NANO LETTERS

2005
Vol. 5, No. 12
2373−2378

Nonequilibrium Steady State of a Nanometric Biochemical System: Determining the Thermodynamic Driving Force from Single Enzyme Turnover Time Traces

Wei Min,[†] Liang Jiang,[‡] Ji Yu,[†] S. C. Kou,[§] Hong Qian,*[,||] and X. Sunney Xie*[,†]

Department of Chemistry and Chemical Biology, Department of Physics, and Department of Statistics, Harvard University, Cambridge, Massachusetts 02138, and Department of Applied Mathematics, University of Washington, Seattle, Washington 98195

Received November 3, 2005

ABSTRACT

A single enzyme molecule in a living cell is a nanometric system that catalyzes biochemical reactions in a nonequilibrium steady-state condition. The chemical driving force, $\Delta\mu$, is an important thermodynamic quantity that determines the extent to which the reaction system is away from equilibrium. Here we show that $\Delta\mu$ for an enzymatic reaction in situ can be determined from the nonequilibrium time traces for enzymatic turnovers of individual enzyme molecules, which can now be recorded experimentally by single-molecule techniques. Three different $\Delta\mu$ estimators are presented from principles of nonequilibrium statistical mechanics: fluctuation theorem, Kawasaki identity, and fluctuation dissipation theorem, respectively. In particular, a maximum likelihood estimation method of $\Delta\mu$ has been derived based on fluctuation theorem. The statistical precisions of these three $\Delta\mu$ estimators are analyzed and compared for experimental time traces with finite lengths.

Introduction. Biochemical reactions in living cells occur in nonequilibrium conditions under which chemical kinetics often governs their biological functions. Conventionally, chemical kinetics has been studied outside living cells in two scenarios. The first one is the equilibrium steady-state. For example, the interconversion between two conformers can be measured from its NMR line shape with motional narrowing at high temperature.[1] The second one is a nonequilibrium non-steady-state. For instance, the stop-flow and temperature-jump experiments monitor the relaxation following a sudden perturbation.[2] In living cells, however, enzyme-mediated biochemical reactions are usually under the condition of the nonequilibrium steady state (NESS).[3−5]

Consider a single enzyme molecule, such as an ATPase, undergoing catalytic turnovers at a particular position inside a cell (e.g., on a membrane), surrounded by constant concentrations of substrates and products, [S] and [P], respectively. As illustrated in Figure 1, an open isothermal−isobaric system is denoted by the dashed circle comprising the enzyme. [S] and [P] are kept constant by continuous supplying of the substrate to and withdrawing the product from the system because of certain cellular mechanisms. We

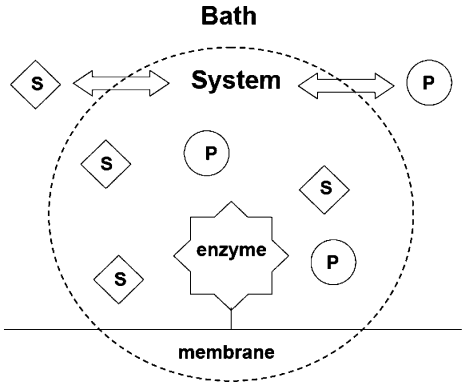

Figure 1. Illustration of single-molecule enzymatic turnovers under a nonequilibrium steady state (NESS). A single enzyme molecule attached on a membrane, surrounded by substrates and products molecules, undergoes catalytic turnovers. The dashed circle denotes an open thermodynamic system in which substrates are continuously converted to products but their concentrations, [S] and [P], are held constant because of the exchange with the bath.

* Corresponding authors. E-mail: xie@chemistry.harvard.edu; qian@amath.washington.edu.
[†] Department of Chemistry and Chemical Biology, Harvard University.
[‡] Department of Physics, Harvard University.
[§] Department of Statistics, Harvard University.
[||] University of Washington.

10.1021/nl0521773 CCC: $30.25
Published on Web 12/14/2005

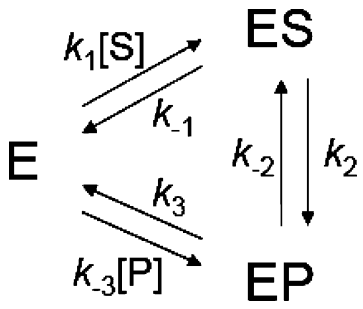

Figure 2. Three-state reversible Michalis−Menten−Briggs−Haldane kinetic scheme with *E*, *ES*, and *EP* denoting free enzyme, enzyme−substrate complex, and enzyme−product complex, respectively. The dependence of forward and backward reaction fluxes, J_+ and J_-, on [S], [P], and rate constants are given by the Briggs−Haldane equation, ref 6. For NESS, the thermodynamic driving force $\Delta\mu = k_BT \ln((k_1k_2k_3[S]/k_{-1}k_{-2}k_{-3}[P]))$.

assume that this enzyme molecule follows reversible Michalis−Menten−Briggs−Haldane cycle kinetics (Figure 2) with *E*, *ES*, and *EP* denoting the free enzyme, enzyme−substrate complex, and enzyme−product complex, respectively.[6] When the enzyme completes a cycle, the system returns back to its initial state with identical thermodynamic state functions of the system. The continuous conversion of substrate to product molecules under this setting represents a NESS in a open system.

