A Residual Structure in Unfolded Intestinal Fatty Acid Binding Protein Consists of Amino Acids That Are Neighbors in the Native State†

Ira J. Ropson,* Joshua A. Boyer,‡ and Paula M. Dalessio

Department of Biochemistry and Molecular Biology, The Pennsylvania State University College of Medicine, Hershey, Pennsylvania 17033

Received October 13, 2005; Revised Manuscript Received January 6, 2006

ABSTRACT: Much of the recent effort in protein folding has focused on the possibility that residual structures in the unfolded state may provide an initiating site for protein folding. This hypothesis is difficult to test because of the weak stability and dynamic behavior of these structures. This problem has been simplified for intestinal fatty acid binding protein (IFABP) by incorporating fluorinated aromatic amino acids during synthesis in *Escherichia coli*. Only the labeled residues give signals by ^{19}F NMR, and the 1D spectra can be assigned in both the native and unfolded states by site-directed mutagenesis. One of the two tryptophans (W82), one of the four tyrosines (Y70), and at least four of the eight phenylalanines (including F68 and F93) of IFABP are involved in a structure that is significantly populated at concentrations of urea that unfold the native structure by fluorescence and CD criteria. These residues are nonlocal in sequence and also contact each other in the native structure. Thus, a template of nativelike hydrophobic contacts in the unfolded state may serve as an initiating site for folding this β-sheet protein.

Understanding the mechanism by which a protein achieves a specific functional three-dimensional structure remains one of the most difficult problems in biochemistry. In order to address this problem, knowledge of the structures of all the states that participate in folding will be required. In particular, a better understanding of the conformations of the unfolded state of a protein is necessary. Although initial descriptions of unfolded proteins emphasized the randomness of the structure, there is ample evidence for nonrandom interactions in the unfolded state for a wide variety of proteins, even under strongly denaturing conditions (*1*, *2*). If these contacts are similar to those of the native state, they could act as initiating sites for folding, greatly simplifying the search for the native structure through conformational space (*1*).

It is difficult to directly determine the structures of these potential initiating sites. Although nuclear magnetic resonance (NMR)[1] is one means to examine these structures, the dynamic nature of the unfolded state has made the complete assignment of an unfolded protein by NMR a difficult task, especially for the side chain resonances (*2*). One way of simplifying the NMR spectra is to replace the aromatic amino acids of a protein with fluorinated derivatives of the same residue (*3*). Fluorine has several advantages for NMR studies, including 100% natural abundance, broad chemical shift distribution, and extreme sensitivity of the chemical shift to the local environment (*3*). Incorporation of the fluoro-labeled amino acids into proteins expressed in *Escherichia coli* is simple, and large quantities of labeled protein can be easily obtained (*4*). Finally, assignment of the resonances to specific residues can usually be done by site-directed mutagenesis, allowing the behavior of specific residues in the sequence to be examined (*3*, *4*).

Several studies have used fluorine labeled proteins and ^{19}F NMR to examine the folding process, both kinetically (*5*−*7*) and at equilibrium (*8*−*10*). The folding of rat intestinal fatty acid binding protein (IFABP) labeled with 6-fluorotryphan (6Ftrp) has been examined previously (*11*). This study examines the folding of mutants of IFABP containing a single tryptophan, and to proteins labeled with 4-fluorophenylalanine (4Fphe-IFABP) or 3-fluorotryosine (3Ftyr-IFABP), providing probes throughout the protein structure. The data support the hypothesis that a cluster of hydrophobic residues that contact each other in the native state participate in a preferred structure in the unfolded state at high concentrations of denaturant.

MATERIALS AND METHODS

Protein Source and Purification. The expression plasmid for IFABP has been previously described (*12*). Mutant expression vectors for W6Y-IFABP, W82Y-IFABP, Y14F-IFABP, Y70F-IFABP, Y117F-IFABP, Y119F-IFABP, F68A-IFABP, and F93A-IFABP were constructed using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). The plasmid DNA for each construct was isolated using a QIA

† This research was supported by NIH Grant GM-57906 to I.J.R.

* To whom correspondence should be addressed. Mailing address: 500 University Dr., Hershey, PA 17033. Tel: 717-531-4064. Fax: 717-531-7072. E-mail: iropson@psu.edu.

‡ Current address: Department of Biochemistry and Biophysics, University of North Carolina, Chapel Hill, NC 27599.

[1] Abbreviations: NMR, nuclear magnetic resonance; IFABP, rat intestinal fatty acid binding protein; 6Ftrp, 6-fluorotryptophan; 4Fphe, 4-fluorophenylalanine; 3Ftyr, 3-fluorotyrosine; 6Ftrp-IFABP, IFABP labeled with 6-fluorotryptophan; 4Fphe-IFABP, IFABP labeled with 4-fluorophenylalanine; 3Ftyr-IFABP, IFABP labeled with 3-fluorotyrosine; W6Y-IFABP, Y14F-IFABP, Y70F-IFABP, Y117F-IFABP, Y119F-IFABP, F68A-IFABP, and F93A-IFABP, various mutants of IFABP; CD, circular dichroism; SDS, sodium dodecyl sulfate; ΔG_{H_2O}, free energy of stabilization in the absence of denaturant; m_G, dependence of ΔG on denaturant concentration.

10.1021/bi052091o CCC: $33.50
Published on Web 02/04/2006

Spin Miniprep kit (Qiagen) and sequenced by the Molecular Genetics Core Facility (Penn State College of Medicine) to verify the presence of the specified mutation and confirm the absence of any other mutations.

6Ftrp-WT-IFABP, 6Ftrp-W6Y-IFABP, and 6Ftrp-W82Y-IFABP were expressed in *E. coli* strain KS463(DE3), a tryptophan auxotroph. 4Fphe-WT-IFABP, 4Fphe-F68A-IFABP, 4Fphe-F93A-IFABP, 3Ftyr-WT-IFABP, 3Ftyr-Y14F-IFABP, 3Ftyr-Y70F-IFABP, 3Ftyr-Y117F-IFABP, and 3Ftyr-Y119F-IFABP were expressed in *E. coli* strain DL39(DE3), a phenylalanine and tyrosine auxotroph. The media used for these preparations were similar to those used for labeling IFABP with 6Ftrp as previously described except for the labeled amino acid (*11*). The yields for the mutant 3Ftyr proteins were low. In order to improve yields and ease purification, the genes for these proteins (3Ftyr-Y14F-IFABP, 3Ftyr-Y70F-IFABP, 3Ftyr-Y117F-IFABP, 3Ftyr-Y119F-IFABP) were transferred to the pET-28b(+) vector (Novagen) and expressed by induction with IPTG in DL39(DE3). The his-tag had no significant effect on stability, folding mechanism, or frequency of the 3Ftyr resonances in the labeled proteins (data not shown) and these proteins were used for assignment purposes only.

All of the labeled proteins were extracted from *E. coli* using the freeze thaw method (*13*). Following lysis of the cells, 6Ftrp-WT-IFABP, 6Ftrp-W6Y-IFABP, 6Ftrp-W82Y-IFABP, 4Fphe-WT-IFABP, 4Fphe-F68A-IFABP, and 3Ftyr-WT-IFABP were purified from the cell supernatant as previously described (*12*, *14*). 4Fphe-F93A-IFABP was isolated from inclusion bodies and purified as previously described (*15*). Protein purity was demonstrated by the presence of a single band on a 20% SDS polyacrylamide gel. The following extinction coefficients were calculated at 280 nm: 1.13 mg^{-1} cm^{-1} for labeled WT-IFABP, 4Fphe-F68A-IFABP, and 4Fphe-F93A-IFABP and 0.87 mg^{-1} cm^{-1} for 6Ftrp-W6Y-IFABP and 6Ftrp-W82Y-IFABP (*16*).

