Quantitative Single-Molecule Conformational Distributions: A Case Study with Poly-(L-proline)

Lucas P. Watkins, Hauyee Chang, and Haw Yang*

Department of Chemistry, University of California at Berkeley, and Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720

Received: October 14, 2005; In Final Form: February 21, 2006

Precise measurement of the potential of mean force is necessary for a fundamental understanding of the dynamics and chemical reactivity of a biological macromolecule. The unique advantage provided by the recently developed constant-information approach to analyzing time-dependent single-molecule fluorescence measurements was used with maximum entropy deconvolution to create a procedure for the accurate determination of molecular conformational distributions, and analytical expressions for the errors in these distributions were derived. This new method was applied to a derivatized poly(L-proline) series, P_nCG_3K-(biotin) ($n = 8$, 12, 15, 18, and 24), using a modular, server-based single-molecule spectrometer that is capable of registering photon arrival times with a continuous-wave excitation source. To account for potential influence from the microscopic environment, factors that were calibrated and corrected molecule by molecule include background, cross-talk, and detection efficiency. For each single poly(L-proline) molecule, sharply peaked Förster type resonance energy transfer (FRET) efficiency and distance distributions were recovered, indicating a static end-to-end distance on the time scale of measurement. The experimental distances were compared with models of varying rigidity. The results suggest that the 23 Å persistence length wormlike chain model derived from experiments with high molecular weight poly(L-proline) is applicable to short chains as well.

Introduction

Knowledge of the free energy surface on which a biological macromolecule resides allows a quantitative understanding of phenomena ranging from folding to catalysis. Its features give important clues to the dynamic structure−function relationship. In addition, accurate experimental characterizations of the free energy surface under physiological conditions provide stringent constraints for tests of theoretical models. These include the identification of conformational species, the determination of their relative population and the heights of the barriers that separate them, and the characterization of their structural flexibility, as indicated by the width of the distribution. The distribution of molecular conformations, an experimentally coarse-grained manifestation of the free energy surface, can in principle be directly measured using single-molecule fluorescence spectroscopy.[1−3]

In determining biomolecular conformational distributions from single-molecule measurements, the experimentalist is faced with the statistical uncertainties associated with low-light detection as well as with other measurement errors.[4,5] This is exacerbated in time-dependent measurements where one relies on only a few photons to determine a molecular parameter. To illustrate the challenges in this area, Figure 1 compares histograms constructed using a commonly adopted approach with the true probability density function (PDF) from a simulated single-molecule Förster resonance energy transfer (FRET) trajectory. To construct these distributions, one first chooses a time period with which to bin the trajectory, computes the distance value (or FRET efficiency) in each time bin, and chooses an interval in which to bin the distance measurements. As illustrated in Figure 1, both the choice of time bins and the distance bins will affect the shape of the distribution, potentially impacting the interpretation of experiments. While the choice of time bins has been discussed previously,[6,7] the choice of distance bins represents yet another obstacle toward realizing the full potential of single-molecule spectroscopy, measurement of the distribution of molecular properties. This article seeks to address this issue by developing a comprehensive method for the extraction of PDFs from single-molecule measurements.

When a trajectory is treated using the maximum information method (MIM), each data point has the same statistical significance.[6] This is superior to equal-time binning, where different data points may have wildly different variances. MIM treatment can thus be regarded as equal-information binning along the time trajectory. This is advantageous in probability density (histogram) estimation and is critical to the use of the maximum entropy method (MaxEnt)[10] for removal of the broadening of the histogram that occurs due to photon-counting statistics. The MaxEnt approach is employed because it offers an unprejudiced framework for extraction of the molecular conformational distribution, constrained by available information and known experimental uncertainties. Prior knowledge about the molecular system can be easily included with proper statistical weighing. When little is known about the density distribution, which is usually the case at the single-molecule level, the MaxEnt approach allows quantitative recovery of the underlying distribution without assuming any models or shapes for the unknown PDF. This is consistent with our previous development of information-based, model-free approaches to analysis of fluorescence single-molecule data.[6,7] To evaluate the accuracy of the deconvolved functions, we have also derived an analytical expression for the covariance matrix of the measured PDF.

As an experimental demonstration, we have measured distances and distance distributions in a series of poly(L-proline).

10.1021/jp055886d CCC: $33.50
Published on Web 03/29/2006

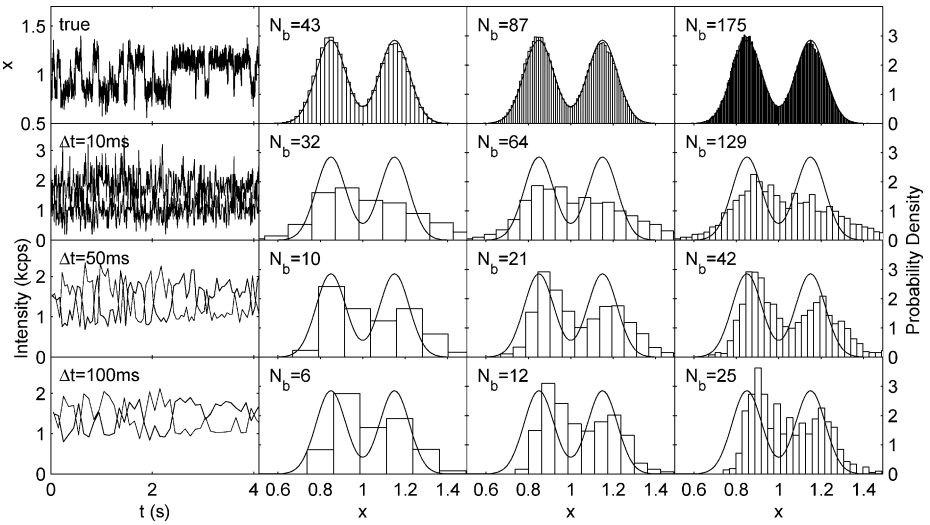

Figure 1. Comparison of histograms constructed from a simulated trajectory by constant-time binning with the true probability density (solid lines), illustrating the challenge in determining conformational distributions from single-molecule measurements. Here, $x \equiv R/R_0$ is the normalized donor−acceptor distance in a FRET measurement (cf. eq 3). The left column shows the true *x*-trajectory, as well as the donor (black) and acceptor (gray) intensity trajectories, binned at 10, 50, and 100 ms. The other three columns compare the underlying probability density (- - -) with histograms computed from the equal-time binned intensity trajectories, using N_b bins in the *x*-coordinate. The numbers of bins for the set of histograms in the third column are generated according to Scott's formula for the optimal bin width.[8] This formula, like most nonparametric density estimations,[9] assumes that there is no error associated with each datum. This assumption is not valid for single-molecule measurements. The columns to the left and right use bin numbers half and twice the optimal value, respectively. Note that different regions of the histogram are broadened differently because of the changes in the variance of the distance estimator as a function of distance.[6]

Previous studies have shown an excellent correspondence between bulk and single-molecule FRET measurements of distance.[11−13] Therefore, the current report focuses on the issues mentioned above. The constituent amino acids of poly(L-proline) are expected to exist primarily in the trans-state and should be fairly static on the time scale of single-molecule measurements, with no complex dynamics.[14] They are thus a good test case for our new approach. Indeed, we recovered narrow distance distributions for donor−acceptor distances in the polyprolines. The means of these distributions are well-explained by application of the wormlike chain (WLC) model with a persistence length of 23 Å.[15] While the short persistence length may have further implications for the structure of proline-rich signaling proteins[16]—which are beyond the scope of this article—it is hoped that the general methologies described herein will aid in the development of quantitative and predictive understanding of the dynamic behavior of biological macromolecules.

Materials and Methods

Server-Based Single-Molecule Microscope. *Microscope Construction.* The design of the microscope is diagrammed in Figure 2. The 532 nm excitation light from a continuous-wave laser—a diode-pumped solid-state laser or DPSS (Coherent, Compass 315M-100)—is passed through a cleanup filter (Chroma, HQ545/10x) and expanded to a diameter of ∼8 mm to match the back aperture of the microscope objective. To minimize sample exposure to light, a shutter is installed in the beam path and is controlled by the control server (see System Software). A polarization element is placed immediately before the entrance of the microscope; it can be a $\lambda/2$ plate, or a $\lambda/4$ plate, a Pockel cell, or any combination thereof. In the experiments reported here, a $\lambda/4$ plate was used to ensure circularly polarized excitation. After it enters the microscope, the excitation beam is reflected from a dichroic mirror (Chroma, Z532rdc) into an 60×, infinity-corrected, N.A. 1.4, oil-immersion objective (Olympus, PlanApo). The objective focuses the light to a diffraction-limited spot on the surface of the sample cover slip, which is secured onto a custom-made, temperature-regulated vacuum chuck mounted on top of a nanometer-resolution piezoelectric stage (Physik Instrumente, P734). The piezoelectric stage is driven by a high-voltage driver (Physik

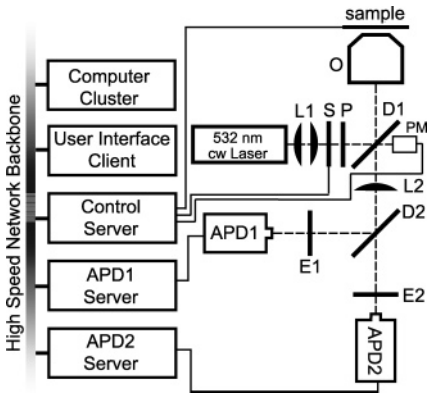

Figure 2. Configuration of a cw-excitation photon-by-photon microscope. Illumination is provided by a continuous-wave DPSS laser. Telescope lens assembly L1 expands the beam to 8 mm in order to fill the back aperture of the objective O. A shutter S is installed in the beam path to minimize unnecessary illumination. A polarization element P is placed immediately before the entrance of the home-built microscope body. Dichroic mirror D1 reflects the excitation light into the back aperture of an infinity-corrected objective O, which focuses it onto the sample mounted on a piezo stage. The excitation light that leaks through D1 is detected by a photodiode PM to monitor laser excitation power. Fluorescence from the sample is collected and collimated by objective O and passes through dichroic D1 before being focused by tube lens L2 and split into donor and acceptor channels by dichroic mirror D2. The emitted photons finally pass through emission filters E1 and E2 before being recorded by APDs APD1 and APD2. The stage and the APDs are each controlled by a separate computer server, all of which are connected via a high-speed network backbone via the TCP/IP protocol to a client computer, which runs the user interface. The modular design of the system is such that if more parameters such as polarization are to be measured, one simply drops in additional APD/server modules.