The Briggs−Haldane equation explicitly gives the dependence of forward and backward reaction fluxes J_+ and J_- on [S], [P], and rate constants.[6] The net reaction flux $J \equiv J_+ - J_-$ from substrate to product is driven by the thermodynamic driving force $\Delta\mu$[4]

$$\Delta\mu = k_BT \ln\left(\frac{k_1k_2k_3[S]}{k_{-1}k_{-2}k_{-3}[P]}\right) \quad (1)$$

which is dependent upon [S] and [P]. The rate constants are defined in Figure 2. For NESS, $\Delta\mu$ is a nonzero constant. When $\Delta\mu = 0$, the system is at equilibrium, in which [S] and [P] satisfy

$$\frac{[P]}{[S]} = \frac{k_1k_2k_3}{k_{-1}k_{-2}k_{-3}} \equiv K_{eq} \quad (2)$$

where K_{eq} is the equilibrium constant. The thermodynamic quantity, $\Delta\mu$, is an intrinsic state function of a system, quantitatively characterizing the tendency toward its equilibrium. The larger the $\Delta\mu$ value, the greater the driving force. Being a constant under NESS, $\Delta\mu$ governs the single-molecule enzymatic kinetics in situ.

At the single enzyme level, chemical reactions are inherently stochastic and can now be followed by single-molecule techniques.[7] For example, time traces of turnovers of single cholesterol oxidase have been recorded in real time;[8] individual steps of 120° rotation of a single F1-ATPase, a rotary molecular motor on a membrane, have been resolved;[9] discrete 8-nm stepping of a single kinesin molecule, a linear motor walking along microtubule, has been monitored.[10] All of these single-molecule measurements in vitro are also conducted in NESS because the reactions of a single enzyme molecule have negligible effects on [S] and [P] of the sounding environment. Given the ever increasing new single-molecule techniques, it is anticipated that such a time trace of a single enzyme under NESS can be recorded in a live cell in the near future.

In this paper, we show the proof of principle that $\Delta\mu$ can be determined directly from NESS time traces of enzymatic turnovers in light of recent advances in both single-molecule techniques and nonequilibrium statistical mechanical theories. In principle, $\Delta\mu$ can be obtained by measuring [S] and [P], together with the separated ensemble kinetic measurements of the rate constants. In contrast, our approach deduces this thermodynamic quantity from nonequilibrium dynamics. From a practical standpoint, it is highly desirable to determine $\Delta\mu$ without relying on sensors to measure local concentrations of metabolites inside the cell as single-molecule experiments are becoming common practice.

In the following sections, we first introduce the stochastic time traces of single enzyme turnover under NESS. Then, three relations for thermodynamic driving force are presented and related to the principles of nonequilirium statistical mechanics, specifically, fluctuation theorem (FT), Kawasaki identity (KI), and fluctuation dissipation theorem (FDT). In particular, we derive a maximum likelihood estimation method of $\Delta\mu$ based on fluctuation theorem. The statistical precisions of the above three theoretical estimators are analyzed and compared for their efficiency in extracting $\Delta\mu$ information from experimental time traces.

Stochastic Time Traces of Single Enzyme Turnover under NESS. Although a particular fluctuating time trace of single enzyme turnovers is not reproducible, its statistical properties are. The stochasticity of the time traces contains the crucial thermodynamic and kinetic information of the system. The kinetic scheme in Figure 2 is a reversible cycle. At equilibrium, the forward flux, J_+, and the backward flux, J_-, cancel out, and the resulting net flux, J, is zero. When $\Delta\mu > 0$, the system exhibits a net flux from substrate to product, but with occasional backward turnovers originating from thermal fluctuations. The direction of the net flux, J, reverses its sign when $\Delta\mu < 0$. Quantitatively, $\Delta\mu = k_BT \ln(J_+/J_-)$, and $\Delta\mu \times J$ is the entropy production rate that is always positive for nonequilibrium processes.[4,5]

Monte Carlo simulations are carried out to mimic the time traces recorded in a real single-molecule experiment. Figure 3 depicts the simulated turnover traces of a single enzyme molecule under NESS with different thermodynamic driving forces. We consider the sequence of events $E \rightarrow ES \rightarrow EP \rightarrow E$ as a full forward cycle and the sequence of events $E \rightarrow EP \rightarrow ES \rightarrow E$ as a full backward cycle. The cumulative number of net cycles (the full forward cycles minus the full backward cycles) is plotted as a function of time.

