Conformational Properties of Unfolded HypF-N

Yujie Chen,[†] Claudia Parrini,[‡] Niccolò Taddei,[‡] and Lisa J. Lapidus*[,†]

Department of Physics and Astronomy, Michigan State University, East Lansing, Michigan 48824, and Department of Biochemical Sciences, University of Florence, Viale Morgagni 50, 50134 Florence, Italy

Received: May 5, 2009; Revised Manuscript Received: October 26, 2009

We have measured the intramolecular diffusion rate between distant residues in the aggregation-prone protein HypF-N under various denaturing conditions. Using the method of cysteine quenching of the tryptophan triplet state, we find that intramolecular diffusion remains roughly constant at high concentrations of denaturant (2−6 M GdnHCl) and slows down at low concentrations of denaturant, but the decrease is not uniform throughout the chain. Extrapolation of these measurements to 0 M GdnHCl gives $D \sim 10^{-7}$ cm^2 s^{-1}, about 1 order of magnitude lower than unstructured peptides and at least 2 orders of magnitude higher than well-behaved proteins. This suggests that there is a dynamic range of conformational reorganization within which partially unfolded states are prone to aggregation.

Introduction

A growing body of evidence suggests that protein aggregation, the basis for many diseases,[1−4] is a capability of all polypeptide chains.[5] However, relative scarcity of aggregants in organisms and the wide range of aggregation rates under the same conditions measured in vitro show that the process of aggregation varies tremendously by sequence.[2,6] This leads to the question of what determines a polypeptide chain's propensity to aggregate. Recent work has shown a strong correlation between aggregation rates and hydrophobicity, hydrophobic patterning, and charge of a sequence;[6−11] additionally, helix and sheet propensity correlate significantly with mutational changes in aggregation rates within a sequence.[8,12−15] From these statistical analyses, a general model of the relevant physicochemical parameters has emerged,[6,8] but it does not fully define the physical basis for how aggregation happens. One possible mechanism is that the aggregation precursor is fairly disordered and transiently exposes hydrophobic residues to the solvent for sufficient time to form bimolecular interactions. Therefore, there may be some correlation between intramolecular diffusion rates of aggregation precursors and aggregation propensity.

To begin to understand what sort of a connection might exist between intramolecular diffusion and aggregation, we have measured the rate of intramolecular contact in the unfolded state of an aggregation-prone protein, HypF-N.[16] This fragment of maturation factor of prokaryotic hydrogenase[17] has been shown to have one of the fastest aggregation rates in vitro,[6] and extensive work has been done to identify the nature of the aggregation precursor.[3,18−23] Campioni et al. have shown that the conformation populated at low pH can form prefibrillar aggregates and can be described as a premolten globule, defined by Uversky as an intermediate between a fully disordered chain and a molten globule.[24] Furthermore, ANS fluorescence measurements during folding have shown that a collapsed state formed within the first 1 ms of folding has a significant number of hydrophobic residues exposed to solvent.[25] These results support the idea that a fairly disordered state, which may even be the unfolded state under folding conditions, is the aggregation precursor for HypF-N, and therefore the rate of intramolecular diffusion should determine the rate of bimolecular association.

Materials and Methods

The HypF-N protein mutants W27F/C7S/C40S and W81F/C7S/C40S were mutated and expressed as described in ref 24. The stock solution of ∼300 μM was stored at −20 °C in sodium acetate buffer (50 mM, pH 5.5) with 1 mM TCEP. Shortly before the experiment, a 300 μL aliquot was thawed and diluted 10:1 into a cuvette containing sodium acetate buffer and the final concentrations of guanidine HCl (GdnHCl) and sucrose that had been thoroughly saturated with N_2O to scavenge solvated electrons and prevent O_2 quenching of the triplet.

