JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Single-Molecule Spectroscopy of Cold Denaturation and the Temperature-Induced Collapse of Unfolded Proteins

Mikayel Aznauryan, Daniel Nettels, Andrea Holla, Hagen Hofmann, and Benjamin Schuler*

Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland

Supporting Information

ABSTRACT: Recent Förster resonance energy transfer (FRET) experiments show that heat-unfolded states of proteins become more compact with increasing temperature. At the same time, NMR results indicate that cold-denatured proteins are more expanded than heat-denatured proteins. To clarify the connection between these observations, we investigated the unfolded state of yeast frataxin, whose cold denaturation occurs at temperatures above 273 K, with single-molecule FRET. This method allows the unfolded state dimensions to be probed not only in the cold- and heat-denatured range but also in between, i.e., in the presence of folded protein, and can thus be used to link the two regimes directly. The results show a continuous compaction of unfolded frataxin from 274 to 320 K, with a slight re-expansion at higher temperatures. Cold- and heat-denatured states are thus essentially two sides of the same coin, and their behavior can be understood within the framework of the overall temperature dependence of the unfolded state dimensions.

Proteins can unfold not only at high but also low temperatures, a behavior that has been quantified thoroughly in terms of the underlying thermodynamics of protein stability.[1] The molecular details of cold denaturation are still under debate, but the most common interpretation of the origin of this phenomenon is the decreasing strength of the hydrophobic effect with decreasing temperature, which is expected to be connected to changes in the structure and dynamics of the hydration layer around the polypeptide chain.[2−6] However, a detailed structural investigation of cold denaturation and possible differences in the structure of cold- and heat-denatured proteins has only become possible rather recently, with the discovery of proteins for which the cold-denatured state is populated above the melting point of water and is thus experimentally accessible without the addition of denaturants.[7−10] Based on NMR experiments, differences in residual secondary structure and the degree of protein hydration in cold- and heat-denatured states have been identified.[11−13] Another remarkable observation, based on NMR measurements on the C-terminal domain of protein L9 in the cold-denatured state, was a trend of increasing radii of hydration as the temperature was reduced.[7,9] This result suggests an interesting possible connection to recent single-molecule and ensemble Förster resonance energy transfer (FRET) experiments that demonstrated a continuous compaction of chemically, heat-, and pH-unfolded proteins with increasing temperature.[14,15] These possibly counterintuitive observations indicate the presence of interactions within the unfolded polypeptide chain that increase with temperature. However, the quantitative link between the NMR and the FRET experiments has remained unclear, especially whether the dimensions of the unfolded state populated upon cold denaturation deviate from the trend observed upon heat denaturation.[16]

Here, we address this question by investigating the temperature dependence of the cold- and heat-denatured states of yeast frataxin (Yfh1), which was chosen specifically as one of the very few available systems that would allow us to probe the properties of the unfolded state both under conditions of complete heat and almost complete cold denaturation (and in between) without extrapolation from high denaturant concentrations or extremes of pH. Even at the temperature of maximum conformational stability, a fraction of the population remains unfolded, which is an essential prerequisite for us to monitor the denatured state over the entire temperature range, from 274 to 332 K, and thus to directly link the behavior of cold- and heat-denatured protein. To probe the dimensions of unfolded frataxin, we use single-molecule FRET, which has been employed for measurements of intramolecular distances and distance distributions for a wide range of proteins denatured by chaotropes and heat, and for intrinsically disordered proteins.[14,17−25] As a key advantage, single-molecule FRET allows a clear separation of folded and unfolded subpopulations, and thus a quantitative analysis of the properties of the unfolded state without influence of the folded state signal, which is often difficult to achieve in ensemble experiments.[26]

For the single-molecule experiments, we created a double cysteine variant (C47S, N16C, S120C) of Yhf1. To quantify the effect of the Cys mutations on the conformational stability of Yhf1, circular dichroism (CD) experiments on the unlabeled protein were performed (Figure 1). The negative ellipticity is maximal at 289 K, corresponding to the temperature of highest conformational stability. The decreasing secondary structure content below and above 289 K correspond to cold and heat denaturation, respectively, as illustrated by the fraction of folded protein, $F_f(T)$, as a function of temperature T (Figure 1C). The data were fitted with $F_f(T) = (1 - \exp(-\Delta G_U/RT))^{-1}$, where the free energy of unfolding, $\Delta G_U(T)$, is parametrized in terms of T_S, the temperature where $\Delta G_U(T)$ is maximal; ΔH_S, the unfolding enthalpy at that temperature; and

Received: July 9, 2013
Published: September 8, 2013

   dx.doi.org/10.1021/ja407009w | J. Am. Chem. Soc. 2013, 135, 14040−14043

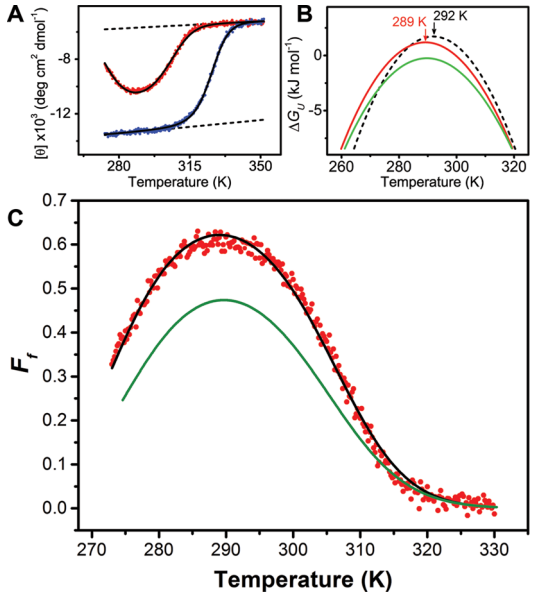

Figure 1. (A) Temperature dependence of the conformational stability of Yfh1 monitored by changes in ellipticity at 222 nm (red dots). The baseline for the unfolded subpopulation (upper dashed line) was obtained by linear extrapolation of the high temperature data. For obtaining a baseline for the folded subpopulation (lower dashed line), the protein was stabilized by adding 0.3 M sodium sulfate (blue dots). The black solid lines show fits with ΔH_S, ΔC_p, and T_S as fit parameters (see SI for details). (B) Temperature dependences of the unfolding free energy, ΔG_U, for the double-Cys variant of Yhf1 based on the thermodynamic parameters derived from CD (red line) and for the labeled protein as derived from single-molecule FRET (green line, Figure 2). The dashed black line represents the stability curve obtained by Pastore et al. for wild-type Yhf1 in 20 mM HEPES buffer.[8] (C) Temperature dependence of the fraction of folded (F_f) Yfh1 (red dots) as obtained from the data in (A), with a fit using eqs S14 and S15 (SI) shown as a solid line. The fit from Figure 3A is shown as a green line for direct comparison with the FRET results.

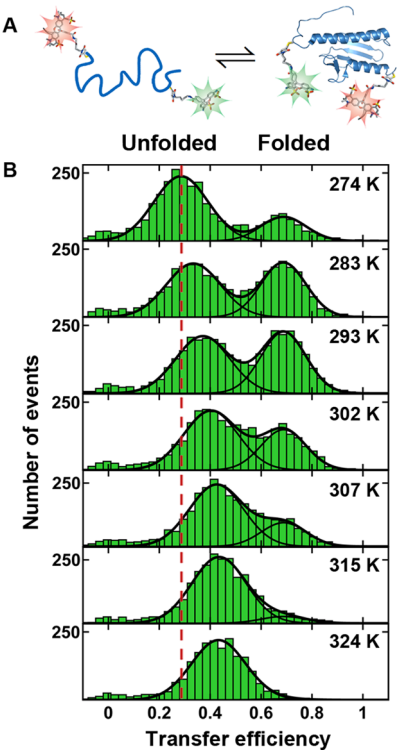

Figure 2. (A) Cartoons of unfolded and folded Yfh1 (based on PDB ID 2GA5) with Alexa Fluor 488 (green) and Alexa Fluor 594 (red) dyes at positions N16C and S120C. (B) Representative examples of single-molecule FRET efficiency histograms of Yfh1 show heat and cold denaturation and a temperature-induced collapse of the unfolded protein. The peak at E = 0.7 corresponds to folded and the peak at lower transfer efficiency to unfolded molecules. To determine mean transfer efficiencies, peaks were fit with Gaussian distributions (black lines); the width of the unfolded-state peak and the position and width of the folded-state peak were used as global fit parameters and are thus identical for all histograms. The red dashed line indicates the position of the unfolded peak at 274 K.

ΔC_p, the change in heat capacity upon unfolding (see SI for details). The $\Delta G_U(T)$ dependence calculated from the resulting thermodynamic parameters is close to the dependence found by Pastore et al.[8,27] for wild-type Yhf1 (Figure 1B, Table S1).

The protein was then labeled with Alexa Fluor 488 and Alexa Fluor 594 as FRET donor and acceptor, respectively (Figure 2A). Confocal single-molecule fluorescence spectroscopy on molecules freely diffusing in solution was used to quantify the equilibrium between folded and unfolded states and to separate the corresponding subpopulations. Bursts of fluorescence photons, detected while single frataxin molecules diffused through the laser focus of the confocal instrument, were used to obtain transfer efficiency histograms between 274 and 332 K. Examples are shown in Figure 2B. To eliminate the contribution from the molecules lacking an active acceptor at a transfer efficiency of $E \approx 0$, we used pulsed interleaved excitation[28] (see SI for details). We observe two peaks, as expected for cooperative cold- and heat-unfolding transitions that can be well approximated by a two-state process.[9] The peak at a transfer efficiency of E = 0.7 corresponds to the subpopulation of folded proteins. The peak centered at transfer efficiencies between E = 0.23 and E = 0.43 corresponds to the subpopulation of unfolded protein. The transfer efficiency histograms illustrate directly that the unfolded state is present over the entire temperature range, while the relative populations of folded and unfolded subpopulations vary.

From the transfer efficiency histograms, the fractions of folded protein molecules are obtained from the relative peak areas. The resulting temperature dependence (Figure 3A) shows that the folded state is maximally populated at about 290 K. Both with decreasing and increasing temperature, we observe the decrease in folded population and concomitant increase in the unfolded population expected for cold and heat denaturation. The good agreement of the corresponding stability curve with the results from unlabeled protein (Figure 1B,C) indicates that the thermodynamic properties of the sample are only marginally affected by FRET labeling.

The separation of native and unfolded subpopulations in the single-molecule experiments allows us to investigate not only the effect of temperature on the equilibrium between the two thermodynamic states, but, more importantly, it allows us to quantify changes in transfer efficiency for the unfolded subpopulation over the entire temperature range, even in the presence of folded molecules. The change in mean transfer efficiency of the unfolded state with temperature is apparent in Figure 2, indicating a change in the average distance between the fluorophores. To analyze the single-molecule data in terms of the dimensions of the unfolded state, we converted the mean transfer efficiencies obtained from the corresponding peaks to

dx.doi.org/10.1021/ja407009w | J. Am. Chem. Soc. 2013, 135, 14040−14043

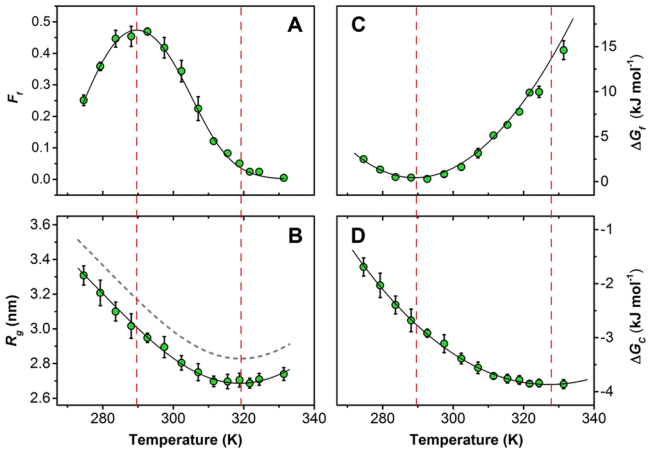

Figure 3. (A) Relative fraction of folded Yhf1 (F_f) calculated as a ratio of the area of the "folded" peak and the sum of the areas of "folded" and "unfolded" peaks with a fit as in Figure 1C (black line). (B) Temperature dependence of the measured radius of gyration (R_g) of unfolded Yfh1 (circles). The R_g values correspond to the interdye segment of the protein chain including the two dye linkers (each estimated to be equivalent to 4.5 peptide bonds[21]). The solid line represents a fit of R_g as a function of temperature (see details in SI). For comparison, we also show the fit of the radii of gyration rescaled to the full-length sequence (gray dashed line, see SI). (C) Temperature dependence of the free energy of folding (ΔG_f) of labeled Yhf1 calculated from the data in (A). (D) Temperature dependence of the free energy of collapse (ΔG_C) of unfolded Yhf1, calculated with respect to the excluded volume limit, where $\varepsilon = 0$ (see SI). Black solid lines in (C) and (D) are fits to the data assuming a constant heat capacity. Error bars indicate standard deviations estimated from three independent measurements.

mean radii of gyration based on polymer physical distance distributions[14,19,21,23,24,29] (see SI for details). Here we use the coil-to-globule transition theory of Sanchez,[19,24,29,30] but the results are robust with respect to the details of the shape of the distance distribution used (Figure S4).