It is evident from Figure 3 that different net reaction fluxes, J, (slope) originate from different thermodynamic driving forces. Because of thermal fluctuations of the heat bath, occasional backward turnover cycles occur when the enzyme takes in ambient heat from the bath to do

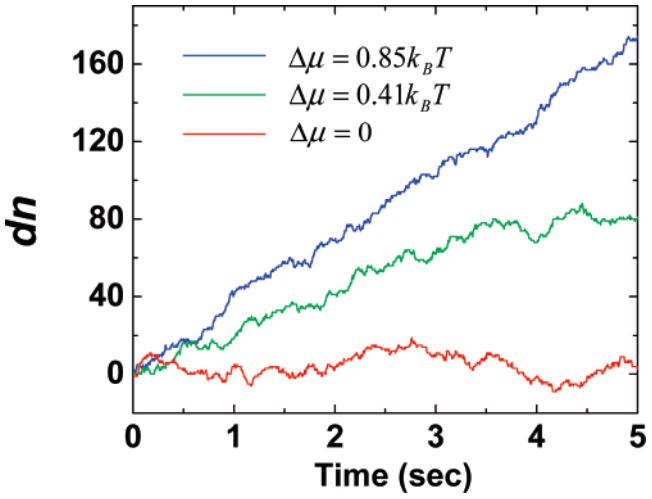

Figure 3. Monte Carlo simulations of turnover traces of a single enzyme molecule under NESS for different $\Delta\mu$ values, mimicking the real time traces recorded in a single-molecule experiment. The cumulative net turnover numbers, *dn*, are plotted against time. Blue curve: $k_1[S] = 368\ s^{-1}$, $k_2 = k_3 = k_{-1} = k_{-2} = 700\ s^{-1}$, $k_{-3}[P] = 158\ s^{-1}$, $J_+ = 70\ s^{-1}$, $J_- = 30\ s^{-1}$, and $\Delta\mu = 0.85\ k_BT$; green curve: $k_1[S] = 315\ s^{-1}$, $k_2 = k_3 = k_{-1} = k_{-2} = 700\ s^{-1}$, $k_{-3}[P] = 210\ s^{-1}$, $J_+ = 60\ s^{-1}$, $J_- = 40\ s^{-1}$, and $\Delta\mu = 0.41k_BT$; red curve: $k_1[S] = 263\ s^{-1}$, $k_2 = k_3 = k_{-1} = k_{-2} = 700\ s^{-1}$, $k_{-3}[P] = 263\ s^{-1}$, $J_+ = 50\ s^{-1}$, $J_- = 50\ s^{-1}$, and $\Delta\mu = 0$.

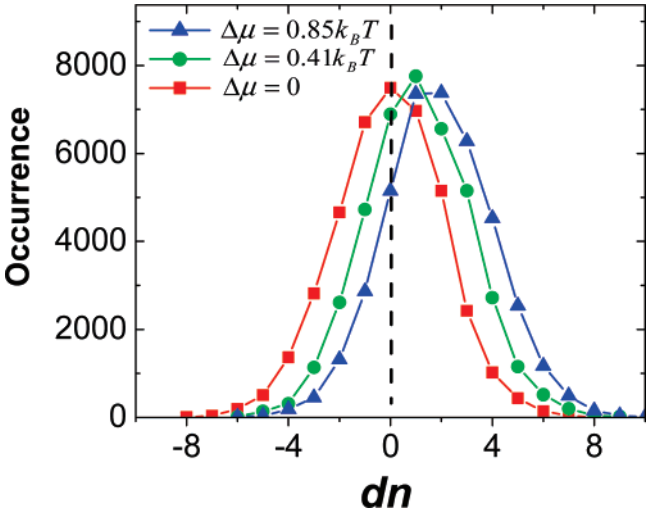

Figure 4. Histogram of enzyme turnovers dn_t for the corresponding stochastic trajectories in Figure 3. The time window, *t*, is chosen to be 0.05 s. It is evident that the dn_t distribution is symmetric for $\Delta\mu = 0$ and biased toward the positive direction for $\Delta\mu > 0$.

chemical work, even though the time-averaged net flux is positive, as required by the second law of thermodynamics. Such fluctuation behavior is highly informative because one can use it to extract the underlying thermodynamic driving force, as we shall show in the following section.

Because of these fluctuations, the number of net turnover cycles accomplished within a given observation time window *t*, dn_t, is a stochastic variable rather than a constant. Figure 4 shows the broad distributions of dn_t for the corresponding time trajectories in Figure 3. We seek to determine $\Delta\mu$ from these distributions of the time traces.

Theoretical Relations between Thermodynamic Driving Force and Nonequilibrium Time Traces. *(a) Fluctuation Theorem (FT) Relation.* From chemical master equations under NESS, it has been shown recently that

$$\frac{\Pr(dn_t = M)}{\Pr(dn_t = -M)} = \exp\left(\frac{\Delta\mu}{k_BT} M\right) \tag{3}$$

where *M* is the positive integer. Equation 3 explicitly gives the ratio between the probability of *M* forward turnovers $\Pr(dn_t = M)$ and that of *M* backward turnover $\Pr(dn_t = -M)$.[11−13] It is instructive to note that eq 3, which is the consequence of microscopic reversibility, is general even when the enzyme molecules exhibit more complex kinetic pathways, as long as the enzyme cycles through a full cycle.[8,14,15]

Equation 3 is intimately related to the remarkable fluctuation theorem,[16−26] which is valid for systems driven arbitrarily far from equilibrium. In its general form, FT provides an analytical expression to relate the probability of observing a process of duration *t* with entropy production $D_t = A$, to that of a process with the same magnitude of entropy change but where the entropy is consumed rather than produced $D_t = -A$

$$\frac{P(D_t = A)}{P(D_t = -A)} = \exp(A) \tag{4}$$

By identifying $(\Delta\mu \cdot dn_t/T)$ in eq 3 as the entropy production, D_t, in isothermal reactions, one can find that eq 3 arises readily from the general form of FT (eq 4), which is the very reason that we term eq 3 as an FT relation.