The instrument to measure tryptophan triplet lifetimes has been described previously.[26] Briefly, the tryptophan triplet is populated simultaneously with the singlet state within a 10 ns pulse of light at 289 nm. The triplet population is monitored by optical absorption at 442 nm. In aqueous conditions, the triplet lives for 40 μs in the absence of quenching but can be observed to decay in as little as 100 ns for a short peptide with Trp and Cys at the ends.[26] The decay in optical absorption was detected by a silicon detector and observed on two oscilloscopes to cover a time range between 10 ns and 100 ms.

The kinetics are modeled as described previously.[26] The observed rate of triplet decay in a disordered chain should depend on three rates, the forward intramolecular diffusion rate, k_{D+}, the backward intramolecular diffusion rate, k_{D-}, and the quenching rate within a close complex, q, as follows

$$k_{obs} = \frac{k_{D+}q}{q + k_{D-}} \tag{1}$$

If $q \gg k_{D-}$, then $k_{obs} = k_{D+}$, but if that condition does not hold, then eq 1 can be rewritten as

$$\frac{1}{k_{obs}} = \frac{1}{qK} + \frac{1}{k_{D+}} = \frac{1}{k_R(T)} + \frac{1}{k_{D+}(T,\eta)} \tag{2}$$

where $K \equiv k_{D+}/k_{D-}$ is the equilibrium constant for forming the Trp-Cys encounter complex and $k_R \equiv qK$. We assume that the reaction-limited rate, k_R, depends exponentially on temperature and that the diffusion-limited rate, k_{D+}, depends on both

* Corresponding author. E-mail: lapidus@msu.edu.
[†] Michigan State University.
[‡] University of Florence.

10.1021/jp904189b
Published on Web 11/23/2009

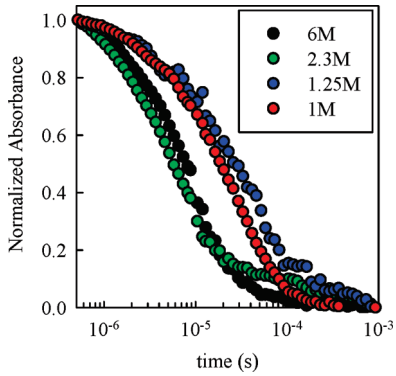

Figure 1. Kinetics of tryptophan triplet lifetime for the HypF mutant W81F at various concentrations of GdnHCl. The fast decays ($\sim 10^5$ s^{-1}) are due to intramolecular contact quenching by C65, and the slow decays ($\sim 10^4$ s^{-1}) are due to the natural lifetime of the triplet in an aqueous environment.

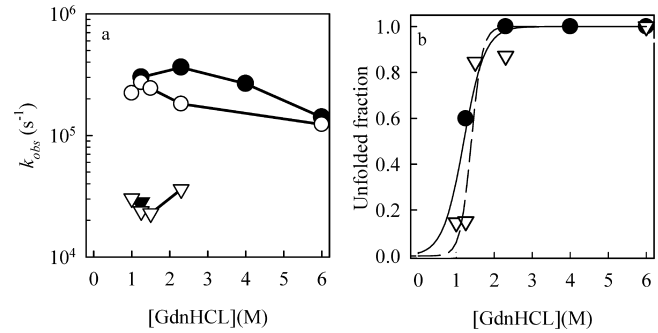

Figure 2. (a) Observed exponential rates of the tryptophan triplet decay versus [GdnHCl] for the two mutants, W81F (white) and W27F (black). The triangles are the slow rate, and the circles are the fast rate as illustrated in Figure 1. (b) The relative amplitude of the fast rate vs [GdnHCl] for W27F (black circles) and W81F (white triangles). The black lines are two-state fits to equilibrium fluorescence measurements with the parameters (W81F, solid line) $m = 9.1$ kJ/mol/M, $C_m = 1.2$ M and (W27F, dashed line) $m = 17.4$ kJ/mol/M, $C_m = 1.4$ M.

