

# Protein Structure and Dynamics from Single-Molecule Fluorescence Resonance Energy Transfer

Dong Wang and Eitan Geva\*

*Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109-1055*

*Received: May 17, 2004; In Final Form: October 12, 2004*

The pros and cons of single-molecule vs ensemble-averaged fluorescence resonance energy transfer (FRET) experiments, performed on proteins, are explored with the help of Langevin dynamics simulations. An off-lattice model of the polypeptide chain is employed, which gives rise to a well-defined native state and two-state folding kinetics. A detailed analysis of the distribution of the donor–acceptor distance is presented at different points along the denaturation curve, along with its dependence on the averaging time window. We show that unique information on the correlation between structure and dynamics, which can only be obtained from single-molecule experiments, is contained in the correlation between the donor–acceptor distance and its displacement. The latter is shown to provide useful information on the free energy landscape of the protein, which is complementary to that obtained from the distribution of donor–acceptor distances.

## 1. Introduction

In recent years, single-molecule spectroscopy has emerged as a powerful tool for studying the structure and dynamics of biological and synthetic polymers.<sup>1–27</sup> Single-molecule fluorescence resonance energy transfer<sup>28–32</sup> (SM-FRET), in particular, has been introduced as a unique probe of conformational dynamics in proteins.<sup>8,14–18,26</sup> In those SM-FRET experiments, one measures the variation over time of the efficiency of energy transfer between a donor dye molecule and an acceptor dye molecule, which label specific sites on the protein. Those variations are then interpreted as reflecting the conformational dynamics of the protein, since the efficiency of energy transfer is very sensitive to the donor–acceptor distance.

Hochstrasser and co-workers have pioneered the application of single-molecule FRET to study protein folding.<sup>14,15</sup> Those authors used SM-FRET in order to explore the folding dynamics of a disulfide cross-linked two-stranded coiled coil from the yeast transcription factor GCN4. In those studies, the protein heterodimers were either surface-immobilized<sup>14,15</sup> or freely diffusing.<sup>15</sup> Distributions of the donor–acceptor distance were measured at three points along the denaturation curve, which correspond to folded, unfolded, and an intermediate state. Besides a clear division into folded and unfolded subpopulations, whose relative weights depend on the denaturant concentration, those authors have also been able to extract information on the correlation between structure and dynamics from the dependence of the distribution of the donor–acceptor distance on the size of the averaging time window as well as from the correlation between the donor–acceptor distance and its displacement (cf. section 2). Weiss and co-workers have used SM-FRET in order to monitor the change in the folded and unfolded subpopulations of the protein chymotrypsin inhibitor 2 (CI2) as a function of denaturant concentration.<sup>16,33</sup> The experiment was performed on freely diffusing protein molecules, and its time resolution was restricted by the relatively short period of time (~milliseconds) during which the molecule is exposed to the laser. A similar technique was also applied to the cold-shock protein (Csp) by Eaton and co-workers.<sup>17</sup> In both cases, the distribution of FRET efficiency has been clearly divided into two compo-

nents, corresponding to the folded and unfolded subpopulations, whose relative weights were dependent on the denaturant concentration. Moerner and co-workers have used FRET between two green fluorescent protein (GFP) mutants to explore the conformational changes of Calmodulin (a Ca<sup>2+</sup> binding protein) as a function of Ca<sup>2+</sup> concentration.<sup>20</sup> In this case, the GFP–Calmodulin construct was immobilized in an agarose gel. Haran and co-workers have used a vesicle-encapsulation technique in order to monitor the distribution of folding and unfolding pathways of single adenylylase (AK) molecules in real time.<sup>18,34</sup> They observed pathways of different types, which were characterized by either rapid or slow transitions between the folded and unfolded states.

The interpretation of SM-FRET experiments has been the subject of several recent theoretical and computational studies.<sup>35–40</sup> Bagchi and co-workers have used Brownian dynamics simulations in order to investigate the distribution of the FRET efficiency at equilibrium and during the collapse of a freely jointed homopolymer,<sup>35,38</sup> the dependence of the end-to-end distance distribution on the stiffness of the polymer,<sup>37</sup> and the FRET survival probability within the context of the Rouse chain model.<sup>36</sup> Yang and Cao have derived analytical expressions for the end-to-end distance distribution as well as the correlation between the end-to-end distance and its displacement in the case of a semiflexible Gaussian chain model.<sup>39</sup> Gopich and Szabo have presented a theoretical analysis of how the end-to-end distance distribution, and the potential of mean force derived from it, depend on the time averaging window.<sup>40</sup> Those authors have also used their theory in order to demonstrate the effect of time averaging in the cases of a Gaussian chain model and a kinetic model for protein folding on a reduced 2D free energy surface.<sup>40</sup>

In the present paper, we employ Langevin dynamics simulations of an off-lattice model protein, with a well-defined native state, to further understand the relationship between SM-FRET observables and the underlying conformational dynamics. The main goal of this study is to explore the unique ability of SM-FRET to probe dynamical aspects associated with the equilibrium fluctuations of the protein at different points along the

denaturation curve. Special attention was also given to the correlation between structure and dynamics.

The organization of the paper is as follows. Section 2 reviews the various observables probed in ensemble-averaged FRET and SM-FRET experiments and the relationship between them. Section 3 describes the model of the polypeptide chain and simulation techniques. The results obtained from the simulations are presented in section 4 and discussed in relation to experiment and previous theoretical work in section 5. The main conclusions are summarized in section 6.

## 2. Single-Molecule vs Ensemble-Averaged FRET

We consider a macromolecule, labeled by a pair of donor–acceptor dye molecules at two specific sites. In a typical FRET experiment, a photoexcited donor either fluoresces or non-radiatively transfers its energy to the acceptor, which then fluoresces at a different (longer) wavelength. The rate constant for nonradiative FRET from donor to acceptor is assumed to be given by the Förster theory, namely<sup>31,41,42</sup>

$$k_{\text{ET}}(R) = k_{\text{D}} \left( \frac{R_0}{R} \right)^6 \quad (1)$$

Here,  $k_{\text{D}}^{-1}$  is the fluorescence lifetime of the free donor,  $R$  is the center-to-center distance between donor and acceptor, and  $R_0$  is a parameter that depends on the temperature, the refractive index of the intervening medium, the donor's fluorescence quantum yield, and the overlap between the donor's emission spectrum and the acceptor's absorption spectrum (the explicit forms of  $R_0$  can be found elsewhere, e.g. in ref 43). The validity of eq 1 is based on the following assumptions, which are believed to be rather reliable in many cases of practical interest:<sup>43</sup> (1) dipole–dipole coupling between the donor and acceptor; (2) weak coupling between the donor and acceptor, in the sense of Fermi's golden rule; (3) homogeneous broadening of donor and acceptor spectra; (4) the wavelength of the exchanged virtual photon is much larger than the donor–acceptor distance; (5) random dipole–dipole orientations; (6) fast orientational dynamics of the dye dipoles on the time scale of the fluorescence lifetime.