Figure 3B shows the resulting values of the radius of gyration, R_g, of the denatured state as a function of temperature over a broad range, bridging the regimes of cold and heat denaturation. These data show a decrease of R_g of the unfolded state of Yhf1 from 273 to ∼320 K, corresponding to a continuous collapse of the unfolded chain with increasing temperature, followed by a slight re-expansion at temperatures above 320 K (Figure 3B). The temperature regime above 335 K is currently not accessible experimentally because of a rapid loss of signal, presumably due to accelerated photodestruction of the fluorophores. For a quantitative comparison of the intramolecular energetics relevant for folding and unfolded state collapse, respectively, the protein stability curve, $\Delta G_f(T)$, and the free energy of collapse, $\Delta G_C(T)$, as obtained from Sanchez theory (see SI for details), are also shown in Figure 3C,D.

The temperature dependence of R_g observed here for the denatured state illustrates an important point: the expansion of the cold-denatured state with decreasing temperature, as observed in NMR experiments,[7,9] and the collapse of unfolded proteins with increasing temperature detected by single-molecule spectroscopy[14] are essentially two sides of the same coin—an overall continuous change in denatured state dimensions with temperature. This observation is in line with the notion of the thermodynamic equivalence of cold- and heat-denatured states,[1] but at the same time demonstrates the non-cooperative shift of the distance distribution in the unfolded state with temperature. This conclusion is further corroborated by experiments with a second variant of frataxin with the FRET dyes in positions that allow us to investigate the unfolded state dimensions without overlap with the folded-state peak, which shows the same behavior (Figure S5). The similarity to the temperature dependence of heat-unfolded Csp*Tm*[16] suggests that what we observe is a rather generic behavior of denatured states, independent of whether they are populated by heating or cooling.

An interesting observation is the non-monotonic dependence of the R_g of unfolded frataxin on temperature. Remarkably, such behavior with a minimum in $R_g(T)$ has been observed even in simulations of hydrophobic homopolymers[3,31−34] and simple heteropolymers[35] in explicit water. In such simulations, the most compact ensemble has typically been assumed to resemble the folded state of a protein, and the expansions at low and high temperatures have been taken to correspond to cold and heat unfolding, respectively. In contrast, our results illustrate that the overall energetics of protein folding on the one hand and unfolded state collapse on the other are quite different (Figure 3): while $\Delta G_f(T)$ has a minimum at ∼290 K, where the native state of frataxin is most stable, the free energy of collapse, $\Delta G_C(T)$, has a minimum at ∼328 K, where the denatured state is most compact. However, $\Delta G_f(T)$ and $\Delta G_C(T)$ share the non-monotonic temperature dependence usually taken as an indication for the involvement of hydrophobic hydration.[2,36] Given the absence of specific interactions in simple polymer models,[3,31−35] the compact state observed in corresponding simulations is thus likely to resemble the compact unfolded state observed here rather than a folded state.

Recent NMR chemical shift experiments confirm the importance of hydration for the properties of the unfolded state and indicate pronounced differences in the degree of hydrogen bonding of backbone amides with water in cold- and heat-denatured yeast frataxin,[13] a process that is probably closely linked to slight changes in secondary structure content,[13] as previously suggested for other unfolded proteins.[14,37] Similarly, molecular dynamics simulations of the temperature dependence of unfolded state dimensions are very sensitive to the water model used, indicating a crucial role of hydration.[23] Both polymer expansion at low temperature and cold denaturation have been rationalized in terms of effects that favor the formation of a hydrated denatured form over a compact state at low temperature. Mechanistic explanations include a reduced probability of cavity formation in the solvent required for local dewetting and concomitant formation of a hydrophobic core;[38] lower hydrogen bond energy for hydration shell water compared to bulk water at low temperature;[3] or, closely connected, a reduction in water entropy at low temperature.[39]

But why is the minimum of $\Delta G_f(T)$ shifted by almost 40 K relative to $\Delta G_C(T)$? $\Delta G_f(T)$ consists of at least three contributions: the difference in internal energy of folded and unfolded states, ΔU_f; the difference in solvation free energy of folded and unfolded states, ΔG_s; and a term due to the difference in configurational entropy of folded and unfolded states,[40] $T\Delta S_{conf}$. ΔU_f may be expected to exhibit little temperature dependence, but we expect a strong effect of $T\Delta S_{conf}$ on the position of the minimum of $\Delta G_f(T)$, since $|\Delta S_{conf}|$ is much larger for folding than for unfolded state

dx.doi.org/10.1021/ja407009w | J. Am. Chem. Soc. 2013, 135, 14040−14043

collapse. Additional contributions to the difference between $\Delta G_f(T)$ and $\Delta G_C(T)$ may come from the change in the temperature dependence of the excess chemical potential of cavity formation in water with the size of the solute in the crossover region of ∼1 nm, as suggested by theory and simulation:[41] the less specific and transiently formed hydrophobic clusters present in a denatured protein are expected to be relatively small, and their hydration may thus exhibit a temperature dependence different from the formation of the folded state, which is larger and has a different surface composition. The recent emergence of force fields and water models that provide a more realistic description of unfolded and disordered proteins[42,43] may allow a more detailed identification of the molecular processes underlying cold denaturation, whose close interrelation with the continuous temperature-induced unfolded-state collapse is demonstrated here. In summary, our work thus connects previously disconnected observations on the unfolded state populated either through cold or heat denaturation by understanding its behavior within the framework of the temperature-dependent dimensions of a single denatured state.

■ ASSOCIATED CONTENT

Supporting Information

Experimental procedures, data analysis, and single-molecule data for a second variant of frataxin. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

schuler@bioc.uzh.ch

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Robert Best and Jeetain Mittal for discussions regarding hydrophobic hydration and conformational entropy, and Andrea Soranno for help with data analysis. This work was supported by the Swiss National Science Foundation and by the Forschungskredit of the University of Zurich.

■ REFERENCES

(1) Privalov, P. L. *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25*, 281.
(2) Privalov, P. L.; Gill, S. J. *Adv. Protein Chem.* **1988**, *39*, 191.
(3) Dias, C. L.; Ala-Nissila, T.; Karttunen, M.; Vattulainen, I.; Grant, M. *Phys. Rev. Lett.* **2008**, *100*, No. 118101.
(4) Yoshidome, T.; Kinoshita, M. *Phys. Rev. E* **2009**, *79*, No. 030905(R).
(5) Graziano, G. *Phys. Chem. Chem. Phys.* **2010**, *12*, 14245.
(6) Dias, C. L. *Phys. Rev. Lett.* **2012**, *109*, No. 048104.
(7) Li, Y.; Shan, B.; Raleigh, D. P. *J. Mol. Biol.* **2007**, *368*, 256.
(8) Pastore, A.; Martin, S. R.; Politou, A.; Kondapalli, K. C.; Stemmler, T.; Temussi, P. A. *J. Am. Chem. Soc.* **2007**, *129*, 5374.
(9) Luan, B.; Shan, B.; Baiz, C.; Tokmakoff, A.; Raleigh, D. P. *Biochemistry* **2013**, *52*, 2402.
(10) Buchner, G. S.; Shih, N.; Reece, A. E.; Niebling, S.; Kubelka, J. *Biochemistry* **2012**, *51*, 6496.
(11) Adrover, M.; Esposito, V.; Martorell, G.; Pastore, A.; Temussi, P. A. *J. Am. Chem. Soc.* **2010**, *132*, 16240.
(12) Shan, B.; McClendon, S.; Rospigliosi, C.; Eliezer, D.; Raleigh, D. P. *J. Am. Chem. Soc.* **2010**, *132*, 4669.
(13) Adrover, M.; Martorell, G.; Martin, S. R.; Urosev, D.; Konarev, P. V.; Svergun, D. I.; Daura, X.; Temussi, P.; Pastore, A. *J. Mol. Biol.* **2012**, *417*, 413.
(14) Nettels, D.; Muller-Spath, S.; Kuster, F.; Hofmann, H.; Haenni, D.; Ruegger, S.; Reymond, L.; Hoffmann, A.; Kubelka, J.; Heinz, B.; Gast, K.; Best, R. B.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 20740.
(15) Sadqi, M.; Lapidus, L. J.; Munoz, V. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12117.
(16) Nettels, D.; Müller-Späth, S.; Küster, F.; Hofmann, H.; Haenni, D.; Rüegger, S.; Reymond, L.; Hoffmann, A.; Kubelka, J.; Heinz, B.; Gast, K.; Best, R. B.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 20740.
(17) Deniz, A. A.; Laurence, T. A.; Beligere, G. S.; Dahan, M.; Martin, A. B.; Chemla, D. S.; Dawson, P. E.; Schultz, P. G.; Weiss, S. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5179.
(18) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419*, 743.
(19) Sherman, E.; Haran, G. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 11539.
(20) Michalet, X.; Weiss, S.; Jager, M. *Chem. Rev.* **2006**, *106*, 1785.
(21) Hoffmann, A.; Kane, A.; Nettels, D.; Hertzog, D. E.; Baumgartel, P.; Lengefeld, J.; Reichardt, G.; Horsley, D. A.; Seckler, R.; Bakajin, O.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 105.
(22) Merchant, K. A.; Best, R. B.; Louis, J. M.; Gopich, I. V.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 1528.
(23) Muller-Spath, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Ruegger, S.; Reymond, L.; Nettels, D.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 14609.
(24) Hofmann, H.; Soranno, A.; Borgia, A.; Gast, K.; Nettels, D.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 16155.
(25) Soranno, A.; Buchli, B.; Nettels, D.; Cheng, R. R.; Muller-Spath, S.; Pfeil, S. H.; Hoffmann, A.; Lipman, E. A.; Makarov, D. E.; Schuler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2012**, *109*, 17800.
(26) Schuler, B.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2008**, *18*, 16.
(27) Martin, S. R.; Esposito, V.; Rios, P. D. L.; Pastore, A.; Temussi, P. A. *J. Am. Chem. Soc.* **2008**, *130*, 9963.
(28) Muller, B. K.; Zaychikov, E.; Brauchle, C.; Lamb, D. C. *Biophys. J.* **2005**, *89*, 3508.
(29) Ziv, G.; Haran, G. *J. Am. Chem. Soc.* **2009**, *131*, 2942.
(30) Sanchez, I. C. *Macromolecules* **1979**, *12*, 980.
(31) Paschek, D.; Nonn, S.; Geiger, A. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2780.
(32) De los Rios, P.; Caldarelli, G. *Phys. Rev. E* **2001**, *63*, No. 031802.
(33) Das, P.; Matysiak, S. *J. Phys. Chem. B* **2012**, *116*, 5342.
(34) Matysiak, S.; Debenedetti, P. G.; Rossky, P. J. *J. Phys. Chem. B* **2012**, *116*, 8095.
(35) Jamadagni, S. N.; Bosoy, C.; Garde, S. *J. Phys. Chem. B* **2010**, *114*, 13282.
(36) Southall, N. T.; Dill, K. A.; Haymet, A. D. J. *J. Phys. Chem. B* **2002**, *106*, 521.
(37) Yang, W. Y.; Larios, E.; Gruebele, M. *J. Am. Chem. Soc.* **2003**, *125*, 16220.
(38) ten Wolde, P. R.; Chandler, D. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 6539.
(39) Oshima, H.; Yoshidome, T.; Amano, K.; Kinoshita, M. *J. Chem. Phys.* **2009**, *131*, No. 205102.
(40) Thompson, J. B.; Hansma, H. G.; Hansma, P. K.; Plaxco, K. W. *J. Mol. Biol.* **2002**, *322*, 645.
(41) Lum, K.; Chandler, D.; Weeks, J. D. *J. Phys. Chem. B* **1999**, *103*, 4570.
(42) Best, R. B.; Mittal, J. *J. Phys. Chem. B* **2010**, *114*, 14916.
(43) Best, R. B.; Hummer, G. *J. Phys. Chem. B* **2009**, *113*, 9004.