temperature and viscosity.[27] Therefore, a plot of $1/k_{obs}$ vs viscosity (η) (increased by the addition of sucrose) at a constant temperature is a line with the intercept equal to $1/k_R$ and the slope equal to $1/\eta k_{D+}$. We model all the data at a single denaturant concentration globally using the following relations[27]

$$k_R(T) = k_{R0} \exp\left(\frac{E_0(T - T_0)}{RTT_0}\right) \quad (3)$$

$$k_{D+}(T, \eta) = \frac{k_{D+0}T}{\eta T_0} \exp(\gamma(T - T_0)) \quad (4)$$

where k_{R0}, k_{D+0}, E_0, and γ are fitting parameters and $T_0 = 273$ K. To model each reaction-limited and diffusion-limited rate with eqs 5 and 6 (below), we use a wormlike chain model with excluded volume to generate a probability distribution ($P(r)$ in eqs 5 and 6) of Trp-Cys distances. Two million wormlike chains were numerically generated as described previously[28] for each mutant and chain parameters, and the measurement of the Trp-Cys distance in each chain is combined into a normalized histogram.

Results

The wild-type sequence of HypF-N contains three cysteines at positions 7, 40, and 65 and two tryptophans at positions 27 and 81. Two mutants were made, W27F/C7S/C40S (referred to as W27F) and W81F/C7S/C40S (W81F), to remove two cysteines and one of these tryptophans so that intramolecular contact is observed between positions 27 and 65 (long loop) and positions 65 and 81 (short loop) independently. The method of Trp-Cys contact quenching exploits the fact that tryptophan can be excited to a long-lived ($\tau > 40$ μs) triplet state that is only significantly quenched by cysteine upon close ($r \sim 4$ Å) contact.[26,29] For both mutants at every concentration of denaturant, we observe one or two exponential decays of the tryptophan triplet state (see Figure 1). The fast rate represents intramolecular contact of the unfolded chain, and the slow rate is the natural lifetime of the triplet state without contact. The relative amplitude of the fast rate gives the equilibrium population of the unfolded state. Figure 1 shows that the relative amplitude of the slow rates (fraction of folded states) increases with decreasing concentration of denaturant. Figure 2 shows the observed contact rates (a) and fraction of the unfolded state (b) for both W81F and W27F. From Figure 2b, we conclude the midpoint of folding for HypF is at about 1.25 M GdnHCL, in reasonable agreement with that measured by equilibrium fluorescence. We also note that the slow rates for both mutants are similar to that observed for free Trp in water ($\sim 10^4$ s^{-1}).[26] This suggests that the tryptophans in both mutants are still solvated in their folded or intermediate states, not buried in the hydrophobic core of the folded protein.

Figure 3 shows the measured decay time as a function of viscosity for different concentrations of GdnHCl. Each temperature is plotted as a different color, and the viscosity is increased by the addition of sucrose. For both mutants at 6 M GdnHCl, each temperature lies on a different line, indicating that k_R depends strongly on temperature. This behavior has been seen for other proteins in strong denaturant and unstructured peptides in all solvent conditions.[28,30] However, at 1.25 M GdnHCl, W81F still shows separate lines for each temperature, while W27F has converged on a single line for all temperatures, suggesting both k_R and k_{D+} have lost any dependence on temperature. Earlier work on proteins L and G showed a loss of temperature dependence on k_R at 2.3 M GdnHCl which was argued to be due to the fact that the equilibrium rate, K, had increased far above 1 at low denaturant.[28] That may still be the case for W27F, but we must also conclude that k_{D+} has lost any dependence on temperature as well. Thus the 65−81 (short) loop appears to be highly collapsed at low denaturant, and the intramolecular diffusion rate is slow enough to be independent of temperature.

