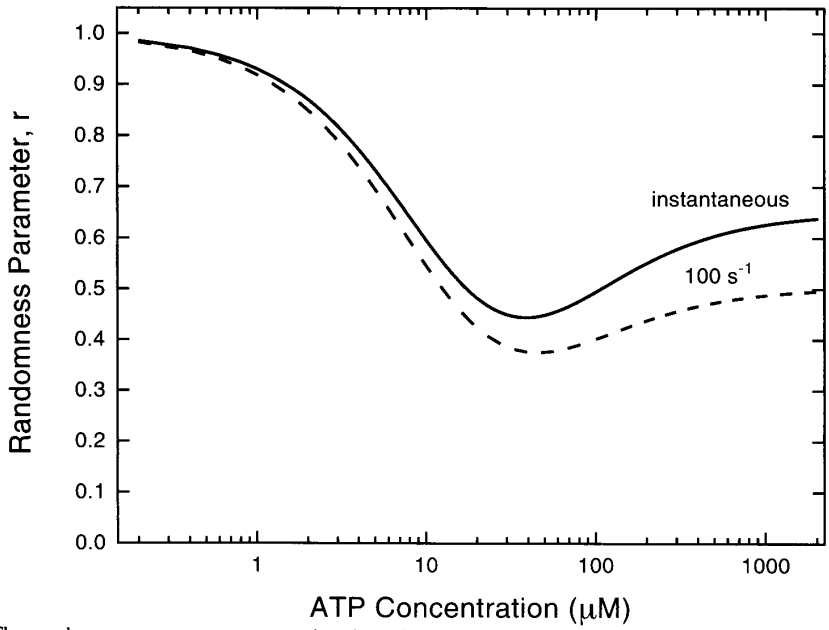

Figure 3. The randomness parameter, r, as a function of ATP concentration, computed from Eq. 30 and plotted semi-logarithmically, for the reaction scheme of either Eqs. 32 or 33. The rate constants published by Gilbert et al. (1995) for a single kinesin head were used. The solid curve shows the calculation subject to the assumption that reactions indicated as "fast" occur instantaneously. The dashed curve shows the same calculation, but with an assumed collective rate of 100 s^{-1} for these reactions.

complete passage through the enzymatic pathway of either Equations 32 or 33. From Equation 1, it is clear that if one 8-nm cycle corresponds instead to two full passages through the scheme (one per head), then the value of r drops to half of that shown in Figure 3, at each concentration of ATP. Predictions for one and two hydrolysis events per 8-nm step are compared in Figure 4.

We are currently in the process of determining r experimentally over a broad range of ATP concentrations. It may seem that the one prior measurement of $r \sim 1/2$ at 2 mM ATP would already weigh in favor of a single ATP hydrolysis per 8 nm of travel, but this finding relies heavily on the particular numerical values for the kinetic constants used. Because these values carry uncertainties of factors of two (perhaps more), it is not yet clear to what degree such comparisons are trustworthy. More reliable evidence would come from a determination of r at rate-limiting concentrations of ATP. In this regime, the ATP-binding step makes the dominant contribution to Equation 19. In the event that one step resulted from each ATP hydrolysis, r would approach unity. If, instead, one step derived from two such hydrolyses, r would approach one-half. The latter analysis should prove more trustworthy, in that it relies neither on numerical values of kinetic rate constants measured in solution, nor on details of how the passage of the two heads through the pathways of Equations 32 or 33 may be coordinated. Note that these reaction schemes refer to one head only. In the case of cooperative heads, each intermediate in the overall reaction pathway must specify a state for each of the two heads simultaneously. Unless the head coordination is such that one begins its trip through the pathway as soon as the other finishes, the theoretical predictions of Figures 3 and 4, which rely on this tacit assumption, will partly fail. The degree of failure depends on details of the coordination between the heads, which must be determined experimentally. Indeed, preliminary evidence indicates that ATP binding to one head does appear to promote the progress of the other head (Hackney 1994a; S.P. Gilbert, pers. comm.). Head coordination affects the predictions for intermediate and high ATP concentrations only, however. At low ATP levels, the predicted values of r remain one or one-half, respectively, for one or two ATP molecules hydrolyzed per 8-nm advance.

DISCUSSION

New experimental techniques for studying single-enzyme molecules demand appropriate methods of kinetic analysis. In traditional biochemical studies with large numbers of proteins, the customary approach has been to measure reaction rates and, thereby, to obtain kinetic constants. Such rates are the reciprocals of the first moment of the probability distribution for the lifetime of the particular intermediate state in question.

