

## The Intestinal Fatty Acid Binding Protein: The Role of Turns in Fast and Slow Folding Processes<sup>†</sup>

Krishnananda Chattopadhyay,<sup>‡</sup> Shi Zhong,<sup>§</sup> Syun-Ru Yeh,<sup>§</sup> Denis L. Rousseau,<sup>§</sup> and Carl Frieden<sup>\*‡</sup>

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, Missouri 63110, and Department of Physiology and Biophysics, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461

Received November 12, 2001; Revised Manuscript Received January 29, 2002

**ABSTRACT:** The intestinal fatty acid binding protein is one of a family of proteins that are composed of two  $\beta$ -sheets surrounding a large interior cavity into which the ligand binds. Glycine residues occur in many of the turns between adjacent antiparallel  $\beta$ -strands. In previous work, the effect of replacing these glycine residues with valine has been examined with stopped flow instrumentation using intrinsic tryptophan fluorescence spectroscopy [Kim and Frieden (1998) *Protein Sci.* 7, 1821–1828]. To resolve the burst phase missing in the stopped flow measurements, these valine mutants have been reexamined with sub-millisecond continuous flow instrumentation. Some of the glycine residues have also been replaced with proline, and the folding reactions of these proline mutants have been compared with those of their valine counterparts. In all cases, the stability of the protein is decreased, but some turns appear to be more critical for final structure stabilization than others. Surprisingly, the rate constants observed for all the mutants measured by sub-millisecond continuous flow methods are quite similar (1400–3000 s<sup>-1</sup>), and in all the mutants, there is a shift in the fluorescence emission maximum from that of the unfolded protein to lower wavelengths, suggesting some collapse of the unfolded state within 200  $\mu$ s. In contrast to the rate constants observed for the initial folding events measured by the sub-millisecond continuous flow method, the rate constants for the slower phase observed in the stopped flow instrument vary widely for the different mutants. The latter step appears to be related to side chain stabilization rather than secondary structure formation. It is also shown that the ligand binds tightly only to the native protein and not to any intermediate forms.

The intestinal fatty acid binding protein (IFABP)<sup>1</sup> is an excellent model system for studying the role of turns in protein folding (1–8). It is small (15 kDa), monomeric, highly soluble, and contains no proline or cysteine residues. Both the crystal structure (9, 10) and the NMR structure (11) of apo-IFABP have been determined. The protein consists of two  $\beta$ -sheets, each composed of five strands surrounding a large interior cavity into which the fatty acid ligand binds. Figure 1 shows the X-ray crystal structure of IFABP with the strands labeled A–J. In previous studies of this protein (1), it was shown, using <sup>19</sup>F-labeled tryptophan, that a region near one turn (the D–E turn) appeared to form a hydrophobic cluster of residues in equilibrium with the folded and unfolded forms of the protein. In a later study, residues in the D–E turn were randomly mutated, and it was concluded that a specific residue (Leu64) in that turn appeared to be

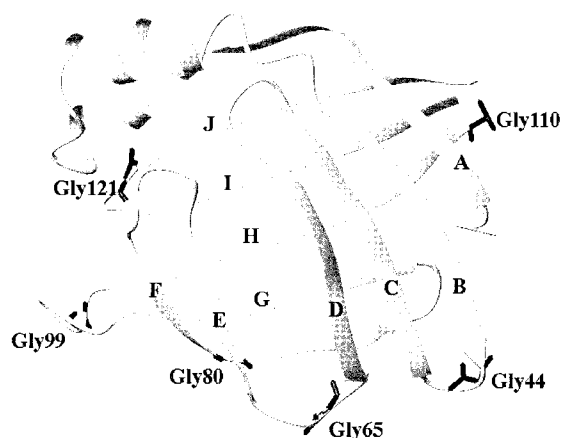


FIGURE 1: X-ray crystal structure of intestinal fatty acid binding protein (10). Positions and the numbering of the glycine residues, which have been mutated, are shown in the figure. This figure has been prepared using MOLMOL (18).

important for the stabilization of the final structure (4). Since many of the turns between antiparallel  $\beta$ -strands contain a glycine residue, a systematic study was performed to change each of the residues to valine and measure the stability and folding rates for each mutant (5). This study showed that a mutation of Gly121 in the I–J turn had a dramatic effect on the rate of refolding, slowing down that rate by over 100-fold (5). More recently, NMR studies as a function of

<sup>†</sup> Supported by NIH Grants DK13332 (to C.F.) and GM54812 (to D.L.R.).

<sup>\*</sup> To whom correspondence should be addressed at the Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, 660 S. Euclid Ave., St. Louis, MO 63110. Telephone: (314) 362 3344. Fax: (314) 362 7183. Email: [frieden@biochem.wustl.edu](mailto:frieden@biochem.wustl.edu).

<sup>‡</sup> Washington University School of Medicine.

<sup>§</sup> Albert Einstein College of Medicine.

<sup>1</sup> Abbreviations: IFABP, rat intestinal fatty acid binding protein; DAUDA, 11-[(5-dimethylaminonaphthalene-1-sulfonyl)amino]undecanoic acid; Gdn, guanidine hydrochloride.

denaturant concentration found that residues in certain turns appeared to exhibit transient but persistent structure even at high denaturant concentrations. These were near the D–E turn, the I–J turn, and the turn between the two short helical structures (7).

The kinetics of refolding of IFABP are complicated due to the presence of an unresolved burst phase observed in the stopped flow measurements. Recent sub-millisecond measurements of a Trp6Tyr mutant of IFABP have resolved the folding reaction into three phases with rate constants of  $>10\,000$  (burst phase),  $\sim 1500$ , and  $\sim 5\text{ s}^{-1}$  (8). This work prompted us to reexamine the folding reactions of wild-type and mutant proteins that contain either a valine or a proline residue substituted for the glycine of particular turns (see Figure 1 for the glycine residues mutated). The data reveal a rapid phase ( $1400\text{--}3000\text{ s}^{-1}$ ) that is relatively unchanged from wild-type to all of the mutants examined, in contrast to the large changes in the slower kinetic phase. These data suggest that there is no correlation between the early folding events, which primarily represent collapse of the unfolded protein, and the slower phase observed in the stopped flow measurements which primarily represents side chain stabilization. In addition, we compare the results to those of two mutants in which a glycine (Gly80) in a  $\beta$ -strand that connects the two  $\beta$ -sheets is mutated to either valine or proline. Finally, we show that ligand binds only to the native structure and not to any intermediate forms.