(b) Kawasaki Identity (KI) Relation. We can derive the following relation by directly integrating eq 3

$$\left\langle \exp\left(-\frac{\Delta\mu}{k_BT} dn_t\right)\right\rangle = \sum_{dn_t=-\infty}^{\infty} \Pr(dn_t) \exp\left(-\frac{\Delta\mu}{k_BT} dn_t\right) = \sum_{dn_t=-\infty}^{\infty} \Pr(-dn_t) = 1 \tag{5}$$

This formula is similar to the so-called Kawasaki identity $\langle\exp(-D_t)\rangle = 1$.[27,28] Positive dn_t trajectories contribute to the Kawasaki average function frequently, but each contribution is small in magnitude because of the negative sign in the exponential. In contrast to their positive counterpart, the infrequent negative dn_t trajectories contribute rarely to the average, but each one is exponentially significant. The exponential rarity of observing negative dn_t trajectories is exactly compensated by the negative exponential in the Kawasaki function. The consequence is that the Kawasaki function has a constant value of unity for all times *t*. It is obvious that, without the occurrence of negative dn_t trajectories, it is impossible for KI to hold.

We note that the KI relation is closely related to other recent advances in the theory of nonequilibrium statistical mechanics, in particular, Jarzynski's nonequilibrium work

relation for equilibrium states[29−32] and Hatano and Sasa's nonequilibrium steady-state equality.[33,34] Recently, eq 5 has been derived based on the generalization of Haldane equation for enzyme kinetics.[35]

(c) Fluctuation Dissipation Theorem (FDT) Relation. Applying cumulant expansion on the exponential function of eq 5 up to the second order, we arrive at the following simplification

$$\frac{\Delta\mu\ \sigma_{dn}^2}{2k_BT\langle dn_t\rangle} = 1 \tag{6}$$

where $\langle dn_t\rangle$ and σ_{dn}^2 are the mean and variance of the dn_t distribution, respectively. We will show that, after neglecting the higher order cumulants, eq 6 can actually be seen as a form of fluctuation dissipation theorem (FDT), which is valid for small fluctuations in the linear response regime.

The first and second laws of thermodynamics state that the mean work $\langle w\rangle$ done to an isothermal−isobaric system to bring about an increase in Gibbs free energy, ΔG, is always no less than the ΔG increase for any arbitrary processes $\langle w\rangle \geq \Delta G$ because of the energy dissipation. The equality holds only when the process is carried out reversibly. For systems that are near equilibrium in which FTD holds, the energy dissipated is proportional to the system's fluctuations. FDT gives an estimate of ΔG for the irreversible process, after correcting the dissipation

$$\Delta G \approx \langle w\rangle - \frac{\sigma_w^2}{2k_BT} \tag{7}$$

where σ_w^2 is the variance of the work distribution.[36,37]

In the NESS situation discussed here, the beginning and the final states are the same; hence, ΔG of the system is zero. Meanwhile, $w = \Delta\mu dn_t$, thus $\langle w\rangle = \Delta\mu\langle dn_t\rangle$ and $\sigma_w^2 = \sigma_{dn}^2\ \Delta\mu^2$. Then, one can solve eq 7 to obtain $\Delta\mu$, which leads to eq 6.

$\Delta\hat{\mu}$ Estimators under Finite Sampling. Let us consider a real single-molecule experiment. For given time window t, let N be the total number of repeated measurements for dn_t that an experimenter can collect. Thus, the product of Nt is then the total length of the single-molecule time trace. It is important to note that the above-presented three theoretical relations are exact only for infinite (perfect) sampling, namely, $N \rightarrow \infty$. However, the finite sampling of single-molecule trajectory is usually the bottleneck of data processing in real experiments. Therefore, there is a strong motivation to understand the statistical errors associated with finite numbers of trials in order to choose appropriate an estimator under given certain experimental conditions.[38−40] In the following, we introduce three $\Delta\hat{\mu}$ estimators for a given set of N measurements of dn_t according to the corresponding theoretical relations discussed in the preceding section.