Supporting Information

Single-Molecule Spectroscopy of Cold Denaturation and the Temperature-Induced Collapse of Unfolded Proteins

Mikayel Aznauryan, Daniel Nettels, Andrea Holla, Hagen Hofmann, and Benjamin Schuler*

University of Zurich, Department of Biochemistry, Winterthurerstrasse 190,
8057 Zurich, Switzerland

Corresponding Author
***** Benjamin Schuler
Email: schuler@bioc.uzh.ch

Preparation and Labeling of Protein

The coding sequence for Yhf1, comprising residues 1-123 of the mature gene product (Swiss-Prot entry: Q07540) with V1M and S120C, was codon optimized for expression in *E.* coli and cloned into pET47b(+) using the SmaI and PacI sites. On the resulting expression plasmid pFra1, frataxin is encoded as a fusion protein with an N-terminal His_6-tag and recognition site for HRV 3C protease. Gene synthesis and construction of the expression plasmid for recombinant Yhf1 was carried out by EZBiolab (Carmel, USA). Based on the pFra1 expression plasmid, we constructed a variant with C47S, N16C and S120C. We used the following procedure for the expression and purification of both Yfh1-16/120 and Yfh1-47/120 variants. His_6-frataxin was produced in *E. coli* BL21(DE3) harboring pFra1. 1 L cultures were grown in LB media for 4 hours at 37°C and then induced by 0.5 mM IPTG for 16 hours at 20°C. After harvesting by centrifugation, the cells were resuspended in Lysis Buffer (20 mM Tris-HCl, pH 8.0, 250 mM NaCl, 20 mM imidazole, 1 mM PMSF) and disrupted by sonication on ice. His_6-frataxin was purified from the soluble fraction of the lysate using a 5 ml HisTrap column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Bound protein was eluted with a linear gradient of 20 to 500 mM imidazole in 20 mM Tris-HCl, pH 7.4, 150 mM NaCl after washing the column with 20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 20 mM imidazole. Purified His_6-frataxin was detected in wash and elution fractions. Pooled wash fractions were subjected to digestion with HRV 3C protease at 4 °C for 16 h at a Yhf1 to protease molar ratio of about 100 to 1. After addition of β-mercaptoethanol to a final concentration of 1 mM, His_6-tagged HRV 3C protease (produced in-house), His_6-tag containing peptide, and uncut protein were separated from frataxin using a 5 ml HisTrap column. Yhf1 in the flow-through was denatured in 4 M guanidine, reduced with 50 mM DTT, concentrated and subjected to gel filtration using a Superdex 75 10/300 GL column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) in 100 mM NaH_2PO_4, pH 7.1, 6 M guanidinium chloride. Yhf1-containing fractions were labeled with Alexa Fluor 488 maleimide (Molecular Probes, Eugene, USA) at a dye:protein molar ratio of 0.7:1 for 2 h at room temperature and subsequently with Alexa Fluor 594 maleimide at a dye:protein molar ratio of 2:1 overnight at 4 °C. Unreacted dye was removed by gel filtration, again using a Superdex 75 10/300 GL column in 100 mM NaH_2PO_4, pH 7.1, 6 M guanidinium chloride. The resulting sample contained recombinant Yhf1 labeled with Alexa Fluor 488 and 594, and frataxin doubly-labeled with either Alexa Fluor 488 or Alexa Fluor 594, as shown by mass spectrometry.

Single-Molecule Fluorescence Spectroscopy

Observations of single molecule fluorescence were made using a MicroTime 200 confocal microscope equipped with a HydraHarp 400 counting module (Picoquant, Germany) and an Olympus UplanApo 60×/1.20W objective. The excitation of the dyes was achieved using pulsed interleaved excitation (PIE)[1]. The wavelength range used for acceptor excitation was selected with a z582/15 band pass filter (Chroma) from the emission of an SC-450-4 supercontinuum fiber laser (Fianium, UK) driven at 20 MHz, which triggers the diode laser used for donor excitation (LDH-D-C-485, PicoQuant, Germany). Emitted photons were collected through the microscope objective, focused onto a 100 μm pinhole, and then separated into four channels with a polarizing beam splitter and two dichroic mirrors

(585DCXR, Chroma). Photons were additionally filtered by bandpass filters (ET525/50M and HQ650/100, Chroma) before being focused onto one of four single photon avalanche detectors (SPADs; Optoelectronics SPCMAQR-15, PerkinElmer, Wellesley, MA). Samples of labeled protein were diluted to a concentration of about 20 pM in 50 mM sodium phosphate buffer, pH 7.0. Tween 20 (0.001%; Pierce) was added to prevent surface adhesion of the protein. To reduce data acquisition time and thus minimize damage to the chromophores at high temperatures, the photoprotective additive β-mercaptoethanol (143 mM) was added. Measurements were performed by placing the confocal spot at a position 50 μm deep in the solution with a total acquisition time of 30 minutes.

For the identification of photon bursts caused by proteins diffusing through the confocal volume, we used only the photons detected after donor excitation, i.e. detected in time intervals of ~25 ns following pulses from the diode laser. Successive photons separated by inter photon times <100 μs were combined into one burst with n_D and n_A photons counted in the donor and acceptor detection channels, respectively. These numbers were corrected for background. Transfer efficiency histograms were obtained from bursts with $n'_A + n_D > 80$ according to:

$$E = \frac{n'_A}{n'_A + \gamma n_D} \qquad \textbf{(S1)}$$

$$\text{with } n'_A = n_A - \frac{\alpha}{1+\alpha}\left(n_A + \gamma n_D\right) - \beta n_D$$

where $\alpha = 0.05$ is the ratio of the extinction coefficients of Alexa 594 and Alexa 488 at the laser donor excitation wavelength ($\lambda_{ex,D} = 482$ nm), β is the ratio of donor photons detected in the acceptor channel (crosstalk), and $\gamma = \eta_A Q_A / \eta_D Q_D$ is the ratio of the products of quantum yields and detection efficiencies for acceptor and donor. We obtained values for α, β, and γ at room temperature (T=295 K) as described previously[2]. Whereas α and β are virtually independent of temperature in the range investigated here, we observed a significant decrease of molecular brightness of the fluorophores with increasing temperature. From the additional observation that also the fluorescence lifetimes of the dyes decreases with increasing temperature, we conclude that the effect of brightness decrease is well explained by a temperature dependent increase of collisional quenching of the dyes. Hence, the changes in fluorescence lifetimes can be used to correct $\gamma = \gamma(T)$ according to:

$$\gamma(T) = \gamma_0 \frac{\tau_{D0} \cdot \tau_A(T)}{\tau_{A0} \cdot \tau_D(T)} \qquad \textbf{(S2)}$$

where $\gamma_0 = \gamma(295\text{ K})$ is the value determined at room temperature. $\tau_D(T)$ and $\tau_A(T)$ are the donor and acceptor fluorescence lifetimes at temperature T, and τ_{D0} and τ_{A0} are the corresponding lifetimes at T=295 K. The lifetimes were derived from the same single-molecule photon data from which the transfer efficiency histograms were obtained. The mean donor and acceptor fluorescence lifetimes were obtained from the burst photons' arrival times

measured with respect to the corresponding donor or acceptor excitation pulses. For determining the donor lifetime, only photons of bursts with E<0.1, emitted by molecules lacking an active acceptor were used[3]. The resulting temperature dependence of the correction factor γ is shown in Figure S1.

Similarly, the temperature dependence of the Förster radius $R_0(T)$ was derived from the one at room temperature $R_0(295\ \mathrm{K})$ (5.4 nm for Alexa Fluor 488 and Alexa Fluor 594 dye pair) according to:

$$R_0^6(T) = R_0^6(295\ \mathrm{K})\frac{\tau_D(T)}{\tau_{D0}}\left(\frac{n(295\ \mathrm{K})}{n(T)}\right)^4 \qquad \textbf{(S3)}$$

where $n(T)$ is the refractive index of the sample solution at temperature T, which was measured for each temperature with a digital Abbe refractometer (Krüss, Germany). For each burst, we also determined the stoichiometry ratio $S = \frac{n_D + n_A}{n_D + n_A + n_{AA}}$, where n_{AA} is the number of photons detected after acceptor excitation, i.e. detected in the 25 ns time intervals following the 582-nm pulses. Figure S2A shows the 2D histogram of stoichiometry ratio and transfer efficiency values of the bursts detected at room temperature (T=295 K). The histogram is dominated by a peak corresponding to molecules lacking an active acceptor (E=0 and S=1). For building the transfer efficiency histograms (Fig. 2 in main text), we only used bursts with S<0.7. In Figure S2B, we compare such a histogram (green) with a transfer efficiency histogram obtained from all bursts (grey), including those with S>0.7.

A custom-built temperature-controlled sample holder employing Peltier elements and a digital temperature controller (TED 4015, Thorlabs) with a Pt100 temperature sensor was used[4]. The thermal contact of the cover slide to the objective via the immersion water results in a temperature gradient in the sample. To keep the gradient small, the microscope objective was also temperature-regulated using an aluminum collar to which a water-cooled Peltier element was mounted. Below 313 K, the sample holder and the objective were heated or cooled to the same temperature. Only the holder could be heated to higher temperatures while we kept the objective at 313 K to avoid damage. The temperature at the confocal volume was calibrated using the temperature-dependent fluorescence lifetime of rhodamine B[4,5]. The resulting temperature calibration is shown in Figure S3.

Determination of the radius of gyration from mean transfer efficiencies

The calculation of the radius of gyration of unfolded protein chains from the measured mean transfer efficiencies was performed assuming two different inter-dye distance distributions: the Gaussian chain distribution[3,4,6] and a distribution obtained from a variation of coil-to-globule transition theory of Sanchez[7-9]. Experimentally determined mean transfer efficiencies of the unfolded state can be calculated from a dye-to-dye distance distribution $P(r)$ according to:

$$\langle E\rangle = \int_0^L E(r)P(r)dr \Big/ \int_0^L P(r)dr \qquad \textbf{(S4)}$$

with

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

The distribution for a Gaussian chain is:

$$P_{Gaussian}(r) = 4\pi r^2 \left(\frac{3}{2\pi\langle r^2\rangle}\right)^{3/2} \exp\left(-\frac{3r^2}{2\langle r^2\rangle}\right) \qquad \textbf{(S5)}$$

Given a value of $\langle E\rangle$, the mean-squared end-to-end distance $\langle r^2\rangle$ can be calculated numerically, which can be converted to the radius of gyration by $R_g = \sqrt{\langle r^2\rangle/6}$ [10].

The theory of Sanchez provides an expression for the probability density function of the radius of gyration, r_g, of the probed segment as a Boltzmann-weighted Flory-Fisk distribution:

$$P_{Sanchez}\left(r_g, \varepsilon, R_{g\Theta}\right) = Z^{-1} r_g^6 \exp\left\{-\frac{7r_g^2}{2R_{g\Theta}^2} + n\left[\frac{1}{2}\varepsilon\phi - \frac{1-\phi}{\phi}\ln\left(1-\phi\right)\right]\right\} \qquad \textbf{(S6)}$$

Here, $n = N+1$ is the number of amino acids (monomers) in the chain segment probed, N is the number of peptide bonds, ε is the mean-field interaction free energy between amino acids, $\phi = \left(R_C/r_g\right)^3$ is the volume fraction of the chain with respect to the most compact state of radius $R_C = \left[3n\mathrm{v}/4\pi\right]^{1/3}$, with the average volume of one amino acid being $\mathrm{v} = 0.13$ nm^3 as shown in[9], and Z is a normalization factor. A realistic value of the radius of gyration of the chain at the Θ-state, $R_{g\Theta}$, has recently been estimated using a generalized expression for the length-scaling of polymers derived by Hammouda[11]:

$$R_{g\Theta} = \sqrt{\frac{2l_p b}{\left(2\nu+1\right)\left(2\nu+2\right)}} N_{dd}^{\nu} \qquad \textbf{(S7)}$$

In Eq. S7, $b = 0.38$ nm is the distance between two C_α-atoms[12], $\nu = 1/2$, $l_p = 0.4$ nm is the persistence length* of unfolded proteins in good solvent extracted from SAXS[9,13], and $N_{dd} = N+9$ takes into account the length of the fluorophore linkers approximately[3,9,14,15].

To relate the distribution of radii of gyration $P(r_g, \varepsilon, R_{g\Theta})$ to a segment end-to-end distance distribution $P(r, \varepsilon, R_{g\Theta})$, which is required to describe the mean transfer

* Here we assume that l_p is to a good approximation independent of temperature. Calculations show that a variation of l_p by 25% result in a minor change of R_g (±0.03 nm).

efficiencies $\langle E \rangle$, we used the conditional probability density function $P(r|r_g)$ suggested by Ziv and Haran[16], which describes the distance distribution of two random points inside the sphere of radius $\delta \cdot r_g$:

$$P\left(r\middle|r_g\right)=\frac{1}{\delta\cdot r_g}\left[3\left(\frac{r}{\delta\cdot r_g}\right)^2-\frac{9}{4}\left(\frac{r}{\delta\cdot r_g}\right)^3+\frac{3}{16}\left(\frac{r}{\delta\cdot r_g}\right)^5\right] \qquad 0\le r<2\delta\cdot r_g \qquad \textbf{(S8)}$$

Here, δ was obtained from the condition that $6\langle R_g^2\rangle=\langle r^2\rangle$ at the Θ-state. The mean transfer efficiency is finally given by:

$$\langle E\rangle=\int_0^L E(r)P(r,\varepsilon,R_{g\Theta})dr=\int_0^L E(r)\int_{R_C}^{L/2}P(r|r_g)\,P(r_g,\varepsilon,R_{g\Theta})dr_g dr\,, \qquad \textbf{(S9)}$$

where R_0 is the Förster radius, and L is the contour length of the protein segment probed. Eq. S9 was solved for ε numerically for each measured temperature, and the mean radii of gyration R_g (dots in Fig. 3B and Fig. S5C) were finally determined according to:

$$R_g=\sqrt{\int_{R_C}^{L/2}r_g^2P(r_g,\varepsilon,R_{g\Theta})dr_g} \qquad \textbf{(S10)}$$

For fitting the R_g temperature dependencies, we substituted the mean-field interaction free energy in Eq. S10 by $\varepsilon(T)=\Delta h_0+\Delta c_p\left(T-T_0\right)-T\left(\Delta s_0+\Delta c_p\ln\left(T/T_0\right)\right)$, where Δh_0, Δs_0, and Δc_p are the changes of enthalpy, entropy, and heat capacity per monomer (amino acid residue) relative to the random coil state ($\varepsilon=0$); T_0 was chosen to be 295 K. From the fit to the data of Yhf1-16/120 we obtained $\Delta h_0=-7.0\pm0.1$ kJ mol^{-1}, $\Delta s_0=-31.5\pm1.1$ J mol^{-1} K^{-1}, and $\Delta c_p=402\pm8$ J mol^{-1} K^{-1}.

