

Ultrafast dynamics of protein collapse from single-molecule photon statistics

Daniel Nettels*, Irina V. Gopich[†], Armin Hoffmann*, and Benjamin Schuler*[§]

*Biochemisches Institut, Universität Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland; and [†]Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

Communicated by William A. Eaton, National Institutes of Health, Bethesda, MD, December 13, 2006 (received for review November 1, 2006)

We use the statistics of photon emission from single molecules to probe the ultrafast dynamics of an unfolded protein via Förster resonance energy transfer. Global reconfiguration of the chain occurs on a time scale of ≈ 50 ns and slows down concomitant with chain collapse under folding conditions. These diffusive dynamics provide a missing link between the phenomenological chemical kinetics commonly used in protein folding and a physical description in terms of quantitative free energy surfaces. The experiments demonstrate the potential of single-molecule methods in accessing the biologically important nanosecond time scales even in heterogeneous populations.

correlation | fluorescence | Hanbury Brown and Twiss | photon bunching | protein folding

The discovery of proteins that fold rapidly in the absence of intermediates (1) has substantially advanced our mechanistic understanding of protein folding. The simplicity of their folding behavior allows thermodynamic and kinetic analyses that have led, e.g., to the characterization of transition states for folding (2) and the prediction of folding rates from native structure (3-5). However, because of their limited experimental accessibility, energetic and dynamic differences between the unfolded states of different proteins have usually been ignored. Similarly, the application of rigorous theories of protein folding in terms of statistical mechanics (6-8) has been hampered by our lack of understanding of structure and dynamics in the unfolded state ensemble. The role of unfolded states in determining protein folding mechanisms is thus largely unknown. Although considerable information about the overall dimensions and residual structure of unfolded proteins has been obtained from methods such as small-angle x-ray scattering and NMR, their dynamics in the submicrosecond range have largely eluded experimental determination. The importance of these time scales has become particularly obvious through the identification of proteins that fold in a few microseconds (9). In this regime, the free energy barrier to folding is assumed to be extremely low or even absent, and diffusive chain dynamics become the dominate factor in folding kinetics. Here, we determine these dynamics for an unfolded protein and investigate how they are affected by the collapse of the unfolded chain under near-physiological conditions.

An ideal way to probe the dynamics of the heterogeneous ensemble of unfolded protein conformations is single-molecule spectroscopy (10). The absence of averaging over many molecules allows spontaneous intramolecular distance fluctuations to be observed at equilibrium, without the need for perturbations to synchronize the ensemble. Förster resonance energy transfer (FRET) between two chromophores attached to the polypeptide chain has been suggested as an approach for investigating its submicrosecond dynamics (11, 12), but has eluded experimental implementation. Here, we use this method to directly probe the unfolded state dynamics of the cold shock protein (Csp) from *Thermotoga maritima*, a small, 7.5-kDa β -barrel protein that exhibits two-state thermodynamics and kinetics (13-17). The protein was labeled terminally with a green fluorescent donor and a red fluorescent acceptor dye via amino- and carboxyl-terminal cysteine residues, and freely diffusing molecules were observed in confocal

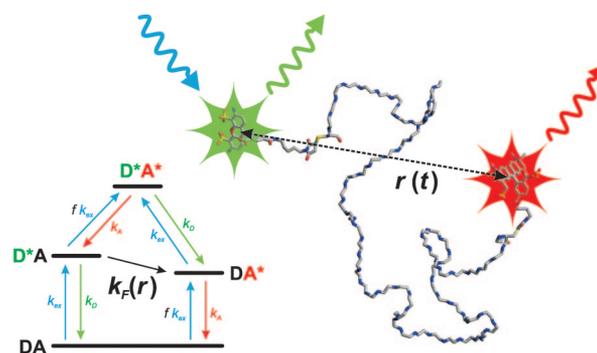


Fig. 1. Principle of measuring distance fluctuations in unfolded proteins with FRET. Energy level diagram (*Left*) of the FRET process between two fluorophores (A and D) attached to the termini of a protein (*Right*) with the relevant transitions between the electronic states. The Förster transfer rate $k_F(r) = k_D(R_0/r)^6$ fluctuates on the same time scale as the distance $r(t)$ between the chain ends and thus allows a direct measurement of chain dynamics. k_{ex} is the excitation rate of the donor D, $f k_{ex}$ is the rate of direct excitation of the acceptor A, and k_D and k_A are the intrinsic decay rates of donor and acceptor, respectively. The possible role of annihilation processes (66-68) in the state D^*A^* is to the best of our knowledge currently unknown for these chromophores and has thus not been included in our photophysical model. Note, however, that these processes would be expected to occur on a time scale shorter than or similar to the fluorescence lifetime of the dyes and thus would not affect our analysis of the correlation functions on the time scale of protein dynamics.

single-molecule experiments (Fig. 1). During the transit of the protein through the observation volume, its donor chromophore is excited by the laser beam. Depending on the distance r to the acceptor, energy transfer results with a rate that determines the relative probabilities of photon emission from donor and acceptor. Correspondingly, distance dynamics within the protein can be measured by fluctuations in the transfer efficiency and thus in the fluorescence emission of the chromophores (Fig. 1).

Results

Measurement of Unfolded State Dynamics. In our experiments, we combine the separation of folded and unfolded subpopulations by single-molecule spectroscopy with the high time resolution available from the photon statistics of a FRET-coupled dye pair. First, a transfer efficiency histogram is created from the photon bursts of individual molecules diffusing through the focus (15,

Author contributions: D.N. and B.S. designed research; D.N., I.V.G., A.H., and B.S. performed research; D.N., I.V.G., and B.S. contributed new reagents/analytic tools; D.N. and B.S. analyzed data; and D.N., I.V.G., and B.S. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: FRET, Förster resonance energy transfer; Csp, cold shock protein; GdmCl, guanidinium chloride; APD, avalanche photodiode.

[§]To whom correspondence should be addressed. E-mail: schuler@bioc.unizh.ch.

This article contains supporting information online at www.pnas.org/cgi/content/full/0611093104/DC1.

