Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy

Armin Hoffmann*, Avinash Kane[†‡], Daniel Nettels*, David E. Hertzog[†], Peter Baumgärtel[§], Jan Lengefeld[§], Gerd Reichardt[¶], David A. Horsley[‖], Robert Seckler[§], Olgica Bakajin[†]**, and Benjamin Schuler*,**

*Biochemisches Institut, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland; [†]BioSecurity and Nanosciences Laboratory, Lawrence Livermore National Laboratory, Livermore, CA 94550; Departments of [‡]Electrical and Computer Engineering and [‖]Mechanical and Aeronautical Engineering, University of California, Davis, CA 95616; [§]Physikalische Biochemie, Universität Potsdam, 14476 Potsdam-Golm, Germany; and [¶]Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung, 12489 Berlin, Germany

Edited by William A. Eaton, National Institutes of Health, Bethesda, MD, and approved November 6, 2006 (received for review May 26, 2006)

We have used the combination of single-molecule Förster resonance energy transfer and kinetic synchrotron radiation circular dichroism experiments to probe the conformational ensemble of the collapsed unfolded state of the small cold shock protein Csp*Tm* under near-native conditions. This regime is physiologically most relevant but difficult to access experimentally, because the equilibrium signal in ensemble experiments is dominated by folded molecules. Here, we avoid this problem in two ways. One is the use of single-molecule Förster resonance energy transfer, which allows the separation of folded and unfolded subpopulations at equilibrium and provides information on long-range intramolecular distance distributions. From experiments with donor and acceptor chromophores placed at different positions within the chain, we find that the distance distributions in unfolded Csp*Tm* agree surprisingly well with a Gaussian chain not only at high concentrations of denaturant, where the polypeptide chain is expanded, but also at low denaturant concentrations, where the chain is collapsed. The second, complementary approach is synchrotron radiation circular dichroism spectroscopy of collapsed unfolded molecules transiently populated with a microfluidic device that enables rapid mixing. The results indicate a β-structure content of the collapsed unfolded state of $\approx$20% compared with the folded protein. This suggests that collapse can induce secondary structure in an unfolded state without interfering with long-range distance distributions characteristic of a random coil, which were previously found only for highly expanded unfolded proteins.

Gaussian chain | microfluidic mixing | protein folding | random coil | secondary structure

With the discovery of small proteins that fold in the absence of populated intermediates (1), our quantitative understanding of the elementary properties of protein folding reactions has made significant advances, including the structural characterization of transition states for folding (2) and the prediction of folding rates from native structure (3–5). One of the most severe limitations for the further development of these approaches is our ignorance about the energetic or structural properties of unfolded[††] states of proteins. Because of the structural heterogeneity and complexity of the ensembles of conformations populated by unfolded proteins, their experimental characterization has proven extremely difficult. Traditional methods, such as small-angle scattering techniques (6), provide only global physical properties, e.g., the radius of gyration. In some cases, more detailed structural information can be obtained from NMR (7–10), but these studies usually provide information about the denatured state only under nonnative conditions, typically in the presence of large concentrations of denaturant, or through severe destabilization of the native state induced by covalent modification or mutations. The most interesting and physiologically relevant situation, however, is that of an unfolded state of a stable protein under native conditions. Unfortunately, the great majority of molecules will then be present in their native conformation, thus overwhelming the signal from unfolded molecules.

We avoid this problem by using two complementary optical techniques: single-molecule fluorescence and kinetic synchrotron radiation circular dichroism (SRCD) spectroscopy. Single-molecule spectroscopy has the inherent ability to separate the signals from subpopulations in heterogeneous mixtures and equilibria, which makes it ideally suited to analyze protein folding reactions (11, 12). Specifically, by using single-molecule Förster resonance energy transfer (FRET), intramolecular distances of the unfolded state can be measured even in the presence of a majority of folded molecules (13, 14). Recently, the collapse of unfolded molecules of the small cold shock protein Csp*Tm* at close to native conditions was discovered with this approach (14). This collapsed unfolded form also can be populated kinetically (15, 16) and has by now been found for a range of small proteins (16–20). It is unclear, however, whether this collapse is a nonspecific random heteropolymer collapse (21) or whether it is accompanied by the formation of specific structure. By placing FRET dye pairs in various positions of the protein, we obtain information on distance distributions in different segments of the unfolded polypeptide chain and their dependence on denaturant concentration.

A versatile method to complement such distance constraints with information about the secondary structure content is circular dichroism (CD) spectroscopy. In this case, however, we have to transiently populate the unfolded state under native conditions by using rapid mixing experiments. For this purpose, we use SRCD spectroscopy with a specifically designed microfluidic continuous-flow mixing system with millisecond dead time. In this way, the collapsed unfolded state, which is populated on a microsecond time scale or faster (14, 16), can be studied

Author contributions: A.H., A.K., and D.N. contributed equally to this work; O.B. and B.S. designed research; A.H., A.K., D.E.H., P.B., J.L., O.B., and B.S. performed research; A.K., D.N., D.E.H., P.B., G.R., D.A.H., R.S., O.B., and B.S. contributed new reagents/analytic tools; A.H., A.K., D.N., and B.S. analyzed data; and A.H., O.B., and B.S. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS direct submission.

Abbreviations: FRET, Förster resonance energy transfer; CD, circular dichroism; SRCD, synchrotron radiation CD.

**To whom correspondence may be addressed. E-mail: bakajin1@llnl.gov or schuler@bioc.unizh.ch.

[††]We denote all nonnative conformations as ''unfolded'' to stress that for true two-state proteins, such as Csp*Tm*, there are no thermodynamically or kinetically distinguishable denatured states.

This article contains supporting information online at www.pnas.org/cgi/content/full/0604353104/DC1.

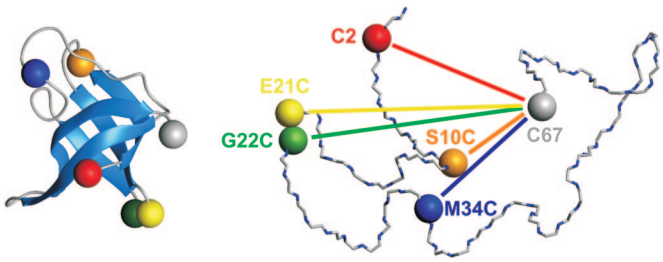

Fig. 1. Schematic of folded (*Left*) and unfolded (*Right*) Csp*Tm* with the sites for dye attachment for FRET indicated by colored spheres. For every variant investigated, one dye was reacted with Cys at position 67, and a second dye was reacted with a Cys at one of the other positions shown.

spectroscopically without interference from the signal of folded molecules, which form on a slower time scale. Importantly, SRCD gives us access to the far-UV wavelength range, where β-structure can be well distinguished from random coil (22, 23) but which is inaccessible with commercial stopped-flow instruments.