Instrumente, E509.C2 and E503.00) that is interfaced to the control server computer (see below).

Emitted light from individual molecules immobilized on the surface of the slide is collected by the same objective and passes through the dichroic mirror. It then passes through a tube lens and is separated by another dichroic mirror (Chroma, Q645LP) into donor and acceptor channels. The photons on their respective beam paths are spectrally filtered by band-pass optics (Chroma, HQ600/80m for donor emission and HQ705/130m for acceptor emission) before being focused on a pair of single-photon-counting avalanche photodiodes (APDs, Perkin-Elmer, SPCM-AQR13). Each APD is connected to a photon registration server and outputs a TTL pulse upon detecting a photon. The modular design of this microscope is such that if more parameters such as polarization were to be measured, one simply drops in additional APD/server modules and the software (see below) will take care of coordination.

The filters and dichroic mirrors used were chosen to match the absorption and fluorescence properties of the fluorophores used in these experiments. Because cross-talk can be fully treated—accounting for the changing intensities on each channel—by including the cross-talk coefficients in the signal-to-background ratio, as previously shown,[6] much broader band-pass filters can be used, allowing for more effective collection of photons.

System Software. To coordinate the complicated tasks of real-time data acquisition and analysis, the system software is split into its core functionalities, which are composed of instrument control, photon registration, and user interface. The piezoelectric stage and each APD are controlled by server programs running on separate computers. All servers are connected by a high-speed TCP/IP network to the client computer running the user interface program. The current implementation utilizes a 1 gigabit per second intranet backbone. No data stream latency was observed.

The instrument control server interfaces with the microscope via a multifunction I/O card (National Instruments, PCI-6052E) to perform a variety of control and measurement functions. These functions include setting and measurement of the position of the piezoelectric stage, control of the shutter, measurement of the laser power, and operation of any other physical components necessary to a particular experiment. The photon registration servers measure and record absolute arrival times of photons at the APDs via a timing/counting interface (National Instruments, PCI-6602). The arrival times are measured against an internal 80 MHz clock on the board, providing 12.5 ns resolution for photon arrival times, which, with the ~100 ns dead time on the APDs, is more than sufficient for the CW excitation being used. Note that the 12.5 ns time resolution refers only to the timing on individual photons, not to the time required to make a distance measurement. Theoretical limits on time resolution in the measured distance trajectories have been presented before,[6] and the experimental realization of those results is expanded upon below. To ensure that no detected photons are missed, the measured TTL stream is polled via a direct memory access channel and buffered before sending out to the client. While it is possible to run them all on the same computer, this generally results in dropped photons at high count rates. Performance is significantly improved when each APD is monitored by a different computer. For high count rate applications, the TTL pulses from the APD are monitored simultaneously by two counters on the same card. The data from each of these are then compared, and errors are corrected before the data are sent to the user interface. Tests showed that such a dual counter configuration should be used when the average count rate is higher than ~20 kilocounts per second (kcps). This configuration was tested using a pulse generator (Stanford Research, DG535) for constant-rate detection and a light-illuminated APD for exponentially distributed interphoton duration at average 10−1000 kcps. No missing photons were observed up to an average count rate of 100 kcps for ~20 s. At greater count rates, impractical for single-molecule fluorescence experiments, the setup misses one photon per 3 s at 500 kcps and seven photons per 0.3 s at 1000 kcps. The standard deviation of chronological time registration was found to be ~3−40 ns. Both autocorrelation and cross-correlation analyses were performed on these test trajectories. No correlations were found in the entire photon detection and data registration process.

The client user interface, which may be run on yet another computer, controls the operations of all of the data and control servers, as well as providing real-time data analysis capabilities by networking with a computer cluster that offers parallel computation. The entire software suite is coded in C++, and the client runs under both the Windows and the GNU/Linux operating systems.

Imaging. To locate molecules for time-resolved observation and on-line analysis, a fluorescence image of the surface that contains immobilized single molecules must be obtained. The stage is raster scanned across a predefined area, generally 10 μm × 10 μm, and photon arrival times are recorded on all channels, along with the position of the stage as a function of time. The reference times for position measurements and photon arrival times on each channel are all synchronized by a trigger pulse generated by the shutter upon opening. With these data, each photon can be assigned a specific origin on the sample cover slip. For viewing, the photon origins are spatially binned into pixels and displayed on screen. These images are generally acquired at the lowest possible excitation power to guard against irreversible photochemical reactions or photobleaching. The data in this report were collected at an excitation power of 780 nW (or 350 W/cm^2 assuming a diffraction-limited focal disk). Inspection of the two channel image allows selection of doubly labeled molecules suitable for recording single-molecule trajectories. Molecules labeled only with the acceptor will not be visible, whereas molecules labeled only with the donor or doubly labeled molecules with a bleached acceptor will be visible only on the donor channel. Correctly labeled molecules will be visible on both channels.

Time Trajectories. Once an appropriately labeled molecule is found, the stage is moved (with shutter closed) so that the selected molecule is at the focus of the objective. The counters are then armed, and the shutter is opened. As in the imaging algorithm, the opened shutter generates a pulse that simultaneously triggers the counters, synchronizing their zero times. A typical intensity trajectory is shown in Figure 3, where the characteristic bleaching pattern of a single-molecule FRET trajectory can be observed. The acceptor usually bleaches first, causing the intensity on the acceptor channel to drop to the level of the background plus cross-talk from the donor channel, while the intensity on the donor channel increases to its value in the absence of the acceptor. When the donor fluorophore bleaches, the intensities on both channels drop to their respective background levels. The high background level on the acceptor channel is caused by plastic coverwells (Molecular Probes, C18139) that were used to prevent sample evaporation. Subsequent experiments have found that the use of plastic spacers (Molecular Probes, P18178) combined with quartz covers reduces the background level to less than 200 cps. The

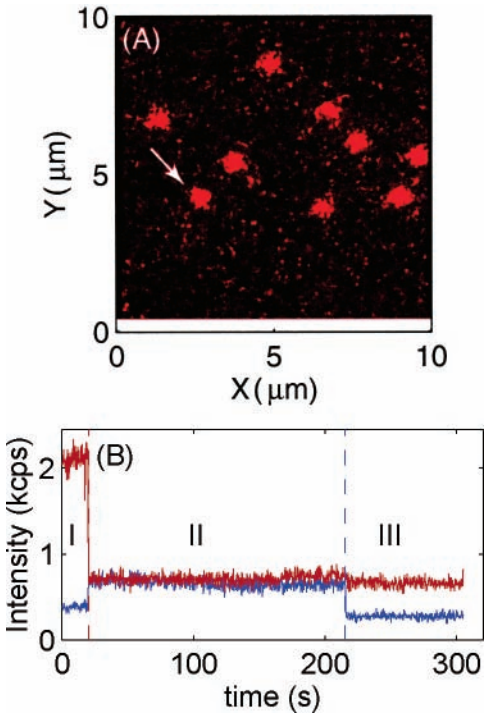

Figure 3. (A) Raster-scanned image of doubly labeled $P_{12}CG_3K$-(biotin), where the donor and acceptor emissions are represented in blue and red false colors, respectively. The image was binned at 40 nm and filtered with a 5-pixel by 5-pixel Gaussian averager. (B) A representative intensity trajectory acquired from the spot indicated by the arrow in panel A. The segmentation of different regions for data analysis is shown by the vertical dashed lines.

maximum intensity on the donor channel (region II) is lower than that on the acceptor channel because, while our filter sets were optimized to include the tail of the acceptor emission, the tail region of the donor's emission spectrum overlaps considerably with the acceptor's emission spectrum. Photons are thus much more efficiently collected from the acceptor than they are from the donor. Data acquired with this microscope may be analyzed using one of the various powerful statistical methods available.[6,7,17−29]

Sample Preparation and Characterization. A series of peptides with the sequence P_nCG_3K(biotin) (n = 8, 12, 15, 18, and 24) were synthesized using the Fmoc solid-phase synthesis technique. The C-terminal lysine was prefunctionalized with biotin on the amine (Nova Biochem, 04-12-1237). The poly-(L-proline) peptides are expected to be predominantly in the trans-isomer (forming a polyproline-II helix) under experimental conditions.[30,31] The peptides were labeled with Alexa Fluor 647 C_2-maleimide (Invitrogen/Molecular Probes, 20347) on the cysteine and Alexa Fluor 555 carboxylic acid C_5-succinimidyl ester (Invitrogen/Molecular Probes, 20009) on the N-terminal proline (cf. Figure 4). The free dye was removed by the addition of 0.2 mg/mL streptavidin (Invitrogen/Molecular Probes, S888) and subsequent centrifugal filtration. Streptavidin-bound proline−dye conjugates were retained by the filter, while free dyes passed through the filter. Unfortunately, the chemical structures of these dyes are proprietary and not available. These two fluorophores form a FRET pair with an R_0 of 51 Å, calculated from the absorption and fluorescence spectra. The ensemble-averaged steady-state anisotropies (measured on a SPEX Fluorolog) are given in Table 1. Large anisotropies are an indication that the orientation of the excited optical dipole does not have time to randomize before it relaxes back to the ground state. The fluorescence lifetime of Alexa 555 is 0.27 ns,[32] much

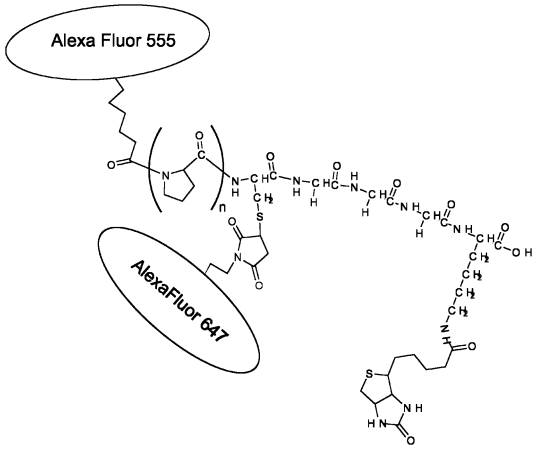

Figure 4. Chemical structure of doubly labeled P_nCG_3K(biotin), where n = 8, 12, 15, 18, and 24.

