Intestinal Fatty Acid Binding Protein: Characterization of Mutant Proteins Containing Inserted Cysteine Residues†

Nan Jiang and Carl Frieden*

Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110

Received June 15, 1993; Revised Manuscript Received August 12, 1993®

ABSTRACT: Site-directed mutagenesis was used to introduce cysteine residues into the rat intestinal fatty acid binding protein, an almost all β-sheet protein that in the wild-type contains neither cysteine nor proline residues. Six mutants (I23C, S53C, V60C, L72C, L89C, and A104C) with a single cysteine residue substituted for a hydrophobic residue were characterized by their stability toward denaturants at pH 7.2 and 9.6, by their fluorescent properties, and by their reactivity toward the sulfhydryl modifying reagents 5,5′-dithiobis(2-nitrobenzoic acid) (DTNB) and 4,4′-dipyridyl disulfide (4-PDS). In terms of protein stability, the substitutions were reasonably conservative with only two (V60C and L89C) being somewhat less stable than the wild-type. The mutant proteins differed considerably, however, in their reactivity toward the modifying reagents. One residue, Cys89, located in a hydrophobic core near a turn between two β-strands, was unreactive, while two residues, Cys60 and Cys104, located in the middle of β-strands in the cavity into which fatty acid binds, reacted only very slowly and were further protected by oleate. Cys53, located near a turn and partially buried, appeared to have an unusually low p*K* value. Two residues, Cys23 and Cys72, reacted more rapidly in the native protein than in the unfolded protein. Both residues are located near the portal for the fatty acid binding, and one, Cys72, was strongly protected from modification by the presence of oleate. Examination of the crystal structure indicates that Cys72 is not easily solvent-accessible. We conclude that this high reactivity for this residue may be a consequence of rapid conformational flexibility in this region of the structure.

Rat intestinal fatty acid binding protein (IFABP)[1] is one of a family of small (15 kDa), single-chain proteins that bind fatty acids, retinol, retinoic acid, or bile salts (Gordon et al., 1991; Jones, P. D., et al., 1988; Sacchettini et al., 1990; Matarese et al., 1989). IFABP is synthesized abundantly in the epithelial cells of the small intestine, and although its function remains unclear (Sweetser et al., 1987), it is an excellent protein for structural studies because of its small size, stability, and efficient overexpression in *Escherichia coli* (Lowe et al., 1987). It is of particular interest in protein folding studies because it is a mostly β-sheet protein containing neither proline nor cysteine residues and can be reversibly unfolded by denaturing agents such as urea or GdnHCl (Ropson et al., 1990). The three-dimensional structures of both the apo- and holo(oleate) protein have been determined to a resolution of 1.2 and 1.75 Å, respectively (Scapin et al., 1992; Sacchettini et al., 1992), and it (IFABP), similar to other proteins in this family that have been structurally characterized (Sacchettini et al., 1987; Jones, T. A., et al., 1988; Scapin et al., 1990; Benning et al., 1992; Zanotti et al., 1992; Xu et al., 1993; Winter et al., 1993), contains 10 antiparallel β-strands organized into 2, nearly orthogonal, β-sheets. These sheets enclose a rather large cavity into which the fatty acid binds. There are two short segments of α-helix connected to the first two β-strands which appear to be near the entrance for fatty acid binding. In a previous investigation using ^{19}F-labeled tryptophan, we showed that when the protein was partially unfolded with denaturant there appears to be a fairly high concentration of an intermediate in equilibrium with the native and unfolded forms (Ropson & Frieden, 1992). This intermediate was more similar to the unfolded form than to the folded protein and may contain an initiation site for the folding process.

Since IFABP does not contain any cysteine residues in its normal sequence, we decided to insert such residues at specific positions in the protein and then examine the reactivity of these site-directed mutants toward thiol modifying reagents. While such measurements do not give direct information about folding properties, other investigators have measured the reactivity of inserted cysteine groups during folding (e.g., Ballery et al., 1993) or their ability, upon modification with alkylating reagents, to block the folding process or to pack incorrectly (Wynn & Richards, 1993). Additionally, this approach provides a method to examine properties of various regions of the protein in terms of the reactivity of the cysteine residues toward thiol modifying reagents. It is well-known that the reactivity of cysteine residues in proteins varies considerably. It is not always clear, however, what factors might be involved in this difference in reactivity. Furthermore, in proteins in which thiol reactivity has been explored, investigators have generally been limited to those cysteine residues present in the native structure. The availability of a protein that contains no cysteines but for which site-directed mutants are readily made and easily overexpressed allows for a more systematic investigation of cysteine reactivity. The present report examines this issue. We have made six relatively conservative substitutions and have found that the cysteine reactivity varies greatly depending on its position in the native structure. Of most interest are two substitutions (positions 23 and 72) where the reactivity is greater in the folded form of the protein than in the unfolded state. While one of these residues (23) is solvent-exposed, the crystal structure of the protein indicates