In a time-resolved ensemble-averaged FRET experiment, one measures the donor fluorescence, following the simultaneous photoexcitation of a large number of donor molecules.<sup>44</sup> The ensemble-averaged fluorescence decay can then be described by

$$\bar{I}(\tau) = \int_0^\infty dR P(R) e^{-k_{\text{D}}[1 + (R_0/R)^6]\tau} \quad (2)$$

Here,  $P(R)$  corresponds to the probability density of  $R$  at the time of excitation, and  $(k_{\text{D}}[1 + (R_0/R)^6])^{-1}$  is the  $R$ -dependent lifetime of the donor's excited state. It is important to note that  $R$  changes on the time scales of microseconds, which is considerably longer than the donor's fluorescence lifetime (nanoseconds), thereby justifying the assumption that  $R$  is frozen during the fluorescence process. Thus, at least in principle, one may obtain the distribution of donor–acceptor distance,  $P(R)$ , from the ensemble-averaged fluorescence decay,  $\bar{I}(\tau)$ , via the relationship in eq 2.

It is interesting to note that the above-mentioned inversion of  $\bar{I}(\tau)$  to obtain  $P(R)$  is numerically unstable. This can be demonstrated by writing  $\bar{I}(\tau)$  in terms of the probability density of  $k_{\text{ET}}$ :<sup>45</sup>

$$\bar{I}(\tau) = e^{-k_{\text{D}}\tau} \int_0^\infty dk_{\text{ET}} P(k_{\text{ET}}) e^{-k_{\text{ET}}\tau} \quad (3)$$

which implies that  $\bar{I}(\tau)$  corresponds to the Laplace transform of  $P(k_{\text{ET}})$ . Thus, extracting  $P(k_{\text{ET}})$  from  $\bar{I}(\tau)$  corresponds to calculating its inverse Laplace transform, which is known to be numerically unstable (i.e., small errors in  $\bar{I}(\tau)$  will be exponentially amplified in  $P(k_{\text{ET}})$ ). This problem is usually bypassed by assuming a certain ad-hoc functional form of  $P(R)$ , such as a linear combination of Gaussians, and best fitting the parameters via a least-squares procedure.

The ensemble-averaged FRET efficiency can be measured by using CW excitation, together with a time-integrated two-channel detection scheme. In this case, one selectively detects the fluorescence intensities that come from the donor and acceptor and uses them for evaluating the ensemble-averaged FRET efficiency:

$$\bar{E}_{\text{ET}} = \frac{I_{\text{A}}}{I_{\text{D}} + I_{\text{A}}} \quad (4)$$

where  $I_{\text{D}}$  and  $I_{\text{A}}$  correspond to the donor and acceptor fluorescence intensity, respectively. (For the sake of simplicity, we assume that the fluorescence photophysics of the two dyes are similar.) In the case where the donor–acceptor distance  $R$  is the same for all molecules, one finds that  $\bar{E}_{\text{ET}} = [1 + (R/R_0)^6]^{-1}$ . The same relationship can be used in order to define an ensemble-averaged  $R$ , which we will denote  $\bar{R}$ , in cases when  $R$  is not fixed:

$$\bar{E}_{\text{ET}} \equiv \frac{1}{1 + (\bar{R}/R_0)^6} \quad (5)$$

It should be noted that, in practice, the value of  $\bar{E}_{\text{ET}}$  obtained in this way will typically be lower than the actual value due to contributions from macromolecules with a missing or photo-bleached acceptor.<sup>16</sup>

Time-resolved and time-integrated ensemble-averaged FRET experiments can be used in order to monitor either the equilibrium or nonequilibrium conformational dynamics of macromolecules, as reflected by variations in the donor–acceptor distance. For example, in the case of proteins, one may use the time-resolved FRET technique in order to probe the equilibrium site–site distance probability density,  $P_{\text{eq}}(R)$ , at different points along the denaturation curve.  $P_{\text{eq}}(R)$  may then be used in order to construct the potential of mean force along the coordinate  $R$ :

$$V_{\text{mf}}(R) = -k_{\text{B}}T \ln[P_{\text{eq}}(R)] \quad (6)$$

Alternatively, one may follow relatively slow nonequilibrium processes, such as protein folding, by monitoring the variation in  $P(R)$  or  $\bar{E}_{\text{ET}}$  as a function of time.

SM-FRET experiments are performed by using a dual-channel detection scheme. In this case, one photoexcites the donor with CW radiation, while simultaneously detecting the fluorescence photons from the donor and acceptor in a selective manner. The fraction of photons detected in the acceptor channel, over a given time averaging window of length  $T_{\text{W}}$ , provides a direct measure of the time-averaged FRET efficiency, which we will denote by  $E_{\text{ET}}(T_{\text{W}})$ . One may then define a time-averaged and  $T_{\text{W}}$ -dependent donor–acceptor distance, which will be denoted by  $\langle R \rangle_{T_{\text{W}}}$ , such that

$$E_{\text{ET}}(T_{\text{W}}) \equiv \frac{1}{1 + [\langle R \rangle_{T_{\text{W}}}/R_0]^6} \quad (7)$$

In the actual experiment, one divides the overall time interval during which the single molecule is monitored into as many time windows as possible, as long as  $T_W$  is long enough so as to gather the statistics necessary for establishing  $E_{ET}(T_W)$  with a reasonable signal-to-noise ratio. If  $T_W$  is very short in comparison to the time scale on which  $R$  changes, then  $R$  remains fixed during the measurement and  $E_{ET}(T_W) = \{1 + [R/R_0]^6\}^{-1}$ . In such a case, there is a direct and exact relationship between the probability distribution of  $E_{ET}$  and  $P(R)$ :

$$P(R) = P(E_{ET} = (1 + [R/R_0]^6)^{-1}) \frac{d}{dR} (1 + [R/R_0]^6)^{-1} \quad (8)$$

On the other hand, if  $T_W$  is longer than the time scale on which  $R$  changes, then  $E_{ET}(T_W)$  will coincide with its ensemble average,  $\bar{E}_{ET}$ , and have the same value regardless of which time window it was measured on. In practice, one is usually in between those two limits, such that there is a distribution of values of  $E_{ET}(T_W)$ , which is narrower than in the limit  $T_W \rightarrow 0$ , but does not have a zero width, as in the limit  $T_W \rightarrow \infty$ . Thus, using this procedure can yield the probability density of  $\langle R \rangle_{TW}$ , i.e.,  $P(\langle R \rangle_{TW})$ . The dual-channel single-molecule procedure therefore provides a direct route to  $P(\langle R \rangle_{TW})$ , whereas the ensemble-averaged procedure can yield  $P(R)$  via an indirect route (cf. eq 2).