The resulting values of R_g correspond to the protein chain segment probed, including the two fluorophore linkers (N_{dd} = N+9). For obtaining the radius of gyration of the full-length protein, $R_{g,full}$ (dashed line in Figure 3B, main text), the relation $R_{g,full}=R_g\left(N_{full}/N_{dd}\right)^{\nu}$ was used, where the scaling exponents ν, corresponding to each probed temperature, were obtained from Eq. S7 using the calculated R_g values instead of $R_{g\Theta}$.

The free energy of collapse, ΔG_C, was calculated with respect to the chain at the excluded volume limit (with volume fraction ϕ_{ex}), where $\varepsilon=0$:

$$\Delta G_C=k_BT\left[\ln(P(\phi,\varepsilon))-\ln(P(\phi_{ex},0))\right] \qquad \textbf{(S11)},$$

where

$$P(\phi,\varepsilon) = Z^{-1}\left(\frac{\phi_0}{\phi}\right)^2 \exp\left\{-\frac{7}{2}\left(\frac{\phi_0}{\phi}\right)^{2/3} + n\left[\frac{\phi}{2}\varepsilon - \frac{1-\phi}{\phi}\ln(1-\phi)\right]\right\} \qquad \textbf{(S12)}$$

and $\phi_0 = \left(R_C / R_{g\Theta}\right)^3$. To obtain the values of ΔG_C for the full-length protein, we first calculated the interaction energy ε for the full-length protein using the scaled values of gyration radii, $R_{g,full}$. These ε values were then used to calculate the ΔG_C corresponding to the full-length protein. The volume fraction at the excluded volume limit (ϕ_{ex}) was obtained by numerically solving

$$-\frac{d\ln\left(P\left(\phi_{ex},\varepsilon\right)\right)}{d\phi} = 0 \qquad \textbf{(S13)}$$

for ϕ with $\varepsilon = 0$.

Circular Dichroism Spectroscopy

CD experiments were performed using a J-810 CD spectrometer (Jasco) equipped with temperature-controlled water bath programmed to raise the temperature at a rate of 1 K/min. Measurements were carried out in cylindrical quartz cells with 1 mm path length and at a protein concentration of 0.35 mg/ml. The unfolding curves of Yhf1 were obtained by monitoring the CD ellipticities at 222 nm as a function of temperature in 50 mM sodium phosphate buffer, pH 7.0. The baselines for folded and unfolded subpopulations were approximated by linear functions of temperature.

Fluorescence Spectroscopy

Ensemble fluorescence measurements were performed on a Fluorolog fluorimeter (Jobin Yvon) equipped with a peltier-controlled cell holder, which was used to raise the temperature at a rate of 1 K/min. The actual temperature of the sample was monitored using a temperature probe directly immersed in the cuvette. Measurements were carried out in quartz cuvette with 3 mm path length. The protein concentration was 1 μM; 0.001% Tween 20 were added to the solution to prevent surface adhesion of the protein. Temperature-induced unfolding of Yhf1 was monitored by the change in the tryptophan fluorescence intensity at 350 nm after excitation at 295 nm. Since the intrinsic Trp fluorescence quantum yield itself strongly depends on temperature, and this dependence is nonlinear, it is difficult to estimate the temperature dependences of the folded and unfolded baselines. To get an estimate of the temperature dependencies of the folded and unfolded states, we used *N*-acetyltryptophanamide (NATA), following the approach of Makhatadze and coworkers[17]. The temperature dependence of the fluorescence intensity change for 2 μM NATA (the same molar concentration as that of Trp residues in frataxin) was recorded under identical conditions. The changes of the fluorescence intensity of NATA with temperature are well described by a second-order polynomial. This polynomial was rescaled to define the temperature dependencies of the fluorescence intensity of the folded, $I_f(T)$, and unfolded, $I_u(T)$ Yhf1.

Thermodynamic analysis of unfolding curves

The unfolding curves of Yhf1 obtained both from CD and fluorescence measurements were fitted assuming a two-state system, for which the experimental observable, $Y(T)$, (CD signal at 222 nm or Trp fluorescence intensity at 350 nm) can be written as:

$$Y(T) = F_f(T) \cdot Y_f(T) + [1 - F_f(T)] \cdot Y_u(T) \qquad \textbf{(S14)}$$

where $Y_f(T)$ and $Y_u(T)$ are the baselines of folded and unfolded subpopulations, respectively. Folded-state baselines in CD and ensemble fluorescence measurements (Fig. 1A, Fig. S6C) were obtained by stabilization of the protein variants with 0.3 M sodium sulfate (pH 7.0), a commonly used kosmotrope for non-specifically stabilizing the folded state by preferential hydration[18]. Pastore and coworkers had previously shown that sodium sulfate strongly stabilizes frataxin[19], and we follow established procedures used e.g. by Shortle and coworkers[20], who employed ammonium sulfate for obtaining native state baselines, or Makhatadze and coworkers[17], who used changes in pH in an analogous way.

The fraction of folded protein $F_f(T)$ is given by:

$$F_f(T) = 1/(1 + \exp(-[\Delta H(T) \; - \; T\Delta S(T)]/RT)) \qquad \textbf{(S15)}$$

where the change in enthalpy $\Delta H(T)$ and entropy $\Delta S(T)$ given by:

$$\Delta H(T) = \Delta H_S \; + \; (T \; - \; T_S) \cdot \Delta C_p \qquad \text{and} \qquad \Delta S(T) = \Delta C_p \cdot \ln\left(\frac{T}{T_S}\right), \qquad \textbf{(S16)}$$

where ΔH_S is the enthalpy change at temperature T_S,[†] which is defined as the temperature at which the fraction of folded protein is maximal, i.e. where the Gibbs free energy reaches its maximum and the entropy change between folded and unfolded state is zero ($\Delta S(T_S) = 0$). ΔC_p is the change in heat capacity upon unfolding[17,21]. The fraction of folded protein was calculated from:

$$F_f = 1 - F_u = \frac{Y(T) - Y_u(T)}{Y_f(T) - Y_u(T)} \qquad \textbf{(S17)}$$

Table S1 summarizes the thermodynamic parameters derived from our CD and FRET measurements in comparison with the values for wild-type Yhf1 calculated using the data by Pastore et al.[22].

[†]The reason for using T_S (instead of the melting temperature, T_m) as a reference temperature is that in cases where the population of the folded state is less than 50% at all temperatures, T_m cannot be defined.

REFERENCES

(1) Muller, B. K.; Zaychikov, E.; Brauchle, C.; Lamb, D. C. *Biophys J* **2005**, *89*, 3508.

(2) Schuler, B. *Methods Mol Biol* **2007**, *350*, 115.

(3) Hoffmann, A.; Kane, A.; Nettels, D.; Hertzog, D. E.; Baumgartel, P.; Lengefeld, J.; Reichardt, G.; Horsley, D. A.; Seckler, R.; Bakajin, O.; Schuler, B. *P Natl Acad Sci USA* **2007**, *104*, 105.

(4) Nettels, D.; Muller-Spath, S.; Kuster, F.; Hofmann, H.; Haenni, D.; Ruegger, S.; Reymond, L.; Hoffmann, A.; Kubelka, J.; Heinz, B.; Gast, K.; Best, R. B.; Schuler, B. *P Natl Acad Sci USA* **2009**, *106*, 20740.

(5) Benninger, R. K. P.; Koc, Y.; Hofmann, O.; Requejo-Isidro, J.; Neil, M. A. A.; French, P. M. W.; deMello, A. J. *Anal Chem* **2006**, *78*, 2272.

(6) O'Brien, E. P.; Morrison, G.; Brooks, B. R.; Thirumalai, D. *J Chem Phys* **2009**, *130*, 124903.

(7) Sherman, E.; Haran, G. *P Natl Acad Sci USA* **2006**, *103*, 11539.

(8) Muller-Spath, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Ruegger, S.; Reymond, L.; Nettels, D.; Schuler, B. *P Natl Acad Sci USA* **2010**, *107*, 14609.

(9) Hofmann, H.; Soranno, A.; Borgia, A.; Gast, K.; Nettels, D.; Schuler, B. *P Natl Acad Sci USA* **2012**, *109*, 16155.

(10) Grosberg, A. Y.; Khokhlov, A. R. *Statistical Physics of Macromolecules*; AIP Press: Woodbury, NY, 1994.

(11) Hammouda, B. *Adv Polym Sci* **1993**, *106*, 87.

(12) Zhou, H. X. *Biochemistry* **2004**, *43*, 2141.

(13) Kohn, J. E.; Millett, I. S.; Jacob, J.; Zagrovic, B.; Dillon, T. M.; Cingel, N.; Dothager, R. S.; Seifert, S.; Thiyagarajan, P.; Sosnick, T. R.; Hasan, M. Z.; Pande, V. S.; Ruczinski, I.; Doniach, S.; Plaxco, K. W. *P Natl Acad Sci USA* **2004**, *101*, 12491.

(14) Schroder, G. F.; Alexiev, U.; Grubmuller, H. *Biophys J* **2005**, *89*, 3757.

(15) McCarney, E. R.; Werner, J. H.; Bernstein, S. L.; Ruczinski, I.; Makarov, D. E.; Goodwin, P. M.; Plaxco, K. W. *J Mol Biol* **2005**, *352*, 672.

(16) Ziv, G.; Haran, G. *J Am Chem Soc* **2009**, *131*, 2942.

(17) Richardson, J. M., 3rd; Lemaire, S. D.; Jacquot, J. P.; Makhatadze, G. I. *Biochemistry* **2000**, *39*, 11154.

(18) Timasheff, S. N. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 67.

(19) Adinolfi, S.; Nair, M.; Politou, A.; Bayer, E.; Martin, S.; Temussi, P.; Pastore, A. *Biochemistry* **2004**, *43*, 6511.

(20) Shortle, D.; Stites, W. E.; Meeker, A. K. *Biochemistry* **1990**, *29*, 8033.

(21) Ibarra-Molero, B.; Makhatadze, G. I.; Sanchez-Ruiz, J. M. *Biochim Biophys Acta* **1999**, *1429*, 384.

(22) Pastore, A.; Martin, S. R.; Politou, A.; Kondapalli, K. C.; Stemmler, T.; Temussi, P. A. *Journal of the American Chemical Society* **2007**, *129*, 5374.

(23) Sanchez, I. C. *Macromolecules* **1979**, *12*, 980.

(24) Flory, P. J. *Statistical mechanics of chain molecules*; Hanser Publishers: Munich, Vienna, New York, 1989.

(25) Chung, H. S.; Louis, J. M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 11837.

Table S1. Thermodynamic parameters of cold and heat unfolding of Yfh1-16/120

	CD measurement	Single-molecule FRET measurement	Data from Pastore et al.[22]
T_S (K)	289.2 ± 0.1	289.7 ± 0.3	291.9 ± 0.5
ΔH_S (kJ mol^{-1})	1.19 ± 0.02	- 0.25 ± 0.05	1.72 ± 0.08
ΔC_p (kJ mol^{-1} K^{-1})	6.28 ± 0.21	5.62 ± 0.19	7.57 ± 0.15

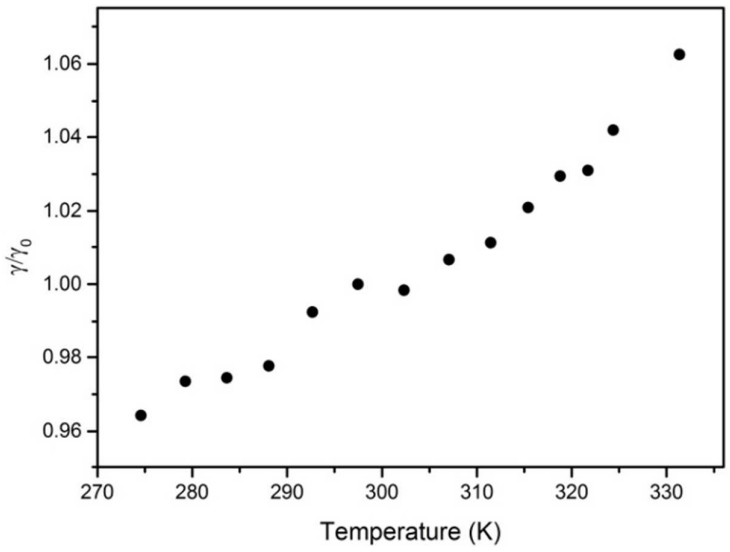

Figure S1. Temperature dependence of the correction factor γ.

