J|A|C|S
ARTICLES
Published on Web 07/23/2002

Fast Folding of a Four-Helical Bundle Protein

Neelan J. Marianayagam, Farid Khan, Louise Male, and Sophie E. Jackson*

Contribution from the University of Cambridge, Centre for Protein Engineering, Department of Chemistry, Lensfield Road, Cambridge, CB2 1EW, United Kingdom

Received June 22, 2001. Revised Manuscript Received May 29, 2002

Abstract: The FK506−FKBP12 binding-domain of the kinase FRAP (FRB) forms a classic up−down four-helical bundle. The folding pathway of this protein has been investigated using a combination of equilibrium and kinetic studies. The native state of the protein is stable with respect to the unfolded state by some 7 kcal mol^{-1} at pH 6.0, 10 °C. A kinetic analysis of unfolding and refolding rate constants as a function of chemical denaturant concentration suggests that an intermediate state may be populated during folding at low concentrations of denaturant. The presence of this intermediate state is confirmed by refolding experiments performed in the presence of the hydrophobic dye 8-anilinonaphthalene-1 sulfonate (ANS). ANS binds to the partially folded intermediate state populated during the folding of FRB and undergoes a large change in fluorescence that can be detected using stopped-flow techniques. Analysis of the kinetic data suggests that the intermediate state is compact and it may even be a misfolded species that has to partially unfold before it can reach the transition state. Folding and unfolding rate constants in water are approximately 150−200 s^{-1} and 0.005−0.06 s^{-1}, respectively, at neutral pH and 10 °C. The folding of FRB is somewhat slower than for other all-helical proteins, probably as a consequence of the formation of a metastable intermediate state. The folding rate constant in the absence of any populated intermediate can be estimated to be 8800 s^{-1}. Despite the presence of an intermediate state, which effectively slows folding, the protein still folds rapidly with a half-life of 5 ms at 10 °C. The dependence of the rate constants on denaturant concentration indicates that the transition state for folding is compact with some 80% of the surface area exposed in the unfolded state buried in the transition state. Data presented for FRB is compared with kinetic data obtained for other all-helical proteins.

Introduction

Folding studies on proteins consisting entirely of α-helices have been of great interest over the past few years. One of the best characterized systems is cytochrome *c*, which adopts a three-helical folded-leaf motif with a heme group covalently bound between the helices in the core of the protein. The folding pathway of holocytochrome *c* has been studied using a variety of experimental approaches.[1−7] Initial experiments showed that the protein populated an intermediate state during folding; however, later experiments established that this was a misfolded state resulting from a non-native ligand interaction with the heme group.[8] Subsequent experiments, under conditions where this non-native ligation was prevented, showed that cytochrome *c* folded with simple two-state kinetics. Recently, however, ultrafast rapid mixing devices and laser-induced temperature-jump methods have shown that there is an additional very fast refolding phase not detected in previous studies.[9,10] It was initially suggested that this fast phase represented the rapid collapse of the polypeptide chain associated with a change in the unfolded state on dilution of denaturant; however, further experiments have established that there is a thermal barrier associated with this phase, leading to the conclusion that this phase represents the formation of an intermediate state.[9,10] In addition to cytochrome *c*, the folding of other heme-containing cytochromes have also been studied. The folding of both holo and apo forms of cytochrome b_5 has been studied by the Whitford group.[11,12] Whereas, the folding of the apo form is simple and follows two-state kinetics,[11] the folding of the holo form is complex, with several kinetic phases in both the unfolding and folding reaction.[12] The folding of cytochrome b_{562}, which forms a four-helical bundle in which the heme group is noncovalently bound to the protein, has been studied using electron-transfer-initiated folding techniques.[13,14] These studies

* To whom correspondence should be addressed. Tel: (44) 1223 762011, Fax: (44) 1223 336362. E-mail: sej13@cam.ac.uk.

(1) Roder, H.; Elöve, G. A.; Englander, S. W. *Nature* **1988**, *335*, 694−699.
(2) Elove, G. A.; Chaffotte, A. F.; Roder, H.; Goldberg, M. E. *Biochemistry* **1992**, *31*, 6876−6883.
(3) Mines, G. A.; Pascher, T.; Lee, S. C.; Winkler, J. R.; Gray, H. B. *Chem. Biol.* **1996**, *3*, 491−497.
(4) Pascher, T.; Chesick, J. P.; Winkler, J. R.; Gray, H. B. *Science* **1996**, *271*, 1558−1560.
(5) Chan, C. K.; Hu, Y.; Takahashi, S.; Rousseau, D. L.; Eaton, W. A.; Hofrichter, J. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 1779−1784.
(6) Shastry, M. C. R.; Sauder, J. M.; Roder, H. *Acc. Chem. Res.* **1998**, *31*, 717−725.
(7) Sauder, J. M.; Roder, H. *Folding Des.* **1998**, *3*, 385−392.
(8) Elove, G. A.; Bhuyan, A. K.; Roder, H. *Biochemistry* **1994**, *33*, 6925−6935.
(9) Hagen, S. J.; Eaton. W. A. *J. Mol. Biol.* **2000**, *297*, 781−789.
(10) Shastry, M. C. R.; Roder, H. *Nat. Struct. Biol.* **1998**, *5*, 385−392.
(11) Manyusa, S.; Whitford, D. *Biochemistry* **1999**, *38*, 9533−9540.
(12) Manyusa, S.; Mortuza, G.; Whitford, D. *Biochemistry* **1999**, *38*, 14352−14362.
(13) Wittung-Stafshede, P.; Winkler, J. R.; Gray, H. B. *J. Am. Chem. Soc.* **1997**, *119*, 9562−9563.
(14) Wittung-Stafshede, P.; Lee, J. C.; Winkler, J. R.; Gray, H. B. *Proc. Natl Acad. Sci. U.S.A.* **1999**, *96*, 6587−6590.

