Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary information

Fast single molecule FRET spectroscopy: theory and experiment

Hoi Sung Chung and Irina V. Gopich

I. Derivation of the average transition path time

In this section we consider the transition path time for a two-state system. The dynamics of a molecule are modeled by one-dimensional (1D) diffusion in a double-well potential $G(q)$, which depends on the reaction coordinate q (see Fig. S1). The potential has two minima corresponding to the folded and unfolded states and a barrier. We are interested in the time required to cross this barrier.

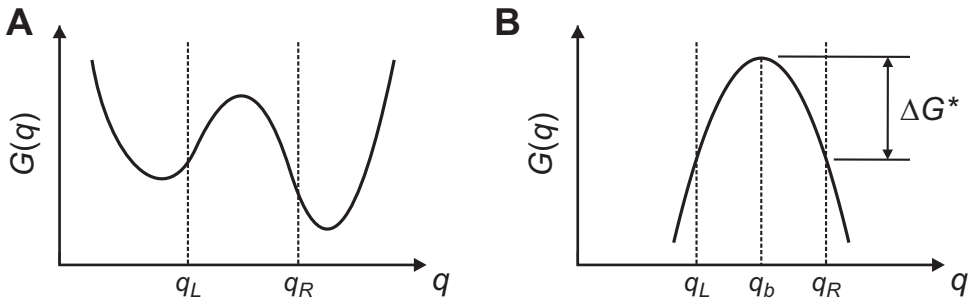

Figure S1. One-dimensional free energy profile along a reaction coordinate q. (A) A transition path from q_L to q_R is defined as a trajectory that crosses q_L and reaches q_R without re-crossing q_L. (B) A parabolic free energy barrier that is used to derive Eq. (18).

The probability density $p(q,t|q_0)$ that a molecule is at q at time t given it was at q_0 initially obeys the diffusion equation:

$$\frac{\partial}{\partial t}p(q,t\,|\,q_0) = D\frac{\partial}{\partial q}e^{-\beta G(q)}\frac{\partial}{\partial q}e^{\beta G(q)}p\,. \qquad (1)$$

For simplicity, we assume that the diffusion coefficient, D, does not depend on the coordinate. Consider a molecule, which is initially in the left well. When the barrier is high enough, the diffusing molecule spends long time in this well, and eventually passes the barrier when it reaches q_R in the right well. A transition path is defined as a trajectory that crosses a point q_L in the left well and reaches q_R without going back to q_L. The time required to reach a certain point (the first passage time) is usually studied by imposing the absorbing boundary at that point, and therefore we set $p(q_R,t) = 0$. To forbid the return to the left well, we set another absorbing boundary at $q = q_L$. Thus a molecule diffuses between q_L and q_R (see Fig. S1A) and disappears when it reaches the boundary at q_L or q_R. Initially, the molecule is assumed to be at $q = q_0$ between the two boundaries. We are looking for the mean time that the molecule spends in this

region on condition that it exits through the right boundary, q_R, $T_R(q_0)$. The average transition path time is the limit of $T_R(q_0)$ at $q_0 \to q_L$.

The quantity that describes the exit through the right boundary is the flux through q_R, $J_R(q_0, t)$, which depends on the starting point q_0. This flux is related to the gradient of the probability density $p(q_R,t|q_0)$:

$$J_R(q_0,t) = -De^{-\beta G(q)} \frac{\partial}{\partial q} e^{\beta G(q)} p(q,t \mid q_0)\Big|_{q=q_R} . \tag{2}$$

Since the probability to exit through q_R at a time later than t when the particle was initially at q_0 is $\int_t^\infty J_R(q_0,t')dt'$,[1] the distribution of the time required to reach the right boundary is $J_R(q_0,t) / \int_0^\infty J_R(q_0,t')dt'$. The mean time is the first moment of this distribution:

$$T_R(q_0) = \frac{\int_0^\infty t J_R(q_0,t)dt}{\int_0^\infty J_R(q_0,t)dt} . \tag{3}$$

