Protein Science (1998), 7:1821–1828. Cambridge University Press. Printed in the USA.

Turn scanning by site-directed mutagenesis: Application to the protein folding problem using the intestinal fatty acid binding protein

KEEHYUK KIM[1] AND CARL FRIEDEN[2]

[1]Department of Food Science, Woosong University, Daejeon, Korea
[2]Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110

(RECEIVED March 23, 1998; ACCEPTED May 22, 1998)

Abstract

We have systematically mutated residues located in turns between β-strands of the intestinal fatty acid binding protein (IFABP), and a glycine in a half turn, to valine and have examined the stability, refolding rate constants and ligand dissociation constants for each mutant protein. IFABP is an almost all β-sheet protein exhibiting a topology comprised of two five-stranded sheets surrounding a large cavity into which the fatty acid ligand binds. A glycine residue is located in seven of the eight turns between the antiparallel β-strands and another in a half turn of a strand connecting the front and back sheets. Mutations in any of the three turns connecting the last four C-terminal strands slow the folding and decrease stability with the mutation between the last two strands slowing folding dramatically. These data suggest that interactions between the last four C-terminal strands are highly cooperative, perhaps triggered by an initial hydrophobic collapse. We suggest that this trigger is collapse of the highly hydrophobic cluster of amino acids in the D and E strands, a region previously shown to also affect the last stage of the folding process (Kim et al., 1997). Changing the glycine in the strand between the front and back sheets also results in a unstable, slow folding protein perhaps disrupting the D-E strand interactions. For most of the other turn mutations there was no apparent correlation between stability and refolding rate constants. In some turns, the interaction between strands, rather than the turn type, appears to be critical for folding while in others, turn formation itself appears to be a rate limiting step. Although there is no simple correlation between turn formation and folding kinetics, we propose that turn scanning by mutagenesis will be a useful tool for issues related to protein folding.

Keywords: β-sheet protein folding and stability; fatty acid binding; turn mutations

The intestinal fatty acid binding protein (IFABP) is a small (15,000 Da) monomeric protein that belongs to a class of proteins that are primarily β-sheet and bind a diverse group of ligands (fatty acids, retinoids, and bile salts) into a large central cavity in the interior of the protein. The structure of the protein has been examined both by X-ray (Scapin et al., 1992; Sacchettini & Gordon, 1993; Banaszak et al., 1994) and NMR (Hodsdon et al., 1996; Hodsdon & Cistola, 1997a, 1997b; Zhang et al., 1997) methods. It is an excellent model protein for folding studies, not only because of its small size and β-sheet structure, but also because it contains no cysteine or proline residues. In a previous paper (Kim et al., 1997), we investigated the role of the turn between two antiparallel β-stands (the D-E turn) by mutating three of the residues including the Gly65 and its two neighbors, Leu64 and Val66. This particular turn was of interest because it is highly conserved in other members of the family and because there are no hydrogen bonds between the main-chain atoms of the two adjacent D and E strands, the space being occupied by side chains of hydrophobic residues (Scapin et al., 1992). From the data obtained, we showed that the crucial residue for final stabilization of the protein upon refolding was Leu64. Modification of this residue to smaller less hydrophobic ones or to hydrophilic ones resulted in proteins that were very unstable and collapsed rapidly to a form that could bind fatty acid (albeit weakly) but did not undergo the final phase of stabilization. It was concluded that the interaction of Leu64 with hydrophobic residues in other strands was the crucial interaction for the final stabilization and possibly in an initial hydrophobic collapse as well (Kim et al., 1997).

These results suggested that systematic modification of other turns might be useful in determining not only the importance of turns in folding, but also might give information about the mechanism of folding of β-sheet proteins in general. There have been

Reprint requests to: C. Frieden, Department of Biochemistry and Molecular Biophysics, Box 8231, Washington University School of Medicine, 660 South Euclid Avenue, St. Louis, Missouri 63110; e-mail: frieden@biochem.wustl.edu.

Abbreviations: IFABP, rat intestinal fatty acid binding protein; Gdn, guanidine hydrochloride; EDTA, ethylenediamine tetraacetic acid; DAUDA, 11-((5-dimethylaminonaphthalene-1-sulfonyl)amino)undecanoic acid.

different views about the role of turns in protein folding as either initiation points or just passive structures linking secondary structures (for a review, see Rose et al., 1985). Aside from its unique structure, IFABP is an excellent model system for such an examination since there are eight turns in the protein that are between antiparallel β-strands and a glycine residue in seven of these turns. In most of these turns, the glycine residue occupies an unusual region of ϕ,ψ space and, therefore, substitution with another amino acid should result in some disruption of the turn. In addition, a glycine residue appears in a half turn that connects the two faces of the β-sheets that surround the interior cavity. In this paper, we have systematically examined the effect of mutations in these turns mutating glycine (or in one case, asparagine) to valine.

The results presented here represent one of the few systematic investigations of the role of turns in the folding process in proteins (Hynes et al., 1989; Serrano et al., 1992; Garrett et al., 1996; Ybe & Hecht, 1996; Zhou et al., 1996; Gu et al., 1997) or peptides (Stroup & Gierasch, 1990; Dealba et al., 1996; Ramirez-Alvarado et al., 1996). The present results, however, indicate the complexity of turn formation and function and may serve as an introduction to a more complete study of both theoretical and experimental aspects of the role of turns in the folding process.

Results

Figure 1 shows the X-ray structure of the intestinal fatty acid binding protein (Scapin et al., 1992) indicating what residues have been mutated to valine. Except for the turn between the small helical region, every turn and the one-half turn (G80) has been mutated. Table 1 gives a description of these turns both by ϕ and ψ angles and turn type from both X-ray and NMR data. The X-ray data represent a closed form of the apo protein while the NMR data represent the mean coordinates for an ensemble of structures that adopt a more open and flexible form in solution. It is clear from this table that there can be striking differences in the ϕ and ψ angles for glycines in turns as determined by these two methods. According to the X-ray data, all the glycine residues to be mutated have positive values of ϕ, placing the conformation in a disfavored region of a Ramachandran plot while the NMR data clearly places Gly65 in a favorable region. Similarly, the ϕ and ψ angles for Gly75 and Gly86 are markedly different between the X-ray and NMR data. The characteristic properties of each mutant have been examined with respect to (1) reversible denaturation by Gdn, (2) rate of refolding as a function of denaturant concentration, and (3) ligand binding.

