THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

A Two-Step Method for smFRET Data Analysis

Jixin Chen,[*,†] Joseph R. Pyle,[†] Kurt Waldo Sy Piecco,[†] Anatoly B. Kolomeisky,[‡] and Christy F. Landes[‡,§]

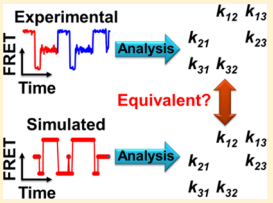
[†]Department of Chemistry and Biochemistry, Ohio University, Athens, Ohio 45701, United States

[‡]Department of Chemistry, Rice University, Houston, Texas 77251, United States

[§]Department of Electrical and Computer Engineering, Rice University, Houston, Texas 77251, United States

Supporting Information

ABSTRACT: We demonstrate a two-step data analysis method to increase the accuracy of single-molecule Förster Resonance Energy Transfer (smFRET) experiments. Most current smFRET studies are at a time resolution on the millisecond level. When the system also contains molecular dynamics on the millisecond level, simulations show that large errors are present (e.g., > 40%) because false state assignment becomes significant during data analysis. We introduce and confirm an additional step after normal smFRET data analysis that is able to reduce the error (e.g., < 10%). The idea is to use Monte Carlo simulation to search ideal smFRET trajectories and compare them to the experimental data. Using a mathematical model, we are able to find the matches between these two sets, and back guess the hidden rate constants in the experimental results.



INTRODUCTION

Single-molecule FRET (smFRET) is an experimental method to measure the dynamic states and rate constants in a molecular system.[1] Its ability to measure rate constants under equilibrium conditions is essential in many biological systems where nonequilibrium states are short-lived and difficult to be isolated from the background, and synchronizing the movement of the molecules is difficult. In the past decade, smFRET has been widely applied in research areas, such as protein conformational change upon thermal or chemical agitations;[2−4] folding/unfolding of DNAs,[5−7] RNAs,[8,9] and proteins;[10−13] and the interactions between protein−DNA,[14,15] protein−RNA,[16] and protein−protein.[17] Such studies have fundamentally changed our understanding of these systems.

This technique is particularly powerful in measuring single-molecule movement at the millisecond to minute time scale with recent efforts pushing the detection limit to the submillisecond level.[18] Faster dynamics at the picosecond to microsecond scales are averaged during the measurement as part of the signal noise. A scheme of the system and data is shown in Figure 1. The FRET value of each molecule changes over time representing the motion of the molecule that has been stochastically locked in a dynamic "state", giving a dwell time of that state followed by a transition to another state. Many measurements focus on first-order activities, where the histogram of the length of the dwell times of each state exponentially decays with increasing of the dwell time.[3−6,19] The average dwell time (the lifetime of each state τ_i) is the reciprocal of the sum of the rate constants to all other states, $\tau_i = 1/\Sigma_{j=1}^{N} k_{ij}$, where N is the total number of states and k_{ij} is the rate constant from state S_i to S_j.[20−22] The lifetimes of interesting biological motions are often in the microsecond to second range in many biological systems, the "slow" time scale (tier-0 dynamics).[23] Accurately measuring the dwell time is critical for smFRET experiments, yet the current available technology has limited the accuracy of measuring relatively fast transitions.[18,24]