Results

Collapse from FRET Efficiency Histograms. For the single-molecule FRET experiments, donor (Alexa Fluor 488) and acceptor (Alexa Fluor 594) dyes were placed specifically at different solvent-exposed positions of Csp*Tm* via pairs of Cys residues introduced by site-directed mutagenesis (Fig. 1). Upon unfolding of the protein, the average distance between the chromophores increases. As a consequence, the rate of energy transfer between them decreases, resulting in a reduced transfer efficiency $E = n_A/(n_A + n_D)$, where n_A and n_D are, respectively, the number of acceptor and donor photons emitted by the molecule [including corrections (24), see *Materials and Methods*]. In confocal single-molecule experiments, transfer efficiencies are determined from photon bursts originating from individual molecules freely diffusing through the focal spot of the laser beam (25). A histogram from a large number of such events shows distinct maxima corresponding to the subpopulations present in the sample (Fig. 2). The peak at high E corresponds to folded molecules, and the peak at intermediate E corresponds to unfolded molecules. The

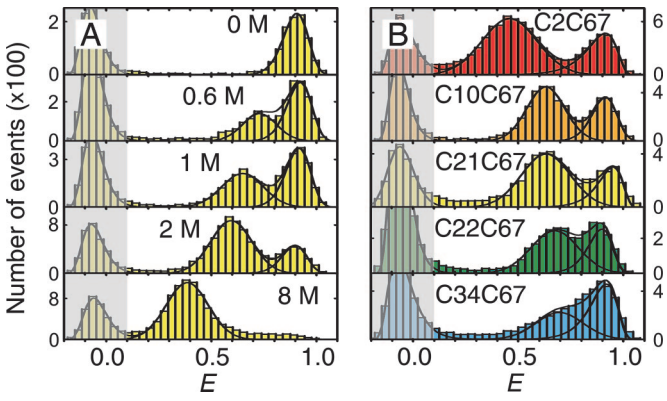

Fig. 2. Energy transfer efficiency (E) histograms from single-molecule FRET measurements. (*A*) Examples from a GdmCl titration of variant C21C67, illustrating unfolded state collapse. See SI Movie 1 for a complete data set of C2C67. (*B*) Histograms of all variants at 1.5 M GdmCl. The peak at $E \approx 0.9$ corresponds to folded molecules, and the peak at intermediate E corresponds to unfolded molecules. The peak at $E \approx 0$ (shaded) originates from molecules with an inactive acceptor (26). To determine mean transfer efficiencies, the unfolded peak was fit to a normal distribution, and the other two peaks were fit to log normal functions (black lines) (14). The colors used correspond to those used in Fig. 1.

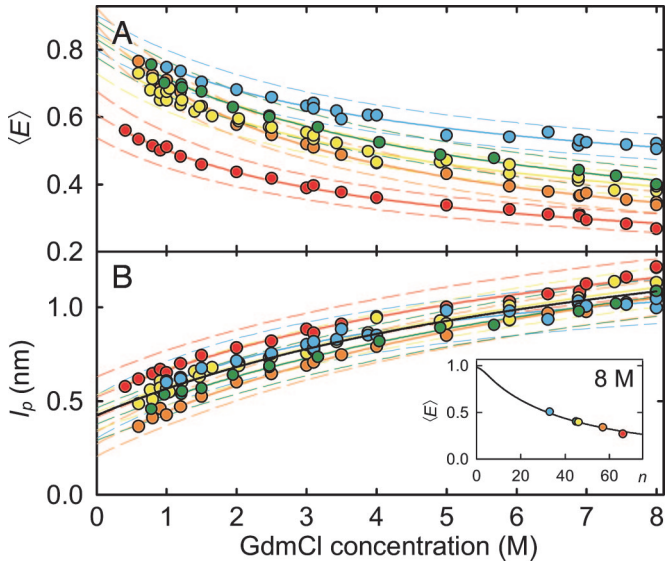

Fig. 3. Denaturant dependence of the mean transfer efficiencies $\langle E \rangle$ and persistence lengths l_p. (*A*) $\langle E \rangle$ of the unfolded state of all variants as a function of GdmCl concentration. (*B*) l_p calculated from $\langle E \rangle$ using Eqs. **1**–**4**. (*Inset*) Shows the measured values of $\langle E \rangle$ for all variants at 8 M GdmCl, and as a black line $\langle E \rangle$ for a Gaussian chain calculated from Eqs. **1**–**4** for $l_p = 1.1$ nm (the mean value of all variants) as a function of sequence separation n (number of peptide bonds excluding linkers).‡‡ The solid lines in the GdmCl titrations are fits to the empirical equation $y = y_0 [1 + \Delta y K x l (1 + Kx)]$ used for interpolation; the solid black line is a fit to all data. Estimated error ranges for $\langle E \rangle$ and l_p are indicated by dashed lines (see *SI Materials and Methods*). The colors used correspond to those used in Fig. 1.

additional peak near $E = 0$ is thought to be caused by molecules lacking an active acceptor chromophore (24, 26) but does not interfere with our analysis.

With increasing concentration of the denaturant guanidinium chloride (GdmCl), we observe a change in the relative amplitudes of the signal from native and unfolded molecules, corresponding to the expected redistribution of the two populations (Fig. 2*A*). However, whereas the mean transfer efficiency $\langle E \rangle$ of the folded state is invariant, the peak from unfolded molecules is continuously shifting to higher transfer efficiencies with decreasing GdmCl concentration. By comparison with stiff polyproline peptides, it has previously been shown that this increase in $\langle E \rangle$ corresponds to a collapse of the unfolded state in response to the altered solvent conditions (14). To address the question of whether this collapse is a global process that is evenly distributed across the polypeptide chain or whether it is due to compaction of a specific part of the molecule, we measured single-molecule transfer efficiency histograms of the labeled Csp*Tm* variants at a wide range of denaturant concentrations. At a given GdmCl concentration, $\langle E \rangle$ in the unfolded state shows an overall increase with decreasing sequence separation, as expected (Fig. 2*B*) . Moreover, all variants exhibit the characteristic continuous collapse at low concentrations of denaturant, resulting in an increase in $\langle E \rangle$ [Fig. 3*A* and supporting information (SI) Movie 1].

Distance Distributions from Mean Transfer Efficiencies. To analyze the mean transfer efficiencies in terms of distance distributions in the unfolded state, we use the Gaussian chain model, the simplest realistic model for describing large-scale properties of macromolecules (21), such as highly unfolded peptides and

‡‡Note that our analysis does not take into account the effect of excluded volume on the length scaling of the end-to-end distance. Such differences become discernible only if unfolded proteins are investigated over a much wider range of chain lengths than here (30).

proteins (6, 27–31). $\langle E \rangle$ of the unfolded state at a given GdmCl concentration can be expressed in terms of the end-to-end distance probability distribution function of a Gaussian chain $P(r)$ and the distance dependence of the transfer efficiency $E(r)$ according to[§§]

$$\langle E \rangle = \int_a^{l_c} E(r)P(r)dr \Big/ \int_a^{l_c} P(r)dr \qquad [1]$$

with

$$E(r) = \frac{1}{1 + (r/R_0)^6} \qquad [2]$$

and

$$P(r) = 4\pi r^2 \left(\frac{3}{2\pi\langle r^2 \rangle}\right)^{3/2} \exp\left(-\frac{3r^2}{2\langle r^2 \rangle}\right), \qquad [3]$$

where r is the end-to-end distance, a is the distance of closest approach of the chain ends, l_c is the contour length of the labeled polypeptide segment (with dyes and linkers, see *Materials and Methods*), and R_0 is the Förster radius of the dye pair at the GdmCl concentration of the solution (R_0 = 5.4 nm at 0 M GdmCl; see *SI Materials and Methods* for details). The mean squared end-to-end distance $\langle r^2 \rangle$ of a Gaussian chain can be expressed as

$$\langle r^2 \rangle = 2l_p l_c = 2l_p nl, \qquad [4]$$

where n is the number of peptide bonds between the chromophores, l is the projection of the distance between two consecutive α-carbon atoms on the axis of the fully extended chain (0.38 nm), and l_p (the only free parameter in the model) is a measure for the effective chain stiffness, which in our limit of $l_c \gg l_p$ is equivalent to the persistence length commonly used in the worm-like chain model (34). Fig. 3*B Inset* shows the measured $\langle E \rangle$ of our Csp*Tm* variants at 8 M GdmCl. The continuous line shows $\langle E \rangle$ as a function of n calculated according to Eqs. **1**–**4** with l_p = 1.1 nm. All variants can be described with the same value of l_p, indicative of isotropic distance distributions within unfolded Csp*Tm*, and in agreement with previous results on other proteins unfolded at high denaturant concentrations (27–30).

