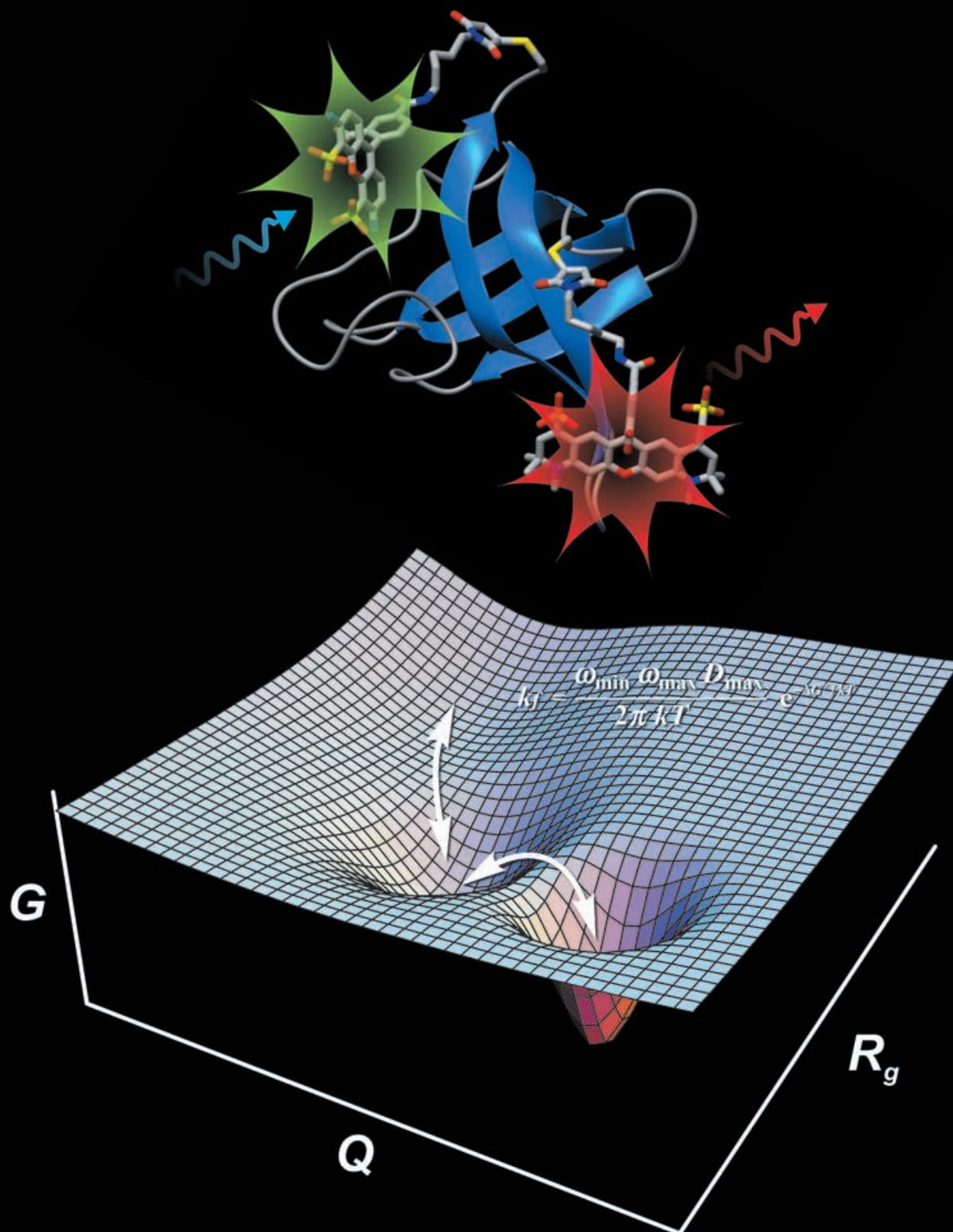


## Single-molecule protein folding



# Single-Molecule Fluorescence Spectroscopy of Protein Folding

Benjamin Schuler\*<sup>[a]</sup>

*Dedicated to Professor Rainer Jaenicke on the occasion of his 75th birthday.*

*Single-molecule spectroscopy is an important new approach for studying the intrinsically heterogeneous process of protein folding. This Review illustrates how different single-molecule fluores-*

*cence techniques have improved our understanding of mechanistic aspects in protein folding, exemplified by a series of recent experiments on a small protein.*

## 1. Introduction

Proteins are the major components in the living cell that translate the genetic information into the whole repertoire of constituents of cellular organization. Their biosynthesis takes place on free or membrane-bound ribosomes, yielding linear sequences of amino acids. In a second step, they acquire their singular, unique three-dimensional structure, which is the central prerequisite for their specific functions. The rules relating the linear information stored in the DNA of an organism's genome to the amino acid sequences of the corresponding proteins are well understood. However, the complexity of the many-body interactions that govern the spontaneous transition of a polypeptide chain into its spatially ordered native conformation has to date prevented a comprehensive solution of the protein-folding problem. Nevertheless, our understanding of protein-folding mechanisms has increased dramatically, and single-molecule spectroscopy has started to provide new perspectives of this fundamental process.

## 2. Protein Folding

Experiments in the 1930s and 40s indicated that protein folding is autonomous and reversible, that is, that a protein's specific structure can be restored after denaturation.<sup>[1–4]</sup> Anfinsen's classical work on the reduction and reoxidation of ribonuclease supplied clear evidence that the amino acid sequences selected through evolution contain all the necessary information for attaining their native structures.<sup>[5–7]</sup> Owing to the marginal conformational stability of folded proteins and the often low activation barriers involved, protein folding is usually not a unique event in the life of a protein, but may occur many times before a protein is finally degraded in the cell. The conceptual simplicity of the folding problem attracted contributions from a wide range of disciplines, especially from the physical sciences. However, the initial hope that the rules for protein folding would be as simple to decode as those for translation soon had to be abandoned.

The hierarchic patterns found in crystal structures of proteins suggested that folding might also be a hierarchical process, and the developing concept of nucleation events emphasized the importance of local interactions in the early stages of

protein folding, with a subsequent condensation of substructures.<sup>[7–10]</sup> The 1970s and 80s were dominated by the hunt for such folding intermediates on the pathway from the unfolded to the native state, and led to the identification of structurally and spectroscopically distinct intermediate states in the folding of a large number of proteins.<sup>[11–13]</sup> This was accompanied by the development of a range of phenomenological models of folding<sup>[11,14–19]</sup> that differ in the degree of parallel formation of substructures, their interactions on the way to more complex assemblies, and the assignment of the rate-limiting step, but which are difficult to distinguish experimentally.<sup>[20]</sup>

There were two very important developments in the 1990s: The discovery of small proteins that fold at very high rates ( $\approx 10^3 \text{ s}^{-1}$  and faster) without populating stable intermediates,<sup>[21,22]</sup> and the introduction of the concepts of statistical mechanics into the theory of protein folding.<sup>[23–27]</sup> The identification of "two-state folders" very clearly demonstrated that stable intermediates are not a general prerequisite for folding, and demanded different approaches to solving the protein-folding problem. One such strategy is the characterization of transition states for folding by protein engineering methods,<sup>[28,29]</sup> another the analysis of elementary processes of protein folding, such as the formation of loops<sup>[30–35]</sup> or isolated secondary structure elements.<sup>[36–39]</sup> Although the "new view"<sup>[40]</sup> of protein folding introduced from statistical thermodynamics<sup>[23–26]</sup> was sometimes perceived to be in contradiction to the "classical view" of folding pathways, it can rather be considered more general,<sup>[41]</sup> and it contributed to a range of new questions, which are currently under intense investigation. These include the hunt for a protein-folding "speed limit"<sup>[30,32,42]</sup> and "downhill"—or barrierless—folding,<sup>[43–45]</sup> the importance of unfolded-state dynamics,<sup>[46–50]</sup> and the prediction of protein-folding rates.<sup>[51–54]</sup>

Another boom for the field of protein folding has come from the realization that misfolding and aggregation<sup>[55]</sup> of

[a] Prof. B. Schuler

Department of Biochemistry, University of Zürich  
Winterthurerstr. 190, 8057 Zürich (Switzerland)  
Fax (+41) 44-635-5907  
E-mail: schuler@bioc.unizh.ch

many proteins—for years disregarded as an irrelevant side reaction—are of great medical relevance and a probable cause of a wide range of diseases, especially neurodegenerative disorders<sup>[56]</sup> such as Alzheimer's, Parkinson's, and Huntington's disease. Common to protein folding and misfolding is the large degree of structural or conformational heterogeneity, both during structure formation and—particularly in the case of aggregates—of the final structures. In many cases, obtaining a detailed structural understanding of the processes involved is difficult or even impossible with classical methods investigating large ensembles of molecules. Therefore, it is a promising new opportunity in the search for a more fundamental understanding of protein folding to turn to the study of individual molecules, with the ultimate goal of identifying the distributions of microscopic pathways an unfolded protein can take to its final state, be it correctly or incorrectly folded.

### 3. Single-Molecule Spectroscopy

The direct investigation of the folding of single protein molecules has only become feasible by means of new methods such as atomic force microscopy (AFM)<sup>[57,58]</sup> and optical single-molecule spectroscopy.<sup>[59–65]</sup> These techniques offer a fundamental advantage beyond our mere fascination for the direct depiction of molecular processes: They can resolve and quantify the properties of individual molecules or subpopulations inaccessible in classical ensemble experiments, which average over many particles. The very first experiments that demonstrated the unfolding of individual protein domains in the large muscle protein titin employed AFM<sup>[57]</sup> or laser tweezers.<sup>[66]</sup> AFM, in particular, has since been used to investigate the mechanical stability of a number of proteins.<sup>[67–69]</sup> In combination with protein engineering methods<sup>[68,70]</sup> and simulations,<sup>[71,72]</sup> these studies will continue to provide an important complement to other approaches, especially for proteins involved in the transduction of mechanical force. In the following, however, I will focus on fluorescence spectroscopy, which is a particularly appealing technique, owing to its extreme sensitivity and versatility.<sup>[61,73,74]</sup> In combination with Förster resonance energy transfer (FRET),<sup>[75–77]</sup> it enables us to investigate

intramolecular distances and conformational dynamics of single proteins, including their folding, misfolding, and function.

#### 3.1. Förster Resonance Energy Transfer

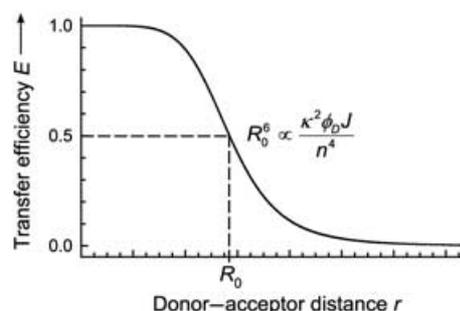
The first quantitative test of Förster's theory<sup>[75,78]</sup> and the crucial experiment that put FRET on the map of biochemistry was published by Stryer and Haugland in 1967.<sup>[79]</sup> They attached dansyl and naphthyl groups to the termini of polyproline peptides and measured the transfer efficiency between them as a function of the length of the peptide. As predicted for the dipole–dipole coupling between the donor and the acceptor dye,<sup>[75]</sup> they found that the transfer efficiency  $E$  depended on the inverse sixth power of the interchromophore distance  $r$ , in agreement with Theodor Förster's famous equation [Eq. (1)]:

$$E = \frac{R_0^6}{R_0^6 + r^6} \quad (1)$$

where  $R_0$  is the Förster radius, the characteristic distance that results in a transfer efficiency of 50% (Figure 1).  $R_0$  is calculated in Förster's theory according to Equation (2):

$$R_0^6 = \frac{9000(\ln 10)\kappa^2\phi_D J}{128\pi^5 n^4 N_A} \quad (2)$$

where  $J$  is the overlap integral between the donor emission and the acceptor absorption spectra,  $\phi_D$  is the donor's fluorescence quantum yield,  $\kappa^2$  is a factor depending on the relative



**Figure 1.** Distance dependence of the Förster resonance energy transfer efficiency between a suitable pair of chromophores, calculated according to Equation (1). The characteristic Förster distance  $R_0$  is calculated from the orientational factor  $\kappa^2$ , the donor quantum yield  $\phi_D$ , the overlap integral  $J$ , and the refractive index of the medium  $n$ .