© 2007 by The National Academy of Sciences of the USA

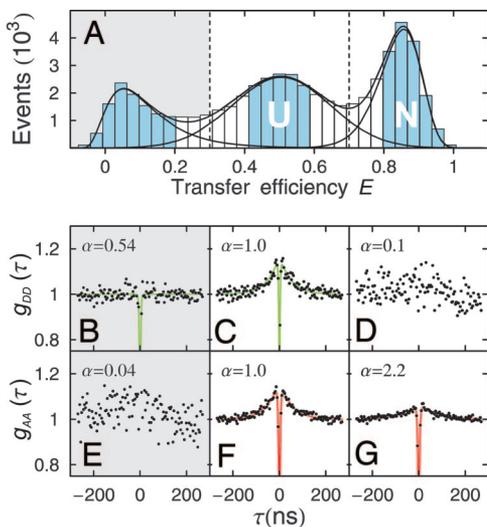


Fig. 2. Intensity autocorrelation functions of the subpopulations identified in single-molecule experiments. (A) Transfer efficiency histogram at 1.4 M GdmCl fit with three peaks (black lines) corresponding to the folded or native (N; high E) and unfolded (U; $E \approx 0.5$) subpopulations, and molecules lacking an active acceptor ($E \approx 0$). (B–G) The ranges of E used for extracting the specific normalized donor and acceptor autocorrelation functions g_{DD} (B–D) and g_{AA} (E–G) are shaded in blue. Fits to g_{DD} are shown in green, and fits to g_{AA} are shown in red (for fit functions, see *Materials and Methods*). The normalized signal amplitude α for each correlation function is given relative to the unfolded state ($\alpha = 1$). g_{DD} and g_{AA} were calculated from the corresponding interphoton time distributions by correcting for pile-up and triplet-state components (see *Materials and Methods*).

18). Fig. 2A shows an example at a concentration of the denaturant guanidinium chloride (GdmCl) of 1.4 M, where three subpopulations are resolved: folded protein molecules with a transfer efficiency E close to 1, unfolded molecules with $E \approx 0.5$, and molecules lacking an active acceptor chromophore with $E \approx 0$ (15, 19). A second synchronized counting card continuously records the time intervals τ between consecutive detected photons with picosecond time resolution. Dead times of detectors and counting electronics are avoided by using a Hanbury Brown and Twiss detection scheme (20–22). Under our conditions, normalized histograms of the resulting interphoton times are essentially equivalent to the donor or acceptor intensity autocorrelation functions $g_{ii}(\tau)$ ($i = A, D$) (23), which report directly on the time scales of fluctuations in fluorescence intensity (see *Materials and Methods*). We can obtain $g_{ii}(\tau)$ for each of the subpopulations shown in Fig. 2A by using only the interphoton times from molecules assigned to a single subpopulation.

The resulting intensity autocorrelation functions of donor [$g_{DD}(\tau)$] and acceptor [$g_{AA}(\tau)$] emission for every subpopulation are shown in Fig. 2B–G. In all cases, $g_{ii}(\tau)$ exhibits a drop in amplitude at $\tau = 0$. This photon antibunching is characteristic of individual quantum systems that cannot emit two photons simultaneously; it decays within a few nanoseconds (22, 24). More interestingly, pronounced photon bunching, i.e., an additional component in $g_{ii}(\tau)$ with positive amplitude, is observed in the 50-ns range. The decay of this bunching signal in the unfolded subpopulation ($E \approx 0.5$) is described by similar time constants for donor and acceptor autocorrelation functions (Fig. 2C and F). The behavior of $g_{ii}(\tau)$ agrees with theoretical predictions for a flexible polypeptide chain (11, 12): if, for example, a donor photon is emitted at $\tau = 0$, the chain ends are likely to be far apart at that instant, corresponding to a low rate of energy transfer, k_F . A very short time later, the ends will still be far apart, and the likelihood of emitting another donor photon will still be increased. However, at times much greater than the reconfiguration time τ_r of the chain, the molecule will have lost the

“memory” of its initial configuration at $\tau = 0$, and the probability of donor emission will be determined by the average transfer efficiency. In other words, we expect an increased autocorrelation of the emission intensity around $\tau = 0$ that decays approximately on the time scale of chain reconfiguration.[†]

The acceptor autocorrelation function $g_{AA}(\tau)$ of the native subpopulation also exhibits a small, but significant, relative bunching signal (Fig. 2G). This signal may in part be caused by the imperfect separation of subpopulations (Fig. 2A), but the slower relaxation compared with the unfolded state could indicate an additional contribution from interactions of the chromophores with the specific environment presented by the native structure. To exclude an influence of native molecules on the determination of unfolded state dynamics, we restrict our measurements and analysis to the donor signal, where only unfolded molecules contribute to bunching (Fig. 2B–D). The optimal signal-to-noise ratio is obtained at a protein concentration of 500 pM, where a separation of subpopulations is no longer possible, but the determination of the correlation times is most accurate. An example for such a measurement is shown in Fig. 3B; the pronounced photon bunching yields an intensity autocorrelation time of 44 ± 3 ns. Under identical conditions, this signal is absent in stiff (19) polyproline peptides with the same average transfer efficiency as unfolded Csp (Fig. 2D), confirming that photon bunching is indeed caused by chain dynamics in the unfolded protein. The bunching signal is also absent in Csp labeled only with a donor chromophore (data not shown), excluding dye–protein interactions and rotational dynamics of the entire chain as origins of the bunching signal we observe. An additional slower component with a relaxation time of several microseconds was observed in all data sets, as expected from the triplet-state lifetimes of the chromophores (25). Fluorophore reorientation occurs on the time scale of ≈ 300 ps (19); resulting fluctuations in the energy transfer rate would thus affect the correlation functions only on time scales much shorter than the range relevant for the dynamics investigated here.

