THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

Cite This: J. Phys. Chem. B 2018, 122, 11095−11099

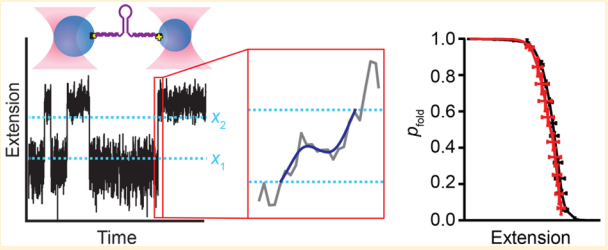
pubs.acs.org/JPCB

Testing Kinetic Identities Involving Transition-Path Properties Using Single-Molecule Folding Trajectories

Krishna Neupane,[†] Noel Q. Hoffer,[†] and Michael T. Woodside*

Department of Physics, University of Alberta, Edmonton, AB T6G 2E1, Canada

ABSTRACT: Recent advances in single-molecule assays have allowed individual transition paths during the folding of single molecules to be observed directly. We used the transition paths of DNA hairpins having different sequences, measured with high-resolution optical tweezers, to test theoretical relations between the properties of the transition paths and the folding kinetics. We showed that folding and unfolding rates were related to the average transition-path times, as expected from theory, for all hairpins studied. We also found that the probability distribution of transition-path occupancies agreed with the profile of the average velocity along the transition paths for each of the hairpins, as expected theoretically. Finally, we used the latter result to show that the committor probability recovered from the velocity profile matches the committor measured empirically. These results validate the proposed kinetic identities.



INTRODUCTION

Transition paths involve the part of a reaction during which the energy barrier between reactants and products is crossed,[1−3] as illustrated in Figure 1a. In biomolecular folding reactions, they are the most interesting part of the folding trajectories, because they encapsulate all the key information about folding mechanisms—the high-energy transition states that are occupied during the transition paths dominate the reaction kinetics. Although transition paths occupy only a very small fraction of the time spent on folding, their central role implies that transition-path properties should be related to macroscopic kinetics such as rates. Indeed, theoretical work[2,4] has shown that for a two-state system, rates and transition-path times for folding and unfolding can be related by

$$2P_F k_U = 2P_U k_F = p(\mathrm{TP})/\tau_{tp} \quad (1)$$

where $P_{F/U}$ is the occupancy of the folded/unfolded state, $k_{F/U}$ is the rate for folding/unfolding, $p(\mathrm{TP})$ is the fraction of the trajectory occupied by transition paths, and τ_{tp} is the average duration of the transition paths. Eq 1 expresses the physical intuition that the number of transitions should be given by the ratio of the total amount of time spent on transition paths to the average time taken to cross a transition path once.

A second relationship that has been proposed[5,6] is that the average velocity profile along the transition paths should be inversely proportional to the position probability distribution within them:

$$P(x|\mathrm{TP}) = [\langle v(x)\rangle \tau_{tp}]^{-1} \quad (2)$$

where $P(x|\mathrm{TP})$ is the probability density for being at a value x of the reaction coordinate when on a transition path and $\langle v(x)\rangle$ is the average velocity at x along the transition path. Again, this relation arises from a simple physical intuition: considering a single transition, the time needed to move from x to $x + dx$ is just $dx/v(x)$, hence the fraction of the transition-path time spent in this reaction-coordinate interval is just $dt/t_{tp} = dx/[v(x)t_{tp}]$, where t_{tp} is the time to complete that particular transition path. Since this relation is true for each transition, the analogous expression holds for the average over all transition paths, and hence the probability for being between x and $x + dx$ is $P(x|\mathrm{TP})dx = dx/[\langle v(x)\rangle \tau_{tp}]$.

These relationships between kinetics and transition-path properties have not yet been fully tested experimentally, because transition paths are very difficult to observe experimentally owing to their brief duration, typically 10^3- to 10^6-fold (or more) shorter than the lifetime of the unfolded or folded states. However, improvements in single-molecule instrumentation in recent years have now allowed transition-path properties to be measured directly. Single-molecule fluorescence spectroscopy has been used to measure the average transition-path time for both proteins[7−9] and DNA hairpins,[10] whereas force spectroscopy has been used to measure the time for individual transition-path crossings,[11−13] the probability distribution along transition paths,[14,15] and the local velocity along the paths.[16] We previously showed that eq 1 holds for a single hairpin,[12] but we did not test its generality by examining different molecules with different unfolding properties. Eq 2 has not yet been tested experimentally at all.