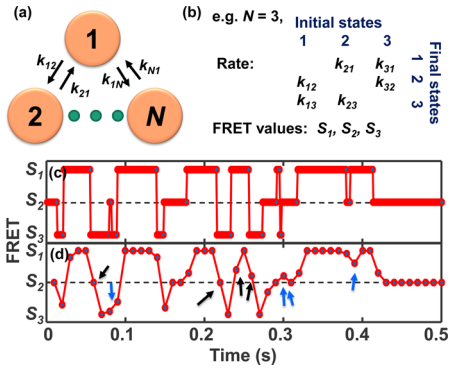


Figure 1. Scheme for simulating a smFRET trajectory. (a) The transition model used for the simulation. (b) The rate constant matrix used in the simulation. (c) A smFRET trajectory simulated to represent the hidden truth (noise-free) at much finer time resolution (1 ms) than the actual experimental time resolution. (d) Experimental observation of the same trajectory (at 10 ms time resolution). Data points were calculated from the integrations of the green and red signals, not from the averages of the FRET values. The arrows show the data points that have large possibilities of being falsely assigned in most software packages. The black arrows indicate the transition edges and the blue arrows indicate fast transitions, both greatly affect the assignment of the transitions and the measurement of the dwell times. Note that experimental noise has not been added to the data, which if added would make the state identification even more difficult.

Most commonly, in a smFRET experiment, the molecule of interest is tagged with two fluorescent dye molecules that have different colors, e.g., green and red, one as a fluorescent energy

Received: June 6, 2016
Revised: June 30, 2016
Published: July 5, 2016

DOI: 10.1021/acs.jpcb.6b05697
J. Phys. Chem. B 2016, 120, 7128−7132

donor and the other as an acceptor. The labeled molecule is immobilized at an interface and the green dye is excited with a laser. The energy absorbed by the green dye initiates fluorescent emission signal of the green dye, or is transferred to the red dye that emits another signal when the red dye is at its adjacent vicinity and at an appropriate orientation. The average physical distance between these two dyes changes with the transitions of the target molecule among different states (e.g., folding/unfolding), and so does the ratio of the fluorescent signals from the donor and from the acceptor: FRET = $I_a/(I_a+I_d)$ where I_a and I_d are the intensities of the acceptor and donor, respectively.[2,25] These signals are often collected by cameras over time with each data point representing the sum signals over a short period of time, and are limited by the camera frame rate.[24] This time interval is the smallest unit used to define a "state" of the target molecule. The accuracy of the state identification depends on the signal-to-noise ratio of the measurement, and at the same time, the probability of transitions happening within this period of time.

The discrete data format has been considered to cause unavoidable errors for smFRET data analysis.[26] A false state is often assigned if one or more transitions have happened in the smallest time unit and the rate constants analyzed after the state identification will contain errors. This problem, called the "camera blurring effect",[27] is severe when these false states have the same FRET values as the real states (an example is shown in Figure 1d). When the transition rates are relatively slow compared to the time resolution of the camera, the current software packages based on Hidden Markov analysis or change point analysis extract the real values, to name several major software, HaMMy,[28] vbFRET,[27] SMART,[29] varBayes-HMM,[30] change point analysis,[31,32] and STaSI.[33] However, when the transition rates are comparable or faster than the time resolution of the camera (real rates faster than the limitation of temporal resolution), significant errors will remain. Several examples are shown later and in the Supporting Information (SI). Statistically, for an exponential decay distribution, when the lifetime τ is the same as the discrete time resolution t_R, $1 - e^{-t_R/\tau} = 1 - e^{-1} = 63\%$; when τ is twice t_R, 39%; and when τ is five times t_R, 18% of the dwell times are shorter than the length of one data point. These dwell times have a large probability to be inaccurately measured and/or falsely assigned, which affects the accurate calculation of the average dwell time. In other words, there is a practical limit on the temporal resolution of the rate constants that is longer than the experimental time resolution.

The smFRET field is pursuing faster time resolution to reveal biological activities and to compare the experimental results with simulations established from a first-principle understanding of atomic and molecular actions.[23] Because many biological molecular motions are at the microsecond region near or below the current limit of smFRET, increasing the accuracy of experimental data analysis for systems with fast dynamics is greatly needed. In this report, we demonstrate a two-step data analysis method to address this need. In addition to the regular data analysis, smFRET trajectories with higher time resolution than the experimental data are simulated and compared to the experimental results. The best matches are confirmed to have a better accuracy than the results of one-step data analysis. The second step is analogous to "fitting".