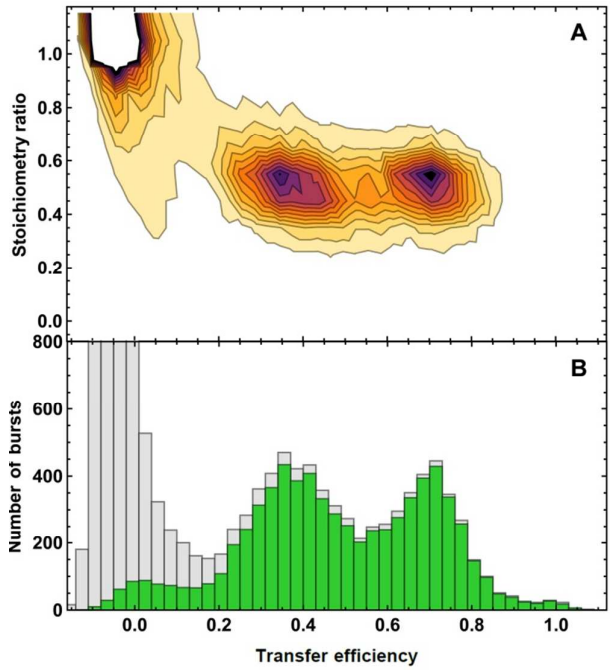

Figure S2. (A) 2D contour plot of stoichiometry ratio vs. transfer efficiency constructed using our data set for the measurement of Yfh1-16/120 at room temperature. The two peaks at 0.5 stoichiometry ratio originate from the protein molecules labeled with both active donor and acceptor fluorophores. The peak at high stoichiometry and $E \approx 0$ corresponds to molecules with an inactive acceptor fluorophore. (B) Transfer efficiency histograms of double-labeled Yhf1-16/120 at room temperature. The grey histogram was obtained using the dataset with the donor excitation only, whereas the green histogram was recorded using pulsed interleaved excitation of donor and acceptor fluorophores and selection of molecules with $S<0.7$.

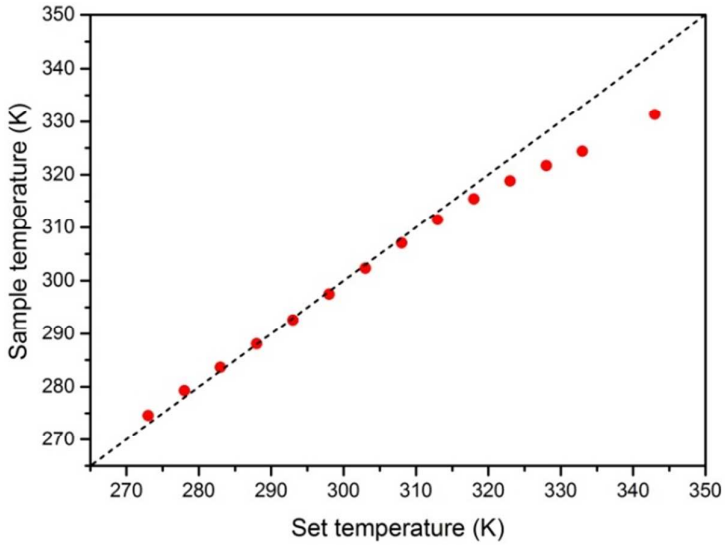

Figure S3. Temperature calibration showing the relation between the temperature of the sample holder and the sample temperature in the confocal volume as determined from rhodamine fluorescence lifetime measurements.

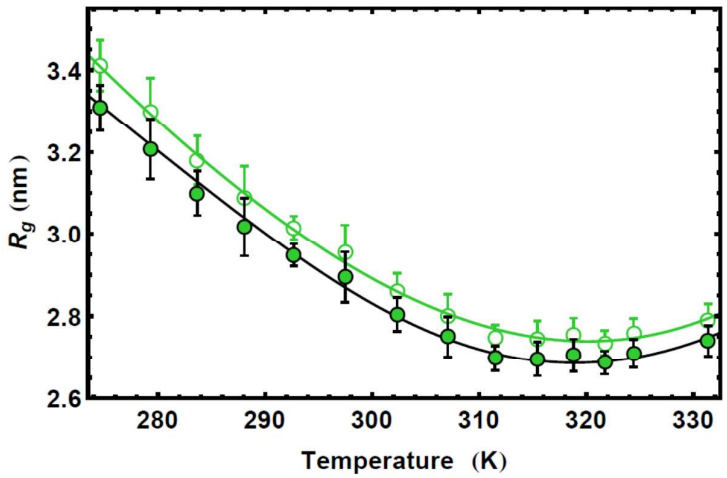

Figure S4. Comparison of the values for gyration radii for Yfh1-16/120 calculated according to the Sanchez model (filled green circles) and assuming a Gaussian chain distribution (open green circles), respectively. To test the robustness of our results, we analyzed the data with two different distribution, the Gaussian chain[3,6] distribution and a distribution obtained from a variant of coil-to-globule transition theory of Sanchez[7,9,16,23], which uses a Flory–Fisk distribution[24] weighted by a potential accounting for effective interactions within the chain. The results from the two types of analysis are very similar, suggesting that the chain dimensions derived from the FRET data under our conditions are robust with respect to the details of the shape of the distance distribution used. In the main text, we thus use the dimensions calculated using the Sanchez model, which is physically most realistic.

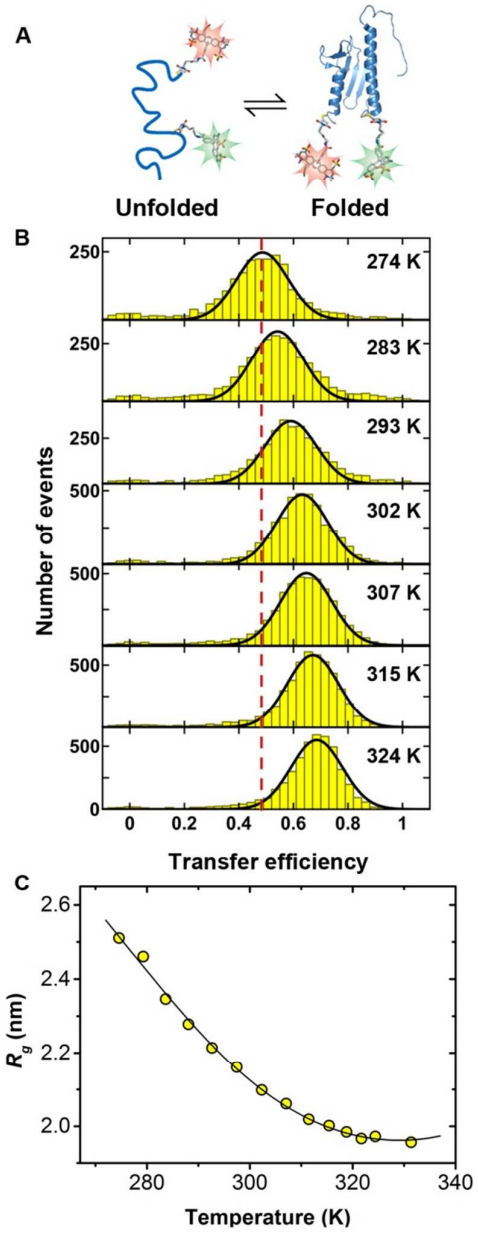

Figure S5. (A) Cartoons of unfolded and folded Yfh1-47/120 (based on PDB ID 2GA5) with Alexa Fluor 488 (green) and Alexa Fluor 594 (red) dyes at positions C47 and S120C. (B) Representative examples of single-molecule FRET efficiency histograms of Yfh1-47/120 show a temperature-induced collapse of the unfolded protein. In this variant, only the unfolded-state peak is detected because of the very close proximity of the fluorophores in the folded structure, resulting in mutual quenching and thus a reduction in molecular brightness of folded molecules[25] (see Fig. S6). As a result of the threshold-based burst selection in free-diffusion experiments, the folded-state population does not appear in the histograms, and there is thus no overlap with the signal from folded molecules that could interfere with the analysis of the unfolded-state population. To determine mean transfer efficiencies, peaks were fit with a single Gaussian distribution (black lines). The red dashed line indicates the position of the unfolded peak at 274 K. (C) Resulting temperature dependence of the radius of gyration (R_g) (circles) determined analogous to Fig. 3B. The solid line represents a fit of R_g as a

function of temperature (see *Determination of the radius of gyration from mean transfer efficiencies* in SI and Eq. S10). The same type of temperature dependence as for Yfh1-16/120 is observed, albeit with a smaller apparent R_g because of the smaller sequence separation of the fluorophores, thus confirming the results shown in Fig. 3.

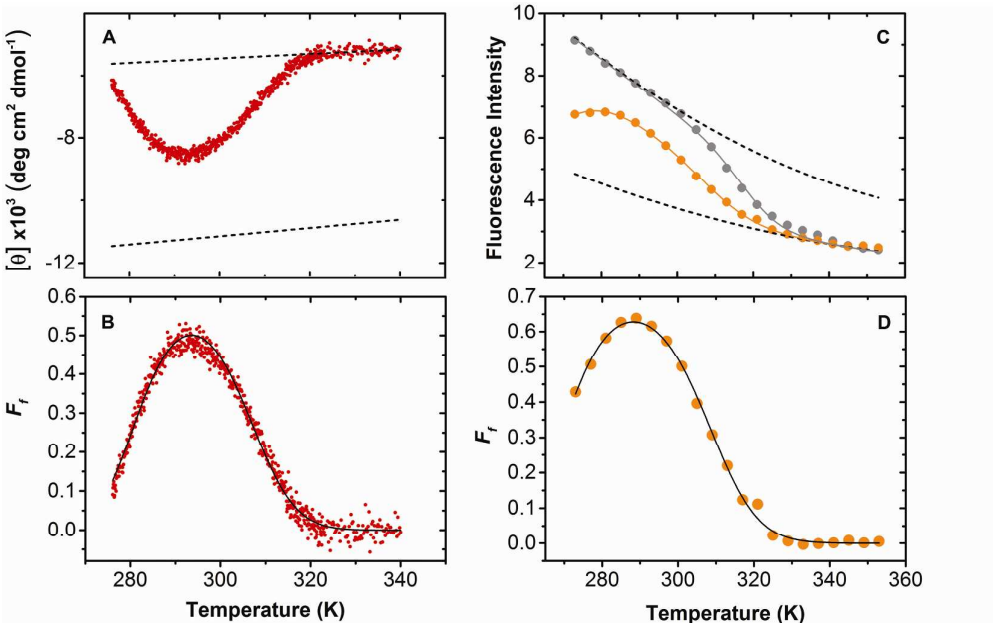

Figure S6. Circular dichroism and ensemble fluorescence analysis showing heat and cold denaturation of Yfh1-47/120. (A) The temperature dependence of the conformational stability in 50 mM sodium phosphate buffer, pH 7.0 (red dots) monitored by changes in ellipticity at 222 nm. The baselines from Fig. 1A were used and scaled linearly by a factor 1.09 to adjust for the slight difference in protein concentration (dashed lines). The black solid line shows a fit with ΔH_S, ΔC_p, and T_S as fit parameters (see main text and *Thermodynamic analysis of unfolding curves* for details). (B) Temperature dependence of the fraction of folded (F_f) Yfh1-47/120 (red dots) obtained from the CD measurements shown in (A), with a fit using Eq. S14 and S15 shown as a solid line. (C) Ensemble Trp fluorescence measurements (excitation 295 nm, emission 350 nm) of Yfh1-47/120 in the absence (orange symbols) and presence (grey symbols) of 0.3 M sodium sulfate. The temperature dependence of the quantum yield of NATA was used to approximate the temperature dependences of the fluorescence intensities of folded and unfolded protein. Following the approach of Makhatadze and coworkers[17], the curve for NATA was scaled to match the unfolded state baseline (lower dashed line) and the folded state baseline in the presence of sodium sulfate (upper dashed line). (D) Temperature dependence of the fraction of folded (F_f) Yfh1-47/120 (orange symbols) obtained from the fluorescence measurements shown in (C), with a fit using Eq. S14 and S15 shown as a solid line. Both the CD and the ensemble fluorescence measurements clearly demonstrate the presence of a folded state with a maximum population and unfolding temperatures similar to variant Yfh1-16/120 (Fig. 1C & 3A), confirming that the absence of a folded-state population in the transfer efficiency histograms (Fig. S5B) is caused by mutual quenching of the FRET dyes owing to their close proximity in the folded structure and their resulting low molecular brightness.