Here we test eqs 1 and 2 using force spectroscopy measurements of DNA hairpins as model two-state folders. Probing the folding dynamics with high-resolution optical

Special Issue: William A. Eaton Festschrift

Received: June 4, 2018
Revised: July 10, 2018
Published: July 13, 2018

DOI: 10.1021/acs.jpcb.8b05355
J. Phys. Chem. B 2018, 122, 11095−11099

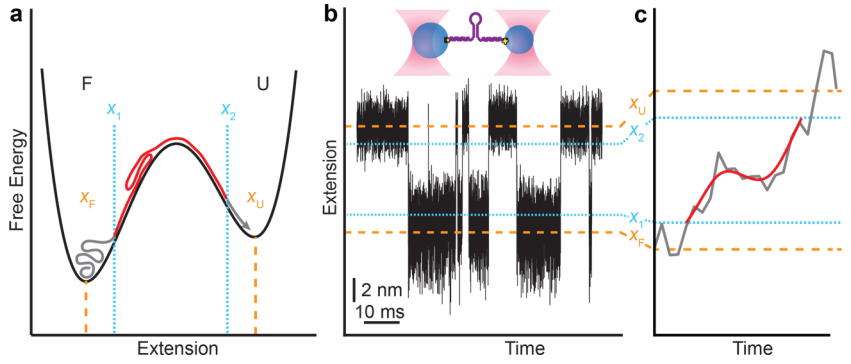


Figure 1. Transition paths in folding reactions. (a) Transition paths are the reactive part of a folding trajectory (red) crossing over the energy barrier separating folded (F) and unfolded (U) states, which excludes the nonproductive fluctuations within the potential wells (gray). (b) Extension trajectory of a single DNA hairpin held in optical traps showing multiple transitions between the unfolded state (at x_U) and folded state (at x_F). Inset: Schematic of measurement geometry showing hairpin attached via DNA handles to beads held in optical traps. (c) Example of a single transition from the extension trajectory, with the transition path traversing from one edge of the barrier region to the other (denoted by the boundaries x_1 and x_2).

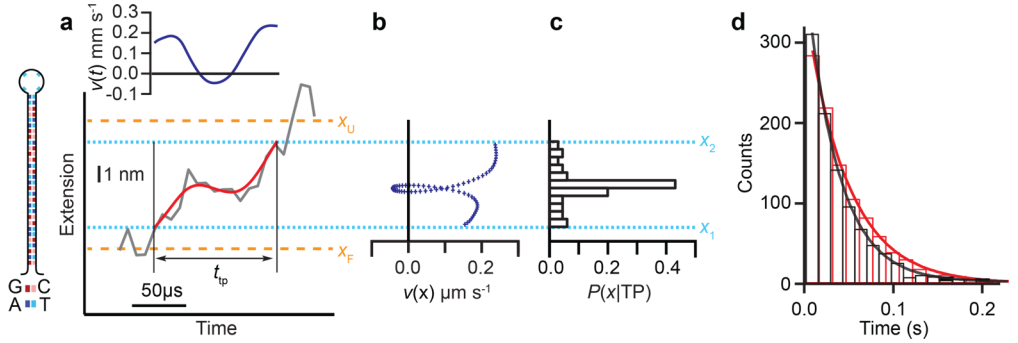


Figure 2. Transition-path analysis. (a) The transit time for an individual transition path, t_{tp}, is measured directly from the trajectory as the time taken to cross between the boundaries x_1 and x_2. The local velocity (upper inset) is calculated from the slope of the transition path after smoothing (red). Left inset: Sequence of DNA hairpin. (b) The velocity profile along the reaction coordinate is multivalued where there is recrossing in the transition path. (c) The transition-path occupancy is found directly from each trajectory. (d) The exponential decay of the dwell times in the folded (black) and unfolded (red) states yields the unfolding and refolding rates.

tweezers,[17] we study hairpins with different sequences, energy-landscape shapes, and unfolding forces in order to test the generality of the relations under a range of conditions. We find that both equations are valid for all of the hairpins studied, despite differences in the unfolding forces, rates, transition times, and average velocity profiles for the different hairpin sequences.