Figure 4 shows the reaction-limited (black circles) and diffusion-limited (white triangles) rates for both mutants. Between 6 and 2.3 M GdnHCl, the rates for both mutants increase slightly, but below 2 M GdnHCl, the diffusion-limited rates start to decrease, more significantly for W27F than for W81F. For the W81F (long loop) mutant, the reaction-limited rate continues to increase, while the W27F (short loop) mutant has no measurable reaction-limited rate and is presumed to be greater than 1×10^7 s^{-1}. Measurements of k_R and k_{D+} were not possible below 1.25 M GdnHCl because the addition of sucrose stabilized the folded state of the protein, preventing the observation of quenching in the unfolded state. However, an extrapolation of k_R and k_{D+} of W81F to 0 M GdnHCl in Figure 4 shows that both of these rates are $5-6 \times 10^5$ s^{-1}.

To determine the underlying polymeric dynamics that produces the observed contact rates, we use a theory by Szabo, Schulten, and Schuten (SSS), which models intramolecular diffusion as diffusion on a one-dimensional potential of mean force. The reaction-limited and diffusion-limited rates are given by[27,31]

$$k_R = \int_a^\infty q(r)P(r)dr \quad (5)$$

$$\frac{1}{k_{D+}} = \frac{1}{k_R^2 D}\int_a^{l_c}\frac{dr}{P(r)}\{\int_r^{l_c}(q(x) - k_R)P(x)dx\}^2 \quad (6)$$

where r is the distance between the tryptophan and cysteine and $P(r)$ is the probability density of finding the polypeptide at that distance. D is the effective intramolecular diffusion coefficient; a is the distance of closest approach (defined to be 4 Å); l_c is the contour length of the peptide; and $q(r)$ is the distance-dependent quenching rate. The distance-dependent quenching rate for the Trp-Cys system has been determined experimentally[32] and drops off exponentially beyond 4 Å, so the reaction-limited rate is mostly determined by the probability