† Supported by USPHS Grant DK13332.

* Author to whom correspondence should be addressed.

® Abstract published in *Advance ACS Abstracts*, October 1, 1993.

[1] Abbreviations: IFABP, intestinal fatty acid binding protein; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); 4-PDS, 4,4′-dipyridyl disulfide; DTT, dithiothreitol; EDTA, ethylenediaminetetraacetic acid; GdnHCl, guanidine hydrochloride.

0006-2960/93/0432-11015$04.00/0

that the other is fairly well protected from solvent by other amino acid residues. Both positions 23 and 72 are near the entrance for the binding of fatty acid to the interior of the protein, and the results imply that there must be rapid conformational flexibility in that region of the protein.

MATERIALS AND METHODS

Materials. [α-^{35}S]ATP was purchased form Amersham. The sequencing kit and Sequenase v2.0 were from United States Biochemical. GdnHCl was obtained from ICN Biochemicals and ultrapure urea from United States Biochemical. They were treated as previously described before use (Ropson et al., 1990). DTNB and 4-PDS were obtained from Sigma. Oleate was made as a 10 mM stock solution by dissolving oleic acid (Nucheck Inc.) in 0.1 N potassium hydroxide. All other reagents used for preparation and purification of mutant proteins were purchased from regular chemical reagent suppliers. Mini-protean II Ready gel (4–20% gradient) from Bio-Rad was used for the electrophoresis.

Protein Preparation and Purification. Cysteine mutations were introduced into the expression vector pMON5840-IFABP recombinant plasmid using the method of Kunkel (1985). The entire gene sequence of each IFABP mutant was determined to ensure that the mutant contained only one altered residue. Wild-type protein and cysteine mutants were overexpressed in *E. coli* strain MG1655 by inducing the *recA* promoter of the vector with nalidixic acid. Purification of mutant proteins was achieved by using a previously published protocol for wild-type IFABP (Sacchettini et al., 1990) except that 1 mM DTT was added to all buffers. Six cysteine mutants were obtained this way: I23C, S53C, V60C, L72C, L89C, and A104C. All six mutant proteins were overexpressed in *E. coli* with yields of 300–800 mg of protein from 4 L of culture. After purification, all mutants showed a single band on SDS–polyacrylamide gels under reducing conditions. To test for the presence of dimer in the absence of DTT, SDS–polyacrylamide nonreducing gel assays were run. For these gels, samples were dissolved in loading buffer (62 mM Tris-HCl, pH 6.8, 10% glycerol, 2% SDS, and 0.025% bromophenol blue) and then loaded onto the gel without boiling. When the gel was overloaded with mutant protein, a faint band with molecular weight equivalent to an IFABP dimer was observed. Mutant I23C had the largest dimer:monomer ratio and mutant V60C the least. However, the presence of the dimer band was negligible when compared to the very intense monomer band.

Protein concentrations were determined by measuring A_{280nm} in 6 M GdnHCl using $\epsilon = 16\,620\ M^{-1}\ cm^{-1}$ [calculated according to Johnson (1988)]. Purified proteins (1 mM) were stored in buffer containing 50 mM potassium phosphate, pH 7.2, 0.1 mM EDTA, 1 mM DTT, and 100 mM KCl at 4 °C. Proteins used for thiol group reactivity studies were further treated with DTT (10–20-fold molar excess) at 4 °C for overnight followed by dialysis in 20 mM potassium phosphate, pH 7.2, and 0.1 mM EDTA at 4 °C, and then stored at −70 °C.