METHODS

Measurements. We analyzed folding trajectories of single DNA hairpins measured under tension reported previously.[13] Briefly, single hairpin molecules of a given sequence were incorporated into tethers containing ~1 kilobase-long handles of double-stranded (ds) DNA on each end of the hairpin via autosticky PCR and ligation, as described.[18] The hairpin sequences used were 20R0/T4, 20R25/T4, 20R55/T4, 20R100/T4, and 30R50/T4 from ref 18 and 20TS06/T4 from ref 19. Tethers were attached to polystyrene beads held in a dual-beam optical trap (Figure 1b, upper inset) similar to one described previously,[20] with the trap stiffnesses kept high (0.75−1.1 pN/nm in one trap and 0.56−0.63 pN/nm in the other) to maximize the time resolution of the measurement. Hairpins were held under tension near $F_{1/2}$, the force at which the folded and unfolded states were equally occupied, and allowed to fluctuate in equilibrium between the two states at constant trap separation while measuring the end-to-end extension of the molecule (Figure 1b). Under these conditions, the time resolution of the measurement was 6−9 μs.[13] For each hairpin, between 12 000 and 55 000 transitions were measured. The end-to-end extension of the molecule was previously shown to be a good reaction coordinate for DNA hairpin folding,[14] an assumption that is implicit in the analysis below.

Transition-Path Analysis. Transition paths were identified as described previously,[12] being those parts of the trajectory passing between two boundaries, x_1 and x_2, demarking the barrier region (defined here as the middle 2/3 of the distance traversed between the folded and unfolded states), as illustrated in Figure 1c. Transit times for individual transition paths (t_{tp}) were measured directly from the trajectories for each transition (Figure 2a) and averaged to obtain τ_{tp}. The velocity at each point along the transition path, $v(x)$, was found as described previously[16] by differentiating the extension trajectory numerically, after first smoothing the trajectories with a smoothing-spline interpolation (Figure 2a, red). The parameters for the smoothing spline were chosen such that the average velocity decreased after smoothing by less than ~20%, as described.[16] Note that because of recrossing events, $v(x)$ was not necessarily single-valued (Figure 2b). The velocity was recorded for each time the path crossed a given x value; counting each crossing as a separate event ensured that

DOI: 10.1021/acs.jpcb.8b05355
J. Phys. Chem. B 2018, 122, 11095−11099

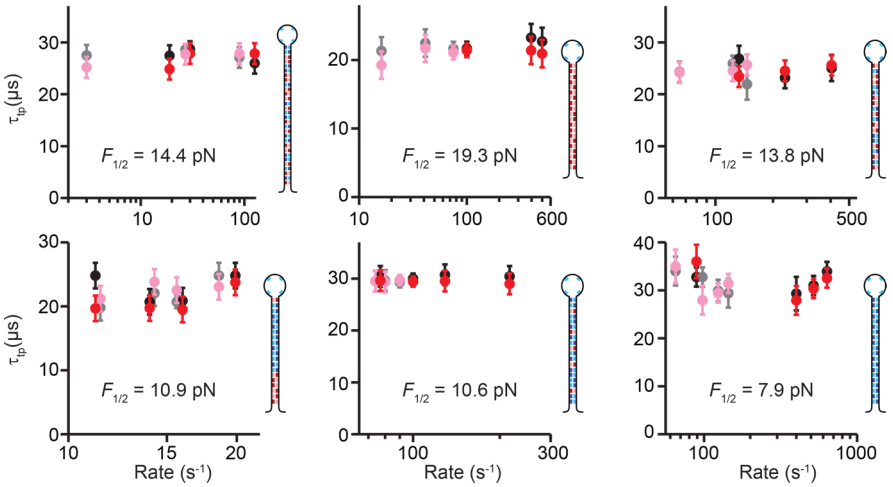


Figure 3. Test of relation between average transition-path times and rates. The average transition-path times measured empirically from folding (black) and unfolding (gray) transitions matched the times predicted from the rates for folding (red) and unfolding (pink) by eq 1, for each of the hairpins studied. $F_{1/2}$ is listed for each hairpin. Insets show hairpin sequences. Error bars represent s.e.m.