Fig. 3*B* shows l_p for every variant as a function of GdmCl concentration calculated from $\langle E \rangle$ (Fig. 3*A*) using Eqs. **1**–**4**. The concomitant renormalization for chain length allows a direct comparison of the variants with different sequence separation of donor and acceptor. The values of l_p averaged over all variants range from 1.1 ± 0.1 nm at 8 M GdmCl, to 0.39 ± 0.07 nm extrapolated to 0 M GdmCl (Fig. 3*B*), close the values found in atomic force microscopy experiments (0.33 to 0.42 nm; summarized in ref. 31) or unstructured peptides [0.64 nm (32)] at 0 M GdmCl. All variants show a similar denaturant dependence of l_p, suggesting isotropic chain collapse. From these data alone we cannot exclude that the more pronounced differences at low denaturant concentrations, especially for variant C10C67, may indicate slight deviations from completely isotropic collapse, but several points suggest otherwise. First, it seems structurally implausible that C2C67 and C21C67, which flank the slightly deviating C10C67, exhibit distance distributions in accord with the remaining variants, whereas a large difference occurs between C2C67 and C10C67. Second, other effects, such as changes in the photophysical properties of the chromophores upon collapse, may play a role; e.g., the two lysine residues directly neighboring Cys-10 may act as fluorescence quenchers (35) whose electrostatic interaction with the negatively charged fluorophores may be shielded at high GdmCl concentrations. In view of the error ranges we estimate for this type of measurement (Fig. 3), we conclude that the data are in agreement with isotropic collapse and do not justify an interpretation beyond this simple model.

[§§]For this analysis, we assume intramolecular distance fluctuations to be slow relative to the fluorescence lifetime of the donor chromophore, as expected for long polypeptide chains (32, 33).

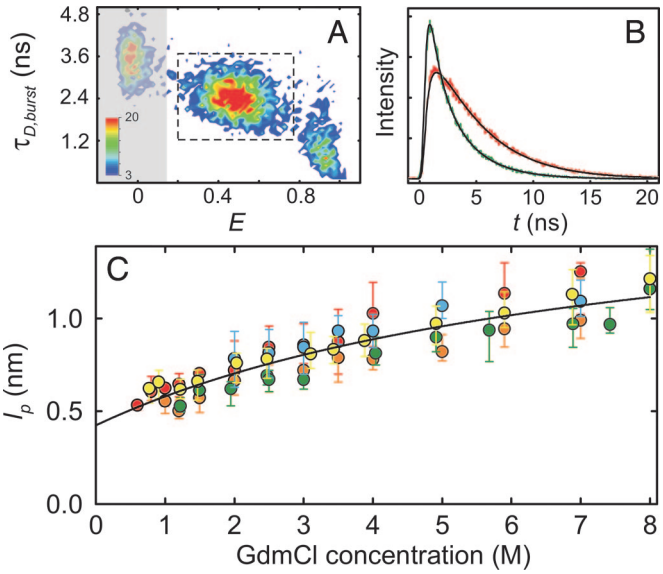

Fig. 4. Fluorescence lifetime distribution analysis. (*A*) All events corresponding to unfolded molecules were selected (dashed box) from a two-dimensional histogram (shown is C2C67 at 1.5 M GdmCl) of the number of bursts (color scale) with transfer efficiency E and donor fluorescence lifetime $\tau_{D,burst}$. (*B*) The photons were combined to generate time-correlated single-photon counting histograms for donor (green) and acceptor (red) (see SI Movie 2 for a complete data set of C2C67). (*C*) The rms end-to-end distance was determined and converted to the apparent persistence length l_p (Eq. **4**) by using a global fit assuming the distance distribution of a Gaussian chain (black lines). Error bars indicate the uncertainty in the fits. The solid black line in is a fit to all data as described in Fig. 3. The colors used correspond to those used in Fig. 1.

Subpopulation-Selective Fluorescence Lifetime Distribution Analysis. Mean transfer efficiencies obtained from transfer efficiency histograms do not provide direct information about the shape of the distance distribution, because the dynamics of inter-dye distance fluctuations are fast compared with the millisecond observation time per burst (14, 36). Fluorescence intensity decays, on the other hand, do report directly on the shape of the distance distribution (37, 38), because distance fluctuations for long polypeptides are slow relative to the fluorescence lifetimes of the dyes (32, 33). Conventional approaches to the analysis of distance distributions from fluorescence lifetimes suffer from interference with signal from native molecules; however, in combination with the separation of subpopulations by virtue of single-molecule detection, we can selectively analyze distance distributions in the unfolded state while remaining unaffected by native signal (17). For every individual fluorescence burst, the mean donor fluorescence lifetime is estimated (39) in addition to E. From the resulting two-dimensional histograms (Fig. 4*A*), we select the bursts corresponding to the unfolded state and, by combining the photons from all unfolded molecules, obtain high-resolution fluorescence decays (Fig. 4*B*). Donor and acceptor decays are then

analyzed in a global fit[¶¶] that assumes a distribution of transfer rates resulting from the distribution of donor–acceptor distances (see *SI Materials and Methods*), analogous to the established procedures for ensemble data (17, 38, 40). Specifically, we use $P(r)$ of a Gaussian chain (Eq. **3**) with $\langle r^2 \rangle$ as the only free parameter.

The time-correlated single-photon counting histograms from unfolded molecules are fit well, if we assume the distance distribution of a Gaussian chain (Fig. 4*B*; see also SI Movie 2), and again we observe a continuous collapse of the chain for all variants, with the expected increase in $\langle r^2 \rangle$ with n (data not shown). To normalize the data for n and to facilitate direct comparison of lifetime and intensity results, we calculate l_p from $\langle r^2 \rangle$ (Eq. **4**) (Fig. 4*C*). The greater demands on the signal-to-noise ratio for this type of analysis reduces the number of suitable data sets, especially at low GdmCl concentrations, where the overlap between native and unfolded subpopulation increases, but nevertheless we again find a similar GdmCl-dependence of l_p for all variants. The good overall agreement of fluorescence intensity (Fig. 3*B*) and lifetime (Fig. 4*C*) results and the lack of consistent deviations from the averages in Figs. 3*B* and 4*C* provide additional evidence that the remaining differences between variants seen in either analysis are within the errors of the methods. The combined results of single-molecule intensity and lifetime data therefore suggest that the chain dimensions of Csp*Tm* are isotropic and Gaussian chain-like on the length scale we are probing, even for strongly collapsed unfolded molecules. But does this random coil behavior exclude the presence of secondary structure in the collapsed unfolded state?