Table 1. *Turns between β-strands in the intestinal fatty acid binding protein*

Residue numbers	Sequence	Turn type based on ϕ,ψ	ϕ,ψ for glycine (X-ray)[a]	ϕ,ψ for glycine (NMR)[a]
43–46[b]	EGNK	II^I	91.1, −117.5	0.9, 5.7
53–56	SNFR	I		
63–66	ELGV	II	81.1, −2.9	−121.9, 103.2
72–75	LADG	IV	93.4, −0.8	−119.2, −90.2
85–88[b]	EGNK	II^I	64.2, −94.5	−135.8, −46.7
96–99[b]	VDNG	I	76.9, 14.6	97.0, 69.9
109–112	SGNE	II^I	71.8, −92	60.1, 99.0
119–122[b,c]	YEGV	I^I	71.4, 20.5	104.1, 21.7
78–82	LTGTW	Half turn	167, −164.6	−101.3, — 156.3

[a]The X-ray values correspond to a closed form of the protein as exhibited by the X-ray crystal structure of apo-IFABP (Scapin et al., 1992). The NMR values represent the solution structure of apo-IFABP that adopts an open form consisting of a manifold of locally disordered states (Hodsdon & Cistola, 1997a, 1997b). The NMR values represent mean coordinates of an ensemble of structures.

[b]The turn type may differ depending on which residues are chosen. Starting at residue 42, 84, or 118 gives a four residue type IV turn.

[c]This turn was listed as residues 120–123 in Scapin et al. (1992).

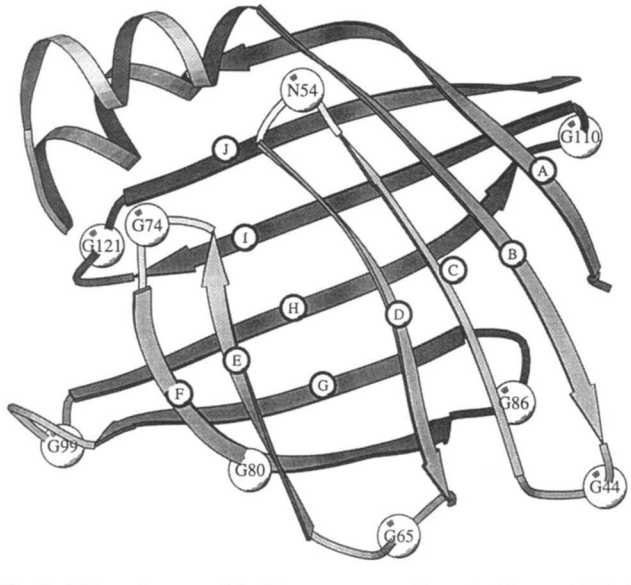

Fig. 1. Ribbon diagram of the X-ray structure of intestinal fatty acid binding protein (Scapin et al., 1992) showing the positions and numbering of the residues that have been mutated to valine. The strands are labeled A–J for identification when discussed in the text. The figure was prepared using Molscript (Kraulis, 1991).

Structural properties

Table 2 presents data on the stability characteristics and refolding rate constants for the mutants examined in this paper. The table also indicates that a number of the mutants are present in inclusion bodies suggesting misfolding during expression. After dissolving these insoluble proteins in Gdn, however, they can be refolded simply by removing the denaturant. In every case, whether the protein is present in inclusion bodies or not, the far UV CD spectrum (from 190 to 250 nm) in the absence of Gdn is superimposible with that of the wild-type (data not shown), indicating that the overall secondary structures of the wild-type and all the mutant proteins are the same. As shown later, all the mutant proteins can bind a fatty acid ligand again indicating that their structures are similar to that of the wild-type. There can be some differences, however, in the near UV CD spectra, an indication of differences in local and/or tertiary interactions. Figure 2 shows near UV CD spectra for wild-type and four mutants (G44V, G54V, G80V, and G121V) with the most dramatic difference in the G80V mutant. Other mutants not shown in this figure have spectra like the G121V mutant and are essentially identical to wild-type. As shown below,

Table 2. *Denaturation characteristics of turn mutants*[a]

Mutant	In inclusion bodies	ΔG^0	C_{mid} (M)[b]	m_G[c]	k_r^o (s^{-1})	Slope ($d \log k/d$ Gdn)
Wild-type	No	5.2	1.36	−3.82	73	−5.2
G44V EGNK	No	5.1	1.0	−5.1	28[d] (<10)	−5.5
N54V SNFR	No	7.5	1.61	−4.6	350[d] (<10)	−6.0
G65V ELGV	Yes	0.73	0.48	−1.5	12	−2.9
G75V LADG	No	4.9	0.91	−5.4	420[d] (<30)	−7.0
G86V EGNK	No	5.3	1.29	−4.1	20	−4.4
D97V RVDNG	Yes	3.4	0.89	−3.8	1.3	−3.3
G99V RVDNGKE	Yes	2.7	0.85	−3.2	1.6	−4.9
G110V SGNE	Partial	4.4	1.15	−3.8	2.6	−3.8
G121V YEGV	Yes	2.1	0.89	−2.4	0.05	—[e]
G80V LTGTW	Yes	0.56	0.48	−1.2	0.37	—[e]

[a]All experiments performed at pH 7.3 in 20 mM potassium phosphate buffer, 0.25 mM EDTA at 20°.
[b]Midpoint of the Gdn denaturation curve.
[c]Slope of the denaturation curve in Gdn assuming a two state model.
[d]This is an extrapolated value assuming the values of the $\log(k_r)$ are linear as a function of Gdn concentration but the data are not linear. The value in parentheses is that approximation of the value including the non-linearity.
[e]See Figure 4. The rate constants are almost independent of the final denaturant concentration.

the G80V mutation, in a half turn of the strand that connects the two faces of the protein, results in a protein that is the most unstable of all the mutants discussed here and the near UV CD data suggest that there have been changes in sidechain interactions. As discussed later, the mutation in this position may disrupt a different region of the protein as well as the half turn. Of interest is the fact that both the far and near UV CD spectra for the G121V mutant are essentially identical to wild-type since, as shown in Table 2, this mutant folds very slowly. Both G44V and G54V show some differences in the near UV CD spectra from wild-type, primarily in the region from 250–270 nm usually assigned to phenylalanine (Strickland, 1974).

