THE JOURNAL OF PHYSICAL CHEMISTRY A

ARTICLE

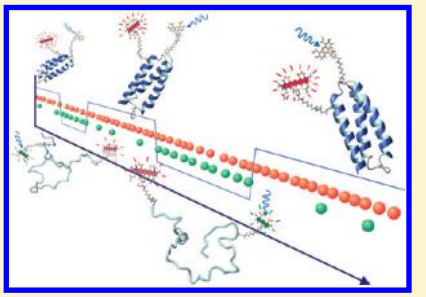
pubs.acs.org/JPCA

Extracting Rate Coefficients from Single-Molecule Photon Trajectories and FRET Efficiency Histograms for a Fast-Folding Protein

Hoi Sung Chung,* Irina V. Gopich, Kevin McHale, Troy Cellmer, John M. Louis, and William A. Eaton*

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, United States

ABSTRACT: Recently developed statistical methods by Gopich and Szabo were used to extract folding and unfolding rate coefficients from single-molecule Förster resonance energy transfer (FRET) data for proteins with kinetics too fast to measure waiting time distributions. Two types of experiments and two different analyses were performed. In one experiment bursts of photons were collected from donor and acceptor fluorophores attached to a 73-residue protein, α_3D, freely diffusing through the illuminated volume of a confocal microscope system. In the second, the protein was immobilized by linkage to a surface, and photons were collected until one of the fluorophores bleached. Folding and unfolding rate coefficients and mean FRET efficiencies for the folded and unfolded subpopulations were obtained from a photon by photon analysis of the trajectories using a maximum likelihood method. The ability of the method to describe the data in terms of a two-state model was checked by recoloring the photon trajectories with the extracted parameters and comparing the calculated FRET efficiency histograms with the measured histograms. The sum of the rate coefficients for the two-state model agreed to within 30% with the relaxation rate obtained from the decay of the donor−acceptor cross-correlation function, confirming the high accuracy of the method. Interestingly, apparently reliable rate coefficients could be extracted using the maximum likelihood method, even at low (<10%) population of the minor component where the cross-correlation function was too noisy to obtain any useful information. The rate coefficients and mean FRET efficiencies were also obtained in an approximate procedure by simply fitting the FRET efficiency histograms, calculated by binning the donor and acceptor photons, with a sum of three-Gaussian functions. The kinetics are exposed in these histograms by the growth of a FRET efficiency peak at values intermediate between the folded and unfolded peaks as the bin size increases, a phenomenon with similarities to NMR exchange broadening. When comparable populations of folded and unfolded molecules are present, this method yields rate coefficients in very good agreement with those obtained with the maximum likelihood method. As a first step toward characterizing transition paths, the Viterbi algorithm was used to locate the most probable transition points in the photon trajectories.



INTRODUCTION

Förster resonance energy transfer (FRET) experiments on single protein molecules can be used to measure the structural and dynamical properties of subpopulations, as well as the kinetics of transitions between subpopulations.[1−3] However, the time scale for the study of kinetics has been limited. The simplest and most frequent single molecule FRET experiment has been to collect photons emitted by donor and acceptor fluorophores attached to the protein freely diffusing through the illuminated volume of a confocal microscope. The photons are binned, and a histogram of the FRET efficiencies for each bin, defined as the fraction of the photons emitted from the acceptor, is constructed. The shape of the histogram determines the minimal number of subpopulations in the sample, while the mean value of the FRET efficiency of each subpopulation yields interdye distances, once the dynamics of the fluorophores and their linkers are properly taken into account.[4,5] If the width of the distribution for a subpopulation is in excess of that expected from shot noise alone, the histogram must be analyzed further. Excess width may be caused by photophysical properties of the fluorophores, such as light-induced spectral changes and quenching of dye fluorescence by mechanisms other than Förster transfer.[6,7] A much more interesting cause of excess width is the interconversion of protein conformations on a time scale comparable to the bin size.[5,8−11] In this case the FRET efficiency histograms change with bin size, a phenomenon with similarities to exchange broadening in NMR experiments, allowing the acquisition of kinetic data from analysis of the histograms.

A second kind of single molecule FRET experiment is to collect photons from molecules immobilized on a surface. The great advantage of this experiment is that molecules can be observed until one of the fluorophores bleaches, which generally is much longer than the observation time in the free diffusion experiment (∼1 ms). The long trajectories also more clearly expose interfering photophysical effects, making it much easier to objectively filter the data. The much greater dynamic range yields dynamical and kinetic information over a wide range of time scales.

Special Issue: Graham R. Fleming Festschrift

Received: February 1, 2010
Revised: May 4, 2010
Published: May 28, 2010

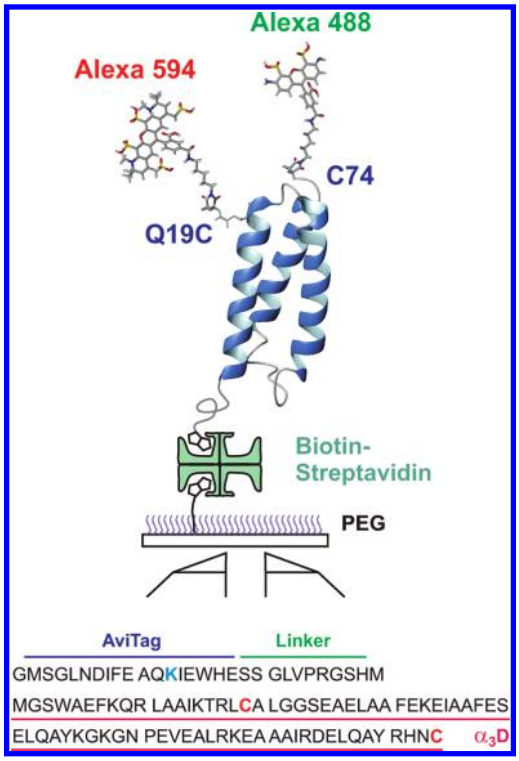


Figure 1. Immobilization of a dye-labeled α_3D via biotin (on the protein)−streptavidin−biotin (on the surface) linkages on a PEG-coated glass surface (PDB code 2A3D). Cysteine residues 19 and 74 of Avi-α_3D$^{Q19C/C74}$ are labeled with donor (Alexa 488) and acceptor (Alexa 594) dyes. The isomers, donor19/acceptor74 and donor74/acceptor19, could not be separated by ion-exchange chromatography, the last step in the purification scheme (see Methods and ref 6).

In this work our objective has been to extract rate coefficients for folding and unfolding from single molecule FRET measurements on a fast folding, two-state protein, α_3D (Figure 1).[12,13] These experiments represent a necessary first step toward the goal of measuring transition path times.[6] Many methods have been developed to obtain kinetics and dynamics from an analysis of photon trajectories, FRET efficiency trajectories, or FRET efficiency histograms in single molecule experiments.[9,11,14−22] We sought simple and analytic methods that could be applied to extract rate coefficients from FRET efficiency histograms and from photon trajectories in both free diffusion and immobilization experiments. Such analytic methods have recently been developed by Gopich and Szabo, and we use them in this work. One is a powerful maximum likelihood method that determines the model parameters that are most consistent with the measured photon trajectories.[20] In this analysis the trajectory is analyzed photon by photon. The method is based on the exact likelihood function corresponding to a two-state kinetic model. It differs from previous maximum likelihood methods in that it is applicable to free diffusion experiments where the emission rate fluctuates due to diffusion within the confocal volume. The ability of the maximum likelihood method to describe the data in terms of a two-state model is checked by recoloring the photon trajectories using the extracted parameters and comparing the calculated FRET efficiency histograms with the measured histograms.[20]

In a more approximate method, these same parameters are obtained by fitting the FRET efficiency histogram to a calculated distribution. For two-state immobilized molecules, the FRET efficiency histograms can be obtained analytically[23,24] from the distribution of the fraction of time spent in one of two interconverting states.[25,26] This distribution has been exploited by Yang and co-workers to extract interconversion rates of enzyme conformations.[9] It has also been incorporated into the probability distribution analyses by Seidel and co-workers.[27] We use a much simpler method where the FRET efficiency histogram is approximated by a sum of three-Gaussian functions with the parameters that are defined by the model of conformational dynamics.[22] This approximation is exact (in the case of immobilized molecules) at short and long bin times.

Both the maximum likelihood and three-Gaussian methods assume a model describing conformational dynamics and thus are not "model-free". In return, these methods permit rate coefficients to be obtained where it is not possible using other methods because the photon detection rate is not sufficiently high compared to the transition frequency to unambiguously assign a bin of photons to a state of the system. Finally, we convert photon trajectories to state trajectories using the Viterbi algorithm.[28,29] The location of transition points between states with the Viterbi algorithm is the first step toward characterizing the transition path between folded and unfolded states.

METHODS

Synthesis, Expression, Purification, and Dye-Labeling of Avi-α_3D$^{Q19C/C74}$. A synthetic gene encoding 73 amino acids of the de novo designed three-helix bundle, α_3D, containing a substitution mutation Q19C, addition of a cysteine residue (C74) at the C-terminus, and an N-terminal biotin acceptor peptide (Avidity LLC, Aurora, CO) and a spacer sequence (Figure 1) was cloned between the Nco1 and BamH1 sites of pET15b vector (Novagen, San Diego, CA). The resulting construct (Avi-α_3D$^{Q19C/C74}$) was verified by DNA sequencing.

The expression construct Avi-α_3D$^{Q19C/C74}$ and a plasmid with an isopropylthiogalactoside (IPTG) inducible birA gene to overexpress the biotin ligase (Avidity LLC) were cotransformed into *E.coli* BL-21 (DE3; Stratagene, La Jolla, CA). Cells were grown in Luria−Bertani medium, and expression was induced at an absorbance of 0.7 monitored at 600 nm with a final concentration of 1 mM IPTG for a period of 3−4 h. A final concentration of 50 μM d-biotin (Sigma, St. Louis, MO) was added to the medium ~30 min before induction. Typically, cells harvested from a 200 mL culture were lysed by uniform suspension in 20 mL of bacterial protein extraction reagent (B-PER, Pierce, Rockford, IL) and sonication. The lysate was centrifuged at 12 800 rpm (SS-34 rotor, ThermoFisher Scientific, Asheville, NC) for 30 min at 4 °C. The supernatant was subjected to affinity chromatography using streptavidin Mutein matrix (Roche Diagnostics GmbH, Mannheim, Germany). The column was equilibrated and washed extensively, after passing the lysate, with 1X PBS (1.7 mM KH_2PO_4, 5 mM Na_2HPO_4, 150 mM NaCl, pH 7.4) and the biotinylated Avi-α_3D$^{Q19C/C74}$ was eluted in 1X PBS containing 2 mM d-biotin. The eluted protein was adjusted to a final concentration of 1 mM DTT, concentrated using centriprep-YM10 devices (Millipore Corp, Bedford, MA) to ~1.5 mL, and loaded onto a Superdex-75 column (1.6 cm ×60 cm; GE HealthCare, Piscataway, NJ) equilibrated in 25 mM Tris−HCl, pH 7.5 and 1 mM DTT at a flow rate of 1.5 mL/min at room temperature. Peak fractions were analyzed by SDS-PAGE, combined and subjected to reversed-phase HPLC on POROS 20 R2

resin (Perseptive Biosystems, Framington, MA) and eluted using a linear gradient from 99.95% water (v/v) and 0.05% TFA to 60% acetonitrile (v/v), 0.05% TFA (v/v), and 39.95% water (v/v) over a period of 16 min at a flow rate of 4 mL/min. Aliquots of the peak fraction were lyophilized and stored at −70 °C. Biotin ligation to the biotin acceptor peptide was confirmed by mass spectrometry. An observed mass of 11 609 g/mol was clearly indicative of the combined mass of biotin (expected mass of 226 g/mol) ligated to acceptor sequence in Avi-$\alpha_3D^{Q19C/C74}$ (expected mass of 11 381 g/mol). Labeling of Alexa Fluor 488 and Alexa Fluor 594 and purification of Avi-$\alpha_3D^{Q19C/C74}$ containing the donor−acceptor pair by size-exclusion chromatography (Superdex-30, 1.6 cm ×60 cm, GE HealthCare) followed by anion-exchange chromatography were performed as described previously with only slight modifications.[6]

Ensemble Tryptophan Fluorescence. Tryptophan fluorescence from unlabeled Avi-$\alpha_3D^{Q19C/C74}$ (23 μM, 50 mM Hepes buffer, pH 7.6, 293 K, 284 nm excitation) was measured as a function of GdmCl concentration under 10-fold excess of dithiothreitol (DTT) to prevent cysteine oxidation. The spectra were filtered to remove noise using singular value decomposition (SVD) and reconstructed from the first two significant SVD components. The peak wavelength of the spectrum was obtained from a polynomial fit.