Ben Schuler (born 1971) studied biochemistry in Regensburg (Germany) and Canterbury (UK), and received his PhD in Physical Biochemistry from Regensburg University. He did his post-doctoral research in Biophysics at the National Institutes of Health (Bethesda, MD, USA) and then started a research group at Potsdam University (Germany) as an Emmy Noether Fellow of the Deutsche Forschungsgemeinschaft. He joined the Biochemistry Institute at Zürich University in 2004 as Assistant Professor, where he is working on the folding and misfolding of proteins, especially using single-molecule fluorescence spectroscopy.



orientation of the chromophores,  $n$  is the refractive index of the medium between the dyes, and  $N_A$  is Avogadro's number.<sup>[75,78]</sup> The idea of such a "spectroscopic ruler"<sup>[79]</sup> has had a huge impact on the investigation of biomolecular structure and dynamics on distances in the range of about 1–10 nm.<sup>[77,80–82]</sup> More recently, renewed interest has come from the realization that FRET can be used to obtain distance information in experiments on single biomolecules,<sup>[76,83]</sup> including proteins.<sup>[61,67,84]</sup>

Experimentally, transfer efficiencies can be determined in a variety of ways,<sup>[78]</sup> but for single-molecule FRET, two approaches are particularly useful. One is the measurement of the fluorescence intensities from both the donor and the acceptor chromophores, and the calculation of the transfer efficiency according to Equation (3):

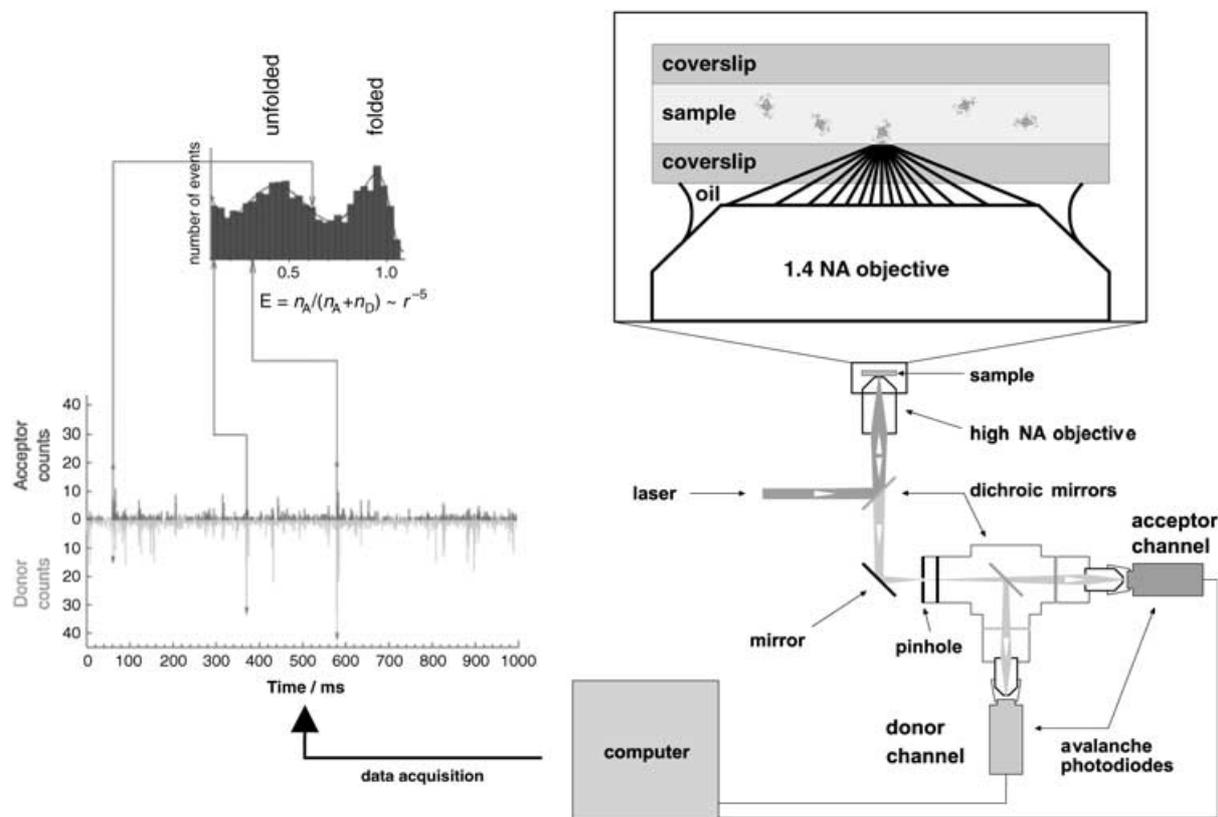
$$E = \frac{n_A}{n_A + \gamma n_D} \quad (3)$$

where  $n_A$  and  $n_D$  are the numbers of photons detected from the acceptor and donor chromophores, respectively, and  $\gamma$  is a correction factor that takes into account the quantum yields of the dyes and the efficiencies of the detection system in the corresponding wavelength ranges.<sup>[85]</sup> A second approach, which can be combined with the first,<sup>[86]</sup> is the determination of the fluorescence lifetime of the donor in the presence ( $\tau_{DA}$ ) and absence ( $\tau_D$ ) of the acceptor, yielding the transfer efficiency as Equation (4):

$$E = 1 - \frac{\tau_{DA}}{\tau_D} \quad (4)$$

### 3.2. Instrumentation

The instrumentation is based on the optics, detectors, and electronics developed for optical single-molecule spectroscopy<sup>[59,73]</sup> and fluorescence correlation spectroscopy.<sup>[67,88]</sup> Experimental setups for single-molecule FRET<sup>[89]</sup> typically involve either confocal excitation and detection using a pulsed or continuous wave laser and avalanche photodiodes (APDs), or wide-field microscopy with two-dimensional detectors such as a charge coupled device (CCD) camera, often in combination with evanescent wave excitation. Wide-field imaging allows the collection of data from many single-molecules in parallel, albeit at a much lower time resolution than in a confocal experiment using APDs. Figure 2 shows a simple schematic with the main optical elements for confocal epifluorescence detection. A laser beam is focused with a high aperture objective to a diffraction-limited focal spot that serves to excite the labeled molecules. In the simplest experiment, the sample molecules are freely diffusing in solution at very low concentrations (typically 10–100 pM), ensuring that the probability of two molecules residing in the confocal volume at the same time is negligible. When a molecule diffuses through the confocal volume, the donor dye is excited, fluorescence from the donor and acceptor is collected through the objective and focused onto the pinhole, a small aperture serving as a spatial filter. A dichroic mirror finally separates donor and acceptor emission into the



**Figure 2.** Scheme of a confocal single-molecule fluorescence experiment on freely diffusing molecules (molecules not to scale). On the left, a typical time trace is shown, with counts detected from the donor chromophore and the acceptor chromophore. For each individual event, a transfer efficiency is calculated and entered into a histogram. Histograms are typically constructed from several thousand bursts.

corresponding detectors, from where the data are collected with multichannel scalers or suitable counting cards. The setup can be extended to sorting photons by additional colors, for example, if more than two chromophores are used,<sup>[90,91]</sup> or by both color and polarization.<sup>[86]</sup> The advantage of observing freely diffusing molecules is that perturbations from surface interactions can largely be excluded, but the observation time is limited by the diffusion times of the molecules through the confocal volume. Typically, every molecule is observed for no more than a few milliseconds. Alternatively, the molecules can be immobilized on the surface and then observed for a more extended period of time, typically a few seconds, until one of the chromophores undergoes photodestruction. The complications in this case are interactions with the surface that can easily perturb the sensitive equilibrium of protein folding (see Section 4.5.). The details of single-molecule instrumentation can be found in several recent reviews.<sup>[61,73,89]</sup> An important development for the wide application of single-molecule methods to the study of biomolecules is the recent availability of comprehensive commercial instrumentation.<sup>[92]</sup>

### 3.3. Protein Labeling

To our misfortune, protein chemistry has not made it easy for us to investigate polypeptides in single-molecule experiments (with the exception of the family of fluorescent proteins<sup>[93,94]</sup>). As of today, even tryptophan, the natural amino acid with the highest fluorescence quantum yield ( $\approx 13\%$ ), is not suitable for single-molecule detection (unless the molecule contains a very large number of tryptophan residues<sup>[95]</sup>) owing to the low photostability of the indole ring. Labeling with extrinsic fluorophores is thus unavoidable, and complicated by the need for suitable reactive groups for site-specific attachment. For FRET, two (or more) chromophores are needed, and their specific placement on the protein ideally requires groups with orthogonal chemistries. For simple systems, such as short peptides, sequences can be designed to introduce only single copies of residues with suitable reactive side chains.<sup>[96,97]</sup> In chemical, solid-phase peptide synthesis, protection groups and the incorporation of non-natural amino acids can be used to increase specificity, but, for longer chains, chemical synthesis becomes inefficient and shorter chains have to be ligated<sup>[98]</sup> to obtain the desired product.<sup>[99]</sup>

Considering the maturity and versatility of the heterologous recombinant protein expression, the production of proteins of virtually any size and sequence in microorganisms is the method of choice for obtaining very pure material in sufficiently large amounts for preparative purposes. However, the number of functional groups that can be used for specific labeling is very limited. Sufficiently specific reactivity is only provided by the sulfhydryl groups of cysteine residues, the amino groups of lysine side chains, and the free  $\alpha$ -amino group of the *N*-terminal amino acid. However, except for small peptides, the statistical and therefore often multiple occurrence of cysteine and especially lysine residues in one polypeptide prevents the specific attachment of exactly one label to a protein. For some applications, such as in-vivo imaging, the degree of

labeling is only of secondary importance, but for FRET, specificity is strictly required.

Currently, the most common approach is to rely exclusively on cysteine derivatization. Increased specificity can be achieved by removing unwanted natural cysteine residues by site-directed mutagenesis or introducing cysteine residues with different reactivity due to the different molecular environments within the protein.<sup>[100]</sup> Labeling is usually combined with multiple chromatography steps to purify the desired adducts. Alternative methods<sup>[101]</sup> are the native chemical ligation of recombinantly expressed and individually labeled protein fragments or intein-mediated protein splicing,<sup>[102]</sup> the specific reaction with thioester derivatives of dyes,<sup>[103]</sup> puromycin-based labeling using in-vitro translation,<sup>[104]</sup> and the introduction of non-natural amino acids.<sup>[105]</sup> Most of the latter methods are not yet used routinely and must be considered as under development. Considering all of these complications, it is very fortunate that a wide variety of excellent organic dyes with various functional groups for protein labeling have become commercially available. Examples of particularly popular chromophores for single-molecule FRET are the cyanine dyes<sup>[106]</sup> and the Alexa Fluor series.<sup>[107]</sup> Semiconductor quantum dots<sup>[108,109]</sup> are promising candidates owing to their extreme photostability, but they are not yet available with single functional groups; to date they have only been used as donors because of their broad absorption spectra, and they are themselves the size of a small protein.