Theory and Analysis of Photon Statistics. To analyze our measurements in terms of protein dynamics, we describe the relative motion of the chain ends as a diffusive process on the potential of mean force that corresponds to the end-to-end distance distribution of the unfolded protein. By combining these diffusive chain dynamics with the distance-dependent stochastic photon emission from the coupled dye pair (Fig. 1), the complete photon statistics of the system can be obtained (12). To avoid time-consuming simulations, we use a recently developed theory (26) that allows us to calculate the intensity correlation functions of donor and acceptor emission numerically. Briefly, if we assume for the unfolded protein the end-to-end distance distribution function of a Gaussian chain [which has recently been shown to be a good approximation for Csp (27)], $p_{eq}(r) = 4\pi r^2 (2\pi \langle r^2 \rangle / 3)^{-3/2} \exp(-3r^2/2\langle r^2 \rangle)$ (see Fig. 5), protein dynamics can be combined with the photophysics of FRET in the rate matrix:

$$\mathbf{K} = D \partial / \partial r p_{eq}(r) \partial / \partial r (p_{eq}(r))^{-1} \mathbf{I} + \mathbf{K}_0(r), \quad [1]$$

where D is the relative diffusion coefficient of the chain ends, $\mathbf{K}_0(r)$ describes the distance-dependent kinetics of interconversion between the four electronic states illustrated in Fig. 1, and \mathbf{I} is the 4×4 -identity matrix. The time dependence of all electronic and conformational transitions in the system is then described by the rate equation $d\mathbf{p}/dt = \mathbf{K}\mathbf{p}$, where \mathbf{p} is the vector of the populations of the four electronic states in Fig. 1. By discretizing the diffusion

[†]Because of the nonlinear dependence of the transfer rate k_F on distance, τ_{DD} , τ_{AA} , and τ_r are not strictly identical, but they can be related accurately using our analysis [see *Theory and Analysis of Photon Statistics* and [supporting information \(SI\) Text](#)]. In our range of distances and end-to-end diffusion coefficients (Fig. 4), the three time constants differ by $<15\%$.

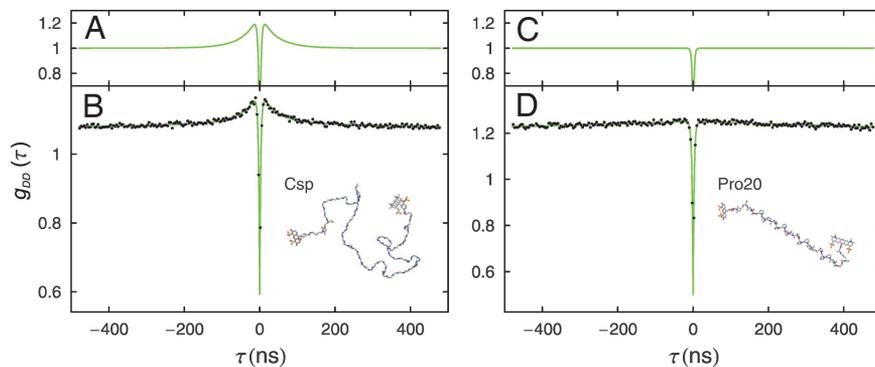


Fig. 3. Calculated (A and C) and measured (B and D) donor intensity autocorrelation functions for unfolded Csp (A and B) and a stiff polyproline peptide (C and D) at 4 M GdmCl, corroborating that the 44-ns component of g_{DD} is caused by chain dynamics. The curve in C was calculated based on the photophysics of the FRET process assuming a fixed distance between the fluorophores, whereas in A diffusive end-to-end distance dynamics of a Gaussian chain were included.

operator, the problem is reduced to matrix algebra, and the intensity correlation functions can be calculated to high accuracy with numerical methods (for details, see *Materials and Methods* and *SI Text*). All parameters needed to define the model in terms of our system are known: the photophysical rate constants of the FRET process (Fig. 1) were measured independently, and $p_{eq}(r)$ is defined uniquely by the mean square end-to-end distance $\langle r^2 \rangle$, which can be calculated from the average transfer efficiency of the unfolded subpopulation (15, 27) (see *SI Text*). We can thus determine the only remaining parameter, the effective end-to-end diffusion coefficient D , by adjusting it such that the calculated intensity autocorrelation function fits the experimental result. Finally, the chain reconfiguration time τ_r (the decay time of the end-to-end distance autocorrelation function) for a Gaussian chain is obtained from $\tau_r \approx \langle r^2 \rangle / 6D$ (see *SI Text*).

In summary, we can thus determine, from the measured intensity correlation function, the dynamics of an unfolded protein in terms of the diffusion coefficient D and the corresponding reconfiguration time τ_r . Examples for calculated intensity correlation functions both with and without chain dynamics are shown in Fig. 3 A and C, respectively. The excellent agreement between the functional forms of experimental and calculated correlation functions (Fig. 3 A and B) suggests that our single reaction coordinate is a reasonable approximation for the dynamics of unfolded Csp. It is worth stressing that what we observe here are large-scale chain dynamics along this reaction coordinate, whereas the wide range of subnanosecond dynamics known to occur in polypeptides (28, 29), such as dihedral angle rotations, essentially enter into the diffusion coefficient D (or the energetic roughness of the free energy surface; see *The Free Energy Surface of Collapse and Implications for Protein Folding*).

Effect of Collapse on Unfolded State Dynamics. The unfolded state of Csp has been found to collapse in response to decreasing GdmCl concentrations (15, 17, 27, 30). This collapse precedes the folding reaction in kinetic experiments (17, 30), and single-molecule FRET can be used to quantify the resulting change in chain dimensions under equilibrium conditions by virtue of the separation of folded and unfolded subpopulations (15, 27). Fig. 4A shows the corresponding decrease in the rms end-to-end distance $(\langle r^2 \rangle)^{1/2}$ determined from the mean transfer efficiencies of the unfolded state (27). For an ideal chain, compaction is expected to lead to faster relaxation of intramolecular distances because of the reduced size of the accessible conformational space (Fig. 4B). This behavior has indeed been observed for some unstructured peptides (31) (however, also see ref. 52), but what happens in the case of a real protein? To answer this question, we measured the change in the decay time of the donor intensity autocorrelation function τ_{DD} in response to decreasing GdmCl concentrations (Fig. 4B). With the above anal-

ysis, we find a decrease in the viscosity-corrected diffusion coefficient D_η from $\approx 0.5 \text{ nm}^2/\text{ns}$ at 8 M GdmCl to $\approx 0.1 \text{ nm}^2/\text{ns}$ at low GdmCl concentrations [corrected for change in solvent viscosity η_s with increasing denaturant concentration according to $D_\eta = D\eta_s^{-1}$ 1 mPa s(31–35)]. The corresponding chain reconfiguration time τ_r increases from $\approx 20 \text{ ns}$ at 8 M GdmCl to $\approx 65 \text{ ns}$ under near-native conditions (Fig. 4B). In conclusion, we find very fast global chain