Single Molecule Spectroscopy. Single molecule FRET experiments were performed using a confocal microscope system (MicroTime200, Picoquant). A dual mode (CW/pulsed) 485 nm diode laser (LDH-D-C-485, PicoQuant) was used to excite donor dyes (Alexa Fluor 488) through an oil-immersion objective (Plan-Apo, NA 1.4, ×100, Olympus). Donor and acceptor (Alexa Fluor 594) fluorescence was collected by the same objective, split into two channels, and focused through optical filters onto photon-counting avalanche photodiodes (PerkinElmer Optoelectronics SPCM-AQR-15). In the free diffusion experiment, a donor dye was excited by the laser in the CW mode to maximize fluorescence emission at an excitation power of 45 μW. Arrival times of donor and acceptor photons were recorded with 1 ps resolution, although the accuracy was limited by the typical response time of the SPADs of ∼350 ps. The protein was diluted to 40 pM, which is sufficient to avoid having two molecules simultaneously in the illuminated volume,[30] in a 50 mM Hepes buffer (pH 7.6) with guanidine hydrochloride (GdmCl). To prevent sticking of proteins on the glass coverslip, 0.01% Tween20 (Thermo Scientific) was used.

In the immobilization experiment, the donor dye was excited by the laser in the pulsed mode (90 ps pulses at 20 MHz, 6 μW) to record both arrival times of photons with 50 ns resolution and the time delay between the laser trigger pulse and the detected photon with 2 ps resolution (∼350 ps jitter) for determining fluorescence lifetimes. The lifetime information was used to remove (∼1%) trajectories produced by impurity molecules based on the observation of fluorescence lifetimes, inconsistent with the apparent FRET efficiency. Protein molecules were immobilized on a biotin-embedded, poly(ethylene glycol) (PEG)-coated glass coverslip (Bio_01, Microsurfaces Inc.) via a biotin (surface)−streptavidin−biotin (protein) linkage (Figure 1). To reduce photobleaching and populating triplet states of the dyes, a cocktail of an oxygen-scavenging system (10 nM protocatechuate 3,4-dioxygenase (PCD, P8279-25UN, Sigma) and 2.5 mM 3,4-dihydroxybenzoic acid (PCA, 37580-25G-F, Sigma)) and a triplet quencher (1 mM Trolox (238813, Sigma)) were included in the Hepes/GdmCl solutions. Additional details for the optical setup and single molecule experiments can be found elsewhere.[5,6]

Extraction of Parameters Using Maximum Likelihood Method. Rate coefficients and mean FRET efficiencies were extracted from photon trajectories using the maximum likelihood method described in detail in ref 20. Briefly, a likelihood function L_j for a photon string of the jth burst in a free diffusion experiment or the jth trajectory of an immobilization experiment was constructed as

$$L_j = \mathbf{1}^{\mathrm{T}} \prod_{k=2}^{N_j} [\mathbf{F}(c_k)\, \exp(\mathbf{K}\tau_k)]\mathbf{F}(c_1)\mathbf{p}_{\mathrm{eq}} \tag{1}$$

where N_j is the number of photons in the burst or trajectory, c_k is the color of the kth photon (donor or acceptor), $\mathbf{K}$ is a matrix of rate coefficients, and τ_k is a time interval between the kth and $(k-1)$th photons. The FRET efficiency matrix $\mathbf{F}$ depends on the color of a photon as $\mathbf{F}$(acceptor) = $\mathbf{E}$ and $\mathbf{F}$(donor) = $\mathbf{I} - \mathbf{E}$, where $\mathbf{E}$ is a diagonal matrix whose elements are the FRET efficiencies of the individual states, $\mathbf{I}$ is the identity matrix, $\mathbf{p}_{\mathrm{eq}}$ is a vector consisting of the equilibrium population of each state, and $\mathbf{1}^{\mathrm{T}}$ is a row vector whose elements are 1.

For a system with only two populations of protein molecules, folded and unfolded, the transitions are described by the two-state kinetic scheme,

$$\mathrm{U} \underset{k_{\mathrm{U}}}{\overset{k_{\mathrm{F}}}{\rightleftarrows}} \mathrm{F} \tag{2}$$

with the FRET efficiencies in the folded and unfolded states ε_{F} and ε_{U}. Here we assume that interdye distance distributions within each state are sampled rapidly compared to the inter-photon interval, so that only mean values of the FRET efficiencies in each state (ε_{F} and ε_{U}) are invoked. In this case eq 1 can be written as

$$\begin{aligned} L_j &= \mathbf{p}_0^{\mathrm{T}} \prod_{k=2}^{N_j} [\Phi(c_k)\, \exp(\Lambda\tau_k)]\Phi(c_1)\mathbf{p}_0 \\ &= \begin{pmatrix} 1 & 0 \end{pmatrix} \prod_{k=2}^{N_j} \left[\Phi(c_k) \times \begin{pmatrix} 1 & 0 \\ 0 & \mathrm{e}^{-(k_{\mathrm{F}}+k_{\mathrm{U}})\tau_k} \end{pmatrix} \right] \Phi(c_1) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{aligned}$$

$$\Phi(\text{acceptor}) = \begin{pmatrix} \varepsilon_{\mathrm{F}} p_{\mathrm{F}} + \varepsilon_{\mathrm{U}} p_{\mathrm{U}} & (\varepsilon_{\mathrm{U}} - \varepsilon_{\mathrm{F}}) p_{\mathrm{U}} \\ (\varepsilon_{\mathrm{U}} - \varepsilon_{\mathrm{F}}) p_{\mathrm{F}} & \varepsilon_{\mathrm{F}} p_U + \varepsilon_{\mathrm{U}} p_{\mathrm{F}} \end{pmatrix}$$

$$\Phi(\text{donor}) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \Phi\,(\text{acceptor}) \tag{3}$$

Here, the rate matrix is given as $\mathbf{K} = \begin{pmatrix} -k_{\mathrm{U}} & k_{\mathrm{F}} \\ k_{\mathrm{U}} & -k_{\mathrm{F}} \end{pmatrix}$, and $\mathbf{\Lambda}$ is a diagonal matrix whose diagonal elements are eigenvalues of $\mathbf{K}$ (i.e., $\mathbf{\Lambda} = \mathbf{U}^{-1}\mathbf{K}\mathbf{U}$). Other matrices and vectors are transformed for calculation as $\mathbf{\Phi} = \mathbf{U}^{-1}\mathbf{F}\mathbf{U}$, $\mathbf{p}_0 = \mathbf{U}^{-1}\mathbf{p}_{\mathrm{eq}}$, and $\mathbf{p}_0{}^{\mathrm{T}} = \mathbf{1}^{\mathrm{T}}\mathbf{U}$. There are four independent parameters, which can be determined by maximizing the product of the likelihood functions ($(L = \Pi L_j)$) for all bursts or trajectories or, equivalently, by maximizing the sum of log-likelihoods. For the sake of convenience, we choose the FRET efficiencies ε_{F} and ε_{U}, the relaxation rate k (= $k_{\mathrm{F}} + k_{\mathrm{U}}$), and the population of the folded

state p_F (= $k_F/(k_F + k_U)$) as independent parameters. Other parameters of interest can be obtained from $p_U = 1 - p_F$, $k_F = kp_F$, and $k_U = kp_U$.

Extraction of Equilibrium and Kinetic Folding Parameters from FRET Efficiency Histogram. Rate coefficients and FRET efficiencies can also be obtained by simply fitting the FRET efficiency histogram (FEH) to an approximate distribution, which, for a two-state system, is the sum of three-Gaussian distributions,[22] with the parameters analytically expressed in terms of the rates and FRET efficiencies, i.e.

$$\mathrm{FEH}(E) \approx A\sum_{i=0}^{2} c_i(2\pi\sigma_i^2)^{-1/2}\exp\left(-\frac{(E-\varepsilon_i)^2}{2\sigma_i^2}\right) \qquad (4)$$

Here, A is the area of the histogram. The subscripts i = 1, 2 represent the folded and unfolded states, respectively. Another Gaussian function (i = 0) accounts for the appearance of FRET efficiency at values intermediate between those of the folded and unfolded peaks due to the transitions between the folded and unfolded states. The parameters for this distribution can be analytically calculated as shown below. The three coefficients are given by

$$\begin{aligned} c_i &= p_i e^{-k_i T} \qquad i = 1, 2 \\ c_0 &= 1 - c_1 - c_2 \end{aligned} \qquad (5)$$

where T is the bin time, $p_1 = p_F$, $p_2 = p_U$, $k_1 = k_U$, and $k_2 = k_F$. Variances of the distributions are given by

$$\sigma_i^2 = \varepsilon_i(1-\varepsilon_i)\langle N^{-1}\rangle \qquad i = 1, 2 \qquad (6)$$

where $\varepsilon_1 = \varepsilon_F$, $\varepsilon_2 = \varepsilon_U$, and $\langle N^{-1}\rangle$ is the average of the inverse of the total number of photons in a bin $N = N_A + N_D$. $\langle N^{-1}\rangle$ is obtained from the experiment. The mean FRET efficiency and the variance of the zeroth distribution are calculated as

$$\begin{aligned} c_0\varepsilon_0 &= \sum_{i=1}^{2}(p_i - c_i)\varepsilon_i \\ c_0\sigma_0^2 &= \langle\varepsilon\rangle_{eq}(1-\langle\varepsilon\rangle_{eq})\langle N^{-1}\rangle \\ &\quad + 2p_1p_2(\varepsilon_2-\varepsilon_1)^2(kT + e^{-kT} - 1)(1-\langle N^{-1}\rangle)/(kT)^2 \\ &\quad + \langle\varepsilon\rangle_{eq}^2 - \sum_{i=0}^{2} c_i\varepsilon_i^2 - \sum_{i=1}^{2} c_i\sigma_i^2 \end{aligned} \qquad (7)$$

where $\langle\varepsilon\rangle_{eq} = p_1\varepsilon_1 + p_2\varepsilon_2$. In eqs 4−7, ε_F, ε_U, k (= $k_F + k_U$), and p_F are fitting parameters.

Converting Photon Trajectories to State Trajectories. In a photon trajectory, each photon is emitted from either the folded or the unfolded state. This photon trajectory can be converted to a state trajectory according to the probabilities calculated from the photon colors, interphoton time intervals, and rate coefficients. The likelihood function in eq 1 accounts for the contribution from all possible state trajectories. We wish to find the specific state trajectory that contributes the most to the likelihood. This most probable trajectory is obtained from the photon trajectory using the Viterbi algorithm[28,29] adjusted for a two-color photon trajectory with measured interphoton times. The procedure consists of four steps: initialization, recursion, termination, and path backtracking.