Secondary Structure Content from SRCD. An ideal method to quantify the secondary structure content of proteins is CD spectroscopy. Collapse of Csp has been shown to occur on a microsecond timescale or faster (12, 14, 16), whereas the folding rate at low GdmCl concentrations is in the tens of milliseconds range (41). In a kinetic mixing experiment with millisecond time resolution, the collapsed unfolded state is therefore formed during the dead time, and only the actual folding process is resolved. The dead time signal amplitude reports on the change in structure during collapse. However, although the formation of α-helical structure is readily observed with stopped-flow CD spectroscopy (typically at 222 nm), formation of β-sheet structure is often not accessible because of the small difference in CD signal between a random coil and β-structure as well as the strong influence of aromatic amino acids at wavelengths of >220 nm (42), the range available in conventional stopped-flow CD. Accordingly, previous attempts to measure secondary structure formation upon collapse of Csp*Tm* with CD spectroscopy have failed (B.S., unpublished results). To solve this problem, we have started to establish a methodology for which SRCD is used with microfabricated rapid mixing devices. SRCD has the advantage of high photon flux in the far UV below 220 nm, where the CD signal differences between random coil and β-structure become very pronounced (43), and the Gaussian beam shape allows focusing into microstructures. Microfluidic mixers (44) permit the implementation of sophisticated mixing strategies, making them the fastest method available for rapid-dilution experiments (45). They can be fabricated in fused silica with path lengths in the micrometer range, thus maximizing transmission in the far UV range. Because of the small feature size, sample consumption is drastically reduced compared with conventional continuous-flow mixing experiments (46), some of which have already led to substantial improvements both in time resolution and accessible wavelength range in kinetic CD experiments (47).

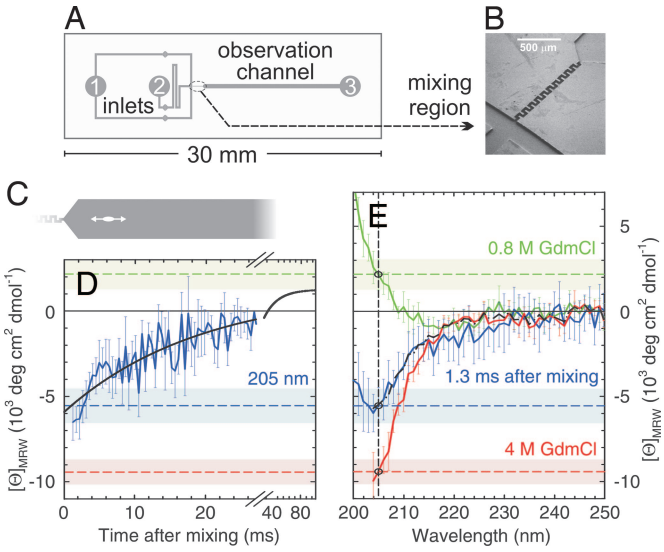

Fig. 5. Secondary structure content of collapsed unfolded Csp*Tm* from SRCD. (*A*) Channel pattern of the microfluidic mixing device. To initiate refolding, unfolded protein injected into inlet 2 is diluted with buffer injected into inlet 1. (*B*) Rapid mixing occurs in the serpentine-shaped channel shown as a scanning electron micrograph. (*C*) The synchrotron radiation beam (white ellipse) is positioned in the observation channel. (*D*) Refolding kinetics of Csp*Tm* at 0.8 M GdmCl measured at 205 nm (blue line; error bars give one standard deviation calculated from eight measurements) and a single-exponential fit to the data (solid black line). (*E*) CD spectrum taken 1.3 ms after mixing (solid blue line), compared with equilibrium spectra under native (solid green line; 0.8 M GdmCl) and unfolding conditions (red line; 4 M GdmCl). The corresponding ellipticities at 205 nm are indicated in the spectra as dashed blue, green, and red lines, respectively. The shaded light blue, green, and red bands indicate one standard deviation from the equilibrium ellipticities of folded and unfolded Csp*Tm*, respectively, at 205 nm. The black dashed–dotted curve is a linear combination of the spectra at 0.8 M and 4 M GdmCl used to estimate the secondary structure content of collapsed unfolded Csp*Tm*.

For refolding, Csp*Tm* in 4 M GdmCl was diluted to a final concentration of 0.8 M GdmCl (Fig. 5). The actual mixing region (Fig. 5*B*) consists of a narrow serpentine-shaped channel that performs mixing in the laminar flow regime by virtue of diffusion and chaotic advection (48). CD kinetics were measured by scanning the synchrotron radiation beam along the broader observation channel (Fig. 5*C*). Different positions in the channel thus correspond to different times after mixing, with a dead time of 1.3 ms. A kinetic trace taken at 205 nm (Fig. 5*D*), fit to a single exponential decay, resulted in a time constant of 19 ± 4 ms, in good agreement with the folding rates measured with stopped-flow fluorescence under identical conditions (41). However, the CD signal of -6×10^3 deg·cm^2·dmol^{-1} at 1.3 ms does not coincide with the CD signal of Csp*Tm* unfolded in 4 M GdmCl (Fig. 5 *D* and *E*, dashed red line), indicating the formation of secondary structure before the actual folding reaction.

For spectral information on collapsed unfolded Csp*Tm*, wavelength scans were taken 1.3 ms after mixing (Fig. 5*E*, solid blue line), when the fraction of native molecules calculated from the folding time is only 7%. Steady-state spectra of Csp*Tm* unfolded in 4 M GdmCl (Fig. 5*E*, solid red line), and native Csp*Tm* in 0.8 M GdmCl (Fig. 5*E*, solid green line) were recorded in the same mixing device and corrected for baselines taken without protein under otherwise identical conditions. These spectra illustrate the advantage gained by extending the wavelength range to <220 nm. Fitting the CD spectrum of collapsed unfolded Csp*Tm* to a linear combination of the spectra in 4 M and 0.8 M GdmCl (Fig. 5*E*, black dash–dotted

[¶¶]Note that a global analysis is substantially more robust than the individual fits and mitigates the common problems of quantitatively analyzing fluorescence decays deviating from single exponential behavior, especially for the short donor lifetime components, which result in a rise term in the acceptor intensity.

curve) results in a signal change of ≈27% relative to the native structure (excluding the contribution from the 7% folded molecules). Relating this amplitude quantitatively to β-structure content is complicated by the contribution of aromatic amino acids to the native state signal, but from CD reference spectra for β-sheet (43) and the β-structure content from the NMR structure of Csp*Tm* (49), we estimate for the collapsed unfolded protein at 0.8 M GdmCl a β-structure content of ≈20% relative to the native protein (assuming no contribution from aromatic amino acids to the CD of collapsed unfolded Csp*Tm*). A very small population of additional α-helical structure cannot be excluded, but from the signal at 222 nm, the α-helix contribution is not more than approximately -1×10^3 deg·cm^2·dmol^{-1}, corresponding to less than ≈3% α-helix (43).