## EXPERIMENTAL PROCEDURES

**Materials.** Mutagenic primers were obtained from Integrated DNA Technologies (Coralville, ID). The QuickChange site-directed mutagenesis kit was obtained from Stratagene (LaJolla, CA). Sodium oleate was obtained from Sigma Chemical Co. (St. Louis, MO). DAUDA, obtained from Molecular Probes (Eugene, OR), was dissolved in 95% ethanol to make a 1 mM stock solution. Gdn was obtained from ICN Biochemicals (Aurora, OH). All other reagents used were analytical grade.

**Mutagenesis.** Mutagenesis, expression, and purification of IFABP have been described elsewhere (5), and the same procedure has been followed here except that the Sephadex G50 column purification step was not used.

**Equilibrium Unfolding Measurements.** These experiments were performed in 20 mM potassium phosphate buffer, pH 7.3, and monitored using a PTI Alphascan fluorometer (Photon Technology International, South Brunswick, NJ). Proteins were incubated at different Gdn concentrations for 1–2 h, and the change in fluorescence intensity at 327 nm (excitation wavelength of 290 nm) was monitored. Protein concentrations used for wild-type and mutant proteins were between 1 and 5  $\mu\text{M}$ . The concentration of Gdn was determined by measuring the refractive index of the solution. The data were fit to a six-parameter equation (12) assuming a simple two-state unfolding transition.

**Sub-millisecond Continuous Flow Measurements.** For the kinetic studies, the folding reaction was initiated by a 6-fold dilution of the unfolded protein with the appropriate buffer solution in a rapid solution mixer with a mixing dead time of  $\sim 100\ \mu\text{s}$ . In this rapid solution mixer, the two parent solutions were pressurized by a syringe pump and introduced

into the T-shaped mixing nozzle, which is in direct contact with a quartz flow cell ( $250\ \mu\text{m} \times 250\ \mu\text{m}$ ). The detailed characterization of this mixer has been described previously (13). The output at 264 nm from a frequency-doubled argon ion laser (Coherent, Santa Clara, CA) is focused to a  $\sim 30\ \mu\text{m}$  spot on the continuously flowing sample in the observation cell. The tryptophan fluorescence is collected and focused on the entrance slit of a 0.27 m polychromator (Spex, Edison, NJ) where it is dispersed and then detected by a charge-coupled device camera (Photometrics). Time-dependent behavior is obtained by moving the quartz cell relative to the laser focusing point along the flow direction. The position of the flow cell is controlled by a micrometer on a translational stage in such a way that an accurate time domain can be calculated from the solution flow rate. The first reliable data point for the spectrum of the fluorescent emission was obtained at  $100\ \mu\text{s}$  although the data shown in the figures are given at  $200\ \mu\text{s}$ . For reliable kinetic traces of the change in the fluorescent intensity, the first point used was that measured at  $200\ \mu\text{s}$ .

**Wavelength Calibration.** The spectral data obtained from the sub-millisecond continuous flow system are in pixel numbers. Conversion of the pixel number into the form of wavelength (in nanometers) was carried out using multiple points to determine the dispersion factor (i.e., nm/pixel) of the spectrophotometer. A linear correlation between pixel number and wavelength was observed. The spectral output for the completely unfolded protein was compared to the steady-state tryptophan fluorescence of the unfolded protein, and the pixel number at the maximum of the sub-millisecond continuous flow spectrum was fixed at 350 nm (steady-state fluorescence maximum for unfolded IFABP). The calibrated pixel numbers are about 10 nm higher than observed with a conventional spectrophotometer and were not used in any quantitative spectral characterization. Instead, the converted spectra generated by this procedure have been used only to qualitatively explain different phases of folding events.

**Stopped Flow Fluorescence.** These experiments were performed using an Applied Photophysics Stopped Flow Spectrophotometer (model SX18MV) and final protein concentrations between 2 and 5  $\mu\text{M}$ . For refolding experiments, mutants were incubated in 2 M Gdn (for valine mutants) or 1.5 M Gdn (for proline mutants) for 1 h and diluted 6-fold to varying Gdn concentrations. All experiments were performed in 20 mM phosphate buffer, pH 7.3 at 20 °C. Unfolding experiments were performed using protein in 25 mM phosphate buffer and diluting 6-fold into Gdn solutions of varying concentrations. Both folding and unfolding kinetics were monitored by observing the change in fluorescence using an excitation wavelength of 290 nm and a cutoff filter of 305 nm. For the folding experiments in the presence of DAUDA, the excitation wavelength used was 350 nm, and a cutoff filter of 450 nm was used.

**Stopped Flow Circular Dichroism.** These experiments were performed using an Applied Photophysics RX1000 rapid kinetics accessory attached to a Jasco J715 spectropolarimeter. The dead time of the mixer is less than 50 ms. The protein samples were incubated in 2 M Gdn for 1 h and then diluted 6-fold in 20 mM potassium phosphate buffer (pH 7.3) at 20 °C. The refolding kinetics were monitored by CD at 216 nm.

Table 1: Turn Mutations, Turn Sequences, and Connecting Strands

mutations	residues	WT sequence	strands involved
G44V/P	43–46	EGNK	B–C
G65V/P	63–66	ELGV	D–E
G99V/P	96–99	VDNG	G–H
G110V/P	109–112	SGNE	H–I
G121V/P	119–122	YEGV	I–J
G80V/P <sup>a</sup>	78–82	LTGTW	half-turn connecting $\beta$ -sheets

<sup>a</sup> No experiments could be performed with G80P because of protein instability.

Table 2: Parameters for the Two-State Unfolding Analysis of Proline Turn Mutants<sup>a</sup>

mutants	$\Delta G^\circ$ (kcal/mol)	midpoint (M)	$m^b$ (kcal mol <sup>-1</sup> M <sup>-1</sup> )
wild-type	5	1.4	-3.8
G44P	3.1	0.8	-4
G65P	2.5	0.9	-2.8
G99P	3.5	0.7	-4.8
G110P	5	1	-5
G121P	1.2	0.6	-2

<sup>a</sup> All experiments were performed at pH 7.3, 20 mM potassium phosphate at 20 °C using Gdn as the denaturant. <sup>b</sup> Slope of the denaturation curve assuming a two-state model.