The probability density  $P(R)$ , which can be obtained from ensemble-averaged time-resolved FRET experiments, reveals structural information. However, the single-molecule FRET technique has the unique ability to provide additional dynamical information. Such dynamical information has been extracted in several ways. Jia et al. have monitored  $P(\langle R \rangle_{TW})$  as a function of  $T_W$ , in the case of surface-immobilized GCN4, and found that conformations corresponding to smaller values of  $\langle R \rangle_{TW}$  move on a faster time scale.<sup>14</sup> Talaga et al.<sup>15</sup> have considered the correlation between the square displacement of the donor–acceptor distance (divided by  $T_W$ )

$$D(j) = [\Delta R(j)]^2/T_W \equiv [\langle R \rangle_{TW}(j+1) - \langle R \rangle_{TW}(j)]^2/T_W \quad (9)$$

and the value of  $\langle R \rangle_{TW}(j)$ . Here,  $\langle R \rangle_{TW}(j)$  and  $\langle R \rangle_{TW}(j+1)$  correspond to the averaged value of  $R$  on the subsequent  $j$  and  $(j+1)$ -th time windows. Those authors found that conformations with larger values of  $\langle R \rangle_{TW}(j)$  are characterized, on average, by smaller values of  $D(j)$ . Rhoades et al.<sup>18</sup> have reported a wide and structured distribution of  $\Delta R(j)$ , with preference to relatively small values, which they attributed to the heterogeneous and multistep nature of the protein folding/unfolding process.

### 3. Protein Model and Simulation Techniques

We employed an off-lattice model of the protein, which was adopted from the work of Thirumalai and co-workers.<sup>46,47</sup> A brief description of the model is given below for the sake of completeness. The polypeptide backbone is described in terms of a chain consisting of  $N$  connected spherical beads of mass  $m$ . It is assumed that each of the beads is centered on the  $\alpha$ -carbon of the corresponding amino acid. Three types of beads are included: hydrophobic (B), hydrophilic (L), and neutral (N). The intramolecular potential energy of the protein is given by

$$V = V_{BL} + V_{BA} + V_{DIH} + V_{NB} \quad (10)$$

where  $V_{BL}$ ,  $V_{BA}$ ,  $V_{DIH}$ , and  $V_{NB}$  correspond to bond length potential, bond angle potential, dihedral angle potential, and nonbonding interactions, respectively.  $V_{BL}$  imposes connectivity between subsequent beads along the chain via a stiff harmonic

spring of equilibrium distance  $a$  and force constant  $k_r = 100\epsilon_h/a^2$  (the parameter  $\epsilon_h$  is considered below). The bending potential,  $V_{BA}$ , is assumed to be harmonic in the bending angle  $\theta_j$  between the  $j$ -th,  $(j+1)$ -th, and  $(j+2)$ -th successive beads, with a force constant  $k_\theta = 20\epsilon_h/\text{rad}^2$  and equilibrium angle of  $105^\circ$ .  $V_{DIH}$  is given by

$$V_{DIH} = \sum_{j=1}^{N-3} [A_j(1 + \cos \phi_j) + B_j(1 + \cos 3\phi_j)] \quad (11)$$

where  $\phi_j$  is the angle between the planes defined by beads  $(j, j+1, j+2)$  and  $(j+1, j+2, j+3)$ . Here,  $A_j = B_j = 1.2\epsilon_h$ , except for when at least two of the beads are neutral, in which case  $A_j = 0$  and  $B_j = 0.2\epsilon_h$ .  $V_{NB}$  accounts for nonbonding interactions between the  $j$ -th and  $k$ -th beads, which are separated by at least two beads along the chain ( $|j-k| \geq 3$ ). The L–L and L–B nonbonding interactions are purely repulsive and given by

$$V_{LL}(r) = V_{LB}(r) = 4 \frac{2\epsilon_h}{3} \left[ \left( \frac{a}{r} \right)^{12} + \left( \frac{a}{r} \right)^6 \right] \quad (12)$$

The N–N, N–L, and N–B nonbonding interactions are also purely repulsive and given by

$$V_{NN}(r) = V_{NL}(r) = V_{NB}(r) = 4\epsilon_h \left( \frac{a}{r} \right)^{12} \quad (13)$$

The B–B interactions are attractive and given by

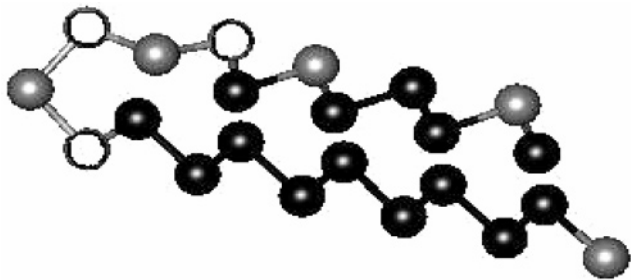
$$V_{BB}(r) = 4\epsilon_h \lambda \left[ \left( \frac{a}{r} \right)^{12} - \left( \frac{a}{r} \right)^6 \right] \quad (14)$$

Here,  $\lambda$  is a dimensionless parameter obtained from a relatively narrow Gaussian distribution centered around  $\lambda = 1$  ( $\lambda$  introduces diversity among hydrophobic beads, which is necessary for rapid folding).<sup>47</sup> Reduced units are used throughout, where the energy unit corresponds to  $\epsilon_h$  ( $\sim 1$  kcal/mol), and the mass and length units correspond to the mass,  $m \sim 3 \times 10^{-22}$  g, and diameter,  $a \sim 5 \times 10^{-8}$  cm, of a single bead (amino acid residue), respectively. We also note that this leads to a time unit of  $\tau_L = (ma^2/\epsilon_h)^{1/2} \sim 3$  ps.

The interaction of the protein with the solvent is modeled within the framework of the Langevin equation:

$$m \frac{d^2 \vec{r}_j}{dt^2} = -\zeta \frac{d\vec{r}_j}{dt} - \nabla_j V + \vec{\Gamma}_j \quad (15)$$

Here,  $\vec{r}_j$  is the position of the  $j$ -th bead ( $j = 1, \dots, N$ ),  $-\nabla_j V$  and  $\vec{\Gamma}_j$  are the systematic and random forces it is subject to, respectively, and  $\zeta$  is the friction coefficient. Equation 15 was integrated numerically via the Verlet algorithm.<sup>46,47</sup> Following ref 47, we set the friction coefficient to  $\zeta = 0.05$  and integration time step to  $\delta t = 0.005$ . The value of  $\zeta$  used here translates to  $\sim 10^{-3}$  of its value in room temperature aqueous solutions.<sup>47</sup> However, the folding mechanism has been observed to be relatively insensitive to the value of the friction coefficient,<sup>48</sup> and using a low value gives rise to a speed-up of the folding process, thereby improving the sampling efficiency in the simulations.<sup>47</sup> Temperature is introduced via the variance of the Gaussian random force, which is assumed to be delta-correlated. Although simplified, this model accounts for some of the main types of interactions that dominate protein folding<sup>46,49</sup> and as such provides a framework for describing protein kinetics which is both reasonable and feasible.



**Figure 1.** Conformation of the native state. The black, gray, and white beads represent hydrophobic (B), hydrophilic (L), and neutral (N) amino acid residues, respectively.

