ELSEVIER

Biophysical Chemistry 101–102 (2002) 565–576

Biophysical Chemistry

www.elsevier.com/locate/bpc

Single-molecule enzymology: stochastic Michaelis–Menten kinetics

Hong Qian[a], Elliot L. Elson[b,*]

[a]*Department of Applied Mathematics, University of Washington, Seattle, WA 98195, USA*
[b]*Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, 660 South Euclid Ave. Box 8231, St. Louis, MO 63110, USA*

Received 23 January 2002; accepted 8 April 2002

Abstract

We provide a stochastic analysis of single-molecule enzymatic reactions that follow Michaelis–Menten kinetics. We show that this system can exhibit oscillatory behavior in the non-equilibrium steady-state at appropriate substrate concentrations. The stochastic model includes both enzyme dynamics and substrate turnover kinetics. The relationship between the probability of substrate survival and the time-correlation of enzyme conformation trajectories is discussed. Deterministic kinetics at large substrate concentrations are obtained as a limit of the stochastic model. We suggest that in addition to fluctuating enzyme conformation, the stochastic nature of substrate concentration fluctuations is another possible source of the complex behavior of single-molecule enzyme kinetics.

Keywords: Stochastic analysis; Oscillatory behaviour; Non-equilibrium; Enzyme dynamics; Deterministic kinetics

1. Introduction

Macromolecular fluctuation has been an essential element in John Schellman's thinking about proteins and DNA [1,2]. Therefore, it comes as no surprise that he has always supported developing methods, both experimental and theoretical, to probe the stochastic nature of the molecular world, as early as the conception of fluorescence correlation spectroscopy (FCS) in the 1970s and as recent as the emergence of single-molecule biophysics. In his own work, John has emphasized quantitative physiochemical approaches, both theoretical and experimental, to study biochemical processes. Following this tradition, we dedicate this work to him.

Laboratory measurements based on confocal fluorescence microscopy, near field optics, two-photon excitation, and single-photon detection have provided the possibility to study enzymology in aqueous solution with single proteins, one molecule at a time [3]. In the past decade, we have witnessed a rapid growth in studying dynamics of single enzymes using FCS and a large repertory of related optical methods [4–8]. The microscopic

*Corresponding author. Tel.: +1-314-362-3346; fax: +1-314-362-7183.
E-mail address: elson@biochem.wustl.edu (E. L. Elson).

0301-4622/02/$ - see front matter
PII: S0301-4622(02)00145-X

motion of an enzyme molecule undergoes rapid thermal fluctuation due to its incessant collisions with the solvent molecules. Therefore, the experimental data obtained from single-molecule enzymology are inevitably stochastic [9]. However, these fluctuations in the measurements are not undesirable noise but are rather valuable data from which the enzyme kinetics can be discerned [4]. The current state of the field is in parallel to that of single-channel recording on membrane proteins in the 1970s [10].

The fluctuating measurements require stochastic models for their data interpretation. Although there is a large amount of literature on membrane channel kinetics based on discrete-state Markov models [10,11], and on fluctuating enzymes and stochastic macromolecule mechanics based on diffusion models [12,13], the cyclic and bimolecular nature of enzyme kinetics has its own unique stochastic properties which have not been fully analyzed. This paper addresses these latter aspects of single-molecule enzymology.

There is a need for a conceptual shift from deterministic kinetics based on molecular concentrations in traditional enzymology to stochastic kinetics based on state probabilities in single-molecule enzymology. With only a single enzyme under measurement at a time, we no longer deal with concentrations of the enzyme. Rather, we must adopt the language of probability. Instead of asking, 'what is the concentration of enzyme X with conformation C at the time t?', it is more natural and logical to ask 'what is the probability of the enzyme X being in conformation C at the time t?' This change in the language is not merely semantic; the stochastic thinking is essential to understanding and characterizing behavior of single molecules.

In principle, there are two types of measurements in single-molecule enzymology. One measures the stochastic conformational dynamics of an enzyme turnover and the resulting data have been called 'trajectories' [14]. One essential feature of such an experiment with a single enzyme is that the concentrations of the substrate(s) and the product(s) are approximately constant over the course of the entire experiment. Thus from the perspective of the enzyme, a trajectory is a stationary stochastic process which has to be analyzed using statistical methods. Recently developed experimental methodology has been focused on this approach. Alternatively, one can measure the number of substrate (or product) molecules and their concentration fluctuations in a small system in which the concentration is sufficiently low or the probe is particularly sensitive. The latter approach, which is closer to the traditional deterministic enzymology [15], could have applications in measuring steady-state metabolites involved in enzyme reactions in living cells [16]. This type of measurement was first carried out to study the random walk movements of ATPase-motor proteins in which single enzyme stochastic turnover is observed in terms of the linear motor stepping [17].

From a theoretical standpoint, as we shall show, these two approaches complement each other. There is a natural mathematical relation between them. Measuring the substrate (or product) concentration is intimately related to the concept of 'survival probability' on which a large amount of literature exists [12,18]. One of the objectives of the present work is to provide a theoretical framework which unifies the quantitative analysis of the enzyme reaction in terms of the survival probability of its substrates and the dynamics of the single enzyme molecule in terms of its stochastic trajectory.

The paper is developed as follows. In the first section the stochastic version of a three-state Michaelis–Menten mechanism is analyzed from the enzyme perspective. The non-equilibrium nature of single enzyme kinetics is shown to lead to oscillatory correlation, consistent with recent measurements on horseradish peroxidase [7]. The range of substrate concentration in which oscillation appears is given, and data analyses based on high-order time correlation functions are discussed. In the second section we analyze stochastic kinetics from the perspective of the substrate, which can be modeled as a random walk. The survival probability is introduced, for which a simple diffusion-like equation is obtained as an approximation. The solution to this equation is a stochastic process, the mean of which is shown to agree with the standard deterministic Michaelis–Menten

kinetics. The equation for stochastic Michaelis–Menten kinetics, therefore, provides a way to calculate the distribution for the survival probability. Finally, we show there is a mathematical relation, a duality between the substrate kinetics, in terms of the waiting time for the next product, and the enzyme dynamics, in terms of the steady-state cyclic reaction. The same relation arises in the kinetics of a single motor protein: its ATPase and its stochastic stepping [19]. The oscillation in enzyme dynamics corresponds to a regularity in the arrival time of the product. Hence, with irreversibility, even a single enzyme molecule can exhibit certain deterministic behavior. The same idea leads to the mechanism of 'kinetic timing' in GTPase signaling [20]. In the discussion section, we briefly review some of the earlier literature on the stochastic treatment of biochemical reactions.