1. Initialization for the first photon

$$\delta_i(1) = [\mathbf{F}(c_1)\mathbf{p}_{eq}]_i \qquad i = 1, 2 \qquad (8)$$

2. Recursion for n = 2, ..., N

$$\begin{aligned} \delta_i(n) &= [\mathbf{F}(c_n)]_{ii}\,\max_j\{[\exp(\mathbf{K}\tau_n)]_{ij}\delta_j(n-1)\} \qquad i = 1, 2 \\ \sigma_i(n) &= \arg\max_j\{[\exp(\mathbf{K}\tau_n)]_{ij}\delta_j(n-1)\} \qquad i = 1, 2 \end{aligned} \qquad (9)$$

Here, i and j denote the indices of the states (1 or 2) and n labels photons. $\delta_i(n)$ is the score of the most probable state trajectory that ends in state i at the time of arrival of the nth photon. $\sigma_i(n)$ is the index of the most probable previous state. State i at the time of arrival of the nth photon is most likely preceded by the state with index $\sigma_i(n)$. $\max_j\{...\}$ and $\arg\max_j\{...\}$ are the maximum value and its index for the variable in the braces over states j. N is the number of photons in the trajectory.

3. Termination

$$s(N) = \arg\max_j\{\delta_j(N)\} \qquad (10)$$

4. Backtracking

$$s(n-1) = \sigma_{s(n)}(n) \qquad n = N, ..., 2 \qquad (11)$$

$s(n)$ is the index of the most probable state that emitted the nth photon. The intervals where the states of two adjacent photons are different become the transition intervals.

Simulation of Photon Trajectories. To check the accuracy of the maximum likelihood and three-Gaussian methods, as well as the determination of transition intervals, folding and unfolding trajectories were simulated with experimental parameters. A total of 200 trajectories of 30 ms duration were simulated for GdmCl concentrations of 1.5, 2, 2.25, 2.5, and 3 M. The FRET efficiencies of the folded and unfolded states were ε_F = 0.93 and ε_U = 0.61 and the total photon count rate was n = 50 ms^{-1} for all GdmCl concentrations. The total photon count rate was not varied for the folded and unfolded states because there is only ~5% difference in the experiment. The equilibrium population of the folded state (p_F) and the sum of rate coefficient ($k = k_F + k_U$) were varied for different GdmCl concentrations, as summarized in Table 2. The initial state was selected randomly according to p_F. When the trajectory starts with a folded state, a string of time intervals (τ_j, j = 1, 2, ...) was generated with a count rate of $(n + k_U)$. In this string, the time interval is exponentially distributed as $\exp[-(n + k_U)\tau]$. The jth photon was assigned at the cumulative time point of $\Sigma_{i=1}^{j}\tau_i$. Photons were colored according to the probability determined by the FRET efficiency (ε_F for the acceptor and 1 − ε_F for the donor). During generation of the photons in the folded state, a point of a transition to the unfolded state was randomly assigned instead of a photon with a probability of $k_U/(n + k_U)$. Then, an unfolded string of photons with a count rate of $(n + k_F)$ was generated and appended until a transition to the folded state was assigned similarly. This procedure was repeated until the trajectory length reached 30 ms.

RESULTS

Effect of Bin Size and Denaturant Concentration on FRET Efficiency Histograms. Figure 2a shows FRET efficiency histograms (FEH) at the midpoint denaturant concentration of 2.25 GdmCl in a free diffusion experiment. The FRET efficiency is defined as $N_A/(N_A + N_D)$, where N_A and N_D are the number of

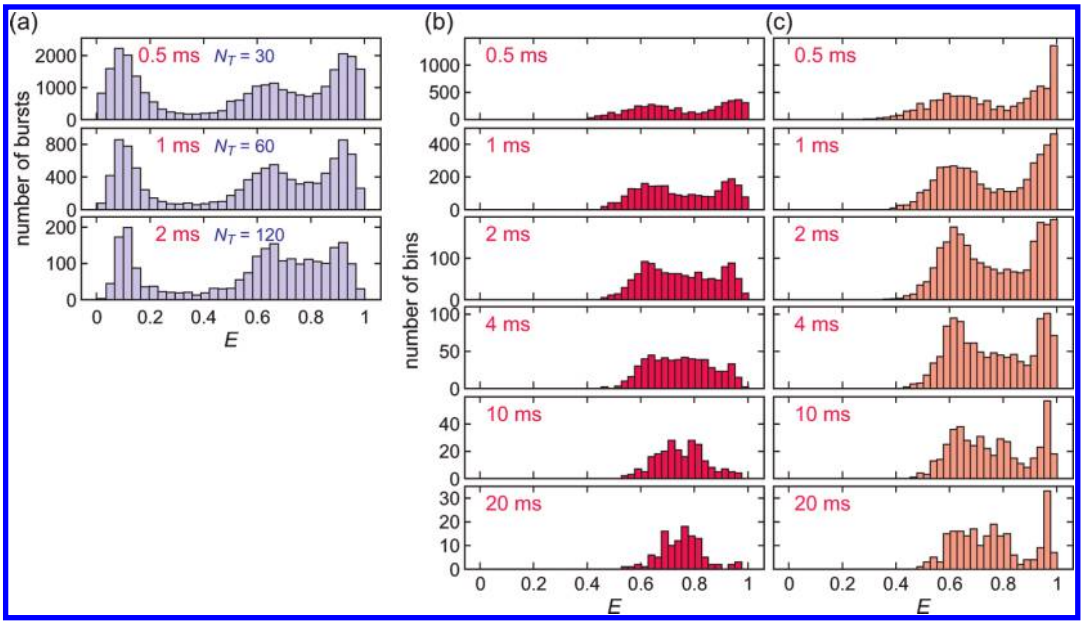


Figure 2. Changes in the FRET efficiency distribution as the bin time increases due to the conformational exchange in the free diffusion (a) and immobilization (b) and (c) experiments at 2.25 M GdmCl. (a) Bins containing more than 30, 60, and 120 photons (N_T) were considered as significant bursts for 0.5, 1, and 2 ms bin times, respectively. (b) and (c) Segments with average photon count rates of >30 ms^{-1} were analyzed. In (c) all data are included. In (b) segments with a FRET efficiency of 0.75 or segments with a constant FRET efficiency of either ~0.6 or ~0.97 lasting longer than 20 ms have been removed (see text).

acceptor and donor photons in a bin, respectively. There are three peaks, one at $E \sim 0$, corresponding to molecules with an inactive acceptor, and peaks at ~0.65 and ~0.95, corresponding to unfolded and folded subpopulations, respectively. Histograms with bin sizes of 0.5, 1, and 2 ms are shown. The bursts of photons last for only 0.5−2 ms, so it was not possible to vary the bin size over a wider range. Nevertheless, the FEH clearly changes as the bin size increases, with the buildup of FRET efficiency at values intermediate between those of the folded and unfolded subpopulations, diagnostic of a dynamical process occurring on a time scale comparable to the bin size.

In the immobilization experiment, the long trajectories permit histograms to be constructed over a much wider range of bin sizes (Figures 2b,c). There are striking changes in the FEHs with bin size, but these histograms are contaminated by several effects and selected trajectories had to be removed prior to further analysis. These contaminating trajectories include those with an extra peak for the unfolded molecule at $E \sim 0.75$ that results from an increased spectral overlap and increased leakage of donor photons into the acceptor channel (from 6% of photons detected in the acceptor channel to 20%) caused by a light-induced, 25 nm red-shift of the donor spectrum (called Alexa 488^R).[6] When the FRET efficiency of the unfolded segment was clearly ~0.75 or the segment with an active acceptor was immediately followed by a high-leak acceptor bleached state ($E > 0.1$), the trajectory was not included in the analysis. These trajectories were about 21% of 1145 trajectories. Since the formation of Alexa 488^R is a light-induced process, the fraction of bursts with Alexa 488^R would be smaller in the free diffusion experiment where a molecule is only very briefly exposed to laser excitation. For example, 12% of the bursts in the free diffusion experiment at 3 M GdmCl belong to this category, as obtained from fitting the data below.

In addition, 7% and 4% of the trajectories in the immobilization experiment show very long segments, corresponding to apparently folded and unfolded molecules, with mean FRET efficiencies of 0.97 and 0.60, respectively. The duration of these segments is more than 10-fold longer than the mean waiting time (determined below) and is therefore so improbable that they cannot be considered as part of the normal folded or unfolded state waiting time distributions. The apparently folded state produces the sharp peak in the measured FEHs close to E = 1.0 (Figure 2c). The origin of these long-lived segments is not clear but may result from formation of a stable complex between the protein and the surface. Parts c and b of Figure 2 show the FEHs before and after removal of the E = 0.75 FRET unfolded state, and the long-lived segments. Changing the bin size from 0.5 to 20 ms in the filtered FEHs (Figure 2b) now clearly shows the changes expected for a two-state system, with the central peak replacing the folded and unfolded peaks as the bin size increases.

Figure 3 shows how changing the denaturant concentration at a fixed bin size affects the FEHs in both the free diffusion and immobilization experiments. The increase of GdmCl from 1.5 to 3 M produces a large population shift from the folded state to the unfolded state. In both experiments the midpoint is closest for the data at 2.25 M GdmCl. Figure 4 shows the ensemble chemical denaturation curve of unlabeled protein measured by tryptophan fluorescence, which results in a comparable denaturation midpoint of 2.50 M (m = 1.75 kcal mol^{-1} M^{-1}).

Extracting Kinetic and Equilibrium Parameters from FRET Efficiency Histograms and Photon Trajectories. The simplest method for extracting rate and equilibrium parameters from the FEHs is to find the parameters of three-Gaussian functions, such that the sum of the Gaussians provides the best least-squares fit to the FEH (see eqs 4−7 of Methods). Parts a and b of Figure 3 show the results of this fit for the free diffusion and immobilization experiments, assuming a two-state model. The four independent parameters of the fit are the FRET efficiencies of the folded and unfolded states, ε_F and ε_U, the sum of the rate coefficients for folding and unfolding ($k = k_F + k_U$), and the fractional population of the folded state, p_F (from which the

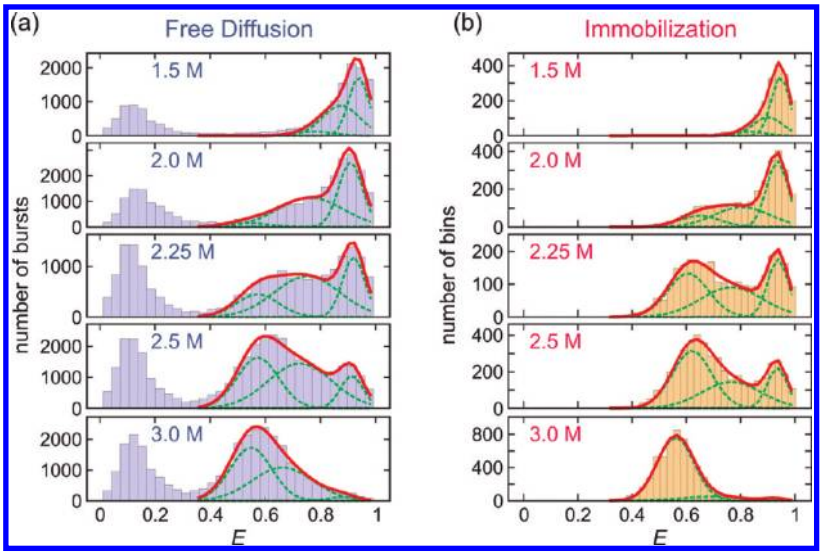


Figure 3. Gausssian fit to FRET efficiency histograms of Alexa 488/Alexa 594-labeled α_3D in the free diffusion experiment (a) and immobilization experiment (b) at constant bin size. In the free diffusion experiment, photons were collected in 2 ms bins and bins containing more than 30 photons were retained. Histograms for the immobilization experiment (b) were constructed by binning photons from trajectories with a photon count rate greater than 15 ms^{-1} (bin time = 1 ms). The average count rate was ~25 ms^{-1} for the free diffusion experiment and ~50 ms^{-1} for the immobilization experiment. Solid and dashed curves are individual Gaussian components and their sum calculated for a two-state model from eqs 4−7. The fitting parameters are given in Table 1.