## RESULTS

Table 1 lists the turn mutations made and the adjacent strands. The rationale for these particular mutations was to examine the importance of turn structures in the stability and refolding kinetics. We previously examined glycine to valine mutations at these turn positions using stopped flow methods (5). Here those data are extended to mutations involving proline as well as measuring the refolding kinetics using sub-millisecond continuous flow methods. This latter method allows determination of rate constants about 20-fold greater than those measured by stopped flow, thereby allowing characterization of the faster folding steps of IFABP (8).

**Equilibrium Unfolding Transitions.** IFABP unfolding is accompanied by a significant loss of fluorescence intensity with a shift in the emission maximum to 350 nm. Thermodynamic parameters using a two-state fit of the unfolding transitions of valine mutants have been reported earlier (5), and the data for proline mutants are summarized in Table 2. Figure 2 shows the equilibrium unfolding transitions of representative valine and proline mutants. From the equilibrium unfolding transitions, these proteins could be divided into two broad categories. For wild-type, Gly44 (Figure 2a), Gly99, and Gly110 mutants, the equilibrium unfolding transitions are reasonably cooperative (Table 2). In contrast, the unfolding transitions of Gly65, Gly80, and Gly121 (Figure 2b) mutants are less cooperative with small  $m$  values, and a simple two-state model may not be appropriate for these mutants. The unfolding transitions of all proline mutants, except G44P, are remarkably similar to their valine counterparts in the shape and extent of cooperativity although the proline mutants are less stable than the corresponding valine mutants. As mentioned, the only exception is G44P which is slightly less cooperative than the G44V mutant (Figure 2a). The reasonably high cooperativity shown by the G99P and G110P mutants may reflect low concentrations of intermediate forms or loss of residual structure in the unfolded state.

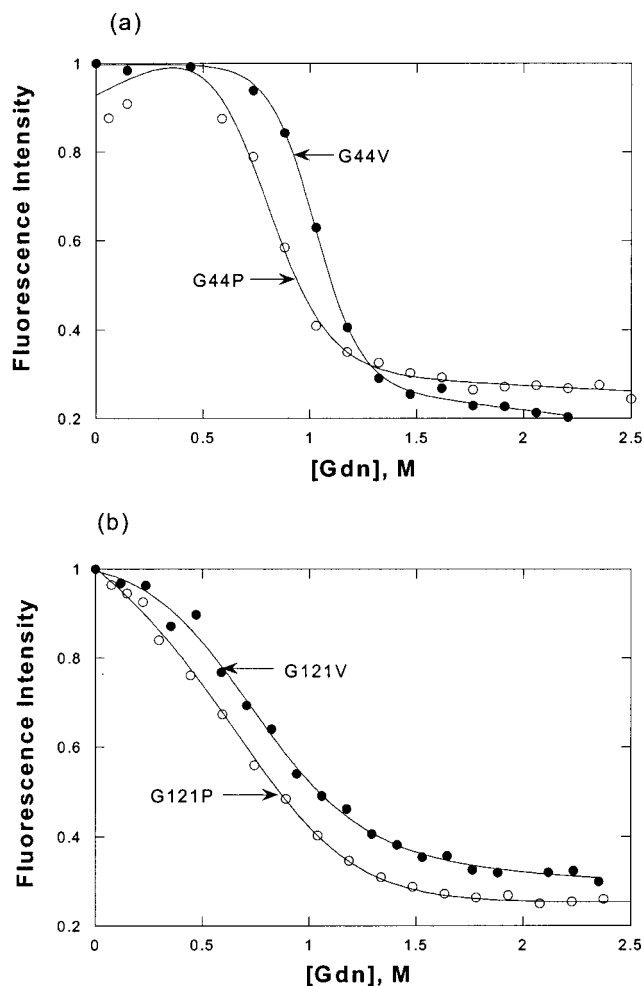


FIGURE 2: Comparisons of equilibrium unfolding transitions between the glycine to valine and glycine to proline mutants: (a) unfolding transitions of G44V (●) and G44P (○); (b) unfolding transitions of G121V (●) and G121P (○). Experiments were performed in 20 mM phosphate buffer at pH 7.3 at 20 °C. The fluorescence intensity at 327 nm was normalized to a maximum value of 1. The solid line through the data corresponds to the fit using a simple two-state unfolding transition (12).

**Unfolding and Refolding Kinetics of IFABP Mutants.** The refolding kinetics of IFABP and its different mutants have been studied by sub-millisecond continuous flow and by stopped flow methods. The same general scheme as used previously (8) will be used to describe the data (eq 1):



where U and N represent the respective unfolded and folded state and  $I_1$  and  $I_2$  represent two kinetic intermediates. The formation of  $I_2$  from the unfolded state (U) is the so-called ‘burst phase’ that takes place in less than 2 ms (the dead time of the stopped flow instrument). The kinetics of this burst phase were monitored by the change in tryptophan fluorescence with time using a sub-millisecond continuous flow mixer while the kinetics of the slower formation of N from  $I_2$  were monitored using stopped flow methods. The unfolding of IFABP does not show any burst phase or missing amplitude in stopped flow measurements, and the unfolding can be fit by a single exponential change in fluorescence intensity with time. This kinetic phase of unfolding makes a typical chevron plot with the stopped flow refolding rate (shown later in Figure 5).