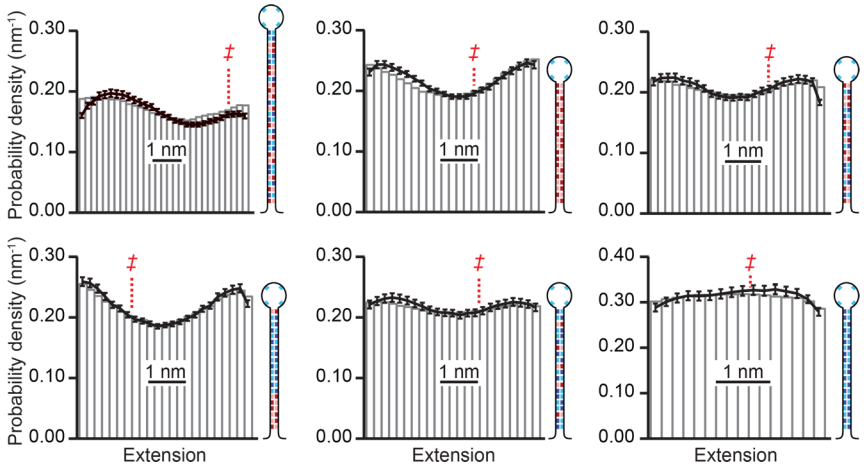


Figure 4. Test of relation between transition-path velocity and occupancy. The average transition-path occupancy (gray) agreed well with the expectation from the transition-path velocity (black) via eq 2 for all hairpins. The location of the barrier top is indicated for each hairpin by the double dagger. Insets show hairpin sequences. Error bars represent s.e.m.

the cumulative time spent near x in each trajectory was correctly counted for comparison to $P(x|\mathrm{TP})$ in eq 2. The average velocity profile over all paths, $\langle v(x)\rangle$, was then found by averaging all velocity values measured at a given position x for all transition-path trajectories. Because the extension change observed in these measurements was only ~1/3 of the value observed in constant-force measurements,[18] owing to the reduction of the distance traveled by the beads compared to the extension change in the hairpin that arises from the compliance of the traps and handles,[21] the observed velocities of the beads were multiplied by the compliance correction factor to recover the velocity of the hairpin ends. Finally, the extension distribution along the transition paths, $P(x|\mathrm{TP})$, was found directly from the trajectories by binning the extension values observed within the transition paths (Figure 2c).

Kinetic Analysis. The distributions of lifetimes for the unfolded and folded states were found from the extension trajectories by partitioning the trajectories into the two states via thresholding as described previously.[18] Folding and unfolding rates were determined from single-exponential fits to the distribution of unfolded- and folded-state lifetimes, respectively (Figure 2d). The occupancies of the folded and unfolded states, $P_{F/U}$, were determined from the fraction of time spent in each state. The barrier-peak location for each hairpin was found previously[13] from energy-landscape reconstructions using committor analysis of the extension trajectories.[22]

RESULTS AND DISCUSSION

From the measured extension trajectories for each hairpin (Figure 1b), we identified each folding and unfolding transition and determined the time for crossing the barrier region in each case, t_{tp}, directly from the trajectory (Figure 2a). The average transition-path time, τ_{tp}, was then calculated, pooling all measurements made at a given force (where the states would have a given occupancy and transition rates). Measurements were repeated for each hairpin over the small range of forces where equilibrium transitions could be reliably observed, typically with state occupancies $P_{F/U} \approx 0.03-0.97$. The resulting values (Figure 3, black: folding, gray: unfolding) were then compared directly with the average transition-path times predicted by eq 1 (Figure 3, red: folding, pink: unfolding). We found excellent agreement between the measured and predicted values across the whole set of hairpins,

DOI: 10.1021/acs.jpcb.8b05355
J. Phys. Chem. B 2018, 122, 11095−11099

covering a wide range of folding/unfolding rates for each hairpin (in most cases at least 1 order of magnitude) and a range of forces from ∼7 up to ∼20 pN, firmly establishing the validity of eq 1.

We next examined the relationship between the average velocity along the transition path and the transition-path occupancy. The average velocity profile was obtained by differentiating smoothing-spline interpolations of the transitions (Figure 2a) to obtain the velocity at each point along the reaction coordinate (Figure 2b). The transition-path occupancy was obtained by isolating the transition paths from the rest of the trajectory and calculating the extension probability distribution at each point of the reaction coordinate within the barrier region (Figure 2c). We then calculated $P(x|\mathrm{TP})$ (Figure 4, gray), finding that the different hairpins had somewhat different transition-path occupancies: for hairpin 20R0/T4, it was effectively flat across the barrier region, whereas for the other hairpins, it was lowest near the top of the barrier (indicated by the ‡ symbol). Despite these differences, when comparing $P(x|\mathrm{TP})$ to $[\langle v(x)\rangle \tau_{tp}]^{-1}$ (Figure 4, black), we found good agreement for all hairpins, validating eq 2.