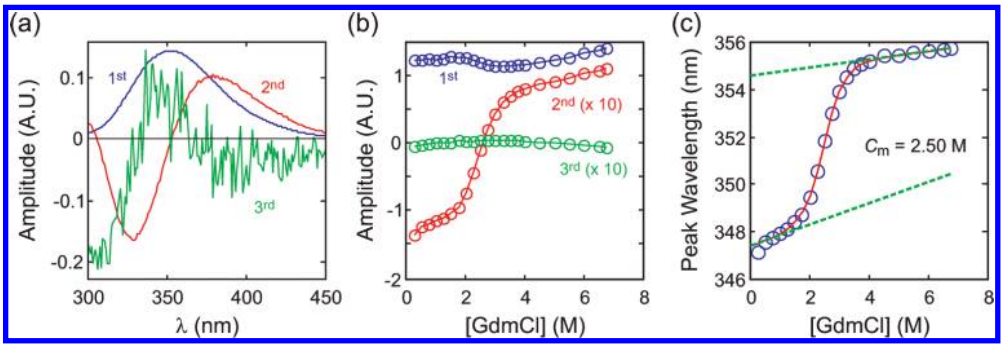


Figure 4. Chemical denaturation of unlabeled α_3D (23 μM) monitored by the tryptophan fluorescence. To prevent the oxidation of free cysteine residues, fluorescence spectra were collected under 10-fold excess of dithiothreitol (DTT). (a) First three singular value decomposition (SVD) component spectra. (b) Changes of the amplitude (multiplied by their singular values) of the three SVD components. (c) Change of the peak wavelength fitted to the equation of the two-state model $\lambda_{peak} = (K(a + b[\text{GdmCl}]) + c + d[\text{GdmCl}])/(K + 1)$, where the equilibrium constant $K = \exp[-m([\text{GdmCl}] - c_m)/RT] = p_F/p_U$. The denaturation midpoint c_m = 2.50 M and m = 1.75 kcal mol^{-1} M^{-1}. Green dashed lines indicate baselines $a + b[\text{GdmCl}]$ and $c + d[\text{GdmCl}]$.

other parameters of interest may be obtained, $p_U = 1 - p_F$, $k_F = kp_F$, $k_U = kp_U$). The values of the parameters are given in Table 1.

Because the fitted relaxation rate k for the immobilized molecules varies with the bin time of the FEH (see below), the FEHs were globally fit to the distributions of three-Gaussian functions with common parameters. The parameters obtained from FEHs of 1 ms bin time and global fitting method are similar except those at 1.5 M GdmCl where the unfolded population is very low (Table 1 and Figure 9). Extracted parameters are presumably most accurate when the bin time is comparable to the inverse of the rate coefficient because there are comparable contributions to the FEH from all three-Gaussian components.

Although FEHs were constructed with a bin time of 2 ms for free diffusion experiments (Figure 3a), the average duration of bursts is shorter than 2 ms. To obtain an effective bin time T for use in eqs 5 and 7, the relative arrival time of each photon in a bin to the mean arrival time was calculated as $t_{j,rel} = t_j - (1/N)\sum_{j=1}^{N} t_j$, where t_j is the arrival time of the jth photon and N is the number of photons in a bin. Then, the distribution of $t_{j,rel}$ averaged for all bins was fit to a Gaussian distribution. The fwhm of this distribution was used for the bin time T. For the data at 1.5, 2, 2.25, 2.5, and 3 M GdHCl, T = 1.08, 1.07, 1.03, 1.05, and 1.06 ms, respectively. However, the estimation of the effective bin time by this method is not always reliable. For example, the dependence of the rate coefficient on the bin time is larger in the free diffusion experiment than in the immobilization experiment. So far, we have not found any more rigorous way to estimate the effective bin time. Therefore, the rate coefficients obtained by the three-Gaussian method, especially from the free diffusion experiment, should be taken as a measure of the time scale and not as an exact parameter.

A more accurate and diffusion-independent determination of the model parameters can be obtained using the maximum likelihood method to analyze photon trajectories using eq 3,

dx.doi.org/10.1021/jp1009669 |J. Phys. Chem. A 2011, 115, 3642–3656

Table 1. Comparison of the Parameters Obtained from the Maximum Likelihood Estimation (MLE) and the Fitting of the FRET Efficiency Distribution with Three-Gaussian Functions (3G)

	GdmCl (M)	ε_F[a]	ε_U[a]	k (ms^{-1})	p_F[a]
immobilization (MLE)	1.5	0.94	0.64	2.64 (±0.23)	0.92
	2	0.93	0.64	1.29 (±0.05)	0.67 (±0.01)
	2.25	0.93	0.61	0.99 (±0.04)	0.43 (±0.01)
	2.5	0.93	0.62	1.06 (±0.04)	0.32
	3	0.92	0.56	1.22 (±0.08)	0.05
immobilization (3G) 1 ms bin time	1.5	0.95	0.83 (±0.04)	1.33 (±0.66)	0.80 (±0.08)
	2	0.93	0.65 (±0.02)	1.10 (±0.15)	0.68 (±0.02)
	2.25	0.94	0.60	0.99 (±0.06)	0.43
	2.5	0.93	0.62	0.96 (±0.06)	0.33
	3	0.93 (±0.02)	0.56	1.24 (±0.32)	0.06 (±0.02)
immobilization (3G) 0.5 − 10 ms bin time[b]	1.5	0.96	0.87 (±0.03)	1.78 (±0.72)	0.70 (±0.10)
	2	0.94	0.66 (±0.02)	1.13 (±0.19)	0.67 (±0.02)
	2.25	0.94	0.61	0.86 (±0.07)	0.43
	2.5	0.94	0.62	0.93 (±0.07)	0.32
	3	0.92 (±0.02)	0.56	1.39 (±0.25)	0.06
free diffusion (MLE)	1.5	0.94	0.57	4.54 (±0.15)	0.84
	2	0.93	0.59	2.19 (±0.04)	0.64
	2.25	0.93	0.58	1.49 (±0.03)	0.50
	2.5	0.92	0.58	1.39 (±0.02)	0.32
	3	0.88	0.55	2.09 (±0.06)	0.12
free diffusion (3G)[c]	1.5	0.94	0.78 (±0.03)	1.70 (±1.76)	0.75 (±0.02)
	2	0.91	0.53 (±0.01)	2.09 (±0.69)	0.75 (±0.06)
	2.25	0.93	0.57 (±0.02)	1.52 (±0.17)	0.55 (±0.02)
	2.5	0.92	0.57	1.50 (±0.11)	0.36 (±0.01)
	3	0.88 (±0.04)	0.55	2.39 (±0.45)	0.19 (±0.03)

[a] Errors (σ) smaller than 0.01 are omitted for ε_F, ε_U, and p_F. [b] FEHs with bin times of 0.5, 1, 2, 4, and 10 ms were globally fitted to three-Gaussian distributions. [c] For the data of 1.5−3 M GdHCl, we used a bin time T = 1.08, 1.07, 1.03, 1.05, and 1.06 ms (see text).

again assuming a two-state model (see Methods). It is first instructive to examine FRET trajectories, obtained by collecting the photons in bins of various sizes at 2 M GdmCl. With the smallest bin size of 0.25 ms, the shot noise is too great to unambiguously partition the FRET trajectory into folded and unfolded segments, while using larger bins to reduce the shot noise does not result in two-state looking FRET trajectories (Figure 5). This result is expected from the Gaussian fit at 2 M GdmCl where the average waiting time in the unfolded state is determined to be 1.3 ms and the average waiting time in the folded state is 2.7 ms. Thus to resolve 90% of the waiting time distribution would require bin sizes of ~0.12 ms and therefore would require a count rate greater than ~240 photons/ms (at 30 photons per bin). It is exactly this situation where the power of the maximum likelihood method can be exploited, and we can use it to analyze photon trajectories where the arrival time of each photon is known to an accuracy of ~0.5 ns.

Figure 6 shows representative photon trajectories from the immobilization experiment that were subjected to this analysis using eq 3 of Methods. The method was modified for the free diffusion data at 3 M GdmCl to account for the potentially significant contribution of molecules with a red-shifted donor (Alexa 488^R) that produces a FRET efficiency of 0.75 for the unfolded state and could not be removed from the ensemble of trajectories as in the immobilization experiment. Since the characteristic time for the light-induced conversion to Alexa 488^R is much longer than the bin time of 1−2 ms,[6] an additional equilibrium between the states with Alexa 488^R, F^R and U^R, was introduced as

$$U \underset{k_U}{\overset{k_F}{\rightleftarrows}} F \qquad U^R \underset{k_U}{\overset{k_F}{\rightleftarrows}} F^R \tag{12}$$

The folding and unfolding rate coefficients between U^R and F^R were kept the same as those between the folded and unfolded subpopulations containing Alexa 488. During the maximization of the likelihood, the FRET efficiency of the unfolded state with Alexa 488^R was fixed to the value from the immobilization experiment ($\varepsilon_U^R = 0.75$). The FRET efficiency of the folded state with Alexa 488^R was calculated from ε_F, ε_U, and ε_U^R. In addition to ε_F, ε_U, k, and p_F, the fractional contribution of the additional equilibrium in eq 12 f^R was a fitting parameter (using the modified likelihood function of eq 1 of Methods as $L_j' = (1 - f^R)L_j + f^R L_j^R$, where L_j^R is the likelihood calculated by parameters with Alexa 488^R). f^R was found to be 0.12. The effect of Alexa 488^R on the determination of the folding and unfolding rates is negligible for the free diffusion data in other GdmCl concentrations. The extracted parameters from the maximum likelihood analysis for both the free diffusion and immobilization experiments are listed in Table 1.

Testing the Maximum Likelihood Method by "Recoloring" the Histograms. Recoloring photon trajectories and reconstructing the FEH is a way of testing the ability of the model used in the maximum likelihood analysis to accurately describe the data.[20] It is important that this method does not require us to specify all parameters that determine a photon trajectory, such as

Table 2. FRET Efficiencies and Folding Parameters Used for the Simulation of Trajectories and Parameters Obtained by the Maximum Likelihood Estimation (MLE) and Global Three-Gaussian Fitting (3G)[a] (Upper Block), and Transition Statistics in the Simulation and Obtained by the Viterbi Algorithm (Lower Block)[b]

		GdmCl				
		1.5 M	2 M	2.25 M	2.5 M	3 M
E_F	simulation	0.93	0.93	0.93	0.93	0.93
	MLE	0.93 (±0.00)	0.93 (±0.00)	0.93 (±0.00)	0.93 (±0.00)	0.93 (±0.00)
	3G	0.94 (±0.00)	0.94 (±0.00)	0.94 (±0.00)	0.94 (±0.00)	0.95 (±0.01)
E_U	simulation	0.61	0.61	0.61	0.61	0.61
	MLE	0.61 (±0.00)	0.61 (±0.00)	0.61 (±0.00)	0.61 (±0.00)	0.61 (±0.00)
	3G	0.68 (±0.15)	0.61 (±0.02)	0.61 (±0.00)	0.62 (±0.00)	0.61 (±0.00)
k (ms^{-1})	simulation	2.5	1.3	1.0	1.0	1.2
	MLE	2.36 (±0.09)	1.26 (±0.03)	0.97 (±0.03)	1.02 (±0.03)	1.20 (±0.09)
	3G	3.54 (±1.44)	1.28 (±0.14)	0.89 (±0.06)	0.84 (±0.06)	0.89 (±0.21)
p_F	simulation	0.9	0.67	0.45	0.3	0.05
	MLE	0.89 (±0.00)	0.66 (±0.01)	0.44 (±0.01)	0.31 (±0.01)	0.05 (±0.00)
	3G	0.89 (±0.04)	0.68 (±0.01)	0.45 (±0.01)	0.31 (±0.01)	0.04 (±0.01)
no. of transitions	simulation	2805	3408	2898	2492	628
	found	1320 (47%)	2131 (63%)	1890 (65%)	1686 (68%)	305 (49%)
exact transitions		480 (36%)	692 (32%)	626 (33%)	528 (31%)	105 (34%)
transitions within ±2 photons		819 (62%)	1241 (58%)	1115 (59%)	939 (56%)	170 (56%)

[a] Global three-Gaussian fitting were performed for the FRET histograms constructed by the bin times of 0.5, 1, 2, 4, 10, and 20 ms. [b] A total of 200 trajectories of 30 ms duration were simulated for each GdmCl concentrations.