2. Non-equilibrium steady-state enzyme turnover

The simplest Michaelis–Menten type model for enzyme kinetics is [15]

$$E+S\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}}ES\underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}}E+P$$

The enzyme in this model has only two conformational states E and ES. In most in vitro experiments with single enzymes, the concentrations of S and P are approximately constant. Therefore, the single-enzyme kinetics is equivalent to the well studied open–close, unimolecular kinetics of single membrane channel proteins: $E\underset{\beta}{\overset{\alpha}{\rightleftharpoons}}ES$ with $\alpha = k_1[S]+k_{-2}[P]$ and $\beta = k_{-1}+k_2$. The kinetics of the two-state system can be found in any standard book on single-channel recording [10] and also in recent work on single-molecule enzymology [21]. The kinetics is simple and there can be only one exponential term, $e^{-(\alpha+\beta)t}$. Some of the essential results for the two-state model are collected in Appendix A.

Few enzymes exhibit single exponential kinetics (cf. [14]). We now consider a Michaelis–Menten type kinetics for a single enzyme in a steady-state [22]:

$$E+S\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}}ES\underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}}EP\underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}}E+P \tag{1}$$

Surprisingly, this basic enzyme kinetic model so far has not been fully investigated in the stochastic terms necessary for single-molecule enzymology. While in an experiment with a single enzyme the concentrations of S and P are approximately constant, usually they are not at chemical equilibrium: $[P]/[S] \neq k_1k_2k_3/k_{-1}k_{-2}k_{-3}$. Because of this condition single enzyme kinetics are fundamentally different from membrane channel kinetics. This is the theoretical justification for studying, in full detail, the three-state model, which is the simplest model capable of exhibiting non-equilibrium steady-state kinetics with the possibility of complex kinetics.

The enzyme in Eq. (1) has three conformational states in a cyclic reaction [19,23,24]. Thus, the kinetics of the single enzyme can be modeled by a three-state Markov process [25]:

$$\frac{d}{dt}\begin{pmatrix}P_E\\P_{ES}\\P_{EP}\end{pmatrix}(t) = \begin{pmatrix}-k_1[S]-k_{-3}[P] & k_{-1} & k_3\\ k_1[S] & -k_{-1}-k_2 & k_{-2}\\ k_{-3}[P] & k_2 & -k_{-2}-k_3\end{pmatrix} \times \begin{pmatrix}P_E\\P_{ES}\\P_{EP}\end{pmatrix} \tag{2}$$

where $P_E(t)$, $P_{ES}(t)$ and $P_{EP}(t)$ are the probabilities of the enzyme being in the E, ES and EP states at time t, respectively. In addition to a time-inde-

pendent term ($\lambda_0=0$), there are two exponential terms, $e^{-\lambda_1 t}$ and $e^{-\lambda_2 t}$, in the kinetics of this enzyme system. However, the λ's need not be real; hence, the kinetics can be oscillatory. If all the rate constants in Eq. (1) as well as $[S]$ and $[P]$ are given, then the λ's can be computed as the eigenvalues of the matrix in Eq. (2):

$$\lambda_{1,2}=-\frac{1}{2}\Big[\big(k_1[S]+k_{-1}+k_2+k_{-2}+k_3 +k_{-3}[P]\big)\pm\sqrt{\Delta}\Big] \quad (3)$$

where

$$\Delta=\big(k_1[S]+k_{-1}-k_2-k_{-2}-k_3+k_{-3}[P]\big)^2 -4\big(k_2-k_{-3}[P]\big)(k_3-k_{-1})$$

For most laboratory measurements, $[P]\approx 0$. Then one immediately sees that $k_{-1}\ll k_3$ yields non-real eigenvalues. In fact, it is easy to verify that when the substrate concentration $[S]$ satisfies

$$\frac{k_{-2}+\left(\sqrt{k_2}-\sqrt{k_3-k_{-1}}\right)^2}{k_1}<[S] <\frac{k_{-2}+\left(\sqrt{k_2}+\sqrt{k_3-k_{-1}}\right)^2}{k_1}$$

$\Delta<0$. Hence, there is a range of substrate concentration under which the enzyme kinetics are oscillatory. Such oscillation has been reported recently [7].

Three-state, non-equilibrium cyclic kinetics has been studied in the past in various biochemical contexts: for example, kinetic proof-reading [26], motor protein kinetics [19], and light-induced chemical steady-states [23]. Nitzan and Ross [23] discovered that a light-induced chemical reaction system could oscillate, but that even in the most favorable condition a three-state system is heavily damped. Whether one can experimentally observe a clear oscillatory behavior depends on the ratio between the imaginary (I) and real (R) parts of the eigenvalues; a low I/R-ratio indicates that only a few oscillations can be observed in the relaxation kinetics. For the three-state kinetic cycle, the I/R-ratio was estimated to be 0.5 [23] and more recently one of us has shown that it is $\frac{1}{\sqrt{3}}$ [27]. With increasing numbers of states within a kinetic cycle, the I/R-ratio increases [27]. This is the mathematical basis for improving the timer quality in GTPase signaling [20].

The oscillatory kinetics reflects the non-equilibrium nature of the single enzyme turnover with a circular flux [19] $J=\big(k_1k_2k_3[S]-k_{-1}k_{-2}k_{-3}[P]\big)/\lambda_1\lambda_2$. The oscillation and the circular flux are intimately related according to the theory of stochastic resonance [27,28] (Ref. [28] is the earliest report on stochastic resonance without forcing as far as we know). Finally, the chemical energy in the transformation of $S\rightarrow P$, on average one turnover per $1/J$ time, becomes the heat dissipated into the aqueous solution. It is important to point out that if the substrate and product are in chemical equilibrium, i.e. $[P]/[S]=k_1k_2k_3/k_{-1}k_{-2}k_{-3}$, then $J=0$. Furthermore, both λ's are necessarily real because $\Delta>0$ (see Appendix B for a proof).