TABLE 1: Steady-State Fluorescence Anisotropy of Doubly Labeled Polyprolines, Excited at 532 nm

n	donor	acceptor
8	0.205 (9)	−0.000 (7)
12	0.244 (8)	−0.003 (6)
15	0.235 (12)	0.016 (11)
18	0.223 (11)	0.023 (14)
24	0.193 (9)	0.040 (20)

shorter than the time scale of rotation of the proline−streptavidin complex (∼10 ns). Thus, in this case, even though the acceptor's anisotropy is low, κ^2 is not in the dynamically averaged regime for single-photon emissions. However, the donor−acceptor distance measurement is made on a photon-by-photon basis over a much longer time scale, ∼1 ms, far longer than the time scales for rotation of the dyes. Averaged over this ∼1 ms time scale (which typically contains 15−25 photons), the value of κ^2 approaches the 2/3 limit. This is demonstrated in detail in a later section (cf. eq 9 and Figure 9).

The labeled peptides were immobilized via biotin−streptavidin chemistry.[33] This immobilization scheme has been shown to exhibit minimal interaction with tethered molecules.[34] Briefly, quartz cover slips (Technical Glass Products, 1 × 1 × 0.17 mm^3) were first cleaned by sequential sonication in 1 M KOH, absolute ethanol, 1 M KOH, and ethanol. They were then dried and silanized with (3-aminopropyl)trimethoxysilane (APS) by soaking for 2 min in a 2% solution of APS in acetone followed by 30 min at 110 °C. The silanized cover slips were functionalized with poly(ethylene glycol) (PEG)−SPA and PEG−biotin by incubation for 3 h in a water solution of 10% PEG−SPA, 0.1% PEG−biotin, and 0.01 M $NaHCO_3$ at pH 8.2. Finally, the streptavidin-bound fluorescently labeled peptide was incubated for 5 min on the active side of the cover slip at a concentration of ∼10 pM. The sample cover slip was then secured on the microscope for observation. No deoxygenation agents were used in the present study. More than 60 valid single-molecule trajectories were acquired for each n.

Data Analysis. *Measuring Time-Dependent FRET Efficiency and Distance Photon by Photon.* The recently developed MIM allows one to quantitatively follow single-molecule FRET efficiency and distance dynamics with the highest time resolution allowed by the information content in an experimental data set.[6] It has, for example, allowed identification of two coexisting conformations of the cdAE1 protein.[35] The time resolution (Δt) for each maximum-information measurement is determined by the expected measurement error $\alpha \equiv \delta x/x$. In a

two-channel FRET setting, it is given by

$$\Delta t = \frac{1}{\alpha^2}\left(\frac{I_d^{\beta}}{\zeta_d(x)}\left[\frac{\partial \zeta_d}{\partial x}\right]^2 + \frac{I_a^{\beta}}{\zeta_a(x)}\left[\frac{\partial \zeta_a}{\partial x}\right]^2\right)^{-1} \quad (1)$$

Here, $x = R/R_0$ is the normalized distance, R is the donor−acceptor distance, and R_0 is the Förster radius—a function of the orientation factor κ^2 that accounts for the relative orientation of the donor and acceptor transition dipoles. I_d^{β} and I_a^{β} are the maximum intensities on each channel (I_d^{β}, donor intensity in the absence of the acceptor, and I_a^{β}, acceptor intensity at $R \ll R_0$), $\zeta(x)$ is the distance-dependent intensity scaling function with $\zeta_d(x) = x^6/(1 + x^6)$ and $\zeta_a(x) = 1/(1 + x^6)$. Within the time period Δt allowed by α, the donor−acceptor energy transfer efficiency and corresponding distance are computed as

$$\hat{E} = \frac{I_d^{\beta} n_a - I_a^{\beta} n_d \beta_a^{-1}}{I_d^{\beta} n_a (1 - \beta_d^{-1}) + I_a^{\beta} n_d (1 - \beta_a^{-1})} \quad (2)$$

and

$$\hat{x} = \left(\frac{\beta_a}{\beta_d} \cdot \frac{I_d^{\beta} n_a - I_a^{\beta} n_d \beta_d}{I_a^{\beta} n_d - I_d^{\beta} n_a \beta_a}\right)^{1/6} \quad (3)$$

In the above equations, n_d and n_a are the number of photons detected within the chosen time interval on the donor and the acceptor channels, respectively, and β_d and β_a are the signal-to-background ratios. This method has been shown to be relatively robust against intermittency or transient variations in the dyes' quantum efficiency.[6] The orientation factor $\kappa^2 = 2/3$ was used for calculating the donor−acceptor distances.[36] Ensemble-averaged steady-state anisotropy measurements of doubly labeled poly(L-proline) conjugated to streptavidin indicate that the fluorescent probes have already experienced depolarization within the ~35 ns protein rotation time. The linkers used for tethering fluorescent probes to the protein surface are expected to exhibit segmental dynamics on the ultrafast to nanosecond time scales.[37,38] Together with the absence of correlation in the single-molecule time trajectories, these considerations lead to the conclusion that 2/3 is a good approximation for κ^2. This assumption was made in several recent studies, showing that accurate distance information can be obtained from immobilized DNA molecules[13] as well as from diffusing single molecules such as DNA[12] and polyproline[11,28] using FRET. These works all point to the importance of carefully considering contributions from background, cross-talk, and other instrumentation factors. Note that in addition to these parameters—time-independent cross-talk has been shown to be a form of background[6]—the correction factor for detector efficiency for the donor and acceptor channels, $(\Phi_a\eta_a)/(\Phi_d\eta_d)$, has been explicitly included in the derivation of eqs 2 and 3. This allows further correction of potential molecule-to-molecule variations in the absorption cross-section or emission spectrum as a result of heterogeneity in the microscopic environment.

Calibration. Before analysis of the photon arrival time data using the MIM algorithm, several calibration values must be determined. These include the background levels on each channel, the maximum intensities on each channel (the background levels and maximum intensities may vary from molecule to molecule), and the cross-talk coefficients (χ_d and χ_a, constant for a given experimental setup). χ_d is the fraction of donor fluorescence that will be observed on the acceptor channel. Likewise, χ_a is the fraction of acceptor fluorescence that will be observed on the donor channel. They may thus be calculated directly from the fluorescence spectra $F_d(\nu)$ and $F_a(\nu)$ of the donor and the acceptor, the transmission curves $T_d(\nu)$ and $T_d(\nu)$ of the emission filters that define the donor and acceptor channels, and the response curve $R_A(\nu)$ of the APD itself by

$$\chi_d = \frac{\int_0^{\infty} F_d(\nu)\, T_a(\nu)\, R_A(\nu)\, d\nu}{\int_0^{\infty} F_d(\nu)\, T_d(\nu)\, R_A(\nu)\, d\nu}$$

$$\chi_a = \frac{\int_0^{\infty} F_a(\nu)\, T_d(\nu)\, R_A(\nu)\, d\nu}{\int_0^{\infty} F_a(\nu)\, T_a(\nu)\, R_A(\nu)\, d\nu}$$

To calculate the other required parameters, each single-molecule time trajectory is divided into three regions, I (FRET), II (donor-only), and III (background) (cf. Figure 3). The intensity changes between regions I and II, due to acceptor photobleaching, and between II and III, due to donor photobleaching, are abrupt. Quantitative segmentation of the time trajectory was accomplished by means of an intensity change point detection algorithm, as detailed previously.[7] Depending on the relative intensities, this method allows one to determine the intensity change point to within a few photons. Regions I, II, and III, as determined by the change point algorithm, were then used to determine the calibration parameters. In region I, the period from the beginning of the trajectory to the time the acceptor photobleaches, the observed intensities on the donor and acceptor channel can be written as

$$\bar{I}_d^{(I)} = I_d^0\, \zeta_d(\bar{x}) + \chi_a\, I_a^0\, \zeta_a(\bar{x}) + B_d$$

$$\bar{I}_a^{(I)} = I_a^0\, \zeta_a(\bar{x}) + \chi_d\, I_d^0\, \zeta_d(\bar{x}) + B_a$$

where $\bar{x}$ is the (unknown) average distance in region I and $\bar{I}_d$ and $\bar{I}_a$ are the average intensities on the donor and acceptor channels. These are computed by applying the maximum likelihood estimator $\bar{I}^{(I)} = N_I/T_I$, where N_I is the total number of photons in the region and T_I is the total time duration of the region.

In region II, the period from the time that the acceptor photobleaches to the time that the donor photobleaches, the effective donor−acceptor distance is $x \to \infty$, and the observed intensities will be

$$\bar{I}_d^{(II)} = I_d^0 + B_d$$

$$\bar{I}_a^{(II)} = \chi_d\, I_d^0 + B_a$$

In region III, the period from the time the donor photobleaches and until the end of the trajectory scan, the intensities are just the background counts:

$$\bar{I}_d^{(III)} = B_d$$

$$\bar{I}_a^{(III)} = B_a$$

Solving these equations, one obtains expressions for the desired calibration values.

$$I_d^0 = \bar{I}_d^{(II)} - \bar{I}_d^{(III)}$$

$$I_a^0 = \frac{I_d^0 + \chi_d\, I_d^0 P}{P + \chi_a}$$

with $P = (\bar{I}_d^{(II)} - \bar{I}_d^{(I)})/(\bar{I}_a^{(I)} - \bar{I}_a^{(II)})$. Once the calibration parameters have been determined for an individual molecule, the photon arrival times from region I can be subjected to the MIM algorithm, generating the desired energy transfer efficiency or distance trajectory. The uncertainties in the determination of these parameters are propagated to assist in the assessment of variations in molecule-to-molecule measurements.

Photon-by-Photon Intensity Correlation. The time correlation function of a photon-by-photon single-molecule intensity trajectory can be directly calculated by representing the time-dependent intensity as a series of Dirac δ functions[20]

$$I_n(t) = \sum_{i=1}^{N} \delta(t - \tau_i^{(n)})$$

where $\{\tau_i^{(n)}\}$ is the set of photon arrival times on channel n. The true correlation function is

$$C_{nm}(t) = \langle I_m(t)\, I_n(0) \rangle - \langle I_m \rangle \langle I_n \rangle$$

where $\langle \cdots \rangle$ indicates an ensemble average.