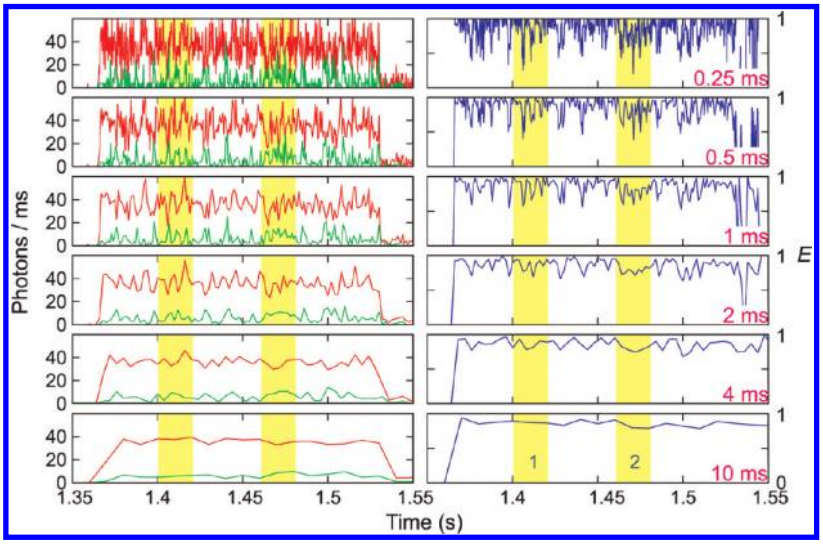


Figure 5. Binned donor (green) and acceptor (red) intensity trajectories (left) and FRET efficiency trajectories (right, blue) at different bin sizes (0.25−10 ms) from immobilization experiment. Regions shaded in yellow are shown as photon trajectories in Figure 6.

photon count rate and, in the case of diffusing molecules, the shape of the laser spot and the diffusion coefficient of the molecule. Therefore, to get accurate FRET efficiency histograms, one does not need to simulate diffusion through the laser spot. The algorithm, described in detail in ref 20, is to use eq 1 or 3 to generate a new photon trajectory from only the measured interphoton time intervals of the experimental trajectory. The recoloring is carried out (photon by photon) using probabilities calculated from the parameters of the model. A new FEH from the binned recolored trajectory is then compared to the experimental FEH. As can be seen from the comparison in Figure 7, the experimental and recolored FEHs are in excellent agreement.

Comparing Methods and Parameters. The parameters extracted from the maximum likelihood analyses and Gaussian fitting method for the solution and immobilization experiments are compared in Figure 8 and Table 1. Overall, the agreement between the FRET efficiencies, fraction folded, and rates obtained from the maximum likelihood method is excellent, indicating that surface immobilization has little effect on folding

dx.doi.org/10.1021/jp1009669 |J. Phys. Chem. A 2011, 115, 3642–3656

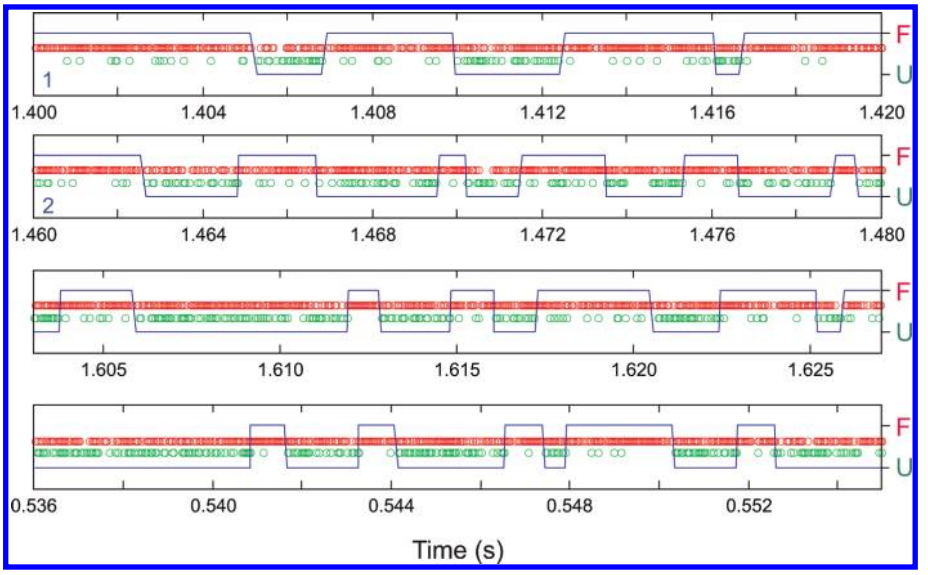


Figure 6. Photon trajectories (green, donor; red, acceptor) converted to state trajectory (blue lines) obtained from the Viterbi algorithm described in Methods. State trajectories show folded and unfolded segments and transition points. The top 2 panels (indicated by 1 and 2 on the lower left corner) are the yellow-shaded regions in Figure 5.

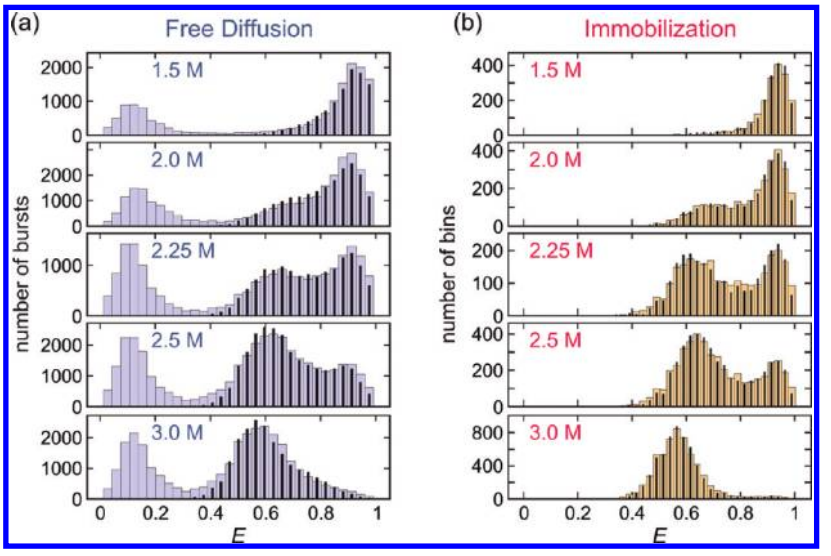


Figure 7. Recolored FRET efficiency histograms of Alexa 488/Alexa 594-labeled α_3D in the free diffusion experiment (a) and immobilization experiment (b) assuming a two-state model. Histograms with wide bars are experimental data (also shown in Figure 3), and those with narrow bars are calculated by recoloring the photons in the trajectories using parameters given in Table 1 that were obtained from the maximum likelihood method as described in the text.

and unfolding of dye-labeled α_3D. The rate coefficients from the maximum likelihood analysis for the two experiments agree to within a factor of 2 and show a roughly linear dependence in a log rate versus denaturant concentration plot (Figure 8d). The midpoint of the equilibrium unfolding curve (Figure 8b) appears near 2.25 M in both free diffusion and immobilization experiments. Impurities in GdmCl and Tween20 are detected primarily in the acceptor channel and account for about 5% of the bursts in the free diffusion experiment and therefore contribute to the differences in the extracted parameters, particularly at the lowest and highest GdmCl concentrations.

The accuracy of the extracted parameters from the maximum likelihood analysis at lower (<1.5 M) and higher (>3 M) GdmCl concentrations is much poorer. In these ranges, the fraction of either folded or unfolded molecule is too small (<5%) to obtain any meaningful parameters, presumably because of background photons (data not shown). In the free diffusion experiment, the background count rate is $b_A = 1.3$ and $b_D = 1.0$ ms^{-1} for acceptor and donor channels, respectively, which is 15% of the 30-photon threshold for a 2 ms bin. In the immobilization experiment the corresponding values are $b_A = 1.0$ and $b_D = 0.7$ ms^{-1}. Therefore, the photons emitted from a brief residence time of the smaller population can be comparable to or smaller than the background photons, which affect the photon statistics for the smaller population.

Table 1 and Figure 9 show that at 2.0, 2.25, 2.5, and 3.0 M GdmCl fitting the FEHs with three-Gaussian functions produces parameters in remarkably good agreement with those obtained

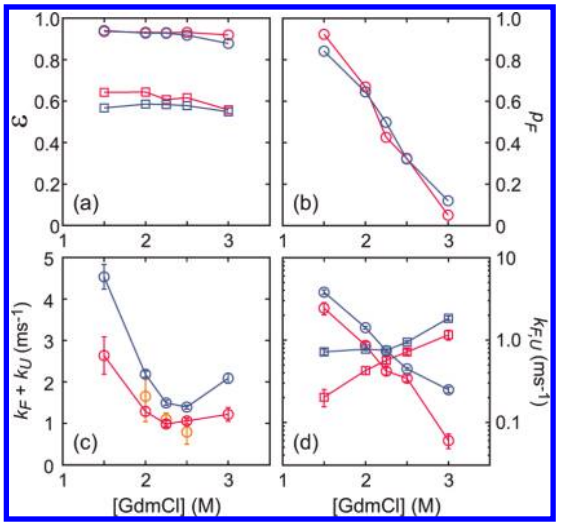


Figure 8. Parameters obtained from the maximum likelihood method in free diffusion (blue) and immobilization (red) experiments. Errors were obtained from the fitting, and error bars in (c) and (d) indicate $\pm 2\sigma$. (a) FRET efficiencies of folded (circle) and unfolded (square) states. (b) Fraction of folded molecules (p_F). (c) Relaxation rates $k_F + k_U$. Rates obtained from the donor−acceptor cross-correlation function are shown in orange. (d) Individual folding (k_F, circle) and unfolding (k_U, square) rate coefficients. The folded and unfolded rate coefficients were fitted (fit not shown) to $RT \ln k_F = RT \ln k_F^0 - m_F[\text{GdmCl}]$ and $RT \ln k_U = RT \ln k_U^0 + m_U[\text{GdmCl}]$, respectively. m_F and m_U are 1.40 and 0.66 kcal mol^{-1} M^{-1} in the immobilization experiment and 1.08 and 0.35 kcal mol^{-1} M^{-1} in the free diffusion experiment, respectively. The sums of the two kinetic m values ($m_F + m_U$) are similar to the equilibrium m value of 1.75 kcal mol^{-1} M^{-1}. Folding rate coefficients at 0 M GdmCl (k_F^0) are 99 and 56 ms^{-1} in immobilization and free diffusion experiments, respectively.