To avoid cluttering, we will denote states E, (ES), and (EP) by 1, 2, 3, respectively, throughout the remainder of this paper.

2.1. Steady-state probability

The dynamics of a single enzyme molecule, a trajectory, is a stochastic process in which the protein jumps among the three states in a seemingly random fashion [14]. The steady-state probabilities for the three states are

$$P_1^{SS}=\frac{k_2k_3+k_3k_{-1}+k_{-1}k_{-2}}{\lambda_1\lambda_2}$$
$$P_2^{SS}=\frac{k_1[S](k_3+k_{-2})+k_{-2}k_{-3}[P]}{\lambda_1\lambda_2}$$
$$P_3^{SS}=\frac{k_1k_2[S]+(k_2+k_{-1})k_{-3}[P]}{\lambda_1\lambda_2} \quad (4)$$

2.2. Transition probability

A unique feature of the single enzyme trajectory is that it contains information on both steady-state and transient kinetics. While the former is expressed in terms of the steady-state probability (Eq. (4)), the latter is contained in transition probabilities, $P_{ij}(t, t+\tau)$, which is the probability

of the enzyme being in state j at time $t+\tau$ when it is in state i at time t. When an enzyme is in a steady-state, the transition probabilities are independent of t. There are nine such functions $P_{ij}(\tau)$, $i, j=1, 2, 3$. For example,

$$P_{13}(\tau)=\frac{k_{-3}[P]+\lambda_2 P_3^{SS}}{\lambda_1-\lambda_2}e^{\lambda_1\tau}+\frac{k_{-3}[P]+\lambda_1 P_3^{SS}}{\lambda_2-\lambda_1}e^{\lambda_2\tau}+P_3^{SS} \quad (5)$$

$$P_{11}(\tau)=\frac{\lambda_2(1-P_1^{SS})+k_1[S]+k_{-3}[P]}{\lambda_2-\lambda_1}e^{\lambda_1\tau}+\frac{k_1[S]+k_{-3}[P]+\lambda_1(1-P_1^{SS})}{\lambda_1-\lambda_2}e^{\lambda_2\tau}+P_1^{SS} \quad (6)$$

$$P_{33}(\tau)=\frac{\lambda_2(1-P_3^{SS})+k_3+k_{-2}}{\lambda_2-\lambda_1}e^{\lambda_1\tau}+\frac{k_3+k_{-2}+\lambda_1(1-P_3^{SS})}{\lambda_1-\lambda_2}e^{\lambda_2\tau}+P_3^{SS} \quad (7)$$

2.3. Time correlation functions

As in any other kinetic studies, the stochastic dynamics of an enzyme molecule can be observed only through appropriate optical or other properties of the molecule. The relation between the spectroscopic signals and the different states of the enzyme, thus, becomes an important component in data analysis. More often than not, the multiple conformational states have the same spectroscopic properties, i.e. they cannot be distinguished by optical measurements. This problem has been extensively discussed in the literature on modeling single-channel recording [11].

While the state probabilities and transition probabilities (Eqs. (4)–(7)) are independent of the signal by which the kinetics are measured, further assumptions or knowledge are needed to analyze a single enzyme trajectory measured in terms of an optical signal. Note that different assumptions at this stage lead to different conclusions about the enzyme dynamics from same experimental data. To illustrate this, we will consider two specific cases motivated by recent work on horseradish peroxidase [7].

Let us first assume that only state 3 (EP) has significant fluorescence signal. Without loss of generality, we then assume that the fluorescence signal $f(i)$:

$$f(1)=f(2)=0, \text{ and } f(3)=1 \quad (8)$$

The mean fluorescence signal is

$$\langle f\rangle=\sum_{i=1}^{3}f(i)P_i^{SS}=P_3^{SS}$$

and time correlation function

$$C_2(\tau)=\langle\Delta f(0)\Delta f(\tau)\rangle=\sum_{i,j=1}^{3}f(i)f(j)P_i^{SS}P_{ij}(\tau)-\langle f\rangle^2 =P_{33}(\tau)P_3^{SS}-(P_3^{SS})^2 \quad (9)$$

One notices that C_2 has two exponential terms from $P_{33}(\tau)$. It is also important to remember that while the dynamics of the enzyme is a Markov process, $f(t)$ is not. Generally speaking, any signal which lumps multiple kinetic states together will be non-Markovian.

With increasing quality of the experimental data, higher-order correlation functions can be introduced to discern the complexity of the kinetics [7,21,29,30]:

$$\begin{aligned}C_3(\tau_1,\tau_2)&=\langle\Delta f(0)\Delta f(\tau_1)\Delta f(\tau_1+\tau_2)\rangle\\&=\sum_{i,j,k=1}^{3}f(i)f(j)f(k)P_i^{SS}P_{ij}(\tau_1)P_{jk}(\tau_2)\\&\quad-\left[C_2(\tau_1)+C_2(\tau_2)+C_2(\tau_1+\tau_2)\right]\langle f\rangle\\&\quad-\langle f\rangle^3\\&=\frac{C_2(\tau_1)C_2(\tau_2)}{P_3^{SS}}-P_3^{SS}C_2(\tau_1+\tau_2)\end{aligned}$$

If we introduce normalized correlation functions $c_2(\tau)=C_2(\tau)/C_2(0)$ and $c_3(\tau_1,\tau_2)=C_3(\tau_1,\tau_2)/C_3(0,0)$, then we have

$$\frac{c_3(\tau_1,\tau_2)}{c_2(\tau_2)}-c_2(\tau_1)=\frac{\langle f\rangle}{2\langle f\rangle-1}\left(\frac{c_2(\tau_1+\tau_2)}{c_2(\tau_2)}-c_2(\tau_1)\right) \quad (10)$$

Except for the simplest two-state kinetics in which $c_1(\tau)$ is a single exponential function (see

Appendix A), in general Eq. (10) is not zero. Hence, Eq. (10) can be used to test the validity of a two-state model.