RESULTS AND DISCUSSION

It is a common way to evaluate a data analysis method using simulated data whose actual parameters are known.[27,30,33] We first confirm that our Monte Carlo simulation program (see SI) generates well-behaved simulated experimental data. As a quick demonstration of the main idea in this report, we have also adapted a classical thresholding method[29,34] to analyze a large number of Monte Carlo simulated smFRET trajectories (see SI). Table 1 shows the results of the same Monte Carlo simulated trajectory analyzed by established methods that are widely used in experimental analysis and our program. To do the overall comparison of all rate constants rather than individually, we use a weighted sum of the absolute difference, the average percentile difference estimator or weighted L1 norm ($wL1$), between the rate constants obtained from the preset true rate constants and those analyzed:

Table 1. Validating Appropriate Data Simulation and Analysis

	set rate			HaMMy[d]			vbFRET		
	1[a]	2	3	1	2	3	1	2	3
1[b]		15[c]	25		14	26		14	26
2	5		30	5		32	5		29
3	10	20		10	21		10	20	
$wL1_{AT}$		N/A			3.7%[e]			2.3%	
				our codes					
1[b]					15			25	
2			5					31	
3			10		19				
$wL1_{AT}$					1.3%				

[a]The initial state of a transition (i). [b]The transition state (j). [c]The rate constant (k_{ij}) of the transition in s^{-1}. [d]The trajectories are shortened for HaMMy analysis. [e]The $wL1_{AT}$ score of the analyzed rate constants with respect to the true values. Please see the sample trajectory in different formats in the subfolder "ExampleTrajct" of the software package in SI.

$$wL1_{AT} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{k_{i,A} - k_{i,T}}{k_{i,T}}\right| \quad (1)$$

where k_i is the i^{th} rate constant, N is the total number of rate constants, the subscripts A and T represent the analyzed and the true values respectively, $k_{i,A}$ represents the i^{th} rate constant of the analyzed values, and $k_{i,T}$ represents the i^{th} true rate constant known in the simulated experimental data. A small difference (of a few percent) between the analyzed rate constants and the preset values confirms that our simulation and our analysis are working properly.

An additional step after the regular data analysis, post-FRET Monte Carlo back simulation (postFRET), is introduced to correct the camera blurring problem. The number of states is determined from the experimental data during the first step of the data analysis using principles such as maximum likelihood (HaMMY),[35] maximum evidence (vbFRET),[27] and Student's t-test (STaSI).[33] When any of the state lifetimes approaches the experimental time resolution (e.g., two times the time resolution), ~40% events are shorter than the time resolution (Table 2). Under this condition, two kinds of false states caused by camera blurring are expected, fast transitions and transition edges (shown by the blue and black arrows respectively in Figure 1d). For an experimental time resolution that is much

DOI: 10.1021/acs.jpcb.6b05697
J. Phys. Chem. B 2016, 120, 7128−7132

Table 2. Confirmation of Improving Accuracy in an Example Model Using postFRET

	simulated true			HaMMy[d]			SMART		
	1[a]	2	3	1	2	3	1	2	3
1[b]		15[c]	25		18	11		18	11
2	5		30	9		27	9		28
3	10	20		4	15		3	15	
$wL1_{AT}$		N/A			42%[e]			43%	
	vbFRET			vbFRET-C			postFRET[f]		
1[b]		17	2		14	16		15	27
2	13		41	9		21	5.8		28
3	0.7	12		5	8		9.5	19	
$wL1_{AT}$		73%			44%			7%	

[a]The initial state of a transition (i). [b]The transition state (j). [c]The rate constant (k_{ij}) of the transition in unit s^{-1} (true values are confirmed to within ±3% by all methods when not binned). [d]The trajectories are shortened for HaMMy analysis. [e]The $wL1_{AT}$ score of the analyzed rate constants with respect to the true values. [f]Time resolution 1 ms binned into 10 ms, each guessed trajectory length is 500 s, scanning time is 21 min. The $wL1_{AT}$ score can be improved to ∼5% at 2000 s length. Please see the sample trajectory in different format in the subfolder "ExampleTrajct" of the software package in SI. The errors of the postFRET analysis have been estimated in the SI.