Reagents. Urea stock solutions (approximately 10 M, ultrapure, Amresco) were prepared and stored at −20 °C as previously described (*17*). On the day of an experiment, a working solution of 9 M urea was prepared by adding the buffer components to a freshly thawed urea stock solution. All buffers contained 25 mM $NaPO_4$, 75 mM NaCl, 0.1 mM EDTA. The pH of working buffer and urea solutions was adjusted to pH 7. Actual denaturant concentrations were determined by refractive index measurements with a Milton Roy Abbe-3 refractometer at 25 °C (*18*). All buffers were filtered through 0.2 mm Whatman nylon membranes. All chemicals were reagent grade unless otherwise noted.

Equilibrium Studies. Unfolding transitions as a function of denaturant concentration were monitored by circular dichroism (CD) and fluorescence. A Jasco J-710 spectropolarimeter was employed to follow changes in secondary structure in the far-UV portion of the CD spectrum (*14*). Fluorescence studies were carried out with a PTI QuantaMaster luminescence spectrometer (*14*). Samples ranging in denaturant concentrations from 0 to 8.5 M urea were prepared using a Hamilton Microlab titrator (*19*). The data were corrected for the background signal of the buffer and urea solutions. For comparison of denaturation curves, the data were normalized from 0 to 1, where 0 and 1 were assigned to the lowest and highest observed intensity, respectively (*14*). Nonlinear least-squares fits to the equilibrium data were generated using KaleidaGraph (Synergy Software) as previously described (*14*). All fits were to a minimum of two independent data sets.

^{19}F NMR Studies. NMR data were collected on a Bruker AMX-500 spectrometer using a 5 mm Bruker hydrogen/fluorine dual channel probe at 470.54 MHz. The ^{19}F spin lattice relaxation times were determined using the inversion recovery method and were <0.5 s for all protein resonances. The pulse interval and pulse width were 2.5 s and 15.2 μs, respectively. The sweep widths ranged from 7000 to 10 000 Hz, and acquisition times ranged between 0.6 and 0.9 s. The spectral offset and sweep width depended on the labeled protein and standard used, but always extended at least 500 Hz beyond the frequencies observed for the native resonances. At least 256 transients were collected for each sample. 6Ftrp was used as an external chemical shift and concentration standard for proteins containing 4Fphe and 3Ftyr. 4Fphe was used as the standard for proteins containing 6Ftrp. Samples were collected with decoupling of the aromatic ring protons during data collection. Proteins were dialyzed against 25 mM PO_4, 75 mM NaCl, 0.1 mM EDTA at pH 7. Initial protein concentrations were greater than 10 mg/mL. Five milligram aliquots of each protein were placed in 2 mL Eppendorf tubes, and sufficient buffer was added to bring the final volume to 0.6 mL. The proteins were lyophilized, and stored at −20 °C until use. Each lyophilized sample was dissolved in the appropriate volumes of deionized water (10% D_2O) and 8 M urea (10% D_2O) for the final concentrations of urea shown.

Spectra were collected at 25 °C and processed with the NUTS software program (Acorn NMR, Livermore, CA) using 5 Hz of line broadening. Line widths were determined with the NUTS software. WINDNMR (*20*, http://www.chem.wisc.edu/areas/reich/plt/windnmr.htm) was used to perform dynamic NMR simulations. The ASC program (*21*, http://mendel.imp.univie.ac.at/mendeljsp/studies/asc.jsp) was used to calculate solvent accessibilities.

RESULTS

Equilibrium Unfolding by Fluorescence and Circular Dichroism. The equilibrium and kinetic mechanism for the folding of IFABP has been the subject of extensive studies (*11*, *14*, *15*, *17*, *22*, *23*). WT-IFABP behaves as a "two-state folder" by fluorescence and CD criteria, suggesting that significant concentrations of only the unfolded and native states are present at equilibrium (*14*, *17*, *22*, *23*). All of the fluoro-aromatic labeled proteins followed a two state model as judged by CD and fluorescence criteria (Table 1, Figure 1). However, the incorporation of fluoro-aromatic amino acids did perturb the equilibrium folding compared to the unlabeled proteins. In every case the dependence of ΔG_{H_2O} on denaturant concentration (m_G) was decreased compared to the unlabeled protein (Table 1, Figure 1). This decrease resulted in a lower ΔG_{H_2O} for the labeled proteins despite small increases in the midpoint for denaturation (Table 1). Unlabeled IFABP has a low value for m_G compared to other proteins of this size (*24*) and in this family (*14*, *25*, *26*), and labeling with fluorine accentuates this property. Fluorinated amino acids are slightly more hydrophobic than their unmodified counterparts (*27*, *28*), and the increased hydrophobicity may stabilize the equilibrium intermediate observed

Table 1: Summary of Equilibrium Unfolding Data

protein	ΔG_{H_2O} (kcal mol^{-1})	m_G (kcal mol^{-1} M^{-1})	midpoint (M)
WT-IFABP	4.79 ± 0.22	−1.13 ± 0.07	4.21 ± 0.04
6Ftrp-IFABP	3.82 ± 0.20	−0.83 ± 0.04	4.61 + 0.04
4Fphe-IFABP	3.98 ± 0.32	−0.85 ± 0.07	4.69 ± 0.08
3Ftyr-IFABP	3.54 ± 0.24	−0.74 ± 0.06	4.77 ± 0.09
W6Y-IFABP	3.86 ± 0.22	−1.09 ± 0.05	3.52 ± 0.04
6Ftrp-W6Y-IFABP	2.93 ± 0.17	−0.77 ± 0.04	3.8 ± 0.05
W82Y-IFABP	4.88 ± 0.44	−1.38 ± 0.12	3.54 ± 0.04
6Ftrp-W82Y-IFABP	3.99 ± 0.10	−1.10 ± 0.03	3.60 ± 0.02
F68A-IFABP	6.17 ± 0.38	−1.76 ± 0.11	3.52 ± 0.02
4Fphe-F68A-IFABP	3.10 ± 0.15	−0.96 ± 0.04	3.23 ± 0.03
F93A-IFABP	2.37 ± 0.16	−1.30 ± 0.06	1.83 ± 0.04
4Fphe-F93A-IFABP	1.94 ± 0.13	−0.77 ± 0.03	2.53 ± 0.07

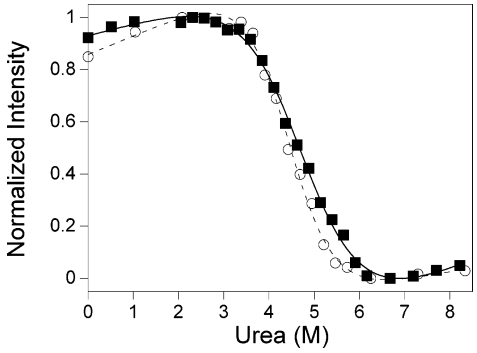

FIGURE 1: Equilibrium fluorescence unfolding of WT-IFABP (open circles) and 4Fphe-WT-IFABP (closed squares). The dashed and solid lines are the fit to a two state model of the equilibrium unfolding of WT-IFABP and 4Fphe-WT-IFABP, respectively.

by NMR during unfolding. Increased stability of the intermediate would decrease the cooperativity of the unfolding transition, as reflected by the decrease in m_G observed for the labeled proteins.

All of the labeled proteins had CD and fluorescence spectra that were very similar to those of their unlabeled counterparts (data not shown). In general, the incorporation of fluoroaromatic amino acids had a smaller effect on stability and folding than conservative amino acid substitutions of those same residues, supporting the use of this method to explore folding at the atomic level.

6Ftrp Labeled IFABP. In a previous study of 6Ftrp-WT-IFABP (*11*), only W82 appeared to participate in an intermediate state whereas W6 did not. Figure 2 shows the dependence of the ^{19}F NMR spectra on urea concentration for 6Ftrp-W6Y-IFABP, 6Ftrp-WT-IFABP, and 6Ftrp-W82Y-IFABP. 6Ftrp-WT-IFABP (middle panel) displayed two peaks in the native (0 M) and unfolded states (7.6 M). This observation emphasizes one of the advantages of ^{19}F NMR. Differences in the local amino acid sequence are sufficient to cause significant chemical shift dispersion for 6Ftrp resonances, even in the unfolded state (*5*). Site-directed mutagenesis clearly identified the residues responsible for each of the resonances. The W82 resonance was downfield from the resonance for W6 for both the native and unfolded states (Figure 2). However, the dependence of the resonance intensities on urea concentration differed for the two tryptophans.