Thirumalai et al. have shown that it is useful to characterize each amino acid sequence in terms of two temperatures: (1) the collapse transition temperature,  $T_\theta$ , which corresponds to the peak of the heat capacity;<sup>50</sup> (2) the folding transition temperature,  $T_F$  ( $\leq T_\theta$ ), which corresponds to the peak of the overlap variance,  $\Delta\chi = \langle \chi^2 \rangle - \langle \chi \rangle^2$ , where

$$\chi = 1 - \frac{2}{(N-2)(N-3)} \sum_{i=1}^{N-3} \sum_{j=i+3}^N \Theta(\epsilon - |r_{ij} - r_{ij}^N|) \quad (16)$$

measures the overlap with the native conformation ( $\Theta(x)$  is the Heaviside function;  $0 \leq \chi \leq 1$ , where  $\chi = 0$  corresponds to a structure that coincides with the native state and  $\chi = 1$  to complete lack of native contacts). Furthermore, the ratio

$$\sigma = \frac{T_\theta - T_F}{T_\theta} \quad (17)$$

was found to correlate well with the folding mechanism. More specifically, sequences with  $\sigma \ll 1$  gave rise to rapid two-state folders, while sequences with large values of  $\sigma$  correspond to slow, multistate, folders. Thus,  $\sigma < 0.4$  typically correlated with a fraction  $\Phi > 0.9$  of the folding trajectories that followed two-state kinetics and reached the native conformation via a rapid nucleation collapse.<sup>47</sup>

In the present study, we consider a 22-residue protein with the following sequence:



The B–B nonbonding interactions are not the same for all B–B pairs and were sampled from a Gaussian distribution of width  $\Delta\lambda = 0.3$ . The procedure starts out by randomly choosing a specific set of values for the 91 values of  $\lambda$ , corresponding to all the possible  $(14 \times 13)/2$  B–B pairs. Specific values of  $\lambda$  are then assigned to specific B–B pairs, so as to minimize the energy of the desired native state, by using the inverse design Monte Carlo procedure of Shakhnovich et al.<sup>51,52</sup> It was verified that the native state of the optimized sequence coincides with the target conformation via the quenching/slow-cooling/simulated-annealing procedure described in ref 47. The resulting sequence corresponds to a fast two-state folder with  $T_\theta = 0.72$ ,  $T_F = 0.60$ , and  $\sigma = 0.17$ . It folds into a unique native conformation, which consists of a single U-turn with two arms of similar length on both sides (see Figure 1). The mean first passage time (or folding time) is estimated to be  $\sim 500\tau_L$  at  $T = 0.4$ .

Thirumalai and co-workers have sampled the conformations that correspond to the unfolded and folded states by setting the simulation temperature above  $T_\theta$  and below  $T_F$ , respectively. However, in actual SM-FRET experiments, one changes the state of the protein by varying the concentration of a denaturant.

One may also explore intermediate points along the denaturation curve, which will be characterized by a dynamical equilibrium between folded and unfolded states. To be more consistent with the SM-FRET experiment, the temperature was held fixed at  $T = 0.40$  ( $< T_F = 0.60$ ) at all the simulations reported in this paper. Denaturation by a chemical agent, such as urea, may be described by adjusting the effective strength of the B–B interactions, which represent the major driving force for folding.<sup>53</sup> The folded state can be sampled by using the B–B interactions in eq 14 (at  $T = 0.40$ ). The unfolded state can be sampled by completely turning off the attractive contribution to the B–B interaction, such that

$$V_{\text{BB}}^{\text{uf}}(r) = 4\epsilon_{\text{B}}\lambda\left(\frac{a}{r}\right)^{12} \quad (19)$$

Intermediate states can be obtained by adjusting the well depth of  $V_{\text{BB}}(r)$ . The “midpoint” of the denaturation curve can then be defined by the following B–B interaction:

$$V_{\text{BB}}^{\text{mid}}(r) = 0.5V_{\text{BB}}(r) \quad (20)$$

The folded and unfolded subpopulations at this midpoint state were found to be 0.64 and 0.36, respectively.

Trajectories were run for an overall number of  $10^8$  time steps, starting from the native state. Equilibration was typically obtained after  $5 \times 10^5$  time steps, following which statistics were collected from the remainder of the trajectory. The fact that the protein is a fast two-state folder implies that the potential surface is relatively smooth. We were therefore able to gather statistics by running a relatively small number of trajectories. Converged results for  $T_W = \delta t = 0.005\tau_L$  were obtained from two such trajectories for the folded, midpoint, and unfolded states (see below). Obtaining results at longer  $T_W$  required running longer trajectories. The largest value of  $T_W$  was dictated by the time scale of conformational motion, which is different at different states (see below). Thus, 16 trajectories were averaged over in order to get converged results for the folded state at  $T_W = 5.0$ , and 128 trajectories were averaged over in order to get converged results for the midpoint and unfolded states at  $T_W = 50.0\tau_L$ .

## 4. Results

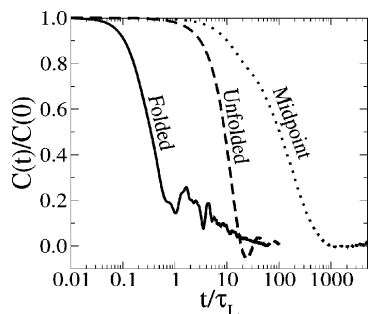
In this section, we present the results obtained from simulations in the folded, unfolded, and midpoint states (see section 3). The discussion focuses on quantities which are relevant to SM-FRET experiments and aims at elucidating the unique ability of those experiments to provide information on the correlations between structure and dynamics.

**A. Time Correlation Functions of the End-to-End Distance.** The characteristic time scales of conformational dynamics at different states along the denaturation curve can be estimated from the corresponding equilibrium correlation functions of the end-to-end distance:

$$C(t) = \langle R(t)R(0) \rangle - \langle R \rangle^2 \quad (21)$$

The correlation functions of the end-to-end distance, as obtained in the folded, midpoint, and unfolded states, are presented in Figure 2.

In the folded state, the decay of the correlation function is characterized by a relatively short time scale, of the order of  $\sim 1\tau_L$ . The initial sharp drop of  $C(t)$  after a period of time  $\tau_L$  is attributed to fluctuations around the native conformation, within the corresponding basin of attraction. The fact that those



**Figure 2.** Correlation function of the end-to-end distance in the folded, unfolded, and midpoint states.

fluctuations are fast is consistent with the fact that this is a rapid two-state folder. More specifically, the native conformation is expected to be significantly more stable in comparison to neighboring conformations, such that small displacements relative to it will give rise to strong returning forces. The longer time tail that follows this initial rapid decay is attributed to transitions to a competing nonnative basin of attraction (see below). The fact that the latter is significantly less stable in comparison to the native state is manifested by the fact that those transitions are rather infrequent.

The decay of the correlation function is significantly slower in the unfolded state and takes place on a time scale of about  $\sim 10\tau_L$ . This is consistent with the fact that the intramolecular nonbonding interactions in the unfolded state are of the excluded volume type, which are short-ranged, and therefore lead to relatively slow dynamics.