According to the KI relation, we can solve eq 5 and define a KI estimator

$$\Delta\hat{\mu}_{KI} \equiv \text{nonzero root of } \frac{1}{N}\sum_{i=1}^{N}\exp\left(-\frac{\Delta\mu}{k_BT}dn_i\right) = 1 \tag{8}$$

Similarly, based on FDT relation, eq 6, we can define a FDT estimator

$$\Delta\hat{\mu}_{FDT} \equiv \frac{2k_BT\langle dn\rangle_N}{\sigma_{dn}^2} = \frac{\dfrac{2k_BT}{N}\displaystyle\sum_{i=1}^{N} dn_i}{\dfrac{1}{N-1}\displaystyle\sum_{i=1}^{N}(dn_i - \langle dn\rangle_N)^2} \tag{9}$$

However, it is much more challenging to define a FT estimator based on eq 3 because eq 3 only gives the relation for various pairs of forward turnovers and backward turnovers. In other words, unlike eq 5 and eq 6, eq 3 itself is not an equation for the overall dn_t distribution. Therefore, to extract one $\Delta\mu$ value with the best accuracy, we need to define

$$\Delta\hat{\mu}_{FT} \equiv \Delta\mu \text{ value that fits eq 3 best for all } M \tag{10}$$

To achieve the optimal statistical estimate for all M, here we exploit a maximum likelihood estimation (MLE) approach. The result is as follows

$$\Delta\hat{\mu}_{FT} \equiv \text{nonzero root of } \sum_{M=1}^{\infty}\frac{MN_M e^{-(\Delta\mu/k_BT)M}}{1+e^{-(\Delta\mu/k_BT)M}} = \sum_{M=1}^{\infty}\frac{MN_{-M}}{1+e^{-(\Delta\mu/k_BT)M}} \tag{11}$$

where N_M is the occurrence that M turnovers are observed. The detailed derivation of MLE is described in the Appendix.

This $\Delta\hat{\mu}_{FT}$ works for any observation time window t. When t is approaching zero, M can only be -1, 0, and 1 in the dn_t distribution. Under such a condition, $\Delta\hat{\mu}_{FT}$ expressed in eq 11 reduces back to $N_{M=1}/N_{M=-1} = \exp(\Delta\mu/k_BT)$, which is exactly the same as the reaction flux relation $\Delta\mu = k_BT\ln(J_+/J_-)$.[4,5,35] Therefore, $\Delta\hat{\mu}_{FT}$ can be regarded as the generalization of the reaction flux relation for any arbitrary observation time window, t.

Statistical Precisions and Comparison among Three $\Delta\hat{\mu}$ Estimators. Now we characterize the statistical precisions for the above three estimators. We can define the following three important properties associated with each estimator[38]

$$\text{systematic bias } B(N) = \langle\Delta\hat{\mu}(N)\rangle - \Delta\mu$$

$$\text{standard deviation } SD(N) = \sqrt{\langle(\Delta\hat{\mu}(N) - \langle\Delta\hat{\mu}(N)\rangle)^2\rangle}$$

$$\text{root mean square error } RMSE(N) = \sqrt{\langle(\Delta\hat{\mu}(N) - \Delta\mu)^2\rangle} = \sqrt{B^2(N) + SD^2(N)} \tag{12}$$

where $\Delta\mu$ denotes the true value. The expectation value of an estimator $\langle\Delta\hat{\mu}(N)\rangle$ is obtained by averaging over W

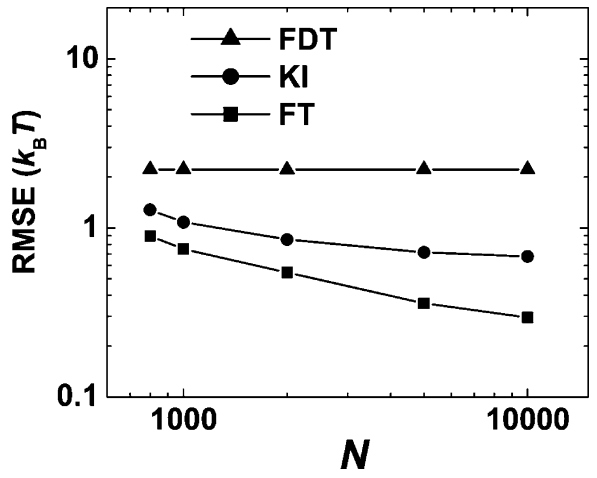

Figure 5. Statistical precisions of three $\Delta\mu$ estimators. Root-mean-square errors (RMSE) for $\Delta\hat{\mu}_{FT}$, $\Delta\hat{\mu}_{KI}$, and $\Delta\hat{\mu}_{FDT}$ are plotted as functions of the total number of repeated measurements, evaluated using eqs 12 and 13. Both *RMSE*(*N*) curves for $\Delta\hat{\mu}_{FT}$ and $\Delta\hat{\mu}_{KI}$ are below the *RMSE*(*N*) curve of $\Delta\hat{\mu}_{FDT}$, indicating that $\Delta\hat{\mu}_{FT}$ and $\Delta\hat{\mu}_{KI}$ behave better than $\Delta\hat{\mu}_{FDT}$. $\Delta\hat{\mu}_{FT}$ is even superior to $\Delta\hat{\mu}_{KI}$, exhibiting a faster convergence of *RMSE* with respect to *N*. In this calculation, $k_1[S] = 430\ s^{-1}$, $k_2 = k_3 = k_{-1} = k_{-2} = 1000\ s^{-1}$, $k_{-3}[P] = 4.3\ s^{-1}$, $J_+ = 100\ sec^{-1}$, $J_- = 1\ s^{-1}$, and $\Delta\mu = 4.6k_BT$. The time window, *t*, for dn_t collection is chosen to be 0.01 s, the time scale of net reaction flux. $W = 5000$ in eq 13.