10.1021/ja016480r CCC: $22.00

showed that the protein folds very rapidly in the presence of heme. H/D exchange studies on apocytochrome b_{562} under equilibrium conditions suggest that the two central α-helices are significantly more stable than the N- and C-terminal helices and that these may form early on the folding pathway.[15] It had earlier been proposed that the apo form of cytochrome b_{562} formed a molten globule-like state, on the basis of solution structural studies.[16] It was speculated that this molten-globule state might be representative of an intermediate state during folding; however, molecular dynamic simulations on the native structure showed that the structure remains highly ordered even in the absence of heme.[17] The folding of cyctochrome *c*′, which also adopts a four-helical bundle structure and which contains a covalently bound heme group, has been studied following triggering of the folding reaction by electron transfer.[18] In this case, the refolding reaction is complex, involving fast and slow phases that have not been definitively defined. In general, it is not clear what effect the heme group has in influencing the folding pathway.

The folding of several non-heme-containing all-helical proteins has also been studied. For example, the folding of a monomeric form of the λ repressor has been studied by NMR line-shape analysis[19] and fluorescence stopped-flow spectroscopy.[20] Together with results from protein engineering and other studies,[21,22] Oas and co-workers have used a diffusion-collision model[23] to calculate the theoretical rate constants for folding of this protein. These are in good agreement with experimental data.[24] NMR spectroscopy and protein engineering techniques have also been used to study the folding of the four-helical bundle protein ACBP.[25−27] In this case, folding is dominated by eight conserved hydrophobic residues that are involved in structure formation during the rate-determining folding step.[27] Recently, the Poulsen group have gone on to show that some amide protons are protected against hydrogen exchange very early in folding, and thus, hydrogen bond formation may precede the rate-limiting step for this protein.[28] Studies on the immunity proteins, Im7 and Im9, have shown that both of these small helical proteins fold rapidly, and both populate intermediate states under suitable experimental conditions.[29−31] Recently, it was established through ultrarapid mixing experiments and ϕ-value analysis that the intermediate state for Im7 is on-pathway and that it is a misfolded species with significant non-native interactions.[32−33] Recently, the all-helical engrailed homeodomain (EnHd) has been shown to fold with rate constants on the order of 10^4−10^5 s^{-1}.[34]

In addition to studies on the folding pathways of small all-helical proteins, the folding pathways of larger helical structures have also been studied. By far the best characterized is the folding pathway of apomyoglobin. The Dyson and Wright groups have performed extensive studies on the folding of this protein and shown that it folds via an obligatory intermediate state in which helical structure is stabilized in the A, G, and H helices and in part of the B helix.[35−38] They have also shown that apomyoglobin forms an equilibrium molten-globule intermediate that is maximally populated at pH 4 and is similar in structure to the kinetic intermediate.[35−38]

A variety of computational approaches for studying folding that use both simplified coarse-grained models and more realistic all-atom models have been developed.[39,40] Due to the complexity of all-atom molecular dynamics simulations, these calculations have to be run at high temperature. Comparison with experimental results, normally acquired at or near room temperature, is therefore problematic. The characterization of small helical proteins that unfold and fold extremely rapidly, for example engrailed homeodomain, enable rate constants to be easily extrapolated to temperatures at which MD simulations are performed.[34] The experimental characterization of the folding pathways of small proteins that fold rapidly is therefore of importance.

The up−down four-helical bundle motif is one of the simplest folds observed in naturally occurring proteins. As a result, there has been much work on the de novo design of this type of structure.[41−43] Ho and DeGrado have designed a four-helical bundle, similar in fold to cytochrome *c*′, comprising of two amphiphilic 16-residue peptides that tetramerized to give the final structure.[41] Guo and Thirumalai have used Langevin dynamics simulations of this protein using an off-lattice model containing the same pattern of hydrophobic and hydrophilic residues.[44] Simulations showed a kinetic partitioning mechanism in which a fraction of the molecules reach the folded state rapidly through a nucleation-collapse mechanism, while the rest of the molecules fold more slowly, populating an intermediate state.[44] Given their simple fold and ubiquitous nature, there is surprisingly little experimental information on the folding of four-helical bundle proteins without prosthetic groups. Here, we present a study of the folding pathway of the four-helical bundle protein FRB. This is a small binding domain from the