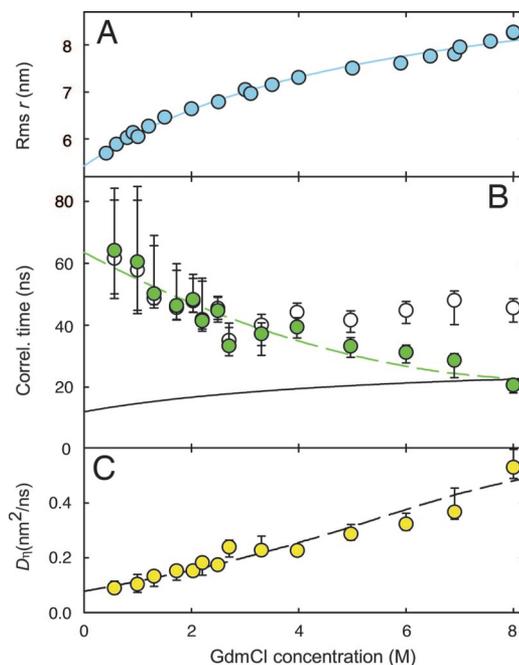


Fig. 4. Denaturant dependence of unfolded state collapse and dynamics. (A) rms end-to-end distance r of the unfolded protein as determined from the mean transfer efficiency of the unfolded state in single-molecule fluorescence experiments. The solid line is an empirical fit to the equation $r = r_0 + Ac/(1 + Bc)$ used to interpolate the values of rms r for further calculations. (B) Donor intensity autocorrelation times τ_{DD} (open circles) and viscosity-corrected end-to-end distance autocorrelation times τ_r (green circles) of unfolded molecules determined from measurements as shown in Fig. 3. The dashed green line is an empirical second-order polynomial fit to the distance correlation times. The black line shows the calculated reconfiguration time of an ideal Gaussian chain assuming a constant diffusion coefficient and the rms r values in A. (C) Effective viscosity-corrected relative diffusion coefficients of the chain ends D_η , illustrating the slowed chain dynamics upon collapse. The dashed line is calculated from the solid and dashed fits in A and B, respectively, according to $D_\eta = \langle r^2 \rangle / 6\tau_r$. Error bars represent our estimate of the combined experimental errors and the systematic uncertainty in data analysis.

reconfiguration in the unfolded state of Csp, but, concomitant with chain collapse at low denaturant concentrations, we observe a deceleration of these dynamics. This finding is opposite to the behavior expected for an ideal chain with invariant D_η (Fig. 4B), suggesting that interactions within the polypeptide chain, or “intrinsic friction” (36), significantly affect unfolded state dynamics upon chain collapse, similar to what has recently been found for unstructured GlySer-repeat peptides (52).

The Free Energy Surface of Collapse and Implications for Protein Folding. With the shape of the end-to-end distance distribution p_{eq} and the effective end-to-end diffusion coefficient D_η as a function of the denaturant concentration, we have obtained the two key parameters for describing the collapse of Csp in terms of a quantitative free energy surface. The change in D_η upon collapse can be expressed in terms of an effective energetic “roughness”, i.e., a distribution of small energy barriers caused by intramolecular interactions that slow down diffusion along the reaction coordinate. Possible physical origins of the dependence of D_η on denaturant concentration, as captured by the roughness, are changes in the packing and interaction strength of the polypeptide backbone and side chains in the unfolded state. These interactions may also contribute to the increase in the β -structure content recently observed upon collapse of unfolded Csp (27). Assuming a random amplitude with a Gaussian distribution independent of r , the rms roughness ε is given by $\varepsilon = k_B T \sqrt{\ln(D_0/D_\eta)}$ (37, 38), where D_0 is the diffusion coefficient in the absence of intramolecular interactions. For D_0 , we choose the viscosity-corrected diffusion coefficient of the unfolded state in 8 M GdmCl. At such high denaturant concentrations, the average chain volume is ≈ 18 times greater than in the folded structure, suggesting complete solvation of the polypeptide chain, and consequently minimal intramolecular interactions. The decrease in D_η we observe upon collapse under native conditions corresponds to an increase in roughness by 1.3 (+0.1/−0.2) $k_B T$. In Fig. 5, we plot the relative energetic roughness superimposed on the potential of mean force $G(r) = -k_B T \ln p_{eq}(r)$ of unfolded Csp as a function of the denaturant concentration to illustrate the influence of collapse on the energy landscape.

Discussion

The free energy surface in the absence of denaturant (Fig. 5) summarizes the global structure and dynamics of unfolded Csp molecules under folding conditions. The rate of folding is determined both by the effective free energy barrier ΔG^\ddagger separating folded and unfolded states, and the “attempt frequency” for crossing the barrier from the unfolded state. In generalized transition state expressions for the folding time τ_f of type $\tau_f = \tau_0 \exp(\Delta G^\ddagger/k_B T)$, the reconfiguration time τ_r of the chain within the unfolded state potential is directly related to the inverse attempt frequency, or the preexponential factor τ_0 . In a simplified Kramers description (39–41), where the curvature of the free energy surface and the effective diffusion coefficient are assumed to be similar in the unfolded well and on the top of the barrier, the preexponential factor is given by $\tau_0 \approx 2\pi\tau_r$. In the absence of a barrier, τ_0 remains as an approximation to the “speed limit” (42), i.e., the minimum time scale of folding. With our value of 65 ns for τ_r of unfolded Csp under native conditions, large-scale chain diffusion thus sets a lower limit of $\tau_0 \approx 0.4 \mu\text{s}$ to the folding time for a protein the size of Csp^{ll}, remarkably similar to estimates based on the length scaling of folding rates (9, 43). It remains to be clarified whether the decrease in the