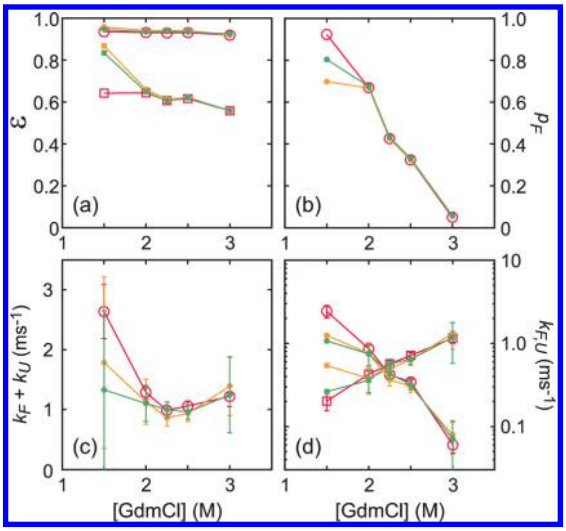


Figure 9. Parameters obtained from the maximum likelihood method (red open symbols), Gaussian fit method for FEHs in the immobilization experiment with 1 ms bin time (green filled symbols), and global Gaussian fit method for FEHs with bin times of 0.5−10 ms (orange filled symbols). (a) FRET efficiencies of folded (circle) and unfolded (square) states. (b) Fraction of folded molecules (p_F). (c) Relaxation rates $k_F + k_U$. (d) Individual folding (k_F, circle) and unfolding (k_U, square) rate coefficients. Errors were obtained from the fitting, and error bars in (c) and (d) indicate $\pm 2\sigma$.

by the more accurate maximum likelihood method. However, at the extremes of concentration (1.5 M) there is such a small contribution to the FEH (Figure 3) from the unfolded state that the parameters deviate considerably from those found by the maximum likelihood method.

As described in ref 20, there is no need to modify the methods to obtain FRET efficiencies corrected for leakage of donor photons into the acceptor channel prior to extraction of parameters by either the three-Gaussian or maximum likelihood method. These corrections for the FRET efficiency can be made after fitting using an equation, $\varepsilon^{cor} = [\langle n\rangle\varepsilon^{fit} - b_A - \chi((1 - \varepsilon^{fit})\langle n\rangle - b_D)]/(\langle n\rangle - b_A - b_D)$, where $\langle n\rangle$ is the average count rate of the sum of donor and acceptor photons (including background photons) and χ (=0.06) is the fraction of donor photons detected in the acceptor channel. For the background photon count rate listed above and the total count rate of 25 ms^{-1} (free diffusion) and 50 ms^{-1} (immobilization), the correction of the FRET efficiencies for the folded ($\varepsilon^{fit} = 0.92$) and unfolded ($\varepsilon^{fit} = 0.60$) results in 0.95 and 0.58 for the free diffusion experiment and 0.93 and 0.58 for the immobilization experiment. These FRET efficiencies can now be corrected for differences in detector sensitivity and quantum yields for the donor and acceptor (the "γ" factor) to obtain accurate FRET efficiencies for purposes of obtaining interdye distance information, an issue which is not of interest here (see refs 4 and 5).

Accuracy of Extracted Rate Coefficients. The accuracy of the extracted rate coefficients using the maximum likelihood method can be critically tested by comparing them to relaxation rates that can be obtained from a simple exponential fit to the decay of an intensity correlation function. The folding/unfolding relaxation time for the immobilized protein was calculated from the average of the donor−acceptor cross correlation functions for the individual trajectories, i.e.

$$C_{DA}(\tau) = \frac{\overline{\langle N_D(t+\tau)N_A(t)\rangle}}{\langle N_D\rangle\langle N_A\rangle} - 1 \qquad (13)$$

where $N_D(t)$ and $N_A(t)$ are the number of donor and acceptor photons in a given bin of 1 μs to 1 ms at time t, respectively. $\langle ...\rangle$ denotes an average in a given trajectory and the upper bar is an average over trajectories. For a two-state system the decay time of the donor−acceptor cross-correlation function should equal the sum of the rate coefficients. Figure 10 shows this cross correlation function at various denaturant concentrations. The correlation functions at 2.0, 2.25, and 2.5 M GdmCl can be well fit by exponential functions. The relaxation rates are plotted in Figure 8c, which shows excellent agreement with the sum of the rate coefficients from the maximum likelihood analysis. At 1.5 and 3.0 M GdmCl the amplitude is too small to obtain any meaningful decay time.

A similar comparison was not made for the free diffusion experiment because the intensity correlation function decay is on the same time scale as the folding/unfolding relaxation time. We therefore attempted to measure the relaxation rate for the dye-labeled protein in a laser temperature jump experiment using a previously described apparatus.[31] However, there was no detectable amplitude because the population changes too little in response to temperature changes near room temperature, the temperature of the single molecule measurements. Walsh et al.[13] showed that α_3D cold denatures below room temperature in the presence of GdmCl and the temperature of maximum stability is (unluckily for us) close to room temperature. We obtained very

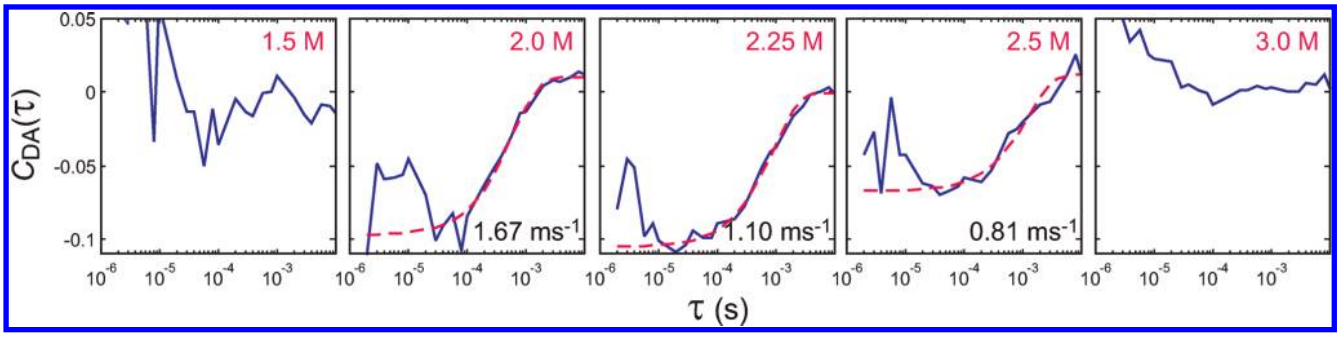


Figure 10. Donor−acceptor cross correlation function obtained from the trajectories of the immobilization experiment. The decay times were obtained from an exponential fit (dashed red line) to the data from 20 μs to 10 ms (continuous blue line). The GdmCl concentrations are indicated. The signal-to-noise at 1.5 M GdmCl and 3.0 GdmCl is too poor to obtain any meaningful fit.

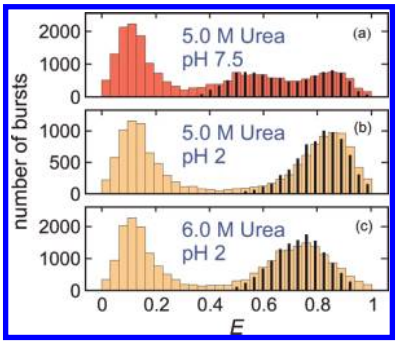


Figure 11. pH and denaturant dependence of α_3D folding (30 ms threshold, 2 ms bins). Black narrow bars were constructed by recoloring the photon trajectories as described in text. The parameters obtained from the maximum likelihood fit to the photn-by-photon trajectories are (a) $k = 1.08$ ms^{-1}, $p_F = 0.46$, $\varepsilon_F = 0.87$, $\varepsilon_U = 0.51$ (b) $k = 7.81$ ms^{-1}, $p_F = 0.68$, $\varepsilon_F = 0.92$, $\varepsilon_U = 0.57$, and (c) $k = 10.1$ ms^{-1}, $p_F = 0.48$, $\varepsilon_F = 0.91$, $\varepsilon_U = 0.57$.

similar results from CD and tryptophan fluorescence measurements on the dye-labeled protein (data not shown).

We can, however, make a qualitative comparison by comparing with ensemble experiments of Zhu et al., who used a rapid mixing method[32] to measure rates at low pH and varying urea concentrations. Figure 11 shows the results from a free diffusion experiment. The rates speed up considerably at low pH, and the folded and unfolded populations exchange so rapidly that the peaks in the FEH are merged. The maximum likelihood method yields a relaxation rate for the 6 M urea photon trajectory of 10 ms^{-1} and a folded fraction of 0.48, to give an unfolding rate coefficient of 5.2 ms^{-1} compared to 4.5 ms^{-1} of Zhu et al. obtained by interpolating between their unfolding rate coefficients at 12 and 31 °C (200 mM phosphoric acid, pH 2.6).

Resolving Transition Points in Photon Trajectories Using Viterbi Algorithm. To locate the transitions between folded and unfolded states in the photon trajectories, we applied the Viterbi algorithm described in the Methods. The input parameters are the FRET efficiencies and rate coefficients obtained from the maximum likelihood method. The Viterbi algorithm is an efficient method of finding a sequence of states from a photon string. The assumption in the application of the algorithm is that each photon belongs to either the folded or unfolded state, so that transitions between the states are assumed to be instantaneous, as was implicitly assumed in the maximum likelihood method used in the preceding analysis.

Figure 6 shows that the algorithm produces a sequence of states, consistent with the visual impression that there are many more green photons in the unfolded compared to the folded state. A more quantitative check on the algorithm can be made by calculating the output FRET efficiencies and rate coefficients, which should be identical to the input if the two-state model adequately describes the data. Figure 12 shows the distributions of the average FRET efficiencies calculated for individual molecules. After converting photon trajectories to state trajectories, the average ε_F and ε_U of each molecule were calculated by combining photons from all the folded and all the unfolded segments in each trajectory, respectively. Only average FRET efficiencies calculated from more than 300 photons were plotted in Figure 12a, and thus the ratio of the folded and unfolded molecules does not correspond to the equilibrium constant. The widths of the unfolded and folded distributions in the FEH are very close to the expected shot-noise limited widths and the mean values are almost identical to the input FRET efficiencies (Figure 12a). Furthermore, in the transition map constructed from the FRET efficiencies of the states before and after transitions (Figure 6), both F → U transitions and U → F transitions are well clustered near the center of the folded and unfolded FRET efficiencies (Figure 12b).

From the inverse of the mean waiting times in the Viterbi-assigned states, the folding and unfolding rate coefficients were obtained as shown in Figure 13. The waiting time distribution generated by the algorithm is exponential, as expected for a two-state system. However, the data in the first bin of the distributions is smaller or comparable to the second bin, which results from the algorithm omitting short-lived states (see next section). Therefore, the first bin was not included in the fitting. The rate coefficients from the mean waiting times (red) are very close to those from the maximum likelihood estimation (black).

Testing Accuracy of the Three-Gaussian Method, Maximum Likelihood Method, and Viterbi Algorithm with Simulations. To assess the accuracy of these methods, we also simulated photon trajectories with experimental parameters, as explained in Methods, and compared these input parameters with those extracted by the three-Gaussian method and the maximum likelihood method. As shown in Figure 14, the extracted rate coefficient from the three-Gaussian method varies with the bin time. The parameters obtained from the longest bin time, where there is only a single peak in the FEH, are clearly not reliable. Therefore, the FRET histograms constructed with bin times of 0.5, 1, 2, 4, 10, and 20 ms were globally fitted to obtain parameters. The results are summarized in Table 2.