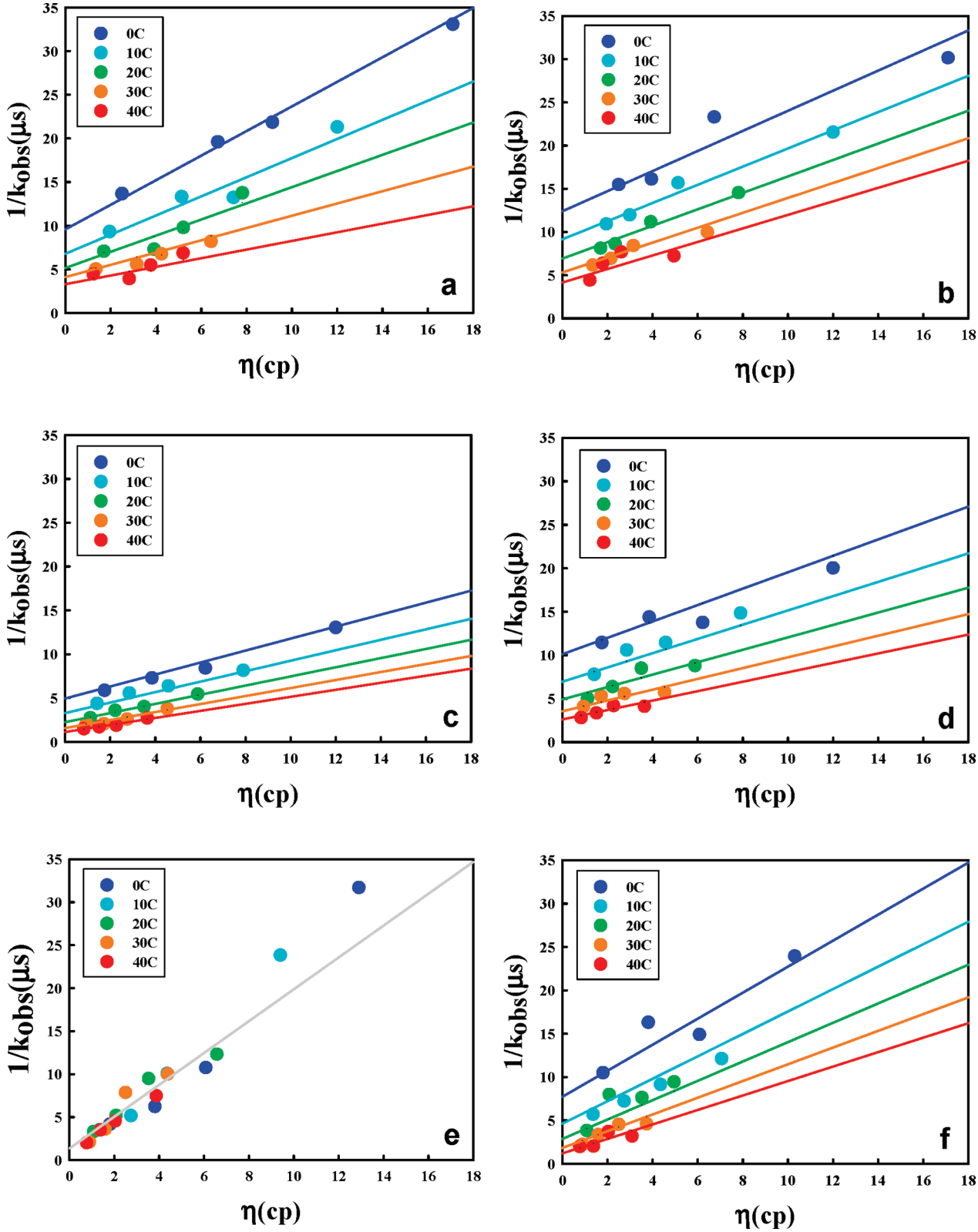

Figure 3. Observed fast rates vs viscosity for W27F (a,c,e) and W81F (b,d,f) in 6 M (a,b), 2.3 M (c,d), and 1.25 M (e,f) GdnHCl. Each temperature is plotted as a separate color. The lines are fits of all the data in each plot to eqs 3 and 4.

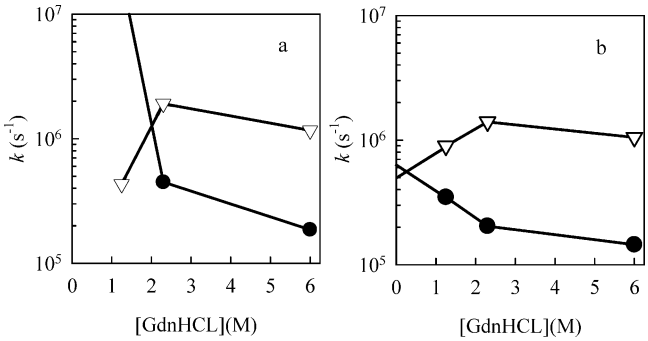

Figure 4. Reaction-limited (black circles) and diffusion-limited (white triangles) rates vs [GdnHCl] for (a) W27F and (b) W81F at 20 °C and 1 cP.

TABLE 1: Wormlike Chain Parameters and Effective Diffusion Coefficients for Both Mutants Studied

[GdnHCL]	k_R (s^{-1})	k_{D+} (s^{-1})	l_p (Å)	d_α (Å)	<r> (Å)	D ($\times 10^{-6}$ cm^2 s^{-1})
		W81F				
0	600000[a]	500000[a]	4	2.3	34.0	0.2
1.25	347950	896190	4	2.75	35.5	0.51
2.3	203920	1401100	4	3.20	37.0	1.15
6	144830	1049700	4	3.50	38.1	1.07
		W27F				
1.25	4.8×10^{14}	433088	4	0	19.9	0.02
2.3	448200	1913100	4	3.20	24.2	0.58
6	185960	1167500	4	4.10	25.1	0.47

[a] Extrapolated data.