Supporting Information

Single-Molecule Spectroscopy of Cold Denaturation and the Temperature-Induced Collapse of Unfolded Proteins

Mikayel Aznauryan, Daniel Nettels, Andrea Holla, Hagen Hofmann, and Benjamin Schuler*

University of Zurich, Department of Biochemistry, Winterthurerstrasse 190,
8057 Zurich, Switzerland

Corresponding Author
***** Benjamin Schuler
Email: schuler@bioc.uzh.ch

Preparation and Labeling of Protein

The coding sequence for Yhf1, comprising residues 1-123 of the mature gene product (Swiss-Prot entry: Q07540) with V1M and S120C, was codon optimized for expression in *E.* coli and cloned into pET47b(+) using the SmaI and PacI sites. On the resulting expression plasmid pFra1, frataxin is encoded as a fusion protein with an N-terminal His_6-tag and recognition site for HRV 3C protease. Gene synthesis and construction of the expression plasmid for recombinant Yhf1 was carried out by EZBiolab (Carmel, USA). Based on the pFra1 expression plasmid, we constructed a variant with C47S, N16C and S120C. We used the following procedure for the expression and purification of both Yfh1-16/120 and Yfh1-47/120 variants. His_6-frataxin was produced in *E. coli* BL21(DE3) harboring pFra1. 1 L cultures were grown in LB media for 4 hours at 37°C and then induced by 0.5 mM IPTG for 16 hours at 20°C. After harvesting by centrifugation, the cells were resuspended in Lysis Buffer (20 mM Tris-HCl, pH 8.0, 250 mM NaCl, 20 mM imidazole, 1 mM PMSF) and disrupted by sonication on ice. His_6-frataxin was purified from the soluble fraction of the lysate using a 5 ml HisTrap column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden). Bound protein was eluted with a linear gradient of 20 to 500 mM imidazole in 20 mM Tris-HCl, pH 7.4, 150 mM NaCl after washing the column with 20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 20 mM imidazole. Purified His_6-frataxin was detected in wash and elution fractions. Pooled wash fractions were subjected to digestion with HRV 3C protease at 4 °C for 16 h at a Yhf1 to protease molar ratio of about 100 to 1. After addition of β-mercaptoethanol to a final concentration of 1 mM, His_6-tagged HRV 3C protease (produced in-house), His_6-tag containing peptide, and uncut protein were separated from frataxin using a 5 ml HisTrap column. Yhf1 in the flow-through was denatured in 4 M guanidine, reduced with 50 mM DTT, concentrated and subjected to gel filtration using a Superdex 75 10/300 GL column (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) in 100 mM NaH_2PO_4, pH 7.1, 6 M guanidinium chloride. Yhf1-containing fractions were labeled with Alexa Fluor 488 maleimide (Molecular Probes, Eugene, USA) at a dye:protein molar ratio of 0.7:1 for 2 h at room temperature and subsequently with Alexa Fluor 594 maleimide at a dye:protein molar ratio of 2:1 overnight at 4 °C. Unreacted dye was removed by gel filtration, again using a Superdex 75 10/300 GL column in 100 mM NaH_2PO_4, pH 7.1, 6 M guanidinium chloride. The resulting sample contained recombinant Yhf1 labeled with Alexa Fluor 488 and 594, and frataxin doubly-labeled with either Alexa Fluor 488 or Alexa Fluor 594, as shown by mass spectrometry.

Single-Molecule Fluorescence Spectroscopy

Observations of single molecule fluorescence were made using a MicroTime 200 confocal microscope equipped with a HydraHarp 400 counting module (Picoquant, Germany) and an Olympus UplanApo 60×/1.20W objective. The excitation of the dyes was achieved using pulsed interleaved excitation (PIE)[1]. The wavelength range used for acceptor excitation was selected with a z582/15 band pass filter (Chroma) from the emission of an SC-450-4 supercontinuum fiber laser (Fianium, UK) driven at 20 MHz, which triggers the diode laser used for donor excitation (LDH-D-C-485, PicoQuant, Germany). Emitted photons were collected through the microscope objective, focused onto a 100 μm pinhole, and then separated into four channels with a polarizing beam splitter and two dichroic mirrors

(585DCXR, Chroma). Photons were additionally filtered by bandpass filters (ET525/50M and HQ650/100, Chroma) before being focused onto one of four single photon avalanche detectors (SPADs; Optoelectronics SPCMAQR-15, PerkinElmer, Wellesley, MA). Samples of labeled protein were diluted to a concentration of about 20 pM in 50 mM sodium phosphate buffer, pH 7.0. Tween 20 (0.001%; Pierce) was added to prevent surface adhesion of the protein. To reduce data acquisition time and thus minimize damage to the chromophores at high temperatures, the photoprotective additive β-mercaptoethanol (143 mM) was added. Measurements were performed by placing the confocal spot at a position 50 μm deep in the solution with a total acquisition time of 30 minutes.

For the identification of photon bursts caused by proteins diffusing through the confocal volume, we used only the photons detected after donor excitation, i.e. detected in time intervals of ~25 ns following pulses from the diode laser. Successive photons separated by inter photon times <100 μs were combined into one burst with n_D and n_A photons counted in the donor and acceptor detection channels, respectively. These numbers were corrected for background. Transfer efficiency histograms were obtained from bursts with $n'_A + n_D > 80$ according to:

$$E = \frac{n'_A}{n'_A + \gamma n_D} \qquad \textbf{(S1)}$$

$$\text{with } n'_A = n_A - \frac{\alpha}{1+\alpha}\left(n_A + \gamma n_D\right) - \beta n_D$$

where $\alpha = 0.05$ is the ratio of the extinction coefficients of Alexa 594 and Alexa 488 at the laser donor excitation wavelength ($\lambda_{ex,D} = 482$ nm), β is the ratio of donor photons detected in the acceptor channel (crosstalk), and $\gamma = \eta_A Q_A / \eta_D Q_D$ is the ratio of the products of quantum yields and detection efficiencies for acceptor and donor. We obtained values for α, β, and γ at room temperature (T=295 K) as described previously[2]. Whereas α and β are virtually independent of temperature in the range investigated here, we observed a significant decrease of molecular brightness of the fluorophores with increasing temperature. From the additional observation that also the fluorescence lifetimes of the dyes decreases with increasing temperature, we conclude that the effect of brightness decrease is well explained by a temperature dependent increase of collisional quenching of the dyes. Hence, the changes in fluorescence lifetimes can be used to correct $\gamma = \gamma(T)$ according to:

$$\gamma(T) = \gamma_0 \frac{\tau_{D0} \cdot \tau_A(T)}{\tau_{A0} \cdot \tau_D(T)} \qquad \textbf{(S2)}$$

where $\gamma_0 = \gamma(295\text{ K})$ is the value determined at room temperature. $\tau_D(T)$ and $\tau_A(T)$ are the donor and acceptor fluorescence lifetimes at temperature T, and τ_{D0} and τ_{A0} are the corresponding lifetimes at T=295 K. The lifetimes were derived from the same single-molecule photon data from which the transfer efficiency histograms were obtained. The mean donor and acceptor fluorescence lifetimes were obtained from the burst photons' arrival times

measured with respect to the corresponding donor or acceptor excitation pulses. For determining the donor lifetime, only photons of bursts with E<0.1, emitted by molecules lacking an active acceptor were used[3]. The resulting temperature dependence of the correction factor γ is shown in Figure S1.

Similarly, the temperature dependence of the Förster radius $R_0(T)$ was derived from the one at room temperature $R_0(295\ \mathrm{K})$ (5.4 nm for Alexa Fluor 488 and Alexa Fluor 594 dye pair) according to:

$$R_0^6(T) = R_0^6(295\ \mathrm{K})\frac{\tau_D(T)}{\tau_{D0}}\left(\frac{n(295\ \mathrm{K})}{n(T)}\right)^4 \qquad \textbf{(S3)}$$

where $n(T)$ is the refractive index of the sample solution at temperature T, which was measured for each temperature with a digital Abbe refractometer (Krüss, Germany). For each burst, we also determined the stoichiometry ratio $S = \frac{n_D + n_A}{n_D + n_A + n_{AA}}$, where n_{AA} is the number of photons detected after acceptor excitation, i.e. detected in the 25 ns time intervals following the 582-nm pulses. Figure S2A shows the 2D histogram of stoichiometry ratio and transfer efficiency values of the bursts detected at room temperature (T=295 K). The histogram is dominated by a peak corresponding to molecules lacking an active acceptor (E=0 and S=1). For building the transfer efficiency histograms (Fig. 2 in main text), we only used bursts with S<0.7. In Figure S2B, we compare such a histogram (green) with a transfer efficiency histogram obtained from all bursts (grey), including those with S>0.7.

A custom-built temperature-controlled sample holder employing Peltier elements and a digital temperature controller (TED 4015, Thorlabs) with a Pt100 temperature sensor was used[4]. The thermal contact of the cover slide to the objective via the immersion water results in a temperature gradient in the sample. To keep the gradient small, the microscope objective was also temperature-regulated using an aluminum collar to which a water-cooled Peltier element was mounted. Below 313 K, the sample holder and the objective were heated or cooled to the same temperature. Only the holder could be heated to higher temperatures while we kept the objective at 313 K to avoid damage. The temperature at the confocal volume was calibrated using the temperature-dependent fluorescence lifetime of rhodamine B[4,5]. The resulting temperature calibration is shown in Figure S3.

Determination of the radius of gyration from mean transfer efficiencies

The calculation of the radius of gyration of unfolded protein chains from the measured mean transfer efficiencies was performed assuming two different inter-dye distance distributions: the Gaussian chain distribution[3,4,6] and a distribution obtained from a variation of coil-to-globule transition theory of Sanchez[7-9]. Experimentally determined mean transfer efficiencies of the unfolded state can be calculated from a dye-to-dye distance distribution $P(r)$ according to:

$$\langle E\rangle = \int_0^L E(r)P(r)dr \Big/ \int_0^L P(r)dr \qquad \textbf{(S4)}$$

with

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

The distribution for a Gaussian chain is:

$$P_{Gaussian}(r) = 4\pi r^2 \left(\frac{3}{2\pi\langle r^2\rangle}\right)^{3/2} \exp\left(-\frac{3r^2}{2\langle r^2\rangle}\right) \qquad \textbf{(S5)}$$

Given a value of $\langle E\rangle$, the mean-squared end-to-end distance $\langle r^2\rangle$ can be calculated numerically, which can be converted to the radius of gyration by $R_g = \sqrt{\langle r^2\rangle/6}$ [10].

The theory of Sanchez provides an expression for the probability density function of the radius of gyration, r_g, of the probed segment as a Boltzmann-weighted Flory-Fisk distribution:

$$P_{Sanchez}\left(r_g, \varepsilon, R_{g\Theta}\right) = Z^{-1} r_g^6 \exp\left\{-\frac{7r_g^2}{2R_{g\Theta}^2} + n\left[\frac{1}{2}\varepsilon\phi - \frac{1-\phi}{\phi}\ln\left(1-\phi\right)\right]\right\} \qquad \textbf{(S6)}$$

Here, $n = N+1$ is the number of amino acids (monomers) in the chain segment probed, N is the number of peptide bonds, ε is the mean-field interaction free energy between amino acids, $\phi = \left(R_C/r_g\right)^3$ is the volume fraction of the chain with respect to the most compact state of radius $R_C = \left[3n\mathrm{v}/4\pi\right]^{1/3}$, with the average volume of one amino acid being $\mathrm{v} = 0.13$ nm^3 as shown in[9], and Z is a normalization factor. A realistic value of the radius of gyration of the chain at the Θ-state, $R_{g\Theta}$, has recently been estimated using a generalized expression for the length-scaling of polymers derived by Hammouda[11]:

$$R_{g\Theta} = \sqrt{\frac{2l_p b}{\left(2\nu+1\right)\left(2\nu+2\right)}} N_{dd}^{\nu} \qquad \textbf{(S7)}$$

In Eq. S7, $b = 0.38$ nm is the distance between two C_α-atoms[12], $\nu = 1/2$, $l_p = 0.4$ nm is the persistence length* of unfolded proteins in good solvent extracted from SAXS[9,13], and $N_{dd} = N+9$ takes into account the length of the fluorophore linkers approximately[3,9,14,15].

To relate the distribution of radii of gyration $P(r_g, \varepsilon, R_{g\Theta})$ to a segment end-to-end distance distribution $P(r, \varepsilon, R_{g\Theta})$, which is required to describe the mean transfer

* Here we assume that l_p is to a good approximation independent of temperature. Calculations show that a variation of l_p by 25% result in a minor change of R_g (±0.03 nm).

efficiencies $\langle E \rangle$, we used the conditional probability density function $P(r|r_g)$ suggested by Ziv and Haran[16], which describes the distance distribution of two random points inside the sphere of radius $\delta \cdot r_g$:

$$P\left(r\middle|r_g\right)=\frac{1}{\delta\cdot r_g}\left[3\left(\frac{r}{\delta\cdot r_g}\right)^2-\frac{9}{4}\left(\frac{r}{\delta\cdot r_g}\right)^3+\frac{3}{16}\left(\frac{r}{\delta\cdot r_g}\right)^5\right] \qquad 0\le r<2\delta\cdot r_g \qquad \textbf{(S8)}$$

Here, δ was obtained from the condition that $6\langle R_g^2\rangle=\langle r^2\rangle$ at the Θ-state. The mean transfer efficiency is finally given by:

$$\langle E\rangle=\int_0^L E(r)P(r,\varepsilon,R_{g\Theta})dr=\int_0^L E(r)\int_{R_C}^{L/2}P(r|r_g)\,P(r_g,\varepsilon,R_{g\Theta})dr_g dr\,, \qquad \textbf{(S9)}$$

where R_0 is the Förster radius, and L is the contour length of the protein segment probed. Eq. S9 was solved for ε numerically for each measured temperature, and the mean radii of gyration R_g (dots in Fig. 3B and Fig. S5C) were finally determined according to:

$$R_g=\sqrt{\int_{R_C}^{L/2}r_g^2P(r_g,\varepsilon,R_{g\Theta})dr_g} \qquad \textbf{(S10)}$$

For fitting the R_g temperature dependencies, we substituted the mean-field interaction free energy in Eq. S10 by $\varepsilon(T)=\Delta h_0+\Delta c_p\left(T-T_0\right)-T\left(\Delta s_0+\Delta c_p\ln\left(T/T_0\right)\right)$, where Δh_0, Δs_0, and Δc_p are the changes of enthalpy, entropy, and heat capacity per monomer (amino acid residue) relative to the random coil state ($\varepsilon=0$); T_0 was chosen to be 295 K. From the fit to the data of Yhf1-16/120 we obtained $\Delta h_0=-7.0\pm0.1$ kJ mol^{-1}, $\Delta s_0=-31.5\pm1.1$ J mol^{-1} K^{-1}, and $\Delta c_p=402\pm8$ J mol^{-1} K^{-1}.