Discussion

Over the past 40 years, many proteins unfolded at high concentrations of denaturant have been shown to obey Gaussian chain statistics on a global scale (6, 27–30, 50, 51). Csp*Tm* is no exception in this respect. Its radius of gyration in 6 M GdmCl calculated from all labeled variants [using Eq. **4** and $R_g^2 = \langle r^2 \rangle/6$ (21)] is 2.9 (±0.1) nm excluding dye linkers, in agreement with the 2.4 (+0.4/−0.2) nm calculated from the scaling law given by Kohn *et al.* (30, 51). But how does the collapse of the unfolded state at low GdmCl concentrations affect the intramolecular distance distributions? Surprisingly, even under near-native conditions, the agreement with Gaussian chain behavior is good, and l_p is similar for all intramolecular distance pairs (Figs. 3*B* and 4*B*). At 1 M GdmCl, for instance, the resulting radii of gyration equal 2.2 (±0.2) nm. The small variance suggests that the distance distributions within the unfolded protein are rather isotropic. We thus have no evidence for native state topology in collapsed unfolded Csp*Tm*, in contrast to NMR experiments on staphylococcal nuclease (8) and eglin (52) at high concentrations of urea and in contrast to suggestions from simulations (53). Similarly, there is no obvious relation to the transition state structure (54). Our observation is in contrast to recent measurements on chymotrypsin inhibitor 2 and acyl-CoA-binding protein, where indications were found for a substantial deviation of the collapsed denatured state from Gaussian chain behavior (17) possibly involving folding intermediates (55). In summary, we conclude from the single-molecule fluorescence data that unfolded Csp*Tm* is close to a random coil in terms of polymer physics, even under near-physiological conditions.

It may therefore come as a surprise that the collapsed unfolded state contains a significant amount of β-structure, as was observed in our kinetic SRCD experiments. How can we reconcile these two observations? Clearly, global random coil behavior does not exclude the presence of short structured segments (30, 56–58), even more so if these are only populated transiently. This argument has been used to resolve the seemingly conflicting views of residual structure observed in proteins under highly denaturing conditions on the one hand and the successful description of global properties of unfolded polypeptides with the random coil model on the other (59, 60). Our results suggest that there are cases for which we must extend this notion of a random coil with residual structure even to the collapsed unfolded state, populated under conditions that have so far evaded confrontation with the "reconciliation problem" (6). Compaction of the chain would be expected to contribute to the formation of local structure, because the increase in excluded volume effects will introduce more steric interference with nonnearest-neighbor residues (29, 60). As a result, the backbone will be forced even more into the core regions of the Ramachandran map, corresponding to extended structures that avoid such steric conflicts. The largest one of these regions, and thus entropically the most favorable one, is the extended structure of β-strands. The particular preference of Csp*Tm* for this conformation is reflected by its extremely low propensity for the formation of α-helices [<1% helical content predicted with AGADIR (61)].

Currently, we have no direct evidence for the detailed conformation and average length of β-strand segments populated in collapsed unfolded Csp*Tm*. Because Gaussian-distributed intramolecular distances can only be observed if the segment length is considerably less than the contour length (21), the stretches of β-conformation must be short relative to the sequence separation of our dye pairs. Details about more local distance distributions could be addressed by using FRET pairs with much smaller Förster radii, which are currently inaccessible to single-molecule spectroscopy. In a very recent study, the dead time amplitudes of stopped-flow ensemble FRET experiments were used to probe the first β-hairpin of a closely related cold shock protein for local conformational preferences in the collapsed unfolded state (62). Magg *et al.* (62) observed a stretching of the second β-strand upon collapse, suggesting that the CD signal observed in our experiments could be due to strands of lengths approaching those in the native state. The question of segment length could possibly also be addressed with new methods for the analysis of CD spectra in terms of the number and size of structured segments in proteins (63, 64), but the current quality of our CD spectra for collapsed unfolded Csp*Tm* does not yet warrant such detailed deconvolution. The further development of kinetic SRCD and its combination with single-molecule fluorescence will be an important complementation of NMR methods in clarifying these structural details and the question of whether the behavior of Csp*Tm* is an exception or possibly a more general characteristic of such small all-β proteins.

Materials and Methods

Synthesis and Labeling of Csp*Tm* Variants. Cysteine residues were introduced by site-directed mutagenesis to provide functional groups for the specific attachment of the dyes essentially as described previously (14). Expression and purification of Cys variants and wild-type protein were performed as described by Kremer *et al.* (49) (for details, see *SI Materials and Methods*).

Confocal Fluorescence Spectroscopy. Observations of single-molecule fluorescence were made with a MicroTime 200 confocal microscope (PicoQuant, Berlin, Germany) equipped with a 470-nm pulsed diode laser (LDH 470) operated at 40 MHz (average power, 130 μW) and an Olympus (Tokyo, Japan) UplanApo ×60, 1.20-W objective. Sample fluorescence was separated into donor and acceptor components with a dichroic mirror (Chroma 585DCXR) and two final filters (Chroma HQ525/50 and Omega 600ALP). Each component was focused onto an avalanche photodiode (SPCM-AQR-15; PerkinElmer Optoelectronics, Fremont, CA), and the arrival time of every detected photon was recorded relative to the exciting laser pulse with a time resolution of 38 ps. Samples of labeled protein were diluted to a concentration of ≈20 pM in 50 mM sodium phosphate buffer at the appropriate GdmCl (Pierce, Rockford, IL) concentration and individually adjusted to pH 7. Tween 20 (0.001%; Pierce) was added to prevent surface adhesion of the protein (14). Data were taken for 30–60 min.

Single-Molecule Data Reduction and Analysis. Successive photons detected in either channel separated by <100 μs were combined into one burst. A burst was retained as a significant event if the total number of counts exceeded 50. Identified bursts were corrected for background, differences in quantum yields, the different collection efficiencies of the detection channels, cross-talk, and direct acceptor excitation with the matrix approach (24) (*SI Materials and Methods*). For determining l_p, the length of dyes and linkers were assumed to be equivalent to an additional 9 aa total, comparable with previous estimates (14, 17, 51, 65). Fluorescence lifetime distribution analysis was performed as described by Laurence *et al.* (17), except that the time-correlated single-photon counting histograms obtained from the unfolded state subpopulations were analyzed in terms of donor–acceptor

distance distributions of a Gaussian chain (38, 40) (for details, see *SI Materials and Methods*).

Microfluidic Mixing Devices and Synchrotron Radiation Circular Dichroism. Mixers were fabricated by deep reactive ion etching of fused silica substrates (HPFS Standard Grade, Corning code 7980; Corning, Corning, NY) to a depth of 14.5 μm. Mixers were sealed by direct fusion wafer bonding to another fused silica substrate. A serpentine-shaped channel after the T region joining the inlet channels performs mixing in the laminar flow regime by diffusion and chaotic advection (48). Dean vortices in the transverse plane and corner vortices in the longitudinal plane accomplish mixing by stretching and folding the fluid streamlines. The widening of the channel after mixing slows down the flow, resulting in an accessible time window of $\approx$27 ms before the solution reaches the exit port (labeled "3" in Fig. 5*A*).