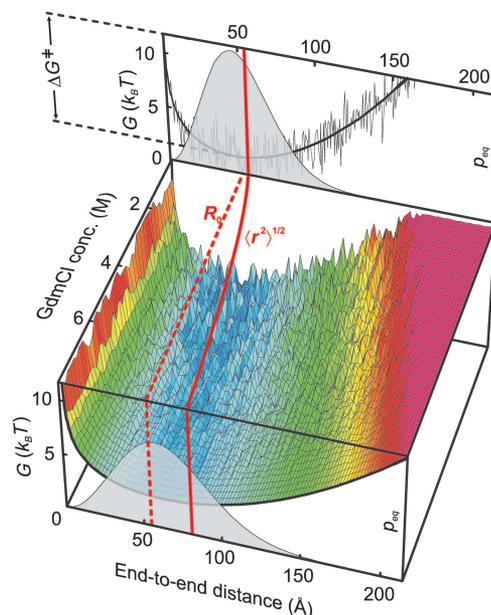


Fig. 5. Schematic free energy surface of Csp collapse. End-to-end distance distribution functions p_{eq} (gray filled functions) and the corresponding potentials of mean force G (thick black lines) of unfolded Csp at 8 M GdmCl (Lower) and 0 M GdmCl (Upper) are shown, assuming a Gaussian chain model (27). A 3D plot of the potential of mean force in the range of 8 to 0 M GdmCl illustrates the collapse and increasing roughness calculated from the data in Fig. 4. The Förster radius R_0 (dashed red line) and the rms end-to-end distance (rms r , red line) are indicated. The thin line superimposed on the average free energy surface (thick line) in Upper (0 M GdmCl) shows one representative distribution of energies compatible with the model for the roughness used (37). Note, however, that the discretization of the end-to-end distance is arbitrary and was introduced only for the purpose of illustration. G is given in units of $k_B T$ and scaled to a maximum value of 10 $k_B T$ relative to the minimum of the average potential at 0 M GdmCl, where 10 $k_B T$ is the height of the free energy barrier for folding.

intramolecular diffusion coefficient we observe upon collapse (Fig. 4) continues toward the transition state region, as suggested from theory (44–47). The rapid folding observed for some proteins [folding times ranging from 0.7 μs for a variant of villin headpiece (48) to several microseconds (9)] suggests that this effect might not be large. This issue could be addressed with a combination of single-molecule experiments of the type described here, and laser temperature-jump experiments, which have been used to determine the molecular time scale of $\approx 2 \mu\text{s}$ for the relaxation from the transition state region to the unfolded state in variants of the five-helix bundle protein λ_{6-85} (49, 50).

Note, however, that the preexponential factor for Csp folding as approximated above is virtually invariant compared with the nine orders of magnitude decrease in its folding time τ_f from 8 to 0 M GdmCl (13). The large change in τ_f is thus almost entirely caused by a change in ΔG^\ddagger , supporting the assumption of a linear dependence of ΔG^\ddagger on denaturant concentration frequently made in the analysis of protein folding kinetics (2, 13).

Comparison to Peptide Dynamics. Short unstructured peptides have been investigated intensively as model systems for the chain dynamics of real proteins, using ensemble lifetime fluorometry (51, 52) and quenching methods (32–35, 53, 69, 70). Extrapolating the rates of contact formation observed in quenching experiments of peptides to the length of the Csp chain yields values between $\approx 0.5 \times 10^6 \text{ s}^{-1}$ and $5 \times 10^6 \text{ s}^{-1}$. For a direct comparison to our results, we estimate the end-to-end contact formation rate for unfolded Csp assuming a Gaussian chain as $k_c = 4\pi Da / (2\pi \langle r^2 \rangle / 3)^{3/2}$ (54), where a is the contact distance. With our values for D and $\langle r^2 \rangle$ (Fig. 4), and

^{ll}The folding time of 12 ms for labeled CspTm(15) results in a free energy barrier for folding of 10 kT at 0 M GdmCl (Fig. 5).

$a = 0.4$ nm (34, 52), we calculate contact rates of $(0.8 \pm 0.1) \times 10^6$ s⁻¹, within the extrapolated values from peptide dynamics. Similarly, the end-to-end diffusion coefficients between 0.04 and 0.2 nm²/ns, obtained for peptides under a variety of denaturant concentrations (34, 52), are close to the values we find for unfolded Csp**. The similarity of the dynamic time scales observed in unfolded Csp and unstructured peptides, and the single exponential decay of our measured intensity correlation functions (Fig. 3C), indicate the absence of specific interactions and corresponding large energy barriers (55), even in collapsed unfolded Csp under native conditions.

Rate of Collapse. Finally, we point out that, according to Onsager's regression hypothesis, or, more generally, the fluctuation dissipation theorem, small fluctuations decay on the average in exactly the same way as macroscopic deviations from equilibrium (56). In other words, the time correlation functions of the spontaneous end-to-end distance fluctuations we observe in single molecules (Figs. 2 and 3) decay with the same time constants as the macroscopic signal would in ensemble perturbation experiments, such as rapid changes in solution conditions or laser-induced temperature jumps. The reconfiguration times measured in our experiments are thus equivalent to the collapse time of unfolded Csp. The similarity in time scale to the dynamics of unstructured peptides shows that the collapse of Csp is a purely diffusive, "downhill" process. Conversely, previous observations of denatured state dynamics in the microsecond range (57, 58) thus probably involve the crossing of substantial barriers.