The resulting values of R_g correspond to the protein chain segment probed, including the two fluorophore linkers (N_{dd} = N+9). For obtaining the radius of gyration of the full-length protein, $R_{g,full}$ (dashed line in Figure 3B, main text), the relation $R_{g,full}=R_g\left(N_{full}/N_{dd}\right)^{\nu}$ was used, where the scaling exponents ν, corresponding to each probed temperature, were obtained from Eq. S7 using the calculated R_g values instead of $R_{g\Theta}$.

The free energy of collapse, ΔG_C, was calculated with respect to the chain at the excluded volume limit (with volume fraction ϕ_{ex}), where $\varepsilon=0$:

$$\Delta G_C=k_BT\left[\ln(P(\phi,\varepsilon))-\ln(P(\phi_{ex},0))\right] \qquad \textbf{(S11)},$$

where

$$P(\phi,\varepsilon) = Z^{-1}\left(\frac{\phi_0}{\phi}\right)^2 \exp\left\{-\frac{7}{2}\left(\frac{\phi_0}{\phi}\right)^{2/3} + n\left[\frac{\phi}{2}\varepsilon - \frac{1-\phi}{\phi}\ln(1-\phi)\right]\right\} \qquad \textbf{(S12)}$$

and $\phi_0 = \left(R_C / R_{g\Theta}\right)^3$. To obtain the values of ΔG_C for the full-length protein, we first calculated the interaction energy ε for the full-length protein using the scaled values of gyration radii, $R_{g,full}$. These ε values were then used to calculate the ΔG_C corresponding to the full-length protein. The volume fraction at the excluded volume limit (ϕ_{ex}) was obtained by numerically solving

$$-\frac{d\ln\left(P\left(\phi_{ex},\varepsilon\right)\right)}{d\phi} = 0 \qquad \textbf{(S13)}$$

for ϕ with $\varepsilon = 0$.

Circular Dichroism Spectroscopy

CD experiments were performed using a J-810 CD spectrometer (Jasco) equipped with temperature-controlled water bath programmed to raise the temperature at a rate of 1 K/min. Measurements were carried out in cylindrical quartz cells with 1 mm path length and at a protein concentration of 0.35 mg/ml. The unfolding curves of Yhf1 were obtained by monitoring the CD ellipticities at 222 nm as a function of temperature in 50 mM sodium phosphate buffer, pH 7.0. The baselines for folded and unfolded subpopulations were approximated by linear functions of temperature.

Fluorescence Spectroscopy

Ensemble fluorescence measurements were performed on a Fluorolog fluorimeter (Jobin Yvon) equipped with a peltier-controlled cell holder, which was used to raise the temperature at a rate of 1 K/min. The actual temperature of the sample was monitored using a temperature probe directly immersed in the cuvette. Measurements were carried out in quartz cuvette with 3 mm path length. The protein concentration was 1 μM; 0.001% Tween 20 were added to the solution to prevent surface adhesion of the protein. Temperature-induced unfolding of Yhf1 was monitored by the change in the tryptophan fluorescence intensity at 350 nm after excitation at 295 nm. Since the intrinsic Trp fluorescence quantum yield itself strongly depends on temperature, and this dependence is nonlinear, it is difficult to estimate the temperature dependences of the folded and unfolded baselines. To get an estimate of the temperature dependencies of the folded and unfolded states, we used *N*-acetyltryptophanamide (NATA), following the approach of Makhatadze and coworkers[17]. The temperature dependence of the fluorescence intensity change for 2 μM NATA (the same molar concentration as that of Trp residues in frataxin) was recorded under identical conditions. The changes of the fluorescence intensity of NATA with temperature are well described by a second-order polynomial. This polynomial was rescaled to define the temperature dependencies of the fluorescence intensity of the folded, $I_f(T)$, and unfolded, $I_u(T)$ Yhf1.

Thermodynamic analysis of unfolding curves

The unfolding curves of Yhf1 obtained both from CD and fluorescence measurements were fitted assuming a two-state system, for which the experimental observable, $Y(T)$, (CD signal at 222 nm or Trp fluorescence intensity at 350 nm) can be written as:

$$Y(T) = F_f(T) \cdot Y_f(T) + [1 - F_f(T)] \cdot Y_u(T) \qquad \textbf{(S14)}$$

where $Y_f(T)$ and $Y_u(T)$ are the baselines of folded and unfolded subpopulations, respectively. Folded-state baselines in CD and ensemble fluorescence measurements (Fig. 1A, Fig. S6C) were obtained by stabilization of the protein variants with 0.3 M sodium sulfate (pH 7.0), a commonly used kosmotrope for non-specifically stabilizing the folded state by preferential hydration[18]. Pastore and coworkers had previously shown that sodium sulfate strongly stabilizes frataxin[19], and we follow established procedures used e.g. by Shortle and coworkers[20], who employed ammonium sulfate for obtaining native state baselines, or Makhatadze and coworkers[17], who used changes in pH in an analogous way.

The fraction of folded protein $F_f(T)$ is given by:

$$F_f(T) = 1/(1 + \exp(-[\Delta H(T) \; - \; T\Delta S(T)]/RT)) \qquad \textbf{(S15)}$$

where the change in enthalpy $\Delta H(T)$ and entropy $\Delta S(T)$ given by:

$$\Delta H(T) = \Delta H_S \; + \; (T \; - \; T_S) \cdot \Delta C_p \qquad \text{and} \qquad \Delta S(T) = \Delta C_p \cdot \ln\left(\frac{T}{T_S}\right), \qquad \textbf{(S16)}$$

where ΔH_S is the enthalpy change at temperature T_S,[†] which is defined as the temperature at which the fraction of folded protein is maximal, i.e. where the Gibbs free energy reaches its maximum and the entropy change between folded and unfolded state is zero ($\Delta S(T_S) = 0$). ΔC_p is the change in heat capacity upon unfolding[17,21]. The fraction of folded protein was calculated from:

$$F_f = 1 - F_u = \frac{Y(T) - Y_u(T)}{Y_f(T) - Y_u(T)} \qquad \textbf{(S17)}$$

Table S1 summarizes the thermodynamic parameters derived from our CD and FRET measurements in comparison with the values for wild-type Yhf1 calculated using the data by Pastore et al.[22].

[†]The reason for using T_S (instead of the melting temperature, T_m) as a reference temperature is that in cases where the population of the folded state is less than 50% at all temperatures, T_m cannot be defined.

REFERENCES

(1) Muller, B. K.; Zaychikov, E.; Brauchle, C.; Lamb, D. C. *Biophys J* **2005**, *89*, 3508.

(2) Schuler, B. *Methods Mol Biol* **2007**, *350*, 115.

(3) Hoffmann, A.; Kane, A.; Nettels, D.; Hertzog, D. E.; Baumgartel, P.; Lengefeld, J.; Reichardt, G.; Horsley, D. A.; Seckler, R.; Bakajin, O.; Schuler, B. *P Natl Acad Sci USA* **2007**, *104*, 105.

(4) Nettels, D.; Muller-Spath, S.; Kuster, F.; Hofmann, H.; Haenni, D.; Ruegger, S.; Reymond, L.; Hoffmann, A.; Kubelka, J.; Heinz, B.; Gast, K.; Best, R. B.; Schuler, B. *P Natl Acad Sci USA* **2009**, *106*, 20740.

(5) Benninger, R. K. P.; Koc, Y.; Hofmann, O.; Requejo-Isidro, J.; Neil, M. A. A.; French, P. M. W.; deMello, A. J. *Anal Chem* **2006**, *78*, 2272.

(6) O'Brien, E. P.; Morrison, G.; Brooks, B. R.; Thirumalai, D. *J Chem Phys* **2009**, *130*, 124903.

(7) Sherman, E.; Haran, G. *P Natl Acad Sci USA* **2006**, *103*, 11539.

(8) Muller-Spath, S.; Soranno, A.; Hirschfeld, V.; Hofmann, H.; Ruegger, S.; Reymond, L.; Nettels, D.; Schuler, B. *P Natl Acad Sci USA* **2010**, *107*, 14609.

(9) Hofmann, H.; Soranno, A.; Borgia, A.; Gast, K.; Nettels, D.; Schuler, B. *P Natl Acad Sci USA* **2012**, *109*, 16155.

(10) Grosberg, A. Y.; Khokhlov, A. R. *Statistical Physics of Macromolecules*; AIP Press: Woodbury, NY, 1994.

(11) Hammouda, B. *Adv Polym Sci* **1993**, *106*, 87.

(12) Zhou, H. X. *Biochemistry* **2004**, *43*, 2141.

(13) Kohn, J. E.; Millett, I. S.; Jacob, J.; Zagrovic, B.; Dillon, T. M.; Cingel, N.; Dothager, R. S.; Seifert, S.; Thiyagarajan, P.; Sosnick, T. R.; Hasan, M. Z.; Pande, V. S.; Ruczinski, I.; Doniach, S.; Plaxco, K. W. *P Natl Acad Sci USA* **2004**, *101*, 12491.

(14) Schroder, G. F.; Alexiev, U.; Grubmuller, H. *Biophys J* **2005**, *89*, 3757.

(15) McCarney, E. R.; Werner, J. H.; Bernstein, S. L.; Ruczinski, I.; Makarov, D. E.; Goodwin, P. M.; Plaxco, K. W. *J Mol Biol* **2005**, *352*, 672.

(16) Ziv, G.; Haran, G. *J Am Chem Soc* **2009**, *131*, 2942.

(17) Richardson, J. M., 3rd; Lemaire, S. D.; Jacquot, J. P.; Makhatadze, G. I. *Biochemistry* **2000**, *39*, 11154.

(18) Timasheff, S. N. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 67.

(19) Adinolfi, S.; Nair, M.; Politou, A.; Bayer, E.; Martin, S.; Temussi, P.; Pastore, A. *Biochemistry* **2004**, *43*, 6511.

(20) Shortle, D.; Stites, W. E.; Meeker, A. K. *Biochemistry* **1990**, *29*, 8033.

(21) Ibarra-Molero, B.; Makhatadze, G. I.; Sanchez-Ruiz, J. M. *Biochim Biophys Acta* **1999**, *1429*, 384.

(22) Pastore, A.; Martin, S. R.; Politou, A.; Kondapalli, K. C.; Stemmler, T.; Temussi, P. A. *Journal of the American Chemical Society* **2007**, *129*, 5374.

(23) Sanchez, I. C. *Macromolecules* **1979**, *12*, 980.

(24) Flory, P. J. *Statistical mechanics of chain molecules*; Hanser Publishers: Munich, Vienna, New York, 1989.

(25) Chung, H. S.; Louis, J. M.; Eaton, W. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 11837.

Table S1. Thermodynamic parameters of cold and heat unfolding of Yfh1-16/120

	CD measurement	Single-molecule FRET measurement	Data from Pastore et al.[22]
T_S (K)	289.2 ± 0.1	289.7 ± 0.3	291.9 ± 0.5
ΔH_S (kJ mol^{-1})	1.19 ± 0.02	- 0.25 ± 0.05	1.72 ± 0.08
ΔC_p (kJ mol^{-1} K^{-1})	6.28 ± 0.21	5.62 ± 0.19	7.57 ± 0.15

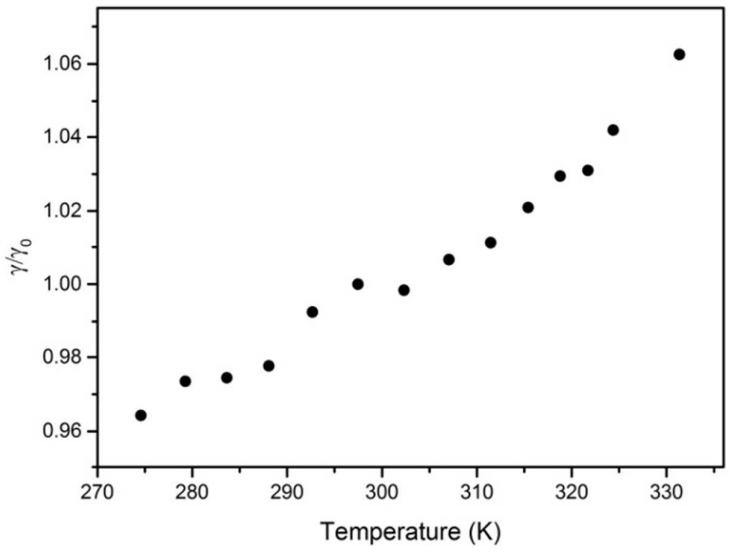

Figure S1. Temperature dependence of the correction factor γ.