independent trials of $\Delta\mu$ estimates and letting $W \to \infty$

$$\langle \Delta\hat{\mu}(N) \rangle \equiv \lim_{W\to\infty} \frac{1}{W} \sum_{k=1}^{W} \Delta\hat{\mu}_k(N) \qquad (13)$$

The comprehensive evaluation of the quality of an estimator is *RMSE*(*N*), which is a combination of the systematic error associated with *B*(*N*) and the statistical error associated with *SD*(*N*). The smaller the *RMSE*(*N*) is, the more precise the estimate is.

Here, enzyme turnover cycles are monitored by Monte Carlo simulation. As an illustration, we choose the kinetic parameters to have $J_+ = 100\ sec^{-1}$, $J_- = 1\ s^{-1}$, and $\Delta\mu = 4.6k_BT$ (Figure 5). The time window, *t*, for dn_t collection is chosen to be 0.01 s, the time scale of the net reaction flux. We only discuss the situations in which backward turnovers can be observed in the sampling, permitting the utilization of $\Delta\hat{\mu}_{FT}$ and $\Delta\hat{\mu}_{KI}$.[41] The statistical evaluation results are plotted in Figure 5. The trend is general, even though the specific error values depend on the chosen kinetic parameters and time window *t*.

Both *RMSE*(*N*) curves for $\Delta\hat{\mu}_{FT}$ and $\Delta\hat{\mu}_{KI}$ are below the *RMSE*(*N*) curve of $\Delta\hat{\mu}_{FDT}$, which indicates that $\Delta\hat{\mu}_{FT}$ and $\Delta\hat{\mu}_{KI}$ behave better than $\Delta\hat{\mu}_{FDT}$. This trend is due to the fact that FT and KI are theoretically accurate even far from equilibrium, whereas FDT is valid only in the near equilibrium regime (small $\Delta\mu$). $\Delta\mu = 4.6k_BT$ used in the simulation here is large enough to discriminate the far from equilibrium from the near equilibrium regime. Our simulation demonstrates that it will be more accurate to use $\Delta\hat{\mu}_{FT}$ and $\Delta\hat{\mu}_{KI}$ than $\Delta\hat{\mu}_{FDT}$, when the $\Delta\mu$ is much larger than k_BT.

More interestingly, $\Delta\hat{\mu}_{FT}$ is superior to $\Delta\hat{\mu}_{KI}$, exhibiting a faster convergence of *RMSE* with respect to *N*. One can track the source of the difference between $\Delta\hat{\mu}_{FT}$ and $\Delta\hat{\mu}_{KI}$ by looking at their *B*(*N*) and *SD*(*N*) (see eq 12). Take $N = 10\,000$ for example, the calculated *RMSE*(*N*) for $\Delta\hat{\mu}_{KI}$ and $\Delta\hat{\mu}_{FT}$ are $0.678k_BT$ and $0.295k_BT$, respectively. The corresponding *SD*(*N*) are $0.180k_BT$ and $0.179k_BT$, respectively; the *B*(*N*) are $0.6535k_BT$ and $0.2338k_BT$, respectively. Therefore, it is the smaller systematic error rather than the statistical error that makes $\Delta\hat{\mu}_{FT}$ superior to $\Delta\hat{\mu}_{KI}$. We note that this phenomenon is generally true for other *N*.

We find that such different systematic errors, *B*(*N*), between $\Delta\hat{\mu}_{FT}$ and $\Delta\hat{\mu}_{KI}$ are deeply rooted in their distinct estimation algorithms. $\Delta\hat{\mu}_{FT}$ mainly utilizes the turnover events with $M = +1$ and -1 in the *dn* distribution (see eq 11). However, the major contribution to $\Delta\hat{\mu}_{KI}$ comes from the very rare backward turnovers with large negative *M* because of the exponential averaging (see eq 8). Compared to sampling $M = +1$ and -1 in the *dn* distribution for $\Delta\hat{\mu}_{FT}$, the insufficient sampling on the very rare backward turnovers with large negative *M* for $\Delta\hat{\mu}_{KI}$ is more error-prone, generating bigger systematic error for finite sampling.

Conclusions. Life processes are always far-from-equilibrium. Macromolecular machineries common in biology, such as molecular motors, manifest intrinsic fluctuation at the single-molecule level. Such nonequilibrium fluctuations are investigated in this paper in the context of biochemical reactions of individual enzymes under NESS, in light of recent statistical mechanical theory and single-molecule techniques. Specifically, we have shown that $\Delta\mu$, the thermodynamic driving force of an enzymatic cycle, can be extracted by the nonequilibrium turnover time traces of single enzyme molecules in living cells that might be measurable experimentally. We hope that the theoretical principles and statistical techniques presented in this paper could provide a general methodology for single-molecule nonequilibrium dynamics in living systems.