Even for the smaller organic dyes, however, interactions with the protein can interfere both with the photophysics of the chromophores and the stability of the protein. This needs to be taken into account both in the design of the labeled variants and the control experiments. Owing to the substantial size of the fluorophores, they can only be positioned on the solvent-exposed surface of the protein if the folded structure is to be conserved. Even then, the use of hydrophobic dyes can lead to aggregation of the protein, or interactions with the protein surface can cause a serious reduction in fluorescence quantum yield, a problem that has been minimized by the introduction of charged groups in many of the popular dyes.<sup>[106,107]</sup> Important control experiments are equilibrium or time-resolved fluorescence anisotropy measurements,<sup>[96,97,99]</sup> which are sensitive to the rotational flexibility of the dyes and can therefore provide indications for undesirable interactions with the protein surface. It is also essential to ensure, by direct comparison with unmodified protein, that labeling has not substantially altered the protein's stability or folding mechanism.<sup>[96,99]</sup>

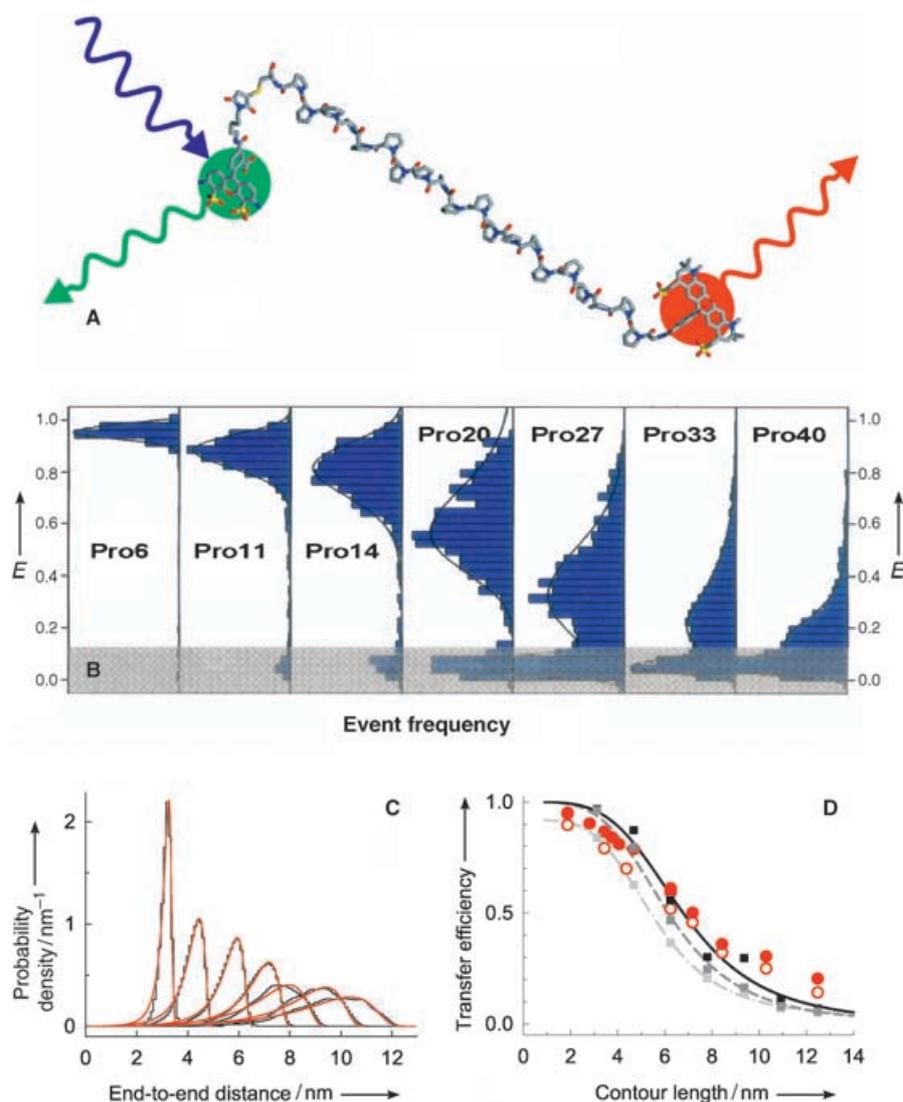
### 3.4. Spectroscopic Controls

The factors influencing single-molecule fluorescence experiments<sup>[110]</sup> are often sufficiently diverse to preclude a complete anticipation of the results. Complications include: photophysical and photochemical effects, such as optical saturation and photobleaching; the influence of diffusion; possible interactions of the chromophore with the polypeptide, resulting in a reduction of quantum yields or lack of fast orientational aver-

aging of the dyes; and the change of solvent conditions. A suitably labeled molecule with well-characterized properties that can be prepared at high purity can thus be extremely valuable for avoiding misinterpretation of the results.

For practical reasons, the first single-molecule FRET experiments<sup>[76]</sup> were performed on DNA duplexes as stiff linkers, because they allowed the individual labeling of the two complementary oligonucleotide strands and subsequent annealing. For FRET experiments on proteins, however, it is desirable to use a polypeptide-based reference molecule, because the type of attachment chemistry and the characteristics of the immediate molecular environment can influence the photophysical properties of the fluorophores. For instance, strong interactions such as the stacking of fluorophores to the ends of the DNA double helix<sup>[111,112]</sup> or strong electrostatic repulsion of a charged dye due to the polyanionic character of nucleic acids<sup>[111,113]</sup> may lead to biopolymer-specific orientational and photophysical effects.

The classic candidate for a suitable polypeptide is poly-L-proline, which was used in Stryer and Haugland's experiment<sup>[79]</sup> owing to its stiff helical structure in aqueous solution.<sup>[114,115]</sup> Oligomers of proline in water form a type II helix with a pitch of 0.312 nm per residue. By including an amino terminal glycine residue and a carboxy terminal cysteine residue in the synthesis, the resulting  $\alpha$ -amino group and the cysteine residue's sulfhydryl group can be labeled specifically with derivatives of suitable reactive dyes such as succinimidyl esters and maleimides, respectively.<sup>[96,97]</sup> Recently, we studied the transfer efficiency between Alexa Fluor 488 and Alexa Fluor 594 attached to the termini of polyproline peptides as a function of the number of proline residues in single-molecule FRET experiments (Figure 3).<sup>[97]</sup> With increasing peptide length, a monotonic decrease in transfer efficiency was observed, but for peptides with more than about 20 residues, the transfer efficiency



**Figure 3.** A) Schematic structure of an icosaproline helix labeled with donor (Alexa 488) and acceptor (Alexa 594) dyes. B) Transfer efficiency histograms from confocal single-molecule measurements on polyproline peptides of various lengths.<sup>[97]</sup> C) End-to-end distance distributions obtained from molecular dynamics simulations of polyproline peptides containing 10, 15, 20, 25, 30, 35, and 40 proline residues plus terminal glycine and cysteine residues (black lines) and the least-squares fits to a wormlike chain model (red lines).<sup>[97]</sup> D) Mean transfer efficiencies from single-molecule measurements (filled red circles) and ensemble time-correlated single-photon counting measurements (empty red circles) as a function of the contour lengths of the peptides (Gly-Pro<sub>n</sub>-Cys), assuming the geometry of polyproline found in the crystal structure,<sup>[114]</sup> in comparison to the dependences calculated for different dynamic regimes (squares) using the normalized end-to-end distance distributions from the molecular dynamics simulations of the peptides. The corresponding lines are empirical fits to the data.<sup>[97]</sup>

was significantly higher than expected for a rigid rod with the structure of a type II helix (Figure 3B). From a quantitative analysis of the end-to-end distance distributions obtained from molecular dynamics simulations (Figure 3C), we were able to assign this increased FRET efficiency to the bending of the chains on a nanosecond time scale (Figure 3D). The persistence length obtained from the molecular dynamics simulations was about 5 nm.<sup>[97]</sup> Polyproline peptides are therefore suitable reference molecules, but their chain dynamics have to be taken into account, especially for higher oligomers.

## 4. Single-Molecule Protein Folding

The basic idea of a protein-folding experiment using FRET is shown in Figure 4. A donor dye and an acceptor dye are attached to the termini of a protein. If a folded protein molecule resides in the volume illuminated by the focused laser beam, excitation of the donor dye results in rapid energy transfer to the acceptor dye, because the termini are in close proximity, and the majority of the fluorescence photons are emitted by the acceptor. Upon addition of chemical denaturant, the protein unfolds, resulting in a larger average distance between the donor and acceptor dyes. Consequently, the energy transfer rate is decreased, and the fraction of photons emitted by the acceptor is lower.

The first experiment of this kind was published by Jia et al.<sup>[116]</sup> in 1999. In a joint effort between the groups of Robin Hochstrasser and William DeGrado, a variant of the GCN4 coiled coil peptide was labeled, then crosslinked to form dimers, and nonspecifically immobilized on a surface to obtain a construct suitable for two-color confocal fluorescence microscopy. Variations in the concentration of the denaturing agent urea resulted in changes of the transfer efficiency and thus showed that it was indeed possible to study the folding of single protein molecules using optical methods. The same group later provided a more quantitative analysis of their experiments, including a direct comparison of GCN4 molecules freely diffusing in solution and immobilized on a aminopropyl-silanized microscope coverslip.<sup>[117]</sup> In a correlation analysis of their data, they observed folding dynamics on a time scale in agreement with ensemble experiments, but also showed that unfolded peptide had its conformational fluctuations slowed, probably because of interactions with the surface.

Another breakthrough was the study of chymotrypsin inhibitor 2 (CI2) by Deniz et al. in 2000.<sup>[99]</sup> Again, a spectroscopy and a biochemistry group had teamed up, managed to produce suitably labeled protein samples using a combination of solid-phase synthesis and native chemical ligation, and studied

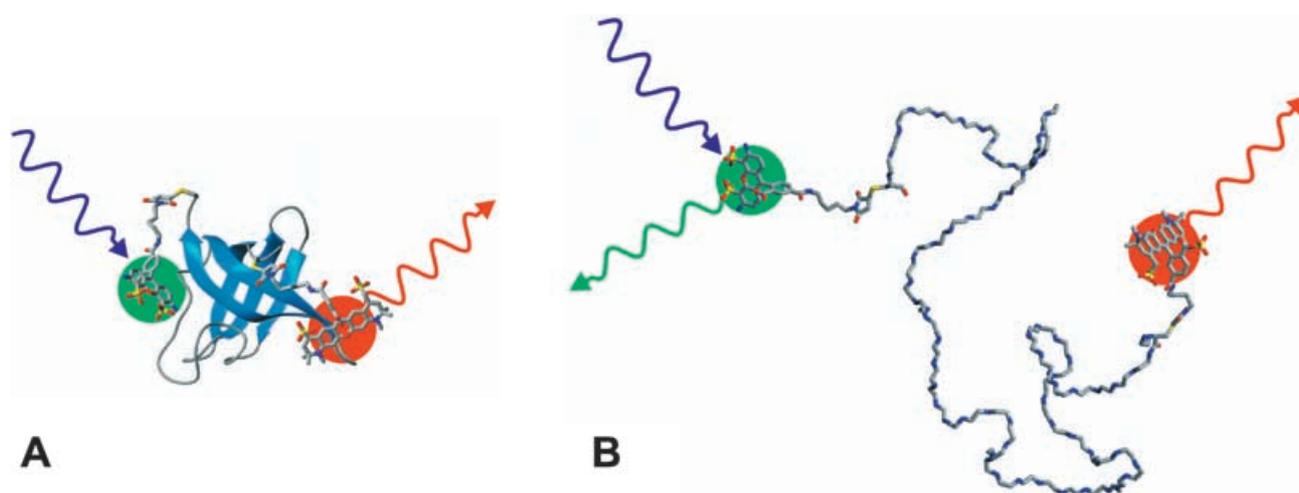
them freely diffusing in solution by confocal fluorescence microscopy. For the first time, they clearly separated the FRET signal from unfolded and folded molecules coexisting in solution, and thus demonstrated the power of single-molecule spectroscopy for separating subpopulations of heterogeneous mixtures for the case of protein folding. However, the lack of a suitable control prevented a rigorous analysis of some of the results.<sup>[96]</sup>