Another quantity which will be used in this derivation is the splitting probability, $\pi_R(q_0)$, i.e., the probability to exit through q_R (rather than through q_L) starting at q_0:

$$\pi_R(q_0) = \int_0^\infty J_R(q_0,t')dt' . \tag{4}$$

The straightforward way to find the time $T_R(q_0)$ would be to solve the diffusion equation, Eq. (1), and to use the solution in Eq. (3). In general, the diffusion equation can be solved only numerically. However, it is possible to find the time for arbitrary profile $G(q)$ without actually solving the diffusion equation. To do so, we use the backward diffusion equation,[2] which involves the initial coordinate q_0:

$$\frac{\partial}{\partial t} p(q,t \mid q_0) = De^{\beta G(q_0)} \frac{\partial}{\partial q_0} e^{-\beta G(q_0)} \frac{\partial}{\partial q_0} p . \tag{5}$$

The backward equation allows one to get closed equations for $T_R(q_0)$ and $\pi_R(q_0)$. First, we get the diffusion equation for the flux by multiplying Eq. (5) by $-De^{-\beta G(q_R)} \frac{\partial}{\partial q_R} e^{\beta G(q_R)}$ according to Eq. (2):

$$\frac{\partial}{\partial t} J_R(q_0,t) = De^{\beta G(q_0)} \frac{\partial}{\partial q_0} e^{-\beta G(q_0)} \frac{\partial}{\partial q_0} J_R . \tag{6}$$

The equation for π_R (Eq. (7a)) is obtained by integrating both sides of Eq. (6) with respect to time t from 0 to ∞ according to Eq. (4) and using $\int_0^\infty \partial J_R / \partial t \, dt = J_R(q_0,\infty) - J_R(q_0,0) = 0$ because the flux is zero both initially and at long times. The equation for T_R is obtained by multiplying both sides of Eq. (6) by t and integrating with respect to t according to Eq. (3) using $\int_0^\infty t \partial J_R / \partial t \, dt = -\int_0^\infty J_R(q_0,t)dt = -\pi_R(q_0)$. Thus the equations for π_R and T_R are:[1]

$$De^{\beta G(q_0)}\frac{\partial}{\partial q_0}e^{-\beta G(q_0)}\frac{\partial}{\partial q_0}\pi_R(q_0)=0\,. \tag{7a}$$

$$De^{\beta G(q_0)}\frac{\partial}{\partial q_0}e^{-\beta G(q_0)}\frac{\partial}{\partial q_0}T_R(q_0)\pi_R(q_0)=-\pi_R(q_0)\,. \tag{7b}$$

The boundary conditions for these equations are straightforward. When the molecule is initially at the boundary, it immediately exits through the same boundary. Therefore, $\pi_R(q_R) = 1$, $\pi_R(q_L) = 0$, $\pi_R(q_L)T_R(q_L) = \pi_R(q_R)T_R(q_R) = 0$.

The above equations for $T_R(q_0)$ and $\pi_R(q_0)$ can be solved by integrating twice with respect to q_0. The integration constants are found using the boundary conditions. Solving the first equation, Eq. (7a), we get the splitting probability:

$$\pi_R(q_0)=\frac{\int_{q_L}^{q_0}\exp(\beta G(q))dq}{\int_{q_L}^{q_R}\exp(\beta G(q))dq}\,. \tag{8}$$

Integrating the second equation, Eq. (7b), we find

$$T_R(q_0)\pi_R(q_0)=-\int_{q_0}^{q_R}\left[C+\frac{1}{D}\int_q^{q_R}\pi_R(q')e^{-\beta G(q')}dq'\right]e^{\beta G(q)}dq\,, \tag{9}$$

where constant C is found later. This expression is further simplified by integrating by parts using $\exp(\beta G(q)) = \alpha d\pi_R(q)/dq$, where $\alpha=\int_{q_L}^{q_R}\exp(\beta G(q'))dq'$ is the normalization factor (see Eq. (8)):

$$T_R(q_0)\pi_R(q_0)=-\alpha\pi_L(q_0)C+\frac{\alpha}{D}\int_{q_0}^{q_R}[\pi_L(q)-\pi_L(q_0)]\pi_R(q)e^{-\beta G(q)}dq\,. \tag{10}$$