Our measurement of the reconfiguration time of ≈ 50 ns is in good agreement with the relaxation time expected for an ideal chain (59–61) and theoretical estimates of the collapse times for proteins (43, 62). Additional evidence for the connection between reconfiguration dynamics and the collapse time comes from laser-induced temperature jump experiments on the 40-residue protein BBL; the collapse of its acid-denatured state occurs on a time scale of ≈ 60 ns (63), remarkably similar to the reconfiguration time of Csp. Future experiments will have to address in more detail issues such as the dependence of collapse times on chain length, temperature, and other parameters (43). For BBL, the enthalpy change involved in collapse evidently is large enough to allow temperature-induced perturbations of chain compactness. For a closely related Csp, laser-induced temperature jump experiments have not resulted in an observable signal (64), raising the possibility that its collapse does not involve a sufficient enthalpy change. Other perturbation methods, such as pressure jump and even the fastest mixing methods, are currently limited to time scales in the tens of microsecond range and above. Spontaneous fluctuations of individual molecules at equilibrium, however, are a universal property of sufficiently thermalized systems. Single-molecule photon statistics thus provide a versatile approach for investigating the dynamics of macromolecules on the biologically important nanosecond time scales.

Materials and Methods

Samples and Instrumentation. Labeled Csp and polyproline peptides were produced as described (15, 19, 27). All experiments were performed in 50 mM sodium phosphate buffer adjusted to pH 7, containing 0.001% Tween 20 to prevent surface adhesion of the polypeptides. High-purity GdmCl solutions (Pierce, Rockford, IL) were used for denaturation experiments. Single-molecule fluorescence was observed by using a PicoQuant MicroTime 200 confocal microscope equipped with a continuous wave solid-state diode-pumped laser (Sapphire 488-200; Coherent, Santa Clara, CA)

operating at 488 nm (average radiant power at the sample: 100 μ W), a 1.2 NA, $\times 60$ microscope objective (UplanApo $\times 60/1.20$ W; Olympus, Melville, NY), and a 100- μ m confocal pinhole. A dichroic mirror (585DCXR; Chroma, Rockingham, VT) separated donor and acceptor fluorescence. Subsequently, each fluorescence component was divided randomly by a 50/50 beam splitter between a pair of two avalanche photodiodes (APDs; Optoelectronics SPCM-AQR-15; PerkinElmer, Wellesley, MA). Additional interference filters [HQ525/50 (Chroma) and 525AF45 (Omega Optical, Brattleboro, VT) for the donor APDs, 600HQLP (Chroma) and HQ640/100 (Chroma) for the acceptor APDs] completed spectral separation of the sample fluorescence and served to suppress the mutual detection of APD breakdown flashes in the infrared (65). Each of the two APD pairs was connected to a photon counter: the donor detectors to a PicoHarp 300 and the acceptor detectors to a TimeHarp 200 (both PicoQuant, Berlin, Germany). In their histogram mode, which we used for the measurements at 500 pM Csp and Pro20, the two input channels of these cards operate as start and stop channels. The measured time intervals Δt between start and stop were histogrammed in 256-ps (PicoHarp) or 304-ps (TimeHarp) bins. To avoid cross-talk between the two channels at short time intervals and simplify data analysis, an electronic time delay Δt_0 was imposed onto the stop channel (SI Fig. 6). Mean transfer efficiency data for calculating rms end-to-end distances (Fig. 4A) were taken from ref. 27. 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. For more details on data acquisition and analysis and the measurement of subpopulation-specific correlation functions (Fig. 2), see SI Text.

Analysis of Interphoton Time Distributions. The histograms shown in Fig. 3B and C (integration time 10 h) represent interphoton time distributions $\bar{\phi}_{ii}(\Delta t)$ ($i = A, D$), which are, in the limit of very low mean photon detection rates τ_i^{-1} , proportional to the intensity autocorrelation function $g_{ii}(\tau)$ with the lag time $\tau = \Delta t - \Delta t_0$. At higher count rates, shorter interphoton times are detected more frequently than longer ones. To correct for this pile-up effect, we fitted the data to:

$$\bar{\phi}_{ii}(\Delta t) = A e^{-\Delta t/\tau_i} g_{ii}(\Delta t - \Delta t_0), \quad [2]$$

where $g_{ii}(\tau)$ was approximated by

$$\begin{aligned} g_{ii}(\tau) &= g_{AB}(\tau) \cdot g_{CD}(\tau) \cdot g_T(\tau) \\ &= (1 - c_{AB} e^{-|\tau|/\tau_{AB}})(1 + c_{CD} e^{-|\tau|/\tau_{ii}})(1 + c_T e^{-|\tau|/\tau_T}), \end{aligned} \quad [3]$$

assuming separation of time scales. A is an overall amplitude, $\exp(-\Delta t/\tau_i)$ accounts for the pile-up effect (compare Eq. 8), g_{AB} , g_{CD} , and g_T correspond to the contributions of photon antibunching, chain dynamics, and triplet-state dynamics, respectively, to the overall autocorrelation function $g_{ii}(\tau)$. For each sample, we determined the triplet-state correlation time τ_T from independent conventional fluorescence correlation spectroscopy measurements. Typical values were in the range of 1–4 μ s. All remaining parameters of $\bar{\phi}_{ii}(\Delta t)$ were determined by least-square fitting to the histogram data. The values obtained for τ_i^{-1} ($\approx 10^5$ s⁻¹) are in good agreement with the mean photon count rates during a photon burst, i.e., during the passage of a molecule through the observation volume. The pile-up-corrected correlation data shown in Fig. 3 were obtained by dividing each histogram value by $A \exp(-\Delta t/\tau_i)$.

Calculation of Intensity Correlation Functions and Interphoton Time Distributions. The intensity autocorrelation functions of fluorescence from donor and acceptor are defined as:

**The somewhat lower intramolecular diffusion coefficients compared to Csp reported recently for GlySer-repeat peptides (52) may be caused by the strong intramolecular hydrogen bonding that has been suggested to occur within these very compact peptides.

$$g_{ii}(\tau) = \frac{\langle n_i(0)n_i(\tau) \rangle}{\langle n_i \rangle^2}, \quad i = A, D, \quad [4]$$

where $n_i(\tau)$ is the fluorescence count rate (photon counts per unit time), which fluctuates because of the stochastic processes of photophysics and protein dynamics. The correlation functions are normalized such that $g_{ii}(\tau) \rightarrow 1$ as $\tau \rightarrow \infty$. The distribution of the times between donor or between acceptor photons are denoted by $\phi_{DD}(\tau)$ and $\phi_{AA}(\tau)$ and are normalized according to $\int_0^\infty \phi_{ii}(\tau) d\tau = 1$, $i = A, D$.