Edman and Rigler [7] have introduced an expression they called the 'non-Markovian function', ϕ, which is similar to Eq. (10) but is expressed using different normalized correlation functions: $G_1(\tau)=\langle f(0)f(\tau)\rangle/\langle f\rangle^2=P_{33}(\tau)/P_3^{SS}$ and $G_2(\tau_1,\tau_2)=\langle f(0)f(\tau_1)f(\tau_2)\rangle/\langle f\rangle^3=$ $P_{33}(\tau_1)P_{33}(\tau_2)/(P_3^{SS})^2$. In terms of the G's the non-Markovian function can be expressed as

$$\phi=\frac{G_2(\tau_1,\tau_2)}{G_1(\tau_2)}-G_1(\tau_1)=0 \tag{11}$$

for the three-state model with Eq. (8). In fact, it is easy to show that $\phi=0$ holds true for Markov processes with any number of states, if only one of the states has a non-zero signal.

In general, however, ϕ is not zero, even for a Markovian system. Let us consider an example, still using the kinetic scheme (Eq. (1)), but now with $f(1)=0$, and $f(2)=f(3)=1$. The mean fluorescence signal then is $\langle f\rangle=P_2^{SS}+P_3^{SS}=1-P_1^{SS}<1$. The correlation functions are

$$\langle f(0)f(\tau)\rangle=1-2P_1^{SS}+P_1^{SS}P_{11}(\tau)$$

$$\langle f(0)f(\tau_1)f(\tau_2)\rangle=1-3P_1^{SS}+P_1^{SS}\big[P_{11}(\tau_1)+P_{11}(\tau_2)+P_{11}(\tau_1+\tau_2)-P_{11}(\tau_1)P_{11}(\tau_2)\big]$$

Hence

$$\phi=\frac{G_2(\tau_1,\tau_2)}{G_1(\tau_2)}-G_1(\tau_1)=\frac{(1-\langle f\rangle)\big[\zeta(\tau_1+\tau_2)-\zeta(\tau_1)\zeta(\tau_2)\big]}{\langle f\rangle(1-\langle f\rangle)\zeta(\tau_2)+\langle f\rangle^2}$$

where $\zeta(\tau)=\big[P_{11}(\tau)-P_1^{SS}\big]/(1-P_1^{SS})$. For complex $\lambda_{1,2}=-\mu+i\upsilon$, $\zeta(t)=e^{-\mu t}\cos(\upsilon t)$. Thus, $\zeta(t_1+t_2)-\zeta(t_1)\zeta(t_2)=-e^{-\mu(t_1+t_2)}\sin(\upsilon t_1)\sin(\upsilon t_2)\neq 0$ in general. Neither terms in the numerator is zero. Fig. 1 shows an example for the ϕ function with complex $\lambda_{1,2}$.

3. Stochastic substrate kinetics

In the previous section, it is assumed that the concentrations of the substrate and the product of the enzymatic reaction are constant. Hence, the enzyme kinetics became pseudo-first-order. In enzymology, one also measures the kinetics of the substrate which is of course intimately related to the dynamics of the enzyme turnover. If the changes in the substrate (and product) concentration are significant, then we shall ask the following probabilistic question: 'What is the probability of the enzyme being in conformation C and the number of substrate molecules being n at the time t?' In other words, we are interested in the joint probability distribution $P_X(n,t)$ where X can be 1(E), 2(ES), or 3(EP). From the joint perspective of the substrate and the enzyme, (S,X), the reaction can be represented schematically as

$$\begin{array}{l}(N,1)\underset{k_{-1}}{\overset{Nk_1}{\rightleftharpoons}}(N-1,2)\underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}}(N-1,3)\underset{k_{-3}}{\overset{k_3}{\rightleftharpoons}}\cdots\underset{(N-n)k_{-3}}{\overset{k_3}{\rightleftharpoons}}\\ (n+1,1)\underset{k_{-1}}{\overset{(n+1)k_1}{\rightleftharpoons}}(n,2)\underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}}(n,3)\underset{(N-n)k_{-3}}{\overset{k_3}{\rightleftharpoons}}\cdots\underset{(N-1)k_{-3}}{\overset{k_3}{\rightleftharpoons}}\\ (1,1)\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}}(0,2)\underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}}(0,3)\underset{Nk_{-3}}{\overset{k_3}{\rightleftharpoons}}(0,1)\end{array} \tag{12}$$

where N is the total number of substrate and product molecules. The substrate kinetics, therefore, behave as a one-dimensional biased random walk. The bias of the random walk arises from the systematic transformation of substrate into product. If we start with N substrate and 0 product molecules, the 'walker' begins at the far left at time zero. The number of substrate molecules, n, is a

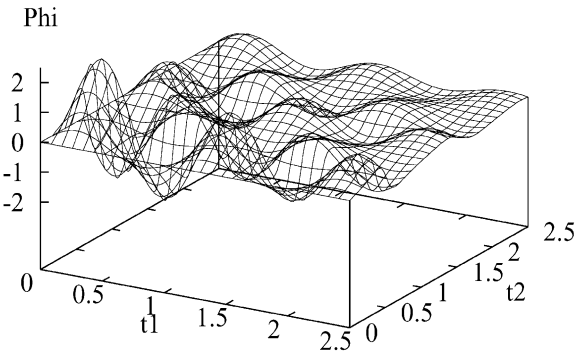

Fig. 1. An example of the ϕ function with $\langle f\rangle=0.5$, $\lambda_{1,2}=-\mu+i\upsilon=-0.5+6i$, and $\phi=-2e^{-\mu(t_1+t_2)}\sin(\upsilon t_1)\sin(\upsilon t_2)/[1+e^{-\mu t_2}\cos(\upsilon t_2)]$. Note that $-\phi$ are being plotted, and $\phi=0$ at the origin.

random variable, and the function $P(n,t)=\sum_{i=1}^{3}P_i(n,t)$ is the probability of having n substrate molecules at time t. Obviously, the fraction of unreacted substrate $\xi=n/N$ is also a random variables and ξ and n are simply related. ξ is known as the 'survival probability'[1] [12,18,19]. In general, the expectation of ξ, as a function of time t, cannot be expressed in terms of simple exponentials for Michaelis–Menten kinetics.

