

Evaluation of rate constants for conformational transitions using single-molecule fluorescence spectroscopy

Alexander M. Berezhkovskii^{a,1}, Marián Bogoñá^a, George H. Weiss^{a,*}

^a *Mathematical & Statistical Computer Lab., Division of Computer Research and Technology, Center for Information Technology, National Institutes of Health, Bldg. 12 A Room 2007, Bethesda, MD 20892, USA*

Received 2 October 2000; in final form 29 November 2000

Abstract

We develop formulae for translating single-molecule fluorescence spectroscopic data into estimates of the rate constants for slow conformational transitions between two states with different lifetimes of the fluorescent probe. These rate constants cannot be determined separately from the bulk experiment. © 2001 Published by Elsevier Science B.V.

There is by now a large set of techniques grouped under the rubric of single-molecule spectroscopy (SMS), that enable an experimenter to monitor dynamic behavior of single molecules in solution, [1–5]. The advantage of this class of methods is that it provides information otherwise unavailable from bulk experiments. In the present note we show how the rate constants for transitions between two conformational states can be evaluated using SM fluorescence spectroscopy. These rate constants cannot be determined by bulk experiments which can only provide their ratio.

Consider a molecule which randomly interconverts between two conformational states according to the kinetic scheme



where the k 's are rate constants. A fluorescent probe attached to the molecule has different fluorescence lifetimes in the two states, τ_1 and τ_2 , respectively. When interconversion occurs much more slowly than the fluorescence, the decay of the fluorescence intensity measured in the bulk experiment is the biexponential

$$\frac{I_{\text{bulk}}(t)}{I_{\text{bulk}}(0)} = P_{\text{eq}}(1)e^{-t/\tau_1} + P_{\text{eq}}(2)e^{-t/\tau_2}, \quad (2)$$

where $P_{\text{eq}}(i)$ is the equilibrium population of state i ($i = 1, 2$),

$$P_{\text{eq}}(1) = k_2/k, \quad P_{\text{eq}}(2) = 1 - P_{\text{eq}}(1) = k_1/k \quad (3)$$

and $k = k_1 + k_2$.

Thus, only the ratio k_1/k_2 can be found from the bulk experiment, rather than the separate rate constants.

The two rate constants can be determined from the SM fluorescence experiments described in [6–9]. In these experiments a randomly chosen molecule is repeatedly excited by a train of laser pulses for a period of time T . The fluorescence

* Corresponding author. Fax: +1-301-402-4544.

E-mail address: ghw@helix.nih.gov (G.H. Weiss).

¹ Permanent address: Karpov Institute of Physical Chemistry, Ul. Vorontsovo Pole 10, 103064, Moscow K-64, Russian Federation.

decay kinetics is again found to be biexponential with the same lifetimes τ_1 and τ_2 as in the bulk experiment

$$\frac{I_{\text{SM}}(t)}{I_{\text{SM}}(0)} = xe^{-t/\tau_1} + (1-x)e^{-t/\tau_2}, \quad (4)$$

where x is the amplitude found by curve-fitting to the data. This amplitude can be interpreted as the fraction of time, out of the total time T , that the molecule has spent in state 1, [10,11]. Since the time spent in state 1 is random the amplitude, x , is also random.

After the experiment is repeated several times, one can estimate the probability density for the amplitude which depends on the duration of the measurement, T , as well as the rate constants. This density will be denoted by $p(x|T)$. The theory of such SM fluorescence experiments has been developed in [10–12] under the assumption that measurement times for each molecule are equal. A generalization of the theory to molecules with an arbitrary number of conformational states with different fluorescence lifetimes of the probe is given in [13], again assuming that the measurement times are identical.

The bulk result in Eq. (2) can be recovered from Eq. (4) in the limit $T \rightarrow \infty$ since, due to ergodicity,

$$\lim_{T \rightarrow \infty} p(x|T) = \delta(x - P_{\text{eq}}(1)). \quad (5)$$

When T is finite $p(x|T)$ is no longer a delta function, and contains information which can be used to find the k_i . Since there are two rate constants two independent equations are required to determine their values. Here we choose these to be the equations for the experimentally accessible mean value of x and the associated variance.