of the shortest distances. Thus the only free parameters in the equations above are D and $P(r)$. Using a wormlike chain model with excluded volume,[28,30] we have found a $P(r)$ that best matches the experimental reaction-limited rate measured for each mutant at each denaturant concentration using eq 5. As with earlier work,[28] we find that the reaction-limited rate is relatively insensitive to the persistence length, so it is kept fixed at l_p = 4 Å, and the excluded diameter (d_α) is varied with denaturant concentration. The parameters of these distributions are listed in Table 1. Then the appropriate $P(r)$ and experimental k_{D+} is used in eq 6 to determine an effective diffusion coefficient, D, also listed in Table 1. Remarkably, between fully denaturing conditions and native conditions, D only decreases by a factor of ~5 for W81F.

Discussion

In this work, we have observed intramolecular diffusion of the unfolded state of HypF under various concentrations of denaturant. Previous use of this technique in the B1 domains of proteins L and G found that: (1) the reaction-limited and diffusion-limited rates monotonically increased and decreased, respectively, with decreasing denaturant; (2) the same qualitative behavior was observed for two different loops in the chain; and (3) the same quantitative behavior was observed for the same loop in both proteins.[28] These observations led to the conclusions that both these homologous proteins are globally compact and have low diffusivity throughout the chain. In contrast, HypF does not follow these rules. While the reaction-limited rate does monotonically increase as denaturant decreases, the diffusion-limited rate slightly increases from 6 M GdnHCl to 2.3 M GdnHCl and then begins to decrease. Below 2.3 M GdnHCl, the degree of compaction and loss of diffusivity is different in separate loops in the chain. The C65−W81 loop compacts and loses diffusivity significantly more than the W27−C65 loop, which only slows down by about a factor of 2. The equilibrium fluorescence of each tryptophan (see Figure 2b) also shows that the folding of Trp81 is much more cooperative than the folding of Trp27. The W27−C65 loop encompasses one of the putative amyloidogenic regions of the protein, based on mutational studies of the homologous protein acylphosphatase, while the C65−W81 loop does not;[20,33] therefore, these results suggest that the amyloidgenic region of the chain is more diffusive than surrounding regions.

What is the physical basis for the increased diffusivity of the long loop? This could be due to the fact that there is significant net charge in the long loop, while the short loop, like the whole chain, is almost neutral. This would force the long loop to adopt more extended conformations. However, the long loop still contains many hydrophobic residues. Using the Kyle−Dolittle scale, the mean hydrophobicity (−0.41) and the standard deviation (3.36) of the long loop are almost the same as the short loop or the entire chain. Also, the hydrophobic pattern of the entire chain oscillates rapidly, with no large clusters of hydrophobic residues to direct collapse. Therefore, the W27−C65 region of the chain, forced into extended conformations by electrostatic repulsions that are not counterbalanced by large hydrophobic clusters pulling the chain together, will have hydrophobes more often exposed to solvent and other, nearby chains making bimolecular hydrophobic interactions more likely.

Thus, from comparing the intramolecular diffusion measurements of HypF with proteins L and G, there appears to be two modes of dynamics for unfolded proteins under conditions that favor folding: (1) The chain compacts to a highly nondiffusive conformation, presumably due to numerous non-native hydrophobic interactions, and (2) the chain remains fairly diffusive. The C65−W81 loop falls into the first category, along with proteins L and G, and the W27−C65 loop falls into the second category. One might assume from this conclusion that there is a linear relation between diffusivity and aggregation propensity, but much work on unstructured peptides shows that there is an extremely fast regime that also avoids aggregation.[27,30,34,35] Therefore, we propose that there is a dangerous dynamic range in which a freely diffusing chain reconfigures just fast enough to expose hydrophobic residues to solvent long enough to form bimolecular associations. Polypeptides that diffuse faster do not leave hydrophobes exposed for long, destabilizing bimolecular complexes, and well-behaved proteins diffuse slow enough that most hydrophobes are exposed to solvent infrequently. This is illustrated by a simple kinetic model