### 4.1. The Cold-Shock Protein

The protein most extensively studied with optical single-molecule methods is the cold-shock protein from the hyperthermophilic bacterium *Thermotoga maritima* (CspTm, Figure 4). It forms a simple, five-stranded beta barrel structure, and by all criteria investigated to date it behaves as a perfect two-state folding system.<sup>[118–120]</sup> In other words, all known equilibrium and kinetic folding data for this protein can be analyzed quantitatively with a model assuming only two thermodynamic states, folded and unfolded, separated by a single activation barrier. Additionally, the high stability of this thermophilic protein made it a promising candidate for mutagenesis and fluorophore attachment without interfering with the folding mechanism. For single-molecule studies, the protein sequence was modified using site-directed mutagenesis, expressed with two cysteine residues close to the termini and labeled with maleimide derivatives of the Alexa Fluors 488 and 594 as donor and acceptor, respectively.<sup>[96]</sup> Specific labeling was achieved by sequential reactions with the two dyes and chromatographic purification after each labeling step, allowing even the separation of the labeling permutants.

### 4.2. Equilibrium Experiments on Freely Diffusing Molecules

In a first series of equilibrium experiments, freely diffusing labeled CspTm molecules were observed with a two-color confocal instrument as illustrated in Figure 2.<sup>[96]</sup> The fluorescence



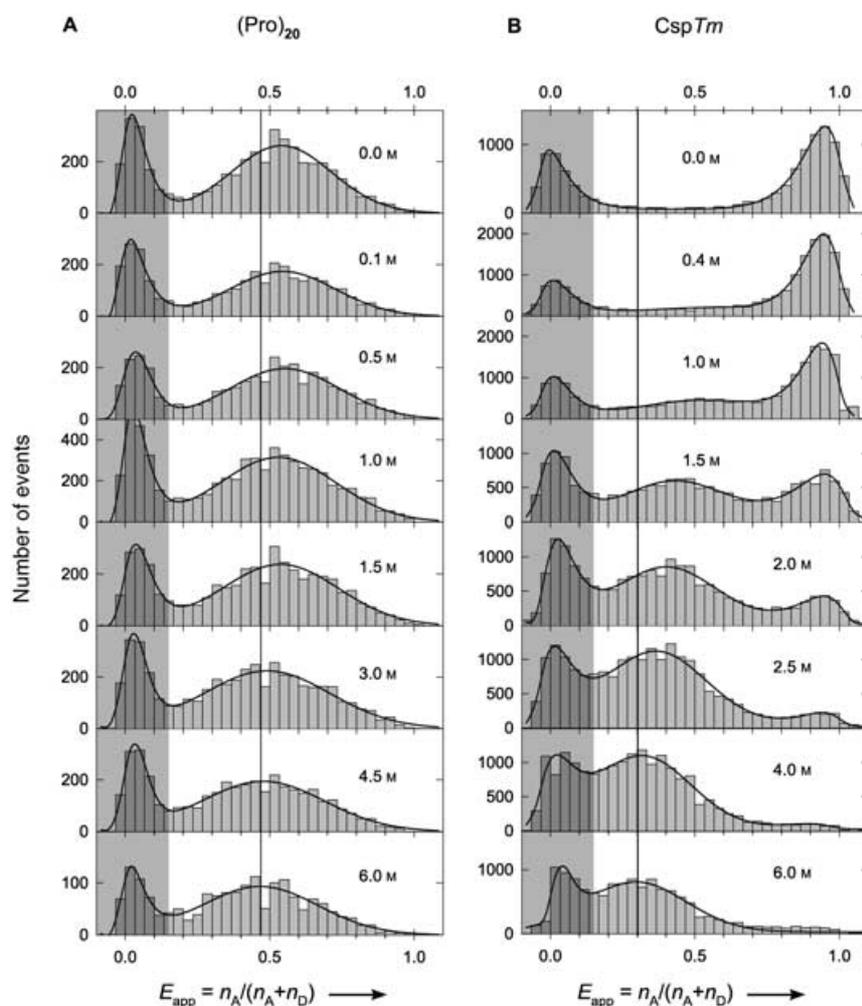
**Figure 4.** Schematic structures<sup>[162]</sup> of folded and unfolded protein labeled with donor (Alexa 488) and acceptor (Alexa 594) dyes. A) Folded CspTm, a 5-stranded, 66-residue  $\beta$ -barrel protein (PDB-code 1G6P),<sup>[163]</sup> B) unfolded CspTm. A blue laser excites the green-emitting donor dye, which can transfer excitation energy to the red-emitting acceptor dye.

emission of the donor and acceptor dyes is separated by a dichroic mirror and detected separately using avalanche photodiodes. The type of resulting signals is shown in Figure 2. Intermittent photon bursts, corresponding to individual molecules diffusing through the confocal volume, are identified using a suitable combination of thresholds. From each burst, a transfer efficiency is calculated according to Equation (2) after background subtraction, and entered into a histogram. Typically, a few thousand bursts are used to ensure reliable statistics.

Figure 5B shows a series of such histograms measured at different concentrations of the denaturant guanidinium chloride (GdmCl), which is used to shift the equilibrium between folded and unfolded molecules. At low denaturant concentrations, the histograms are dominated by events with high FRET efficiency from folded molecules, but at higher GdmCl concentrations, a second peak is observed at lower efficiencies, corresponding to unfolded molecules. This unfolding transition nicely illustrates the existence of two thermodynamic states separated by a free-energy barrier, similar to the results ob-

tained with Cl2.<sup>[99]</sup> More importantly, however, the mean transfer efficiency of the subpopulation of unfolded proteins shows a clear shifting of the peak with decreasing GdmCl concentration, whereas the mean efficiency of folded molecules remains constant. In order to account for possible effects of the solvent on the photophysics or photochemistry of the fluorophores, the same denaturant dependence was analyzed with a control molecule, an icosaprolin peptide (Figure 3A) labeled with the same dyes as the protein (Figure 5B). In this case, only a very slight shift of the efficiency peak was observed, which can be accounted for by the change in refractive index of the solution. Consequently, the much larger decrease in transfer efficiency observed for unfolded *CspTm* molecules at low GdmCl concentrations must correspond to a real distance change—a collapse in response to the change in solvent conditions. Collapsed *CspTm* only becomes noticeable under conditions where the large majority of molecules are folded, and therefore it could not be observed in a corresponding equilibrium ensemble experiment. The strength of the single-molecule approach is that

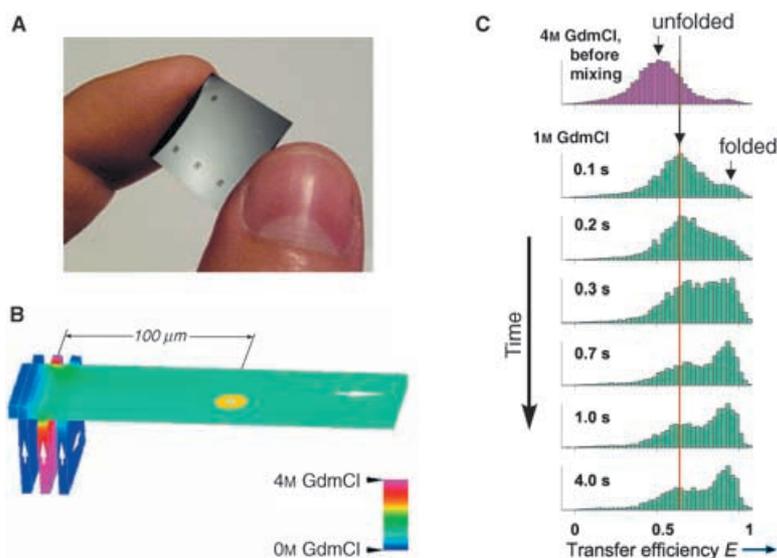
the subpopulations can be separated, and investigated independently.



**Figure 5.** Histograms of measured FRET efficiencies at various GdmCl concentrations for labeled *CspTm* (A) and  $(Pro)_{20}$  (B). The black curves are the best fits to the data using lognormal and/or Gaussian functions. The vertical lines indicate the mean transfer efficiency of A) the unfolded subpopulation or B)  $(Pro)_{20}$  at 6 M GdmCl. The peak at transfer efficiencies near zero (shaded in grey) is due to molecules lacking an acceptor chromophore, owing either to residual impurities or to photobleaching.

### 4.3. Folding Kinetics Using Microfluidic Mixing

A shortcoming of the equilibrium experiment is that, owing to the small population of unfolded molecules at low denaturant concentrations, their mean transfer efficiency cannot be determined with high accuracy over the entire concentration range. A way to solve this problem is the transient population of the unfolded state in a kinetic experiment. In view of the optical requirements of confocal detection, microfluidic laminar-flow mixing devices<sup>[121,122]</sup> provide an ideal technical solution to this problem.<sup>[123]</sup> Several micrometer-sized channels are etched into a silicon chip, which is then bonded to a glass cover slide (Figure 6A,B), allowing excitation and emission in the visible wavelength range. Solutions containing protein, denaturant, and buffer can then be driven through the channels using compressed air and mixed in the region where the channels merge. The Reynolds numbers in these experiments are small enough (approximately  $10^{-2}$ ) to



**Figure 6.** Protein-folding kinetics studied by microfluidic mixing. A) Photograph of the mixing device. B) Schematic of the mixing area, with the three inlet channels coming in from the bottom. GdmCl concentrations calculated from its diffusion coefficient and the flow rate are indicated on a color scale. The  $1/e^2$  intensity contour of the laser beam (light blue) is illustrated along with the cone of fluorescence emission collected by the microscope objective (yellow). C) Histograms of measured FRET efficiencies before (topmost panel) and at different distances downstream of the mixing area, corresponding to different times after mixing. The vertical red line indicates the mean FRET efficiency value in the unfolded state after mixing.

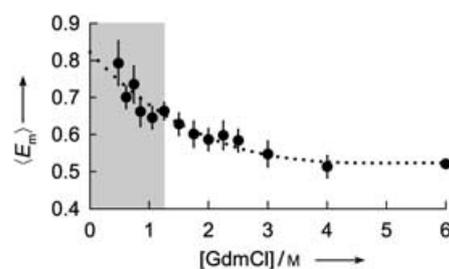
guarantee laminar flow, and the microscopic dimensions of the mixer channels ensure that rapid mixing can occur by diffusion alone. Consequently, the concentrations of all solutes can be calculated in the entire mixing region from their diffusion coefficients (Figure 6B). Under the conditions used by Lipman et al.,<sup>[123]</sup> displacements due to diffusion and flow are comparable over a period of a few milliseconds, and the solutions are mixed by the time they have reached a point  $50\ \mu\text{m}$  beyond the mixing region. The laser beam is positioned downstream of the mixing region at distances chosen to correspond, via the flow rate, to the desired delays following mixing. Data are acquired in an analogous manner to the experiments on freely diffusing molecules (see Section 3.2.), and transfer efficiency histograms are accumulated at different times after mixing (Figure 6C).