Microfluidic devices were mounted in the SRCD sample chamber (see *SI Materials and Methods*) via a custom-designed holder with connections to two syringe pumps (PHD22/2000; Harvard Apparatus, Holliston, MA). Motorized translation stages (M-111.1; PI, Karlsruhe, Germany) allowed reproducible positioning of the mixer relative to the beam. For refolding, a solution of 6.7 mM unfolded Csp*Tm* in 4 M GdmCl/50 mM sodium phosphate buffer (pH 7.0) was injected through inlet 2 (Fig. 5) at a flow rate of 30 μl/min and mixed with buffer solution without GdmCl injected through inlet 1 at a flow rate of 120 μl/min. Complete mixing was assessed via absorbance and CD scans orthogonal to the flow direction across the observation channel. Flow rates $\geq$150 μl/min [corresponding to flow velocities of 6.9 m/s in the mixing channel (25 μm wide) and 0.43 m/s in the observation channel (400 μm wide)] were found to result in uniform concentrations after the mixing region. The dead time calculated from these flow rates and the position of the synchrotron beam is 1.3 ms.

We thank Jana Kramer for excellent technical assistance. This work has been supported by the Human Frontier Science Program (A.H., D.E.H., O.B., and B.S.), the Schweizerische Nationalfonds (B.S.), and the Deutsche Forschungsgemeinschaft (J.L., B.S., and R.S.). O.B., A.K., and D.E.H. performed work under the auspices of the U.S. Department of Energy at University of California Lawrence Livermore National Laboratory under Contract W-7405-Eng-48, with funding from the Laboratory Research and Development Program and additional support by the National Science Foundation Center for Biophotonics Science and Technology through Cooperative Agreement PHY 0120999 (managed by the University of California, Davis). A.K. also was supported by the Student Employee Graduate Research Fellowship Program at Lawrence Livermore National Laboratory. A.K. and D.E.H.'s travel was supported by the International Institute for Complex Adaptive Matter. R.S., P.B., and the SRCD station at Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung were supported by German Federal Ministry of Education and Research Contract 05 KS4IP1/2.

1. Jackson SE (1998) *Fold Des* 3:R81–R91.
2. Fersht AR (1998) *Structure and Mechanism in Protein Sci* (Freeman, New York).
3. Galzitskaya OV, Finkelstein AV (1999) *Proc Natl Acad Sci USA* 96:11299–11304.
4. Baker D (2000) *Nature* 405:39–42.
5. Muñoz V, Eaton WA (1999) *Proc Natl Acad Sci USA* 96:11311–11316.
6. Millett IS, Doniach S, Plaxco KW (2002) *Adv Protein Chem* 62:241–262.
7. Shortle DR (1996) *Curr Opin Struct Biol* 6:24–30.
8. Shortle D, Ackerman MS (2001) *Science* 293:487–489.
9. Dyson HJ, Wright PE (2004) *Chem Rev* 104:3607–3622.
10. Religa TL, Markson JS, Mayor U, Freund SM, Fersht AR (2005) *Nature* 437:1053–1056.
11. Jia YW, Talaga DS, Lau WL, Lu HSM, DeGrado WF, Hochstrasser RM (1999) *Chem Phys* 247:69–83.
12. Schuler B (2005) *ChemPhysChem* 6:1206–1220.
13. Deniz AA, Laurence TA, Beligere GS, Dahan M, Martin AB, Chemla DS, Dawson PE, Schultz PG, Weiss S (2000) *Proc Natl Acad Sci USA* 97:5179–5184.
14. Schuler B, Lipman EA, Eaton WA (2002) *Nature* 419:743–747.
15. Lipman EA, Schuler B, Bakajin O, Eaton WA (2003) *Science* 301:1233–1235.
16. Magg C, Schmid FX (2004) *J Mol Biol* 335:1309–1323.
17. Laurence TA, Kong XX, Jager M, Weiss S (2005) *Proc Natl Acad Sci USA* 102:17348–17353.
18. Kuzmenkina EV, Heyes CD, Nienhaus GU (2006) *J Mol Biol* 357:313–324.
19. Sherman E, Haran G (2006) *Proc Natl Acad Sci USA* 103:11539–11543.
20. Qi PX, Sosnick TR, Englander SW (1998) *Nat Struct Biol* 5:882–884.
21. Grosberg AY, Khokhlov AR (1994) *Statistical Physics of Macromolecules* (Am Inst Phys, Woodbury, NY).
22. Sutherland JC (1996) in *Circular Dichroism and the Conformational Analysis of Biomolecules*, ed Fasman GD (Plenum, New York), pp 599–633.
23. Wallace BA, Janes RW (2001) *Curr Opin Chem Biol* 5:567–571.
24. Schuler B (2006) in *Protein Folding Protocols*, eds Bai Y, Nussinov R (Humana, Totowa, NJ), Vol 366.
25. Deniz AA, Laurence TA, Dahan M, Chemla DS, Schultz PG, Weiss S (2001) *Annu Rev Phys Chem* 52:233–253.
26. Schuler B, Lipman EA, Steinbach PJ, Kumke M, Eaton WA (2005) *Proc Natl Acad Sci USA* 102:2754–2759.
27. Tanford C (1968) *Adv Protein Chem* 23:121–282.
28. Damaschun G, Damaschun H, Gast K, Zirwer D (1998) *Biochemistry (Moscow)* 63:259–275.
29. Zhou HX (2002) *J Phys Chem B* 106:5769–5775.
30. Kohn JE, Millett IS, Jacob J, Zagrovic B, Dillon TM, Cingel N, Dothager RS, Seifert S, Thiyagarajan P, Sosnick TR, *et al*. (2004) *Proc Natl Acad Sci USA* 101:12491–12496.
31. Zhou HX (2004) *Biochemistry* 43:2141–2154.
32. Lapidus LJ, Steinbach PJ, Eaton WA, Szabo A, Hofrichter J (2002) *J Phys Chem B* 106:11628–11640.
33. Möglich A, Joder K, Kiefhaber T (2006) *Proc Natl Acad Sci USA* 103:12394–12399.
34. Doi M, Edwards SF (1988) *The Theory of Polymer Dynamics* (Oxford Univ Press, New York).
35. Ratner V, Amir D, Kahana E, Haas E (2005) *J Mol Biol* 352:683–699.
36. Gopich IV, Szabo A (2003) *J Phys Chem B* 107:5058–5063.
37. Haas E, Wilchek M, Katchalskikatzir E, Steinberg IZ (1975) *Proc Natl Acad Sci USA* 72:1807–1811.
38. Van Der Meer BW, Coker, G, III, Chen SYS (1994) *Resonance Energy Transfer: Theory and Data* (VCH, New York).
39. Eggeling C, Berger S, Brand L, Fries JR, Schaffer J, Volkmer A, Seidel CA (2001) *J Biotechnol* 86:163–180.
40. Beechem JM, Haas E (1989) *Biophys J* 55:1225–1236.
41. Perl D, Welker C, Schindler T, Schröder K, Marahiel MA, Jaenicke R, Schmid FX (1998) *Nat Struct Biol* 5:229–235.
42. Woody RW, Dunker AK (1996) in *Circular Dichroism and the Conformational Analysis of Biomolecules*, ed Fasman GD (Plenum, New York), pp 109–158.
43. Greenfield NJ (2004) *Methods Enzymol* 383:282–317.
44. Knight JB, Vishwanath A, Brody JP, Austin RH (1998) *Phys Rev Lett* 80:3863–3866.
45. Hertzog DE, Michalet X, Jager M, Kong XX, Santiago JG, Weiss S, Bakajin O (2004) *Anal Chem* 76:7169–7178.
46. Shastry MG, Luck SD, Roder H (1998) *Biophys J* 74:2714–2721.
47. Kimura T, Uzawa T, Ishimori K, Morishima I, Takahashi S, Konno T, Akiyama S, Fujisawa T (2005) *Proc Natl Acad Sci USA* 102:2748–2753.
48. Chamarthy P, Wereley ST (2004) *2004 Proceedings of the ASME International Mechanical Engineering Congress and Exposition*, 3 Vols (ASME Press, New York), CD-ROM, art no 61902.
49. Kremer W, Schuler B, Harrieder S, Geyer M, Gronwald W, Welker C, Jaenicke R, Kalbitzer HR (2001) *Eur J Biochem* 268:2527–2539.
50. Wilkins DK, Grimshaw SB, Receveur V, Dobson CM, Jones JA, Smith LJ (1999) *Biochemistry* 38:16424–16431.
51. McCarney ER, Werner JH, Bernstein SL, Ruczinski I, Makarov DE, Goodwin PM, Plaxco KW (2005) *J Mol Biol* 352:672–682.
52. Ohnishi S, Lee AL, Edgell MH, Shortle D (2004) *Biochemistry* 43:4064–4070.
53. Zagrovic B, Snow CD, Khaliq S, Shirts MR, Pande VS (2002) *J Mol Biol* 323:153–164.
54. Perl D, Holtermann G, Schmid FX (2001) *Biochemistry* 40:15501–15511.
55. Teilum K, Poulsen FM, Akke M (2006) *Proc Natl Acad Sci USA* 103:6877–6882.
56. Fitzkee NC, Rose GD (2004) *Proc Natl Acad Sci USA* 101:12497–12502.
57. Jha AK, Colubri A, Freed KF, Sosnick TR (2005) *Proc Natl Acad Sci USA* 102:13099–13104.
58. Ding F, Jha RK, Dokholyan NV (2005) *Structure (London)* 13:1047–1054.
59. McCarney ER, Kohn JE, Plaxco KW (2005) *Crit Rev Biochem Mol Biol* 40:181–189.
60. Fitzkee NC, Fleming PJ, Gong HP, Panasik N, Street TO, Rose GD (2005) *Trends Biochem Sci* 30:73–80.
61. Muñoz V, Serrano L (1997) *Biopolymers* 41:495–509.
62. Magg C, Kubelka J, Holtermann G, Haas E, Schmid FX (2006) *J Mol Biol* 360:1067–1080.
63. Pancoska P, Janota V, Keiderling TA (1999) *Anal Biochem* 267:72–83.
64. Sreerama N, Venyaminov SY, Woody RW (1999) *Protein Sci* 8:370–380.
65. Schröder GF, Alexiev U, Grubmüller H (2005) *Biophys J* 89:3757–3770.