Acknowledgment. We thank Chris Jarzynski and Binny Cherayil for helpful discussions, and supports from NSF and DOE, Office of Science, Office of Basic Energy Science, Chemical Sciences.

Appendix

Fluctuation theorem, eq 3, says that

$$p_{-M} = p_M \exp\left(-\frac{\Delta\mu}{k_BT} M\right) \quad M = 1, 2, 3, \ldots \qquad (A.1)$$

We want to estimate $\Delta\mu$ by the method of maximum likelihood. To do so, we first write down the likelihood $L((\Delta\mu/k_BT), p_0, p_1, p_2, \ldots)$, the probability of observing the outcome of $\{N_0, N_{\pm1}, N_{\pm2}, N_{\pm3}, \ldots\}$

$$L\left(\frac{\Delta\mu}{k_BT}, p_0, p_1, p_2, ...\right) = \frac{N!}{\prod_{M=-\infty}^{\infty} N_M!} \prod_{M=-\infty}^{\infty} p_M^{N_M} \propto p_0^{N_0} \prod_{M=1}^{\infty} p_M^{N_M} \left[p_M \exp\left(-\frac{\Delta\mu}{k_BT} M\right)\right]^{N_{-M}} \quad \text{(A.2)}$$

The MLE $\Delta\hat{\mu}$ of $\Delta\mu$ can be found by maximizing the likelihood

$$\Delta\hat{\mu} = \arg\max_{\Delta\mu} L\left(\frac{\Delta\mu}{k_BT}, p_0, p_1, p_2, ...\right) \quad \text{(A.3)}$$

Or equivalently

$$\Delta\hat{\mu} = \arg\max_{\Delta\mu} \log\left(L\left(\frac{\Delta\mu}{k_BT}, p_0, p_1, p_2, ...\right)\right) = \arg\max_{\Delta\mu}\left(N_0 \log p_0 + \sum_{M=1}^{\infty} [(N_M + N_{-M}) \log p_M] - \frac{\Delta\mu}{k_BT} \sum_{M=1}^{\infty} MN_{-M}\right) \quad \text{(A.4)}$$

subject to the constraint

$$p_0 + \sum_{M=1}^{\infty} p_M \left[1 + \exp\left(-\frac{\Delta\mu}{k_BT} M\right)\right] = 1 \quad \text{(A.5)}$$

that is, the total probability should be 1.

Denote $l\left(\frac{\Delta\mu}{k_BT}, p_0, p_1, p_2, ...\right) = N_0 \log p_0 + \sum_{M=1}^{\infty} [(N_M + N_{-M}) \log p_M] - \frac{\Delta\mu}{k_BT} \sum_{M=1}^{\infty} MN_{-M}$

To find its maximizer, we use the Lagrange multiplier

$$Lag = l\left(\frac{\Delta\mu}{k_BT}, p_0, p_1, p_2, ...\right) - \lambda\left\{p_0 + \sum_{M=1}^{\infty} p_M \left[1 + \exp\left(-\frac{\Delta\mu}{k_BT} M\right)\right] - 1\right\} \quad \text{(A.6)}$$

The MLEs are the solutions of

$$\left.\frac{\partial Lag}{\partial \Delta\mu}\right|_{(\Delta\hat{\mu}, \hat{p}_0, \hat{p}_1, \hat{p}_2, ...)} = 0, \quad \left.\frac{\partial Lag}{\partial p_0}\right|_{(\Delta\hat{\mu}, \hat{p}_0, \hat{p}_1, \hat{p}_2, ...)} = 0,$$
$$\left.\frac{\partial Lag}{\partial p_M}\right|_{(\Delta\hat{\mu}, \hat{p}_0, \hat{p}_1, \hat{p}_2, ...)} = 0, \quad \left.\frac{\partial Lag}{\partial \lambda}\right|_{(\Delta\hat{\mu}, \hat{p}_0, \hat{p}_1, \hat{p}_2, ...)} = 0 \quad \text{(A.7)}$$

The final solution results in

$$\Delta\hat{\mu}_{FT} \equiv \text{positive root of} \sum_{M=1}^{\infty} \frac{MN_M e^{-(\Delta\mu/k_BT)M}}{1 + e^{-(\Delta\mu/k_BT)M}} = \sum_{M=1}^{\infty} \frac{MN_{-M}}{1 + e^{-(\Delta\mu/k_BT)M}} \quad \text{(A.8)}$$