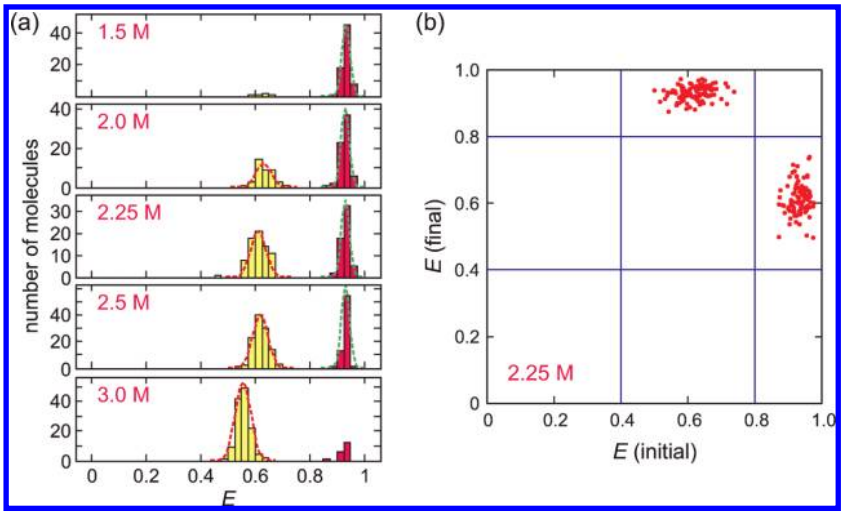


Figure 12. FRET efficiencies of the folded and unfolded states obtained using the Viterbi algorithm. (a) Photon trajectories from immobilized molecules were converted into folded and unfolded states at a single photon level using the Viterbi algorithm. For each molecule, the average ε_F and ε_U were calculated by combining photons from all the folded and all the unfolded segments in a single trajectory, respectively. Dashed lines indicate the shot-noise limited (upper bound) distributions calculated from $\sigma = [E(1-E)/(N_A + N_D)]^{1/2}$, where $N_A + N_D = 300$. The input parameters (Table 1) and mean values from the state FRET efficiencies are as follows. 1.5 M: input, $\varepsilon_U = 0.64$, $\varepsilon_F = 0.94$; output, $\langle E_U \rangle = 0.63$, $\langle E_F \rangle = 0.93$. 2.0 M: input, $\varepsilon_U = 0.64$, $\varepsilon_F = 0.93$; output, $\langle E_U \rangle = 0.63$, $\langle E_F \rangle = 0.93$. 2.25 M: input, $\varepsilon_U = 0.61$, $\varepsilon_F = 0.93$; output, $\langle E_U \rangle = 0.61$, $\langle E_F \rangle = 0.93$. 2.5 M: input, $\varepsilon_U = 0.62$, $\varepsilon_F = 0.93$; output, $\langle E_U \rangle = 0.62$, $\langle E_F \rangle = 0.93$. 3.0 M: input, $\varepsilon_U = 0.56$, $\varepsilon_F = 0.92$; output, $\langle E_U \rangle = 0.56$, $\langle E_F \rangle = 0.93$. (b) Transition map. Each dot indicates a transition from the folded to the unfolded states and vice versa. Transitions are plotted only when the segments before and after the transition contain more than 100 photons.

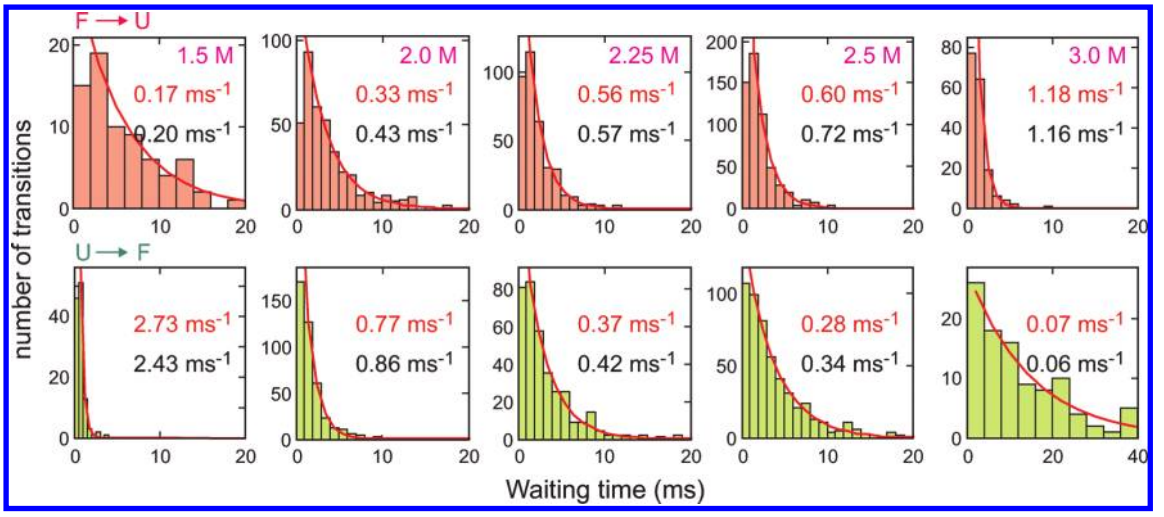


Figure 13. Waiting time distributions in folded and unfolded states obtained from the Viterbi algorithm. Unfolding and folding rate coefficients obtained from either the exponential fit (red) or the maximum likelihood method (black) are compared.

In general, the parameters obtained from the maximum likelihood estimation (upper block) are closer to the input parameters for simulations than those obtained from the three-Gaussian method. The only noticeable deviation is the relaxation rate k at 1.5 M in the maximum likelihood method, which was underestimated by 6% probably due to the fast folding rate and low unfolded population. On the other hand, the rate coefficients are more significantly overestimated or underestimated in the three-Gaussian method.

Compared to the high accuracy of determining parameters using the maximum likelihood estimation, a significant fraction of transitions were missed (half of the transitions at 1.5 and 3 M GdHCl and one-third of the transitions at 2−2.5 M GdHCl), as shown in the lower block of Table 2. As deduced from the higher missing percentage at 1.5 and 3 M GdHCl, many transitions were missed because of the low photon count rates compared to the inverse of the length of segments for the minor population. Since a fair number of transitions are not detected, the folding and unfolding rates obtained from the distributions of the waiting times shown in Figure 12 may not be accurate. However, 73% (58%) of the missed segments are shorter than 1 ms (0.5 ms), indicating that the rates obtained from the exponential fit by excluding the first bin shown in Figure 12 are only slightly affected.

Another issue in identifying transitions is the accuracy of locating the exact transition intervals by assuming instantaneous transitions. Previously, we have reported 50% accuracy to locate transition intervals for folding of protein G B1 domain.[6] However,

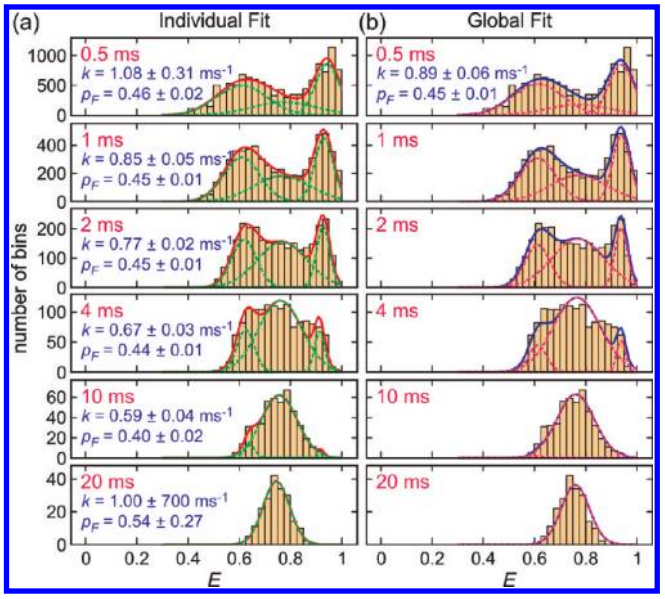


Figure 14. Extraction of parameters by three-Gaussian fitting of simulated data. FRET efficiency histograms were obtained by binning the simulated photon trajectories (2.25 M GdmCl; see Table 2 for simulation parameters) with different bin times. (a) Three-Gaussian fit of individual histograms of different bin times. Quoted relaxation rate k and fraction of folded population p_F are obtained by fitting FEH to three-Gaussian distributions calculated for a two-state model from eqs 4−7. Three individual Gaussian components and their sum are shown with dashed green and solid red curves, respectively. (b) Global three-Gaussian fit of histograms of 0.5−20 ms bin times. Three individual Gaussian components and their sum are plotted with dashed magenta and solid blue curves, respectively.

as shown in Table 2, the accuracy is just 33%. In principle, the accuracy should be high when the separation of the FRET efficiency of two states is large. The two FRET efficiencies for the two proteins are similar. For protein G, however, the photon count rate in the folded state was almost half that in the unfolded state due to mutual fluorescence quenching by the dyes and this difference could be incorporated into the search for the most probable transition interval, while only photon colors have been used for α_3D. Nevertheless, if the accuracy window is expanded to $\pm 50\ \mu$s, which corresponds to transition intervals differing by less than 3 photons, the accuracy increases to 60%.

It should be noted that the Viterbi algorithm considered here has some analogy to the intensity change point analysis by Watkins and Yang.[17] Both methods recover a state trajectory. Transition points in the intensity change point analysis are found by monitoring variation of fluorescence intensity, whereas we monitor variation in color patters. That is why the Viterbi algorithm can in principle be applied to diffusing molecules. The advantage of the intensity change point analysis is that it is model free.

DISCUSSION

The most frequent single molecule FRET experiment is one in which photons are collected from donor and acceptor fluorophores attached to a protein or other macromolecule freely diffusing through the illuminated volume of a confocal microscope system and a histogram of the FRET efficiencies for each burst is constructed.[1−3] Such experiments have been used to obtain distance information from mean FRET efficiencies of subpopulations or dynamical information of one of the subpopulations.[33] To our knowledge, there have been only a few studies in which rate coefficients have been extracted from FRET data in free diffusion experiments.[8,10,34] In these studies, rate coefficients have been obtained by comparing the FRET efficiency distributions from simulations and experiments (of course relaxation rates can be obtained using FCS methods, but equilibrium experiments are required to determine the fraction of each subpopulation to obtain individual rate coefficients).

In this work we have used two different methods of Gopich and Szabo described in refs 20 and 22 to extract folding and unfolding rate coefficients from FRET efficiency data for the fast-folding protein α_3D in both free diffusion and immobilization experiments. One of the methods uses a simple formula to fit FRET efficiency histograms (FEH) with a sum of three-Gaussian functions, while a second, more rigorous and accurate maximum likelihood method, analyzes trajectories photon by photon. The accuracy of the methods was assessed from the long trajectories of α_3D immobilized on a glass surface. Immobilization allows filtering to eliminate trajectories with photophysical artifacts such as the light induced change in the emission spectrum of the donor dye. It also allows calculation of relaxation rates by a totally independent method: the decay of the donor−acceptor cross-correlation function.

In the free diffusion experiment, the molecule resides in the confocal volume for ∼1 ms, so kinetic data can only be obtained from analysis of FEHs in this kind of experiment when the kinetics are on a time scale comparable to the bin size. The diagnostic feature that exposes folding and unfolding kinetics is the appearance of FRET efficiency at values intermediate between the folded and unfolded peaks as the bin size is increased even though the shot noise is decreasing (Figure 2a). The effect of kinetics is much more clearly demonstrated in the FEHs for the immobilized molecules (Figure 2b), where the bin size could be varied over a much wider range: from 0.5 to 20 ms. There are similarities as well as important differences with NMR exchange broadening. In the NMR experiment the peaks first broaden and move closer together in frequency as the exchange rate between the two conformations increases. For a fixed bin size and photon count rate, the positions and widths of the folded and unfolded FRET peaks do not change as the rate increases. Instead, there is a buildup of a new peak at intermediate values of the FRET efficiency as the rate increases.[22] In both NMR and FRET experiments there is a single peak when the exchange rate becomes much faster than the frequency difference or the reciprocal of the bin size, respectively. In the FRET experiment of Figure 2, the bin size changes at fixed conformational exchange and count rates, resulting in a sharpening of the folded and unfolded peaks due to decreased shot noise and the growth of an intermediate peak to produce a "Batman"-like histogram with 2 and 4 ms bin sizes, and eventually a single peak at the largest bin size of 20 ms.

Figure 3a,b, and Table 1 show the results of fitting the FEHs with a sum of three-Gaussian functions (see eqs 4−7 of Methods) for the free diffusion and immobilization experiments, assuming a two-state model. The fit at each denaturant concentration yields the mean FRET efficiencies of the folded and unfolded states and the folding and unfolding rate coefficients.