Here $\pi_L(q_0) = 1 - \pi_R(q_0)$. The constant C can be found at this stage using the boundary condition at $q_0 = q_L$, $\pi_R(q_L)T_R(q_L) = 0$. After some manipulations we arrive at

$$T_R(q_0)=\frac{\alpha}{D}\int_{q_0}^{q_R}\pi_L(q)\pi_R(q)e^{-\beta G(q)}dq+\frac{\alpha}{D}\pi_L(q_0)\int_{q_L}^{q_0}\frac{\pi_R^2(q)}{\pi_R(q_0)}e^{-\beta G(q)}dq\,. \tag{11}$$

When $q_0 \to q_L$, the second term in the above equation vanishes, and we get the transition path time $t_{TP} \equiv T_R(q_L)$

$$t_{TP}=\frac{\int_{q_L}^{q_R}\left(\int_{q_L}^{x}e^{\beta G(z)}dz\right)\left(\int_x^{q_R}e^{\beta G(z)}dz\right)e^{-\beta G(x)}dx}{D\int_{q_L}^{q_R}e^{\beta G(z)}dz}\,. \tag{12}$$

$$=\frac{1}{D}\int_{q_L}^{q_R}\pi_L(x)\pi_R(x)e^{-\beta G(x)}dx\int_{q_L}^{q_R}e^{\beta G(z)}dz$$

This expression for the average transition path time[3] is a special case of the more general expression derived by Berezhkovskii et al.[4] for the radiation (instead of absorbing) boundary condition at q_L and q_R. One of the remarkable consequences of Eq. (12) is that the transition path time from the left to the right boundary is the same as that from the right to the left, even if the right well, for instance, has a lower energy (as in Fig. S1A). In this case the transitions from the right to the left well occur less frequently, but the transition path time is the same.[4]

To proceed further, we assume that the free energy profile at the barrier top near $q = q_b$ is quadratic (see Fig. S1B),

$$G(q) = G(q_b) - (\omega^*)^2(q - q_b)^2/2, \tag{13}$$

where $(\omega^*)^2$ is the curvature at the top. By using this in Eq. (12) and replacing variables as $\beta(\omega^*)^2(z - q_b)^2/2 = z_1^2$, $\beta(\omega^*)^2(x - q_b)^2/2 = x_1^2$, $\beta(\omega^*)^2(q_R - q_b)^2/2 = \beta(\omega^*)^2(q_L - q_b)^2/2 = q_1^2 = \beta\Delta G^*$ (see Fig. S1B), we have

$$t_{TP} = \frac{2}{\beta D^*(\omega^*)^2} \frac{\int_{-q_1}^{q_1} \left(\int_{-q_1}^{x_1} e^{-z_1^2} dz_1 \right) \left(\int_{x_1}^{q_1} e^{-z_1^2} dz_1 \right) e^{x_1^2} dx_1}{\int_{-q_1}^{q_1} e^{-z_1^2} dz_1}, \tag{14}$$

where star at D^* indicates the diffusion coefficient at the barrier top. In the next step we evaluate the integrals in the brackets in the numerator using $\int_{\pm a}^{b} \exp(-z^2) dz = [\mathrm{erf}(b) \mp \mathrm{erf}(a)]\sqrt{\pi}/2$, where erf($x$) is the error function, and simplify the expression using the fact that the integrated functions are even:

$$t_{TP} = \frac{\sqrt{\pi}}{\beta D^*(\omega^*)^2 \mathrm{erf}(q_1)} \int_0^{q_1} [\mathrm{erf}^2(q_1) - \mathrm{erf}^2(x_1)] e^{x_1^2} dx_1 . \tag{15}$$

The integral in the above expression can be further simplified by integrating by parts:

$$t_{TP} = \frac{4}{\beta D^*(\omega^*)^2 \mathrm{erf}(q_1)} \int_0^{q_1} \mathrm{erf}(x_1) F(x_1) dx_1 , \tag{16}$$

where $F(z) = \exp(-z^2) \int_0^z \exp(x^2) dx$ is the Dawson integral.