If we identify probabilities with concentrations, it is straightforward to write the kinetic equations for $P_i(n,t)$ according to Eq. (12):

$$\begin{aligned}\frac{dP_2(n,t)}{dt}&=-(k_{-1}+k_2)P_2(n)+k_{-2}P_3(n)\\&\quad+(n+1)k_1P_1(n+1)\\\frac{dP_3(n,t)}{dt}&=k_2P_2(n)-(k_{-2}+k_3)P_3(n)\\&\quad+(N-n)k_{-3}P_1(n)\\\frac{dP_1(n,t)}{dt}&=k_{-1}P_2(n-1)+k_3P_3(n)\\&\quad-\left[nk_1+(N-n)k_{-3}\right]P_1(n)\end{aligned}\tag{13}$$

The equilibrium distribution for $P_i(n)$ is readily obtained (see Appendix C). Furthermore for large N, asymptotic analysis [31] yields a diffusion equation for the survival probability (Qian and Elson, manuscript in preparation):

$$\frac{\partial}{\partial t}P(\xi,t)=\frac{1}{N}\frac{\partial}{\partial\xi}\left[J(\xi)P(\xi,t)\right]+\frac{1}{2N^2}\frac{\partial^2}{\partial\xi^2}\times\left[D(\xi)P(\xi,t)\right]\tag{14}$$

where

$$J(\xi)=Nk_1\xi P_1^{SS}(\xi)-k_{-1}P_2^{SS}(\xi)=\frac{Nk_1k_2k_3\xi-Nk_{-1}k_{-2}k_{-3}(1-\xi)}{Nk_1(k_2+k_{-2}+k_3)\xi+k_2k_3+k_3k_{-1}+k_{-1}k_{-2}+Nk_{-3}(k_{-1}+k_{-2}+k_2)(1-\xi)}$$

and

$$D(\xi)=Nk_1\xi P_1^{SS}(\xi)+k_{-1}P_2^{SS}(\xi)$$

$P(\xi,t)$ is the probability of the (unreacted) substrate fraction being ξ, thus the product being $1-\xi$, at time t. Neglecting the $1/N^2$ term, Eq. (14) can be shown to be mathematically equivalent to the non-linear ordinary differential equation [32]:

$$d\xi/dt=-J(\xi)/N\tag{15}$$

which is precisely what one obtains from deterministic steady-state Michaelis–Menten kinetics [15,22]. The non-linearity stems from the bimolecular nature of enzyme kinetics. In general with non-constant $J(\xi)$, ξ is not a simple exponential.

Eq. (13) and its asymptotic representation Eq. (14), as the stochastic version of the Michaelis–Menten model, constitute a generalization of the deterministic kinetics. The deterministic kinetics are recovered from Eq. (14) at large N in the form of Eq. (15) that agrees exactly with the standard deterministic enzyme kinetics for this system based on the law of mass action. The significance of the random walk-like Eq. (13) and diffusion-like Eq. (14), however, is that they provide a framework for calculating fluctuations and correlation function for small N as well large N. As we have shown, the standard deterministic theory is applicable only to systems with very large N.

3.1. A relation between cyclic enzyme dynamics and substrate kinetics

While an enzyme undergoes stochastic transitions among its conformational states (E, ES, EP), the number of product molecules continues to increase on average. Every time a cycle $E\rightarrow ES\rightarrow EP\rightarrow E$ is completed, a substrate is converted to a product. The making of a product is associated with each $ES\rightarrow EP$ transition. Let us again consider the situation in which $[S]$ and $[P]$ are constant, and $\hat{k}_1=k_1[S]$, $\hat{k}_{-3}=k_{-3}[P]$ are pseudo-first order

[1] We see that calling ξ a probability can be misleading and confusing, especially when one deals with function $P(\xi,t)$, which is the probability for the fraction of unreacted substrate at time t. Calling ξ the fraction of unreacted substrate could obviate the possible confusion.

rate constants. Then the addition and reduction of one product molecule can be represented by [19]

$$(-)\overset{k_{-2}}{\leftarrow}EP\underset{\hat{k}_{-3}}{\overset{k_3}{\rightleftharpoons}}E\underset{k_{-1}}{\overset{\hat{k}_1}{\rightleftharpoons}}ES\overset{k_2}{\rightarrow}(+) \tag{16}$$

According to this scheme, after each forward turnover $(+)$, the enzyme is in *EP*, state 3. Hence, the probability of having another turnover forward after time t is $P(+,t|+)=P(+,t|3)$ and having a turnover backward after time t is $P(-,t|+)=P(-,t|3)$. Similarly, after each backward turnover $(-)$, the enzyme is in *ES*, state 2. Hence, the probability $P(+,t|-)=P(+,t|2)$ and $P(-,t|+)=P(-,t|2)$. These conditional probabilities are readily calculated according to the formulae given by Qian [19] based on the eigenvalues of the matrix

$$\begin{pmatrix} -k_1-k_{-3} & k_{-1} & k_3 \\ k_1 & -k_{-1}-k_2 & 0 \\ k_{-3} & 0 & -k_{-2}-k_3 \end{pmatrix} \tag{17}$$

Note, this matrix and the one in Eq. (2) are related by connecting the transitions on the left and on the right of Eq. (16): the enzyme dynamics corresponds to a periodic boundary condition while the product waiting time correspond to an absorbing boundary condition for the same reaction in Eq. (16). The oscillatory kinetics in the cyclic reaction is intimately related to a pseudo-periodicity in the product formation [27].[2] Fig. 2 shows the distribution for the waiting time of the arrival of next product molecule, calculated according to Eq. (16). It is seen that the sharper distribution, which means the products arrive with a more regular time interval, corresponds to the oscillatory kinetics. Therefore, the irreversibility, which is necessary for the oscillatory kinetics, can make even a single enzyme behave more as a deterministic machine. See Appendix D for a simple illustration of the duality between mean arrival time and the steady-state flux.