Equilibrium denaturation

The characteristics of Gdn denaturation are given in Table 2, which presents data for ΔG^0, m_G, the slope of the equilibrium denaturation curve and C_{mid}, the midpoint of the denaturation curve. As implied above, denaturation by Gdn is a reversible process and unfolding and refolding data are superimposible. From an analysis of the equilibrium denaturation data a value of the slope, m_G, was

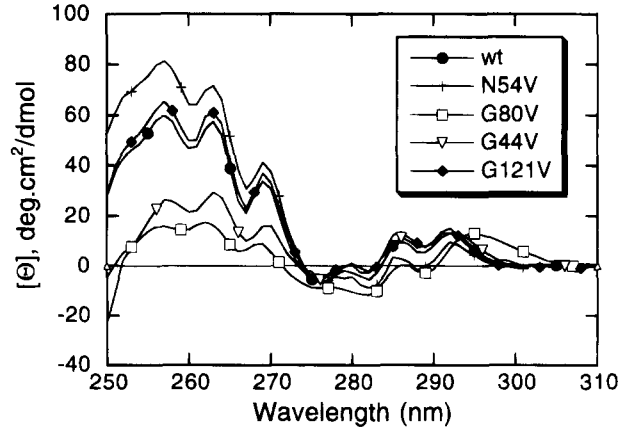

Fig. 2. Near UV CD data for wild-type and several glycine to valine mutants. All experiments performed in 20 mM potassium phosphate buffer 0.25 mM EDTA pH 7.3 at 20–23° using a 1 cm pathlength cell except for G80V that used a 10 cm pathlength cell. The protein concentration used varied between 0.49–1.26 mg/mL except for G80V for which the concentration used was 0.053 mg/mL. The spectra of the wild-type protein and G121V mutant are essentially superimposible.

obtained assuming a two state process (Santoro & Bolen, 1988). As shown by the data in Table 2, there is an interesting correlation between stability and expression of the protein in inclusion bodies with those mutants that are less stable occurring in inclusion bodies. The data show that mutation at position 65 (G65V) yields a very unstable protein in agreement with our previous data involving this turn (Kim et al., 1997). Other relatively unstable proteins are those in turns at positions 99 (G99V) and 121 (G121V) and the half turn at position 80 (G80V). Figure 3 shows the denaturation curves as a function of Gdn concentration for these mutants. It is

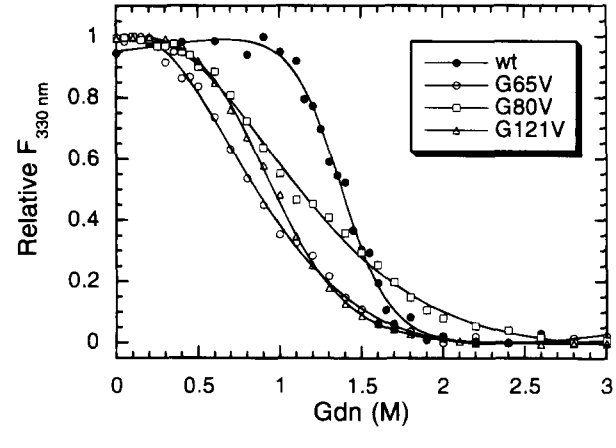

Fig. 3. Denaturation of wild-type and several glycine to valine mutants as measured by changes in intrinsic fluorescence as a function of Gdn concentration. Experiments performed in 20 mM potassium phosphate buffer, 0.25 mM EDTA, pH 7.3 at 20°. The protein concentration used was 2–4 mM (0.03–0.06 mg/mL). The data were normalized to a maximum value of 1.0. The solid lines through the data are fits to data as described in Materials and methods assuming a two state model. Other mutant proteins not shown in this figure showed cooperative behavior similar to that shown for the wild-type. Except for the G44V mutant the data are for the same mutant proteins as shown in Figure 2.

clear from the figure and the m_G values given in Table 2 that there is a loss of the high degree of cooperativity in the unfolding process for these mutants compared to the wild-type protein. In these cases, it is likely that the calculation of ΔG^0 is incorrect, and is only an indication of the low stability. On the other hand, mutations in turns at positions 54 (G54V), 86 (G86V), and 110 (G110V) have very little effect either on stability or cooperativity (data not shown in Fig. 3).

Refolding rate constants

Figure 4 shows the values of the refolding rate constants (k_r), plotted on a log scale as a function of the final denaturant concentration. For some of the mutants the plot is linear as a function of the final Gdn concentration as noted for a number of other proteins. For several other mutants, particularly N54V, the rate constant appears to level off at lower Gdn concentrations. While such behavior is typified by proteins for which proline isomerization is a rate limiting step in refolding, it should be reemphasized that IFABP contains no proline. Of particular interest are the G121V mutant, for which the rate of refolding is much slower than for any other mutant, and the G80V mutant. In both cases, the refolding rate constants are almost independent of the final denaturant concentration. Other slow folding mutants are G99V and G110V. Also of note in these figures is the slope of the linear portion to the refolding rates which varies considerably from mutant to mutant.

In all cases, there was an initial phase that was too fast to be measured by stopped-flow techniques.

Ligand binding by turn mutants

As indicated above, all the turn mutants fold to a native-like structure based on the similarity of the far UV CD spectra. As a measure of whether the folded proteins were functional, ligand binding studies were performed. In these experiments the fatty acid analog DAUDA was used. Figure 5 shows typical titration data using wild-type protein. As with wild-type and all the mutant proteins, the data reflect a simple binding process. Assuming one binding site, dissociation constants are easily determined because of the large change in fluorescent properties for this compound. From the data in Figure 5, a value of 21 nM can be calculated for the wild-type protein, very similar to that for oleate determined by others (Cistola et al., 1996; Kurian et al., 1996; Richieri et al., 1996). Thus, DAUDA binding should be a good indicator of functionality of the mutants. Table 3 lists these dissociation constants. Interestingly, three mutants, G44V, G86V and G110V, that show stabilities (ΔG^0 values) close to that of the wild-type protein (5.2 kcal/mol) bind ligand poorly, while two unstable mutants, G80V and G65V, bind ligand reasonably tightly.