faster than (e.g., 1/10 of) the shortest lifetime of the states, the errors in the rate constant analysis can be corrected by an existing method.[1,27] Under this condition, the major false states are within those data points that last only one data point (mainly transition edges shown in Figure 1d by the black arrows). This type of false state has been well understood and addressed,[1] e.g., the software package vbFRET has incorporated an option to correct the results using a one-data-point correction/reassigning method (vbFRET-C, SI, Table S2).[27] But this method cannot correct the errors when the transitions are faster than the frame time.

In order to increase the accuracy of experimental data analysis for the system with fast transitions, the idea reported here is to simulate smFRET trajectories at finer time resolution than the experimental conditions (to simulate the hidden truth), then bin the data (sum a fixed number of data points into one) to mimic the camera blurring effect in real experiments (Figure 1d). We find that the blurring effect can be simulated by just a few times better time resolution than the experimental resolution because the information of interest does not depend on the simulation time resolution (SI, Table S3). Thus, the need and cost of simulating molecular motion at the nanosecond or even faster times is not necessary for this purpose. Note for those motions belonging to activities whose lifetimes are shorter than a few microseconds, the current techniques cannot observe the information in the smFRET experiment but rather it is a source of noise.

We hypothesize that a simulated trajectory that has the same blurring effect as the experimental trajectory represents the hidden truth of the experiment. Thus, we generate a series of simulated trajectories and compare them to the experimental results. To do this comparison, a model criterion is proposed to be the same weighted sum of absolute difference:

$$wL1_{SE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{k_{i,S} - k_{i,E}}{k_{i,E}}\right| \quad (2)$$

where k_i is the rate constant, N is the total number of rate constants, and the subscripts S and E represent the results analyzed with any of the same established methods from the simulated trajectory and the experimental trajectory, respectively. Two identical trajectories give a $wL1_{SE}$ = 0% and postFRET is set to search for a simulated trajectory with the minimum $wL1_{SE}$ score. Note that the rate constants analyzed from experimental and simulated results all contain errors, and the truth is hidden in the experiment but known in the simulated results.

The postFRET method is able to find Monte Carlo simulated smFRET trajectories that regenerate the experimental results. The true rate constants are known for the simulated experimental data no matter how we blurred the trajectory. Thus, for these simulated experiments, we are able to calculate the score $wL1_{AT}$ and to evaluate the guesses by comparing the rate constants analyzed by different methods to the true values. An example is shown in Table 2, where postFRET is able to reduce the average difference $wL1_{AT}$ to 7% from HaMMy (44%), SMART (45%), vbFRET (77%), and vbFRET-C (43%). Many other simulations also confirm this improvement is general to many models (see SI). A software package with the trajectory simulation codes (MATLAB), the postFRET codes (MATLAB), and an example trajectory; estimated analysis errors/uncertainties are included in the SI.

In postFRET, we use a local search algorithm to find a group of Monte Carlo simulated trajectories that matches the experimental results by minimizing $wL1_{SE}$ as the evaluation function. Figure 2 explains the search procedure to find the minimum $wL1_{SE}$.[36] We sequentially search the local neighbor-