[2] The waiting time for the next product and the cycle time of enzyme are not exactly the same since from the enzyme conformation point of view, converting a product to a substrate molecule also completes a cycle.

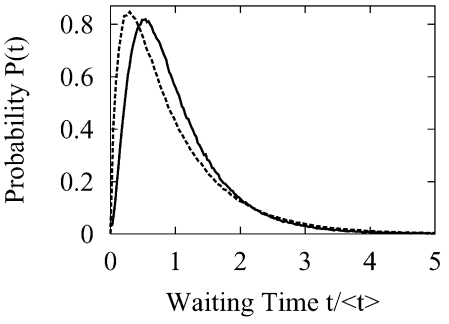

Fig. 2. Comparison between two distributions for the waiting time of the next arrival product catalyzed by the single enzyme according to stochastic Michaelis–Menten kinetics, Eq. (1). For the solid line: $\hat{k}_1=1$, $k_2=2$, $k_3=3$, $k_{-1}=0.3$, $k_{-2}=0.2$, $\hat{k}_{-3}=0.1$ and $\lambda_{1,2}=-3.3\pm1.23i$ according to Eq. (3). For the dashed line: $\hat{k}_1=1$, $k_2=2$, $k_3=3$, $k_{-1}=0.9$, $k_{-2}=1.8$, $\hat{k}_{-3}=2.7$ and $\lambda_{1,2}=-4.06,\ -7.34$. Both distributions are normalized by their respective mean waiting time $\langle t\rangle$, 1.98 and 1.18, respectively. The relative standard deviations, $\sqrt{\langle(\Delta t)^2\rangle}/\langle t\rangle$, for the two distributions are significantly different, 0.73 and 2.1, respectively.

4. Discussion

The stochastic treatment of chemical reactions was initiated by Kramers in 1940 [33]. Since then, his approach to chemical reaction rates has grown into a main focus of theoretical chemistry [34]. The Kramers' theory, however, addresses mainly unimolecular reactions and bimolecular collisions in aqueous solution. It does not address the relation between the stochastic molecular reactions on the level of individual molecules and the behavior of an ensemble of molecules, when there are molecular interactions. Also in 1940, Delbrück [35] lucidly posed the essential problem of the stochastic treatment of biochemical reactions and identified the non-linear and non-equilibrium characteristics of such reactions and the mathematical analogy to a random walk. Since this seminal contribution, much work has been devoted to this subject [36–39] and many results are summarized in excellent books [31,40–42]. Only in the past decade, however, have experimental

techniques been improved sufficiently to permit the study of single-molecule enzymology (for a recent review, see [6]). This progress on the experimental front calls for the revival of the earlier theoretical works and their integration with the existing literature on single-channel recording [11,43]. More importantly, single-molecule enzymology requires a stochastic reasoning in its data interpretation and kinetic modeling.

In summary, we have provided a basis in stochastic modeling for quantitative kinetic analysis in single-molecule enzymology. Measurements on stochastic protein conformational dynamics and the stochastic substrate/product kinetics are interpreted within a single mathematical treatment. Oscillatory kinetics is shown to be expected in the stochastic, Markovian models for non-linear biochemical kinetics. This analysis suggests an alternative but complementary mechanism to the fluctuating enzyme model for analyzing complex kinetics of single-molecule enzymology.

Appendix A: Higher-order correlation functions for two-state kinetics

To facilitate comparison with the three-state model, related formulae for two-state kinetics $A \underset{\beta}{\overset{\alpha}{\rightleftharpoons}} B$ are collected in this appendix. We assume the fluorescence signal for the two states are f_A and f_B, respectively. Then $\langle f\rangle = \frac{f_A + Kf_B}{1+K}$ where $K=\alpha/\beta$. Time correlation function:

$$C_2(\tau) = \frac{(f_A - f_B)^2 K}{(1+K)^2} e^{-\lambda\tau} \tag{18}$$

where $\lambda = \alpha + \beta$, and

$$C_3(\tau_1,\tau_2) = \frac{(f_A - f_B)^3 K(K-1)}{(1+K)^3} e^{-\lambda(\tau_1+\tau_2)} \tag{19}$$

The C_3 is related to the C_2 after normalization:

$$\frac{c_3(\tau_1,\tau_2)}{c_2(\tau_2)} - c_2(\tau_1) = 0 \tag{20}$$

where $c_2(\tau) = C_2(\tau)/C_2(0)$ and $c_3(\tau_1,\tau_2) = C_3(\tau_1,\tau_2)/C_3(0,0)$. Note the similarity between this expression and Eq. (11), even though the normalization factors used for G's are different from those for c's. Eq. (20) requires no assumption on f_A and f_B; hence, it is a general result for any two-state process with a single exponential.

With sufficient data, even higher order correlation functions can be computed:

$$\begin{aligned} C_4(\tau_1\tau_2\tau_3) &= \langle \Delta f(0)\Delta f(\tau_1)\Delta f(\tau_1 + \tau_2)\Delta f(\tau_1+\tau_2+\tau_3)\rangle \\ &= \left(\frac{f_A - f_B}{1+K}\right)^4 \{K(1-K)^2 e^{-\lambda(\tau_1+\tau_2+\tau_3)} + K^2 e^{-\lambda(\tau_1+\tau_3)}\} \end{aligned} \tag{21}$$

A special case of C_4 with $\tau_3 = \tau_1$ gives

$$C_4(\tau_1\tau_2\tau_1) = \left(\frac{f_A - f_B}{1+K}\right)^4 \{K(1-K)^2 e^{-\lambda(2\tau_1+\tau_2)} + K^2 e^{-2\lambda\tau_1}\}$$

which has been derived by Schenter et al. [21] if one substitutes K with $p_B/(1-p_B)$.