(15) Fuentes, E. J.; Wand, A. J. *Biochemistry* **1998**, *37*, 3687−3698.
(16) Feng, Y.; Sligar, S. G.; Wand, A. J. *Nat. Struct. Biol.* **1994**, *1*, 30−35.
(17) Laidig, K. E.; Daggett, V. *Folding Des.* **1996**, *1*, 335−346.
(18) Lee, J. C.; Gray, H. B.; Winkler, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 7760−7764.
(19) Huang, G. S.; Oas, T. G. *Biochemistry* **1995**, *34*, 3884−3892.
(20) Ghaemmaghami, S.; Word, J. M.; Burton, R. E.; Richardson, J. S.; Oas, T. G. *Biochemistry* **1998**, *37*, 9179−9185.
(21) Burton, R. E.; Huang, G. S.; Daugherty, M. A.; Fullbright, P. W.; Oas, T. G. *J. Mol. Biol.* **1996**, *263*, 311−322.
(22) Burton, R. E.; Huang, G. S.; Daugherty, M. A.; Calderone, T. L.; Oas, T. G. *Nat. Struct. Biol.* **1997**, *4*, 305−310.
(23) Karplus, M.; Weaver. D. C. *Nature* **1976**, *260*, 404−406.
(24) Burton, R. E.; Myers, J. K.; Oas, T. G. *Biochemistry* **1998**, *37*, 5337−5343.
(25) Kragelund, B. B.; Robinson, C. V.; Knudsen, J.; Dobson, C. M.; Poulsen, F. M. *Biochemistry* **1995**, *34*, 7217−7224.
(26) Kragelund, B. B.; Hojrup, P.; Jensen, M. S.; Scherling, C. K.; Juul, E.; Knudsen, J.; Poulsen, F. M. *J. Mol. Biol.* **1996**, *256*, 187−200.
(27) Kragelund, B. B.; Osmark, P.; Neergaard, T. B.; Schiodt, J.; Kristiansen, K.; Knudsen, J.; Poulsen, F. M. *Nat. Struct. Biol.* **1999**, *6*, 594−601.
(28) Teilum, K.; Kragelund, B. B.; Knudsen, J.; Poulsen, F. M. *J. Mol. Biol.* **2000**, *301*, 1307−1314.
(29) Ferguson, N.; Capaldi, A. P.; James, R.; Kleanthous, C.; Radford, S. E. *J. Mol. Biol.* **1999**, *286*, 1597−1608.
(30) Gorski, S. A.; Capaldi, A. P.; Kleanthous, C.; Radford, S. E. *J. Mol. Biol.* **2001**, *312*, 849−863.
(31) Ferguson, N.; Li, W.; Capaldi, A. P.; Kleanthous, C.; Radford, S. E. *J. Mol. Biol.* **2001**, *307*, 393−405.
(32) Capaldi, A. P.; Kleanthous, C.; Radford, S. E. *Nat. Struct. Biol.* **2002**, *9*, 209−216.
(33) Capaldi, A. P.; Shastry, M. C. R.; Kleanthous, C.; Roder, H.; Radford, S. E. *Nat. Struct. Biol.* **2001**, *8*, 68−72.
(34) Mayor, U.; Johnson, C. M.; Daggett, V.; Fersht, A. R. *Proc. Natl Acad. Sci. U.S.A.* **2000**, *97*, 13518−13522.
(35) Jennings, P. A.; Wright, P. E. *Science* **1993**, *262*, 892−896.
(36) Jamin, M.; Baldwin, R. L. *J. Mol. Biol.* **1998**, *276*, 491−504.
(37) Tsui, V.; Garcia, C.; Cavagnero, S.; Siuzdak, G.; Dyson, H. J.; Wright, P. E. *Protein Sci.* **1999**, *8*, 45−49.
(38) Eliezer, D.; Yao, J.; Dyson, H. J.; Wright, P. E. *Nat. Struct. Biol.* **1998**, *5*, 148−155.
(39) Dobson, C. M.; Sali, A.; Karplus, M. *Angew. Chem.* **1998**, *37*, 868−893.
(40) Shea, J.-E.; Brooks, C. L., III. *Annu. Rev. Phys. Chem.* **2001**, *52*, 499−535.
(41) Ho, S. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **1987**, *109*, 6751−6758.
(42) Betz, F. B.; Liebman, P. A.; DeGrado, W. F. *Biochemistry* **1997**, *36*, 2450−2458.
(43) Schafmeister, C. E.; LaPorte, S. L.; Miercke, L. J. W.; Stroud, R. M. *Nat. Struct. Biol.* **1997**, *4*, 1039−1045.
(44) Guo, Z.; Thirumalai, D. *J. Mol. Biol.* **1996**, *263*, 323−343.

Figure 1. Ribbon diagram representation of the structure of the four-helical bundle protein FRB. The four tryptophan residues that give rise to the fluorescence of the protein and are used to monitor folding are shown in stick representation.

large, 289 kDa S6 kinase FRAP—the target of the binary complex formed between FKBP12 and rapamycin.[45] The structure of the 95-residue FRB domain is shown in Figure 1.[46] It has an up−down motif similar to that of cytochrome b_5 and tobacco mosaic virus coat protein.

Here, we present a characterization of the folding pathway of FRB. Experiments conducted under equilibrium and non-equilibrium conditions, which measure the stability, unfolding, and refolding rate constants of FRB, are presented. We establish that an intermediate state is populated on the folding pathway at low concentrations of chemical denaturants. Results on the folding of FRB are compared to those obtained for other all-helical proteins.

Experimental Section

Materials. Guanidinium chloride (GdnHCl) was purchased from Melford Laboratories. All other materials were analytical grade and purchased from Sigma.

Expression and Purification of FRB. DNA encoding FRB was cloned into both pGEX (Pharmacia) and pRSET (Invitrogen) vectors. The protein was expressed from these vectors in *Escherichia coli* as an N-terminal fusion with either glutathione-S-transferase or a hexa-histidine affinity tag. Standard methods were used to express these fusion proteins at either 25 or 30 °C (see Pharmacia and Invitrogen manuals). Fusion protein was purified using affinity chromatography and the affinity tag removed by cleavage with thrombin (Sigma). The cleaved protein was then purified using FPLC on a Superdex G75 gel-filtration column (Amersham-Pharmacia). The resulting protein was pure as judged by SDS−PAGE and the molecular weight confirmed by mass spectrometry. Pure protein was flash frozen and stored at −80 °C.

Equilibrium Studies*:* FRB was reversibly unfolded using guanidinium chloride. A 7.8 M stock solution of GdnHCl was made and serially diluted using a Hamilton Microlab to give 60 different concentrations ranging from 0 to 5.9 M. FRB was added to each sample to give a final protein concentration of 1 μM, final buffer conditions were 50 mM TrisHCl, 1 mM DTT, pH 7.5. Samples were equilibrated at the desired temperature for at least 2 h before measurements were taken. The fluorescence of each sample was measured using a SLM Aminco Bowman spectrometer. The excitation wavelength was 280 nm, and the emission was monitored between 300 and 400 nm.

The program Kaleidagraph was used to fit the fluorescence data at 330 nm (the maximal difference in fluorescence between the folded and unfolded states) to eq 1

$$F = \frac{(\alpha_F + \beta_F[D]) + (\alpha_U + \beta_U[D]\exp\{m_{U-F}([D] - [D]_{50\%})/RT\}}{1 + \exp\{m_{U-F}([D] - [D]_{50\%})/RT\}} \quad (1)$$

where F is the fluorescence signal, α_F is the spectroscopic signal at 0 M denaturant for the folded state, and β_F is the slope of the folded baseline (α_U and β_U are the corresponding values for the unfolded state). The $[D]_{50\%}$ is the denaturant concentration at which 50% of the protein is unfolded. From the best fit of the data to eq 1, m_{U-F} and $[D]_{50\%}$ were obtained with their standard errors.