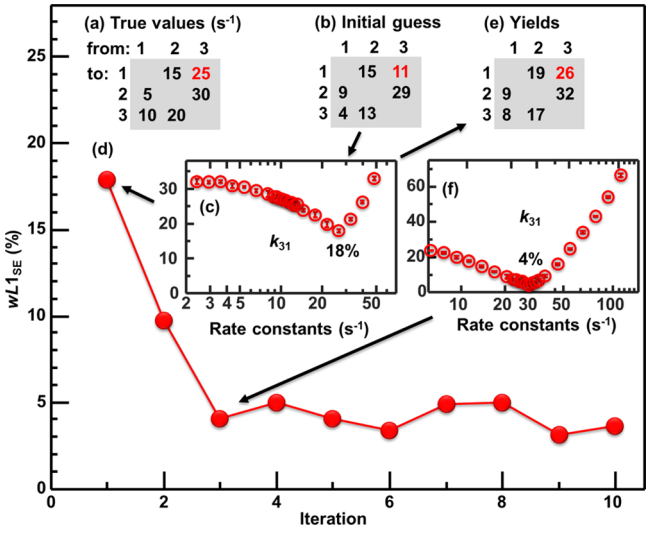


Figure 2. An example of postFRET search. (a) True rate constants. (b) A set of initial guesses from experimental results. (c) Scanning for the k_{31} value at its neighborhood for the local minimum $wL1_{SE}$ = 18% (error bars representing the standard deviation of 5 simulations). The search neighborhood is designed to be exponentially near the initial guess value, while the true value can be outside of this search range. (d) All six rate constants are searched around their neighborhoods sequentially, which yields (e) the best rate constants after the first round (1st iteration). These values are then used as the starting guesses for the next iteration. (f) After 3 rounds of scanning, the best $wL1_{SE}$ score for a scanning of k_{31} is reduced to 4%, yielding one set of accepted final guesses. The standard deviation of several acceptable guesses can be used to estimate the reproducibility of postFRET simulations. See SI for the algorithm and the MATLAB codes.

DOI: 10.1021/acs.jpcb.6b05697
J. Phys. Chem. B 2016, 120, 7128−7132

hoods of the rate constants in each iteration. The resulting best guesses are then used as the initial guesses of the next search iteration. The theoretical minimum $wL1_{SE}$ is the uncertainty of the Monte Carlo simulation, which is dependent on the length of the simulation and the stability of the random function used in the simulation. Thus, we see a fast drop of the $wL1_{SE}$ score at the first few iterations but then stabilized at ∼4% rather than 0%. The variation of the rate constants in the tail comprehensively reflects the reproducibility and the error of the search algorithm.

The search neighborhood is set to a relatively large scale (Figure 2c,f), which reduces the probability of getting stuck at local optimums. The variation of the Monte Carlo simulation also helps the search algorithm to overcome local optima. Thus, we find that simulated annealing search (commonly used to overcome local optimums) or machine learning is not needed in the models we have tested. However, they are potentially needed in more complicated models or for higher precision levels.

The postFRET method works generally in several other sample models that have four states, five states, six states, with the sequential reaction model (e.g., $A \overset{K_1}{\leftrightarrow} B \overset{K_2}{\leftrightarrow} C$), or simply a two-state model (see SI, Tables S4−S10). For example, the simulated trajectory for the two-state model is shown in Figure 3, when ∼40% of the well times of state S_1 and ∼60% of the dwell times of state S_2 are shorter than one-data-point. This lack of information makes it difficult to extract the true rate constants from this trajectory. However, if we force a two-state analysis, postFRET is able to accurately extract the rate constant at $wL1_{AT}$ = 1% (SI). Future improvements for postFRET can be investigated, such as different evaluation functions, scanning methods, more transition models, and extension to other data format such as single photon counting data. For proof-of-concept purpose in this report, noise is not discussed to reduce the complexity, whose effect is expected to depend on the structure of the noise and the signal-to-noise ratio.[26,29] Preliminary results show that no obvious effect is observed for a signal-to-noise ratio of ∼10 (data not shown). This method is expected to be more useful for noisy experiments than less noisy experiments if the noise can be properly measured and simulated.

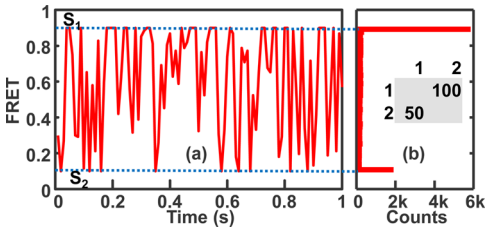


Figure 3. A two-state trajectory with true rate constants 100 s^{-1} and 50 s^{-1} simulated at 0.1 ms, and then binned into an experimental time resolution at 10 ms (i.e., the noise-free red and green signals were binned, then the FRET values were calculated). (a) FRET values of the first 1 s data of the 135 s long trajectory. (b) Histogram of the occurrence of the FRET values of the total trajectory (inset shows the rate constants). True state values are set to FRET = 0.9 and 0.1, respectively.