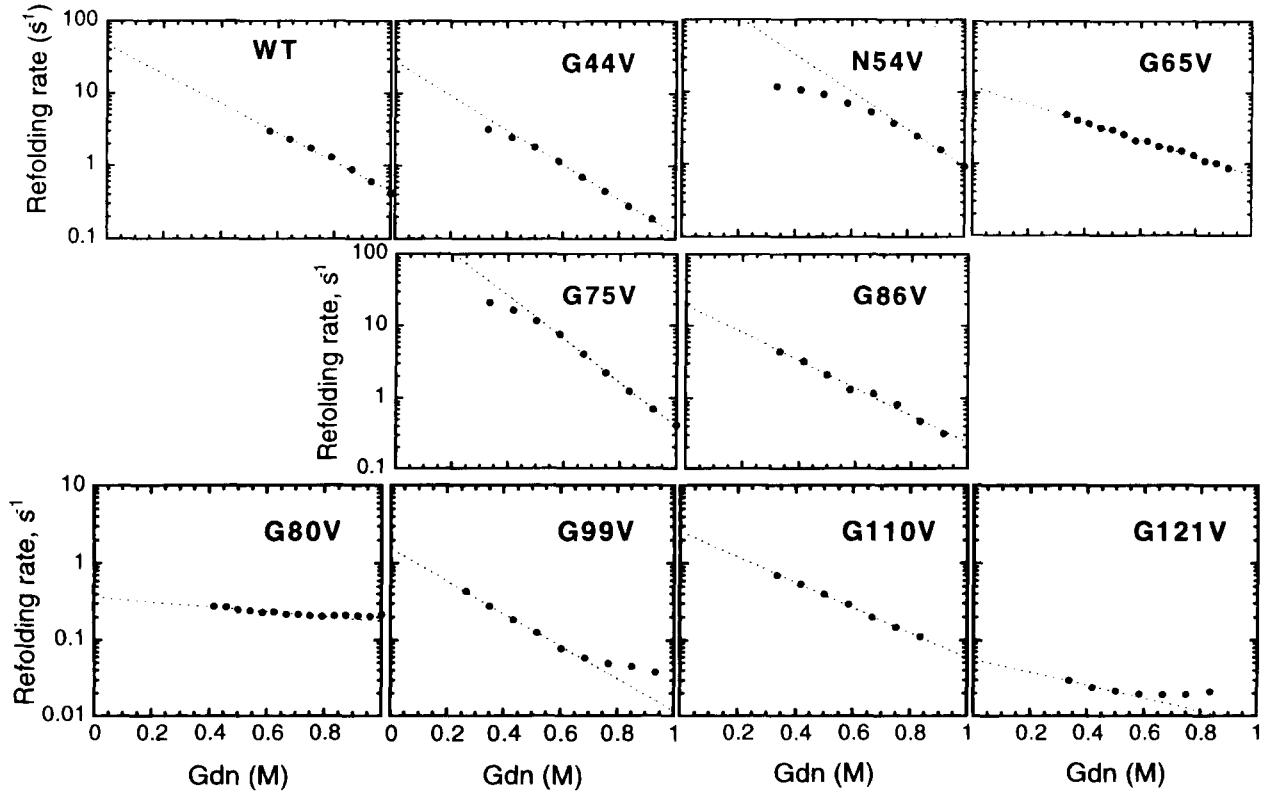

Fig. 4. Plots of the refolding rate constants (k_r) plotted on a log scale as a function of the final denaturant concentration. The initial Gdn concentration prior to dilution was 2 M for all mutants except G99V and G75V (1.6 M) and G80V (2.5 M). The final protein concentrations ranged from 1.3–5.2 μM (0.02–0.078 mg/mL). Linear regions of the data shown in these plots were fit with a simple exponential function to obtain the slope shown in the last column of Table 2. Experiments performed in 20 mM potassium phosphate buffer, 0.25 mM EDTA, pH 7.3 at 20°. Note that the y-axis scale is different in the bottom row from that in the top two rows.

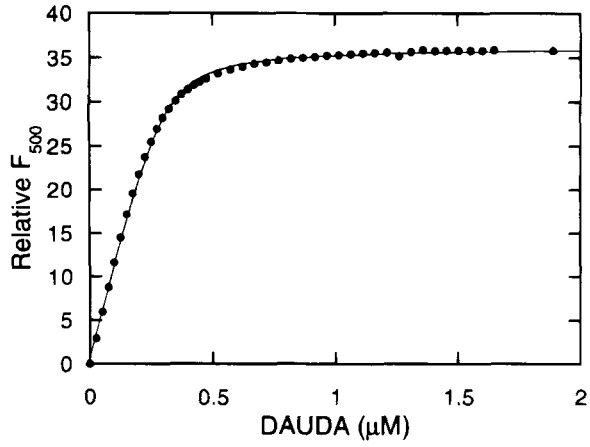

Fig. 5. Typical fluorescence titration data for wild-type protein using the fluorescence fatty acid analog, DAUDA. The concentration of the protein used was 0.3 μM and the solid line is the fit to a simple binding equation assuming a single binding site. Experiments performed in 20 mM potassium phosphate buffer, pH 7.3, 0.25 mM EDTA, at 20°.

Discussion

In general, investigators have considered two differing views on the role of turns in the folding process: that an individual turn may serve as an initiator of folding as, for example, a nucleation site or that the turn may be passive serving only to link β-strands that themselves contain all the necessary stabilizing forces for structure formation. Within this generalization, there are three possible effects of mutations in turns: (1) little or no effect on either stability or kinetic properties; (2) an effect on stability with or without an effect on the kinetics of refolding; or (3) an effect on rates of refolding with or without an effect on stability. In this paper, we have mutated glycine residues in turns to valine and have observed all the permutations. The data show that there do not appear to be any simple generalizations between turn type and either stability or refolding kinetics. As noted in Table 1, the X-ray data suggest that all glycine residues mutated have positive values of the backbone dihedral angle ϕ, a conformation that is disfavored in regular secondary structures (α-helices or β-strands) while the NMR data give somewhat different results. The NMR data are probably more applicable since they represent the structures most highly populated in solution.