The decay of the correlation function in the midpoint state is characterized by a relatively wide range of time scales, spreading from  $\sim \tau_L$  to  $\sim 10^3\tau_L$ . The short time scales,  $\sim (1-10)\tau_L$ , are attributed to the dynamics within the folded and unfolded subpopulations. The slower time scale,  $\sim (10^2-10^3)\tau_L$ , is comparable to the folding first passage time of  $\sim 500\tau_L$  and is therefore attributed to transitions between the unfolded and folded states.

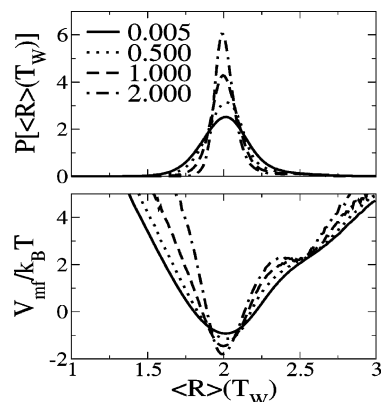
It should be noted that we have previously estimated  $\tau_L$  at  $\sim 3$  ps, which means that SM-FRET would have been unable to resolve the conformational dynamics if this was a real protein molecule. This rapid dynamics is attributed to the small number of beads and low value of the friction used in the actual simulations. Unfortunately, running the simulations with much larger protein molecules and frictions, which would lead to slower conformational dynamics that can be resolved by SM-FRET, is not computationally feasible. It is nevertheless reasonable to assume that many of the observations made for this model will also be relevant in the case of much slower conformational dynamics. The relevance of the following analysis is based on the validity of this assumption.

### B. Equilibrium Distributions of the End-to-End Distance.

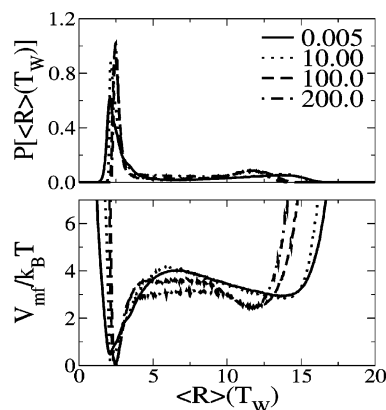
The distributions of the end-to-end distance, as obtained from the simulations, are presented in Figures 3, 4, and 5, for the folded, midpoint, and unfolded states, respectively. The distributions are shown for different averaging time windows and as a function of the window-time-averaged value of  $R$ , which is defined by

$$\langle R \rangle_{T_W} = T_W^{-1} \int_0^{T_W} d\tau R(\tau) \quad (22)$$

Although this convenient definition of  $\langle R \rangle_{T_W}$  is somewhat different from that dictated by eq 7, one expects that this distinction will not qualitatively modify the main observations reported below. The actual (non-time-averaged) distribution of



**Figure 3.** Distribution of the end-to-end distance at the folded state, for the indicated values of the averaging time window,  $T_W$  (upper panel). Also shown are the corresponding potentials of mean force (lower panel).

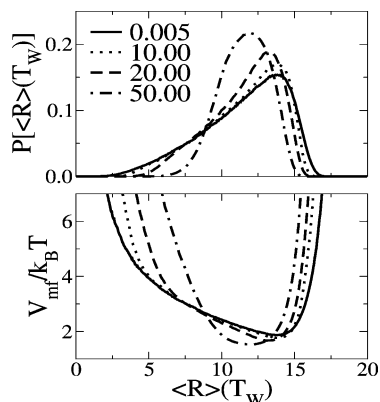


**Figure 4.** Same as Figure 3, for the midpoint state.

end-to-end distances corresponds to the smallest possible averaging time window, which is equal to the simulation time step,  $T_W = \delta t = 0.005$ . Also shown in Figures 3–5 are the potentials of mean force that can be obtained from the end-to-end distributions.

The folded state (Figure 3) is characterized by a relatively narrow distribution, which appears to be unimodal and peaked at the relatively small value of  $R \sim 2.0$ . Accordingly, the corresponding potential of mean force has the shape of a single well. The distribution becomes narrower as the averaging time window approaches the characteristic time scale of conformational dynamics in the folded state ( $\sim \tau_L$ ). As a result, the corresponding potential of mean force is distorted relative to the actual one. In fact, the potential of mean force develops a double-well feature, where the deep well corresponds to the native state and the shallow well corresponds to a competing basin of attraction. It is interesting to note that resolving this metastable state, which is manifested by a hardly noticeable shoulder at  $T_W = \delta t$ , is in fact facilitated by the time averaging inherent to SM-FRET experiments. This is made possible by the fact that the time scale for intrabasin dynamics is shorter in comparison to that which characterizes the interbasin dynamics. Thus, setting  $T_W$  to a value which is intermediate between those two time scales leads to “motional narrowing” of the two peaks, which helps in exposing the metastable conformation.

The midpoint state (Figure 4) is characterized by a bimodal distribution, with one narrow peak at  $R \sim 2$ , which corresponds to the folded subpopulation, and another, wider, peak at larger values of  $R \sim 10-17$ , which correspond to the unfolded subpopulation. Distinguishing between the two peaks becomes easier as the averaging time window increases, as long as  $T_W$

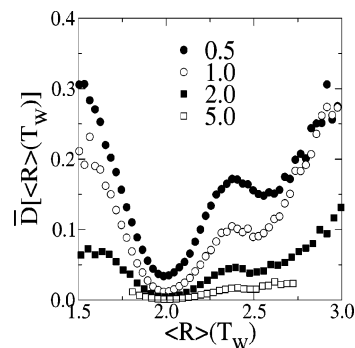


**Figure 5.** Same as Figure 3, for the unfolded state.

is smaller than the characteristic time scale for transitions between the folded and unfolded states ( $\sim 500\tau_L$ ). At the longest time averaging window,  $T_W = 200\tau_L$ , one observes the emergence of a finite probability density at values of  $R$  that fall between the folded and unfolded subpopulations. Computing the distributions that correspond to even longer values of  $T_W$  is prohibitively expensive and was not attempted in practice. However, one expects the probability density in the intermediate region to further increase with  $T_W$ , at the expense of the folded and unfolded subpopulations, and the width of the overall distribution to further decrease. This would transform the distribution from a bimodal one into a unimodal one, with the peak at a value of  $R$  that corresponds to an average over the folded and unfolded subpopulations.

The unfolded state (Figure 5) is characterized by a wide unimodal and asymmetrical distribution of the end-to-end distance, with a slow rise at  $R \sim 3-13$  and a sharp drop at  $R \sim 15$ . The distribution becomes narrower and more symmetrical as the averaging time window becomes comparable to, or larger than, the characteristic time scale of conformational dynamics in the unfolded state ( $\sim 10\tau_L$ ). This indicates a correlation between structure and dynamics in the unfolded state, such that conformations with shorter  $R$  move on a faster time scale and are therefore averaged over at smaller  $T_W$ . This is consistent with the observation that motion in the folded state, where  $R$  is small, occurs on a time scale faster than that of the unfolded state, where  $R$  is relatively large (cf. Figure 2).