More efficient search algorithms and parallel computational methods are highly desired in the future to reduce the data analysis time. Generating one Monte Carlo time trajectory is fast, e.g., ∼0.3 s for a trajectory with 1×10^5 data points. After the trajectory is binned into 1×10^4 data points, the computational time for the rate constant analysis is ∼0.2 s. However, because of the massive semi-exhaustive search algorithm adapted in this report, the overall searching time for the 3 state example is ∼10 min, and the 5-state and 6-state examples take a few hours to a day to converge.

CONCLUSIONS

We have demonstrated the proof-of-concept that adding an additional Monte Carlo data analysis step significantly increases the accuracy of experimental data analysis for smFRET. It allows the retrieval of the information that is blurred in a typical data collection. We believe that this additional analysis will improve smFRET research when studying biological systems with fast transition rates, whose dynamic state lifetimes are at or below the limitation of many current measurements.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcb.6b05697.

Methods, different models, sample trajectories, and error analysis (PDF)

Open MATLAB codes (ZIP)

AUTHOR INFORMATION

Corresponding Author

*Telephone: +1-740-593-9768; Email: chenj@ohio.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

J.C. thanks the Ohio University faculty startup funding, OURC award, CMSS, and NQPI. J.C. thanks Dr. Hugh H. Richardson and Dr. Alexander Govorov for beneficial discussions. C.F.L. thanks Welch Foundation (Grant C-1787). A.B.K. acknowledges Welch Foundation (Grant C-1559), NSF (Grant CHE-1360979), and the Center for Theoretical Biological Physics sponsored by the NSF (Grant PHY-1427654).

REFERENCES

(1) Roy, R.; Hohng, S.; Ha, T. A Practical Guide to Single-Molecule FRET. *Nat. Methods* **2008**, *5*, 507−516.

(2) Ha, T.; Ting, A. Y.; Liang, J.; Caldwell, W. B.; Deniz, A. A.; Chemla, D. S.; Schultz, P. G.; Weiss, S. Single-Molecule Fluorescence Spectroscopy of Enzyme Conformational Dynamics and Cleavage Mechanism. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96* (3), 893−898.

(3) Landes, C. F.; Rambhadran, A.; Taylor, J. N.; Salatan, F.; Jayaraman, V. Structural Landscape of Isolated Agonist-Binding Domains from Single AMPA Receptors. *Nat. Chem. Biol.* **2011**, *7* (3), 168−173.

(4) Vafabakhsh, R.; Levitz, J.; Isacoff, E. Y. Conformational Dynamics of a Class C G-Protein-Coupled Receptor. *Nature* **2015**, *524* (7566), 497−501.

(5) McKinney, S. A.; Declais, A.-C.; Lilley, D. M. J.; Ha, T. Structural Dynamics of Individual Holliday Junctions. *Nat. Struct. Biol.* **2003**, *10* (2), 93−97.

(6) Chen, J.; Poddar, N. K.; Tauzin, L. J.; Cooper, D.; Kolomeisky, A. B.; Landes, C. F. Single-Molecule FRET Studies of HIV TAR−DNA Hairpin Unfolding Dynamics. *J. Phys. Chem. B* **2014**, *118* (42), 12130−12139.