SI Materials and Methods

Protein Labeling. Labeling was carried out according to the procedures supplied by the manufacturer (Molecular Probes/Invitrogen) under nitrogen atmosphere. First, Alexa Fluor 488 maleimide was reacted with the protein, and singly labeled protein was separated from unlabelled and doubly labeled protein by ion exchange chromatography (Mono Q HR 5/5; GE Healthcare Bio-Sciences AB, Uppsala, Sweden). The fractions containing singly labeled Csp*Tm*, as confirmed by electrospray ionization mass spectroscopy, were labeled with Alexa Fluor 594 maleimide after concentration by ultrafiltration. Doubly labeled protein was again separated from singly labeled protein by ion exchange chromatography. The conformational stabilities ΔG_u and unfolding midpoints $c_{1/2}$ of all labeled variants are given in the following table (cooperativity of the transition for all fits: $m = 6.0\,kJ\,mol^{-1}\,M^{-1}$):

	C2C67	**C10C67**	**C21C67**	**C22C67**	**C34C67**
ΔG_u (kJ/mol)	12 ± 1	10 ± 1	14 ± 1	11 ± 1	14 ± 1
$c_{1/2}$ (M)	2.0 ± 0.1	1.7 ± 0.2	2.3 ± 0.1	1.9 ± 0.2	2.4 ± 0.1

Ensemble steady-state polarization measurements of the attached dyes resulted in anisotropies of ≈0.1 or less for all samples under unfolding conditions, indicating sufficient rotational averaging during the fluorescence lifetime of the dyes to justify using $\kappa^2 = 2/3$ (1).

Calibration and Error Analysis of Single Molecule Data. Identified bursts were corrected for background, differences in quantum yields, the different collection efficiencies of the detection channels, cross-talk (acceptor emission detected in the donor channel and donor emission detected in the acceptor channel), and direct excitation of the acceptor with the following matrix approach (2). The relation between the raw photon counts per identified burst $n_{D,0}\, n_{D,0}$ and $n_{A,0}$, as measured in the two detection channels for acceptor and donor emission, respectively, and the corrected values n'_A and n_D can be expressed by the matrix equation

$$\begin{pmatrix} n_{A,0} \\ n_{D,0} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} n'_A \\ n_D \end{pmatrix} + \begin{pmatrix} b_A \\ b_D \end{pmatrix},$$

where the matrix a_{ij} describes the cumulative effect of the differences in quantum yields, the different collection efficiencies of the detection channels, and cross-talk, i.e. acceptor emission detected in the donor channel and donor emission detected in the acceptor channel. b_A and b_D are the background count rates in the acceptor and the donor channel, which can be estimated from a measurement on blank buffer solutions. The elements of matrix a_{ij} were determined for our instrument (except for a scaling factor ζ) from a measurement of two samples containing protein singly labeled with donor or acceptor dye, respectively, in the micromolar range, with a concentration ratio equal to the ratio of the dyes' extinction coefficients at the excitation wavelength (3) (ensuring that, at identical laser power, the same mean number of excitation events take place per unit time in both samples). By inverting the resulting matrix, the correction matrix $c_{ij} = a_{ij}^{-1}$ is obtained, which transforms the background-corrected raw counts $n_{A,0} - b_A$ and $n_{D,0} - b_D$ into the corrected values n'_A and n_D. Note that the factor ζ remains unknown, but cancels if intensity ratios are computed, as in the case of the transfer efficiency. Finally, n'_A has to be corrected for direct excitation of the acceptor according to $n_A = n'_A - (n'_A + n_D)\ \varepsilon_A/(\varepsilon_A + \varepsilon_D)$, where ε_D and ε_A are the extinction coefficients of donor and acceptor, respectively, at the excitation wavelength. These corrections were also taken into account for burst identification. The dependence of the Förster radius R_0 on denaturant concentration was determined by measuring changes in spectral overlap, donor quantum yield, and the refractive index of the solvent, and was found to be dominated by the change in refractive index.

The error ranges given in Fig. 3 are estimates of combined random and systematic errors. Random errors were derived from multiple independent measurements, and range from about $\sigma_{Exp} = 0.02$ at $\langle E \rangle \approx 0.3$ to $\sigma_{Exp} = 0.04$ at $\langle E \rangle \approx 0.8$. The largest sources of systematic error are the changes in extinction coefficients and quantum yields of donor and acceptor upon attachment to different positions in the protein, and especially their change upon collapse. In lack of a method to measure extinction coefficients and quantum yields independently (which contribute differently to the observed value of $\langle E \rangle$), only changes in the emission of samples with identical optical density can be determined. We find standard deviations for this "emissivity" calculated from all singly labeled variants at 8 M GdmCl (normalized to a maximum of 1) of 0.04 and 0.09 for donor and acceptor, respectively, and we assume that changes in extinction coefficients and quantum yields contribute equally. Additional changes in extinction coefficients and quantum yields upon chain collapse are even more difficult to assess. We assume for our error estimates that these changes are less than the average measured differences between the emissivities of singly labeled samples at 8 M and 0 M

GdmCl, 0.17 and 0.24 for acceptor and donor, respectively. From these assumptions, we obtain estimates for the uncertainties of E and l_p using error propagation.