6Ftrp-W82Y-IFABP contains only a single tryptophan (W6) in its sequence. The loss of the native state W6 resonance intensity and gain of the W6 unfolded resonance

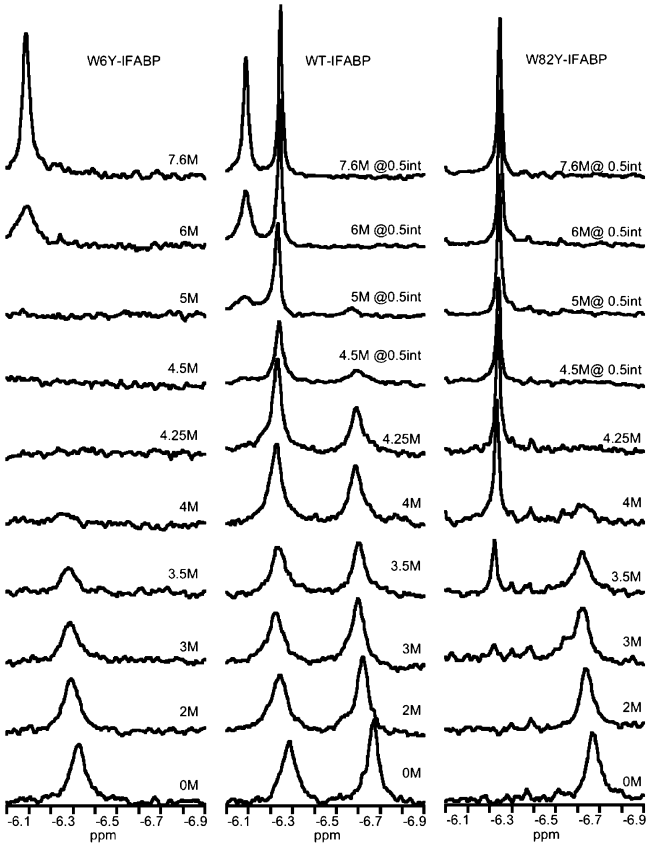

FIGURE 2: Spectra of 6Ftrp-W6Y-IFABP, 6Ftrp-WT-IFABP, and 6Ftrp-W82Y-IFABP at the listed concentrations of urea. Some of the spectra have been multiplied by 0.5 for easier visualization. The chemical shift scale is from 4Fphe.

intensity at increasing concentrations of denaturant corresponded to the equilibrium unfolding change observed by fluorescence (Table 1). The line widths of the native and unfolded resonances (30−35 Hz and 9−11 Hz, respectively) were nearly constant over the entire range of denaturant concentrations, suggesting that the two conformations were in slow exchange on the NMR time scale. Slow exchange kinetics for this process is consistent with the rates observed for the transition between the native and unfolded states (between 0.1 and 10 s^{-1}, depending on the denaturant concentration) by stopped-flow kinetics for W82Y-IFABP (*23*) and the rate observed for the loss of native state intensity for the W6 resonance during unfolding of 6Ftrp-WT-IFABP by stopped-flow NMR (*5*). The assigned resonances for the native and unfolded conformations of W6 in 6Ftrp-WT-IFABP demonstrated behavior similar to that of 6Ftrp-W82Y-IFABP. Thus, W6 exchanges between only the native and unfolded conformations during unfolding.

6Ftrp-W6Y-IFABP has only W82 in its sequence, and the loss of the native state resonance intensity for W82 corresponded to the equilibrium unfolding change observed by fluorescence (Table 1). However, the gain in intensity for the unfolded conformation could not be explained by a simple N ⇔ U transition. Instead, the unfolded state resonance was only observed at much higher concentrations of denaturant, similar to those concentrations where the W82 resonance appeared for 6Ftrp-WT-IFABP. No additional peaks were observed throughout the entire spectral range, even when the sweep width was tripled. At 7.6 M urea, the line width of the resonance associated with the unfolded

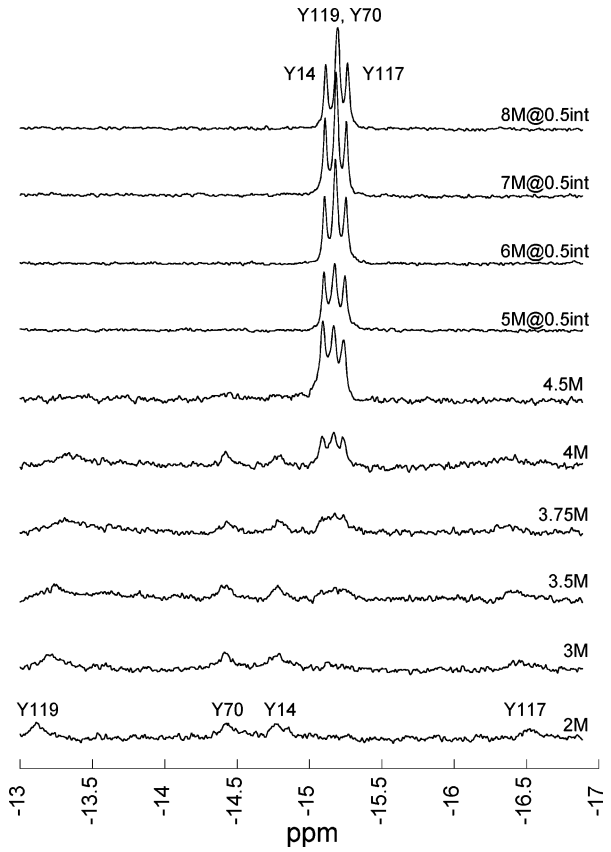

FIGURE 3: Spectra of 3Ftyr-WT-IFABP at the listed concentrations of urea. Some of the spectra have been multiplied by 0.5 for easier visualization. The chemical shift scale is from 6Ftrp.

conformation of W82 was broader than that for W6 in 6Ftrp-WT-IFABP, about 19 Hz for W82 compared to 10 Hz for W6 at 7.6 M urea. As discussed below, the simplest explanation for these results requires that W82 pass through at least one additional conformation during unfolding.

Both of the mutant proteins were less stable than WT-IFABP (Table 1). This decreased stability is reflected in the reduction of the native resonance intensities of 6Ftrp-W6Y-IFABP and 6Ftrp-W82Y-IFABP at lower concentrations of denaturant than in 6Ftrp-WT-IFABP. However, the unfolded resonance for W82 in 6Ftrp-W6Y-IFABP appeared over a similar range of denaturant concentrations as that observed for the W82 resonance in 6Ftrp-WT-IFABP, leading to a broader range of denaturant concentrations where no resonance intensity was observed.

3Ftyr Labeled IFABP. The dependence of the NMR spectra of 3Ftyr-WT-IFABP on urea concentration is shown in Figure 3. The protein yields were lower for the 3Ftyr labeled proteins than for the 6Ftrp and 4Fphe labeled proteins. However, sufficient protein was available to assign the resonances and to obtain a complete equilibrium unfolding profile for 3Ftyr-WT-IFABP. The line widths of the native state 4Ftyr resonances are broader than for most other fluorine labels. The increased line width in the native state is probably due to the nonsymmetrical labeling of the tyrosine ring. The label is next to the OH group of tyrosine, and ring flips would place the fluorine nucleus in different environments that are in chemical exchange. The line widths of the native resonances increased considerably at concentrations of urea that were not sufficient to unfold the protein by CD or fluorescence criteria (e.g., 3 M urea, Figure 3). This

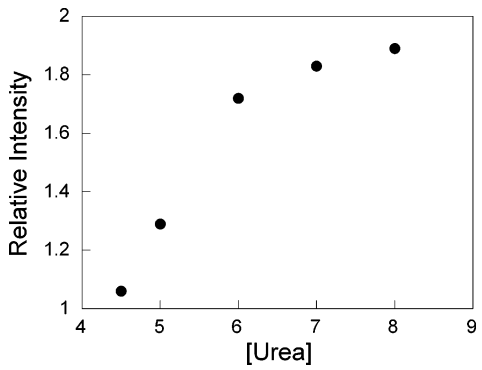

FIGURE 4: Relative intensity of the middle 3Ftyr unfolded state resonance. The relative intensity of the middle resonance was compared to the average intensities of the two flanking resonances, assuming that the line widths of all peaks were identical (13 Hz).

phenomenon is likely due to increases in the rate of ring flipping, resulting in line broadening.