Appendix B: Real eigenvalues in equilibrium

Let $\hat{k}_1 = k_1[S]$ and $\hat{k}_{-3} = k_{-3}[P]$ be the pseudo-first order rate constants. If $\hat{k}_1 k_2 k_3 = k_{-1}k_{-2}\hat{k}_{-3}$, then

$$\begin{aligned} \Delta &= \left[(\hat{k}_1 - k_{-2}) + (k_{-1} - k_3) + (\hat{k}_{-3} - k_2)\right]^2 - 4(k_{-1} - k_3)(\hat{k}_{-3} - k_2) \\ &= \left[k_{-2}\left(\frac{k_{-1}\hat{k}_{-3}}{k_2 k_3} - 1\right) + (k_{-1} - k_3) + (\hat{k}_{-3} - k_2)\right]^2 - 4(k_{-1} - k_3)(\hat{k}_{-3} - k_2) \\ &\geq \begin{cases} [k_{-2}(\frac{k_{-1}\hat{k}_{-3}}{k_2 k_3} - 1) + (k_{-1} - k_3) + (\hat{k}_{-3} - k_2)]^2 & \text{if} \quad (k_{-1} - k_3)(\hat{k}_{-3} - k_2) \leq 0 \\ [(k_{-1} - k_3) + (\hat{k}_{-3} - k_2)]^2 - 4(k_{-1} - k_3)(\hat{k}_{-3} - k_2) & \text{if} \quad (k_{-1} - k_3)(\hat{k}_{-3} - k_2) \geq 0 \end{cases} \\ &\geq 0 \end{aligned}$$

Appendix C: Distribution and moments of substrate fluctuation

The equilibrium distribution of Eq. (13), a variance of the birth-and-death process, can be readily obtained [Taylor and Karlin, [25]]:

$$P_1(n)=\frac{Q^{-1}N!K^{N-n}}{(N-n)!n!},$$
$$P2(n)=\frac{Q^{-1}N!K^{N-n-1}}{(N-n-1)!n!}\frac{k_1}{k_{-1}},$$
$$P_3(n)=\frac{Q^{-1}N!K^{N-n-1}}{(N-n-1)!n!}\frac{k_1k_2}{k_{-1}k_{-2}}$$

where $K=k_1k_2k_3/(k_{-1}k_{-2}k_{-3})$ and

$$Q(K)=(1+K)^{N-1}\left(1+K+\frac{NKk_{-3}}{k_3}+\frac{NKk_{-2}k_{-3}}{k_2k_3}\right)$$

Furthermore, the mean and variance of the number of substrate molecules in equilibrium can be computed from the function Q:

$$\langle n\rangle=N-\frac{\partial\ln Q(K)}{\partial\ln K}=\frac{N+K}{1+K}-\frac{K+\frac{Nk_1}{k_{-1}}+\frac{Nk_1k_2}{k_{-1}k_{-2}}}{1+K+\frac{Nk_1}{k_{-1}}+\frac{Nk_1k_2}{k_{-1}k_{-2}}}$$
$$\left\langle(\Delta n)^2\right\rangle=\frac{\partial\ln Q(K)}{\partial(\ln K)^2}=\frac{NK}{(1+K)^2}+\frac{K+\frac{Nk_1}{k_{-1}}+\frac{Nk_1k_2}{k_{-1}k_{-2}}}{\left(1+K+\frac{Nk_1}{k_{-1}}+\frac{Nk_1k_2}{k_{-1}k_{-2}}\right)^2}$$

For large N, the results reduce to $N/(1+K)$ and $NK/(1+K)^2$, as expected from a simple binomial distribution.

Appendix D: Duality between mean arrival time and steady-state cyclic flux

There is an intimate relation between the cyclic reaction of a single enzyme and the substrate turnover kinetics. Let us consider the simple kinetics

$$A\underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}}B\overset{k_2}{\rightarrow}C$$

for a single molecule starting in state A. The probability density function for the time of the molecule arriving C can be easily computed:

$$P_c(t)=\frac{k_1k_2}{\gamma_2-\gamma_2}[e^{-\gamma_1t}-e^{-\gamma_2t}]$$

in which

$$\gamma_{1,2}=\frac{1}{2}\times\left[(k_1+k_{-1}+k_2)\pm\sqrt{(k_1+k_{-1}+k_2)^2-4k_1k_2}\right].$$

Therefore, the mean arrival time is

$$\langle t\rangle=\int_0^\infty tP_C(t)dt=\frac{k_1k_2}{\gamma_2-\gamma_1}\times\left(\frac{1}{\gamma_1^2}-\frac{1}{\gamma_2^2}\right)=\frac{k_1+k_{-1}+k_2}{k_1k_2}$$

and the relative variance is $1-\frac{2k_1k_2}{(k_1+k_{-1}+k_2)^2}$. The dual of this problem is to compute the steady-state flux under a periodic boundary condition. That is, we assume state C is equivalent to state A. We, therefore, have a cyclic kinetics $A\rightleftharpoons B\rightarrow A$. The steady-state flux is readily obtained: $J=\frac{k_1k_2}{k_1+k_{-1}+k_2}$ which is precisely $\frac{1}{\langle t\rangle}$. In other words, J, the number of turnover per unit time, matches the mean turnover (arrival) time.

References

[1] J.A. Schellman, Solvent denaturation, Biopolymers 17 (1978) 1305–1322.
[2] J.A. Schellman, Fluctuation and linkage relations in macromolecular solution, Biopolymers 29 (1990) 215–224.
[3] C.-L. Bai, C. Wang, X.S. Xie, P.G. Wolynes, Single molecule physics and chemistry, Proc. Natl. Acad. Sci. USA 96 (1999) 11075–11076.

[4] E.L. Elson, W.W. Webb, Fluorescence correlation spectroscopy: a new biophysical probe based on occupation number fluctuations, Ann. Rev. Biophys. Bioeng. 4 (1975) 311–334.

[5] S. Nie, D.T. Chiu, R.N. Zare, Probing individual molecules with confocal fluorescence microscopy, Science 266 (1994) 1018–1021.

[6] X.S. Xie, J.K. Trautman, Optical studies of single-molecule at room temperature, Ann. Rev. Phys. Chem. 49 (1998) 441–480.