These results are seemingly quite simple, but they contain some interesting implications. Considering first the transition-path times observed in Figure 3, we see that they have little, if any, dependence on the rates for the transitions (or equivalently on the relative occupancy of the folded and unfolded states) and hence on the applied force. The lack of force-dependence for the transition time contrasts starkly with the strong force-dependence of the rates, which vary roughly exponentially with force,[23,24] as illustrated in Figure 5 for the hairpin 20R100/T4. The exponential force-dependence of the rates arises from the fact that the barrier height, $\Delta G^{\ddagger}$, varies linearly with force to a first approximation,[23−25] and rates are exponentially sensitive to barrier heights.[26] In contrast, τ_{tp} depends only logarithmically on the barrier height:[4,27]

$$\tau_{tp} \approx \frac{\ln(2e^{\gamma}\beta\Delta G^{\ddagger})}{\beta D \kappa_b} \tag{3}$$

where D is the diffusion coefficient, κ_b is the barrier stiffness, β is the inverse thermal energy, and γ is Euler's constant. Hence the force-dependence of $\Delta G^{\ddagger}$ should cause only small changes in τ_{tp}. Given that to first order the barrier curvature should be insensitive to force, the fact that τ_{tp} is roughly constant with force also implies that D, too, has little to no force-dependence. In eq 1, then, the strong force-dependence of $k_{F/U}$ must be balanced precisely by an inverse dependence of $P_{F/U}$ on force to produce a force-independent τ_{tp}.

Turning to the validation of eq 2, we note that the agreement between the velocity calculated from the smoothed trajectories and the probability distribution measured directly from the original data indicates that the smoothing procedure is not introducing artifacts into the transition-path shape: the average velocity profile is the same as what would be expected on the basis of the local occupancy statistics. However, this work shows that measurements of the transition-path occupancy are sufficient to determine the average velocity profile across the transition paths, since these two quantities are inversely proportional. $P(x|\mathrm{TP})$ thus offers an alternative route to measuring $\langle v(x)\rangle$ directly, one that is less stringent technically: it could in principle be used if the trajectories are too noisy to obtain $v(x)$ via differentiation or even if it is not possible to sample the transition paths sufficiently finely to define their shapes adequately (e.g., owing to sampling-rate limitations). Of course, $P(x|\mathrm{TP})$ is more limited than direct measurement of velocities, as it yields only the average velocity profile. If the folding is dominated by a single type of transition path, this approach may be sufficient to characterize the transition behavior,[5] but if multiple, distinct types of transition paths are present, then it may provide a misleading picture because it reflects the average behavior.

An intriguing consequence of eq 2 is that $\langle v(x)\rangle$ can be used to evaluate the committor probability, $p_{fold}(x)$, through the relationship of the latter to $P(x|\mathrm{TP})$. By Bayes' theorem, $P(x|\mathrm{TP}) = P(x)p(\mathrm{TP}|x)/p(\mathrm{TP})$, where $p(\mathrm{TP}|x)$ is the conditional probability of being on a transition path at extension x, and $P(x)$ is the equilibrium extension probability (where the effects of measurement compliance have been removed by deconvolution[19]). Given that we also have $p(\mathrm{TP}|x) = 2p_{fold}(x)[1 - p_{fold}(x)]$ for ideal diffusion,[28] we obtain

$$p_{fold}(x)[1 - p_{fold}(x)] = N/[2T\langle v(x)\rangle P(x)] \tag{4}$$

where N is the number of transitions observed in the trajectory, and T is the total trajectory duration. We verified this relation using transition-path measurements of hairpin 30R50/T4. Calculating p_{fold} from the average transition-path velocity via eq 4 (Figure 6, red), we compared the result to the

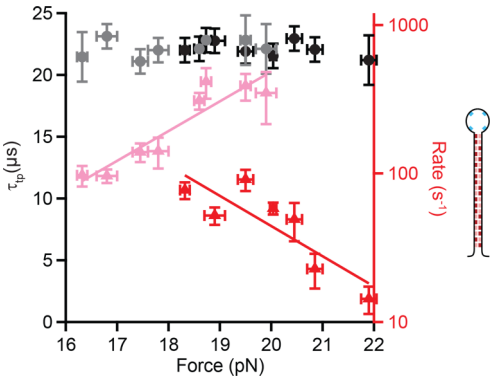


Figure 5. Force-dependence of transition-path times compared to rates. As illustrated using measurements of hairpin 20R100/T4, the transition-path times for folding (black) and unfolding (gray) are independent of force, whereas the rates for folding (red) and unfolding (pink) vary exponentially with force. Error bars represent s.e.m.