As expected for a protein with four tyrosines, four peaks were observed for the native state at low concentrations of denaturant. Three peaks were observed for the unfolded state at high urea concentrations, with the central peak having twice the intensity of the other two peaks, suggesting that two of the resonances for 3Ftyr had nearly identical frequencies in the fully unfolded state. Site directed mutagenesis (replacing each tyrosine with phenylalanine) was used to assign individual resonances to specific residues in the native and unfolded states (Figure 3). At 5 M urea, where the protein was about 65% unfolded by fluorescence criteria, the intensity of the middle peak was only slightly larger than the intensity of the flanking peaks (Figure 4). The central peak became much more intense over the same denaturant concentration range for which the W82 unfolded resonance appeared in 6Ftrp-WT-IFABP, while the two flanking peaks did not increase as much at higher urea concentration (Figures 3 and 4). Thus, one of the tyrosine resonances shows the same behavior as that of W82. Site-directed mutagenesis identified Y70 as the residue causing this behavior. 3Ftyr-Y70F-IFABP displayed three peaks of similar intensity at 4.8 M urea that maintained the same proportions in 9 M urea (data not shown).

4Fphe Labeled IFABP. The dependence of the fluorine spectrum of 4Fphe-WT-IFABP on urea concentration is shown in Figure 5. As expected for a protein with eight phenylalanines, eight peaks were observed for the native state. The native resonances were assigned by mutagenesis (F68A-IFABP, F93A-IFABP) or by site specific incorporation of 4Fphe (*4*, *29*). The differences in line shape and intensity for the various resonances are due to local structure differences for each residue in the native state (*29*). The spectra for 4Fphe-WT-IFABP at high concentrations of denaturant are shown in Figure 6. At 8 M urea, 8 peaks of approximately equal line width and intensity can be deconvoluted. Two resonances overlapped near 5.5 ppm, and two other resonances overlapped between 5.3 and 5.4 ppm. Significant missing intensity was observed at 5.5 M urea, where the protein was about 75% unfolded by optical methods (Table 1). Integration of the spectra at 5.5 M urea indicated that at least 50% of the expected intensity was absent, suggesting that at least 4 of the 8 phenylalanines showed equilibrium unfolding behavior similar to those of W82 and Y70.

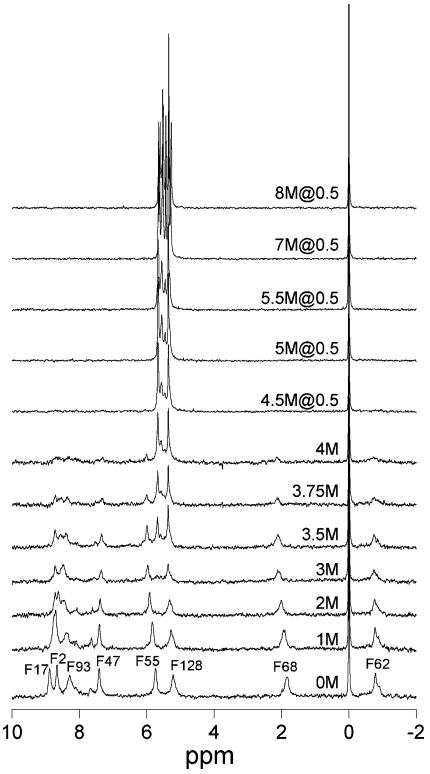

FIGURE 5: Spectra of 4Fphe-WT-IFABP at the listed concentrations of urea. Some of the spectra have been multiplied by 0.5 for easier visualization. The chemical shift scale is from 6Ftrp. The resonances associated with F68 and F93 in the native state were assigned by mutagenesis. Site-specific incorporation of 4Fphe has assigned the other resonances in native IFABP (29).

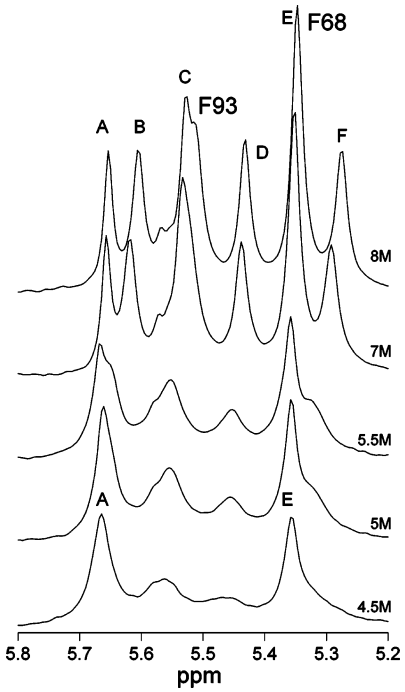

FIGURE 6: Spectra of 4Fphe-WT-IFABP at high concentrations of urea. Assignments of the resonances for F68 and F93 in the unfolded state are labeled. Unassigned resonances are labeled A−F. The chemical shift scale is from 6Ftrp.

Two mutant proteins (F68A-IFABP and F93A-IFABP) provided useful assignment information for the unfolded state. Figures 7 and 8 show the urea concentration dependence of the spectra for 4Fphe-F68A-IFABP. The missing

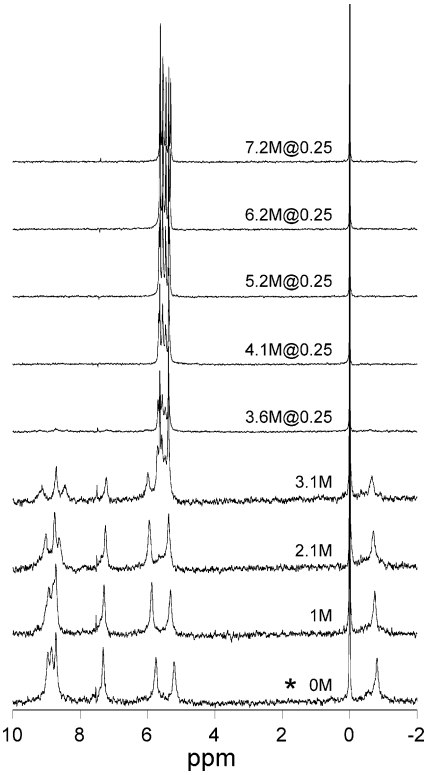

FIGURE 7: Spectra of F68A-4Fphe-IFABP at the listed concentrations of urea. Some of the spectra have been multiplied by 0.5 for easier visualization. The chemical shift scale is from 6Ftrp. The asterisk (*) is the frequency of the peak missing from the 4Fphe-WT-IFABP spectra.