(7) Cosa, G.; Harbron, E. J.; Zeng, Y.; Liu, H. W.; O'Connor, D. B.; Eta-Hosokawa, C.; Musier-Forsyth, K.; Barbara, P. F. Secondary Structure and Secondary Structure Dynamics of DNA Hairpins

DOI: 10.1021/acs.jpcb.6b05697
J. Phys. Chem. B 2016, 120, 7128−7132

Complexed with HIV-1 NC Protein. *Biophys. J.* **2004**, *87* (4), 2759−2767.
(8) Zhuang, X.; Rief, M. Single-Molecule Folding. *Curr. Opin. Struct. Biol.* **2003**, *13* (1), 88−97.
(9) Keller, B. G.; Kobitski, A.; Jäschke, A.; Nienhaus, G. U.; Noé, F. Complex RNA Folding Kinetics Revealed by Single-Molecule FRET and Hidden Markov Models. *J. Am. Chem. Soc.* **2014**, *136* (12), 4534−4543.
(10) Margittai, M.; Widengren, J.; Schweinberger, E.; Schröder, G. F.; Felekyan, S.; Haustein, E.; König, M.; Fasshauer, D.; Grubmüller, H.; Jahn, R.; et al. Single-Molecule Fluorescence Resonance Energy Transfer Reveals a Dynamic Equilibrium between Closed and Open Conformations of Syntaxin 1. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100* (26), 15516−15521.
(11) Chung, H. S.; McHale, K.; Louis, J. M.; Eaton, W. A. Single-Molecule Fluorescence Experiments Determine Protein Folding Transition Path Times. *Science* **2012**, *335* (6071), 981−984.
(12) McLoughlin, S. Y.; Kastantin, M.; Schwartz, D. K.; Kaar, J. L. Single-Molecule Resolution of Protein Structure and Interfacial Dynamics on Biomaterial Surfaces. *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110* (48), 19396−19401.
(13) Li, C.-B.; Yang, H.; Komatsuzaki, T. Multiscale Complex Network of Protein Conformational Fluctuations in Single-Molecule Time Series. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (2), 536−541.
(14) Uphoff, S.; Holden, S. J.; Le Reste, L.; Periz, J.; van de Linde, S.; Heilemann, M.; Kapanidis, A. N. Monitoring Multiple Distances within a Single Molecule Using Switchable FRET. *Nat. Methods* **2010**, *7* (10), 831−836.
(15) Hwang, W. L.; Deindl, S.; Harada, B. T.; Zhuang, X. Histone H4 Tail Mediates Allosteric Regulation of Nucleosome Remodelling by Linker DNA. *Nature* **2014**, *512* (7513), 213−217.
(16) Kim, H.; Abeysirigunawarden, S. C.; Chen, K.; Mayerle, M.; Ragunathan, K.; Luthey-Schulten, Z.; Ha, T.; Woodson, S. A. Protein-Guided RNA Dynamics during Early Ribosome Assembly. *Nature* **2014**, *506* (7488), 334−338.
(17) Keller, A. M.; Benítez, J. J.; Klarin, D.; Zhong, L.; Goldfogel, M.; Yang, F.; Chen, T.-Y.; Chen, P. Dynamic Multibody Protein Interactions Suggest Versatile Pathways for Copper Trafficking. *J. Am. Chem. Soc.* **2012**, *134* (21), 8934−8943.
(18) Juette, M. F.; Terry, D. S.; Wasserman, M. R.; Zhou, Z.; Altman, R. B.; Zheng, Q.; Blanchard, S. C. The Bright Future of Single-Molecule Fluorescence Imaging. *Curr. Opin. Chem. Biol.* **2014**, *20*, 103−111.
(19) Zhuang, X.; Bartley, L. E.; Babcock, H. P.; Russell, R.; Ha, T.; Herschlag, D.; Chu, S. A Single-Molecule Study of RNA Catalysis and Folding. *Science* **2000**, *288* (5473), 2048−2051.
(20) Raff, L. M. *Principles of Physical Chemistry*; Prentice Hall: Upper Saddle River, NJ, 2001.