Table 3. *Dissociation constants for wild-type and mutant IFABP proteins*[a]

Mutant	K_d (nM)
Wild-type	20.9 ± 0.6
G44V EGNK	343 ± 17
N54V SNFR	23 ± 1
G65V ELGV	78 ± 12
G75V LADG	79 ± 3
G86V EGNK	276 ± 13
D97V RVDNG	61 ± 6
G99V RVDNGKE	106 ± 2
G110V SGNE	490 ± 15
G121V YEGV	120 ± 25
G80V LTGTW	200 ± 15

[a]All experiments performed at pH 7.3 in 20 mM potassium phosphate buffer, 0.25 mM EDTA at 20°.

For glycines with positive ϕ values, it is a good assumption that the substitution with valine will distort the structure of the turn. For example, using Sleuth (written by Dr. Jay Ponder and available at http://dasher.wustl.edu/pub/sleuth), we find only 10 valine residues with nonstandard angles in the $i + 1$ position of a turn and 67 in the $i + 2$ position listed in the protein database (out of ~266,000 residues and 1,339 proteins). These numbers are almost certainly high since the structural resolution for some proteins may not sufficiently great for accurate determination of the ϕ and ψ angles. Additionally, several of the glycines in turns of IFABP are highly conserved in members of the family of these binding proteins (Banaszak et al., 1994). These include not only G44 and G65, but also the GN or NG motif that occurs near position Gly86, and the GD or DG motif near position Gly110 suggesting the importance of these glycine residues in those turns. In spite of that, and based on the CD and ligand binding data, the overall structure appears to be maintained in all the mutants described here. It seems clear that the valine mutation does not prevent turn formation nor does it produce any large changes in overall structure. The mutation, however, favors a different turn than in the wild-type structure, and the critical issue becomes how this structural change in the turn manifests itself in the folding process.

Cooperativity in the folding of the C-terminal face: The G99V, G110V, and G121V mutant proteins fold slowly

It is of interest to consider the effect of mutations in turns between the last four strands in this face of the protein (turns between strands G-H, H-I, and I-J). The glycine residues in the three turns between these strands have positive ϕ values in both the NMR and X-ray structures. In each case, the mutation gives rise to a slow folding protein. Clearly these turn mutations interfere with strand-strand interactions such that the final stabilization of this face of the protein becomes a rate limiting step. Beyond that, however, the results suggest that folding of these strands is a highly cooperative process since mutation in any one of the turn gives similar results. The trigger for the cooperative folding may relate to which interactions between adjacent strands are the strongest or to an initial hydrophobic collapse. In an earlier study (Ropson & Frieden, 1992), we postulated that the region around Trp82 (in the F-strand and in a highly hydrophobic environment) was involved in a populated intermediate in equilibrium with native and unfolded protein suggesting that this is the region that might be involved in an initial collapse in the folding process. If this were true, then this region may serve as the initiator for the β-sheet

formation. It is of interest that our attempts to express a peptide containing only residues 84–131 failed because the product was proteolytically digested (K. Kim & C. Frieden, unpubl. results), also suggesting that residues in other portions of the protein (i.e., those involved in forming the hydrophobic cluster) are important in the cooperative folding of the last four strands.

In regard to a triggering effect, the effect of the G80V mutation may also be relevant. Gly80 is in a half turn in the strand that links one face of the protein to the other. Modification of this residue has a dramatic effect. The resulting protein is unstable and binds ligand rather poorly. Because residues around Gly80 are in close proximity to residues in the D-E turn, it may be that mutations in the half-turn may disrupt the same portion of the protein as mutations in the D-E turn. As discussed above, we have postulated a hydrophobic cluster in the folding process centered around the D-E turn (Ropson & Frieden, 1992) and the valine at position 80 may affect this cluster. Curiously, the refolding rate for this mutant (and for G121V) is essentially independent of the final denaturant concentration as well as being very much slower than that of the wild-type. This point will be further discussed below for the G121V mutant. The importance of hydrophobic interactions in early folding steps may also relate to a study of the folding of cellular retinoic acid binding protein I, a protein with a similar overall structure as IFABP, in which Clark et al. (1997) concluded that hydrogen bonding was a late step in folding. Smith et al. (1994), Smith and Regan (1995), and Minor and Kim (1994a, 1994b) have discussed β-sheet propensities relative to specific amino acids.

The mutation at Gly121, which is in the turn between the last two strands of the C-terminal face, results in the most dramatic changes in the properties of the protein. Not only is the protein quite unstable and the denaturation curve the least cooperative of all other mutant proteins described here, but the protein folds very slowly and the rate constant for refolding is essentially independent of the final denaturant concentration. In another context we have presented evidence that diffusion of denaturant into a protein (in that case the *Escherichia coli* dihydrofolate reductase) during unfolding proceeds through at least two steps, the first being instability on the outer surface of the protein followed by penetration of the denaturant into the interior, the latter being slow (Hoeltzli & Frieden, 1995). It is likely that diffusion of the denaturant from the interior of a partially folded protein is rate limiting (and therefore explains the dependence of the refolding rate constant on the final denaturant concentration). If so, then a step slower than denaturant diffusion would be rate limiting. In this case, that would be the correct turn formation. There are alternate explanations for the very slow refolding process. One arises from molecular dynamics calculations (M. Hodsdon & J. Ponder, pers. comm.). Starting with a 6 amino acid peptide simulation of turn formation for the G121V mutant shows that a turn forms very rapidly compared to the rate of formation of the wild-type turn. It is, however, an incorrect turn type being a type I instead of a type I^I turn. Thus, it may be that an incorrect turn forms rapidly between the last two strands and the slow refolding rate reflects the search for the correct turn to allow interaction between the strands. On the other hand, simulations of the mutant turn structure using LINUS (Srinivasan & Rose, 1995) suggest a β-strand-like propensity throughout the I-J region suggesting that the slow refolding is due to the difficulty of forming any turn.