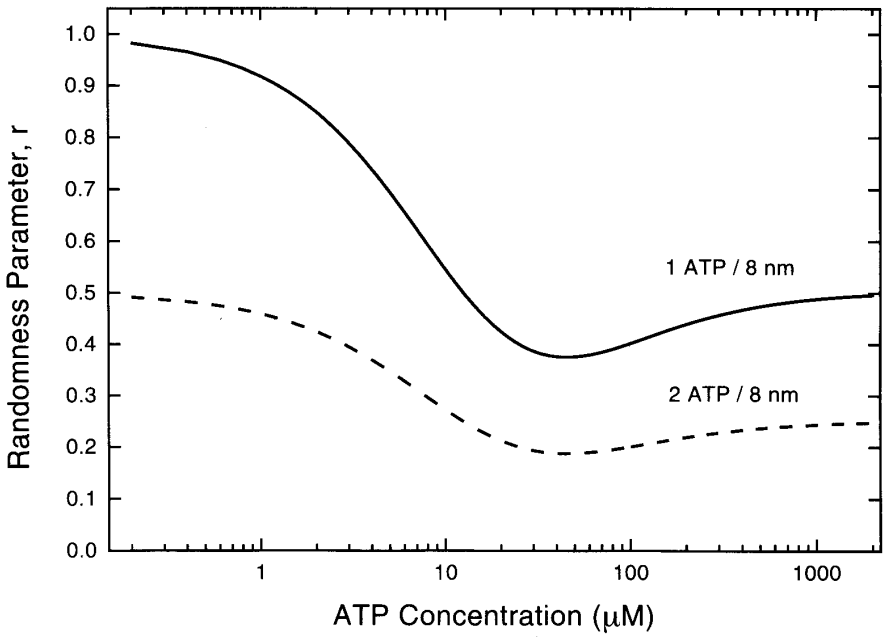

Figure 4. The randomness parameter, r, as a function of ATP concentration, computed from Eq. 30 and plotted semi-logarithmically, for the reaction scheme of Eqs. 32 or 33. The rate constants published by Gilbert et al. (1995) for a single kinesin head were used. The solid curve shows the calculation subject to the assumption that a single ATP is hydrolyzed to produce an 8-nm step. The dashed curve shows the calculation subject to the assumption that two ATPs are hydrolyzed to produce an 8-nm step. Values of the dashed curve are precisely half those of the solid one. Both curves are computed by assuming that "fast" reactions occur at a collective rate of 100 s^{-1}.

In the new single-molecule studies, it becomes useful to study some of the higher moments of these probability distributions. As we have shown here, the second moment can carry especially useful information about the reaction pathway.

Analysis involving the randomness parameter, r, for single processive enzymes has the important advantage that thermal and measurement noise can be rendered insignificant in sufficiently long experiments. In this sense, the statistical analysis is robust. However, a disadvantage to this approach is that the interpretation of results is model-dependent and can prove difficult, especially when the number of rate-limiting steps in the reaction pathway grows large. Furthermore, any given value of r can be produced by a variety of reaction schemes, although it does provide a lower bound on the number of intermediate states. Despite these difficulties, when combined with additional kinetic information derived from biochemical or other data, fluctuation analysis may prove helpful in untangling difficult questions about mechanochemical coupling, as well as providing additional corroborating evidence about a variety of enzymatic pathways.

ACKNOWLEDGMENTS

S.M.B. is supported by a grant from the National Institute of General Medical Sciences. M.J.S. is supported by a National Science Foundation predoctoral fellowship. We gratefully acknowledge Partha Mitra and Karel Svoboda for seminal contributions to this work. We also thank Susan Gilbert and Andrew McKellips for helpful discussions, and Susan Gilbert and Ken Johnson for generously providing data in advance of publication.

REFERENCES

Block, S.M. 1995. Nanometers and piconewtons: The macromolecular mechanics of kinesin. *Trends Cell Biol.* **5:** 169.

Block, S.M. and K. Svoboda. 1995. Analysis of high resolution recordings of motor movements. *Biophys. J.* **68:** 230s.

Block, S.M., L.S.B. Goldstein, and B.J. Schnapp. 1990. Bead movement by single kinesin molecules studied with optical tweezers. *Nature* **348:** 348.

Burton, K. 1992. Myosin step size: Estimates from motility assays and shortening muscle. *J. Muscle Res. Cell Motil.* **13:** 590.

Carrier, G.F., M. Krook, and C.F. Pearson. 1983. *Functions of a complex variable.* Hod Books, Ithaca, New York.