To obtain the transition path time in the limit of large q_1 (or large ΔG^*), we present the error function in the above expression in terms of the complementary error function, erf(x_1) = 1 – erfc(x_1), and evaluate the integral $\int_0^q F(x) dx = q^2 {}_2F_2(1,1;\tfrac{3}{2},2;-q^2)/2$, where ${}_2F_2(a_1,a_2;b_1,b_2;\, c)$ is the generalized hypergeometric function:

$$t_{TP} = \frac{4}{\beta D^*(\omega^*)^2 \mathrm{erf}(q_1)} \left[q_1^2 {}_2F_2(1,1;\tfrac{3}{2},2;-q_1^2)/2 - \int_0^{q_1} \mathrm{erfc}(x_1) F(x_1) dx_1 \right]. \tag{17}$$

In the limit $q_1 >> 1$, erf(q_1) $\rightarrow$ 1, the integral in the above expression is evaluated using Mathematica (Wolfram Research, Inc.) as $\int_0^{\infty} \mathrm{erfc}(x_1) F(x_1) dx_1 = (\ln 2)/4$, and the generalized hypergeometric function is approximated using the asymptotic expansion, $z^2 {}_2F_2(1,1;3/2,2;-z^2) \rightarrow \ln(2z) + \gamma/2$, where $\gamma \approx 0.577$ is the Euler constant. Collecting all terms and replacing q_1^2 by $\beta\Delta G^*$, we get the average transition path time in the high barrier limit:

$$t_{TP} = \frac{1}{\beta D^*(\omega^*)^2} \ln(2e^{\gamma} \beta\Delta G^*) . \tag{18}$$

The above transition path time was first derived by A. Szabo. The derivation is based on the assumption of high barrier, $\beta\Delta G^* >> 1$, and quadratic potential at the barrier top. Since the dependence on the barrier height is weak, $\sim \ln\beta\Delta G^*$, this time is not sensitive to the choice of the boundaries q_L and q_R that define the transition path.

II. Comparison of the transition path times calculated using Szabo's equation (Eq. (18)) and numerical integration (Eq. (12)).

In this section, we discuss the effect of the shape of a free-energy barrier and the curvature at the barrier top on the transition path time. The average transition path times are calculated by the numerical integration of Eq. (12) ($t_{TP,num}$) and compared with those obtained from Szabo's equation, Eq. (18) ($t_{TP,Szabo}$).

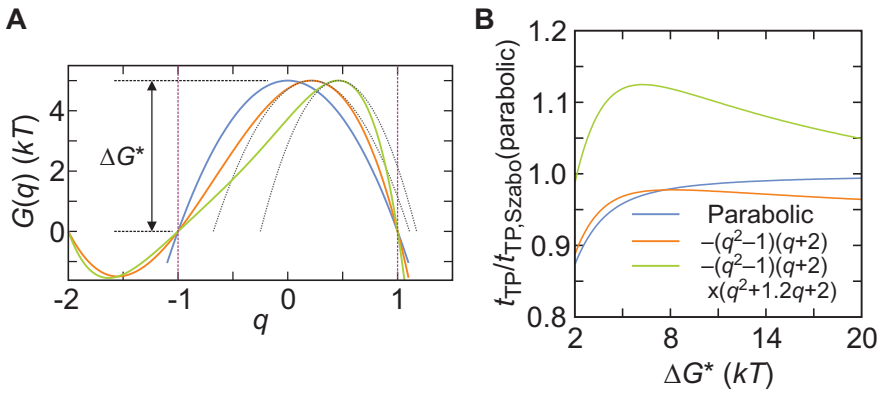

Figure S2. (A) Three free energy barriers with parabolic (blue) and two asymmetric shapes (red, green). The equations for the blue, red, and green barriers are $-A(q^2 - 1)$, $-B(q^2 - 1)(q + 2)$, and $-C(q^2 - 1)(q + 2)(q^2 + 1.2q + 2)$, respectively. The boundaries of the transition path are q = -1 and 1 (purple dashed lines). The constants A, B, and C are determined according to the barrier height, ΔG^* (5 kT in this figure). The black dotted lines show the parabolas with the curvatures at the maxima of the two asymmetric barriers, which are 1.25 (red) and 1.98 (green) times larger than that of the parabolic barrier (blue). (B) The numerically calculated transition path times, $t_{TP,num}$, of three barriers relative to $t_{TP,Szabo}$ of a parabolic barrier as a function of ΔG^*. Note that the barrier height is changed without changing the boundaries for the transition path, and therefore, the curvature of the barrier top increases linearly with the increasing barrier height.