[7] L. Edman, R. Rigler, Memory landscapes of single-enzyme molecules, Proc. Natl. Acad. Sci. USA 97 (2000) 8266–8271.

[8] A.A. Heikal, S.T. Hess, G.S. Baird, R.Y. Tsien, W.W. Webb, Molecular spectroscopy and dynamics of intrinsically fluorescent proteins: Coral red (dsRed) and yellow (Citrine), Proc. Natl. Acad. Sci. USA 97 (2000) 11996–12001.

[9] X.S. Xie, H.P. Lu, Single-molecule enzymology, J. Biol. Chem. 274 (1999) 15967–15970.

[10] B. Sakmann, E. Neher, Single-Channel Recording, 2nd Edition, Plenum Press, New York, 1995.

[11] F.G. Ball, J.A. Rice, Stochastic models for ion channels: introduction and bibliography, Math. Biosci. 112 (1992) 189–206.

[12] N. Agmon, J.J. Hopfield, Transient kinetics of chemical reactions with bounded diffusion perpendicular to the reaction coordinate: intramolecular processes with slow conformational changes, J. Chem. Phys. 78 (1983) 6947–6959.

[13] H. Qian, Equations for stochastic macromolecular mechanics of single proteins: equilibrium fluctuations, transient kinetics and non-equilibrium steady-state, J. Phys. Chem. 106 (2002) 2065–2073.

[14] H.P. Lu, L.-Y. Xun, X.S. Xie, Single-molecule enzymatic dynamics, Science 282 (1998) 1877–1882.

[15] G.G. Hammes, Enzyme Catalysis and Regulation, 2nd Edition, Academic Press, New York, 1982.

[16] A. Arkin, P. Shen, J. Ross, A test case of correlation metric construction of a reaction pathway from measurements, Science 29 (1997) 1275–1279.

[17] K. Svoboda, P.P. Mitra, S.M. Block, Fluctuation analysis of motor protein movement and single enzyme kinetics, Proc. Natl. Acad. Sci. USA 91 (1994) 11782–11786.

[18] J. Wang, P.G. Wolynes, Intermittency of single molecular reaction dynamics in fluctuating environments, Phys. Rev. Lett. 74 (1995) 4317–4320.

[19] H. Qian, A simple theory of motor protein kinetics and energetics II, Biophys. Chem. 83 (2000) 35–43.

[20] G.P. Li, H. Qian, Kinetic timing: a novel mechanism for improving the accuracy of GTPase timers in endosome fusion and other biological processes, Traffic 3 (2002) 249–255.

[21] G.K. Schenter, H.P. Lu, X.S. Xie, Statistical analysis and theoretical models of single-molecule enzymatic dynamics, J. Phys. Chem. 103 (1999) 10477–10488.

[22] I.H. Segel, Enzyme Kinetics, John Wiley & Sons, New York, 1993.

[23] A. Nitzan, J. Ross, Oscillations, multiple steady-states, and instabilities in illuminated systems, J. Chem. Phys. 59 (1973) 241–250.

[24] T.L. Hill, Free Energy Transduction and Biochemical Cycle Kinetics, Springer-Verlag, New York, 1989.

[25] H.M. Taylor, S. Karlin, An Introduction to Stochastic Modeling, 3rd Edition, Academic Press, New York, 1998.

[26] J.J. Hopfield, Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity, Proc. Natl. Acad. Sci. USA 71 (1974) 4135–4139.

[27] H. Qian, M. Qian, Pumped biochemical reactions, nonequilibrium circulation, and stochastic resonance, Phys. Rev. Lett. 84 (2000) 2271–2274.

[28] H.-S. Hahn, A. Nitzan, P. Ortoleva, J. Ross, Threshold excitations, relaxation oscillations, and effect of noise in an enzyme reaction, Proc. Natl. Acad. Sci. USA 71 (1974) 4067–4071.

[29] A.G. Palmer, N.L. Thompson, High-order fluorescence fluctuation analysis of model protein clusters, Proc. Natl. Acad. Sci. USA 52 (1989) 6148–6152.

[30] H. Qian, On the statistics of fluorescence correlation speetroscopy, Biophys. Chem. 38 (1990) 49–57.

[31] J. Keizer, Statistical Thermodynamics of non-equilibrium Processes, Springer-Verlag, New York, 1987.

[32] R. Haberman, Mathematical Models: Mechanical Vibrations, Population Dynamics, and Traffic Flow, SIAM Press, Philadelphia, 1998.

[33] H.A. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, Physica 7 (1940) 284–304.

[34] P. Hänggi, P. Talkner, M. Borkovec, Reaction-rate theory: fifty years after Kramers, Rev. Mod. Phys. 62 (1990) 251–341.

[35] M. Delbrück, Statistical fluctuations in autocatalytic reactions, J. Chem. Phys. 8 (1940) 120–124.

[36] C.J. Jachimowski, D.A. McQuarrie, M.E. Russell, A stochastic approach to enzyme-substrate reactions, Biochemistry 3 (1964) 1732–1736.

[37] I.G. Darvey, B.W. Ninham, P.J. Staff, Stochastic models for second-order chemical reaction kinetics: the equilibrium state, J. Chem. Phys. 45 (1966) 2145–2155.

[38] I.G. Darvey, P.J. Staff, The application of the theory of Markov processes to the reversible one substrate-one intermediate-one product enzymic mechanism, J. Theor. Biol. 14 (1967) 157–172.

[39] P.J. Staff, Approximation method for equilibrium distributions in second-order chemical reaction kinetics, J. Chem. Phys. 46 (1967) 2209–2212.

[40] D.A. McQuarrie, Stochastic Approach to Chemical Kinetics, Methuen, London, 1968.
[41] D.T. Gillespie, Markov Processes: An Introduction for Physical Scientists, Academic Press, San Diego, 1992.
[42] N.G. van Kampen, Stochastic Processes in Physics and Chemistry, Elsevier, Amsterdam, 1992.
[43] F. Conti, Noise analysis and single-channel recordings, Curr. Top. Membr. Transp. 22 (1984) 371–405.