Subpopulation-Selective Fluorescence Lifetime Distribution Analysis. Bursts assigned to the unfolded subpopulation were identified in a plot of $\tau_{D,burst}$ (the donor fluorescence lifetime estimated with a maximum likelihood algorithm (4)) versus $\langle E \rangle$ (Fig. 4*A*). Photons from these bursts were combined to generate time correlated single photon counting (TCSPC) histograms, $\tilde{I}_D(t)$ and $\tilde{I}_A(t)$, for donor and acceptor, respectively. TCSPC histograms $b_D(t)$ and $b_A(t)$ from all photons not in bursts (5) were used to calculate background-corrected histograms $I_D(t)$ and $I_A(t)$ as

$$I_{D,A}(t) = \tilde{I}_{D,A}(t) - b_{D,A}(t)\,(\text{Total subpopulation burst time}) / (\text{Total non-burst time}) . \qquad \textbf{[5]}$$

The resulting decays were fit globally with the coupled equations

$$\begin{aligned} I_D(t) &= IRF_D * h_D(t - t_{0,D}) \\ I_A(t) &= IRF_A * h_A(t - t_{0,A}), \end{aligned} \qquad \textbf{[6]}$$

where

$$h_D(t) = a_D \int_0^\infty P(r) e^{-k(r)t} dr \quad \text{and} \qquad \textbf{[7]}$$

$$h_A(t) = a_A \int_0^\infty P(r) \frac{k(r)-k_D}{k(r)-k_A} \left(e^{-k_A t} - e^{-k(r)t} \right) dr + \alpha\, h_D(t) + \frac{\varepsilon_A}{\varepsilon_A + \varepsilon_D} I_{tot}\, k_A\, e^{-k_A t} \qquad \textbf{[8]}$$

are convolved with the instrument response functions $IRF_{D,A}$ of the donor and acceptor detection channels, respectively. $t_{0,D}$ and $t_{0,A}$ denote the time origins of the decays, and $a_{D,A}$ their amplitudes. The integrands in Eqs. **7** and **8** are the solutions of the rate equations describing donor and acceptor decay including Förster transfer with the chromophores at distance *r*, weighted by the distance distribution $P(r)$ for a Gaussian chain (Eq. **3**). Distance fluctuations for long polypeptides are expected to be slow relative to the fluorescence lifetime of donor and acceptor (6); we thus assume the inter-dye distance *r* to be constant during the donor fluorescence life time (≤ 4 ns). k_D is the intrinsic decay rate of the donor (without energy transfer), and $k(r) = k_D \left(1 + (R_0 / r)^6\right)$ is the rate of transfer; k_A is the decay rate of the acceptor. The second term on the right hand side of Eq. **8** corrects for the fraction $\alpha = 0.08$ of donor photons detected in the acceptor channel (cross-talk); the relative amount of acceptor photons observed in the donor channel is negligible. The last term in Eq. **8**

accounts for direct excitation of the acceptor; I_{tot} is the total integrated intensity of $I_D(t)$ and $I_A(t)$. Eq. **6** was fit to the background-corrected TCSPC histograms with $t_{0,D}$, $t_{0,A}$, a_D, a_A, k_A, and $\langle r^2 \rangle$ as free parameters. All remaining parameters were determined independently, and a GdmCl concentration-dependent Förster radius R_0 was used as described above.

k_D was obtained from single exponential fits to background-corrected TCSPC histograms of donor photons in bursts with transfer efficiencies less then 0.2, corresponding to molecules lacking an active acceptor chromophore. The resulting values of k_D increase with denaturant concentration. As the population of the native molecules at low GdmCl concentrations hampers the direct measurement of k_D for the unfolded subpopulation, we estimated k_D by linear extrapolation of the data at denaturant concentrations >3 M. In our range of GdmCl concentrations, the resulting k_D varies over a range of 0.27–0.33 s^{-1}. Error bars in Fig. 4 were obtained by taking into account both the maximum plausible range of k_D and the uncertainty in R_0. Remaining non-systematic deviations of the TCSPC fits close to t_0 may be due to uncertainties in the background subtraction caused by the count rate-dependence of the detector response functions.

SRCD Setup. SRCD measurements were implemented essentially as described previously (7). Undulator beam line U125/2-10m NIM at BESSY II synchrotron (Berlin) (8) was coupled to the sample chamber via a LiF window (Korth Kristalle GmbH, Altenholz, Germany) separating the ultra high vacuum of the beam line from the experimental chamber under atmospheric pressure. As the pressure difference causes strain birefringence in the LiF window, a MgF_2 Rochon polarizer (B. Halle Nachfl., Berlin, Germany) follows in the optical path to ensure linear polarization. A photoelastic modulator (Hinds Instruments, Hillsboro, NC; model I/CF50) is used to convert linearly polarized into circularly polarized light. A Suprasil lens (B. Halle Nachfl., Berlin, Germany) with a focal length of 10 mm (at $\lambda = 200$ nm) focuses the beam to a spot size of 60 μm and 25 μm (full width at half maximum) parallel and perpendicular to the direction of sample flow, respectively. Transmitted light is detected with a low noise, solar-blind channel photomultiplier (Perkin Elmer, model CPM 1321). The photon flux at the sample is about $5 \cdot 10^{10}$ photons per second at a bandwidth of 0.3 nm. The absolute CD sensitivity was calibrated with (1S)-(+)-camphor-10-sulfonic acid (9). The accessible wavelength range in Fig. 5*E* is limited by GdmCl absorption.

1. Schuler B, Lipman EA, Steinbach PJ, Kumke M, Eaton WA (2005) *Proc Natl Acad Sci USA* 102**:**2754-2759.
2. Schuler B (2006) in *Protein Folding Protocols*, eds Bai Y, Nussinov R (Humana, Totowa, NJ), Vol 366.
3. Schuler B, Lipman EA, Eaton WA (2002) *Nature* 419**:**743-747.
4. Eggeling C, Berger S, Brand L, Fries JR, Schaffer J, Volkmer A, Seidel CA (2001) *J Biotechnol* 86**:**163-80.
5. Laurence TA, Kong XX, Jager M, Weiss S (2005) *Proc Natl Acad Sci USA* 102**:**17348-17353.
6. Lapidus LJ, Steinbach PJ, Eaton WA, Szabo A, Hofrichter J (2002) *J Phys Chem B* 106**:**11628-11640.
7. Sutherland JC (1996) in *Circular Dichroism and the Conformational Analysis of Biomolecules*, ed Fasman GD (Plenum, New York), pp 599-633.
8. Reichardt G, Bahrdt J, Schmidt JS, Gudat W, Ehresmann A, Muller-Albrecht R, Molter H, Schmoranzer H, Martins M, Schwentner N*, et al.* (2001) *Nucl Inst Methods Phys Res Sec a* 467**:**462-465.
9. Chen GC, Yang JT (1977) *Anal Lett* 10**:**1195-1207.