After initiating the reaction, a redistribution of the populations was observed (Figure 6C). As the number of unfolded molecules decreased, there was a corresponding increase in the number of folded molecules, and eventually the distribution converged to the equilibrium at the final concentration of denaturant. The positions of the peaks, however, remained constant, indicating that the average end-to-end distances of the molecules in the subpopulations did not change during the course of the reaction. This is the behavior expected for a two-state system. However, although the peak positions are constant during the folding reaction, the mean FRET efficiency of the unfolded molecules after mixing was increased compared to the situation before mixing. The lower denaturant concentration after dilution results in more compact unfolded molecules, which are populated within  $50\ \text{ms}$  (the dead time of the mixer) and exhibit higher FRET efficiencies.<sup>[123]</sup> It is note-

worthy that qualitatively different signal changes result from collapse and folding. Collapse of the unfolded state causes a shift of the corresponding peak to higher efficiency, whereas folding increases the folded-state population (as measured by peak area) and depopulates the unfolded state. In an ensemble FRET experiment, both collapse and folding would result in an overall increase in the transfer efficiency, and the respective contributions could be identified only indirectly by kinetic modeling. With pressure-jump experiments, it has now been shown that the collapse process of cold-shock proteins occurs on a time scale shorter than  $50\ \mu\text{s}$ ,<sup>[124]</sup> in agreement with the upper limit of unfolded chain dynamics obtained from single-molecule spectroscopy (see Section 4.4.). It remains to

be established whether this reaction is merely a nonspecific heteropolymer collapse in response to altered solvent conditions or whether specific parts of the polypeptide chain become structured under native conditions prior to the actual folding process.

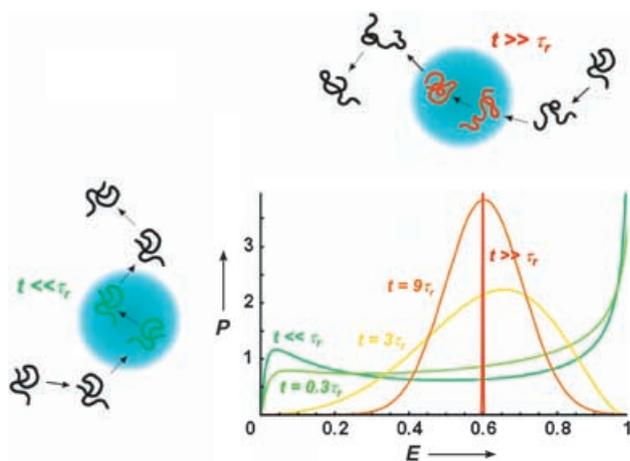
The transient population of unfolded molecules under native conditions also allows a more accurate determination of their mean FRET efficiencies at low denaturant concentrations, which—compared to the equilibrium experiment (see Section 4.2.)—has allowed us to extend the accessible concentration range substantially (Figure 7).<sup>[123]</sup> This brings us closer to the physiologically relevant conditions, that is, in the absence of denaturants. The structure and dynamics of the unfolded state under native conditions are a potentially very important determinant of the folding reaction and are therefore under intense investigation.<sup>[49,125]</sup>



**Figure 7.** Dependence of the mean transfer efficiencies for unfolded CspTm as a function of GdmCl concentration obtained in microfluidic mixing experiments. The dotted line shows a third-degree polynomial fit to the data. The shaded region indicates the range of denaturant concentrations where reliable data could not be obtained from corresponding equilibrium experiments.

#### 4.4. Chain Dynamics and the Protein-Folding “Speed Limit”

Even equilibrium single-molecule experiments provide the opportunity to obtain information on protein dynamics, particularly those of the unfolded state.<sup>[96,126,127]</sup> Figure 8 illustrates



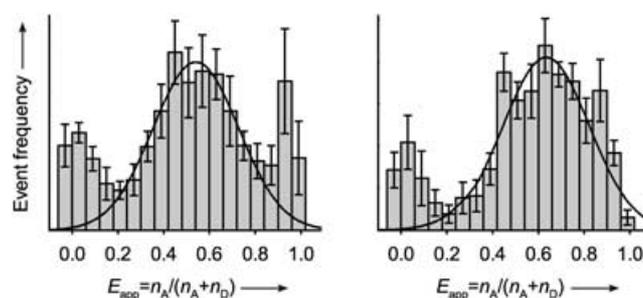
**Figure 8.** Chain dynamics and the time scale of observation. The diffusion of molecules through the focal spot is shown schematically. In one case (green), the reconfiguration time of the chain  $\tau_r$  is large relative to the observation time  $t$ , resulting in a very broad transfer efficiency distribution. At the other extreme (red),  $\tau_r$  is much less than  $t$ , and theoretically a delta function is expected for the transfer efficiency distribution of all molecules. Intermediate cases are shown in light green, yellow, and orange.

two limiting scenarios. If, on the one hand, chain dynamics are very fast relative to the observation time for a single-molecule (determined by the diffusion time through the confocal volume), the chain explores a large fraction of its conformational space during this time, resulting in complete averaging of end-to-end distances, and a single observed value for the transfer efficiency for all molecules. If, on the other hand, chain dynamics are very slow relative to the observation time, every molecule may enter the confocal volume with a different conformation, which essentially remains constant during the time the molecule is observed. As a result, a different FRET efficiency is measured for every molecule, yielding a very broad distribution of transfer efficiencies for many molecules. For intermediate cases, distributions between these two extremes are expected. This type of “line broadening” can be used to extract dynamic information from transfer efficiency histograms.<sup>[126,128]</sup>

However, the issue is complicated by the fact that, even for a molecule with a single fixed distance or very rapid conformational averaging, the resulting FRET efficiency histogram does not resemble a delta function. Experimental distributions are broadened by shot noise (the variation in count rates about fixed means due to the discrete nature of the signals) and other as yet unidentified sources, possibly nonrandom photon emission intervals caused by triplet state formation or intensity variation across the focal volume. Quantifying these contributions will require detailed experimental and theoretical investi-

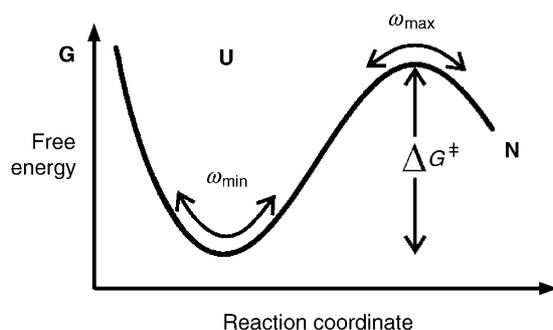
gation, but, with a suitable reference molecule, dynamic information can be obtained even without a detailed knowledge of such intrinsic broadening. For this purpose, we have used polyproline peptides (see Section 3.4.), which have a narrow end-to-end distance distribution due to their large persistence length.<sup>[97]</sup> The transfer efficiency distribution determined for such a rodlike molecule in a free diffusion measurement thus reflects the intrinsic width of the distribution and can serve as a reference. For a molecule with conformational dynamics that are slow on the time scale of diffusion through the focal volume, we would expect a broadening of the distribution relative to this reference. Conversely, a lack of broadening relative to the polyproline peptide would indicate fast conformational averaging and allows us to place limits on the autocorrelation time of the end-to-end distance dynamics.

It was shown by Gopich and Szabo<sup>[126]</sup> that, for a Gaussian chain with a mean squared end-to-end distance  $\langle r^2 \rangle = R_0^2$ , the polypeptide reconfiguration time is  $\tau_0 = 9.8t(\sigma_{\text{app}}^2 - \sigma_0^2)$ , where  $\sigma_{\text{app}}^2$  is the variance in the measured transfer efficiencies for the unfolded protein, and  $\sigma_0^2$  is the variance due to noise and other, non-interdye distance effects. Equating  $\sigma_0^2$  with the variance determined for icosaproline, and recognizing that, to within experimental error,  $\sigma_{\text{app}}$  cannot be more than 25% larger than  $\sigma_0$ , a maximum value for  $\tau_0$  of 200  $\mu\text{s}$  was determined by Schuler et al.<sup>[96,129a]</sup> More recent measurements using a shorter observation period of 100  $\mu\text{s}$  (Figure 9) show that  $\tau_0 < 30 \mu\text{s}$ ,<sup>[129b]</sup> consistent with estimates from other methods.<sup>[42]</sup> Recently, it was shown that similar information can be obtained directly from fluorescence correlation spectroscopy.<sup>[127]</sup>



**Figure 9.** Analysis of the width of the transfer efficiency histogram of unfolded protein (left panel,  $\sigma_{\text{CSP}}$ ) molecules relative to the polyproline reference (right panel,  $\sigma_{\text{PPD}}$ ). As  $\sigma_{\text{CSP}}$  is not significantly greater than  $\sigma_{\text{PPD}}$ , chain dynamics in the unfolded state must be fast relative to the observation time of 100  $\mu\text{s}$ .

The conformational dynamics in the unfolded state are closely related to the maximum rate at which the protein could fold in the absence of an activation barrier, the protein-folding “speed limit”, and can therefore provide an estimate for the elusive pre-exponential factor in a description of protein-folding kinetics using reaction-rate theory, as illustrated in Figure 10. According to Kramers’ theory of unimolecular reaction rates in solution,<sup>[130,131]</sup> a reaction can be described as a diffusive process along a reaction coordinate on a free-energy surface. Assuming parabolic profiles for the free-energy minimum corresponding to the unfolded state and the activation



**Figure 10.** One-dimensional free-energy surface for a protein-folding reaction described with Kramers' theory.

barrier, the folding rate coefficient is given by Equation (5):

$$k_f = \frac{\omega_{\min} \omega_{\max} D_{\max}}{2\pi kT} \exp(-\Delta G^\ddagger / kT) \quad (5)$$

where  $\omega_{\min}$  and  $\omega_{\max}$  are frequencies that characterize the curvature of the free-energy profile at the unfolded well and (inverted) barrier top, respectively;  $D_{\max}$  is the diffusion constant at the barrier top;  $\Delta G^\ddagger$  is the height of the free-energy barrier;  $k$  is Boltzmann's constant; and  $T$  is the absolute temperature (a unit mass has been assumed for the fictitious particle diffusing on this surface). If  $\omega_{\min} \approx \omega_{\max}$  and  $D_{\max} \approx D_{\min}$ , the pre-exponential factor in Equation (5) is essentially given by the reconfiguration rate of the unfolded state,  $k_r = \omega_{\min}^2 D_{\min} / kT$ , yielding  $k_f = k_r / 2\pi \exp(-\Delta G^\ddagger / kT)$ .<sup>[132,133]</sup> This equation can be solved for  $\Delta G^\ddagger$ , if  $k_f$  and  $k_r$  are known, resulting in a way to calculate the free-energy barrier, the central parameter determining the reaction rate. Note that the free-energy barrier contains the activation entropy and is therefore notoriously difficult to determine for any solution reaction. The above result, combined with measurements on contact formation in unstructured peptides,<sup>[134]</sup> yields lower and upper bounds on the free-energy barrier for folding of CspTm of 4 and 11  $kT$ , respectively.<sup>[96]</sup> The determination of bounds on a free-energy barrier height provides a previously unavailable benchmark for theoretical free-energy surfaces.<sup>[135]</sup>