References

(1) Anderson, P. W. *J. Phys. Soc. Jpn.* **1954**, *9*, 316.
(2) Steinfeld, J. I.; Francisco, J. S.; Hase, W. L. *Chemical Kinetics and Dynamics*, 2nd ed.; Prentice-Hall: New Jersey, 1999.
(3) Wyman, J. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 3983.
(4) Hill, T. L. *Free Energy Transduction and Biochemical Cycle Kinetics;* Springer: New York, 1989.
(5) Qian, H.; Beard, D. A. *Biophys. Chem.* **2005**, *114*, 213.
(6) Segel, I. H. *Enzyme Kinetics*; John-Wiley-Interscience: New York, 1975.
(7) For a recent review, see Tinnefeld, P.; Sauer, M. *Angew. Chem., Int. Ed.* **2005**, *44*, 2642.
(8) Lu, H. P.; Xun, L.; Xie, X. S. *Science* **1998**, *282*, 1877.
(9) Yasuda, R.; Noji, H.; Yoshida, M.; Kinosita, K.; Itoh, H. *Nature* **2001**, *410*, 898.
(10) Asbury, C. L.; Fehr, A. N.; Block, S. M. *Science* **2003**, *302*, 2130.
(11) Gaspard, P. *J. Chem. Phys.* **2004**, *120*, 8898.
(12) Seifert, U. *J. Phys. A: Math. Gen.* **2004**, *37*, L517.
(13) Seifert, U. *Europhys. Lett.* **2005**, *70*, 36.
(14) Flomenbom, O.; Velonia, K.; Loos, D.; Masuo, S.; Cotlet, M.; Engelborghs, Y.; Hofkens, J.; Rowan, A. E.; Nolte, R. J. M.; van der Auweraer, M.; de Schryver F. C.; Klafter, J. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2368.
(15) Kou, S. C.; Cherayil, B. J.; Min, W.; English, B. P.; Xie, X. S. *J. Phys. Chem. B* **2005**, *109*, 19068.
(16) Evans, D. J.; Cohen, E. G. D.; Morriss, G. P. *Phys. Rev. Lett.* **1993**, *71*, 2401.
(17) Gallavotti, G.; Cohen, E. G. D. *Phys. Rev. Lett.* **1995**, *74*, 2694.
(18) Kurchan, J. *J. Phys. A: Math. Gen.* **1998**, *31*, 3719.
(19) Crooks, G. E. *Phys. Rev. E* **1999**, *60*, 2721.
(20) Evans, D. J.; Searles, D. J. *Adv. Phys.* **2002**, *51*, 1529.
(21) van Zon, R.; Cohen, E. G. D. *Phys. Rev. Lett.* **2003**, *91*, 110601.
(22) Wang, G. M.; Sevick, E. M.; Mittag, E.; Searles, D. J.; Evans, D. J. *Phys. Rev. Lett.* **2002**, *89*, 050601.
(23) Carberry, D. M.; Reid, J. C.; Wang, G. M.; Sevick, E. M.; Searles, D. J.; Evans, D. J. *Phys. Rev.* Lett. **2004**, *92*, 140601.
(24) Jarzynski, C.; Wojcik, D. K. *Phys. Rev. Lett.* **2004**, *92*, 230602.
(25) Bustamante, C.; Liphardt, J.; Ritort, F. *Physics Today* **2005**, *58*, 43.
(26) Collin, D.; Ritort, F.; Jarzynski, C.; Smith, S. B.; Tinoco, I., Jr.; Bustamante, C. *Nature* **2005**, *437*, 231.
(27) Yamada, T.; Kawasaki, K. *Prog. Theor. Phys.* **1967**, *38*, 1031.
(28) Carberry, D. M.; Williams, S. R.; Wang, G. M.; Sevick, E. M.; Evans, D. J. *J. Chem. Phys.* **2004**, *121*, 8179.
(29) Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690.
(30) Hummer, G.; Szabo, A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3658.
(31) Liphardt, J.; Dumont, S.; Smith, S. B.; Tinoco, I., Jr.; Bustamante, C. *Science* **2002**, *296*, 1832.
(32) Schurr, J. M.; Fujimoto, B. S. *J. Phys. Chem. B* **2003**, *107*, 14007.
(33) Hatano, T.; Sasa, S. *Phys. Rev. Lett.* **2001**, *86*, 3463.
(34) Trepagnier, E. H.; Jarzynski, C.; Ritort, F.; Crooks, G. E.; Bustamante, C.; Liphardt, J. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 15038.
(35) Qian, H.; Xie, X. S. Submitted for publication, 2005.
(36) Callen, H. B.; Welton, T. A. *Phys. Rev.* **1951**, *83*, 34.
(37) Herman, J. *J. Phys. Chem.* **1991**, *95*, 9029.
(38) Gore, J.; Ritort, F.; Bustamante, C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12564.
(39) Zuckerman, D. M.; Woolf, T. B. *Phys. Rev. Lett.* **2002**, *89*, 180602.
(40) Shirts, M. R.; Bair, E.; Hooker, G.; Pande, V. S. *Phys. Rev. Lett.* **2003**, *91*, 140601.

(41) We note that, for the use of $\Delta\hat{\mu}_{FDT}$, such a backward turnover sampling is not required, because $\Delta\hat{\mu}_{FDT}$ needs only the mean and the variance of the dn_t distribution (see eq 9). Such a straightforward algorithm of $\Delta\hat{\mu}_{FDT}$ provides a simple and quick estimation of $\Delta\mu$ even the backward turnover is not observed in the sampling.

NL0521773