Fluorescence Studies of Cysteine Mutants. Fluorescence emission spectra (300–400 nm) on excitation at 290 nm were collected on a PTI spectrofluorometer (Alphascan). Fluorescence emission spectra of native proteins were collected using solutions containing 1 μM protein, 1 mM DTT, and 20 mM potassium phosphate, 0.1 mM EDTA, pH 7.2, or 20 mM sodium pyrophosphate, pH 9.6. For unfolded proteins, 2.5 M GdnHCl was present in buffer. All experiments were done at 20 °C.

Equilibrium Denaturation Studies of IFABP Mutants. Denaturation curves were determined as a function of GdnHCl concentration using the value of the fluorescence emission change at 322 nm for the L89C mutant and at 327 nm for all other mutants. Equilibrium data were fit to a two-state model as previously described (Ropson et al., 1990) using an equation adapted from Bolen and Santoro (1988). All measurements were done at 20 °C.

Reactivity Studies of Thiol Groups. The kinetics of the cysteine reaction were followed with an Applied Photophysics stopped-flow apparatus using either DTNB or 4-PDS and recording the absorption change at 412 or 324 nm, respectively. DTNB was made as a 10 mM stock solution in 20 mM potassium phosphate, pH 7.0, and was diluted to the working concentration before use. 4-PDS was prepared as 1 mM stock solution in the same buffer each time before use and then diluted to the desired concentration. A 10 mM stock solution of glutathione, used for cysteine control reactions, was also made in 20 mM potassium phosphate buffer, pH 7.0. All stock solutions and dilution buffers were filtered through a 0.22-μm filter before use. To measure the cysteine reactivity, 4–8 μM protein in 20 mM buffer (potassium phosphate for pH 6.4–8.0, sodium pyrophosphate for pH 8.4–9.6) was mixed with an equal volume of 200 μM DTNB in H_2O. For the unfolded proteins, IFABP was preincubated in 9 M urea buffered with 20 mM sodium pyrophosphate, pH 9.6, and then mixed with an equal volume of DTNB in 10 M urea buffered with 5 mM potassium phosphate, pH 6.9. The final pH was 9.6. All reactions were done at 20 °C. Similar experiments were conducted with 4-PDS. Kinetic data were fit by nonlinear regression using either a one- or a two-exponential function. The reactivity of the holoproteins was measured in the presence of a 2-fold molar excess of oleate.

Difference Distance Map Calculation. Coordinates for the crystal structures of the apo- and holo(oleate) protein (Sacchettini et al., 1992; Scapin et al., 1992) were obtained from Dr. James Sacchettini. The method used here for assessing conformational changes on ligand addition is not based upon ensemble averages, like the center-of-mass, but on calculating a difference distance map. In our procedure, provided by Dr. George Rose (personal communication), corresponding distances between every two atoms are measured in both the apoprotein and the holoprotein, and their differences are taken. That is, for all i and j, the quantity $\Delta\delta_{ij}$ is determined, where

$$\Delta\delta_{ij} = \delta_{ij}^{apo} - \delta_{ij}^{holo}$$

Here, δ_{ij}^{apo} is the distance from i to j in the apoprotein, δ_{ij}^{holo} is the distance from i to j in the holoprotein, and $\Delta\delta_{ij}$ is the difference between these two distances. $\Delta\delta_{ij}$ does not depend on the center-of-mass.

In Figure 3, the quantity $\Delta\delta_{ij}$ has been calculated for all α-carbon pairs. Each square in the figure indicates that the distance between $C_\alpha(i)$ and $C_\alpha(j)$ has changed by at least 1.0 Å upon binding fatty acid. Open squares annotate an (i,j) distance that has decreased in the holoprotein; filled squares annotate a distance that has increased.

RESULTS

Fluorescence Spectra of Native and Unfolded IFABP Mutants. Fluorescence data, reflecting the two tryptophan residues at positions 6 and 82, for wild-type IFABP and cysteine mutants are given in Table I. When denatured in 2.5 M GdnHCl in 20 mM potassium phosphate, 1 mM DTT, and 0.1 mM EDTA, pH 7.2, all proteins at the same concentration

Table I: Comparison of the Fluorescence of Wild-Type IFABP and Cysteine Mutants[a]

protein	F_{rel}, pH 7.2	F_{rel}, pH 9.6
WT	1.00	0.82
I23C	1.01	0.84
S53C	1.16	1.14
V60C	1.02	0.88
L72C	1.02	1.09
L89C[b]	0.42	0.33
A104C	0.86	0.63
WT in 2.5 M GdnHCl[c]	0.23	0.23

[a] Relative fluorescence at the fluorescence maximum of 327 nm upon excitation at 290 nm. Conditions used: 1 μM IFABP in 0.1 mM EDTA and 1 mM DTT in 20 mM potassium phosphate buffer, pH 7.2, or 20 mM sodium pyrophosphate buffer, pH 9.6. [b] Value given here is at 322 nm, the wavelength of maximum fluorescence for L89C. [c] The maximum fluorescence is at 350 nm. Value given here is at 327 nm.