Kinetic Studies. All kinetic experiments were carried out on an Applied Photophysics stopped-flow reaction analyzer (Model SF.17MV) in fluorescence mode. An excitation wavelength of 280 nm was used and the fluorescence above 320 nm measured. The dead-time of the stopped-flow apparatus was determined using a *N*-bromosuccinimide (NBS) quenching assay[47] under identical conditions to those employed in the folding studies and was typically found to be 2 ms. All refolding data were fit within appropriate limits.

Unfolding Experiments. A stock solution of folded FRB in native-buffer was mixed rapidly in a ratio of 1:10 with an unfolding solution containing high concentrations of chemical denaturant. The final protein concentration was 2 μM FRB, and final buffer conditions were either 50 mM MES, 1 mM DTT, pH 6.0 or 50 mM Tris, 1 mM DTT, pH 7.5. The final denaturant concentration was varied in 0.5 M increments from 3 to 7 M GdnHCl.

Unfolding is a single-exponential process, and the kinetic traces were fit to a single-exponential function plus a linear drift, a result of baseline instability, using the Applied Photophysics software. Rate constants obtained from this fitting procedure were usually the average from at least four measurements.

Refolding Experiments. A stock solution of FRB, unfolded in 4 M GdnHCl, was mixed in a 1:10 ratio with refolding buffer. Final protein concentrations were 2 μM, and final buffer conditions were either 50 mM MES, 1 mM DTT, pH 6.0 or 50 mM Tris, 1 mM DTT, pH 7.5. The final denaturant concentration was varied between 0.36 and 2 M GdnHCl. Refolding traces were fit to double-exponential functions.

In addition to the [denaturant]-jump experiments outlined above, pH-jump experiments were also performed. A stock solution of alkali-unfolded FRB at pH 12.0 was mixed in a 1:1 ratio with low pH buffer to give final conditions of 5 μM FRB, 50 mM MES, 1 mM DTT, pH 6.0. The kinetic traces were fit to a triple-exponential process.

Multiple refolding phases are often observed, due to heterogeneity in the unfolded state, which is a result of slow proline isomerization.[48,49] Only the fast, major phase that corresponds to the folding of the fraction of molecules that have all their proline residues in a native-conformation in the unfolded state is considered here.

For many small proteins it has been established that the kinetics of folding are independent of the method of denaturation and only depend on the final conditions. Thus, results from the pH-jump experiment are identical to those obtained from the [denaturant]-jump experiment, if the final conditions are the same.

Data Analysis. Unfolding and refolding data were fit to a three-state model of folding where an intermediate is rapidly formed and

(45) Chen, J.; Zheng, X.; Brown, E. J.; Schreiber, S. L. *Proc. Natl Acad. Sci. U.S.A.* **1995**, *92*, 4947−4951.
(46) Choi, J.; Chen, J.; Schreiber, S. L.; Clardy, J. *Science* **1996**, *273*, 239−242.
(47) Peterman, B. F. *Anal. Biochem.* **1979**, *93*, 442−444.
(48) Brandts, J. F.; Halvorson, H. R.; Brennan, M. *Biochemistry* **1975**, *14*, 4953−4963.
(49) Jackson, S. E.; Fersht, A. R. *Biochemistry* **1991**, *30*, 10436−10443.

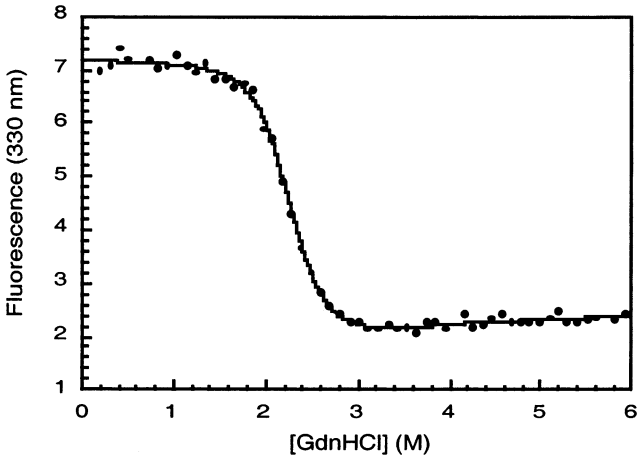

Figure 2. Typical GdnHCl-induced denaturation curve carried out on wild-type FRB at pH 6.0, 10 °C. The best fit of the data to eq 1 is shown by the solid line.

populated, using eq 2.[29,32]

$$\ln k_{obs} = \ln\left(\left(K_{UI}\exp\left(-\frac{m_{U-I}[D]}{RT}\right)\right)k_{IN}\exp\left(-\frac{m_{I-\ddagger}[D]}{RT}\right)\Big/ \left(1 + K_{UI}\exp\left(-\frac{m_{U-I}[D]}{RT}\right)\right) + k_{NI}\exp\left(\frac{m_{F-\ddagger}[D]}{RT}\right)\right) \quad (2)$$

where K_{UI} is the equilibrium constant between the unfolded and intermediate states, m_{U-I} is the m value between unfolded and intermediate states, k_{IN} is the rate constant for the formation of the native state from the intermediate state, $m_{I-\ddagger}$ is the m value between the intermediate and transition states, k_{NI} is the unfolding rate constant, and $m_{F-\ddagger}$ is the m value between the native and transition states.[29,32]

1-Anilinonaphthalene-8-sulfonate (ANS) experiments: [Denaturant]-jump refolding experiments were also performed in the presence of the hydrophobic dye ANS. ANS is known to bind to clusters of exposed hydrophobic groups often present in partially structured intermediate states.[50] On binding it undergoes a large change in fluorescence (the signal increases dramatically as ANS fluorescence is quenched in the aqueous environment, yield increases, and λ_{max} shifts to lower wavelengths characteristic of an apolar environment) and it is therefore a very sensitive probe of metastable intermediate states. It does not bind to the unfolded or native states of FRB (data not shown). An excitation wavelength of 372 nm was used and the emission was monitored above 405 nm. Final conditions were 3 μM FRB, 50 mM MES, pH 6.0, 1 mM DTT, 110 mM ANS, 0.36 M GdnHCl. The experiment was performed at 10 °C.