Colquhoun, D. 1971. *Lectures on biostatistics.* Clarendon Press, Oxford.

Colquhoun, D. and A.G. Hawkes. 1977. Relaxation and fluctuations of membrane currents that flow through drug-operated channels. *Proc. R. Soc. Lond. B Biol. Sci.* **199:** 231.

Feller, W. 1968, 1971. *An introduction to probability theory*

and its applications, vol. 1 and vol. 2. Wiley, New York.

Finer, J.T., R.M. Simmons, and J.A. Spudich. 1994. Single myosin molecule mechanics: Piconewton forces and nanometre steps. *Nature* **368:** 113.

Finzi, L. and J. Gelles. 1995. Measurement of lactose repressor-mediated loop formation and breakdown in single DNA molecules. *Science* **267:** 378.

Funatsu, T., Y. Harada, M. Tokunaga, K. Saito, and T. Yanagida. 1995. Imaging of single fluorescent molecules and individual ATP turnovers by single myosin molecules in aqueous solution. *Nature* **374:** 555.

Gilbert, S.P., M.R. Webb, M. Brune, and K.A. Johnson. 1995. Pathway of processive ATP hydrolysis by kinesin. *Nature* **373:** 671.

Guthold, M., M. Bezanilla, D.A. Erie, B. Jenkins, H.G. Hansma, and C. Bustamante. 1994. Following the assembly of RNA polymerase-DNA complexes in aqueous solutions with the scanning force microscope. *Proc. Natl. Acad. Sci.* **91:** 12927.

Hackney, D.D. 1988. Kinesin ATPase: Rate limiting ADP release. *Proc. Natl. Acad. Sci.* **85:** 6314.

Hackney, D.D. 1994a. Evidence for alternating head catalysis by kinesin during microtubule-stimulated ATP hydrolysis. *Proc. Natl. Acad. Sci.* **91:** 6865.

———. 1994b. The rate-limiting step in microtubule-stimulated ATP hydrolysis by dimeric kinesin head domains occurs while bound to the microtubule. *J. Biol. Chem.* **269:** 16508.

Howard, J., A.J. Hudspeth, and R.D. Vale. 1989. Movement of microtubules by single kinesin molecules. *Nature* **342:** 154.

Johnson, K.A. and S.P. Gilbert. 1995. Pathway of the microtubule-kinesin ATPase. *Biophys. J.* **68:** 173s.

Kornberg, A. and T.A. Baker. 1992. *DNA replication,* 2nd edition. W.H. Freeman, New York.

Molloy, J.E., J.E. Burns, J.C. Sparrow, R.T. Tregear, J. Kendrick-Jones, and D.C.S. White. 1995. Single-molecule mechanics of heavy meromyosin and S1 interacting with rabbit or *Drosophila* actins using optical tweezers. *Biophys. J.* **68:** 298s.

Peskin, C.S. and G. Oster. 1995. Coordinated hydrolysis explains the mechanical behavior of kinesin. *Biophys. J.* **68:** 202s.

Radmacher, M., M. Fritz, H.G. Hansma, and P.K. Hansma. 1994. Direct observation of enzyme activity with the atomic force microscope. *Science* **265:** 1577.

Sadhu, A. and Taylor. 1992. A kinetic study of the kinesin ATPase. *J. Biol. Chem.* **267:** 11352.

Samuel, A.D.T. and H.C. Berg. 1995. Fluctuation analysis of rotational speeds of the bacterial flagellar motor. *Proc. Natl. Acad. Sci.* **92:** 3502.

Schafer, D.A., J. Gelles, M.P. Sheetz, and R. Landick. 1991. Transcription by single molecules of RNA polymerase observed by light microscopy. *Nature* **352:** 444.

Spudich, J.A. 1993. Motor molecules in motion. *Nature* **348:** 284.

Stryer, L. 1995. DNA structure, replication, and repair. In *Biochemistry,* 4th edition, chap. 31, p. 800. W.H. Freeman, New York.

Svoboda, K., P.P. Mitra, and S.M. Block. 1994. Fluctuation analysis of motor protein movement and single enzyme kinetics. *Proc. Natl. Acad. Sci.* **91:** 11782.

Svoboda, K., C.F. Schmidt, B.J. Schnapp, and S.M. Block. 1993. Direct observation of kinesin stepping by optical trapping interferometry. *Nature* **365:** 721.

Yin, H., R. Landick, and J. Gelles. 1994. Tethered particle motion method for studying transcript elongation by a single RNA polymerase molecule. *Biophys. J.* **67:** 2468.