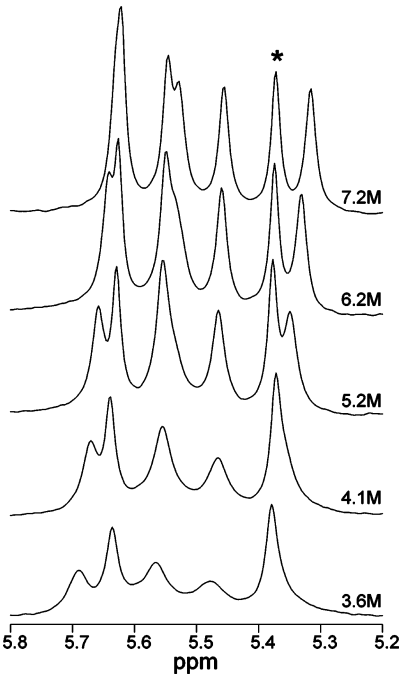

FIGURE 8: Spectra of F68A-4Fphe-IFABP at high concentrations of urea. The chemical shift scale is from 6Ftrp. The asterisk (*) is the frequency of the peak missing from the 4Fphe-WT-IFABP spectra.

resonances from the 4Fphe-WT-IFABP spectra were obvious in both the native (0 M urea) and fully unfolded (7.2 M urea) states. Figures 9 and 10 show similar plots for 4Fphe-F93A-IFABP, assigning the resonances for F93. A comparison of the mutant and wild-type protein spectra shows that both

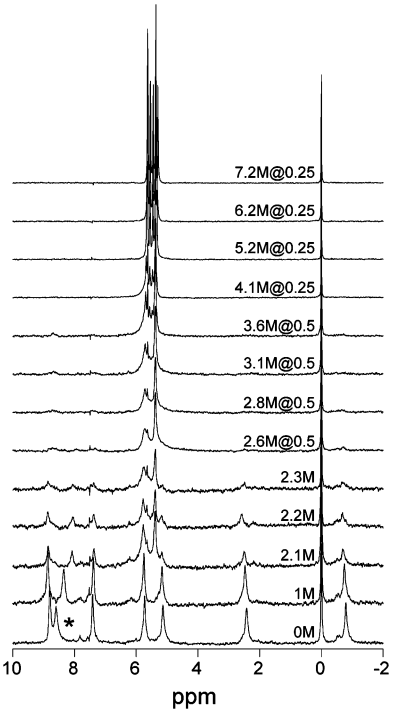

FIGURE 9: Spectra of F93A-4Fphe-IFABP at the listed concentrations of urea. Some of the spectra have been multiplied by 0.5 for easier visualization. The chemical shift scale is from 6Ftrp. The asterisk (*) is the frequency of the peak missing from the 4Fphe-WT-IFABP spectra.

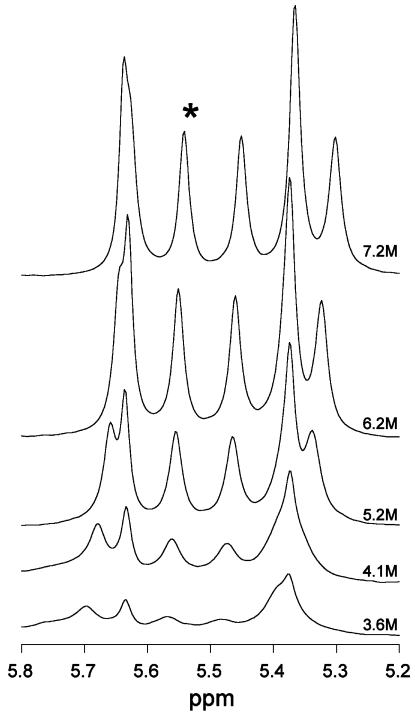

FIGURE 10: Spectra of F93A-4Fphe-IFABP at high concentrations of urea. The chemical shift scale is from 6Ftrp. The asterisk (*) is the frequency of the peak missing from the 4Fphe-WT-IFABP spectra.

F68 and F93 have unfolding behaviors similar to that of W82 and Y70.

Both of these phenylalanine to alanine mutations significantly destabilized the native state, and the loss of intensity for the native state resonances began at lower concentrations of denaturant than observed for 4Fphe-WT-IFABP (Figures 7 and 9). Although there was still significant missing intensity between 4 and 6 M urea for the mutant proteins, the spectra became more similar to the completely unfolded state at lower concentrations of denaturant than observed for 4Fphe-WT-IFABP (Figures 8 and 10), suggesting that these mutations also destabilized the intermediate.

DISCUSSION

IFABP is a predominately β-sheet protein consisting of two five-stranded sheets (*30*). The strands are connected by reverse turns in a simple up and down topology with the exception of the helix−turn−helix motif connecting the first and second β-strands. The β-sheets are arched with respect to each other, creating an internal solvent filled cavity that is isolated from exterior solvent by the two helices. The aromatic residues are scattered throughout the sequence, providing reporter groups for folding from all structural elements. Most of the aromatic side chains have solvent exposed surface areas of less than 10%, and those that expose greater than 10% of their surface area are oriented toward the internal cavity (*21*).

IFABP has been the subject of many folding studies, which have primarily focused on the kinetic mechanism for the folding of this β-sheet protein. The folding of this protein is complex, with a burst phase and at least two intermediates on the folding pathway (*11*, *14*, *15*, *17*, *22*, *23*, *31*, *32*). One of these intermediates lacks secondary structure by CD criteria, suggesting that at least one of the two tryptophans of the protein is involved in some tertiary contact in the absence of stable secondary structure (*14*, *17*, *22*, *23*). Kinetic studies of the folding of mutant proteins each having a single tryptophan have shown that W82 participates in the kinetic intermediate, but W6 does not (*23*).

There are several possible explanations for the "missing intensity" phenomenon observed at some but not all of the labeled residues. It is unlikely that the missing intensity is due to aggregation. There is no evidence for higher molecular weight aggregates at similar denaturant and protein concentrations by NMR measurements of diffusion (*33*). No aggregates have been detected by analytical ultracentrifugation, small-angle X-ray scattering, or size exclusion chromatography under similar conditions (data not shown). Further, it is hard to imagine how aggregation might affect residues in the same protein molecule so differently. This behavior is not specific to urea, since missing intensity for the same resonances occurs when guanidine hydrochloride is the denaturant (*11*, data not shown). Missing intensity has been observed in HSQC experiments for ^{15}N-IFABP for some of the amide proton resonances at similar urea concentrations, although the resonances were not assigned (*34*). As such, the missing intensity was not caused by fluorine labeling.

Chemical Exchange. Chemical exchange on the intermediate exchange time scale is the most likely explanation for this phenomenon. Chemical exchange occurs when a nucleus is in conformational exchange between two or more chemical environments (*35*). The simplest case of chemical exchange assumes two conformations for a particular nucleus, each of which has a unique chemical shift. If the exchange rate is slow compared to the difference in chemical shift (k in s^{-1} $\ll \Delta$ frequency in s^{-1}), two peaks will appear in the NMR

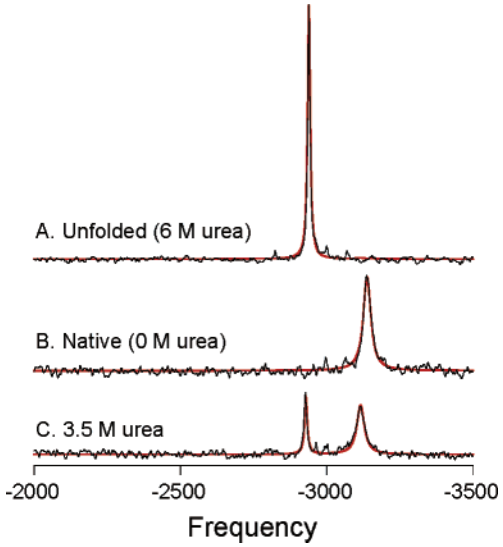

FIGURE 11: Spectra of 6Ftrp-W82Y-IFABP at selected concentrations of urea with simulations of chemical exchange. All plots are at the same intensity for the internal standard at 0 Hz (4Fphe). The simulations for A and B assume 100% unfolded and native, respectively. The simulation in C assumes a N ⇔ U model for exchange, with 77% population of the native state and 23% population of the unfolded state. The rate of exchange was 1 s^{-1}.

spectra, one for each conformation. The intensity of each peak is proportional to the population of that conformation. If the rate of exchange is fast compared to the chemical shift difference ($k \gg \Delta$ frequency), a single peak will appear at the population weighted average frequency. The fast exchange regime is expected for a fully unfolded protein, with many conformations, each with a different chemical shift, in rapid exchange with each other. Finally, if the rate of exchange is similar to the chemical shift difference ($0.02k < \Delta$ frequency $< 50k$), significant line broadening will occur. Near coalescence ($k \approx \Delta$ frequency) the resonance intensity can be completely absent (*35*, *36*). The use of chemical exchange to explain the dependence of NMR spectra on denaturant concentration is not unprecedented and has been used to measure the rate of folding in rapid folding proteins (*36*−*39*).