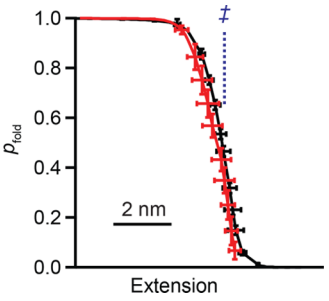


Figure 6. Comparison of empirical p_{fold} and p_{fold} calculated from transition-path velocity. The committor determined empirically from the full extension trajectories (black, error bars represent s.e.m.) matches the result for p_{fold} obtained from transition-path velocities via eq 4 (red, error bars represent standard error from bootstrapping analysis).

DOI: 10.1021/acs.jpcb.8b05355
J. Phys. Chem. B 2018, 122, 11095−11099

committor calculated directly from the extension trajectories using the definition of p_{fold},[29] as done previously[22] (Figure 6, black). The agreement was very good within error. We note that the relation between the committor and $p(TP|x)$ underlying eq 4 assumes folding is a diffusive process (as verified experimentally in previous work[14]), but eqs 1 and 2 do not; hence, this validation of eq 4 underlines the consistency of the diffusive model of folding with the identities being tested.

CONCLUSIONS

We tested theoretical relationships between rates and transition-path times, transition-path occupancies and velocities, and transition-path velocities and the committor function. In each case, we found quantitative agreement, validating the corresponding identities. This work illustrates how experimental measurements of transition-path properties can help validate physical theories of folding by probing the microscopic dynamics during folding reactions.

AUTHOR INFORMATION

Corresponding Author

*E-mail: michael.woodside@ualberta.ca.

ORCID

Michael T. Woodside: 0000-0003-4695-0397

Author Contributions

†K.N. and N.Q.H. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Natural Sciences and Engineering Research Council Canada and Alberta Innovates Technology Futures.

REFERENCES

(1) Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L. Transition Path Sampling: Throwing Ropes Over Rough Mountain Passes, in the Dark. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291−318.

(2) Hummer, G. From Transition Paths to Transition States and Rate Coefficients. *J. Chem. Phys.* **2004**, *120*, 516−523.

(3) E, W.; Vanden-Eijnden, E. Transition-Path Theory and Path-Finding Algorithms for the Study of Rare Events. *Annu. Rev. Phys. Chem.* **2010**, *61*, 391−420.

(4) Chaudhury, S.; Makarov, D. E. A Harmonic Transition State Approximation for the Duration of Reactive Events in Complex Molecular Rearrangements. *J. Chem. Phys.* **2010**, *133*, 034118.

(5) Makarov, D. E. Shapes of Dominant Transition Paths from Single-Molecule Force Spectroscopy. *J. Chem. Phys.* **2015**, *143*, 194103.

(6) Berezhkovskii, A. M.; Makarov, D. E. Communication: Transition-Path Velocity as an Experimental Measure of Barrier Crossing Dynamics. *J. Chem. Phys.* **2018**, *148*, 201102.

(7) Chung, H. S.; McHale, K.; Louis, J. M.; Eaton, W. A. Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times. *Science* **2012**, *335*, 981−984.

(8) Chung, H. S.; Eaton, W. A. Single-Molecule Fluorescence Probes Dynamics of Barrier Crossing. *Nature* **2013**, *502*, 685−688.

(9) Chung, H. S.; Piana-Agostinetti, S.; Shaw, D. E.; Eaton, W. Structural Origin of Slow Diffusion in Protein Folding. *Science* **2015**, *349*, 1504−1510.

(10) Truex, K.; Chung, H. S.; Louis, J. M.; Eaton, W. A. Testing Landscape Theory for Biomolecular Processes with Single Molecule Fluorescence Spectroscopy. *Phys. Rev. Lett.* **2015**, *115*, 018101.

(11) Yu, H.; Dee, D. R.; Liu, X.; Brigley, A. M.; Sosova, I.; Woodside, M. T. Protein Misfolding Occurs by Slow Diffusion across Multiple Barriers in a Rough Energy Landscape. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 8308−8313.