The maximum likelihood analysis of the photon trajectories (Figure 6) provides a more accurate determination of these parameters. In these trajectories the photon detection rate is too low to obtain rate coefficients from waiting time distributions, since the binned trajectories are much too noisy to assign a bin to

a state (Figure 5). It is therefore necessary to employ statistical methods, and this is exactly the situation in which maximum likelihood methods can be exploited (eq 3 of Methods). The results of applying the maximum likelihood analysis in eq 3 are presented in Table 1. An interesting test of the ability of the method to describe the data is to recolor the FEHs. In this test, the interphoton time intervals are the same as those of the experimental trajectory and the photons are colored according to the probability of a donor or acceptor photon using the parameters from the maximum likelihood analysis.[20] As shown in Figure 7, the recolored histograms virtually superimpose on the FEHs calculated by binning the photon trajectories, demonstrating the adequacy of a two-state model.

We more critically tested the accuracy of the parameters obtained from the maximum likelihood analysis by comparing the sum of the rate coefficients with the relaxation rate obtained from the decay of the donor−acceptor cross-correlation function. At 2.0, 2.25, and 2.5 M GdmCl, the agreement is excellent, with the sum of the rate coefficients found to be the same as the rate of decay of the correlation function to better than 30% (Figure 8c). Relaxation rates could not be determined at either 1.5 or 3.0 M GdmCl where the amplitudes are too small to obtain any meaningful rate. Interestingly, as judged by linear dependence of the rate coefficient on denaturant concentration (Figure 8d), the maximum likelihood method yields what appear to be reliable parameters at these concentrations where populations of the minor state are less than 10% (Table 1). A major difference, of course, between the maximum likelihood method and the FCS analysis is that the former assumes a model, while the latter does not.

The observation time is too short to make a similar comparison for the free diffusion experiment. We therefore attempted a laser temperature jump experiment, but no signal was detectable. Unluckily, the temperature of maximum stability is very close to the room temperature of our single molecule experiments, so the population change with temperature is too small. However, a comparison with the results of rapid mixing experiments on the *unlabeled* protein in urea and at low pH, where the rates increase considerably (Figure 11), shows very good agreement (see Results for details).

Table 1 and Figures 8 and 9 compare the results from the two different types of experiments and the two methods of analysis. Immobilization appears to have only a small effect, as judged by the agreement between the parameters obtained from the more accurate maximum likelihood method. The somewhat surprising result is the ability of the simple Gaussian fitting method to extract apparently accurate parameters, except at the extremes of concentration (1.5 M GdmCl), where there is such a small contribution to the FEH (Figure 3) from the unfolded state. When one of the populations falls below 5%, the maximum likelihood method also fails because background photons become comparable to the photons being emitted by the minor population.

Both maximum likelihood and Gaussian fitting methods used in this paper are based on the assumption that the total photon count rate (i.e., donor plus acceptor) is the same in the folded and unfolded states, which is valid since only a 5% difference in count rate was observed in the immobilization experiment. In addition, the Gaussian fitting method assumes that the total count rate does not change during the bin time, which is not the case in free diffusion experiments. This causes a problem in determining the appropriate bin time for the Gaussian fitting method. To address this issue, we employed an effective bin time based on the distribution of the number of photons in time bins. However, the agreement between the two methods is only approximate. In addition, we found that the extracted relaxation rates depend on the bin time in both the free diffusion and immobilization experiments. Both of these problems are not issues in the maximum likelihood method because it analyzes the trajectories photon by photon and does not require binning. Although the Gaussian fitting method works remarkably well, the maximum likelihood method is overall the more robust and accurate.

The experiments and analysis reported here are a prelude to directly observing transition paths, the actual barrier crossing process between the folded and unfolded states. We previously showed that for protein G, the transition path time is less than 200 μs, more than 10 000-fold less than the mean first passage time between folded and unfolded states at the midpoint denaturant concentration.[6] Assuming a harmonic free energy barrier with a curvature equal to that of the wells and a 1 μs preexponential factor[35] (see eq 3 of ref 6), the barrier crossing time is predicted to be 500 ns, much too short to observe with the current photon detection rate. In future experiments it should be possible to measure at least the average transition path time, by increasing the viscosity to slow the transition and analyzing a very large number of trajectories. Such an analysis will first require location of the transition path region in the photon trajectories.[36] We therefore explored the use of the Viterbi algorithm, which locates the most probable transition points assuming an instantaneous transition and converts the photon trajectory to a state trajectory (Figure 6). Analysis of simulated photon trajectories with this algorithm are a useful guide for determining how the accuracy of locating the transition path region depends on FRET efficiency differences and count rate and will therefore be helpful in designing experiments.

AUTHOR INFORMATION

Corresponding Author

*E-mail: H.S.C., chunghoi@niddk.nih.gov; W.A.E., eaton@helix.nih.gov.

ACKNOWLEDGMENT

We thank Attila Szabo for numerous helpful discussions, and Annie Aniana for technical assistance. This work was supported by the Intramural Research Program of the NIDDK, NIH.

REFERENCES

(1) Michalet, X.; Weiss, S.; Jäger, M. Single-molecule fluorescence studies of protein folding and conformational dynamics. *Chem. Rev.* **2006**, *106*, 1785.

(2) Joo, C.; Balci, H.; Ishitsuka, Y.; Buranachai, C.; Ha, T. Advances in single-molecule fluorescence methods for molecular biology. *Annu. Rev. Biochem.* **2008**, *77*, 51.

(3) Schuler, B.; Eaton, W. A. Protein folding studied by single-molecule FRET. *Curr. Opin. Struct. Biol* **2008**, *18*, 16.

(4) Best, R. B.; Merchant, K. A.; Gopich, I. V.; Schuler, B.; Bax, A.; Eaton, W. A. Effect of flexibility and cis residues in single-molecule FRET studies of polyproline. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18964.

(5) Merchant, K. A.; Best, R. B.; Louis, J. M.; Gopich, I. V.; Eaton, W. A. Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1528.

(6) 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.

(7) Chung, H. S.; Louis, J. M.; Eaton, W. A. Distinguishing between protein dynamics and dye photophysics in single molecule FRET experiments. *Biophys. J.* **2010**, *98*, 696.

(8) 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*, 22103.

(9) Hanson, J. A.; Duderstadt, K.; Watkins, L. P.; Bhattacharyya, S.; Brokaw, J.; Chu, J. W.; Yang, H. Illuminating the mechanistic roles of enzyme conformational dynamics. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 18055.

(10) Santoso, Y.; Joyce, C. M.; Potapova, O.; Le Reste, L.; Hohlbein, J.; Torella, J. P.; Grindley, N. D. F.; Kapanidis, A. N. Conformational transitions in DNA polymerase I revealed by single-molecule FRET. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 715.

(11) Margittai, M.; Widengren, J.; Schweinberger, E.; Schroder, G. F.; Felekyan, S.; Haustein, E.; Konig, M.; Fasshauer, D.; Grubmuller, H.; Jahn, R.; Seidel, C. A. M. 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*, 15516.

(12) Walsh, S. T. R.; Cheng, H.; Bryson, J. W.; Roder, H.; DeGrado, W. F. Solution structure and dynamics of a de novo designed three-helix bundle protein. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 5486.

(13) Walsh, S. T. R.; Sukharev, V. I.; Betz, S. F.; Vekshin, N. L.; DeGrado, W. F. Hydrophobic core malleability of a de novo designed three-helix bundle protein. *J. Mol. Biol.* **2001**, *305*, 361.

(14) Andrec, M.; Levy, R. M.; Talaga, D. S. Direct determination of kinetic rates from single-molecule photon arrival trajectories using hidden Markov models. *J. Phys. Chem. A* **2003**, *107*, 7454.

(15) Schröder, G. F.; Grubmüller, H. Maximum likelihood trajectories from single molecule fluorescence resonance energy transfer experiments. *J. Chem. Phys.* **2003**, *119*, 9920.

(16) Watkins, L. P.; Yang, H. Information bounds and optimal analysis of dynamic single molecule measurements. *Biophys. J.* **2004**, *86*, 4015.

(17) Watkins, L. P.; Yang, H. Detection of intensity change points in time-resolved single-molecule measurements. *J. Phys. Chem. B* **2005**, *109*, 617.

(18) Milescu, L. S.; Yildiz, A.; Selvin, P. R.; Sachs, F. Maximum likelihood estimation of molecular motor kinetics from staircase dwell-time sequences. *Biophys. J.* **2006**, *91*, 1156.

(19) Witkoskie, J. B.; Cao, J. S. Analysis of the entire sequence of a single photon experiment on a flavin protein. *J. Phys. Chem. B* **2008**, *112*, 5988.

(20) Gopich, I. V.; Szabo, A. Decoding the pattern of photon colors in single-molecule FRET. *J. Phys. Chem. B* **2009**, *113*, 10965.

(21) McKinney, S. A.; Joo, C.; Ha, T. Analysis of single-molecule FRET trajectories using hidden Markov modeling. *Biophys. J.* **2006**, *91*, 1941.

(22) Gopich, I. V.; Szabo, A. Theory of single-molecule fluorescence resonance energy transfer histograms. *Adv. Chem. Phys.*, in press.

(23) Gopich, I. V.; Szabo, A. Single-macromolecule fluorescence resonance energy transfer and free-energy profiles. *J. Phys. Chem. B* **2003**, *107*, 5058.

(24) Gopich, I. V.; Szabo, A. Theory of photon statistics in single-molecule Förster resonance energy transfer. *J. Chem. Phys.* **2005**, *122*, No. 014707.

(25) Geva, E.; Skinner, J. L. Two-state dynamics of single biomolecules in solution. *Chem. Phys. Lett.* **1998**, *288*, 225.

(26) Berezhkovskii, A. M.; Szabo, A.; Weiss, G. H. Theory of single-molecule fluorescence spectroscopy of two-state systems. *J. Chem. Phys.* **1999**, *110*, 9145.

(27) Kalinin, S.; Valeri, A.; Antonik, M.; Felekyan, S.; Seidel, C. A. M. Detection of structural dynamics by FRET: A photon distribution and fluorescence lifetime analysis of systems with multiple states. *J. Chem. Phys.* **2010**, DOI: 10.1021/jp102156t.

(28) Viterbi, A. J. Error bounds for convolution codes and an asymptotically optimum decoding algorithm. *IEEE Trans. Information Theory* **1967**, *13*, 260.

(29) Rabiner, L. R. A tutorial on hidden Markov models and selected applications in speech. *Proc. IEEE* **1989**, *77*, 257.

(30) Gopich, I. V. Concentration effects in "single-molecule" spectroscopy. *J. Phys. Chem. B* **2008**, *112*, 6214.

(31) Thompson, P. A.; Munoz, V.; Jas, G. S.; Henry, E. R.; Eaton, W. A.; Hofrichter, J. The helix-coil kinetics of a heteropeptide. *J. Phys. Chem. B* **2000**, *104*, 378.

(32) Zhu, Y.; Alonso, D. O. V.; Maki, K.; Huang, C. Y.; Lahr, S. J.; Daggett, V.; Roder, H.; DeGrado, W. F.; Gai, F. Ultrafast folding of alpha(3): A de novo designed three-helix bundle protein. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15486.

(33) Nettels, D.; Hoffmann, A.; Schuler, B. Unfolded protein and peptide dynamics investigated with single-molecule FRET and correlation spectroscopy from picoseconds to seconds. *J. Phys. Chem. B* **2008**, *112*, 6137.

(34) Pljevaljcic, G.; Millar, D. P.; Deniz, A. A. Freely diffusing single hairpin ribozymes provide insights into the role of secondary structure and partially folded states in RNA folding. *Biophys. J.* **2004**, *87*, 457.

(35) Kubelka, J.; Hofrichter, J.; Eaton, W. A. The protein folding 'speed limit'. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76.

(36) Lee, T.-H.; Lapidus, L. J.; Zhao, W.; Travers, K. J.; Herschlag, D.; Chu, S. Measuring the folding transition time of single RNA molecules. *Biophys. J.* **2007**, *92*, 3275.