Recently, a general theory of these quantities has been developed (26). In this theory, the rate matrix \mathbf{K} describes all electronic and conformational transitions in the system, and the matrix elements of the off-diagonal matrices \mathbf{V}_A and \mathbf{V}_D specify those transitions that result in the emission of acceptor or donor photons. Let \mathbf{p}_{ss} be the vector of normalized steady-state probabilities obtained by solving $\mathbf{K}\mathbf{p}_{ss} = 0$ ($\mathbf{1}^\top \mathbf{p}_{ss} = \sum_i \mathbf{p}_{ss}(i) = 1$, where $\mathbf{1}$ is the unit vector). The mean time between detected acceptor (donor) photons, τ_i , and the mean fluorescence count rate $\langle n_i \rangle$ ($i = A, D$) are then given by:

$$\tau_i^{-1} = \langle n_i \rangle = \mathbf{1}^\top \mathbf{V}_i \mathbf{p}_{ss} \quad [5]$$

(see equation 2.16 of ref. 26). The intensity correlations are:

$$g_{ii}(\tau) = \tau_i^2 \mathbf{1}^\top \mathbf{V}_i e^{\mathbf{K}\tau} \mathbf{V}_i \mathbf{p}_{ss} \quad [6]$$

(see equation 2.18 of ref. 26). The interphoton time distributions are:

$$\phi_{ii}(\tau) = \tau_i \mathbf{1}^\top \mathbf{V}_i e^{(\mathbf{K}-\mathbf{V}_i)\tau} \mathbf{V}_i \mathbf{p}_{ss} \quad [7]$$

(see equation 2.25 of ref. 26). The interphoton time distributions can be related to the corresponding intensity correlation functions if the mean time between photons τ_i is much longer (in our case on the microsecond time scale) than the relaxation time of the correlation function (in our case on the nanosecond time scale). Then, using the separation of time scales, the interphoton time distribution can be approximated as:

$$\phi_{ii}(\tau) \approx \tau_i^{-1} g_{ii}(\tau) e^{-\tau/\tau_i}. \quad [8]$$

We confirmed numerically that this approximation holds very well under the conditions of our experiments. Thus, the intensity correlation function can be obtained by multiplying the interphoton time distribution by an exponential (pile-up effect correction; compare Eq. 2).^{††} Details of the theoretical model used for calculating $g_{ii}(\tau)$ and the procedures for comparing measured and calculated correlation functions are given in *SI Text*.

Note Added in Proof. Two papers (71, 72) published since submission of this manuscript are closely related to the present work.

^{††}Note that the experimentally determined $\tilde{\phi}_{ii}$ of Eq. 2 and the calculated ϕ_{ii} of Eq. 8 are related according to $\tilde{\phi}_{ii}(\Delta t) = A e^{-\Delta t/\tau_i} \tau_i^{-1} \phi_{ii}(\Delta t - \Delta t_0)$.

We thank Attila Szabo for guidance on theoretical issues, and William Eaton, Peter Hamm, Christian Hübner, Gerhard Hummer, and Rudolf Rigler for discussion and helpful comments on the manuscript. This work was supported by the Swiss National Science Foundation and the Human Frontier Science Program. I.V.G. was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Diabetes and Digestive and Kidney Diseases.