(12) Neupane, K.; Foster, D. A. N.; Dee, D. R.; Yu, H.; Wang, F.; Woodside, M. T. Direct Observation of Transition Paths during the Folding of Proteins and Nucleic Acids. *Science* **2016**, *352*, 239−242.

(13) Neupane, K.; Wang, F.; Woodside, M. T. Direct Measurement of Sequence-Dependent Transition Path Times and Conformational Diffusion in DNA Duplex Formation. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 1329−1334.

(14) Neupane, K.; Manuel, A. P.; Lambert, J.; Woodside, M. T. Transition-Path Probability as a Test of Reaction-Coordinate Quality Reveals DNA Hairpin Folding Is a One-Dimensional Diffusive Process. *J. Phys. Chem. Lett.* **2015**, *6*, 1005−1010.

(15) Neupane, K.; Manuel, A. P.; Woodside, M. T. Protein Folding Trajectories Can Be Described Quantitatively by One-Dimensional Diffusion over Measured Energy Landscapes. *Nat. Phys.* **2016**, *12*, 700−703.

(16) Neupane, K.; Hoffer, N. Q.; Woodside, M. T. Measuring the Local Velocity along Transition Paths during the Folding of Single Biological Molecules. *Phys. Rev. Lett.* **2018**, *121*, 018102.

(17) Ritchie, D. B.; Woodside, M. T. Probing the Structural Dynamics of Proteins and Nucleic Acids with Optical Tweezers. *Curr. Opin. Struct. Biol.* **2015**, *34*, 43−51.

(18) Woodside, M. T.; Behnke-Parks, W. M.; Larizadeh, K.; Travers, K.; Herschlag, D.; Block, S. M. Nanomechanical Measurements of the Sequence-Dependent Folding Landscapes of Single Nucleic Acid Hairpins. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 6190−6195.

(19) Woodside, M. T.; Anthony, P. C.; Behnke-Parks, W. M.; Larizadeh, K.; Herschlag, D.; Block, S. M. Direct Measurement of the Full, Sequence-Dependent Folding Landscape of a Nucleic Acid. *Science* **2006**, *314*, 1001−1004.

(20) Neupane, K.; Yu, H.; Foster, D. A. N.; Wang, F.; Woodside, M. T. Single-Molecule Force Spectroscopy of the Add Adenine Riboswitch Relates Folding to Regulatory Mechanism. *Nucleic Acids Res.* **2011**, *39*, 7677−7687.

(21) Greenleaf, W. J.; Woodside, M. T.; Abbondanzieri, E. A.; Block, S. M. Passive All-Optical Force Clamp for High-Resolution Laser Trapping. *Phys. Rev. Lett.* **2005**, *95*, 208102.

(22) Manuel, A. P.; Lambert, J.; Woodside, M. T. Reconstructing Folding Energy Landscapes from Splitting Probability Analysis of Single-Molecule Trajectories. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112*, 7183−7188.

(23) Bell, G. I. Models for the Specific Adhesion of Cells to Cells. *Science* **1978**, *200*, 618−627.

(24) Dudko, O. K.; Hummer, G.; Szabo, A. Intrinsic Rates and Activation Free Energies from Single-Molecule Pulling Experiments. *Phys. Rev. Lett.* **2006**, *96*, 108101.

(25) Woodside, M. T.; Block, S. M. Reconstructing Folding Energy Landscapes by Single-Molecule Force Spectroscopy. *Annu. Rev. Biophys.* **2014**, *43*, 19−39.

(26) Hänggi, P.; Talkner, P.; Borkovec, M. Reaction-Rate Theory: Fifty Years after Kramers. *Rev. Mod. Phys.* **1990**, *62*, 251−341.

(27) Chung, H. S.; Louis, J. M.; Eaton, W. A. Experimental Determination of Upper Bound for Transition Path Times in Protein Folding from Single-Molecule Photon-by-Photon Trajectories. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 11837−11844.

(28) Best, R. B.; Hummer, G. Reaction Coordinates and Rates from Transition Paths. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6732−6737.

(29) Du, R.; Pande, V. S.; Grosberg, A. Y.; Tanaka, T.; Shakhnovich, E. S. On the Transition Coordinate for Protein Folding. *J. Chem. Phys.* **1998**, *108*, 334−350.

DOI: 10.1021/acs.jpcb.8b05355
J. Phys. Chem. B 2018, 122, 11095−11099