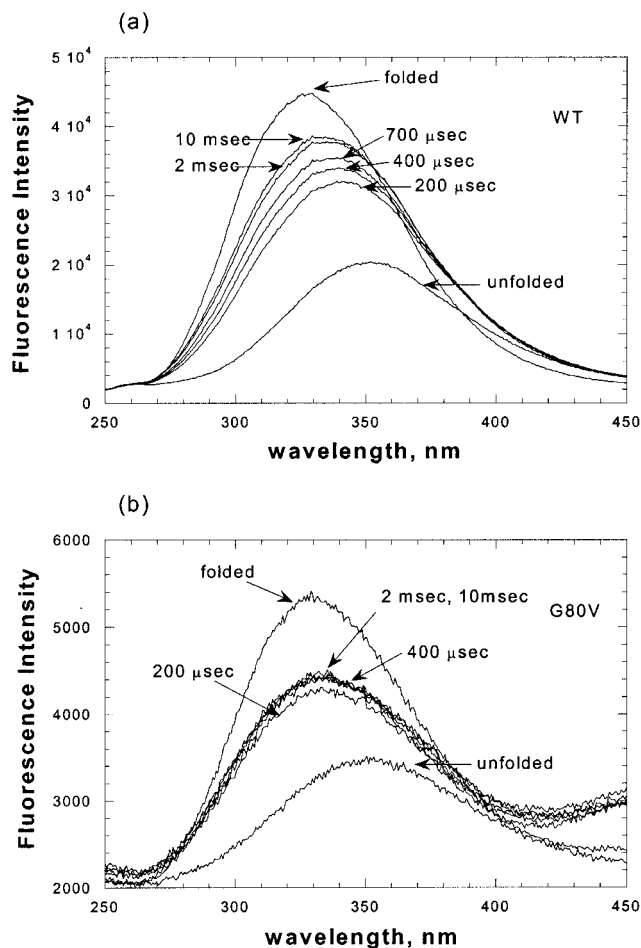


FIGURE 3: Fluorescence emission spectra of (a) wild-type and (b) G80V at different times obtained in sub-millisecond continuous flow experiments. The time points are labeled in the figure. The experiments were performed in 20 mM potassium phosphate buffer at pH 7.3 at room temperature. The protein samples were taken up in 2 M Gdn solution in buffer and diluted 6-fold with buffer. The folded spectra for both the proteins were obtained after waiting a sufficiently long time to complete the folding. The unfolded spectra for both the proteins were determined from protein samples (in 2 M Gdn) at the same final protein concentration.

**Sub-millisecond Continuous Flow Measurements.** Figure 3a shows representative fluorescence spectra of wild-type IFABP at different time points monitored by sub-millisecond continuous flow experiments. Fluorescence spectra of unfolded and folded IFABP are shown in the figure for comparison. Figure 3a indicates that a significant increase in fluorescence intensity takes place before 200  $\mu$ s and that this phase (the U to  $I_1$  transition in eq 1) is too fast to be measured even by the sub-millisecond continuous flow method. This fluorescence intensity change, which occurs within 200  $\mu$ s, is accompanied by a significant blue shift in the fluorescence emission maximum from 350 nm in the unfolded state (U) to 340 nm at 200  $\mu$ s ( $I_1$ ). A systematic shift in the fluorescence emission maximum of wild-type IFABP takes place between 200  $\mu$ s and 2 ms (the  $I_1$  to  $I_2$  transition). The fluorescence emission maximum of wild-type at 2 ms is at 330 nm, and little change is observed between 2 and 10 ms. The completely folded state (N) with an emission maximum at 327 nm is formed in a slow kinetic phase ( $I_2$  to N) as observed in stopped flow experiments.

Figure 3b shows fluorescence spectra of G80V at different times of refolding. In contrast to wild-type and most other

Table 3: Results from the Sub-millisecond Continuous Flow and Stopped Flow Experiments on Turn Mutants of IFABP

mutants	$k$ ( $s^{-1}$ ) <sup>a</sup>	$I(200\mu s)^{max}$ (nm) <sup>b</sup>	$I(2ms)^{max}$ (nm) <sup>b</sup>	$k_u^0$ ( $s^{-1}$ ) <sup>c</sup>	$k_f^0$ ( $s^{-1}$ ) <sup>c</sup>	$k_u$ ( $s^{-1}$ ) <sup>d</sup>	$k_f$ ( $s^{-1}$ ) <sup>d</sup>
WT	2000	341	334	0.006	75	0.34	7
G44V	2040	340	332	0.02	33		
G44P	2500	337	332			2.6	4.9
G65V	1900	339	334	0.38	32		
G65P	3000	344	342				Ns <sup>e</sup>
G80V	Nd <sup>f</sup>	332	332	0.008	0.74		
G99V	1400	339	335	0.005	2.8		
G99P	2650	336	332			1	0.59
G110V	2000	339	334	0.007	4.2		
G110P	2400	345	345			0.11	0.94
G121V	Nd <sup>f</sup>	345	341	0.002	0.11		
G121P	2500	342	340			0.35	0.2
							0.009 <sup>g</sup>

<sup>a</sup>  $k$  is the rate of folding measured by sub-millisecond continuous flow experiments. Initial Gdn concentration for the valine mutants is 2 M and for proline mutants is 1.5 M. Experiments performed in 20 mM potassium phosphate at pH 7.3. Errors involved in the rate calculation are within 10%. <sup>b</sup>  $I(200\mu s)^{max}$  and  $I(2ms)^{max}$  are the emission maxima of intermediate  $I(200\mu s)$  and  $I(2ms)$ , respectively. These values are for qualitative comparisons only. In all cases, the value of the fluorescence maximum for the unfolded protein was assumed to be 350 nm. <sup>c</sup>  $k_u^0$  and  $k_f^0$  are the rate of unfolding and folding, respectively, of valine mutants extrapolated to zero Gdn concentration. These rates were measured by stopped flow fluorescence. <sup>d</sup>  $k_u$  and  $k_f$  are the rate of unfolding and folding, respectively, of the proline mutants (1.5 and 0.25 M Gdn, respectively). Errors involved in the rate calculation are less than 10%. <sup>e</sup> Ns: not seen; no fluorescence change observed in stopped flow measurements. <sup>f</sup> Nd: not detected; unable to detect this phase (see text). <sup>g</sup> G121P has an additional kinetic phase in refolding. Amplitudes of these two phases are almost identical.

turn mutants, the emission maximum of G80V at 200  $\mu$ s is at 330 nm and does not shift with time. Furthermore, the change in fluorescence intensity between 200  $\mu$ s and 2 ms ( $I_1$  to  $I_2$ ) is very small for this mutant protein. Similarly, only small changes in fluorescence intensity between 200  $\mu$ s and 2 ms have been found for three more mutants: G121V, G121P, and G110P.

Table 3 lists emission maxima at 200  $\mu$ s and 2 ms for the wild-type protein and different turn mutants. The spectra at 200  $\mu$ s [defined as  $I(200\mu s)$ ] and at 2 ms [defined as  $I(2ms)$ ] are qualitatively similar for wild-type, G44V, G44P, G65V, G99V, G99P, and G110V (Figure 3a and Table 3). Five mutant proteins, G65P, G80V, G121V, G121P, and G110P, show significant differences in emission maxima from wild-type for either  $I(200\mu s)$  or  $I(2ms)$  although as noted before (Experimental Procedures) these values are rather qualitative.