#### 4.5. Folding Trajectories

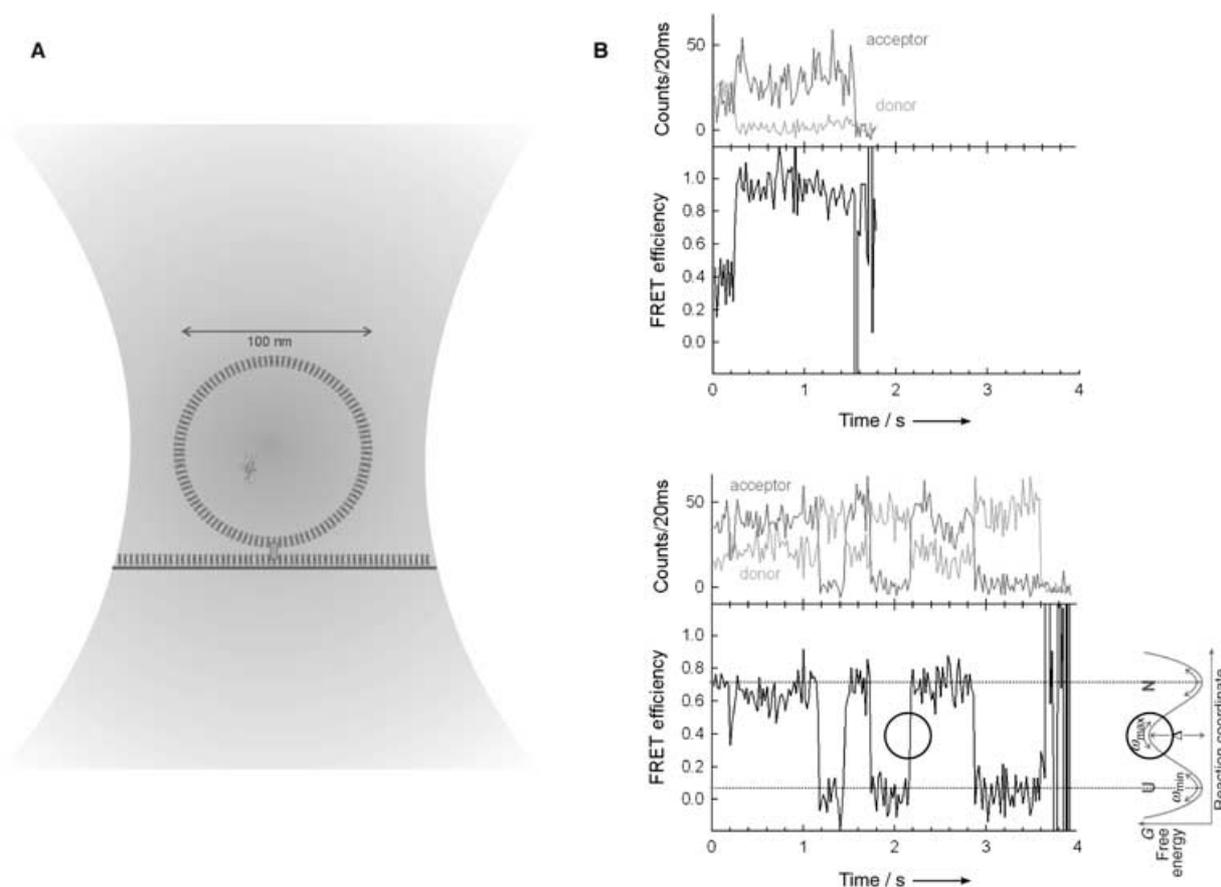
The above examples show that single-molecule experiments on freely diffusing protein molecules can provide a wealth of information on the elementary processes of protein folding, but the observation period for every molecule is limited to a few milliseconds by the diffusion time through the focal volume. One approach for observing individual proteins for an extended period of time is their immobilization on a surface, a technique that has been very successfully utilized for RNA folding experiments.<sup>[67,136]</sup> However, owing to the usually very low conformational stability of proteins, interactions with the surface can easily disturb the folding reaction.<sup>[117]</sup> Strategies for minimizing such interactions include the optimization of surface functionalization<sup>[137]</sup> and the encapsulation of individual

protein molecules in surface-tethered lipid vesicles.<sup>[138]</sup> Both methods have allowed the observation of single-molecule protein-folding reactions using RNase H and adenylate kinase, respectively.<sup>[139,140]</sup> In all of the cases where surface-bound vesicles were used successfully, no indications were found for interactions of the molecules under study with the membrane, as shown by single-molecule polarization measurements,<sup>[138,140,141]</sup> the absence of binding to corresponding supported bilayers, and the agreement with results obtained from experiments not using vesicles.<sup>[141,142]</sup> Herein, I want to illustrate the power of the vesicle encapsulation method in experiments on the cold-shock protein CspTm.<sup>[141]</sup>

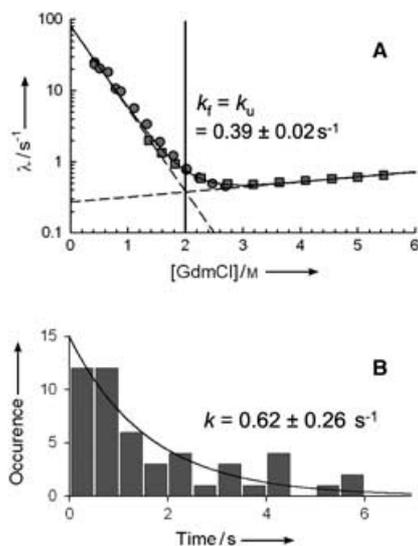
Following a change in conditions that alters the equilibrium population ratio of a two-state system, its relaxation is characterized by an exponential time course. Although the macroscopic relaxation rate may be slow because of the low probability of crossing the barrier, the actual barrier-crossing process for individual molecules is expected to be much shorter than the residence time in each state. Consequently, with the current time-resolution in single-molecule experiments, one expects an individual protein molecule to reside in either the unfolded or the folded state for times given by the inverse mean values of the folding and unfolding rate coefficients, with intermittent and irresolvable jumps between the two states.<sup>[43]</sup> However, a direct observation of this bistable behavior had not been made, and previous reports of folding trajectories of individual proteins have shown an unexpected degree of complexity.<sup>[140,143–145]</sup> This raises the question whether the simple kinetic properties derived from classical experiments on large ensembles of molecules are reflected in the folding paths taken by individual proteins.

To test this idea, individual, labeled CspTm molecules were encapsulated in unilamellar lipid vesicles in aqueous buffer containing 2 M GdmCl, and were then surface-tethered using biotin-avidin chemistry (Figure 11 A). Under these solution conditions, the rates of the folding and unfolding reactions are equal, and consequently folded and unfolded states are equally populated<sup>[96]</sup> (Figure 12 A). A sample-scanning confocal microscope was used to locate individual molecules of CspTm and record fluorescence intensity traces. Donor and acceptor photons were collected separately, allowing the calculation of FRET efficiencies as a function of time. Two such measurements are shown in Figure 11 B. Steady levels of FRET efficiency are followed by rapid jumps, until photobleaching of one of the dyes occurs. Based on the previous FRET experiments with CspTm,<sup>[96]</sup> high transfer efficiencies were identified with the folded state and low transfer efficiencies with the unfolded state of the protein. Consequently, abrupt changes in the FRET efficiency represent folding or unfolding events. The actual transitions were too rapid to be time-resolved, even with a sampling period of 100  $\mu$ s. This behavior is exactly what would be predicted from the notion of a barrier-crossing process between two well-defined thermodynamic states, as described above.

The simplicity of the trajectories and the stochastic occurrence of folding and unfolding permitted a quantitative comparison with ensemble experiments. From the trajectories ob-



**Figure 11.** Experiments on protein molecules immobilized in surface-tethered lipid vesicles. A) Schematic of *CspTm* encapsulated within a surface-tethered vesicle (relative dimensions not to scale). The laser beam is indicated in grey. B) Folding trajectories of individual protein molecules showing the characteristic steplike transitions between the folded and unfolded states, corresponding to the actual barrier-crossing process (○).



**Figure 12.** Comparison of folding kinetics obtained from ensemble stopped-flow experiments and single-molecule folding trajectories. A) Dependence of the observed relaxation rate  $\lambda$  of the folding/unfolding reaction of *CspTm* on GdmCl concentration observed in an ensemble fluorescence stopped-flow experiment with the corresponding fit to a two-state model (black line). At the unfolding midpoint (vertical line), the folding and unfolding rate constants  $k_f$  and  $k_u$  are equal. B) Histogram of transition times obtained from 43 protein molecules (54 transitions) fitted to a single exponential decay (black curve). The rate constant obtained from the fit agrees well with the ensemble-averaged folding rate constant.

served, a histogram of all time intervals preceding folding and unfolding transitions was constructed (Figure 12B). An exponential fit to this histogram gave a rate constant in agreement with the folding rate constant determined in an ensemble stopped-flow experiment under identical solution conditions, validating the single-molecule result. Consequently, the simple two-state behavior of *CspTm* inferred from ensemble-averaged experiments is reflected quantitatively in the folding trajectories of individual molecules. Extensions of studies of this type might allow a measurement of the heterogeneity of protein stabilities and folding kinetics within a population, essentially a test of the ergodic hypothesis for protein folding, or the investigation of protein folding in vivo.

## 5. Summary and Outlook

From the successes in the past five years, it is to be expected that single-molecule studies will play an increasingly important role in our understanding of protein folding. Single-molecule FRET experiments on freely diffusing protein molecules can be used to investigate equilibrium and dynamic properties that are difficult to obtain from ensemble experiments. They allow the direct enumeration of thermodynamic states, the investigation of intramolecular distances in unfolded proteins at equilib-

rium (even under conditions favoring the folded state), and they provide information about the reconfigurational dynamics in the unfolded state. To study single-molecule folding kinetics, microfabricated laminar-flow mixers can be coupled to a confocal fluorescence detection system. This way, folding and unfolding can be initiated by an abrupt change in denaturant concentration under continuous flow, which allows us to follow the evolution of the intramolecular distance distribution as folding progresses. Finally, studies on molecules immobilized in surface-tethered lipid vesicles enable the observation of folding trajectories of individual proteins and have provided the first, model-free demonstration of two-state protein folding.

A next step will be to study molecules with more complicated folding mechanisms, or folding under conditions where the reaction to the correctly folded native state competes with misfolding and aggregation. These processes are closely connected to a range of diseases, such as Alzheimer's, Parkinson's, and Huntington's disease, and type II diabetes (where protein misfolding and aggregation lead to the accumulation of pathological deposits, or amyloid). The mechanisms involved in their formation can be expected to be even more heterogeneous than those of protein folding and will profit greatly from single-molecule fluorescence detection, which allows the separation of subpopulations<sup>[146]</sup> or the assembly of individual aggregates to be observed in real time.<sup>[147,148]</sup> For these problems, it is particularly important to control the effect of the dyes on the stability and aggregation behavior of the polypeptides under study.

Another area for which single-molecule approaches hold a lot of promise is the study of accessory proteins and other factors that influence protein folding in vivo.<sup>[149,150]</sup> Single-molecule studies on the mechanism of the molecular chaperone GroE in vitro have already been reported,<sup>[151–153]</sup> and the coming-of-age of single-molecule studies in vivo<sup>[154]</sup> might soon allow us to address issues such as the effect of co-translational folding on the ribosome,<sup>[155]</sup> macromolecular crowding,<sup>[156]</sup> and the entire range of molecular chaperones on protein folding and stability in a cellular environment.