Figure 11 shows simulations for chemical exchange for 6Ftrp-W82Y-IFABP using WinDNMR (*20*). The simulation procedure requires considerable knowledge of the system, including the frequencies, line widths, populations, and rates of exchange for all of the participating states. Fortunately, much of this information can be obtained from spectra where no chemical exchange is occurring (see Supporting Information). The model for chemical exchange at this site assumed a simple N ⇔ U transition, using the values for frequency and line width for the N and U states from the spectra in Figure 11. All of the spectra fit well to this model as long as the rate of exchange between N and U was <5 s^{-1}. This rate agrees with the rates observed by stopped-flow fluorescence and stopped-flow CD for WT-IFABP and W82Y-IFABP (*17*, *23*). Further, stopped-flow NMR has been used to directly measure the rate of decrease in intensity for the native 6Ftrp-W6 resonance ($k \approx 0.1$ s^{-1}) for an unfolding transition from 0 to 4.5 M urea (*5*). Only the population of the N and U states changed with urea concentration. Figure 12 shows a comparison of the dependence of the unfolding of 6Ftrp-W6Y-IFABP by fluorescence intensity and by NMR

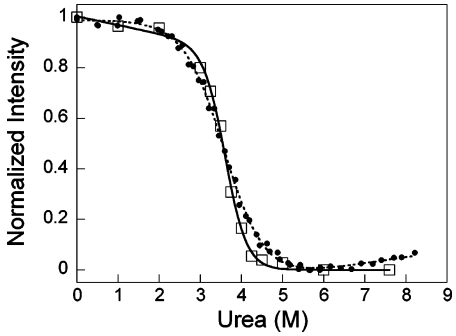

FIGURE 12: Normalized fluorescence and NMR intensities. The dependence of the unfolding of 6Ftrp-W82Y-IFABP on denaturant concentration by fluorescence (●) and NMR (□) native state intensities predicted by simulation. The lines through the data are the fit to a two-state model for unfolding at equilibrium for each data set.

simulations to a two-state model for conformational exchange for this nucleus. The midpoints for unfolding were identical by these methods (3.60 vs 3.61 M urea). The transition may be slightly more cooperative by NMR intensity measurements than by fluorescence (m_G of 1.8 vs 1.1), but there are not enough data points to be certain of this result. Further, the m_G value for IFABP is unusually low for a protein of this size (*24*). If the low m_G for WT-IFABP is due to the presence of the intermediate state, then this site may show a more cooperative transition (higher m_G value), because it does not participate in the intermediate.

Figure 13 shows simulations to chemical exchange models for 6Ftrp-W6Y-IFABP. The first simulated model assumes a simple N ⇔ U model for the transition, using the values for frequency and line width for the N and U states from the spectra in Figure 13A and Figure 13B. Figure 13C shows the best fit simulation for the N ⇔ U model to the actual data at 4.75 M urea. The populations of N and U were systematically varied between 10% and 90% and the rate of exchange was varied over 3 orders of magnitude (1−1000 s^{-1}) without finding a better fit. Thus, the equilibrium unfolding of 6Ftrp-W6Y-IFABP by NMR is not two state, even though this protein fits well to a two state model for unfolding at equilibrium by optical methods. As such, although some residues (W6) significantly populate only the native and unfolded conformations at equilibrium, other residues (W82) must populate additional conformations.

Figure 13D shows the simulation of the chemical exchange process for a more complex model, N ⇔ I ⇔ U for 6Ftrp-W6Y-IFABP at 4.75 M urea. There are two rate constants, the first for exchange between N and I, and the second between I and U. The line width of the intermediate state is not known, but should be between those of the native state (35 Hz) and the unfolded state (10 Hz). Simulations at these extremes indicated that any spectral differences caused by variations in line width of the intermediate state were small compared to the other simulated variables, so the line width was set to 20 Hz. The population of the N state was set to 5%, based on the predicted concentration of the N state at 4.75 M urea (Table 1), leaving four variables to simulate: the populations of the I and U states, the rate of exchange between I and U, and the frequency of the W82 resonance in the I conformation.

In order to obtain a flat spectrum for chemical exchange at this site, the populations of both the intermediate state

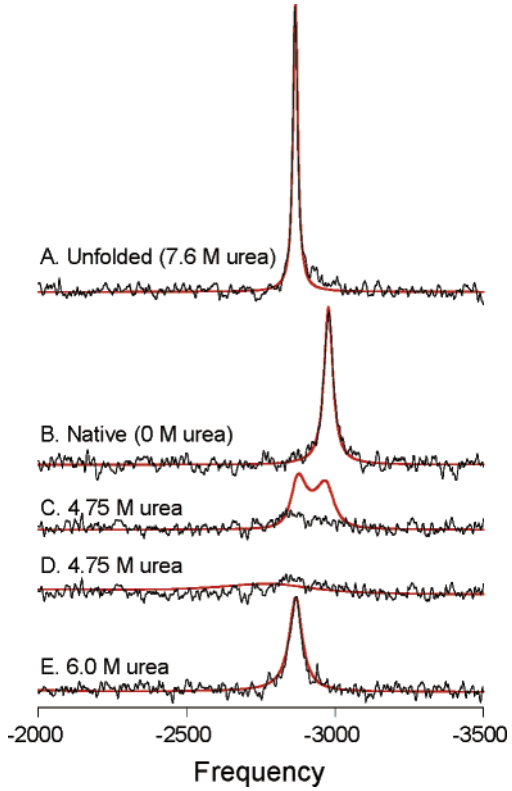

FIGURE 13: Spectra of 6Ftrp-W6Y-IFABP at selected concentrations of urea with simulations of chemical exchange. All plots are at the same intensity for the internal standard at 0 Hz (4Fphe). The simulations for A and B assume 100% unfolded and 100% native, respectively. The simulation in C assumes a N ⇔ U model for exchange, with 50% population of each state. The rate of exchange was 110 s^{-1}. The simulation in D assumes a N ⇔ I ⇔ U model. The populations of N, I, and U were 5%, 40%, and 55%, respectively. The rate of exchange for N to I and for I to U was 1 s^{-1} and 800 s^{-1}, respectively. The simulation of spectrum E assumes a N ⇔ I ⇔ U model. The populations of N, I, and U were 0.1%, 10%, and 89.9%, respectively. The rate of exchange was 1000 s^{-1}.

and the unfolded state must be between 30% and 70%, regardless of the values for the rate of exchange and frequency of W82 in the I conformation. For example, in Figure 12D, the populations of the intermediate and unfolded states were set to 40% and 55%, respectively, resulting in a good fit to the data. If the population of either the intermediate or unfolded state is greater than 70%, a peak appears near the frequency of the more highly populated state. This result is shown in Figure 13E, where the protein is 90% in the unfolded state, and a peak is observed at the unfolded state frequency.

The frequency of the intermediate state must be >300 Hz away from the unfolded state resonance, or a broad but still obvious peak would be visible. The simulation in Figure 13C is similar to those observed when the frequency of the intermediate state in a N ⇔ I ⇔ U model is less than 300 Hz away. The frequency for the intermediate state can be either upfield or downfield of the unfolded state frequency without changing the fit. The exchange rate between the intermediate and unfolded states must be between 50% and 300% of the frequency difference to broaden the signal into the baseline. If the frequency of the intermediate resonance is 800 Hz away from the unfolded resonance, the rate of exchange must be between 400 and 2400 s^{-1}, as shown in Figure 13D. Increasing the difference in frequency results in equally good fits so long as the rate of exchange is increased in proportion. Future experiments will use different field strengths, temperatures, and pH values in an attempt to modulate the rate of exchange between the intermediate and unfolded states in order to better determine the actual frequency for the W82 resonance in the intermediate state.