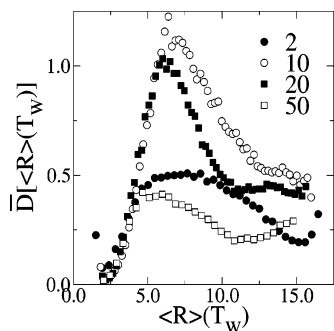
**C. Correlation between the Donor–Acceptor Distance and Its Displacement.** We next consider the correlation between structure and dynamics, as reflected in the correlation between the value of the displacement  $D(j)$  and the value of  $\langle R \rangle_{TW}(j)$  (cf. eq 9). The smallest possible averaging time window is given by  $T_W = \delta t = 0.005$ . In this case  $D(t) = [R(t + \delta t) - R(t)]^2 / \delta t = \delta t \dot{R}^2(t)$ . According to statistical mechanics, the probability distribution of the displacement,  $R$ , and corresponding instantaneous velocity,  $\dot{R}$ , are uncorrelated [ $P(R, \dot{R}) \propto \exp[-V_{mf}(R)/k_B T] \exp[-m\dot{R}^2/2k_B T]$ ]. Thus, the average value of  $D$ , over the subset of conformations that correspond to the same end-to-end distance  $R$ , will be constant and independent of  $R$ :  $\bar{D}(R) = \dot{R}^2 \delta t = k_B T \delta t / m$ . It should also be noted that in this case  $\bar{D}(R)$  becomes vanishingly small for a truly infinitesimal  $\delta t$ . The opposite extreme corresponds to the case where  $T_W$  is very large in comparison to the characteristic time scale of conformational dynamics. In this case, the ergodic hypothesis implies that averaging  $R$  over the time window is equivalent to taking the ensemble average, namely  $\langle R \rangle_{TW} = \bar{R}$ . The value of  $D$  will vanish at that single value of  $\langle R \rangle_{TW}$  since  $\langle R \rangle_{TW}(j+1) = \langle R \rangle_{TW}(j) = \bar{R}$ . Thus, no information regarding correlations between structure and dynamics can be obtained in the extreme



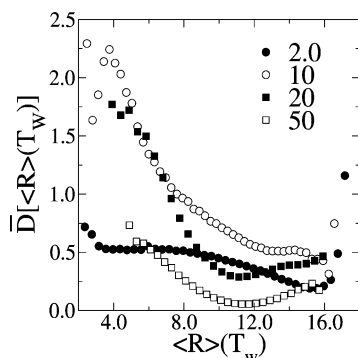
**Figure 6.** Dependence of  $\bar{D}(\langle R \rangle_{TW})$  on  $\langle R \rangle_{TW}$  in the folded state for the indicated values of the averaging time window,  $T_W$ .

limits of very short or very long averaging time windows. However, such information can still be obtained if one chooses an intermediate averaging time window, which is comparable to the characteristic time scale of the conformational dynamics. The facts that  $D$  is nonnegative and vanishes at the limits  $T_W \rightarrow 0$  and  $T_W \rightarrow \infty$  also imply that the dependence of  $\bar{D}(\langle R \rangle_{TW})$  on  $T_W$  is nonmonotonic and will exhibit a turnover.

The dependence of  $D(\langle R \rangle_{TW})$  on  $\langle R \rangle_{TW}$ , as obtained from the simulations, is presented in Figures 6, 7, and 8 for the folded, midpoint, and unfolded states, respectively. The dependence is shown in those figures for different averaging time windows. In the folded state (Figure 6)  $\bar{D}(\langle R \rangle_{TW})$  is observed to decrease when  $T_W$  increases, which implies that the values of  $T_W$  used correspond to the decreasing end of the turnover curve. This decrease signals one's approach to the ergodic limit as  $T_W$  increases. One also observes structure in  $\bar{D}(\langle R \rangle_{TW})$ , particularly when  $T_W$  is comparable to the characteristic time scale for conformational dynamics in this state ( $\sim \tau_L$ ). Furthermore, the structure is reminiscent of that of the distribution of end-to-end distances and potential of mean force for the same state (cf. Figure 3). More specifically, the positions of the two minima in  $\bar{D}(\langle R \rangle_{TW})$  coincide with those of the two maxima in the distribution of the end-to-end distances. As discussed in section 4.2, the two maxima in Figure 3 correspond to the native state and a competing metastable state. The structure observed in Figure 6 indicates that  $\bar{D}(\langle R \rangle_{TW})$  provides an independent and rather sensitive probe of those states, which is based on the underlying correlation between structure and dynamics. The fact that the lowest value of  $\bar{D}(\langle R \rangle_{TW})$  corresponds to the native state suggests that the characteristic time scale for conformational dynamics is faster in the native conformation and its close vicinity. This is because the ergodic limit will be approached more rapidly, i.e., for lower values of  $T_W$ , in this case. (The correlation between a smaller value of  $\bar{D}(\langle R \rangle_{TW})$  and faster dynamics should be noted.) This observation is consistent with the point of view that associates the native state with the minimum of a deep and narrow well on the protein's potential energy surface. Similarly, the higher value of  $\bar{D}(\langle R \rangle_{TW})$  in the second minimum, which corresponds to the competing metastable state, suggests that the latter correspond to a relatively shallow well in the energy landscape of the protein.<sup>54,55</sup> Finally, the larger value of  $\bar{D}(\langle R \rangle_{TW})$  between the two minima is the result of the slower time scale which characterizes the interconversion between the native and metastable basins of attraction, which implies that the ergodic limit will be reached at longer  $T_W$ . It should also be noted that the effective range of  $\langle R \rangle_{TW}$  decreases as  $T_W$  increases and that  $\bar{D}(\langle R \rangle_{TW})$  becomes smaller and structureless at  $T_W = 5\tau_L$ , which is significantly longer than the characteristic time for conformational dynamics in the folded state.



**Figure 7.** Same as Figure 6, for the midpoint state.



**Figure 8.** Same as Figure 6, for the unfolded state.

The correlation between structure and dynamics also persists in the midpoint and unfolded states. In the midpoint state (Figure 7)  $\bar{D}(\langle R \rangle_{T_W})$  clearly shows two minima at the positions of the maxima in the distribution of end-to-end distance in Figure 4. The two minima are particularly pronounced when  $T_W = 2.0\tau_L$ , which is intermediate between the time scales of conformational dynamics in the folded and unfolded states. In this case,  $D(\langle R \rangle_{T_W})$  is also observed to go through a turnover as  $T_W$  is increased. The same turnover behavior of  $\bar{D}(\langle R \rangle_{T_W})$ , as a function of  $T_W$ , is also observed in the case of the unfolded state (Figure 8). Here too the position of the minimum in  $\bar{D}(\langle R \rangle_{T_W})$  coincides with that of the maximum of the corresponding end-to-end distance distribution (Figure 5) as long as  $T_W$  is comparable to the characteristic time scale of conformational dynamics in this state ( $\sim 10\tau_L$ ). It should also be noted that, for the same value of  $T_W$ ,  $\bar{D}(\langle R \rangle_{T_W})$  is smaller in the folded state than in the unfolded state (e.g., compare the results for  $T_W = 2.0$  in Figures 6 and 8). This observation is consistent with the fact the dynamics in the folded state is in fact faster than that in the unfolded state.