The kinetic traces for the transition from 200  $\mu$ s to 2 ms for wild-type and all turn mutants can be fit to a single-exponential function, and the rate constants are also given in Table 3. Figure 4 shows a representative example of the time-dependent fluorescence change using wild-type protein. For wild-type and most of the turn mutants, the rate constant for this phase ranges between 1400 and 3000  $s^{-1}$  (Table 3). For G80V and G121V, the intensity change of this phase is very small so that the transition appears to be too fast to be resolved or the two states have identical fluorescence properties.

**Stopped Flow Measurements: Fluorescence.** The slowest refolding phase ( $I_2$  to N) has been studied by stopped flow experiments. The stopped flow data can be fit to a single-exponential function for all the proteins (except G121P and

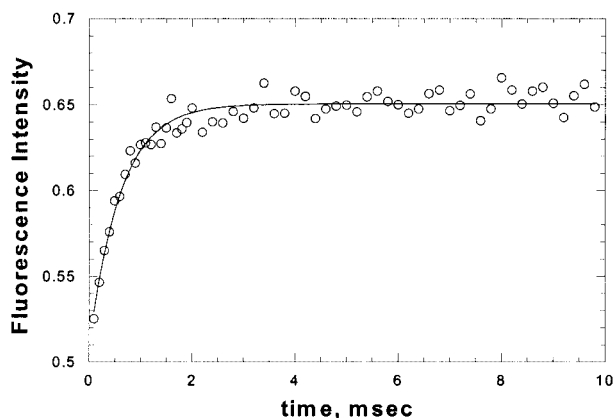


FIGURE 4: Typical kinetic trace derived from the sub-millisecond continuous flow data for wild-type IFABP. The fluorescence intensity was calibrated using *N*-acetyltryptophanamide at each time point. The experimental conditions are the same as those given in Figure 3. The data were obtained at the wavelength of maximum fluorescence change.

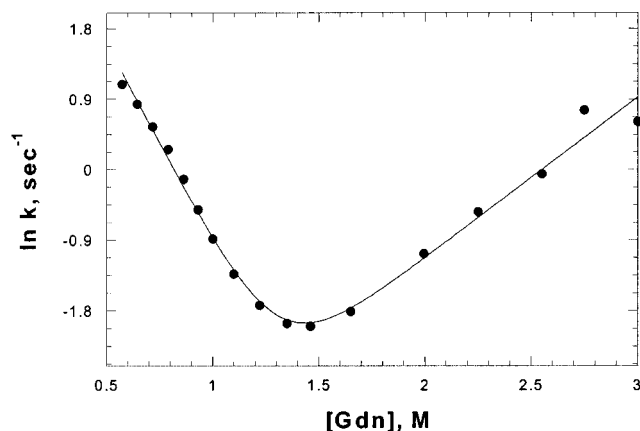


FIGURE 5: [Gdn] dependence of unfolding and refolding rate constants ( $I_2$  to N) determined by stopped flow fluorescence experiments show chevron behavior. A typical chevron plot observed for wild-type IFABP is shown. Refolding experiments were carried out by diluting the unfolded protein (in 2 M Gdn) 6-fold into varying concentrations of Gdn. For the unfolding experiments, protein samples were taken in buffer and then mixed with varying concentrations of the denaturant. The final Gdn concentrations after mixing are shown in the figure. All the experiments were performed in 20 mM phosphate buffer at pH 7.3 at 20 °C.

G65P) for both folding and unfolding. Where we have measured the unfolding rates as a function of denaturant concentration, the rate constants show typical chevron behavior. Figure 5 shows a representative chevron plot obtained for wild-type IFABP. Rates of refolding and unfolding extrapolated to zero Gdn ( $k_u^0$  and  $k_f^0$ ) for valine mutants have been determined from the fit and shown in Table 3.  $k_f^0$  values obtained by this method are similar to previous results (5).

The refolding properties of some of the proline mutants are more complicated than those of the wild-type or valine mutants. Refolding of G121P is biphasic with a very slow phase possibly due to proline isomerization. On the other hand, G65P does not exhibit any fluorescence change observable by stopped flow methods. As a result of these complications, we have shown rate constants for folding (Table 3) of the proline mutants at a single Gdn concentration of 0.25 M instead of using the extrapolated values. The

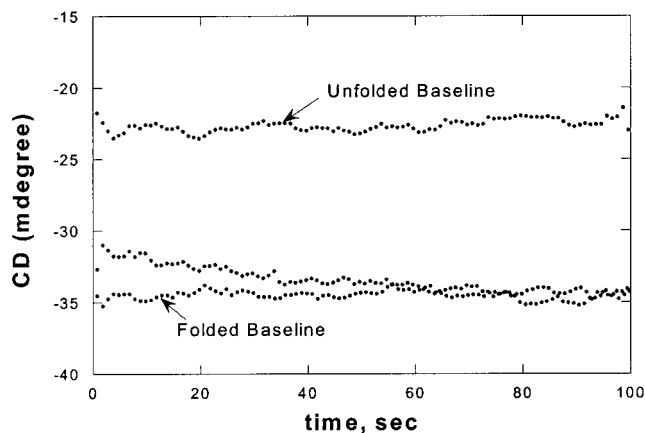


FIGURE 6: Kinetic trace observed in the stopped flow CD experiments on G121V. G121V was unfolded in 2 M Gdn and then refolded by diluting 6-fold. The kinetics were monitored at 216 nm. A folded baseline (observed at zero Gdn concentration) and an unfolded baseline (at 2 M Gdn concentration) are also shown in the figure for comparison. The experiments were performed using 20 mM potassium phosphate buffer, pH 7.3 at 20 °C.

unfolding rate constants of the proline mutants are also given for a single Gdn concentration of 1.5 M.