If the single-molecule trajectory is long enough, the ensemble average may be converted to a time average

$$C_{nm}(t) = \frac{1}{(T-t)} \int_0^{T-t} I_n(\tau) \cdot I_m(t+\tau)\, \mathrm{d}\tau - \bar{I}_n \bar{I}_m$$

$$= \frac{1}{(T-t)} \sum_{i=1}^{N_n} \sum_{j=1}^{N_m} \delta(t + \tau_i^{(n)} - \tau_j^{(m)}) - \bar{I}_n \bar{I}_m$$

where T is the duration of the entire trajectory (cf. region I in Figure 3). The correlation function is just a sum of scaled δ functions, which can be computed directly from the photon arrival sequence. The correlation function $C_{mn}(t)$ may be further averaged over a time interval $[t_a, t_b]$ to reduce the stochastic noise

$$\bar{C}_{nm}(t_a, t_b) = \frac{1}{t_b - t_a} \int_{t_a}^{t_b} C_{nm}(t)\, \mathrm{d}t$$

$$= \frac{1}{(t_b - t_a)} \sum_{i=1}^{N_n} \sum_{j=1}^{N_m} \frac{\mathbf{1}_{t_a \le \tau_j^{(n)} - \tau_i^{(n)} \le t_b}}{T - (\tau_j^{(m)} - \tau_i^{(n)})} - \bar{I}_n \bar{I}_m$$

The term $\mathbf{1}_{\mathrm{expr}}$ is the indicator function, equal to 1 when expr is true and 0 otherwise.

Because the correlation function is calculated as a time average over a single trajectory, errors may arise due to incomplete sampling of the conformational space. These errors are estimated using the method of Zwanzig and Ailawadi.[39] In the averaging of correlation functions from multiple trajectories, errors may be propagated in the usual manner. This allows calculation of intensity autocorrelation and cross-correlation in FRET trajectories on a photon-by-photon basis and is analogous to most implementations of fluorescence correlation spectroscopy (FCS).[40] As such, it is straightforward to use this microscope in FCS type applications.

Recovering the Underlying PDF. Distributions measured using single-molecule fluorescence methods are commonly visualized by constructing a histogram from a binned time trajectory (averaged over every, e.g., 50 or 100 ms to reduce Poisson counting noise) and have already allowed researchers to uncover many new features in various systems,[41−46] including studies on the dynamics and folding of short peptides—one of the first treatments of the potential of mean force and photon statistics in relation to single molecule measurements—and the discovery of the dynamic equilibrium between closed and open forms of syntaxin I.[47] A quantitative assessment of the underlying PDF is therefore expected to provide further insight for the systems of interest.

When constructed from a fluorescence single-molecule time trajectory, the PDF contains contributions from both the molecular property and the photon detection statistics.[43,48,49] An information-based method such as MIM, in addition to its exact accounting of time resolution and measurement uncertainty, is advantageous for quantitative construction of the molecular distribution. Because the information content in each measurement (be it efficiency or distance) is constant, MIM can also be understood as equal-information binning, in contrast to the commonly used equal-time binning. That the information content is the same for every measurement is an important property that allows one to construct statistically robust distribution functions. This further affords model-free deconvolution to uncover the sought molecular property distribution in a least-biased, objective way. While the ideas contained in the following discussion are general, the development focuses on statistical methods that are applicable to experiments with immobilized single molecules. Such an experimental scheme can in principle provide dynamical information on a time scale covering several decades.

Gaussian Kernel Density Estimation. To estimate the distance distribution from a single-molecule time trajectory, one starts by constructing the raw experimental PDF, $\hat{r}(x)$—containing contributions from photon counting-related measurement uncertainties—from MIM-extracted distances $\hat{x}_i$. The maximum likelihood estimators $\hat{x}_i$ are asymptotically normal (Gaussian distributed) and are centered around the true but unknown distance, x_i. By virtue of the equal-information binning (cf. eq 1), each $\hat{x}_i$ has the same variance α^2.[6] This naturally leads to the use of the Gaussian kernel estimator for $r(x)$

$$\hat{r}(x) = \frac{1}{T} \sum_{i=1}^{N} \Delta t_i\, k(x; \alpha^2) \tag{4}$$

where N is the number of MIM measurements made from the trajectory and Δt_i and $\hat{x}_i$ are the duration and distance estimate from the i-th measurement.

To illustrate how experimental measurements yield overly broad density functions, raw PDFs calculated from three trajectories simulated under different conditions are compared in Figure 5 with their respective true PDFs (see Computational Validation for simulation methods). As can be seen, the raw PDF is an entirely inadequate measurement of the underlying PDF. It should be emphasized, however, that due to the statistically uniform nature of MIM measurements, these raw PDFs already represent an improvement over histograms constructed from constant-time binned trajectories. In equal-time measurements, each time bin contributes to the overall histogram with different significance levels, bringing additional bias and skewness to the resulting histogram.

Mathematically, Figure 5 can be understood by considering the raw density as the underlying molecular PDF, $h(x)$, twice convoluted with the Gaussian kernel $k(x;\alpha^2)$. One convolution is due to the measurement error in $\hat{x}$, normally distributed with a variance α^2 by virtue of maximum likelihood estimation, and one is due to use of the Gaussian kernel estimator in eq 4,

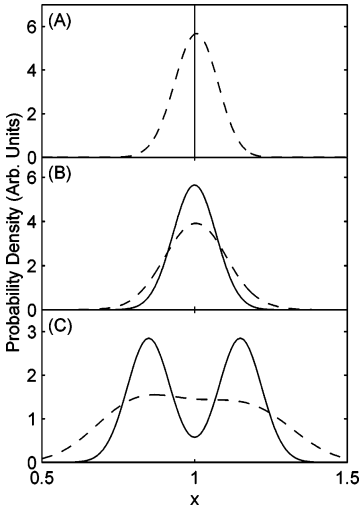

Figure 5. Comparison of true (- - -) and raw (—) PDFs for trajectories simulated (A) at constant x, (B) on a harmonic potential, and (C) on a bimodal potential.

introducing an additional variance α^2. That is

$$\hat{r}(x) = [h(x) \otimes k(x;\alpha^2)] \otimes k(x;\alpha^2)$$
$$= h(x) \otimes k(x;2\alpha^2)$$

where $\otimes$ denotes the convolution operation. The task at hand is then to recover the true molecular PDF, $h(x)$, from knowledge of the raw PDF $r(x)$ and the convolution kernel $k(x)$.

Covariance of the Raw Density. Because of the limited duration of single-molecule trajectories, the raw PDF will contain errors resulting from the nonzero relaxation time of the distance correlation function and from lack of suitable sampling of the raw histogram. Errors of the second sort can be assessed by application of Efron's bootstrap method, explained further in the Appendix. The use of the bootstrap, however, requires the data to be independent and identically distributed. In general, a single-molecule time trajectory may exhibit significant time correlation. That is, the discrete distance measurements (coarse-grained in time) made by the MIM are not necessarily independent, although they should be identically distributed.

To calculate the covariance matrix for the raw density, the notation is changed slightly from that of the previous section. Instead of writing the raw density as a weighted sum of Gaussians (cf. eq 4), it is written as an integral over time. This makes the treatment more general, since it is not constrained to situations with discrete distance measurements. Given knowledge of the estimated trajectory $\hat{x}(t)$, the raw PDF at a particular x obtained from a trajectory $\hat{x}(t)$ of duration T can be written as

$$\hat{r}(x) = \frac{1}{T}\int_0^T k[\hat{x}(t) - x,\alpha^2]\, dt$$

Note that in the case of discrete measurements, this reduces to eq 4. This PDF is estimated from a trajectory of limited duration, so it may contain statistical errors due to insufficient sampling of the conformational space, as has been discussed by Zwanzig and Ailawadi.[39] However, its ensemble average will be the true raw PDF

$$r(x) = \langle \int_0^T k[\hat{x}(t) - x,\alpha^2]\, dt \rangle$$
$$= \int_0^T \langle\, k[\hat{x}(t) - x,\alpha^2] \,\rangle\, dt$$

The second equality holds because the ensemble average and the time integral operations commute. Writing the difference between the measured raw PDF and the true raw PDF as $\delta\hat{r}(x) \equiv \hat{r}(x) - r(x)$, the covariance matrix of the raw PDF is

$$\sigma_r^2(x_1, x_2) = \langle\, \delta\hat{r}(x_1)\, \delta\hat{r}(x_2) \,\rangle$$

$$= \frac{1}{T^2}\int_0^T\int_0^T \langle\, k[\hat{x}(t_1) - x_1,\alpha^2]\, k[\hat{x}(t_2) - x_2,\alpha^2] \,\rangle -$$
$$\langle\, k[\hat{x}(t_1) - x_2,\alpha^2] \,\rangle \langle\, k[\hat{x}(t_2) - x_2,\alpha^2] \,\rangle\, dt_1\, dt_2$$

The integrand is a time correlation function and, to a very good approximation, will only be a function of $|t_2 - t_1|$. When the conformational space projected on the x-coordinate is appropriately sampled, this correlation should decay on a time scale much shorter than the length of the trajectory, T. This allows simplification to a more convenient form

$$\sigma_r^2(x_1, x_2) = \frac{2}{T}\int_0^T \langle\, k[\hat{x}(0) - x_1,\alpha^2]\, k[\hat{x}(t) - x_2,\alpha^2] \,\rangle -$$
$$\langle\, k[\hat{x}(0) - x_1,\alpha^2] \,\rangle \langle\, k[\hat{x}(t) - x_2,\alpha^2] \,\rangle\, dt \quad (5)$$

This formula is just the integral of a correlation function and is simple to evaluate. The ensemble average in the integrand may be converted to a time average for calculation of the correlation function. Given multiple trajectories from the same sample, the correlation function should be averaged across trajectories before integration. Once again, errors in this correlation function may be evaluated by the method of Zwanzig and Ailawadi.[39] To find the covariance of a raw PDF that has been averaged over multiple trajectories, T in eq 5 should be the sum of the durations of all the trajectories.