An ultimate goal is the time-resolved observation of the barrier-crossing processes of protein folding. Ideally, we could then watch parts of the protein chain come together and form the native structure, where each molecule might be expected to take a slightly different route to the free-energy minimum. However, an intrinsic limitation of fluorescence detection from individual molecules is that the photon emission rate cannot be greater than the decay rate of the electronically excited state of the chromophores used, typically of the order of  $10^9 \text{ s}^{-1}$ . In view of other complications, such as the population of triplet states or nonideal photon collection efficiencies, observing processes on microsecond time scales and faster will require new ideas from spectroscopy, organic chemistry, biochemistry and theory.<sup>[126,128,157–161]</sup> The marriage of these disciplines has proven very successful for protein folding in the past and will certainly continue to provide new insights and challenges.

## Acknowledgements

I thank Olgica Bakajin, William Eaton, Irina Gopich, Gilad Haran, Everett Lipman, and Attila Szabo for many inspiring discussions; Everett Lipman for providing parts of Figure 2, Irina Gopich for providing part of Figure 8; and the Deutsche Forschungsgemeinschaft (Emmy Noether-Programm) and the VolkswagenStiftung for financial support.

**Keywords:** emission spectroscopy · FRET (fluorescence resonance energy transfer) · protein folding · proteins · single-molecule studies

- [1] H. Wu, *Chin. J. Physiol.* **1931**, *5*, 321–344.
- [2] M. L. Anson, A. E. Mirsky, *J. Gen. Physiol.* **1934**, *17*, 393–408.
- [3] M. L. Anson, *Adv. Protein Chem.* **1945**, *2*, 361–384.
- [4] H. Neurath, J. P. Greenstein, F. W. Putnam, J. O. Erickson, *Chem. Rev.* **1944**, *34*, 157–265.
- [5] M. Sela, F. H. White, Jr., C. B. Anfinsen, *Science* **1957**, *125*, 691–692.
- [6] C. B. Anfinsen, *Biochem. J.* **1972**, *128*, 737–749.
- [7] C. B. Anfinsen, *Science* **1973**, *181*, 223–230.
- [8] C. Levinthal, *J. Chim. Phys. Phys.-Chim. Biol.* **1968**, *65*, 44–45.
- [9] D. B. Wetlauffer, *Proc. Natl. Acad. Sci. USA* **1973**, *70*, 697–701.
- [10] T. Y. Tsong, R. L. Baldwin, *J. Mol. Biol.* **1972**, *63*, 453–469.
- [11] P. S. Kim, R. L. Baldwin, *Annu. Rev. Biochem.* **1982**, *51*, 459–489.
- [12] *Protein Folding* (Ed.: R. Jaenicke), Elsevier, Amsterdam, **1980**.
- [13] T. E. Creighton, *Protein Folding*, W. H. Freeman and Company, New York, **1992**.
- [14] O. B. Ptitsyn, A. A. Rashin, *Biophys. Chem.* **1975**, *3*, 1–20.
- [15] M. Karplus, D. L. Weaver, *Nature* **1976**, *260*, 404–406.
- [16] M. I. Kanehisa, T. Y. Tsong, *Biopolymers* **1979**, *18*, 1375–1388.
- [17] N. Go, H. Abe, H. Mizuno, H. Taketomi in *Protein folding* (Ed.: R. Jaenicke), Elsevier, Amsterdam, **1980**, pp. 167–181.
- [18] S. C. Harrison, R. Durbin, *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 4028–4030.
- [19] A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 10869–10873.
- [20] M. Karplus, D. L. Weaver, *Protein Sci.* **1994**, *3*, 650–668.
- [21] S. E. Jackson, A. R. Fersht, *Biochemistry* **1991**, *30*, 10428–10435.
- [22] S. E. Jackson, *Folding Des.* **1998**, *3*, R81–91.
- [23] J. D. Bryngelson, P. G. Wolynes, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 7524–7528.
- [24] J. N. Onuchic, Z. Luthey Schulten, P. G. Wolynes, *Annu. Rev. Phys. Chem.* **1997**, *48*, 545–600.
- [25] K. A. Dill, H. S. Chan, *Nat. Struct. Biol.* **1997**, *4*, 10–19.
- [26] E. I. Shakhnovich, *Curr. Opin. Struct. Biol.* **1997**, *7*, 29–40.
- [27] C. M. Dobson, A. Sali, M. Karplus, *Angew. Chem.* **1998**, *110*, 908–935; *Angew. Chem., Int. Ed.* **1998**, *37*, 868–893.
- [28] L. Serrano, A. Matouschek, A. R. Fersht, *J. Mol. Biol.* **1992**, *224*, 805–818.
- [29] A. R. Fersht, *Structure and Mechanism in Protein Science*, W. H. Freeman and Company, New York, **1998**.
- [30] O. Bieri, J. Wirz, B. Hellrung, M. Schutkowski, M. Drewello, T. Kiefhaber, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 9597–9601.
- [31] L. J. Lapidus, W. A. Eaton, J. Hofrichter, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 7220–7225.
- [32] S. J. Hagen, J. Hofrichter, A. Szabo, W. A. Eaton, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 11615–11617.
- [33] O. Bieri, T. Kiefhaber, *Biol. Chem.* **1999**, *380*, 923–929.
- [34] H. Neuweiler, M. Sauer, *Curr. Pharm. Biotechnol.* **2004**, *5*, 285–298.
- [35] H. Neuweiler, A. Schulz, M. Böhmer, J. Enderlein, M. Sauer, *J. Am. Chem. Soc.* **2003**, *125*, 5324–5330.
- [36] W. A. Eaton, V. Muñoz, S. J. Hagen, G. S. Jas, L. J. Lapidus, E. R. Henry, J. Hofrichter, *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 327–359.
- [37] V. Muñoz, P. A. Thompson, J. Hofrichter, W. A. Eaton, *Nature* **1997**, *390*, 196–199.
- [38] P. A. Thompson, W. A. Eaton, J. Hofrichter, *Biochemistry* **1997**, *36*, 9200–9210.