Although the rate of refolding obtained by sub-millisecond continuous flow kinetics remains mostly unaltered for different turn mutants, significant variation has been observed in the refolding rate as measured by stopped flow methods (Table 3). The refolding rates of G121V, G121P, and G80V are very much slower than the wild-type protein. Mutation at G99 and at G110 also results in slow refolding compared to wild-type. On the other hand, no significant change in the refolding rate has been observed for any mutation at position 44, and the G65V refolding rate constant is also similar to wild-type. The rate of unfolding of turn mutants is more or less identical to that of wild-type except for mutations at position 44 and 65 mutants. Unfolding of both G44V and G44P, as well as G65V, is faster than that of the wild-type protein.

*Stopped Flow Measurements: Circular Dichroism.* To understand the kinetics of secondary structure formation, stopped flow CD experiments were performed on wild-type and a number of turn mutants. For wild-type, all the CD change was observed within the dead time of the stopped flow mixer which indicates that the major extent of secondary structure has been formed in less than 50 ms. Since refolding of wild-type protein under the conditions used was fast (half-time around 100 ms, and the dead time of the stopped flow mixer for the CD experiments is about 50 ms), the CD experiments were performed for some slow mutant proteins for more reliable observations. Figure 6 shows stopped flow CD traces of refolding of G121V at 216 nm. Under these conditions, the half-time of refolding as measured by fluorescence changes would be about 25 s. Figure 6 also indicates the unfolded (at 2 M Gdn) and folded (in the absence of Gdn) baseline at the same wavelength for comparison. As can be seen, more than 80% of the CD change at 216 nm takes place in the dead time of this instrument. Similar results have been obtained for G80V (data not shown).

*Folding of IFABP in the Presence of Ligands: Formation of Ligand-Bound Structure.* To understand how IFABP folds

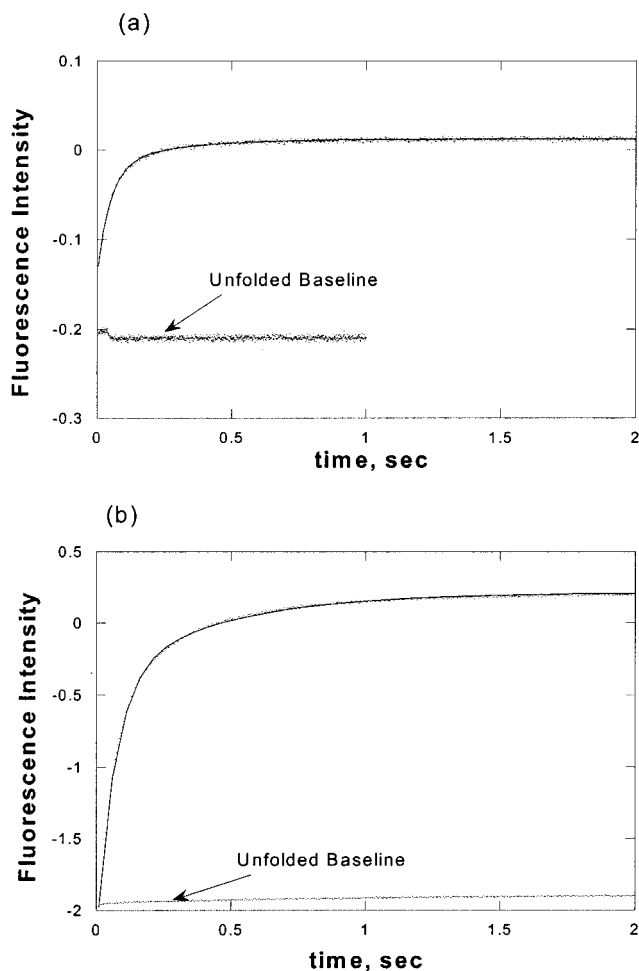


FIGURE 7: Folding kinetics in the presence of fatty acids. (a) Unfolded wild-type IFABP (in 2 M Gdn) was diluted 6-fold in 20 mM potassium phosphate buffer containing excess sodium oleate at pH 7.3 and at 20 °C. The kinetics were monitored by changes in the fluorescence intensity using an excitation wavelength of 290 nm with a 310 nm cutoff filter. (b) Unfolded wild-type IFABP (in 2 M Gdn) was diluted 6-fold in the same buffer containing excess DAUDA at pH 7.3 at 20 °C. The kinetics were monitored by the change in fluorescence intensity using an excitation wavelength of 350 nm and a 450 nm cutoff filter. Unfolded baselines (in 2 M Gdn) are shown for comparison.

in the presence of fatty acids and to determine whether fatty acids bind to the unfolded, folded, or intermediate forms, we have carried out stopped flow kinetic measurements in the presence of fatty acids. A 5-fold molar excess of fatty acid was used for all experiments to ensure complete and rapid (3) binding. Figure 7a shows stopped flow kinetics of folding of IFABP in the presence of excess sodium oleate. The protein was unfolded with 2 M Gdn and mixed with buffer containing oleate, and the resulting kinetics were monitored by changes in tryptophan fluorescence. The baseline fluorescence corresponding to the unfolded protein (at 2 M Gdn) is shown in the figure for comparison. Two kinetic phases were detected in this experiment. The first is a burst phase that corresponds to the missing amplitude of fluorescence while the second kinetic phase can be fit to a single exponential with a rate constant of  $12 \text{ s}^{-1}$ . The same experiment was repeated using a fatty acid with a dansyl group attached (DAUDA). DAUDA has earlier been used successfully to measure the dissociation constant of fatty acid binding to wild-type and different mutants, and the binding

characteristics of DAUDA and oleate have been found to be identical (5). Thus, DAUDA can be used to observe ligand binding during refolding. Figure 7b shows the change in fluorescence with time upon refolding (excitation wavelength 350 nm using a 450 nm cut-off filter). No burst phase was observed in this measurement. The observed rate constants using tryptophan or DAUDA ( $\sim 14 \text{ s}^{-1}$ ) fluorescence measurements are similar.

## DISCUSSION

**Stability.** All the turn mutants described here show decreased stability relative to wild-type. In some cases, the equilibrium unfolding curve shows little cooperativity (G65V, G65P, G80V, G121V, and G121P). Except for position Gly44, the unfolding transitions for the valine and proline mutants at the same position look remarkably similar. The rate constants for the slow refolding phase, as discussed below, can be quite different, perhaps reflecting different stability of intermediate species.

**Sub-millisecond Continuous Flow Data.** As shown previously (8), the refolding process can be conveniently divided into three phases: a very fast phase which is completed within 200  $\mu\text{s}$  and could not be observed even by the sub-millisecond continuous flow method, a kinetic phase between 200  $\mu\text{s}$  and 2 ms as observed by sub-millisecond continuous flow, and a slow phase observed by standard stopped flow measurements.