MaxEnt. To deconvolve the raw PDF, a one-dimensional form of the MaxEnt can be used.[10,50,51] A merit function $\mathcal{M}$ is constructed for a trial molecular PDF $h(x)$

$$\mathcal{M}[h(x),\lambda] = \chi^2 + \lambda H \quad (6)$$

χ^2 is a measure of the goodness-of-fit between the raw PDF and the convolution of the proposed molecular PDF

$$\chi^2 = \int_{-\infty}^{\infty} \frac{[r(x) - h(x) \otimes k(x;2\alpha^2)]^2}{\sigma_r^2(x, x)}\, dx \quad (7)$$

H is the negative entropy of the proposed molecular histogram

$$H = \int_{-\infty}^{\infty} h(x) \ln h(x)\, dx \quad (8)$$

and λ is a Lagrange multiplier, adjusted so that the final χ^2 after optimization is within $1 \pm 1/\sqrt{N}$. An initial guess of $h = \hat{r}(x)$ is used and a provisional $h(x)$ is found when $\mathcal{M}[h(x),\lambda]$ is minimized. The minimization is performed numerically using a steepest descent algorithm. The analytical gradients of the merit function are provided in the Supporting Information.

The provisional $h(x)$ is used to find the correct λ and, thus, the experimentally justified $h(x)$. This is the core concept behind MaxEnt. Given a set of underlying PDFs, all of which adequately represent the data, the one with the highest entropy is the only one that is justified by the data. While one of the lower entropy PDFs may be more correct, the data are insufficient to show this. This correct $h(x)$ can be found by varying the Lagrange multiplier λ used in the minimization until χ^2 is within the range $1 \pm 1/\sqrt{N}$.

Covariance of the Deconvolved Density. Because $r(x) = h(x) \otimes k(x)$, $\delta r(x) = \delta h(x) \otimes k(x)$ and the covariance matrix σ_h^2 of

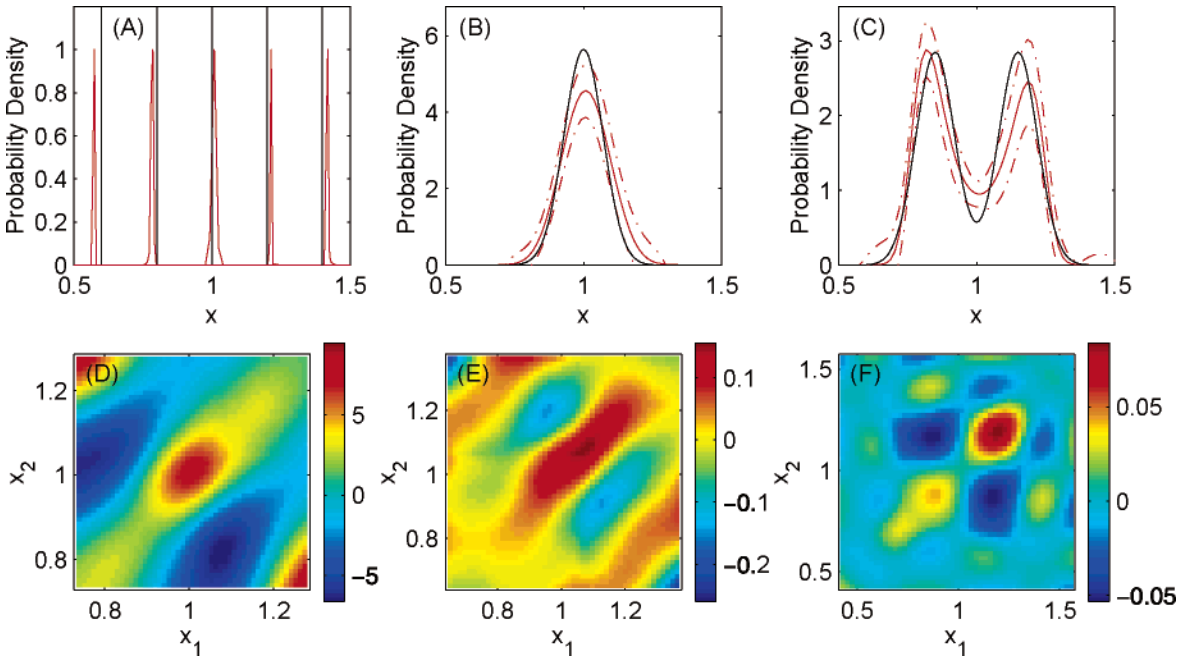

Figure 6. Results from deconvolution of test trajectories: (A) True (black) and deconvolved (red) probability densities for constant trajectories at $x = 0.6$, 0.8, 1.0, 1.2, and 1.4. (B) True (black) and deconvolved (red) probability densities for a trajectory simulated on a harmonic potential centered at $x = 1.0$. The 95% confidence interval for the deconvolved density is indicated by dashed lines. (C) True (black) and deconvolved (red) probability densities for a trajectory simulated on a bimodal potential centered at $x = 1.0$. The 95% confidence interval for the deconvolved density is indicated by broken lines. (D) Covariance matrix for the deconvolved density shown in panel A at $x = 1.0$. (E) Covariance matrix of the deconvolved density shown in panel B. (F) Covariance matrix of the deconvolved density shown in panel C.

the molecular PDF can be calculated by the two-dimensional deconvolution of the covariance matrix of the raw PDF

$$\langle \mathrm{d}r(x_1)\, \delta r(x_2) \rangle = \langle \mathrm{d}h(x_1)\, \delta h(x_2) \rangle \otimes k(x_1) \otimes k(x_2)$$

$$\sigma_r(x_1, x_2) = \sigma_h(x_1, x_2) \otimes k(x_1) \otimes k(x_2)$$

Knowledge of this covariance matrix is important for several reasons. The diagonal term gives the variance of the deconvolved PDF as a function of x, a measure of the overall and point-by-point accuracy. Just as important are the off-diagonal terms, which provide information about the relative accuracy between different regions of the density. For instance, in a trajectory from a bimodal PDF, the time-scale of equilibration between the two high-density regions will be much longer than that within the two regions. This means that the deconvolved densities at values of x within the same potential well should be accurate with respect to one another. That is, they should be positively correlated. This will be reflected by a positive covariance. The densities at values of x in different potential wells, on the other hand, should be negatively correlated, since more time spent in one potential well means less time spent in the other. This will produce a negative covariance. Thus, the covariance between two points in the PDF is primarily determined by the time-scale of equilibration between those two points. Therefore, the deconvolved covariance matrix provides further insights into the dynamics afforded by the experimentally measured PDF.

Computational Validation. To test this procedure, some basic simulations were performed. Three classes of trajectories were produced as follows: constant x, a harmonic potential, and a bimodal potential. The trajectories were produced by simulation of one-dimensional high-friction Langevin dynamics and subsequent conversion to photon arrival time data, as previously described.[6] The trajectories on harmonic potentials were simulated at a temperature of $\beta = 100$, with friction coefficient $\gamma = 1.0$. The bimodal trajectories were simulated at the same temperature, with a friction coefficient of $\gamma = 0.1$. The combined number of photons emitted before bleaching of the dyes was 2×10^5, and the signal-to-background ratio on both channels was 5.0.

As can be seen in Figure 6, the deconvolution procedure performs well. The constant−distance trajectories in Figure 6A all deconvolve to δ functions. Bias in the location of the peak is small (less than 0.023). This is expected based on the bias analysis of the MIM.[6] Trajectories from harmonic potentials produce Gaussian PDFs in Figure 6B that match the true PDF from the underlying trajectory. The deconvolved height and standard deviation of the Gaussian profile are 4.57 ± 0.70 and 0.22 ± 0.06, respectively, agreeing with the true values of 5.64 and 0.17. The covariance matrices for these deconvolved PDFs, shown in Figure 6D,E, are similar. Densities that are close to each other are positively correlated, while points farther away are negatively correlated. If the trajectory spends more time in one part of the density function, it spends less time in the other.

The deconvolution of the trajectory from a bimodal potential is more revealing. The raw PDF is too broad, and it obscures the true separation of the two potential wells. From the raw PDF (cf. Figure 5), it is not possible to measure the width of the individual peaks or the depth of the barrier that separates them. The deconvolved PDF shown in Figure 6C, on the other hand, matches well with the underlying PDF. The heights of the peaks and the depth of the well between them are 2.88 ± 0.37, 2.45 ± 0.57, and 1.71 ± 0.51, respectively; all compare well with the true values of 2.85, 2.85, and 2.28. The deconvolved covariance matrix in Figure 6F is similarly informative. Two distinct regions of positive covariance can be identified, corresponding to the two wells in the potential. The covariance between points in different wells is negative, indicating the slower time-scale of equilibration between the wells. None of these properties would be apparent from the raw histogram alone.

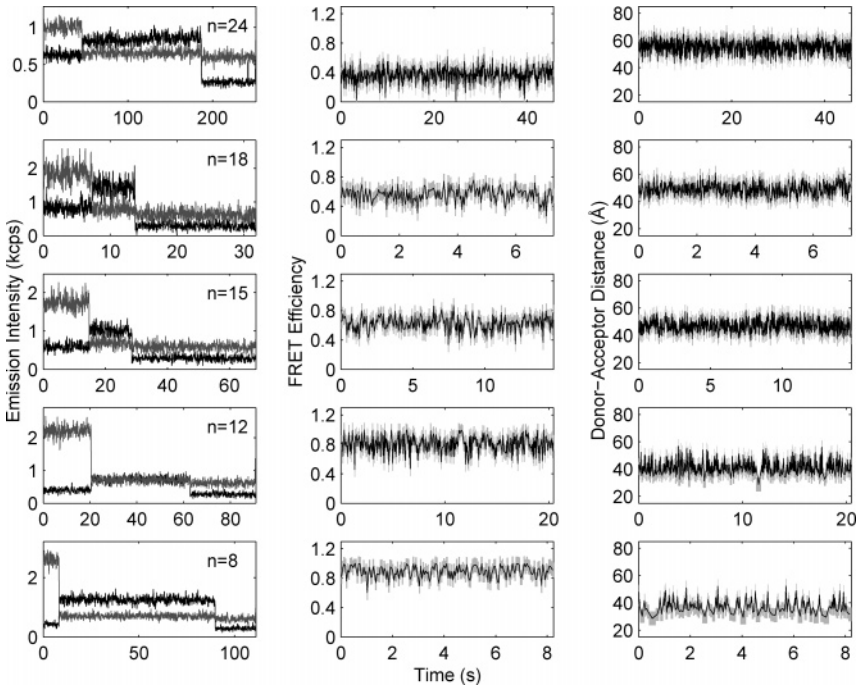

Figure 7. Intensity (left), FRET efficiency (center), and calculated donor−acceptor distance (right) as a function of time. Intensity trajectories are binned at 100 ms on both the donor (black) and the acceptor (gray) channels. Distance and efficiency trajectories are calculated using the MIM and assume $R_0 = 51$ Å. Dimensions of the gray boxes on the distance and efficiency trajectories indicate the time resolution (horizontal dimension) and expected 95% confidence interval in energy transfer efficiency or donor−acceptor distance (vertical dimension).