Mutations in the turn between the D and E strands

Examination of the structure of IFABP show that while there is no hydrophobic core, there is a hydrophobic cluster of amino acids between strands D and E as discussed above. We have discussed elsewhere (Kim et al., 1997) the characteristics of mutations in the turn between these two strands and have shown that mutations at residue Leu64 produced a very unstable protein that appeared to behave like an intermediate in the folding pathway since no slow final step resulting in stabilization occurred. It was concluded that the leucine residue could be involved in both early and late steps in folding, the late step being an interaction with hydrophobic residues in strands other than D and E to promote the final stabilization. Between the D and E strands there are no hydrogen bonds, but only hydrophobic interactions. Although the G65V mutation is not as drastic as that of the mutation of the neighboring leucine, it still produces a very unstable protein. The large change in stability may reflect a change in the positioning of the neighboring leucine sidechain (Leu64) such that interactions with hydrophobic residues in other strands cannot occur. Mutations in this turn would appear to argue against the conclusion of Plaxco et al. (1998) that there is a correlation between folding rate and linear separation in the sequence of contacting residues since folding rates for turn mutations are very different yet the distance between hydrophobic residues remains the same. It is curious that the ϕ and ψ angles for Gly65 differ so much when comparing the X-ray and NMR structures, with the latter being in an allowed region of ϕ and ψ space while the former is not.

Mutations in type II^I turns

There are three turns (the B-C turn (G44V), the F-G turn (G86V), and the H-I turn (G110V)) that are type II^I containing glycine as the second residue (Chou & Fasman, 1977; Sibanda & Thornton, 1991). The glycine-to-valine mutation should almost certainly change the turn type, perhaps into the more common type I turn, yet there is little effect on stability suggesting that these turns do not greatly influence interactions between strands. In all cases, the mutation mostly appears to weaken ligand binding (up to 23-fold) more than any of the other mutations. Large structural changes do not appear to have occurred since the stability of these three mutants is rather similar and all show highly cooperative unfolding behavior, only slightly less than that of the wild-type protein. Interestingly these three turns lie on one side of the protein nearest to the carboxyl group of the fatty acid ligand although none is in direct contact with the ligand. The effect of mutations in this region must be transmitted through the structure in some way as to affect ligand binding. $\text{Log}(k_r)$ is a linear function of the final denaturant concentration, and two of the three (G44V and G86V) fold at rates similar to wild-type. In this regard, the slow rate of folding of the G110V mutant suggests that the interaction of the two adjacent (H and I) strands is controlled by the residues in the turn or related to cooperativity of the β-sheet formation as discussed earlier. Near UV CD data shown in Figure 2 indicate that there are some differences in the tertiary structure of G44V relative to the wild-type protein perhaps related to a nearby phenylalanine (Strickland, 1974) (at position 47) and its interaction with phenylalanine residues on 2 neighboring strands (at positions 62 and 68).

Other turn mutations

The turn between the C and D strands does not contain a glycine residue, so it was not possible to determine the effect of a large structural perturbation. This modest change (N54V) did produce a slightly more stable protein with somewhat weaker ligand binding. This mutation results in only slight differences in the near UV CD

spectrum from wild-type (data not shown). It is of interest to note that this region of the protein is critical for ligand entry (Sacchettini et al., 1992; Cistola et al., 1996) and is flexible in solution (Hodsdon & Cistola, 1997a, 1997b). The small changes in stability and ligand binding and its flexibility suggests that this turn is not an important one in the folding process.

The glycine in the turn between the E and F strands is at the $i + 3$ position. Mutation at this turn (G75V) does not cause as much of a structural perturbation as observed in other turns. The mutation does have some affect on stability but refolding rates are not greatly affected and ligand binding changes only by about fourfold. This glycine is not conserved in the family, and the major effect does appear to be on ligand binding. It should be noted that, in both this mutant and the N54V mutant, $\log(k_r)$ as a function of Gdn concentration appears to be nonlinear leveling off at a rate constant close to that observed for the wild-type protein. While it is not yet clear how to interpret this result, it is of interest that the NMR data show these two regions of the protein, around N54 and G75, to be flexible (Hodsdon & Cistola, 1997a, 1997b) suggesting as above that these turns are not important in the folding process. Perhaps, however, Gdn shifts the order-disorder equilibrium that occurs in these regions.

As with Gly75, the glycine in the turn between the G and H strands is at the $I + 3$ position. Mutation of this glycine (G99V) may not perturb the turn greatly. A companion mutation, D97V, substituted a hydrophobic residue for a hydrophilic one. The effect of either mutation was similar in that the proteins fold slowly, are somewhat less stable than wild-type, and bind fatty acid 3–5 fold less tightly than wild-type.

It is perhaps disappointing that there are no obvious and clear correlations between mutations in turns and folding characteristics. Gu et al. (1997) reach a similar conclusion. Turn scanning by site-directed mutagenesis, however, has not been a general technique used to examine issues of protein folding. In an earlier study (Kim et al., 1997), we examined the role of residues in the D-E turn of IFABP by random mutagenesis of residues surrounding a central turn residue. That study was valuable in showing that an adjacent leucine residue appeared to be critical for final stabilization of the protein. In the present paper we have chosen only a selected number of turn residues to mutate but clearly the study could be extended by scanning techniques to examine the role of all the turn residues. Turn scanning by randomly mutating specific residues should be a valuable technique for examining the importance of specific interactions between adjacent strands relating to the folding mechanism.

Materials and methods

Materials

Mutagenic primers were obtained from Integrated DNA Technologies (Coralville, Idaho). The QuikChange site-directed mutagenesis kit was obtained from Stragene (La Jolla, California) Ultrapure Gdn was obtained from ICN Biochemicals (Aurora, Ohio). DAUDA was obtained from Molecular Probes (Eugene, Oregon) and dissolved in 95% ethanol to make a 1 mM stock solution and kept light-free at 4°. All other reagents used were analytical grade.

Mutagenesis, protein expression, and purification

Site-directed mutagenesis was used to substitute a valine codon (GTN) from a glycine codon (GGN) using primer lengths of 27 nucleotides for both strands. The mutant plasmids were identified by sequencing, and the mutation yields were over 90%.