As shown under Results, the formation of at least one intermediate species is too fast to be measured even by the sub-millisecond continuous flow method (i.e., less than 200  $\mu\text{s}$ ). Since there is a spectral shift associated with this phase, it appears that some collapse of the protein must occur. The observation of a spectral shift toward the blue suggests that the tryptophan residues are at least partially buried in a hydrophobic environment within 200  $\mu\text{s}$ . Two observations are consistent with this conclusion. First, a hydrophobic cluster involving the region around Trp82 in equilibrium with the folded and unfolded forms has been proposed based on  $^{19}\text{F}$  NMR experiments using  $[6\text{-}^{19}\text{F}]$ tryptophan-labeled protein (1), a conclusion strengthened by sub-millisecond experiments with a Trp6Tyr mutant of IFABP (8). Second, equilibrium residue-specific NMR experiments show that some turns appear to retain transient structure even at very high denaturing conditions and an initial folding event has been suggested to involve those amino acid residues (7). These residues are found to center around the turn between strands D and E, the turn between strands I and J, and the turn between the helices (Figure 1). The initial formation of intermediate  $I_1$  with a rate of more than  $10\,000 \text{ s}^{-1}$  may very well represent the hydrophobic collapse around these regions. Hydrophobic residues, Val60, Val61, Leu64, Leu89 in the D–E turn; Tyr119, Val122 near the I–J turn; and Phe17, Ile23, Val122 near the helical region, as suggested by NMR, may be involved in early hydrophobic collapse of IFABP.

Of the 11 turn mutants examined, only the rate constants for the fast phase ( $I_1$  to  $I_2$  of eq 1) for G80V and G121V cannot be followed by sub-millisecond continuous flow experiments. For the wild-type and the other mutants, however, the rate constants for the fast phase were all within 2-fold of the wild type, ranging from 1400 to  $3000 \text{ s}^{-1}$  (Table 3). Since secondary structure, as measured by stopped flow

CD, appears to form in these more rapid phases, this similarity suggests that formation of secondary structure of both wild-type and mutant proteins is cooperative. For many of the mutants, except G65P, G110P, and G121P, there is a further blue shift of the fluorescence spectrum, suggesting an increased hydrophobic environment around the tryptophans. It was suggested above that the first step of the folding is the hydrophobic clustering between the helices, around the D–E turn region and around the I–J turn. It should be noted that, in the G121V mutant, a critical turn residue in the I–J turn is mutated by a more hydrophobic residue, valine. Gly121 is surrounded by three hydrophobic residues, namely, Val122, Tyr119, and Phe17. Similarly, Gly80 is very close to the D–E turn in a cluster of hydrophobic amino acids including Phe78, Phe93 and Phe62, Val60, Tyr70, and Trp82. Therefore, mutation of Gly121 or Gly80 by the more hydrophobic valine might be expected to increase the hydrophobicity around the D–E and I–J turn, regions that have been suggested to be the initiation site of the folding (7). The increase in hydrophobic character in the initiation site of the protein structure might facilitate the hydrophobic collapse and subsequent faster formation of the secondary structure observed for G80V and G121V (Table 3).

In a previous paper (5), we were unable to distinguish whether the slower rates of folding for mutant turns in the C-terminal sheet were associated with a slower burst phase or not. The answer to this is now quite clear since the sub-millisecond continuous flow data show that the observable fast rates are essentially unaffected in the Gly to Val or Gly to Pro mutants. Thus, the slower observed rates, as measured by stopped flow, are due to slower readjustments of the final structure, presumably representing the stabilization of side chains.

*Stopped Flow Data: CD and Fluorescence Measurements.* Previous attempts to measure CD changes on refolding were complicated by the rapid refolding reaction. Both the G80V and the G121V fold slowly enough so that CD measurements can be made. Refolding data from 2–0.3 M Gdn exhibit only two phases: a burst phase and a slow phase. Stopped flow CD measurements show that over 80% of the CD change occurs in the burst phase. Thus, we postulate that a large fraction of the secondary structure forms in the early phase with a rate constant on the order of  $2000\text{ s}^{-1}$ , although the actual secondary structure measurement was limited by the 50 ms dead time of the instrument. Hence, stopped flow CD experiments on wild-type and two other slow folding mutants (G80V and G121V) suggest that in intermediate  $I_2$  a majority of the secondary structure has already formed.

Rate constants for the slow refolding step for different valine and proline turn mutants have been determined by stopped flow experiments. As described previously, large decreases in refolding rate constants are observed for the mutations in turns between the last four C-terminal strands (5). However, the most significant decrease in the rate has been observed for Gly80 and Gly121 mutants. In both, the mutant proteins show noncooperative unfolding transitions (low  $m$  in Table 2) and slow formation of the native structure (Table 3). Noncooperative equilibrium unfolding transitions of these mutants suggest the possibility of stabilization of the intermediate states in equilibrium, which might lead to slower formation of the final structure. Particularly interesting

is G80P, which was expressed in inclusion bodies but could not be purified because of a constant precipitation problem. For G121P, the slowest step is probably due to proline isomerization, but it is interesting to note that only G121P, out of the six proline turn mutants, shows a slow phase possibly due to proline isomerization. Another unusual mutant is G65P, which does not show any slow phase in stopped flow measurements. Moreover, this protein does not bind fatty acid and thus may not adopt the native fold. Similar behavior has been reported with a protein resulting from triple mutation in this same D–E turn (4). As discussed earlier, the decrease in folding rate constants for these mutants is not a consequence of dramatic changes in early folding steps. Rather, these slower rates probably represent the stabilization of side chain interactions or hydrogen bond formation.

*Formation of Ligand Binding Structure.* DAUDA and oleate binding experiments show that, as expected, fatty acid does not bind to the unfolded protein (data not shown). The data of Figure 7 answer the question of whether the ligand binds to the intermediate state,  $I_1$  or  $I_2$ , or only to the native state of the protein. If DAUDA were to bind to an intermediate state, we would expect an additional phase of fluorescence change much faster than the  $14\text{ s}^{-1}$  observed. The observation of only one phase of fluorescence increase (and the lack of any fast or burst phase) suggests that the fatty acid binds only to the native form of the protein. G65P, which does not fold to native state, also does not bind fatty acid (data not shown). Thus, the formation of the native state involving stabilization of side chains is required for ligand binding.