TABLE 2: Mean ($\bar{\tau}$) and Standard Deviation σ_τ of Donor (d) and Acceptor (a) Bleaching Times for P_nCG_3K(Biotin)

n	$\bar{\tau}_a$ (s)	$\sigma_{\tau a}$ (s)	$\bar{\tau}_d$ (s)	$\sigma_{\tau d}$ (s)	$\overline{\Delta t}$ (ms)
8	6.0	6.9	118	154	70
12	7.8	8.9	161	176	70
15	11	14	192	168	26
18	11	16	135	135	30
24	16	19	194	192	46

The differences between the true density and the recovered density can be attributed to the finite length of the trajectories. This root cause manifests itself in two ways. The accuracy of the density at a particular x value depends on how often the trajectory visits that value. Additionally, the accuracy of the covariance estimate depends on the accuracy of the correlation functions, which are themselves strongly dependent on the length of the trajectory.

Results and Discussion

Static Poly(L-Proline) End-to-End Distance on Single-Molecule Experiment Time Scales. Representative intensity trajectories for all peptide lengths are shown in Figure 7 (first column), along with the reconstructed efficiency (second column) and distance trajectories (third column) from MIM analysis with a relative error of $\alpha = 0.1$ (equivalent to a distance uncertainty of $\Delta R = 5.1$ Å). The bleaching times for both the donor (~160 s) and the acceptor (~10 s) appear roughly exponentially distributed (histograms are presented in the Supporting Information) and are summarized in Table 2. The average time resolution (Δt) is ~26−70 ms.

The fluorescence intensities appear constant over time, indicative of constant energy transfer efficiency on the time scales accessible to fluorescence single-molecule experiments. Similarly, the MIM-determined FRET efficiencies and distances appear to fluctuate randomly about their respective mean values. To examine if the extrinsic probes may transiently interact with the peptide in a nonspecific way,[52−55] correlation analyses on the fluorescence intensity, energy transfer efficiency, and distance were carried out on all of the trajectories. The lack of correlation on the experimental time scales indicates that single poly(L-proline) molecules interrogated in this study can be considered as exhibiting static mean end-to-end distances on time scales from milliseconds to tens of seconds (see Supporting Information for correlation results), in contrast to studies where the measurement time scale is comparable to that of molecular motions.[52,53] Higher time resolution data or more extended poly(L-proline) molecules may allow direct observation of conformational dynamics.

For molecules such as poly(L-proline) that presumably exhibit constant energy transfer efficiencies, one expects a sharp distribution peaking at the mean value. The raw distribution functions $\hat{r}_n(E)$ ($n = 8$, 12, 15, 18, and 24) constructed using the Gaussian kernel estimator (cf. eq 4), however, appear very broad (cf. Figure 8). This is not surprising, as they are broadened both by the photon-counting noise and by the density estimation procedure. To recover the underlying efficiency distribution of individual poly(L-proline) molecules, the MaxEnt deconvolution procedure was applied to the raw distribution functions $\hat{r}_n(E)$. As shown in Figure 8, the MaxEnt deconvolution drastically reduces the distribution to sharply peaked $\hat{h}_n(E)$, as one would have expected from poly(L-proline) molecules with time-invariant energy transfer efficiency.

Toward Quantitative FRET Measurement. *Sufficient Sampling of Donor−Acceptor Relative Orientations Using MIM.* The use of $\kappa^2 = 2/3$ in R_0 implies that orientational correlations between the donor and the acceptor dyes disappear on a time-scale shorter than the interphoton timing and that the number of photons used in distance calculations is sufficient to ensure that the distribution of κ^2 is close to normal. For the former, one examines the intensity auto- and cross-correlation functions (shown in the Supporting Information). They show no significant correlation at short time-scales, in support of this randomization assumption.

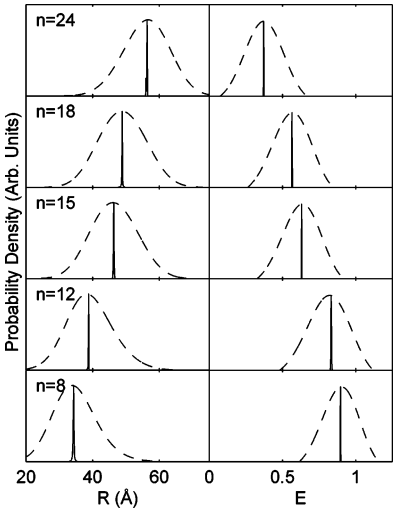

Figure 8. Raw (- - -) and deconvolved (—) distribution functions from single P_nCG_3K(biotin) trajectories shown in Figure 7.

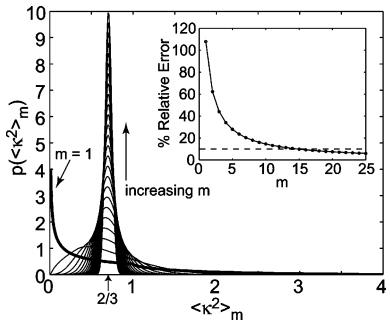

Figure 9. Probability density $p(\langle\kappa^2\rangle_m)$ as a function of the m-measurement mean value $\langle\kappa^2\rangle_m$. The bold line highlights the $m = 1$ function. As m increases, the mean value rapidly peaks at the ensemble-averaged value, $\langle\kappa^2\rangle_{m\to\infty} = 2/3$. The inset shows the relative standard deviation of $\langle\kappa^2\rangle_m$ as a function of m, which quickly decreases to less than 10% (at $m \simeq 15$).

With the operating assumption that the relative donor−acceptor orientation randomizes on a time scale much faster than interphoton timing, each detected photon can be considered as an instantaneous sampling of the PDF for κ^2.[56]

$$p(\kappa^2) = \frac{2}{\sqrt{3\kappa^2}}[\ln(2+\sqrt{3}) - g(\kappa^2)] \qquad (9)$$

where $g(\kappa^2) = 0$ when $0 < \kappa^2 < 1$ and $g(\kappa^2) = \ln(\sqrt{\kappa^2-1} + \sqrt{\kappa^2})$ when $1 < \kappa^2 < 4$. The excitation−emission cycling within a single molecule then allows repeated sampling of different relative orientations. In using a set of photons for a distance measurement, the effective κ^2 for the measurement will be the mean, $\langle\kappa^2\rangle$. As shown by the numerical study presented in Figure 9, as few as ~10 photons are required before the central limit theorem takes effect and the means approach the $\langle\kappa^2\rangle \to 2/3$ limit. This implies that if the spectra of the dyes and the refractive index of the medium between the dyes do not change appreciably over the course of the experiment, the measured distance is linearly related to the actual molecular distance in any one trajectory.

Molecule-to-Molecule Variations Are Dominated by Parameter Calibration Uncertainty. The underlying distribution of FRET efficiency and distance within individual molecules (relative distribution) can be reliably recovered with combined use of MIM and MaxEnt deconvolution. This permits one to begin discussing potential complications related to variations between molecules. The results are summarized in Figure 10.

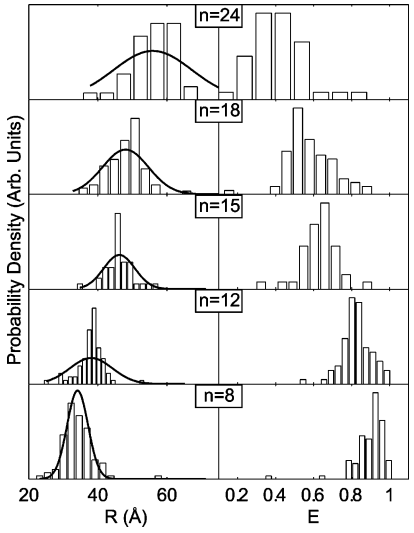

Figure 10. Distributions of donor−acceptor distances and FRET efficiency of P_nCG_3K(biotin) (n = 8, 12, 15, 18, and 24). The solid line is a Gaussian distribution with a variance that is the mean of expected variance of individual molecules by propagating uncertainties in parameters calibration. This indicates that the molecule-to-molecule distributions are dominated by uncertainties in parameter calibration.

For this data set, in general, a broad molecule-to-molecule variation is observed in the measured absolute distances $\hat{h}_n(x)$ and energy transfer efficiencies $\hat{h}_n(E)$. The width of the distance distribution for a given oligopeptide exceeds what would have been expected from statistical errors in the MIM analysis of individual trajectories. Control experiments using linearly polarized excitation light at 0, 45, and 90° at the same molecule resulted in trajectories of constant intensity within measurement uncertainties after correcting for depolarization effects in the optical components. This observation rules out the scenario in which either the donor or the acceptor probe is locked in a fixed orientation during the observation period.