$$A \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} A^*$$

$$A^* + A^* \xrightarrow{k_{bi}} [A^*A^*] \xrightarrow{k_{agg}} \text{Agg}$$

$$[A^*A^*] \xrightarrow{k_{-1}} A + A^*$$

where A is the unfolded ensemble of conformations and A^* is the conformational subset that can make bimolecular associations. $[A^*A^*]$ is the bimolecular encounter complex that can lead to irreversible aggregation, Agg, but can also be returned to monomers by conversion of A^* into A. The intramolecular reconfiguration rates, k_1 and k_{-1}, are related to the intramolecular diffusion coefficient, D. Keeping k_{bi}, k_{agg}, and $K \equiv k_1/k_{-1}$ constant, Figure 5 shows the formation of Agg is substantially slower when k_1 is much higher or much lower than k_{bi} than when they are about the same. Proof of this hypothesis will certainly require more measurements of aggregation-prone

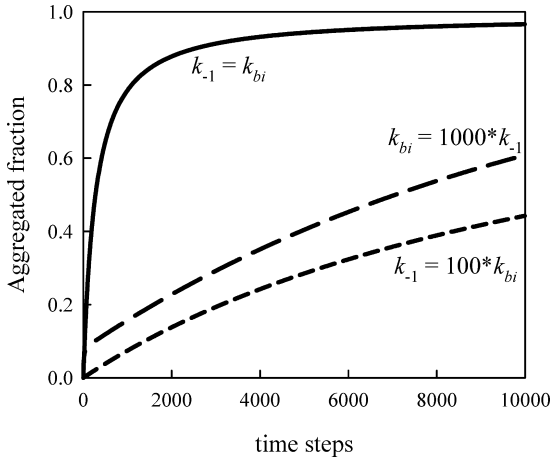

Figure 5. Formation of aggregated species, Agg, according to the model outlined in the text. For these simulations, $k_{bi} = k_{agg} = 1$, $K = 0.1$, and k_{-1} varies as indicated on the plot.

sequences, but a more complete understanding of hydrophobic encounter complexes between proteins will also help to define the limits of this dynamical regime of aggregation.

Acknowledgment. We thank Fabrizio Chiti and William Wedemeyer for helpful discussions. Niccolò Taddei thanks Italian MIUR for support (grant PRIN 2007B57EAB_005). This work was partially supported by NSF grant MCB-0825001. The research of Lisa Lapidus, Ph.D. is supported in part by a Career Award at the Scientific Interface from the Burroughs Wellcome Fund.