*Comparison of Gly to Val Mutations with Gly to Pro Mutations.* No clear correlation between stability, folding, and unfolding rates of the proline mutants exists relative to the valine mutants. Equilibrium unfolding curves and sub-millisecond continuous flow kinetic rate constants for the valine and proline mutations are more or less similar. There are some exceptions, however. The unfolding transition curve of G44P is different from its valine counterpart, and we do not observe the fast folding phase of G121V. In terms of the stopped flow folding rates, there are similarities between the valine and proline changes for some turns (G44, G99, G110) and not for others (two phases for G121P, not for G121V, no slow phase for G65P while there is for G65V). Since our results suggest that the slow kinetic phase observed for the IFABP folding is the side chain packing and stabilization of the overall structure of the proteins, one might conclude from these results that the formation of the turns and their stabilization involving G44, G99, and G110 is passive; that is, that these turns themselves are not directly involved. Rather, interactions between the strands around these turns contain the interacting residues and the turn follows. In contrast, the differences between the valine and proline mutations at Gly65 and Gly121 are large, suggesting the importance of residues in those turns in the folding process, consistent with our previous conclusions (7). That the G65P mutant shows no slow phase suggests that proline sufficiently disrupts the turn such that no side chain stabilization occurs. In contrast, there are two slow phases in the refolding of the G121P mutant, the slower one probably representing proline isomerization. Mutating Gly80, a residue in the strand that separates the two  $\beta$ -sheets, yields very

unstable protein with the Gly80P mutant so unstable that it continually precipitates out of solution so that it is not possible to characterize its stability or folding kinetics.

*Secondary Structure Formation, Side Chain Stabilization, and Ligand Binding.* With at least two mutants (G121V and G80V), we can distinguish secondary structure formation, as measured by CD, from processes associated with the final structure formation as measured by fluorescence changes. In both cases, major CD changes occur well before the slower fluorescence changes. Thus, most of the secondary structure appears to form early and later steps may reflect minor changes in secondary structure associated with side chain stabilization. A similar observation has been made for the *E. coli* dihydrofolate reductase where NMR stopped flow experiments using  $^{19}\text{F}$ -labeled tryptophan showed that stabilization of the tryptophan side chains was a later step in the folding process associated with the slower fluorescence changes (14, 15). As shown in Figure 7a,b, ligand binding is coincident with the slower phase of folding and, by inference, with the stabilization of side chains rather than the formation of secondary structure. The largest difference between valine and proline mutants is at position 65, a turn shown in earlier experiments to be important in the folding process. While G65V binds fatty acid only about 4-fold weaker than wild-type (5), fatty acid binds poorly, if at all, to G65P, a mutant in which there is no slow folding phase, suggesting side chain stabilization and final stabilization of secondary structure do not occur.

## CONCLUSIONS

Site-directed mutagenesis in turns can give clues to the mechanism of protein folding (4, 5, 16, 17). The studies here confirm earlier conclusions, from equilibrium NMR experiments as a function of denaturant concentration, that at least two turns between strands are important factors in folding while other turns are not. These findings support the concept that such equilibrium NMR studies are relevant to kinetic pathways of folding. Surprisingly, mutations in turns which dramatically affect the slower folding steps (processes longer than 2 ms) do not markedly affect the rapid kinetics (that process between 200  $\mu\text{s}$  and 1 ms). All the evidence suggests that in the initial phase ( $k > 10\,000\text{ s}^{-1}$ ) a hydrophobic core is formed. It is followed by a rapid phase with a rate of

1500–3000  $\text{s}^{-1}$ , after which a large fraction of the secondary structure is formed. During the last phase, side chain stabilization is established and the rest of the secondary structure forms. It is only this last phase that is sensitive to mutations in the turn region.

## ACKNOWLEDGMENT

We thank Robert Horton for excellent technical assistance.

## REFERENCES

1. Ropson, I. J., and Frieden, C. (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89, 7222–7226.
2. Ropson, I. J., and Dalessio, P. M. (1997) *Biochemistry* 36, 8594–8601.
3. Kim, K., Cistola, D. P., and Frieden, C. (1996) *Biochemistry* 35, 7553–7558.
4. Kim, K., Ramanathan, R., and Frieden, C. (1997) *Protein Sci.* 6, 364–372.
5. Kim, K., and Frieden, C. (1998) *Protein Sci.* 7, 1821–1828.
6. Dalessio, P. M., and Ropson, I. J. (1998) *Arch. Biochem. Biophys.* 359, 199–208.
7. Hodsdon, M. E., and Frieden, C. (2001) *Biochemistry* 40, 732–742.
8. Yeh, S. R., Ropson, I. J., and Rousseau, D. L. (2001) *Biochemistry* 40, 4205–4210.
9. Sacchettini, J. C., Scapin, G., Gopaul, D., and Gordon, J. I. (1992) *J. Biol. Chem.* 267, 23534–23545.
10. Scapin, G., Gordon, J. I., and Sacchettini, J. C. (1992) *J. Biol. Chem.* 267, 4253–4269.
11. Hodsdon, M. E., Ponder, J. W., and Cistola, D. P. (1996) *J. Mol. Biol.* 264, 585–602.
12. Santoro, M. M., and Bolen, D. W. (1988) *Biochemistry* 27, 8063–8068.
13. Takahashi, S., Yeh, S. R., Das, T. K., Chan, C. K., Gottfried, D. S., and Rousseau, D. L. (1997) *Nat. Struct. Biol.* 4, 44–50.
14. Hoeltzli, S. D., and Frieden, C. (1996) *Biochemistry* 35, 16843–16851.
15. Hoeltzli, S. D., and Frieden, C. (1998) *Biochemistry* 37, 387–398.
16. Frieden, C., Huang, E. S., and Ponder, J. W. (2001) in *Protein Structure, Stability and Folding* (Murphy, K. P., Ed.) 2nd ed., pp 133–158, Humana Press, Totowa, NJ.
17. McCallister, E. L., Alm, E., and Baker, D. (2000) *Nat. Struct. Biol.* 7, 669–673.
18. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics* 14, 51–55, 29–32.

BI012042L