(21) Gillespie, D. T. Exact Stochastic Simulation of Coupled Chemical Reactions. *J. Phys. Chem.* **1977**, *81* (25), 2340−2361.
(22) Benitez, J. J.; Keller, A. M.; Ochieng, P.; Yatsunyk, L. A.; Huffman, D. L.; Rosenzweig, A. C.; Chen, P. Probing Transient Copper Chaperone-Wilson Disease Protein Interactions at the Single-Molecule Level with Nanovesicle Trapping. *J. Am. Chem. Soc.* **2008**, *130*, 2446−2447.
(23) Henzler-Wildman, K.; Kern, D. Dynamic Personalities of Proteins. *Nature* **2007**, *450* (7172), 964−972.
(24) Nir, E.; Michalet, X.; Hamadani, K. M.; Laurence, T. A.; Neuhauser, D.; Kovchegov, Y.; Weiss, S. Shot-Noise Limited Single-Molecule FRET Histograms: Comparison between Theory and Experiments. *J. Phys. Chem. B* **2006**, *110* (44), 22103−22124.
(25) Förster, T. Energiewanderung und Fluoreszenz. *Naturwissenschaften* **1946**, *33* (6), 166−175.
(26) Lee, T.-H. Extracting Kinetics Information from Single-Molecule Fluorescence Resonance Energy Transfer Data Using Hidden Markov Models. *J. Phys. Chem. B* **2009**, *113* (33), 11535−11542.
(27) Bronson, J. E.; Fei, J.; Hofman, J. M.; Gonzalez, R. L., Jr.; Wiggins, C. H. Learning Rates and States from Biophysical Time Series: A Bayesian Approach to Model Selection and Single-Molecule FRET Data. *Biophys. J.* **2009**, *97* (12), 3196−3205.
(28) McKinney, S. A.; Joo, C.; Ha, T. Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling. *Biophys. J.* **2006**, *91* (5), 1941−1951.
(29) Greenfeld, M.; Pavlichin, D. S.; Mabuchi, H.; Herschlag, D. Single Molecule Analysis Research Tool (SMART): An Integrated Approach for Analyzing Single Molecule Data. *PLoS One* **2012**, *7* (2), e30024.
(30) Okamoto, K.; Sako, Y. Variational Bayes Analysis of a Photon-Based Hidden Markov Model for Single-Molecule FRET Trajectories. *Biophys. J.* **2012**, *103* (6), 1315−1324.
(31) Taylor, J. N.; Li, C.-B.; Cooper, D. R.; Landes, C. F.; Komatsuzaki, T. Error-Based Extraction of States and Energy Landscapes from Experimental Single-Molecule Time-Series. *Sci. Rep.* **2015**, *5*, 9174.
(32) Watkins, L. P.; Yang, H. Detection of Intensity Change Points in Time-Resolved Single-Molecule Measurements. *J. Phys. Chem. B* **2005**, *109* (1), 617−628.
(33) Shuang, B.; Cooper, D.; Taylor, J. N.; Kisley, L.; Chen, J.; Wang, W.; Li, C. B.; Komatsuzaki, T.; Landes, C. F. Fast Step Transition and State Identification (STaSI) for Discrete Single-Molecule Data Analysis. *J. Phys. Chem. Lett.* **2014**, *5* (18), 3157−3161.
(34) Blanco, M.; Walters, N. G. Chapter 9 - Analysis of Complex Single-Molecule FRET Time Trajectories. In *Methods in Enzymology: Single Molecule Tools: Fluorescence Based Approaches, Part A*; Walters, N. G., Ed.; Academic Press, 2010; Vol. *472*, pp 153−178.
(35) McKinney, S. A.; Joo, C.; Ha, T. Analysis of Single-Molecule FRET Trajectories Using Hidden Markov Modeling. *Biophys. J.* **2006**, *91*, 1941−1951.
(36) Michalewicz, Z.; Fogel, D. B. *How to Solve It: Modern Heuristics*; Springer Science & Business Media, 2013.

DOI: 10.1021/acs.jpcb.6b05697
J. Phys. Chem. B 2016, 120, 7128−7132