References and Notes

(1) Dobson, C. M. *Nat. Rev. Drug Discovery* **2003**, *2*, 154.
(2) Stefani, M.; Dobson, C. M. *J. Mol. Med.-Jmm* **2003**, *81*, 678.
(3) Uversky, V. N.; Fink, A. L. *Biochim. Biophys. Acta-Proteins Proteom.* **2004**, *1698*, 131.
(4) Selkoe, D. J. *Nature* **2003**, *426*, 900.
(5) Chiti, F.; Taddei, N.; Baroni, F.; Capanni, C.; Stefani, M.; Ramponi, G.; Dobson, C. M. *Nat. Struct. Biol.* **2002**, *9*, 137.
(6) Dubay, K. F.; Pawar, A. P.; Chiti, F.; Zurdo, J.; Dobson, C. M.; Vendruscolo, M. *J. Mol. Biol.* **2004**, *341*, 1317.
(7) Dobson, C. M. *Semin. Cell Dev. Biol.* **2004**, *15*, 3.
(8) Chiti, F.; Stefani, M.; Taddei, N.; Ramponi, G.; Dobson, C. M. *Nature* **2003**, *424*, 805.
(9) Calamai, M.; Taddei, N.; Stefani, M.; Ramponi, G.; Chiti, F. *Biochemistry* **2003**, *42*, 15078.
(10) Schmittschmitt, J. P.; Scholtz, J. M. *Protein Sci.* **2003**, *12*, 2374.
(11) Zbilut, J. P.; Mitchell, J. C.; Giuliani, A.; Colosimo, A.; Marwan, N.; Webber, C. L. *Phys. A* **2004**, *343*, 348.
(12) Chiti, F.; Mangione, P.; Andreola, A.; Giorgetti, S.; Stefani, M.; Dobson, C. M.; Bellottl, V.; Taddei, N. *J. Mol. Biol.* **2001**, *307*, 379.
(13) Tjernberg, L.; Hosia, W.; Bark, N.; Thyberg, J.; Johansson, J. *J. Biol. Chem.* **2002**, *277*, 43243.
(14) Hosia, W.; Bark, N.; Liepinsh, E.; Tjernberg, A.; Persson, B.; Hallen, D.; Thyberg, J.; Johansson, J.; Tjernberg, L. *Biochemistry* **2004**, *43*, 4655.
(15) Calamai, M.; Chiti, F.; Dobson, C. M. *Biophys. J.* **2005**, *89*, 4201.
(16) Rosano, C.; Zuccotti, S.; Bucciantini, M.; Stefani, M.; Ramponi, G.; Bolognesi, M. *J. Mol. Biol.* **2002**, *321*, 785.
(17) Colbeau, A.; Elsen, S.; Tomiyama, M.; Zorin, N. A.; Dimon, B.; Vignais, P. M. *Eur. J. Biochem.* **1998**, *251*, 65.
(18) Kelly, J. W. *Curr. Opin. Struct. Biol.* **1998**, *8*, 101.
(19) Guijarro, J. I.; Sunde, M.; Jones, J. A.; Campbell, I. D.; Dobson, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4224.
(20) Chiti, F.; Taddei, N.; Bucciantini, M.; White, P.; Ramponi, G.; Dobson, C. M. *EMBO J.* **2000**, *19*, 1441.
(21) Ramirez-Alvarado, M.; Merkel, J. S.; Regan, L. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8979.
(22) Torrent, J.; Balny, C.; Lange, R. *Protein Pept. Lett.* **2006**, *13*, 271.
(23) Dumoulin, M.; Kumita, J. R.; Dobson, C. M. *Acc. Chem. Res.* **2006**, *39*, 603.
(24) Campioni, S.; Mossuto, M. F.; Torrassa, S.; Calloni, G.; de Laureto, P. P.; Relini, A.; Fontana, A.; Chiti, F. *J. Mol. Biol.* **2008**, *379*, 554.
(25) Calloni, G.; Taddei, N.; Plaxco, K. W.; Ramponi, G.; Stefani, M.; Chiti, F. *J. Mol. Biol.* **2003**, *330*, 577.
(26) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 7220.
(27) Lapidus, L. J.; Steinbach, P. J.; Eaton, W. A.; Szabo, A.; Hofrichter, J. *J. Phys. Chem. B* **2002**, *106*, 11628.
(28) Singh, V. R.; Kopka, M.; Chen, Y.; Wedemeyer, W. J.; Lapidus, L. J. *Biochemistry* **2007**, *46*, 10046.
(29) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *Phys. Rev. Lett.* **2001**, *87*, 4.
(30) Buscaglia, M.; Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *Biophys. J.* **2006**, *91*, 276.
(31) Szabo, A.; Schulten, K.; Schulten, Z. *J. Chem. Phys.* **1980**, *72*, 4350.
(32) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. *Phys. Rev. Lett.* **2001**, 87.
(33) Chiti, F.; Taddei, N.; White, P. M.; Bucciantini, M.; Magherini, F.; Stefani, M.; Dobson, C. M. *Nat. Struct. Biol.* **1999**, *6*, 1005.
(34) Buckler, D. R.; Haas, E.; Scheraga, H. A. *Biochemistry* **1995**, *34*, 15965.
(35) Hagen, S. J.; Hofrichter, J.; Szabo, A.; Eaton, W. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 11615.

JP904189B