The average value of the amplitude is independent of T and is

$$\langle x \rangle = \int_0^1 xp(x|T) dx = P_{\text{eq}}(1) = k_2/k. \quad (6)$$

However, the variance does depend on the measurement time. Its large- T behavior is

$$\sigma^2(T) = \frac{2P_{\text{eq}}(1)P_{\text{eq}}(2)}{kT} = \frac{2k_1k_2}{k^3T}. \quad (7)$$

Using Eqs. (6) and (7) we obtain expressions for the rate constants as

$$k_1 = \frac{2\langle x \rangle(1 - \langle x \rangle)^2}{T\sigma^2(T)}, \quad k_2 = \frac{2\langle x \rangle^2(1 - \langle x \rangle)}{T\sigma^2(T)}. \quad (8)$$

Thus, to find the rate constants one has to estimate $\langle x \rangle$ and $\sigma^2(T)$ from experimental data to calculate k_1 and k_2 from this last equation.

The theory summarized so far assumes that the experimental output is an infinite set of the amplitudes, $\{x_j\}$, $j = 1, 2, \dots$ and that the durations of all of the measurements are equal. In fact, the number of single molecules sampled in any experiment is finite, $j = 1, 2, \dots, M$. Therefore, only estimates of the average and variance are available. These will be denoted by \bar{x} and $\bar{\sigma}^2(T)$, and are

$$\bar{x} = \frac{1}{M} \sum_{j=1}^M x_j, \quad \bar{\sigma}^2(T) = \frac{1}{M-1} \sum_{j=1}^M (x_j - \bar{x})^2. \quad (9)$$

When these two estimates are found from the data they are inserted into Eq. (8) to provide estimates of the rate constants, \bar{k}_1 and \bar{k}_2 , provided that all of the durations are equal and the time T is sufficiently large. A corresponding inequality will be given later in the paper.

The estimates discussed above presumes that the measurement times for all of the molecules are equal. In practice these times can vary from one molecule to another. The experimental output consists of the set $\{x_1, T_1; x_2, T_2; \dots, x_M, T_M\}$ where x_j is the amplitude obtained in the j 'th measurement which has the duration T_j . It will be assumed that the T_j are large in the sense described below. It will be shown that, even when the T_j differ, it is possible to estimate the rate constants making use of Eq. (8). For this purpose we need to have estimates for $\langle x \rangle$ and the product $D = T\sigma^2(T)$ appearing in the denominator. The parameter \bar{x} in Eq. (9) is to be substituted for $\langle x \rangle$ in the numerator of Eq. (8). The denominator in Eq. (8), to lowest order in $1/M$, is replaced by

$$\bar{D} = \frac{1}{M} \sum_{j=1}^M T_j(x_j - \bar{x})^2. \quad (10)$$

To prove the correctness of this formula, as well as to provide the correction terms we note that the x_j are independent random variables described by the probability density $p(x|T_j)$. All of the x_j have the same average value $\langle x \rangle = P_{\text{eq}}(1) = k_2/k$. The variance, at sufficiently large T_j , is given by $\sigma^2(T_j) = 2k_1k_2/(k^3T_j)$. We would like to construct an exact formula for \bar{D} , rather than the approximation in Eq. (10), which has the property

$$\begin{aligned} \langle \bar{D} \rangle &= \int_0^1 \cdots \int_0^1 \bar{D}(x_1, \dots, x_M) \prod_{j=1}^M p(x_j|T_j) dx_j \\ &= \frac{1}{M} \sum_{j=1}^M T_j \sigma^2(T_j). \end{aligned} \quad (11)$$

When all of the T_j are large the value of $\langle \bar{D} \rangle$ is equal to $2k_1k_2/k^3$ as in the case of equal measurement times.