Results

Equilibrium Denaturation. GdnHCl-induced denaturation experiments were performed under various conditions and a typical unfolding curve is shown in Figure 2. Experimentally, it has been shown that the free energy of unfolding is linearly proportional to the denaturant concentration:[51]

$$\Delta G_{\mathbf{U-F}}{}^{\mathbf{D}} = \Delta G_{\mathbf{U-F}}{}^{\mathbf{H_2O}} - m_{U-F}[D] \quad (3)$$

where $\Delta G_{U-F}{}^{D}$ is the free energy of unfolding at any denaturant concentration, $\Delta G_{U-F}{}^{H2O}$ is the free energy of unfolding in water, and m_{U-F} is a constant of proportionality dependent on the average degree of exposure of each residue upon unfolding.

(50) Jones, B.; Jennings, P.; Pierre, R.; Matthews, C. *Biochemistry* **1994**, *33*, 15250−15258.

Table 1. Thermodynamic Parameters from the Unfolding of FRB under Equilibrium Conditions

protein	conditions	m_{U-F} (kcal mol^{-1} M^{-1})	$[D]_{50\%}$ (M)	$\Delta G_{U-F}{}^{H_2O}$ (kcal mol^{-1})
FRB	25 °C, pH 7.5	2.60 ± 0.20	2.41 ± 0.02	6.27 ± 0.49
FRB	10 °C, pH 7.5	3.02 ± 0.22	2.92 ± 0.02	8.82 ± 0.65
FRB	37 °C, pH 7.5	2.11 ± 0.23	1.94 ± 0.04	4.09 ± 0.45
FRB	25 °C, pH 6.0	3.03 ± 0.70	2.09 ± 0.05	6.33 ± 1.47
HisFRB[a]	10 °C, pH 7.5	2.73 ± 0.12	2.83 ± 0.01	7.73 ± 0.34
HisFRB[a]	10 °C, pH 6.0	3.11 ± 0.15	2.26 ± 0.01	7.03 ± 0.34

[a] The affinity tag was left on for ease of purification.

At the midpoint of unfolding, when $[D] = [D]_{50\%}$, eq 3 simplifies to

$$\Delta G_{U-F}{}^{H_2O} = m_{U-F}[D]_{50\%} \quad (4)$$

Fluorescence data were fitted to eq 1 to yield values for m_{U-F} and $[D]_{50\%}$ that were then used to calculate $\Delta G_{U-F}{}^{H_2O}$. All these values are shown in Table 1.

The protein is reasonably stable at room temperature, approximately 6−7 kcal mol^{-1}, and has a typical m_{U-F} value for a protein of this size. The His tag has little effect on stability (Table 1).

Kinetics. The rate constants for refolding and unfolding of FRB as a function of denaturant concentration were measured by [GdnHCl]- and pH-jump experiments and are shown in Figure 3. Parts A, B, and C of Figure 3 show data for FRB at pH 7.5, His-tagged FRB at pH 7.5, and His-tagged FRB at pH 6.0, respectively. In all three cases, the unfolding limb of the plot is linear, as observed for the unfolding of most small proteins. The unfolding data, combined with the equilibrium data, can be used to predict the refolding rate constants as a function of denaturant concentration, assuming that the folding follows a two-state model, i.e., that there are no intermediate states populated during folding.[52] These predicted rate constants are shown by the solid lines in Figure 3. It is clear that, at low concentrations of denaturant, there is deviation from two-state behavior, and the experimentally measured rate constants are lower than expected. It has been shown that this type of observed "rollover" may be due to the population of an intermediate state during folding.[53] However, such rollover has also been attributed to the transient aggregation of protein during folding.[54] For example, studies on the 102-residue protein U1A show that above 2 μM protein the refolding kinetics appears to deviate from two-state behavior, while below 2 μM, the refolding rate constants fit well to a two-state model.[54] It is possible to distinguish between these two phenomena by measuring the refolding rate constants as a function of protein concentration. If rollover is due to the population of an intermediate state, the refolding rate constants should be independent of protein concentration, while, if transient aggregation is occurring, the refolding rate constants will decrease with increasing protein concentration.

The refolding rate constant for FRB (measured by pH-jump experiment in water) as a function of protein concentration is shown in Figure 4. There is a decrease in the refolding rate

(51) Tanford, C. *Adv. Protein Chem.* **1968**, *23*, 121−282.
(52) Jackson, S. E.; Fersht, A. R. *Biochemistry* **1991**, *30*, 10428−10435.
(53) Matouschek, A.; Kellis, J. T., Jr.; Serrano, L.; Fersht, A. R. *Nature* **1989**, *342*, 122−126.
(54) Oliveberg, M. *Acc. Chem. Res.* **1998**, *31*, 765−772.