The *E. coli* strain, MG1655, transformed with a mutated pMON-IFABP construct, was incubated overnight in 4 mL LB media containing 100 μg/mL ampicillin, and then transferred to 100 mL LB media in 250 mL flasks. Protein expression was induced by the addition of 1 mM nalidixic acid to the growing culture at OD_{600} = 0.6–0.8 and cells were allowed to grow 4–5 h more, harvested, and frozen at −70 °C. The cell pellet, thawed in 5 mL of 20 mM Tris-Cl, pH 8.0, 5 mM EDTA, was sonicated on ice for 2.5 min using a Branson Model 250 sonifier set to 35% and 30 W. The homogenized cell suspension was centrifuged using a Beckman microcentrifuge at 4 °C, 13,000 rpm for 5 min. Since some mutant proteins appeared to be produced as inclusion bodies, both the supernatant and the pellet were analyzed by SDS-PAGE to determine which fraction would contain most of the mutant protein. If the majority of the protein was present in the supernatant, method 1 (below) was employed for further purification. If the protein appeared to form inclusion bodies, method 2 was employed.

Method 1: (Purification of mutant from supernatant solution). The clarified supernatant was stirred at 4 °C while a buffered ammonium sulfate solution (90% saturated in 20 mM Tris-Cl, pH 8.0, 0.25 mM EDTA) was added slowly until 60% saturation was obtained, followed by slow stirring for 1 h. The precipitate was removed by centrifugation (30 min, 7,500 × *g*) and the supernatant was transferred to dialysis tubing (MW cut off = 3,500, Spectra/Por). Dialysis was carried out against multiple changes of 20 mM Tris-Cl, pH 8.0, 0.25 mM EDTA (Buffer A) at 4 °C until the ionic strength of the dialysis buffer was close to that of the original buffer A. The dialyzed protein sample was loaded on to a QAE disk (Zeta prep, CUNO), preequilibriated with Buffer A, and the mutant protein was eluted by applying 20 mM KCl in Buffer A. The eluent was collected and concentrated using an Amicon ultrafiltration kit with a YM-10 membrane. A Sephadex G-50 column (5.2 × 120 cm) and a Lipidex-1000 column (at 37 °C, preequilibriated with 20 mM potassium phosphate, pH 7.3, 0.25 mM EDTA) were used for the further purification and delipidation, respectively. Protein fraction were pooled and concentrated using a YM-10 membrane. Mutant proteins were over 98% homogeneous as determined by the Coomassie Blue–stained SDS-PAGE gel.

Method 2: (Recovery of mutant IFABP from inclusion bodies). The pellet, containing inclusion bodies, was washed five times with Buffer A and centrifuged for 30 min at 4 °C at 10,000 × *g* and dissolved in 30 mL of 6 M Gdn in Buffer A and the insoluble material was removed by centrifugation. The unfolded protein was renatured by diluting 100-fold into Buffer A and then concentrated to about 80 mL using a YM-10 membrane. After thorough dialysis against Buffer A, aggregates formed were removed by centrifugation (10,000 × *g*, 30 min) and the supernatant applied to a Zeta-Prep capsule (CUNO). The rest of the procedure followed was the same as above except delipidation. Since these proteins formed aggregates inside a Lipidex-1000 column at 37 °C during delipidation, the column was run at room temperature.

Protein concentrations for the wild-type and mutant proteins were determined by absorbance at 280 nm using a value of 1.13 A/mg/mL.

Circular dichroism. Spectra were collected using a Jasco J600 or J715 spectrophotometer. The pathlength of the cell was 0.1 cm for far UV-CD or 1.0 cm for near UV-CD measurements. For the G80V mutant, a 10 cm cell was used. All measurements were made

in 20 mM potassium phosphate, pH 7.3, 0.25 mM EDTA at ambient temperature, and the protein concentrations were adjusted to 0.15–0.27 mg/mL for far UV-CD and 0.49–1.26 mg/mL for near UV-CD measurements respectively.

Determination of stability in Gdn. Equilibrium unfolding and refolding as a function of denaturant concentrations in 20 mM potassium phosphate, pH 7.3, 0.25 mM EDTA, were monitored by fluorescence spectroscopy with excitation at 290 nm and emission at 328 nm on a PTI Alphascan fluorometer (Photon Technology International, South Brunswick, New Jersey). Wild-type and mutant proteins incubated in a series of concentrations of Gdn for 6–24 h showed no further changes in fluorescence at which time the data were collected. A detailed procedure has been described elsewhere (Ropson et al., 1990; Kim et al., 1996).

Refolding kinetics. The kinetics of refolding were followed by fluorescence using an Applied Photophysics stopped-flow spectrophotometer (Model SX18MV) interfaced to an Archimedes computer. Changes in fluorescence (λ_{ex} = 290 nm and λ_{em} >305 nm with a WG305 Schott glass cutoff filter) were followed at 20 °C. The unfolded proteins were preincubated in buffered Gdn (1.6–2.5 M), 20 mM potassium phosphate, pH 7.3, 0.25 mM EDTA and diluted sixfold using a syringe ratio of 5:1. The initial Gdn concentrations used are given in the figure legends. Different concentrations of Gdn in the same buffer but in the absence of protein were used to obtain the final Gdn concentration given in Figure 4.

Measurement of DAUDA binding. Experiments were performed using the PTI Alphascan fluorometer (Photon Technology International) with an excitation wavelength of 350 nm and an emission wavelength of 500 nm. The initial volume for the protein solution was 40 mL to reduce nonspecific binding to the walls of the cuvette. The volume of ligand added was less than 1 mL. The protein concentration used was 0.10–0.50 μM depending on whether the binding was loose or tight.

Acknowledgments

We thank Drs. David Cistola, Michael Hodsdon, and Jay Ponder for many helpful discussions and for their preliminary molecular dynamics calculations, George Rose and Rajgopal Srinivasan for their discussions of LINUS and Ran Kim for excellent technical assistance. This work was supported by National Institutes of Health Grant DK13332.

References

Banaszak L, Winter N, Xu Z, Bernlohr DA, Cowan S, Jones TA. 1994. Lipid-binding proteins: A family of fatty acid and retinoid transport proteins. *Adv Protein Chem 45*:89–151.

Chou PY, Fasman GD. 1977. Beta-turns in proteins. *J Mol Biol 115*:135–175.

Cistola DP, Kim K, Rogl H, Frieden C. 1996. Fatty acid interactions with a helix-less variant of intestinal fatty acid-binding protein. *Biochemistry 35*:7559–7565.

Clark PL, Liu ZP, Rizo J, Gierasch LM. 1997. Cavity formation before stable hydrogen bonding in the folding of a beta-clam protein. *Nat Struct Biol 4*:883–886.