For the analysis to follow we will need the average measurement duration

$$\bar{T} = \frac{1}{M} \sum_{j=1}^M T_j. \quad (12)$$

We look for an estimator \bar{D} that is linear in the T_j and reduces to the $T\sigma^2$ in Eq. (9) when all of the T_j are equal to T . Thus, we are led to the form

$$\bar{D} = \frac{1}{M-1} \sum_{j=1}^M T_j (x_j - \bar{x})^2 + \sum_{j=1}^M \beta_j (T_j - \bar{T}). \quad (13)$$

Specifying the value of the β_j 's in this equation requires some slightly more complicated considerations. We can invoke the requirement in Eq. (11) to fix the value of this variable in terms of the x_j . A tedious, but straightforward algebraic calculation leads to

$$\beta_j = \frac{x_j^2}{M(M-1)}. \quad (14)$$

Thus the final expression for \bar{D} is

$$\bar{D} = \frac{1}{M-1} \left\{ \sum_{j=1}^M T_j (x_j - \bar{x})^2 + \frac{1}{M} \sum_{j=1}^M (T_j - \bar{T}) x_j^2 \right\}, \quad (15)$$

which is expressed in terms of the observable parameters x_j and T_j . When $M \gg 1$ this expression

reduces to Eq. (10) which is one of the main results of this Letter. Values of M that are of the order of 100 or more will generally suffice to validate the reduction.

Finally, we clarify what is meant by the large- T limit which allows us to use the approximate expression for the variance, $\sigma^2(T) = 2k_1k_2/(k^3T)$. The theory in [12] can be used to provide an expression for $\sigma^2(T)$ that is exact for all T . This is

$$\sigma^2(T) = \frac{2k_1k_2}{k^3T} \left[1 - \frac{1}{kT} (1 - e^{-kT}) \right]. \quad (16)$$

From this it follows that a value of $kT \geq 10$ is sufficient to comfortably validate an approximation to $T\sigma^2(T)$ as being independent of T . When $kT = 10$ (or even 20) the probability density $p(x|T)$ for a given T can differ noticeably from a Gaussian depending on the values of the rate constants, as seen in illustrative examples in [12]. Nonetheless Eq. (16) is exact.

To summarize, the main results of this note are the expressions for the rate constants given in Eq. (8) together with formulae for \bar{x} and \bar{D} given in Eqs. (9) and (10), respectively. These formulae allow one to estimate the rate constants for slow conformational transitions between two states with different fluorescence lifetimes of the probe using single-molecule fluorescence spectroscopy. These rate constants cannot be determined using a bulk experiment alone.

References

- [1] W.E. Moerner (Ed.), *Acc. Chem. Res.* 29 (1966) 561 (a special issue entirely devoted to SMS).
- [2] S. Nie, R.N. Zare, *Ann. Rev. Biophys. Biomol. Struct.* 26 (1997) 567.
- [3] X.S. Xie, J.K. Trautman, *Ann. Rev. Phys. Chem.* 49 (1998) 441.
- [4] *Science* 283 (5408) (1999) (contains a collection of articles on SMS).
- [5] *Chem. Phys.* 247 (1999) 1 (a special issue entirely devoted to SMS).
- [6] L. Edman, Ü. Mets, R. Rigler, *Exp. Tech. Phys.* 41 (1995) 157.
- [7] L. Edman, Ü. Mets, R. Rigler, *Proc. Natl. Acad. Sci.* 93 (1996) 6710.
- [8] S. Wennmalm, L. Edman, R. Rigler, *Proc. Natl. Acad. Sci.* 94 (1997) 6710.

- [9] Y. Jia, A. Sytnik, L. Li, S. Vladimirov, B.S. Cooperman, R.M. Hochstrasser, *Proc. Natl. Acad. Sci.* 94 (1997) 7932.
- [10] E. Geva, J.L. Skinner, *Chem. Phys. Lett.* 288 (1998) 225.
- [11] L. Edman, S. Wennmalm, F. Tamsen, R. Rigler, *Chem. Phys. Lett.* 292 (1998) 15.
- [12] A.M. Berezhkovskii, A. Szabo, G.H. Weiss, *J. Chem. Phys.* 110 (1999) 9145.
- [13] A.M. Berezhkovskii, A. Szabo, G.H. Weiss, *Phys. Chem. B* 104 (2000) 3776.