This kind of chemical exchange behavior would explain the dependence of the spectra on urea concentration for 3Ftyr-IFABP and 4Fphe-IFABP as well. There is missing intensity for Y70 in 3Ftyr-IFABP and for F68, F93, and at least two other phenylalanines in 4Fphe-IFABP over the same range of denaturant concentration observed for the W82 resonance. Some of these resonances appear to move from the intermediate chemical exchange regime into the fast chemical exchange regime, since both the frequency and line width of a few of the peaks in the 4Fphe-IFABP spectra change dramatically (Figure 6, peaks B and F). However, it is equally clear that not all residues participate in the structure of the intermediate: W6, Y14, Y117, Y119, and at least 2 phenylalanines (Figure 6, peaks A and E) do not show line broadening between 5 and 8 M urea.

The behavior of the mutant proteins also supports the presence of a populated intermediate. The native state of IFABP is destabilized by the W6Y mutation, and the intensity of the native W82 resonance in 6Ftrp-W6Y-IFABP begins to decrease at lower concentrations of denaturant than for 6Ftrp-WT-IFABP (Figure 3). However, W6 does not participate in the intermediate state, and the stability of the intermediate state is not affected. Thus, the unfolded resonance of W82 appears over a similar denaturant concentration range as in 6Ftrp-WT-IFABP. The F68A and F93A mutations destabilize the native conformation. However, these sites participate in the intermediate as well, and these mutations appear to destabilize the intermediate conformation (Figures 5−10). As such, the spectrum of the extended unfolded state appears at lower urea concentrations in 4Fphe-F68A-IFABP and 4Fphe-F93A-IFABP than in 4Fphe-WT-IFABP. Since these mutations affect all of the 4Fphe resonances that participate in the intermediate in the same manner, this conformation may act as a unified thermodynamic unit, with all of these residues participating in the same state at the same time.

This model for chemical exchange is the simplest model that fits the data. More complex models with additional conformational states are also consistent with the data. However, the exchange between these conformations must remain in the intermediate chemical exchange time regime because a flat spectrum was observed for 6Ftrp-W6Y-IFABP at 4.5 M urea when the number of transients was increased to 8192 (data not shown). If exchange between many states was slow, small peaks representing those conformations should have been observed in this spectra. If multiple conformations were in fast exchange, then an average peak would have been observed, like in the unfolded protein.

Comparison of the Properties of the NMR Intermediate to the Kinetic Intermediates. The spectral properties of the NMR intermediate (little if any secondary structure and fluorescence similar if not identical to the unfolded state) are quite similar to those of one of the kinetic intermediates observed during the unfolding of this protein. This intermediate has little if any secondary structure by stopped-flow CD experiments, and the fluorescence spectrum of the kinetic

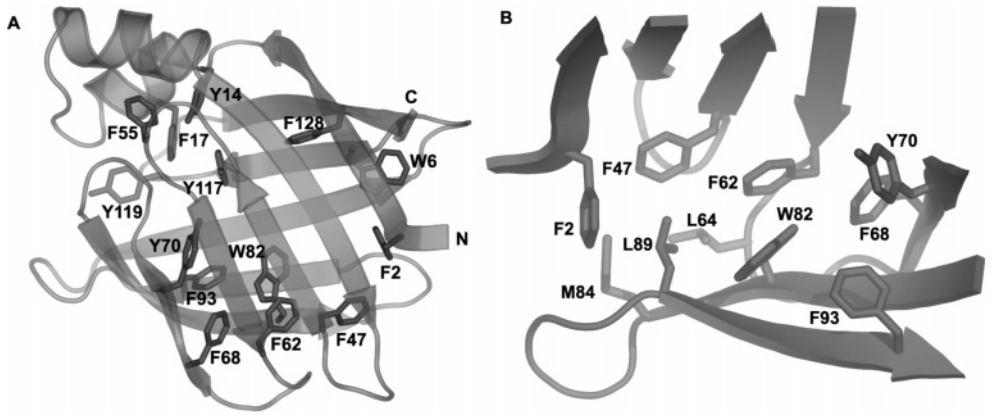

FIGURE 14: (A) Structure of IFABP (1ifc) showing the locations of the N and C termini and all of the aromatic residues. F68, Y70, W82, and F93 participate in the intermediate observed by NMR whereas W6, Y14, Y119, and Y117 do not. (B) The hydrophobic cluster in IFABP. The point of view of this picture is from the vicinity of the helices looking toward the base of the protein between the sheets.

intermediate has the same wavelength of maximal fluorescence as the unfolded state, but is slightly more intense. We have shown that only W82 participates in the kinetic intermediate (*23*), and W82 also shows the intermediate exchange behavior described here. However, even though they have similar spectral properties, the rate of exchange is very different. The stopped-flow data suggests that the kinetic intermediate exchanges with the unfolded state at about 1 s^{-1}, compared to an exchange rate of at least 200 s^{-1} for the NMR intermediate and the unfolded state. As such, they are not identical states. The data suggests that I_{NMR} is more unfolded than the kinetic intermediate, and thus is even more similar to the unfolded state in its spectral properties.

Structure of the Intermediate. Figure 14 shows the hydrophobic core of native IFABP. All of the residues identified as participating in the intermediate detected by ^{19}F NMR are neighbors in the native state. There are three other phenylalanines (F2, F47, F62) in this cluster which might account for the other phenylalanine residues that show chemical exchange behavior. Finally, F55 is between F47 and F62 in the sequence, and represents another residue whose NMR behavior might be affected by the intermediate.

A number of proteins have been shown by NMR to have both native and nonnative interactions involving hydrophobic residues at concentrations of denaturant sufficient to unfold the protein by optical methods (reviewed in ref *2*). For example, there is both native and nonnative clustering of hydrophobic residues in lysozyme (*40*). At this point in time only nativelike interactions have been detected, but that does not mean that all of the contacts in this structure are nativelike. On the other hand, several residues that are not in this cluster in the native state (W6, Y14, Y117, Y119, and at least 2 phenylalanines) do not participate in the intermediate.

The residues in the intermediate are not very close together in the sequence of the protein, only in the native structure. Given the changes in the optical and NMR properties of residues in this cluster, the structure of the intermediate is certainly not identical to the native state. Rather, the intermediate is likely to be an ensemble of related structures that bring these residues together. A peptide corresponding to the central hairpin loop in this sequence (F62−Y70) for another protein in this family has been shown to have significant secondary structure in solution (*41*, *42*). This hairpin might serve as a local hydrophobic cluster upon which additional hydrophobic residues can assemble from further away in the sequence. Hence, gathering these long range interactions together lays down a template upon which secondary structure can form quickly and accurately, greatly reducing the conformational space that needs to be searched to obtain the native fold (*1*). This hypothesis might provide a general mechanism for the folding of β-sheet proteins.

ACKNOWLEDGMENT

We acknowledge useful discussions with Drs. Maria Bewley and John Flanagan at the Penn State University College of Medicine, Drs. David Shortle, Blake Hill, and Joel Tolman at Johns Hopkins University, and Dr. Carl Frieden at Washington University.