- [39] F. Krieger, B. Fierz, O. Bieri, M. Drewello, T. Kiefhaber, *J. Mol. Biol.* **2003**, *332*, 265–274.
- [40] R. L. Baldwin, *J. Biomol. NMR* **1995**, *5*, 103–109.
- [41] V. S. Pande, A. Grosberg, T. Tanaka, D. S. Rokhsar, *Curr. Opin. Struct. Biol.* **1998**, *8*, 68–79.
- [42] J. Kubelka, J. Hofrichter, W. A. Eaton, *Curr. Opin. Struct. Biol.* **2004**, *14*, 76–88.
- [43] W. A. Eaton, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 5897–5899.
- [44] J. Sabelko, J. Ervin, M. Gruebele, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 6031–6036.
- [45] W. Y. Yang, M. Gruebele, *Nature* **2003**, *423*, 193–197.
- [46] D. Shortle, *FASEB J.* **1996**, *10*, 27–34.
- [47] N. C. Fitzkee, G. D. Rose, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12497–12502.
- [48] J. E. Kohn, I. S. Millett, J. Jacob, B. Zagrovic, T. M. Dillon, N. Cingel, R. S. Dothager, S. Seifert, P. Thiagarajan, T. R. Sosnick, M. Z. Hasan, V. S. Pande, I. Ruczinski, S. Doniach, K. W. Plaxco, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 12491–12496.
- [49] B. Zagrovic, C. D. Snow, S. Khaliq, M. R. Shirts, V. S. Pande, *J. Mol. Biol.* **2002**, *323*, 153–164.
- [50] C. Frieden, K. Chattopadhyay, E. L. Elson, *Adv. Protein Chem.* **2002**, *62*, 91–109.
- [51] V. Muñoz, W. A. Eaton, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11311–11316.
- [52] O. V. Galzitskaya, A. V. Finkelstein, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11299–11304.
- [53] E. Alm, D. Baker, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11305–11310.
- [54] K. W. Plaxco, K. T. Simons, D. Baker, *J. Mol. Biol.* **1998**, *277*, 985–994.
- [55] R. Jaenicke, R. Seckler, *Adv. Protein Chem.* **1997**, *50*, 1–59.
- [56] C. M. Dobson, *Nature* **2003**, *426*, 884–890.
- [57] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, H. E. Gaub, *Science* **1997**, *276*, 1109–1112.
- [58] A. Engel, H. E. Gaub, D. J. Müller, *Curr. Biol.* **1999**, *9*, R133–136.
- [59] W. E. Moerner, *J. Phys. Chem. B* **2002**, *106*, 910–927.
- [60] E. Haustein, P. Schwillie, *Curr. Opin. Struct. Biol.* **2004**, *14*, 531–540.
- [61] X. Michalet, A. N. Kapanidis, T. Laurence, F. Pinaud, S. Doose, M. Pflughoefft, S. Weiss, *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 161–182.
- [62] P. Tamarat, A. Maali, B. Lounis, M. Orrit, *J. Phys. Chem. A* **2000**, *104*, 1–16.
- [63] T. Plakhotnik, E. A. Donley, U. P. Wild, *Ann. Rev. Phys. Chem.* **1997**, *48*, 181–212.
- [64] X. S. Xie, *Acc. Chem. Res.* **1996**, *29*, 598–606.
- [65] P. M. Goodwin, W. P. Ambrose, R. A. Keller, *Acc. Chem. Res.* **1996**, *29*, 607–613.
- [66] M. S. Kellermayer, S. B. Smith, H. L. Granzier, C. Bustamante, *Science* **1997**, *276*, 1112–1116.
- [67] X. Zhuang, M. Rief, *Curr. Opin. Struct. Biol.* **2003**, *13*, 88–97.
- [68] R. B. Best, J. Clarke, *Chem. Commun.* **2002**, 183–192.
- [69] M. Carrion-Vazquez, A. F. Oberhauser, T. E. Fisher, P. E. Marszalek, H. Li, J. M. Fernandez, *Prog. Biophys. Mol. Biol.* **2000**, *74*, 63–91.
- [70] P. M. Williams, S. B. Fowler, R. B. Best, J. L. Toca-Herrera, K. A. Scott, A. Steward, J. Clarke, *Nature* **2003**, *422*, 446–449.
- [71] M. Rief, H. Grubmüller, *ChemPhysChem* **2002**, *3*, 255–261.
- [72] E. Paci, J. Clarke, A. Steward, M. Vendruscolo, M. Karplus, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 394–399.
- [73] M. Böhmer, J. Enderlein, *ChemPhysChem* **2003**, *4*, 793–808.
- [74] G. Haran, *J. Phys. Condens. Matter* **2003**, *15*, R1291–R1317.
- [75] T. Förster, *Ann. Phys.* **1948**, *6*, 55–75.
- [76] T. Ha, T. Enderle, D. F. Ogletree, D. S. Chemla, P. R. Selvin, S. Weiss, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 6264–6268.
- [77] P. R. Selvin, *Nat. Struct. Biol.* **2000**, *7*, 730–734.
- [78] B. W. Van Der Meer, G. Coker, III, S. Y. S. Chen, *Resonance Energy Transfer: Theory and Data*, VCH, Weinheim, **1994**.
- [79] L. Stryer, R. P. Haugland, *Proc. Natl. Acad. Sci. USA* **1967**, *58*, 719–726.
- [80] L. Stryer, *Annu. Rev. Biochem.* **1978**, *47*, 819–846.
- [81] E. Haas, E. Katchalskikatzir, I. Z. Steinberg, *Biopolymers* **1978**, *17*, 11–31.
- [82] J. R. Lakowicz, *Principles of Fluorescence Spectroscopy*, 2nd ed., Kluwer Academic Publishers, New York, **1999**.
- [83] A. A. Deniz, M. Dahan, J. R. Grunwell, T. J. Ha, A. E. Faulhaber, D. S. Chemla, S. Weiss, P. G. Schultz, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 3670–3675.
- [84] T. Heyduk, *Curr. Opin. Biotechnol.* **2002**, *13*, 292–296.
- [85] A. A. Deniz, T. A. Laurence, M. Dahan, D. S. Chemla, P. G. Schultz, S. Weiss, *Annu. Rev. Phys. Chem.* **2001**, *52*, 233–253.
- [86] C. Eggeling, S. Berger, L. Brand, J. R. Fries, J. Schaffer, A. Volkmer, C. A. Seidel, *J. Biotechnol.* **2001**, *86*, 163–180.
- [87] S. T. Hess, S. Huang, A. A. Heikal, W. W. Webb, *Biochemistry* **2002**, *41*, 697–705.
- [88] M. Eigen, R. Rigler, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 5740–5747.
- [89] T. Ha, *Methods* **2001**, *25*, 78–86.
- [90] S. Hohng, C. Joo, T. Ha, *Biophys. J.* **2004**, *87*, 1328–1337.
- [91] J.-P. Clamme, A. A. Deniz, *ChemPhysChem* **2005**, *6*, 74–77.
- [92] M. Wahl, F. Koberling, M. Patting, H. Rahn, R. Erdmann, *Curr. Pharm. Biotechnol.* **2004**, *5*, 299–308.
- [93] O. Shimomura, *J. Microsc.* **2005**, *217*, 3–15.
- [94] Y. Sako, T. Uyemura, *Cell Struct. Funct.* **2002**, *27*, 357–365.
- [95] M. Lippitz, W. Erker, H. Decker, K. E. van Holde, T. Basche, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2772–2777.
- [96] B. Schuler, E. A. Lipman, W. A. Eaton, *Nature* **2002**, *419*, 743–747.
- [97] B. Schuler, E. A. Lipman, P. J. Steinbach, M. Kumke, W. A. Eaton, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2754–2759.
- [98] P. E. Dawson, S. B. Kent, *Annu. Rev. Biochem.* **2000**, *69*, 923–960.
- [99] A. A. Deniz, T. A. Laurence, G. S. Beligere, M. Dahan, A. B. Martin, D. S. Chemla, P. E. Dawson, P. G. Schultz, S. Weiss, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5179–5184.
- [100] V. Ratner, E. Kahana, M. Eichler, E. Haas, *Bioconjugate Chem.* **2002**, *13*, 1163–1170.
- [101] A. N. Kapanidis, S. Weiss, *J. Chem. Phys.* **2002**, *117*, 10953–10964.
- [102] R. David, M. P. Richter, A. G. Beck-Sickingler, *Eur. J. Biochem.* **2004**, *271*, 663–677.
- [103] B. Schuler, L. K. Pannell, *Bioconjugate Chem.* **2002**, *13*, 1039–1043.
- [104] J. Yamaguchi, N. Nemoto, T. Sasaki, A. Tokumasu, Y. Mimori-Kiyosue, T. Yagi, T. Funatsu, *FEBS Lett.* **2001**, *502*, 79–83.
- [105] T. A. Cropp, P. G. Schultz, *Trends Genet.* **2004**, *20*, 625–630.
- [106] R. B. Mujumdar, L. A. Ernst, S. R. Mujumdar, C. J. Lewis, A. S. Waggoner, *Bioconjugate Chem.* **1993**, *4*, 105–111.
- [107] N. Panchuk-Voloshina, R. P. Haugland, J. Bishop-Stewart, M. K. Bhalgat, P. J. Millard, F. Mao, W. Y. Leung, *J. Histochem. Cytochem.* **1999**, *47*, 1179–1188.
- [108] C. J. Murphy, *Anal. Chem.* **2002**, *74*, 520A–526A.
- [109] W. C. Chan, D. J. Maxwell, X. Gao, R. E. Bailey, M. Han, S. Nie, *Curr. Opin. Biotechnol.* **2002**, *13*, 40–46.
- [110] D. H. Bunfield, L. M. Davis, *Appl. Optics* **1998**, *37*, 2315–2326.
- [111] A. Hillisch, M. Lorenz, S. Diekmann, *Curr. Opin. Struct. Biol.* **2001**, *11*, 201–207.
- [112] D. G. Norman, R. J. Grainger, D. Uhrin, D. M. Lilley, *Biochemistry* **2000**, *39*, 6317–6324.
- [113] R. M. Clegg, A. I. Murchie, A. Zechel, D. M. Lilley, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 2994–2998.
- [114] P. M. Cowan, S. McGavin, *Nature* **1955**, *176*, 501–503.
- [115] P. R. Schimmel, P. J. Flory, *Proc. Natl. Acad. Sci. USA* **1967**, *58*, 52–59.
- [116] Y. W. Jia, D. S. Talaga, W. L. Lau, H. S. M. Lu, W. F. DeGrado, R. M. Hochstrasser, *Chem. Phys.* **1999**, *247*, 69–83.
- [117] D. S. Talaga, W. L. Lau, H. Roder, J. Tang, Y. Jia, W. F. DeGrado, R. M. Hochstrasser, *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 13021–13026.
- [118] D. Wassenberg, C. Welker, R. Jaenicke, *J. Mol. Biol.* **1999**, *289*, 187–193.
- [119] B. Schuler, W. Kremer, H. R. Kalbitzer, R. Jaenicke, *Biochemistry* **2002**, *41*, 11670–11680.
- [120] D. Perl, C. Welker, T. Schindler, K. Schröder, M. A. Marahiel, R. Jaenicke, F. X. Schmid, *Nat. Struct. Biol.* **1998**, *5*, 229–235.
- [121] J. B. Knight, A. Vishwanath, J. P. Brody, R. H. Austin, *Phys. Rev. Lett.* **1998**, *80*, 3863–3866.
- [122] E. Kauffmann, N. C. Darnton, R. H. Austin, C. Batt, K. Gerwert, *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6646–6649.
- [123] E. A. Lipman, B. Schuler, O. Bakajin, W. A. Eaton, *Science* **2003**, *301*, 1233–1235.
- [124] C. Magg, F. X. Schmid, *J. Mol. Biol.* **2004**, *335*, 1309–1323.
- [125] D. Shortle, *FASEB J.* **1996**, *10*, 27–34.
- [126] I. V. Gopich, A. Szabo, *J. Phys. Chem. B* **2003**, *107*, 5058–5063.

- [127] K. Chattopadhyay, E. L. Elson, C. Frieden, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 2385–2389.
- [128] I. V. Gopich, A. Szabo, *J. Chem. Phys.* **2005**, *122*, 1–18.
- [129] a) B. Schuler, E. A. Lipman, W. A. Eaton, *Nature* **2003**, *421*, 94; b) B. Schuler, E. A. Lipman, W. A. Eaton, unpublished results.
- [130] H. A. Kramers, *Physica* **1940**, *7*, 284–304.
- [131] P. Hänggi, P. Talkner, M. Borkovec, *Rev. Mod. Phys.* **1990**, *62*, 251–341.
- [132] N. D. Socci, J. N. Onuchic, P. G. Wolynes, *J. Chem. Phys.* **1996**, *104*, 5860–5868.
- [133] D. K. Klimov, D. Thirumalai, *Phys. Rev. Lett.* **1997**, *79*, 317–320.
- [134] L. J. Lapidus, P. J. Steinbach, W. A. Eaton, A. Szabo, J. Hofrichter, *J. Phys. Chem. B* **2002**, *106*, 11628–11640.
- [135] J. J. Portman, S. Takada, P. G. Wolynes, *J. Chem. Phys.* **2001**, *114*, 5082–5096.
- [136] B. Onoa, I. Tinoco, *Curr. Opin. Struct. Biol.* **2004**, *14*, 374–379.
- [137] E. V. Amirgoulova, J. Groll, C. D. Heyes, T. Ameringer, C. Rocker, M. Moller, G. U. Nienhaus, *ChemPhysChem* **2004**, *5*, 552–555.
- [138] E. Boukobza, A. Sonnenfeld, G. Haran, *J. Phys. Chem. B* **2001**, *105*, 12165–12170.
- [139] J. Groll, E. V. Amirgoulova, T. Ameringer, C. D. Heyes, C. Rocker, G. U. Nienhaus, M. Moller, *J. Am. Chem. Soc.* **2004**, *126*, 4234–4239.
- [140] E. Rhoades, E. Gussakovsky, G. Haran, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3197–3202.
- [141] E. Rhoades, M. Cohen, B. Schuler, G. Haran, *J. Am. Chem. Soc.* **2004**, *126*, 14686–14687.
- [142] B. Okumus, T. J. Wilson, D. M. J. Lilley, T. Ha, *Biophys. J.* **2004**, *87*, 2798–2806.
- [143] J. M. Fernandez, H. B. Li, *Science* **2004**, *303*, 1674–1678.
- [144] T. R. Sosnick, *Science* **2004**, *306*, 411.
- [145] J. M. Fernandez, H. B. Li, J. Brujic, *Science* **2004**, *306*, 411.
- [146] L. O. Tjernberg, A. Pramanik, S. Bjorling, P. Thyberg, J. Thyberg, C. Nordstedt, K. D. Berndt, L. Terenius, R. Rigler, *Chem. Biol.* **1999**, *6*, 53–62.
- [147] S. Kim, E. A. A. Nollen, K. Kitagawa, V. P. Bindokas, R. I. Morimoto, *Nat. Cell Biol.* **2002**, *4*, 826–831.
- [148] Y. Inoue, A. Kishimoto, J. Hirao, M. Yoshida, H. Taguchi, *J. Biol. Chem.* **2001**, *276*, 35227–35230.
- [149] F. U. Hartl, M. Hayer-Hartl, *Science* **2002**, *295*, 1852–1858.
- [150] S. Walter, J. Buchner, *Angew. Chem.* **2002**, *114*, 1142–1158; *Angew. Chem. Int. Ed.* **2002**, *41*, 1098–1113.
- [151] T. Ueno, H. Taguchi, H. Tadakuma, M. Yoshida, T. Funatsu, *Mol. Cell* **2004**, *14*, 423–434.
- [152] H. Taguchi, T. Ueno, H. Tadakuma, M. Yoshida, T. Funatsu, *Nat. Biotechnol.* **2001**, *19*, 861–865.
- [153] R. Yamasaki, M. Hoshino, T. Wazawa, Y. Ishii, T. Yanagida, Y. Kawata, T. Higurashi, K. Sakai, J. Nagai, Y. Goto, *J. Mol. Biol.* **1999**, *292*, 965–972.
- [154] J. Ichinose, Y. Sako, *Trac-Trend. Anal. Chem.* **2004**, *23*, 587–594.
- [155] A. N. Fedorov, T. O. Baldwin, *J. Biol. Chem.* **1997**, *272*, 32715–32718.
- [156] R. J. Ellis, *Trends Biochem. Sci.* **2001**, *26*, 597–604.
- [157] J. Wang, *J. Chem. Phys.* **2003**, *118*, 952–958.
- [158] V. B. P. Leite, J. N. Onuchic, G. Stell, J. Wang, *Biophys. J.* **2004**, *87*, 3633–3641.
- [159] A. M. Berezhkovskii, A. Szabo, G. H. Weiss, *J. Chem. Phys.* **1999**, *110*, 9145–9150.
- [160] I. V. Gopich, A. Szabo, *J. Chem. Phys.* **2003**, *118*, 454–455.
- [161] H. Yang, X. S. Xie, *J. Chem. Phys.* **2002**, *117*, 10965–10979.
- [162] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graphics* **1996**, *14*, 51–55.
- [163] W. Kremer, B. Schuler, S. Harrieder, M. Geyer, W. Gronwald, C. Welker, R. Jaenicke, H. R. Kalbitzer, *Eur. J. Biochem.* **2001**, *268*, 2527–2539.

---

Received: December 9, 2004

Revised: March 29, 2005