1. Jackson SE (1998) *Fold Des* 3:R81–R91.
2. Fersht AR (1998) *Structure and Mechanism in Protein Science* (Freeman, New York).
3. Muñoz V, Eaton WA (1999) *Proc Natl Acad Sci USA* 96:11311–11316.
4. Galzitskaya OV, Finkelstein AV (1999) *Proc Natl Acad Sci USA* 96:11299–11304.
5. Baker D (2000) *Nature* 405:39–42.
6. Bryngelson JD, Onuchic JN, Socci ND, Wolynes PG (1995) *Proteins* 21:167–195.
7. Dill KA, Chan HS (1997) *Nat Struct Biol* 4:10–19.
8. Dobson CM, Sali A, Karplus M (1998) *Angew Chem Int Ed Engl* 37:868–893.
9. Kubelka J, Hofrichter J, Eaton WA (2004) *Curr Opin Struct Biol* 14:76–88.
10. Talaga DS, Lau WL, Roder H, Tang J, Jia Y, DeGrado WF, Hochstrasser RM (2000) *Proc Natl Acad Sci USA* 97:13021–13026.
11. Haas E, Steinberg IZ (1984) *Biophys J* 46:429–437.
12. Wang ZS, Makarov DE (2003) *J Phys Chem B* 107:5617–5622.
13. Perli D, Welker C, Schindler T, Schröder K, Marahiel MA, Jaenicke R, Schmid FX (1998) *Nat Struct Biol* 5:229–235.
14. Wassenberg D, Welker C, Jaenicke R (1999) *J Mol Biol* 289:187–193.
15. Schuler B, Lipman EA, Eaton WA (2002) *Nature* 419:743–747.
16. Schuler B, Kremer W, Kalbitzer HR, Jaenicke R (2002) *Biochemistry* 41:11670–11680.
17. Lipman EA, Schuler B, Bakajin O, Eaton WA (2003) *Science* 301:1233–1235.
18. Deniz AA, Laurence TA, Dahhan M, Chemla DS, Schultz PG, Weiss S (2001) *Annu Rev Phys Chem* 52:233–253.
19. Schuler B, Lipman EA, Steinbach PJ, Kumke M, Eaton WA (2005) *Proc Natl Acad Sci USA* 102:2754–2759.
20. Hanbury Brown R, Twiss RQ (1956) *Nature* 177:27–29.
21. Berglund AJ, Doherty AC, Mabuchi H (2002) *Phys Rev Lett* 89:068101.
22. Mets Ü (2001) in *Fluorescence Correlation Spectroscopy*, eds Elson ES, Rigler R (Springer, Berlin), pp 346–359.
23. Reynaud S (1983) *Ann Phys* 8:315–370.
24. Ehrenberg M, Rigler R (1974) *Chem Phys* 4:390–401.
25. Widengren J, Mets Ü, Rigler R (1995) *J Phys Chem* 99:13368–13379.
26. Gopich IV, Szabo A (2006) *J Chem Phys* 124:154712.
27. Hoffmann A, Kane A, Nettels D, Hertzog D, Baumgärtel P, Lengefeld J, Reichardt G, Horsley DA, Seckler R, Bakajin O, Schuler B (2007) *Proc Natl Acad Sci USA* 104:105–110.
28. Karplus M, McCammon JA (1983) *Annu Rev Biochem* 52:263–300.
29. Palmer AG (2001) *Annu Rev Biophys Biomol Struct* 30:129–155.
30. Magg C, Schmid FX (2004) *J Mol Biol* 335:1309–1323.
31. Buscaglia M, Lapidus LJ, Eaton WA, Hofrichter J (2006) *Biophys J* 91:276–288.
32. Bieri O, Wirz J, Hellrung B, Schutkowski M, Drewello M, Kiefhaber T (1999) *Proc Natl Acad Sci USA* 96:9597–9601.
33. Lapidus LJ, Eaton WA, Hofrichter J (2000) *Proc Natl Acad Sci USA* 97:7220–7225.
34. Lapidus LJ, Steinbach PJ, Eaton WA, Szabo A, Hofrichter J (2002) *J Phys Chem B* 106:11628–11640.
35. Krieger F, Fierz B, Bieri O, Drewello M, Kiefhaber T (2003) *J Mol Biol* 332:265–274.
36. Hagen SJ, Qiu LL, Pabit SA (2005) *J Phys Cond Mat* 17:S1503–S1514.
37. Zwanzig R (1988) *Proc Natl Acad Sci USA* 85:2029–2030.
38. Bryngelson JD, Wolynes PG (1989) *J Phys Chem* 93:6902–6915.
39. Kramers HA (1940) *Physica* 7:284–304.
40. Socci ND, Onuchic JN, Wolynes PG (1996) *J Chem Phys* 104:5860–5868.
41. Klimov DK, Thirumalai D (1997) *Phys Rev Lett* 79:317–320.
42. Hagen SJ, Hofrichter J, Szabo A, Eaton WA (1996) *Proc Natl Acad Sci USA* 93:11615–11617.
43. Li MS, Klimov DK, Thirumalai D (2004) *Polymer* 45:573–579.
44. Camacho CJ, Thirumalai D (1993) *Proc Natl Acad Sci USA* 90:6369–6372.
45. Portman JJ, Takada S, Wolynes PG (2001) *J Chem Phys* 114:5082–5096.
46. Shea JE, Brooks CL (2001) *Annu Rev Phys Chem* 52:499–535.
47. Best RB, Hummer G (2006) *Phys Rev Lett* 96:228104.
48. Kubelka J, Chiu TK, Davies DR, Eaton WA, Hofrichter J (2006) *J Mol Biol* 359:546–553.
49. Yang WY, Gruebele M (2003) *Nature* 423:193–197.
50. Yang WY, Gruebele M (2004) *Biophys J* 87:596–608.
51. Haas E, Katchalskikatzir E, Steinberg IZ (1978) *Biopolymers* 17:11–31.
52. Möglich A, Joder K, Kiefhaber T (2006) *Proc Natl Acad Sci USA* 103:12394–12399.
53. Hudgins RR, Huang F, Gramlich G, Nau WM (2002) *J Am Chem Soc* 124:556–564.
54. Szabo A, Schulten K, Schulten Z (1980) *J Chem Phys* 72:4350–4357.
55. Hagen SJ (2003) *Proteins* 50:1–4.
56. Chandler D (1987) *Introduction to Modern Statistical Mechanics* (Oxford Univ Press, New York).
57. Kuzmenkina EV, Heyes CD, Nienhaus GU (2005) *Proc Natl Acad Sci USA* 102:15471–15476.
58. Chattopadhyay K, Elson EL, Frieden C (2005) *Proc Natl Acad Sci USA* 102:2385–2389.
59. Zimm BH (1956) *J Chem Phys* 24:269–278.
60. de Gennes PG (1985) *J Phys Lett* 46:L639–L642.
61. Doi M, Edwards SF (1988) *The Theory of Polymer Dynamics* (Oxford Univ Press, New York).
62. Thirumalai D (1995) *J Phys I* 5:1457–1467.
63. Sadqi M, Lapidus LJ, Munoz V (2003) *Proc Natl Acad Sci USA* 100:12117–12122.
64. Magg C, Kubelka J, Holtermann G, Haas E, Schmid FX (2006) *J Mol Biol* 360:1067–1080.
65. Kurtziefer C, Zarda P, Mayer S, Weinfurter H (2001) *J Mod Optic* 48:2039–2047.
66. Bradforth SE, Jinez R, Vanmourik F, Vangrondelle R, Fleming GR (1995) *J Phys Chem* 99:16179–16191.
67. De Schryver FC, Vosch T, Cotlet M, Van der Auweraer M, Müllen K, Hofkens J (2005) *Acc Chem Res* 38:514–522.
68. Hübner CG, Zumofen G, Renn A, Herrmann A, Müllen K, Basché T (2003) *Phys Rev Lett* 91:093903.
69. Neuweiler M, Doose S, Sauer M (2005) *Proc Natl Acad Sci USA* 102:16650–16655.
70. Neuweiler M, Löllmann M, Doose S, Sauer M (2007) *J Mol Biol* 365:856–869.
71. Merchant KA, Best RB, Louis JM, Gopich IV, Eaton WA (2007) *Proc Natl Acad Sci USA* 104:1528–1533.
72. Xu Y, Purkayastha P, Gai F (2006) *J Am Chem Soc* 128:15836–15842.