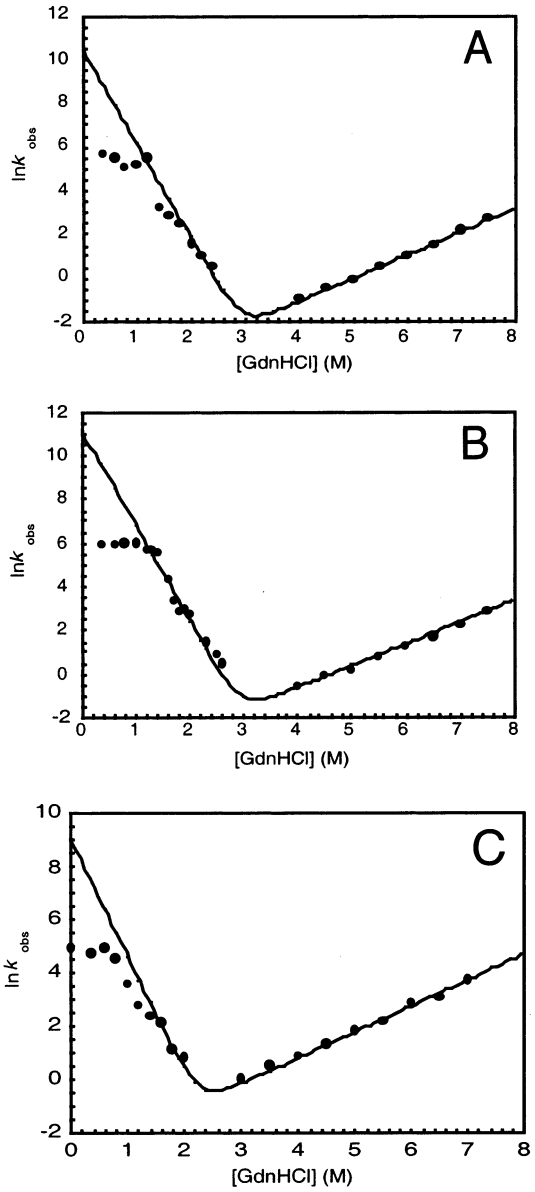

Figure 3. Unfolding and refolding rate constants as a function of denaturant concentration. Rate constants were measured by [GdnHCl]-jump experiments (filled circles). The solid line shows the change of rate constants with denaturant concentration for a simple two-state model, calculated from equilibrium and unfolding data.[52] All experiments were carried out at 10 °C: (A) FRB, pH 7.5; (B) HisFRB, pH 7.5; (C) HisFRB, pH 6.0. Points at 0 M denaturant were determined by pH-jump experiment.

constant with increasing protein concentration between 3 and 20 μM, indicative of transient aggregation during folding under these conditions. However, at protein concentrations of 2 μM and below, the observed rate constant does not vary significantly with protein concentration. Therefore, transient aggregation is not taking place under these conditions. Importantly, data shown in Figure 3 were acquired using a protein concentration of 2 μM; as we have shown that transient aggregation does not occur under these conditions, this suggests that an intermediate state is populated during the folding of FRB.

Recently, it has been noted that curvature in chevron plots often occurs at low concentrations of denaturant where refolding rate constants are frequently high. In these cases, it has been proposed that rollover may result from errors in curve fitting, as the rates approach the limit of stopped-flow instrumentation

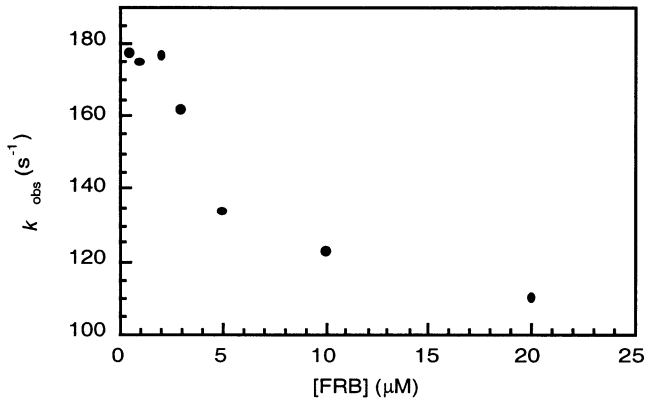

Figure 4. Dependence of the refolding rate constant on protein concentration. The HisFRB construct was used and the rate constant measured in 0 M GdnHCl (at pH 6.0, 10 °C) by pH-jump experiment.

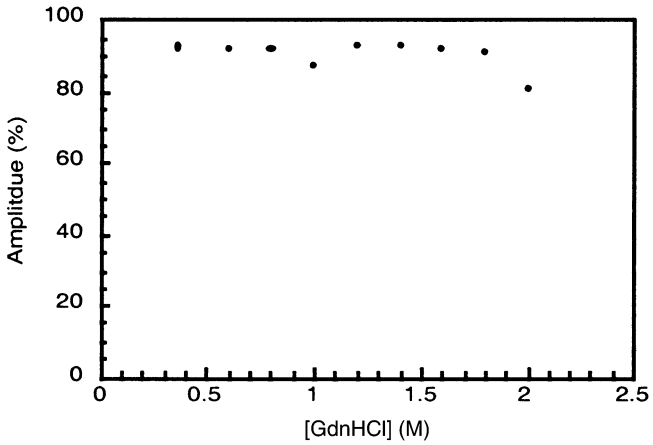

Figure 5. Plot of the amplitude of the major refolding phase of FRB against denaturant concentration. Refolding reactions were monitored using the split time-base facility on the stopped-flow over short and long time scales. This allowed the accurate determination of the percentage of the major refolding phase relative to the total change in amplitude. Slow phases exist due to heterogeneity in the unfolded state resulting from proline isomerization.

and the fast rates are masked by slower proline-isomerization-dependent processes.[55] If this is the case, then one would expect the amplitudes of the fast phase to decrease with decreasing denaturant concentration. The amplitude of the major fast-folding phase was determined as a percentage of the total change in amplitude on folding and is shown in Figure 5. There is no change in amplitude over the denaturant concentration range used; therefore, it is unlikely that the curvature at low [D] is due to errors in data fitting.

To confirm that the rollover observed is due to the presence of a populated intermediate state, refolding experiments were performed in the presence of the hydrophobic dye ANS. ANS binds to patches of exposed hydrophobic residues in partially structured states and in doing so undergoes a large change in fluorescence as the environment changes from polar to hydrophobic.[50] Neither the unfolded or native state of FRB bind ANS (data not shown), so if folding were occurring directly from the unfolded to the folded state, one would expect no change in ANS fluorescence during folding. Figure 6 shows the change in ANS fluorescence detected during the folding of FRB at low concentrations of denaturant; a large exponential decay is observed. The high initial ANS fluorescence corresponds to its

(55) Krantz, B. A.; Sosnick, T. R. *Biochemistry* **2000**, *39*, 11696−11701.