## 5. Discussion

We first comment on the relationship between the results reported in this paper and those obtained for simpler models previously considered by other groups. Bagchi and co-workers have used Brownian dynamics simulations in order to investigate the distribution of the FRET efficiency at equilibrium and its variation during the collapse of a freely jointed homopolymer.<sup>35,38</sup> The corresponding equilibrium distribution of the end-to-end distance<sup>38</sup> is qualitatively similar to ours. The distribution they report is unimodal in the folded and unfolded state, with the latter broader and shifted to longer distances in comparison to the former. The distribution at intermediate temperatures is bimodal, with contributions from the folded and unfolded subpopulations. Cao and co-workers have analyzed SM-FRET within the framework of a semiflexible Gaussian chain model.<sup>39</sup> Similarly to the treatment in the present paper, the dynamics of

the homopolymer was described in terms of a Langevin equation. They found that the displacement of the end-to-end distance in the collapsed structure, which is qualitatively similar to the folded state, is smaller than that in the coiled structure, which is qualitatively similar to the unfolded state. This is consistent with our result regarding the correlation between  $\bar{D}(\langle R \rangle_{T_W})$  and  $\langle R \rangle_{T_W}$ . Gopich and Szabo have calculated the dependence of the distribution of end-to-end distances on  $T_W$  in the case of a freely jointed Gaussian chain model and for a reduced kinetic model with protein-like features.<sup>40</sup> The distribution of the end-to-end distances can be obtained analytically in the case of a Gaussian chain and is qualitatively similar to that in the unfolded state. To model the dynamics of  $R(t)$ , the authors of ref 40 assumed that it is governed by a 1D diffusion equation subject to the corresponding potential of mean force. Similarly to our observation for the unfolded state, they found that the distribution of end-to-end distances narrows and becomes more symmetrical as  $T_W$  becomes comparable to the characteristic decay time of the autocorrelation function of  $R(t)$ . In the case of the 2D potential energy surface, the authors have chosen parameters which gave rise to a separation of time scales between the fast dynamics within the folded and unfolded states and the slow transitions between those states. They also assumed that the slow process can be described by simple rate kinetics and avoided dealing explicitly with the dynamics within the folded and unfolded states by assuming that  $T_W$  is longer than the corresponding times scales. Similarly to our results for the midpoint state, they have observed the “motional narrowing” of the individual peaks which correspond to the folded and unfolded states, followed by the transition from a bimodal into a narrow unimodal distribution when  $T_W$  becomes comparable to, or longer than, the folding time.

In comparison to those previous studies, the protein model analyzed in this paper includes several important new features. More specifically, the model is based on a heteropolymer, rather than a homopolymer, with protein-like bonding and nonbonding (hydrophilic–hydrophobic) interactions, a well-defined native conformation, and a well-characterized two-state folding kinetics. Although those additional features bring the theoretical treatment one step closer to the reality of SM-FRET experiments, the results of the present study cannot be directly related to specific protein molecules employed in SM-FRET experiments. For example, it should be noted that shot noise, spectral diffusion, and dipole angle distributions, which contribute to the width of experimental distributions of FRET efficiencies, were not taken into account in the simulations. Nevertheless, it is interesting to compare our observations regarding the correlation between structure and dynamics to those made in actual SM-FRET measurements. The SM-FRET studies of Hochstrasser and co-workers on the protein GCN4 are particularly suitable for this purpose.<sup>14,15</sup> On the basis of the analysis of the dependence of the experimental distribution of the end-to-end distance on the averaging time window, and similarly to us, the authors of ref 14 have concluded that the dynamics in the folded state is faster than that in the unfolded state. Furthermore, the authors of ref 15 have reported a distribution of  $\bar{D}(\langle R \rangle_{T_W})$  with structure that depends on the denaturant concentration, which may reflect the underlying potential surface. Interestingly, the experimental observations regarding the relative size of  $\bar{D}(\langle R \rangle_{T_W})$  in the folded and unfolded states seem to disagree with the results obtained from our simulations as well as with the results of Cao et al. for semiflexible Gaussian chains.<sup>39</sup> More specifically, the experimentally measured  $\bar{D}(\langle R \rangle_{T_W})$  was found to be larger in the folded state, whereas our simulations and ref

39 predict an opposite trend. It is interesting to note in this context that the experimental results also suggest that the unfolded state is rather sensitive to the fact that the protein was surface-immobilized in the actual experiments.<sup>15</sup> In fact, the authors of ref 15 have interpreted the smaller value of  $\bar{D}(\langle R \rangle_{TW})$  in the unfolded state as resulting from a slowdown of the conformational dynamics due to interactions with the surface. However, at least in the context of the model employed here, smaller values of  $\bar{D}(\langle R \rangle_{TW})$  are actually indicative of faster conformational dynamics (assuming that  $T_W$  is comparable to or larger than the time scale of the corresponding conformational dynamics). Thus, the fact that surface immobilization leads to a smaller values of  $\bar{D}(\langle R \rangle_{TW})$  may be attributed to the fact that the strong interactions with the surface actually lead to more rapid fluctuations around the corresponding equilibrium point, whose time scale may be even faster than that of the fluctuations around the native structure in the folded state.

## 6. Summary

SM-FRET experiments provide a unique opportunity for exploring the relationship between structure and dynamics in biomolecules. In this paper, we explored this aspect of SM-FRET within the context of Langevin dynamics simulations of an off-lattice model for the polypeptide chain. The model of the protein and simulation techniques employed in this paper represent a compromise between prohibitively expensive atomistic simulations and significantly oversimplified lattice models.<sup>56–59</sup> Such a compromise is necessary in light of the relatively long time scale of SM-FRET experiments, which is dictated by the need for collecting enough fluorescence photons to establish a good signal-to-noise ratio. Furthermore, comparison of the experimental and simulation results can shed light on the adequacy of the underlying model. Although direct comparison to experiment was not attempted in the present paper, the general behavior observed in the simulation is for the most part consistent with currently available experimental data. One notable exception is the relative size of  $\bar{D}(\langle R \rangle_{TW})$  in the folded and unfolded states, which follow opposite trends in experiment and simulation, and may provide information on the influence of surface immobilization.

The simulations help to elucidate several ways in which SM-FRET can provide information on the different time scales that characterize protein motion at different points along the denaturation curve. A key role is played by the averaging time window. As demonstrated by the results, the requirement for a finite averaging time window does not necessarily constitute a disadvantage. For example, using an averaging time window longer than the time scale of motion within the native and competing basins of attraction, but shorter than the time scale of transitions between those states, has actually led to improved resolution of the corresponding subpopulations. More generally, the dependence of the distribution of the donor–acceptor distance on the averaging time window provides valuable information on the correlation between structure and dynamics. Even more dramatic is the impact of time averaging on the correlation between the donor–acceptor distance and its displacement. In this case, time averaging is actually necessary in order to resolve the underlying correlations between structure and dynamics.