Dealba E, Jimenez MA, Rico M, Nieto JL. 1996. Conformational investigation of designed short linear peptides able to fold into beta-hairpin structures in aqueous solution. *Fold Design 1*:133–144.

Garrett JB, Mullins LS, Raushel FM. 1996. Are turns required for the folding of ribonuclease T1. *Protein Sci 5*:204–211.

Gu HD, Kim D, Baker D. 1997. Contrasting roles for symmetrically disposed beta-turns in the folding of a small protein. *J Mol Biol 274*:588–596.

Hoeltzli SD, Frieden C. 1995. Stopped-flow NMR spectroscopy: Real-time unfolding studies of 6-F-19-tryptophan-labeled *Escherichia coli* dihydrofolate reductase. *Proc Natl Acad Sci USA 92*:9318–9322.

Hodsdon ME, Cistola DP. 1997a. Discrete backbone disorder in the nuclear magnetic resonance structure of apo intestinal fatty acid binding protein: Implications for the mechanism of ligand entry. *Biochemistry 36*:1450–1460.

Hodsdon ME, Cistola DP. 1997b. Ligand binding alters the backbone mobility of intestinal fatty acid-binding protein as monitored by ^{15}N NMR relaxation and ^{1}H exchange. *Biochemistry 36*:2278–2290.

Hodsdon ME, Ponder JW, Cistola DP. 1996. The NMR solution structure of intestinal fatty acid-binding protein complexed with palmitate: Application of a novel distance geometry algorithm. *J Mol Biol 264*:585–602.

Hynes TR, Kautz RA, Goodman MA, Gill JF, Fox RO. 1989. Transfer of a beta-turn structure to a new protein context. *Nature 339*:73–76.

Kim K, Cistola DP, Frieden C. 1996. Intestinal fatty acid-binding protein: The structure and stability of a helix-less variant. *Biochemistry 35*:7553–7558.

Kim K, Ramanathan R, Frieden C. 1997. Intestinal fatty acid binding protein: A specific residue in one turn appears to stabilize the native structure and be responsible for slow refolding. *Protein Sci 6*:364–372.

Kraulis P. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein stucture. *J Appl Crystallog 24*:946–950.

Kurian E, Kirk WR, Prendergast FG. 1996. Affinity of fatty acid for rat intestinal fatty acid binding protein: Further examination. *Biochemistry 35*:3865–3874.

Minor DL Jr, Kim PS. 1994a. Context is a major determinant of beta-sheet propensity. *Nature 371*:264–267.

Minor DL Jr, Kim PS. 1994b. Measurement of the beta-sheet-forming propensities of amino acids. *Nature 367*:660–663.

Plaxco KW, Simons ST, Baker D. 1998. Contact order, transition state placement and the refolding rates of single domain proteins. *J Mol Biol 277*:985–994.

Ramirez-Alvarado M, Blanco FJ, Serrano L. 1996. De novo design and structural analysis of a model beta-hairpin peptide system. *Nat Struct Biol 3*:604–612.

Richieri GV, Ogata RT, Kleinfeld AM. 1996. Kinetics of fatty acid interactions with fatty acid binding proteins from adipocyte, heart, and intestine. *J Biol Chem 271*:11291–11300.

Ropson IJ, 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 19F NMR. *Proc Natl Acad Sci USA 89*:7222–7226.

Ropson IJ, Gordon JI, Frieden C. 1990. Folding of a predominantly beta-structure protein: Rat intestinal fatty acid binding protein. *Biochemistry 29*:9591–9599.

Rose GD, Gierasch LM, Smith JA. 1985. Turns in peptides and proteins. *Adv Protein Chem 37*:1–109.

Sacchettini JC, Gordon JI. 1993. Rat intestinal fatty acid binding protein. A model system for analyzing the forces that can bind fatty acids to proteins. [Review]. *J Biol Chem 268*:18399–18402.

Sacchettini JC, Scapin G, Gopaul D, Gordon JI. 1992. Refinement of the structure of *Escherichia-coli*-derived rat intestinal fatty-acid binding-protein with bound oleate to 1.75-angstrom resolution: Correlation with the structures of the apoprotein and the protein with bound palmitate. *J Biol Chem 267*:23534–23545.

Santoro MM, Bolen DW. 1988. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry* 27:8063–8068.

Scapin G, Gordon JI, Sacchettini JC. 1992. Refinement of the structure of recombinant rat intestinal fatty acid-binding apoprotein at 1.2 Å resolution. *J Biol Chem 267*:4253–4269.

Serrano L, Matouschek A, Fersht AR. 1992. The folding of an enzyme. III. Structure of the transition state for unfolding of barnase analysed by a protein engineering procedure. *J Mol Biol 224*:805–818.

Sibanda BL, Thornton JM. 1991. Conformation of beta hairpins in protein structures: Classification and diversity in homologous structures. *Methods Enzymol 202*:59–82.

Smith CK, Regan L. 1995. Guidelines for protein design: The energetics of beta sheet side chain interactions. *Science 270*:980–982.

Smith CK, Withka JM, Regan L. 1994. A thermodynamic scale for the beta-sheet forming tendencies of the amino acids. *Biochemistry 33*:5510–5517.

Srinivasan R, Rose GD. 1995. LINUS: A hierarchic procedure to predict the fold of a protein. *Proteins 22*:81–99.

Strickland EH. 1974. Aromatic contributions to circular dichroism spectra of proteins. *CRC Crit Rev Biochem 2*:113–175.

Stroup AN, Gierasch LM. 1990. Reduced tendency to form a beta turn in peptides from the P22 tailspike protein correlates with a temperature-sensitive folding defect. *Biochemistry 29*:9765–9771.

Ybe JA, Hecht MH. 1996. Sequence replacements in the central beta-turn of plastocyanin. *Protein Sci 5*:814–824.

Zhang F, Lucke C, Baier LJ, Sacchettini JC, Hamilton JA. 1997. Solution structure of human intestinal fatty acid binding protein: Implications for ligand entry and exit. *J Biomol NMR 9*:213–228.

Zhou HX, Hoess RH, DeGrado WF. 1996. *In vitro* evolution of thermodynamically stable turns. *Nat Struct Biol 3*:446–451.