Fig. S2A shows three barriers (ΔG^* = 5 kT) with different shapes. The barrier in blue is a parabolic barrier and those in red and green are asymmetric. Fig. S2B shows the ratio of $t_{TP,num}$ calculated for these three barriers to $t_{TP,Szabo}$ of a parabolic barrier (blue curve in Fig. S2A). As derived in the previous section, $t_{TP,num}$ is the exact average transition path time regardless of the shape of the barrier while $t_{TP,Szabo}$ is an approximate value on the assumption that the barrier is sufficiently high and its shape is parabolic. As expected from this assumption for $t_{TP,Szabo}$, in the case of the parabolic barrier (blue curve in Fig. S2A), $t_{TP,Szabo}$ is slightly longer than $t_{TP,num}$ for a low barrier but they converge as the barrier height increases (blue curve in Fig. S2B).

For the other two asymmetric barriers, the curvatures at the top are larger than that of the parabolic barrier (blue), as indicated by the black dotted lines in Fig. S2A. Therefore, as follows

from Eq. (18), the transition path times of these asymmetric barriers are expected to be shorter than that of the parabolic barrier. Interestingly, however, $t_{TP,num}$ for these two barriers are not so short but similar to that of the parabolic barrier (Fig. S2B). (Note that the numerically calculated values for the three barriers are divided by a constant, $t_{TP,Szabo}$ of a parabolic barrier, for a given ΔG^*.) The underestimation of the transition path time in Eq. (18) results from the shape of the barrier. As mentioned above, $t_{TP,Szabo}$ is the transition path time for a parabolic barrier with the curvature at the barrier top, black dotted lines in Fig. S2A. Due to their large curvatures, these parabolas do not cover the tail parts on the left side of the original barriers (red and green curves) while the differences on the right side of the barriers are relatively small. Therefore, the narrower transition path range of $t_{TP,Szabo}$ compared to that of $t_{TP,num}$ results in the underestimation of the transition path time. Another interesting result in this comparison is that the transition path time is not affected by the detailed shape of the barrier so much if the barrier height and the range of the transition path are the same.

In general, it is not easy to extract the shape of the free energy surface from the experimental data. However, there have been several studies to reconstruct a 1-D free energy profile either using statistical analyses of fluorescence photon trajectories[5,6] or by deconvolution of single molecule trajectories obtained from optical force experiments.[7-9] In these cases, the average transition path time can be obtained by both Eq. (12) and Eq. (18). If the shape of a barrier is parabolic, the two values will be similar. However, as discussed above, if the shape of a barrier significantly deviates from the parabolic shape, Eq. (18) will not give an accurate value. In addition, the transition path time calculated numerically using Eq. (12) will be more sensitive to the choice of the boundaries compared to that of the parabolic barrier.

References

1 C. W. Gardiner, *Handbook of stochastic methods for physics, chemistry and the natural sciences*, Springer, Berlin, 1985.

2 A. Szabo, K. Schulten and Z. Schulten, *J. Chem. Phys.*, 1980, **72**, 4350-4357.

3 G. Hummer, *J. Chem. Phys.*, 2004, **120**, 516-523.

4 A. M. Berezhkovskii, M. A. Pustovoit and S. M. Bezrukov, *J. Chem. Phys.*, 2003, **119**, 3943-3951.

5 G. F. Schröder and H. Grubmüller, *J. Chem. Phys.*, 2003, **119**, 9920-9924.

6 K. R. Haas, H. Yang and J.-W. Chu, *J. Phys. Chem. B*, 2013, **117**, 15591-15605.

7 M. T. Woodside, P. C. Anthony, W. M. Behnke-Parks, K. Larizadeh, D. Herschlag and S. M. Block, *Science*, 2006, **314**, 1001-1004.

8 K. Neupane, D. B. Ritchie, H. Yu, D. A. N. Foster, F. Wang and M. T. Woodside, *Phys. Rev. Lett.*, 2012, **109**, 068102.

9 J. C. M. Gebhardt, T. Bornschlögl and M. Rief, *Proc. Natl. Acad. Sci. USA*, 2010, **107**, 2013-2018.