SUPPORTING INFORMATION AVAILABLE

An additional figure, a table of resonance assignments and line widths, and a discussion of the line fitting procedures and assumptions are found in the supplement. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Daggett, V., and Fersht, A. (2003) The present view of the mechanism of protein folding, *Nat. Rev. Mol. Cell. Biol. 4*, 497−502.
2. Dyson, H. J., and Wright, P. E. (2004) Unfolded proteins and protein folding studied by NMR, *Chem. Rev. 104*, 3607−3622.
3. Danielson, M. A., and Falke, J. J. (1996) Use of ^{19}F NMR to probe protein structure and conformational changes, *Annu. Rev. Biophys. Biomol. Struct. 25*, 163−195.
4. Frieden, C., Hoeltzli, S. D., and Bann, J. G. (2004) The preparation of ^{19}F-labeled proteins for NMR studies, *Methods Enzymol. 380*, 400−415.
5. Frieden, C., Hoeltzli, S. D. and Ropson, I. J. (1993) NMR and protein folding: Equilibrium and stopped-flow studies, *Protein Sci. 2*, 2007−2014.
6. Hoeltzli, S. D. and Frieden, C. (1996) Real-time refolding studies of 6-^{19}F-tryptophan labeled *Escherichia coli* dihydrofolate reductase using stopped-flow NMR spectroscopy, *Biochemistry 35*, 16843−16851.
7. Bann, J. G., Pinkner, J., Hultgren, S. J., and Frieden, C. (2002) Real-time and equilibrium ^{19}F-NMR studies reveal the role of domain-domain interactions in the folding of the chaperone PapD, *Proc. Natl. Acad. Sci. U.S.A. 99*, 709−714.
8. Schuler, B., Kremer, W., Kalbitzer, H. R., and Jaenicke, R. (2002) Role of entropy in protein thermostability: Folding kinetics of a

hyperthermophilic cold shock protein at high temperatures using ^{19}F-NMR, *Biochemistry 41*, 11670−11680.
9. Bai, P., Luo, L., and Peng, Z. (2000) Side chain accessibility and dynamics in the molten globule state of α-lactalbumin: A ^{19}F-NMR Study, *Biochemistry 39*, 372−380.
10. Shu, Q., and Frieden, C. (2004) Urea-dependent unfolding of murine adenosine deaminase: Sequential destabilization as measured by ^{19}F NMR, *Biochemistry 43*, 1432−1439.
11. Ropson, I. J., and Frieden, C. (1992) Dynamic NMR spectral analysis and protein folding: Identification of a highly populated folding intermediate of rat intestinal fatty acid binding protein by ^{19}F NMR, *Proc. Natl. Acad. Sci. U.S.A.* 89, 7222−7226.
12. Sacchettini, J. C., Banaszak, L. J., and Gordon, J. I. (1990) Expression of rat intestinal fatty acid binding protein in *E. coli* and its subsequent structural analysis: A model for studying the molecular details of fatty acid-protein interaction, *Mol. Cell. Biochem. 98*, 81−93.
13. Johnson, B. H., and Hecht, M. H. (1994) Recombinant proteins can be isolated from *E. coli* cells by repeated cycles of freezing and thawing, *Biotechnology 12*, 1357−1360.
14. Dalessio, P. M., and Ropson, I. J. (2000) β-Sheet proteins with nearly identical structures have different folding intermediates, *Biochemistry 39*, 860−871.
15. Kim, K., and Frieden, C. (1998) Turn scanning by site-directed mutagenesis: Application to the protein folding problem using the intestinal fatty acid binding protein, *Protein Sci. 7*, 1821−1828.
16. Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption coefficient of a protein, *Protein Sci. 4*, 2411−2423.
17. Ropson, I. J. and Dalessio, P. M. (1997) Fluorescence spectral changes during the folding of intestinal fatty acid binding protein, *Biochemistry 36*, 8594−8601.
18. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol. 131*, 266−280.
19. Gualfetti, P. J., Bilsel, O, and Matthews, C. R. (1999) The progressive development of structure and stability during the equilibrium unfolding of the α-subunit of tryptophan synthase from *Escherichia coli*. *Protein Sci. 8*, 1623−1635.
20. Reich, H. J. (1996) WinDNMR−Dynamic NMR Spectra for Windows, *J. Chem. Educ. Software 3D2.*
21. Eisenhaber, F., and Argos, P. (1993) Improved strategy in analytic surface calculation for molecular systems: Handling of singularities and computational efficiency, *J. Comput. Chem. 14*, 1272−1280.
22. Ropson, I. J., Gordon, J. I., and Frieden, C. (1990) Folding of a predominately β-structure protein: rat intestinal fatty acid binding protein, *Biochemistry 29*, 9591−9599.
23. Dalessio, P. M., Fromholt, S. E., and Ropson, I. J. (2005) The role of Trp-82 in the folding of intestinal fatty acid binding protein, *Proteins 61*, 176−183.
24. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding, *Protein Sci. 4*, 2138−2148.
25. Burns, L. L., Dalessio, P. M., and Ropson, I. J. (1998) Folding mechanism of three structurally similar β-sheet proteins, *Proteins 33*, 107−118.
26. Burns, L. L., and Ropson, I. J. (2001) Folding of intracellular retinol and retinoic acid binding proteins, *Proteins 43*, 292−302.
27. Xu, Z. J., Love, M. L., Ma, L. Y. Y., Blum, M., Bronskill, P. M., Bernstein, J., Grey, A. A., Hofmann, T., Camerman, N., and Wong, J. T. F. (1989) Tryptophanyl-tRNA synthetase from *Bacillus subtilis*, *J. Biol. Chem. 264*, 4304−4311.
28. Lee, K. H., Lee, H. Y., Slutsky, M. M., Anderson, J. T., and Marsh, E. N. G. (2004) Fluorous effect in proteins: *De novo* design and characterization of a four-α-helix bundle protein containing hexafluoroleucine, *Biochemistry 43*, 16277−16284.
29. Li, H., and Frieden, C. (2005) NMR studies of 4-^{19}F-phenylalanine-labeled intestinal fatty acid binding protein: Evidence for conformational heterogeneity in the native state, *Biochemistry 44*, 2369−2477.
30. Scapin, G., Gordon, J. I., and Sacchettini, J. C. (1992) Refinement of the structure of recombinant rat intestinal fatty acid-binding apoprotein at 1.2 Å resolution, *J. Biol. Chem. 267*, 4253−4269.
31. Yeh, S.-R., Ropson, I. J., and Rousseau, D. L. (2001) Hierarchical folding of intestinal fatty acid binding protein, *Biochemistry 40*, 4205−4210.
32. Chattopadhyay, K., Zhong, S., Yeh, S.-R., Rousseau, D. L., and Frieden, C. (2002) The intestinal fatty acid binding protein: the role of turns in fast and slow folding processes, *Biochemistry 41*, 4040−4047.
33. Li, H., and Frieden, C. (2005) Phenylalanine side chain behavior of the intestinal fatty acid binding protein: The effect of urea on backbone and side chain stability, *J. Biol. Chem.*, in press.
34. Hodsdon, M. E., and Frieden, C. (2001) Intestinal fatty acid binding protein: The folding mechanism as determined by NMR studies, *Biochemistry 40*, 732−742.
35. Sandstrom, J. (1982) *Dynamic NMR Spectroscopy*, Academic Press, New York.
36. Myers, J. K., and Oas, T. G. (2001) Preorganized secondary structure as an important determinant of fast protein folding, *Nat. Struct. Biol. 8*, 552−558.
37. Huang, G. S. and Oas, T. G. (1995) Submillisecond folding of monomeric λ repressor, *Proc. Natl. Acad. Sci. U.S.A.* 92, 6878−6882.
38. Burton, R. E., Huang, G. S., Daugherty, M. A., Fullbright, P. W., and Oas, T. G. (1996) Microsecond protein folding through a compact transition state, *J. Mol. Biol. 263*, 311−322.
39. Spector, S., and Raleigh, D. P. (1999) Submillisecond folding of the peripheral subunit-binding domain, *J. Mol. Biol. 293*, 763−768.
40. Klein-Seetharaman, J., Oikawa, M., Grimshaw, S. B., Wirmer, J., Duchardt, E., Ueda, T., Imoto, T, Smith, L. J., Dobson, C. M., and Schwalbe, H. Long-range interactions within a nonnative protein, *Science 295*, 1719−1722.
41. Rotondi, K. S., and Gierasch, L. M. (2003) Role of local sequence in the folding of cellular retinoic acid binding protein I: Structural propensities of reverse turns, *Biochemistry 42*, 7976−7985.
42. Rotondi, K. S., Rotondi, L. F., and Gierasch, L. M. (2003) native structural propensity in cellular retinoic acid-binding protein I 64-88: The role of locally encoded structure in the folding of a β-barrel protein, *Biophys. Chem. 100*, 421−436.

BI052091O