Instead, it was found that the spread was dominated by variation in the observed calibration values (I_d^β, I_a^β, β_d, and β_a in eqs 3 and 2). These calibration-related uncertainties may result from variations in locating individual molecules from the single-molecule image or from variations in the immediate chemical environment of the molecule under investigation. This is visualized in Figure 10 by comparing the distributions of $\hat{h}_n(x)$ from all molecules (bars) with a Gaussian distribution (thick solid lines) having a variance of

$$\sigma^2(\hat{x}) = \frac{1}{M}\sum_{j=1}^{M}\sigma^2_{\hat{x}}(j)$$

where M is the total number of molecules of a given poly(L-proline) length and $\sigma^2_{\hat{x}}(j)$ is the expected variance of the distance measure for the j-th molecule by propagating errors in parameters calibration. More accurate measurements such as those using multispectral methods[32] will be needed in order to address issues such as the shape of the molecule-to-molecule distribution. Indeed, it will be interesting to examine the possibility that individual polyproline molecules exist in different conformations and do not interconvert on the time scale of observation. One likely physical origin is that the number of cis-residues contained in individual polyproline molecules may vary from molecule to molecule, resulting in such a broad end-to-end distance distribution. Work along this direction is underway.

Here, we will focus on the trend of the mean end-to-end distance exhibited by the series of polyprolines, discussed below.

Worm-Like Chain Model for Poly(L-proline) Molecules with a Short Persistence Length. The experimentally determined mean donor−acceptor distances may be compared with those predicted by three different models for polymer chains. In all of these comparisons, a unit length increment of 3.12 Å from C_α to C_α will be used for calculating the contour length, l_c.[57] In order of decreasing rigidity, these models are (A) a rodlike poly(L-proline), which exhibits an effective persistence length $l_p \rightarrow l_c$. This model appears to be implicitly assumed in the original paper for the use of FRET as a bulk-level spectroscopic ruler.[58] A concurrently proposed theoretical model is also consistent with this rodlike picture for short poly(L-proline) chains[59] and is, therefore, included in this category. (B) A less rigid model with the widely used persistence length of $l_p = 220$ Å for all-*trans*-poly(L-proline).[60−62] (C) A flexible model with a $l_p = 23$ Å persistence length, derived from osmometric experiments on high molecular weight poly(L-proline).[15]

These models will be discussed in the framework of a statistical description of stiff-chain polymers, the WLC model. The expected end-to-end distance, $\langle R \rangle$, of the WLC model is calculated using the mean-field expression for its probability density[63,64]

$$p(r;u) = \frac{4\pi \mathcal{N} r^2}{(1-r^2)^{9/2}} \exp\left(-\frac{3u}{4}\frac{1}{(1-r^2)}\right)$$

where $r = R/l_c$ and $u = l_c/l_p$. The normalization constant $\mathcal{N}$ is given by

$$\mathcal{N} = \frac{4s^{3/2}e^s}{\pi^{3/2}(4 + 12s^{-1} + 15s^{-2})}$$

with $s = 3u/4$. Thus, the expectation value of R is

$$\frac{\langle R \rangle}{l_c} = \frac{4\sqrt{3u}(5+u) - 2e^{3u/4}\sqrt{\pi}(-10+3u)\,\mathrm{erfc}[\sqrt{3u}/2]}{\sqrt{\pi}[20 + 3u(4+u)]} \tag{10}$$

where

$$\mathrm{erfc}[z] = \frac{2}{\pi}\int_z^\infty e^{-a^2}\,\mathrm{d}a$$

is the complementary error function.

With the Förster radius $R_0 = 51$ Å determined from the spectral overlap of the donor and acceptor probes, the only remaining parameter to be determined is the distance between the center of the emitting dipole to the C_α to which the donor or the acceptor is tethered. Unfortunately, no chemical structure is available for the Alexa Fluor 555 and 647 dyes. Nevertheless, the structures for the coupling moieties, maleimide and succinimidyl ester (cf. Figure 4), are known and can be used to estimate a lower bound for the linker distance. For this purpose, one counts nine chemical bonds from either the N or the C terminus C_α for the linkers. The linkers are also described within the framework of the WLC model, using a 6.5 Å persistence length for polymethylene as an approximation and a C−C contour increment of 1.26 Å.[65,66] Using eq 10, a lower bound of 12.2 Å for the joint linker distance was obtained. Constrained by this lower bound constraint, the WLC model with the $l_p = 23$ Å persistence length (model C) appears to describe the experimental data well, with a fitted linker distance of 11.23 Å, as summarized in Figure 11.

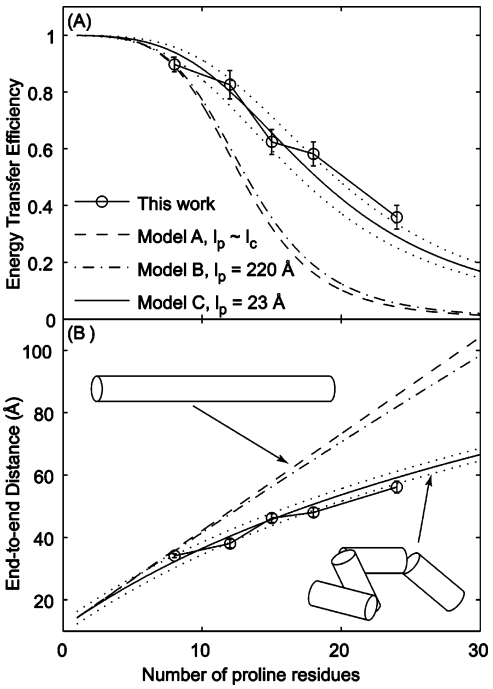

Figure 11. Comparison of experimental results with various models for poly(L-proline). As a reference, ±2 Å ranges for the model are also displayed. Experimental error bars represent 95% confidence interval of the experimental mean.

It is evident that poly(L-proline) exhibits considerable flexibility even for the relatively short chains studied here. These results are consistent with the recent studies from diffusing single molecules[11] and from NMR experiments:[67] Both found shorter-than-expected end-to-end distance if compared with a rigid polyproline model. While poly(L-proline) is believed to exist predominantly in the trans-form in room temperature aqueous solutions,[30,31] theoretical considerations indicate that the inclusion of 5% cis-residues in an otherwise trans-polyproline is sufficient to reduce the apparent persistence length significantly.[68] Therefore, it is very likely that a small number of proline residues exist in the cis-form for the short chains studied here, giving rise to the observed flexibility.

Concluding Remarks

While spectroscopy at the single-molecule level in principle allows the direct measurement of molecular property distributions, a quantitative determination of these distributions remains challenging, especially in time-dependent experiments. Uncertainties associated with low-light detection broaden and sometimes skew the experimentally obtained distribution. To address this issue, a deconvolution procedure has been developed using the distance-dependent FRET as an illustrative example.

An uniformly broadened PDF is first prepared using the previously developed maximum-information approach. This amounts to equal-information binning and ensures that every point in the underlying histogram is broadened by the same amount. Straightforward deconvolution, attempting to make the best fit possible between the experimental data and the reconvoluted PDF, produces an overfit that is not supported by experimental data. The resulting deconvolved PDF is too rough, and its features are too sharp. This leads naturally to the use of a maximum entropy-based method. It has two necessary components: statistical uniformity of the underlying data,

already provided by the MIM, and accurate knowledge of the variance in the raw experimental histogram. The analytical expressions for the variance derived in this work provide the second requirement. It should be emphasized that the calculation of this variance takes into account the time scale of dynamics in the system under observation. Furthermore, the calculation of the full covariance matrix of the deconvolved PDF allows accurate assessment of the relative heights, widths, and importance of each observed mode. Prior assumptions about the functional form of the underlying probability density of the molecular parameter are no longer necessary for its accurate calculation.

For each single poly(L-proline) molecule studied here, sharply peaked distance and FRET efficiency distributions were observed, suggesting a time-independent end-to-end distance on the time scale of fluorescence single-molecule spectroscopy. This, in turn, allows discussion of molecule-to-molecule variations in the measured distance (FRET efficiency) on more quantitative terms. It was found that these variations were dominated by uncertainties in parameter calibration. The systematic study of a series of poly(L-proline) allows one to assess models of differing rigidity. It was found that a WLC model with the $l_p = 23$ Å persistence length (derived from high molecular weight osmometry studies[15]) was in very good agreement with the present single-molecule results. While an all-*trans*-polyproline chain is expected to exhibit a persistence length much longer than oligopeptides of other composition, it has been suggested that a small percentage of cis-residues would be sufficient to allow some flexibility in the otherwise rigid chain.[68] Indeed, the presence of cis-residues cannot be ruled out in room temperature solutions. Therefore, an emerging picture for short poly(L-proline) chains is that they are composed of short trans-repeats interspersed with the occasional cis-residue. Longer time trajectories are expected to allow more detailed examination of this model and to provide insights into the nature of the molecule-to-molecule variations.

The approach presented here for recovering the underlying molecular property distribution is general and is expected to be applicable to other experimental observables. With methods such as this, an understanding of conformational features, as well as the dynamics within, may begin to be developed and placed within a quantitative, predictive theory. As an example for future applications, this approach will allow quantitative comparison of the manner by which molecular property distributions may change as a result of changes in the underlying molecule and to identify subtle yet functionally important molecular conformations.

Acknowledgment. This work was supported by the National Science Foundation, the donors of the Petroleum Research Fund of the American Chemical Society, and the University of California at Berkeley. The Office of Science, U.S. Department of Energy, is acknowledged for the use of specialized equipment (under Contract No. DE-AC03-76SF00098). L.P.W. acknowledges a graduate research fellowship from the National Science Foundation. We thank D. King for assistance with synthesis of poly(L-proline) samples used in this work, A. P. Alivisatos for the generous loan of a fluorometer, and C. C. Hayden and C.-Y. Wang for helpful discussions. We also thak I. Gopich and W. A. Eaton for a critical reading of an earlier version of the manuscript and for many helpful comments.

Appendix

Efron's Bootstrap. An alternative method for the determination of errors in the raw histogram is the bootstrap method by Efron.[69,70] Given the raw histogram, a set of auxiliary histograms is constructed by resampling each original data point from the raw histogram. The standard deviation $\sigma_B(x)$ of this set of auxiliary histograms has been shown, subject to certain assumptions, to be a good estimator of the error in the original histogram.

The assumption required by the bootstrap method is that all data points in the original histogram are independent and identically distributed. In single-molecule time trajectories, though, this assumption may not always be justified. If slow dynamics are being manifested in the trajectories being studied, the data points in the original histogram will not be independent, although they should be identically distributed. This means that if the original number of data points is used to resample the raw histogram, the standard deviation determined will be significantly lower than is justified.