Table 2. **Parameters from the Best Fit of the Unfolding and Refolding Data to a Three-State Model**

protein	K_{U-I}	m_{U-I} (kcal mol^{-1} M^{-1})	k_{IN} (s^{-1})	$m_{I-‡}$ (kcal mol^{-1} M^{-1})	k_{NI} (s^{-1})	$m_{F-‡}$ (kcal mol^{-1} M^{-1})	β_T
FRB, pH 7.5	37 ± 33	2.5 ± 1.1	214 ± 216	0.5 ± 1.2	0.0053 ± 0.0035	0.60 ± 0.06	0.8
HisFRB, pH 7.5	292 ± 150	3.1 ± 0.4	183 ± 93	0.8 ± 0.5	0.0115 ± 0.0055	0.54 ± 0.04	0.8
His FRB, pH 6.0	8.4 ± 4.3	2.7 ± 0.6	139 ± 30	0.8 ± 0.7	0.060 ± 0.016	0.52 ± 0.03	0.8

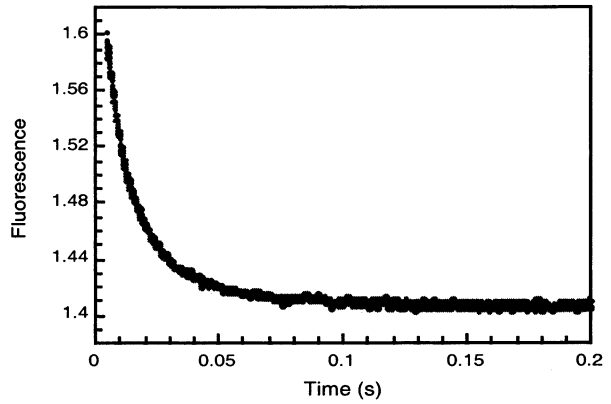

Figure 6. Refolding trace for FRB in the presence of ANS at 0.63 M GdnHCl, pH 6.0, 10 °C.

binding to a partially structured intermediate state formed within the millisecond dead-time of mixing. The fluorescence decreases exponentially as the population of the intermediate state decreases as the folded state is formed. Thus, rollover observed in the chevron plots can be attributed to the population of an intermediate state formed during folding.

Kinetic data can be fitted to a three-state model[29,32] to yield values for K_{UI}, m_{U-I}, k_{IN}, $m_{I-‡}$, k_{NI}, and $m_{F-‡}$ (see Table 2). Values for K_{UI} range between 8 and 292, depending on conditions illustrating that the intermediate state is some 1−3 kcal mol^{-1} more stable than the unfolded state in water. The rate constant for the formation of the native state from the intermediate state ranges from 140 to 214 s^{-1}, consistent with the measured rates. Values of m_{U-I} and $m_{I-‡}$, which are proportional to the change in surface area between the unfolded and intermediate states, and the intermediate and the rate-limiting transition states, have significant errors. However, the values indicate that the intermediate state is compact, and the negative values for $m_{I-‡}$ suggest that the protein may even have to partially unfold before it can attain the transition state. The sum of all three m values (m_{U-I}, $m_{I-‡}$, and $m_{F-‡}$) ranges from 2.4 to 2.8 kcal mol^{-1}, consistent with m_{U-F} obtained from equilibrium experiments (Table 1). The $m_{F-‡}$.value can be used to calculate β_T, a measure of the average solvent accessibility of the transition state relative to unfolded and native states, using eq 5:

$$\beta_T = \frac{m_{F-‡}}{m_{U-F}} \tag{5}$$

For FRB, values are 0.8 under all experimental conditions, indicating that some 80% of the surface area buried in the native state is already buried in the transition state.

Discussion

Levinthal was one of the first to propose that there must be intermediate states on folding pathways in order to bias the polypeptide chain's search through conformational space to allow folding to occur on observed time scales.[56] Although his calculation was simple and made several approximations, for many years it was a commonly held belief that all proteins must fold through intermediate structures. In 1991, however, it was shown that populated intermediate states are not necessary for fast folding,[52] and this has now been shown for many small proteins.[57] In addition, it has now been established that intermediate states observed during the folding of some proteins are partially misfolded species off the direct folding path.[32,58] Such studies have led to a reasssessment of the exact role of intermediate states in folding processes.[59−61] To address this issue, more information is needed on which proteins populate intermediate states, under what conditions these states are observed, and what structures these states adopt. Here we present a characterization of the folding pathway of the four-helical bundle protein FRB and show that at low concentrations of denaturant an intermediate state is populated.

A common method for determining whether a protein populates an intermediate state during folding is to measure the unfolding and refolding rate constants as a function of denaturant concentration.[53] For a protein that does not populate any intermediate states, a plot of ln k versus [D] results in a V-shaped or chevron plot.[52] If an intermediate is populated, however, rollover is observed at low concentrations of denaturant.[53] From Figure 3, it is clear that rollover is observed for FRB under all the conditions measured. Recently, however, it has been established that rollover may also result from transient aggregation of the protein during refolding.[54] To establish whether transient aggregation is occurring during the folding of FRB, the refolding rate constant was measured as a function of protein concentration. Figure 4 shows that, while there is some transient aggregation at high protein concentrations, there is no detectable aggregation at low protein concentrations, which are the conditions used for the experiments shown in Figure 3. Thus, in the case of FRB, it is unlikely that rollover results from transient aggregation. Recently, it has also been proposed that rollover may also result from the incorrect fitting of kinetic data at high rates approaching the limitations of the stopped-flow instrumentation.[55] If the rate is underestimated as a result of the this, then the amplitude of the major refolding phase would decrease. Figure 5 clearly shows that there is no decrease in amplitude on decreasing the denaturant concentration; thus, it is unlikely that curvature results from errors in data fitting. Movements in the transition state with denaturant concentration and specific ionic strength effect have also both been proposed to cause curvature in chevron plots in the absence of a populated intermediate state.[62−63] To verify that an intermediate state is

(56) Levinthal, C. *J. Chim. Phys.* **1968**, *85*, 44−45.
(57) Jackson, S. E. *Folding Des.* **1998**, *3*, R81−R90.
(58) Matagne, A.; Radford, S. E.; Dobson, C. M. *J. Mol. Biol.* **1997**, *267*, 1068−1074.
(59) Baldwin, R. L.; Rose, G. D. *Trends Biochem. Sci.* **1999**, *24*, 77−83.
(60) Baldwin, R. L.; Rose, G. D. *Trends Biochem. Sci.* **1999**, *24*, 26−33.
(61) Fersht, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 10869−10873.
(62) Oliveberg, M.; Tan, Y.-J.; Silow, M.; Fersht, A. R. *J. Mol. Biol.* **1998**, *277*, 933−943.