Other aspects of the correlations between structure and dynamics can be studied within the theoretical and computational framework employed in this paper. Those include (1) the correlation between the unfolded conformation and nonequilibrium folding pathway, (2) the influence of immobilization

and confinement on the equilibrium and nonequilibrium dynamics of the protein, and (3) the signature of the correlation between structure and dynamics on photon statistics. Those aspects are the subject of ongoing work in our group and will be reported in future publications.

**Acknowledgment.** The authors are grateful to Dr. Qiang Shi for helpful suggestions and to the Petroleum Research Fund for financial support (through Grant 36486-G).

## References and Notes

- (1) Edman, L.; Mets, Ü.; Rigler, R. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6710.
- (2) Ha, T.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6264.
- (3) Deniz, A. A.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 3670.
- (4) Nie, S.; Zare, R. N. *Annu. Rev. Biophys. Struct.* **1997**, *26*, 567.
- (5) Jia, Y.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 7932.
- (6) Xi, X. S.; Trautman, J. K. *Annu. Rev. Phys. Chem.* **1998**, *49*, 441.
- (7) Weiss, S. *Science* **1999**, *283*, 1676.
- (8) Weiss, S. *Nature: Struct. Biol.* **2000**, *7*, 724.
- (9) Ha, T.; et al. *Chem. Phys.* **1999**, *247*, 107.
- (10) Ha, T.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 893.
- (11) Ha, T.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9077.
- (12) Ishii, Y.; et al. *Chem. Phys.* **1999**, *247*, 163.
- (13) Geva, E.; Skinner, J. L. *Chem. Phys. Lett.* **1998**, *288*, 225.
- (14) Jia, Y.; et al. *Chem. Phys.* **1999**, *247*, 69.
- (15) Talaga, D. S.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13021.
- (16) Deniz, A. A.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5179.
- (17) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature (London)* **2002**, *419*, 743.
- (18) Rhoades, E.; Gussakovsky, E.; Haran, G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3197.
- (19) Hu, D.; et al. *Nature (London)* **2000**, *405*, 1030.
- (20) Brasselet, S.; Peterman, E. J. G.; Miyawaki, A.; Moerner, W. E. *J. Phys. Chem. B* **2000**, *104*, 3676.
- (21) Ying, L.; Wallace, M. L.; Balasubramanian, S.; Klenerman, D. *J. Phys. Chem. B* **2000**, *104*, 5171.
- (22) Zhuang, X.; et al. *Science* **2000**, *288*, 2048.
- (23) Russel, R.; et al. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 155 2002.
- (24) Zhuang, X.; et al. *Science* **2002**, *296*, 1473.
- (25) Zhuang, L. E. B. X. W.; Das, R.; Chu, S.; Herschlag, D. *J. Mol. Biol.* **2003**, *6*, 923.
- (26) Haran, G. *J. Phys.: Condens. Matter* **2003**, *15*, R1291.
- (27) Zhuang, X. W.; Reif, M. *Curr. Opin. Struct. Biol.* **2003**, *13*, 88.
- (28) Stryer, L.; Haugland, R. P. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *58*, 719.
- (29) Haas, E.; Wilchek, M.; Katchalski-Katzir, E.; Steinberg, I. Z. *Proc. Natl. Acad. Sci. U.S.A.* **1975**, *72*, 1807.
- (30) Selvin, P. R. *Methods Enzymol.* **1995**, *246*, 300.
- (31) Andrews, D. L.; Demidov, A. A. *Resonance Energy Transfer*; Wiley: New York, 1999.
- (32) Selvin, P. R. *Nature: Struct. Biol.* **2000**, *7*, 730.
- (33) Deniz, A. A.; et al. *Annu. Rev. Phys. Chem.* **2001**, *52*, 233.
- (34) Rhoades, E.; Gussakovsky, E.; Haran, G. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7418.
- (35) Srinivas, G.; Yethiraj, A.; Bagchi, B. *J. Phys. Chem. B* **2001**, *105*, 2475.
- (36) Srinivas, G.; Yethiraj, A.; Bagchi, B. *J. Chem. Phys.* **2001**, *114*, 9170.
- (37) Srinivas, G.; Bagchi, B. *J. Chem. Phys.* **2002**, *116*, 837.
- (38) Srinivas, G.; Bagchi, B. *Phys. Chem. Commun.* **2002**, *5*, 59.
- (39) Yang, H.; Witkoskie, J. B.; Cao, J. S. *J. Chem. Phys.* **2002**, *117*, 11010.
- (40) Gopich, I. V.; Szabo, A. J. *Phys. Chem. B* **2003**, *107*, 5058.
- (41) Förster, T. *Ann. Phys.* **1948**, *2*, 55.
- (42) Dexter, D. L. *J. Chem. Phys.* **1953**, *21*, 836.
- (43) Scholes, G. D. *Annu. Rev. Phys. Chem.* **2003**, *54*, 57.
- (44) Ratner, V.; Sinev, M.; Haas, E. *J. Mol. Biol.* **2000**, *299*, 1383.
- (45) Lee, M.; Tang, J.; Hochstrasser, R. M. *Chem. Phys. Lett.* **2001**, *344*, 501.
- (46) Honeycutt, J. D.; Thirumalai, D. *Biopolymers* **1992**, *32*, 695.
- (47) Veithans, T.; Klimov, D.; Thirumalai, D. *Fol. Des.* **1996**, *2*, 1.
- (48) Klimov, D. K.; Thirumalai, D. *Phys. Rev. Lett.* **1997**, *79*, 317.
- (49) Dill, K. A. *Biochemistry* **1990**, *29*, 7133.

(50) de Gennes, P. G.; *Scaling Concepts in Polymer Physics*; Cornell University Press: Ithaca, NY, 1979.

(51) Shakhnovich, E.; Gutin, A. M. *Protein Eng.* **1993**, *6*, 793.

(52) Shakhnovich, E. *Phys. Rev. Lett.* **1994**, *72*, 3907.

(53) Walqwist, A.; Covell, D. G.; Thirumalai, D. *J. Am. Chem. Soc.* **1998**, *120*, 427.

(54) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. *Science* **1991**, *254*, 1598.

(55) Frauenfelder, H.; Wolynes, P. G.; Austin, R. H. *Rev. Mod. Phys.* **1999**, *71*, S419.

(56) Succi, N. D.; Onuchic, J. N. *J. Chem. Phys.* **1994**, *101*, 1519.

(57) Succi, N. D.; Onuchic, J. N. *J. Chem. Phys.* **1995**, *103*, 4732.

(58) Dill, K. A.; et al. *Protein Sci.* **1995**, *4*, 561.

(59) Pande, V. S.; Grosberg, A. Y.; Tanaka, T.; Rokhsar, T. S. *Curr. Opin. Struct. Biol.* **1998**, *8*, 68.