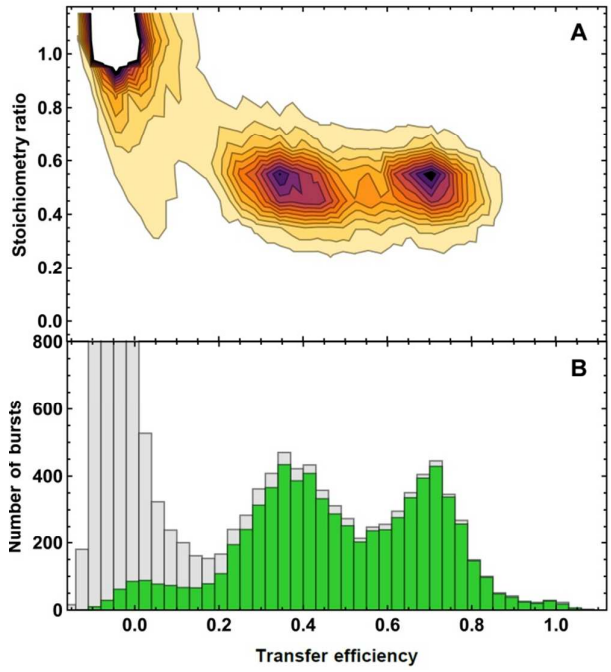

Figure S2. (A) 2D contour plot of stoichiometry ratio vs. transfer efficiency constructed using our data set for the measurement of Yfh1-16/120 at room temperature. The two peaks at 0.5 stoichiometry ratio originate from the protein molecules labeled with both active donor and acceptor fluorophores. The peak at high stoichiometry and $E \approx 0$ corresponds to molecules with an inactive acceptor fluorophore. (B) Transfer efficiency histograms of double-labeled Yhf1-16/120 at room temperature. The grey histogram was obtained using the dataset with the donor excitation only, whereas the green histogram was recorded using pulsed interleaved excitation of donor and acceptor fluorophores and selection of molecules with $S<0.7$.

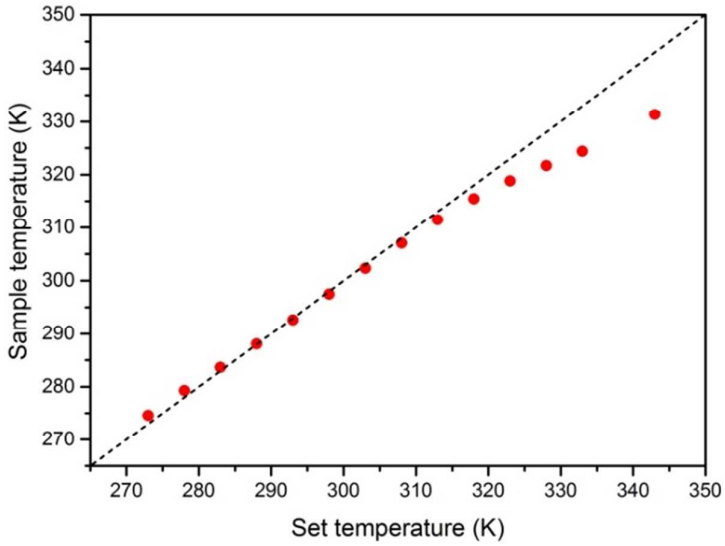

Figure S3. Temperature calibration showing the relation between the temperature of the sample holder and the sample temperature in the confocal volume as determined from rhodamine fluorescence lifetime measurements.

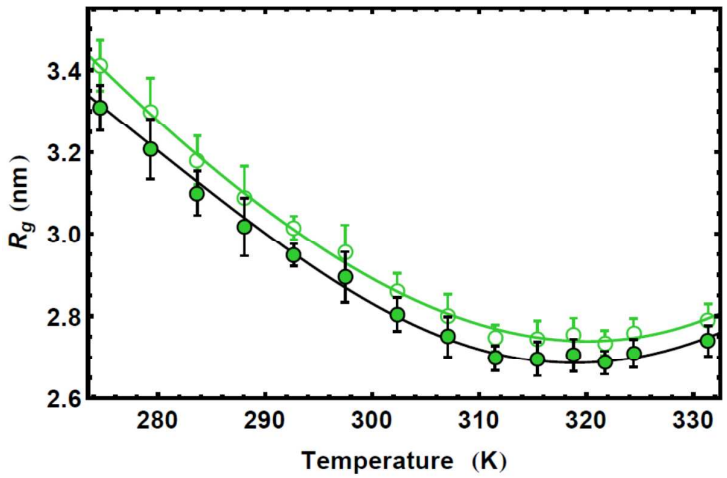

Figure S4. Comparison of the values for gyration radii for Yfh1-16/120 calculated according to the Sanchez model (filled green circles) and assuming a Gaussian chain distribution (open green circles), respectively. To test the robustness of our results, we analyzed the data with two different distribution, the Gaussian chain[3,6] distribution and a distribution obtained from a variant of coil-to-globule transition theory of Sanchez[7,9,16,23], which uses a Flory–Fisk distribution[24] weighted by a potential accounting for effective interactions within the chain. The results from the two types of analysis are very similar, suggesting that the chain dimensions derived from the FRET data under our conditions are robust with respect to the details of the shape of the distance distribution used. In the main text, we thus use the dimensions calculated using the Sanchez model, which is physically most realistic.

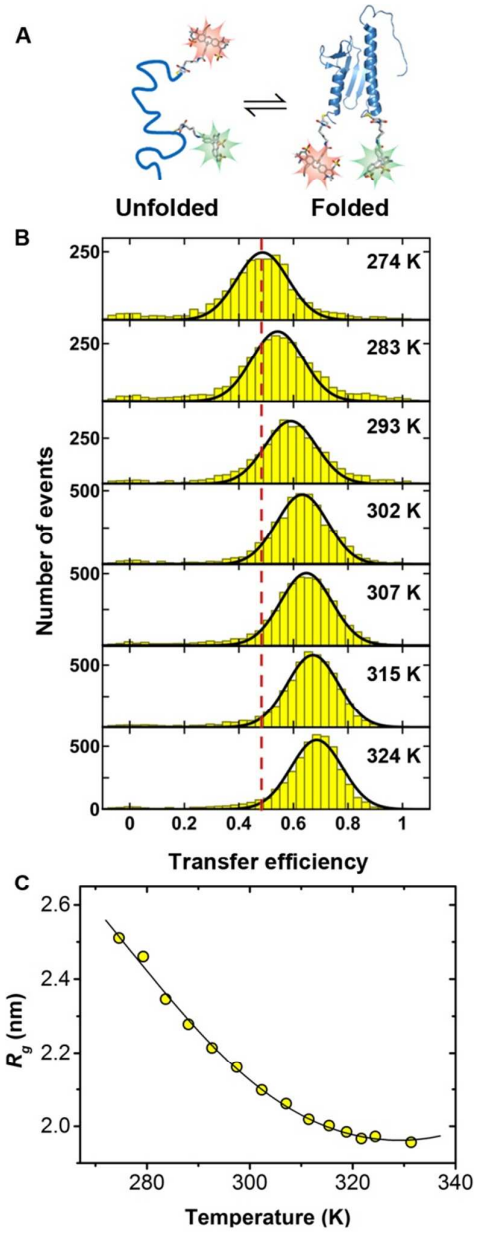

Figure S5. (A) Cartoons of unfolded and folded Yfh1-47/120 (based on PDB ID 2GA5) with Alexa Fluor 488 (green) and Alexa Fluor 594 (red) dyes at positions C47 and S120C. (B) Representative examples of single-molecule FRET efficiency histograms of Yfh1-47/120 show a temperature-induced collapse of the unfolded protein. In this variant, only the unfolded-state peak is detected because of the very close proximity of the fluorophores in the folded structure, resulting in mutual quenching and thus a reduction in molecular brightness of folded molecules[25] (see Fig. S6). As a result of the threshold-based burst selection in free-diffusion experiments, the folded-state population does not appear in the histograms, and there is thus no overlap with the signal from folded molecules that could interfere with the analysis of the unfolded-state population. To determine mean transfer efficiencies, peaks were fit with a single Gaussian distribution (black lines). The red dashed line indicates the position of the unfolded peak at 274 K. (C) Resulting temperature dependence of the radius of gyration (R_g) (circles) determined analogous to Fig. 3B. The solid line represents a fit of R_g as a

function of temperature (see *Determination of the radius of gyration from mean transfer efficiencies* in SI and Eq. S10). The same type of temperature dependence as for Yfh1-16/120 is observed, albeit with a smaller apparent R_g because of the smaller sequence separation of the fluorophores, thus confirming the results shown in Fig. 3.

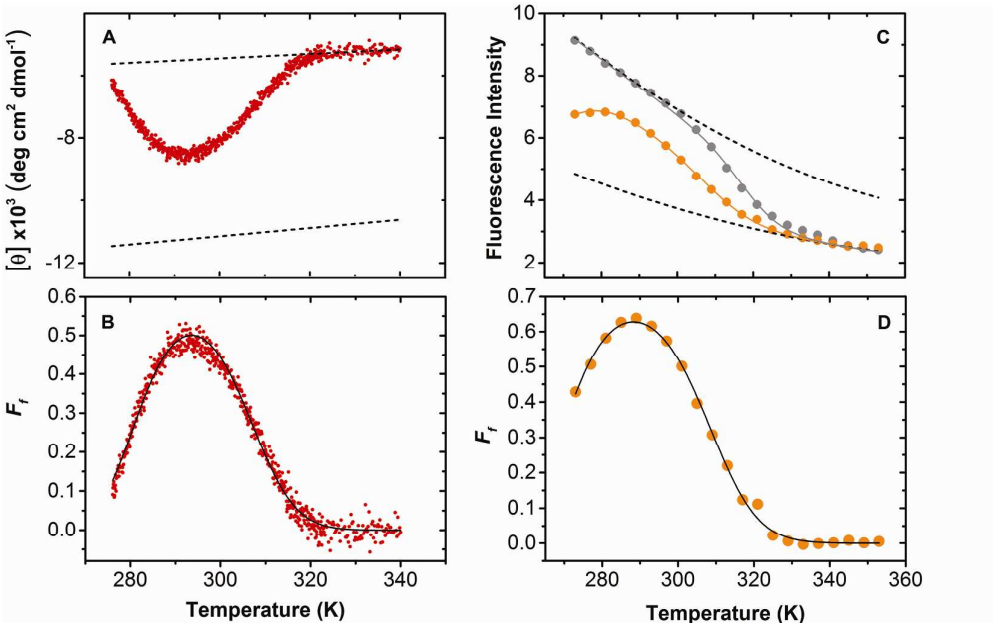

Figure S6. Circular dichroism and ensemble fluorescence analysis showing heat and cold denaturation of Yfh1-47/120. (A) The temperature dependence of the conformational stability in 50 mM sodium phosphate buffer, pH 7.0 (red dots) monitored by changes in ellipticity at 222 nm. The baselines from Fig. 1A were used and scaled linearly by a factor 1.09 to adjust for the slight difference in protein concentration (dashed lines). The black solid line shows a fit with ΔH_S, ΔC_p, and T_S as fit parameters (see main text and *Thermodynamic analysis of unfolding curves* for details). (B) Temperature dependence of the fraction of folded (F_f) Yfh1-47/120 (red dots) obtained from the CD measurements shown in (A), with a fit using Eq. S14 and S15 shown as a solid line. (C) Ensemble Trp fluorescence measurements (excitation 295 nm, emission 350 nm) of Yfh1-47/120 in the absence (orange symbols) and presence (grey symbols) of 0.3 M sodium sulfate. The temperature dependence of the quantum yield of NATA was used to approximate the temperature dependences of the fluorescence intensities of folded and unfolded protein. Following the approach of Makhatadze and coworkers[17], the curve for NATA was scaled to match the unfolded state baseline (lower dashed line) and the folded state baseline in the presence of sodium sulfate (upper dashed line). (D) Temperature dependence of the fraction of folded (F_f) Yfh1-47/120 (orange symbols) obtained from the fluorescence measurements shown in (C), with a fit using Eq. S14 and S15 shown as a solid line. Both the CD and the ensemble fluorescence measurements clearly demonstrate the presence of a folded state with a maximum population and unfolding temperatures similar to variant Yfh1-16/120 (Fig. 1C & 3A), confirming that the absence of a folded-state population in the transfer efficiency histograms (Fig. S5B) is caused by mutual quenching of the FRET dyes owing to their close proximity in the folded structure and their resulting low molecular brightness.