Table 3. A Comparison of the Thermodynamic and Kinetic Parameters for FRB and Other All-Helical Proteins

protein[a]	ref	no. of residues	$\Delta G_{U-F}^{H_2O}$ (kcal mol^{-1})	temp (°C)	$k_f^{H_2O}$ (s^{-1})[c]	β_T
FRB		95	7	10	156[b]	0.8
cytochrome *c*	19	104	7	20	2800	0.5
cytochrome b_5	10,11	104	3	10	5590	0.7
cytochrome *c*′	17	125		22	250	ND
cytochrome b_{562}	12,13	106		20	14000	ND
EnHd	31	61	2	25	37500	0.9
Im 7	28−30	85	46	10	345	0.9
Im 9	28−30	87	6	10	1450	0.9
ACBP	24−27	86	7	20	704	0.6
λ repressor						
wild type	18−21	80	3	37	4900	0.4
G46A,G48A		80	5	37	88000	0.8

[a] FRB and cytochromes *c*′, b_5, and b_{562} are all four-helical bundles with an up−down motif. The immunity proteins, Im7 and Im9, and ACBP are four-helical bundles where one of the helices is shorter than the others. λ repressor is a four-helical protein, whereas cytochrome *c* and engrailed homeodomain (EnHd) have three-helices. [b] Measured directly by pH-jump experiment. [c] Some of these values are estimates of the folding rate constant in water from extrapolation of data acquired at higher denaturant concentrations and assuming two-state behavior.

populated during the folding of FRB, we undertook an experiment aimed at detecting the intermediate species in a more direct manner. For this, refolding experiments with the hydrophobic dye ANS were undertaken. Figure 6 shows the exponential decrease in ANS fluorescence observed during folding at low concentrations of denaturant. The initial high level of ANS fluorescence corresponds to the rapid formation of the intermediate state that binds ANS (within the deadtime of the stopped-flow apparatus), and the decay corresponds to the decrease in population of the intermediate state as the folded state is formed. Neither the unfolded state nor the native state bound ANS (data not shown). We, therefore, have strong evidence that an intermediate state is populated during the folding of FRB. Analysis of the kinetic data with a three-state model of folding (Table 2) shows that the intermediate state formed is compact and that this state may even have to partially unfold before it can reach the transition state. It is interesting to note that similar results have recently been published for the helical bundle protein Im7.[31] The Radford group used ϕ value analysis to show that the intermediate state is a misfolded species with significant non-native interactions.[31]

In this paper, we have shown that a small four-helical bundle protein, FRB, folds fast with a half-life of approximately 5 ms in water. Significantly, it populates an intermediate state on the folding pathway. This intermediate state may be a partially misfolded species that has to unfold before it can reach the transition state. A comparison of the kinetic and thermodynamic data for FRB with those for other all-helical proteins is shown in Table 3. It is clear that there is a wide variation in folding rate constants, ranging from 156 to 88 000 s^{-1}. Although, in comparison to many α/β or all-β proteins, the folding of FRB is fast,[57] in comparison with other all-α proteins, it is one of the slowest. It is interesting to note that the next slowest is Im7, with a rate constant of 345 s^{-1}, and that it too populates an intermediate state. This is in general agreement with Fersht's argument that stable intermediate states slow folding.[61] Whether a difference in rate constant between 100 and 10 000 s^{-1} is significant in vivo remains unclear. It is interesting to note that the predicted rate constant for folding for FRB based on a two-state transition fits well to that predicted on the basis of contact order[64] (data not shown). This suggests that local interactions are important in defining the folding of FRB.

For small proteins, there is some evidence that there is a correlation between stability and population of intermediate states. For example, in the case of ubiquitin, an intermediate state can be detected only in the presence of the stabilizing salt sodium sulfate.[65] Comparison of all-helical proteins (Table 3) shows little correlation. For example, cytochrome *c* and ACBP are both as stable as FRB but fold with apparent two-state kinetics as measured by the [denaturant] dependence of the rate constants. Interestingly, however, in both cases there is some evidence that an intermediate forms very rapidly but is probably not sufficiently well populated for an effect on the chevron plot to be detectable. In comparison, Im7 is less stable but still populates an intermediate state. Even within a family of proteins with the same structure, such as the immunity proteins, there is no correlation. This suggests that intermediate states are stabilized by a specific set of interactions and less influenced by the global stability of the protein. In addition, there is no correlation between the size and population of intermediate states (Table 3). Some large, stable proteins fold in a two-state manner, while smaller, less stable proteins populate intermediate states. Whether an intermediate is populated or not appears to depend on specific interactions, structure, and stability and is therefore likely to vary from protein to protein.

β_T (Table 3) is a measure of the average solvent accessible surface area in the transition state. The value for FRB is 0.8, which is similar to the values obtained the immunity proteins (0.9), which also have a four-helical structure, although their topology is significantly different from that of FRB. It is worth noting that in both cases the intermediate states are highly compact and structured states that may be partially misfolded species.

In conclusion, the four-helical bundle protein FRB populates an intermediate state during folding. This has been definitively shown by a thorough kinetic analysis consisting of protein concentration dependence, amplitude dependence, and refolding in the presence of ANS. There is some evidence that the intermediate state is highly compact and possibly a partially misfolded species. FRB's folding is somewhat slower than for other all-helical proteins, which may be a consequence of the presence of this stable intermediate state. Despite this, it folds efficiently with a half-life of 5 ms in water. The folding of FRB is most similar to the folding of the four-helical proteins belonging to the immunity family.

JA016480R

(63) Went, H. M.; Benitez-Cardoza, C. B.; Jackson, S. E. *Protein Sci.* submitted.

(64) Plaxco, K. W.; Simons, K. T.; Baker, D. *J. Mol. Biol.* **1998**, *277*, 985−994.

(65) Khorasanizadeh, S.; Peters, I. D.; Roder, H. *Nat. Struct. Biol.* **1996**, *3*, 193−205.