This oversight can be remedied by estimation of the dominant time scales of the trajectories under consideration. In the spirit of Zwanzig's use of correlation times for calculation of errors, the number of uncorrelated distance estimates can be estimated by dividing the total duration of the trajectory by the $1/e$ time of the correlation function. The bootstrapped error calculated from this number of independent points is generally comparable with the error calculated by the analytical method described in the main text and may be more expedient in situations where the analytical approach cannot be applied or where fast calculations are required.

Supporting Information Available: Analytical gradients of the maximum entropy merit function, dye bleaching lifetimes, time resolution, and correlation analysis. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Weiss, S. *Science* **1999**, *283*, 1676−1683.
(2) Moerner, W.; Orrit, M. *Science* **1999**, *283*, 1670−1676.
(3) Xie, X. S.; Trautman, J. K. *Annu. Rev. Phys. Chem.* **1998**, *49*, 441−480.
(4) Jung, Y.; Barkai, E.; Silbey, R. *J. Chem. Phys.* **2002**, *117* (24), 10980−10995.
(5) Lippitz, M.; Kulzer, F.; Orrit, M. *ChemPhysChem* **2005**, *6*, 770−789.
(6) Watkins, L. P.; Yang, H. *Biophys. J.* **2004**, *86*, 4015−4029.
(7) Watkins, L. P.; Yang, H. *J. Phys. Chem. B* **2005**, *109*, 617−628.
(8) Scott, D. W. *Biometrika* **1979**, *66* (3), 605−610.
(9) Silverman, B. W. *Density Estimation for Statistics and Data Analysis*; Chapman & Hall/CRC: New York, 1998.
(10) Jaynes, E. T. *Proc. IEEE* **1982**, *70* (9), 939−952.
(11) Schuler, B.; Lipman, E. A.; Steinbach, P. J.; Kumke, M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 2754−2759.
(12) Lee, N. K.; Kapanidis, A. N.; Wang, Y.; Michalet, X.; Mukhopadhay, J.; Ebright, R. H.; Weiss, S. *Biophys. J.* **2005**, *88*, 2939−2953.
(13) Sabanayagam, C. R.; Eid, J. S.; Meller, A. *J. Chem. Phys.* **2005**, *122*, 061103.
(14) Gornick, F.; Mandelkern, L.; Diorio, A. F.; Roberts, D. E. *J. Am. Chem. Soc.* **1971**, *93*, 1769−1777.
(15) Mattice, W. L.; Mandelkern, L. *J. Am. Chem. Soc.* **1971**, *93*, 1769−1777.
(16) Kay, B. K.; Williamson, M. P.; Sudol, M. *FASEB* **2000**, *14* (2), 231−241.
(17) Fries, J. R.; Brand, L.; Eggeling, C.; Köllner, M.; Seidel, C. A. M. *J. Phys. Chem. A* **1998**, *102*, 6601−6613.
(18) Novikov, E.; Hofkens, J.; Cotlet, M.; Maus, M.; De Schryver, F. C.; Boens, N. *Spectrochim. Acta* **2001**, *57* (11), 2109−2133.
(19) Enderlein, J.; Sauer, M. *J. Phys. Chem. A* **2001**, *105*, 48−53.
(20) Yang, H.; Xie, X. S. *J. Chem. Phys.* **2002**, *117*, 10965−10979.
(21) Barsegov, V.; Mukamel, S. *J. Chem. Phys.* **2002**, *116*, 9802−9810.
(22) Schröder, G. F.; Grubmüller, H. *J. Chem. Phys.* **2003**, *119*, 7830−7834.
(23) Andrec, M.; Levy, R. M.; Talaga, D. S. *J. Phys. Chem. A* **2003**, *107* (38), 7454−7464.
(24) Laurence, T. A.; Kapanidis, A. N.; Kong, X.; Chemla, D. S.; Weiss, S. *J. Phys. Chem. B* **2004**, *108*, 3051−3067.

(25) Witkoskie, J. B.; Cao, J. S. *J. Chem. Phys.* **2004**, *121*, 6361−6372.
(26) Witkoskie, J. B.; Cao, J. S. *J. Chem. Phys.* **2004**, *121*, 6373−6379.
(27) Enderlein, J.; Goodwin, P. M.; van Orden, A.; Ambrose, W. P.; Erdmann, R.; Keller, R. A. *Chem. Phys, Lett.* **1997**, *270*, 464−470.
(28) Gopich, I.; Szabo, A. *J. Chem. Phys.* **2005**, *122*, 014707.
(29) Gopich, I.; Szabo, A. *J. Phys. Chem. B* **2005**, *109*, 6845−6848.
(30) Harrington, W. F.; Sela, M. *Biochim. Biophys. Acta* **1958**, *27*, 24−41.
(31) Steinberg, I. Z.; Harrington, W. F.; Berger, A.; Sela, M.; Katchalski, E. *J. Am. Chem. Soc.* **1960**, *82*, 5263−5279.
(32) Luong, A. K.; Gradinaru, C. C.; Chandler, D. W.; Hayden, C. C. *J. Phys. Chem. B* **2005**, *109* (33), 15691−15698.
(33) Ha, T.; Rasnik, I.; Cheng, W.; Babcock, H. P.; Gauss, G. H.; Lohman, T. M.; Chu, S. *Nature* **2002**, *419*, 638−641.
(34) Okumus, B.; Wilson, T. J.; Lilley, D. M. J.; Ha, T. *Biophys. J.* **2004**, *87*, 2798−2806.
(35) Pal, P.; Lesoine, J. F.; Lieb, M. A.; Novotny, L.; Knauf, P. A. *Biophys. J.* **2005**, in press.
(36) dos Remedios, C. G.; Moens, P. D. J. *J. Struct. Biol.* **1995**, *115*, 175−185.
(37) Budzien, J.; Raphael, C.; Ediger, M. D.; de Pablo, J. J. *J. Chem. Phys.* **2002**, *116*, 8209−8217.
(38) He, Y.; Lutz, T. R.; Ediger, M. D.; Ayyagari, C.; Bedrov, R.; Smith, G. D. *Macromolecules* **2004**, *37*, 5032−5039.
(39) Zwanzig, R.; Ailawadi, N. K. *Phys. Rev.* **1969**, *182*, 280−283.
(40) Magde, D.; Elson, E.; Webb, W. W. *Phys. Rev. Lett.* **1972**, *29*, 705−708.
(41) Xie, Z.; Srividya, N.; Sosnick, T. R.; Pan, T.; Scherer, N. F. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101* (2), 534−539.
(42) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419* (6908), 743−747.
(43) Talaga, D. S.; Lau, W. L.; Roder, H.; Tang, J. Y.; Jia, Y. W.; DeGrado, W. F.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97* (24), 13021−13026.
(44) Brasselet, S.; Peterman, E. J. G.; Miyawaki, A.; Moerner, W. E. *J. Phys. Chem. B* **2000**, *104* (15), 3676−3682.
(45) Yang, H.; Luo, G.; Karnchanaphanurach, P.; Louie, T.-M.; Xun, L.; Xie, X. S. *Science* **2003**, *302*, 262−266.
(46) Slaughter, B. D.; Unruh, J. R.; Allen, M. W.; Urbauer, R. J. B.; Johnson, C. K. *Biochemistry* **2005**, *44* (10), 3694−3707.
(47) Margittai, M.; Widengren, J.; Schweinberger, E.; Schröder, G. F.; Felekyan, S.; Haustein, E.; König, M.; Fasshauer, D.; Grubmüller, H.; Jahn, R.; Seidel, C. A. M. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100* (26), 15516−15521.
(48) Lee, M.; Tang, J. Y.; Hochstrasser, R. M. *Chem. Phys. Lett.* **2001**, *344*, 501−508.
(49) Lee, M.; Kim, J.; Tang, J.; Hochstrasser, R. M. *Chem. Phys. Lett.* **2002**, *359*, 412−419.
(50) Jaynes, E. T. *Phys. Rev.* **1957**, *106*, 620−630.
(51) Jaynes, E. T. *Phys. Rev.* **1957**, *108*, 171−190.
(52) Jia, Y.-w.; Sytnik, A.; Li, L.; Vladimirov, S.; Cooperman, B. S.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 7932−7936.
(53) Geva, E.; Skinner, J. L. *Chem. Phys. Lett.* **1998**, *288* (2−4), 225−229.
(54) Edman, L.; Mets, U.; Rigler, R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6710−6715.
(55) Buschmann, V.; Weston, K. D.; Sauer, M. *Bioconjugate Chem.* **2003**, *14*, 195−204.
(56) Dale, R. E.; Eisinger, J.; Blumberg, W. E. *Biophys. J.* **1979**, *26*, 161−194.
(57) Cowan, P. M.; McGavin, S. *Nature* **1955**, *176*, 501−503.
(58) Stryer, L.; Haugland, R. P. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 719−726.
(59) Schimmel, P. R.; Flroy, P. J. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 52−59.
(60) Brant, D. A.; Flory, P. J. *J. Am. Chem. Soc.* **1965**, *87*, 2788−2791.
(61) Brant, D. A.; Flory, P. J. *J. Am. Chem. Soc.* **1965**, *87*, 2791−2800.
(62) Flory, P. J. *Statistical Mechanics of Chain Molecules*; John Wiley & Sons: New York, 1969.
(63) Bhattacharjee, J. K.; Thirumalai, D.; Bryngelson, J. D. arXiv: cond-mat/9709345, 1997.
(64) Grosberg, A., Ed. *Theoretical and Mathematical Models in Polymer Research*; Academic Press: New Yrok, 1998.
(65) Yamakawa, H. *Annu. Rev. Phys. Chem.* **1984**, *35*, 23−47.
(66) Harnau, L.; Winkler, R. G.; Reineker, P. *Europhys. Lett.* **1999**, *45*, 488−494.
(67) Jacob, J.; Baker, B.; Bryant, R. G.; Cafiso, D. S. *Biophys. J.* **1999**, *77*, 1086−1092.
(68) Tanaka, S.; Scheraga, H. A. *Macromolecules* **1975**, *8*, 623−631.
(69) Efron, B. *Ann. Stat.* **1979**, *7*, 1−26.
(70) Efron, B.; Gong, G. *Am. Stat.* **1983**, *37*, 36−48.