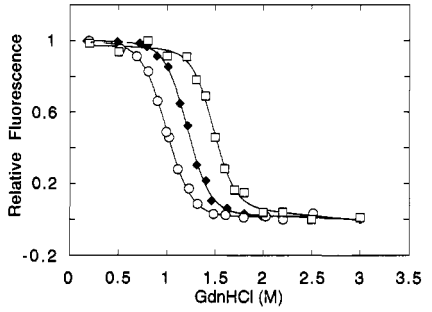

FIGURE 1: Relative fluorescence of three IFABP mutants as a function of GdnHCl concentration. Experiments were performed in 20 mM potassium phosphate buffer, 0.1 mM EDTA, and 1 mM DTT at pH 7.2 and 20 °C. The protein concentration used was 1 μM. Fluorescence was measured at 322 nm for L89C and at 327 nm for S53C and A104C with excitation at 290 nm. Mutant IFABPs: L89C (○); S53C (♦); A104C (□). The solid lines are fits of the data to the two-state model as previously described (Ropson et al., 1990).

exhibited identical emission spectra with a peak at 350 nm. Except for L89C, all native proteins had a maximum emission at 327 nm with intensities much larger than the unfolded proteins. While I23C, V60C, and L72C had the same fluorescence intensity as wild-type IFABP, A104C was about 14% less fluorescent while S53C was 16% more fluorescent than wild-type. L89C, under the same conditions, gave only 42% of the fluorescence signal as wild-type, and the emission peak was shifted to 322 nm. As shown in Table I, similar results were also observed at pH 9.6. Except for mutants S53C and L72C, the fluorescence at pH 9.6 was lower than that of the wild-type protein at pH 7.2.

Equilibrium Denaturation Studies of IFABP Mutants. Figure 1 shows the relative fluorescence at 327 nm (excitation wavelength at 290 nm) of three of the mutants (S53C, L89C, and A104C) as a function of GdnHCl concentration at pH 7.2. The data for all the mutants were characteristic of a two-state model similar to that observed here and earlier (Ropson et al., 1990) for the wild-type protein. As shown in the figure and as true for all the mutants, the decrease in fluorescence intensity at 327 nm between 0.8 to 2.0 M GdnHCl indicated a cooperative unfolding as observed for wild-type IFABP. Except for the V60C and L72C mutants (data not shown), there was little or no effect of denaturant concentration on the fluorescence intensity between 0 and 0.8 M GdnHCl. Unfolding of native V60C showed a moderate decline in fluorescence intensity before reaching the transition. On the other hand, mutant L72C showed an increasing fluorescence intensity at GdnHCl concentrations below 0.8 M and reached a maximum value when the denaturant concentration was increased to 0.8 M. At concentrations above 2.5 M, the effect of denaturant concentration on the fluorescence of unfolded proteins was hardly observed. The data in Table II, showing ΔG_{H_2O} and midpoint values, were obtained from nonlinear fits of unfolding data to a two-state model. R values of a least-squares fit fell within 0.999–0.995. The apparent free energy of unfolding (ΔG_{H_2O}) at 0 M GdnHCl differed for each mutant, ranging from 4.62 to 7.76 kcal mol^{-1} at pH 7.2 and from 3.9 to 7.49 kcal mol^{-1} at pH 9.6. Compared to wild-type IFABP at either pH, mutants I23C, S53C, and L72C had similar midpoints, V60C and L89C had lower midpoints, and A104C had a higher midpoint. The data for the wild-type at pH 7.2 are in good agreement with previously published values (Ropson et al., 1990).

Table II: Stability of Wild-Type and Cysteine Mutants of IFABP[a]

	pH 7.2		pH 9.6	
protein	ΔG_{H_2O}[b] (kcal mol^{-1})	midpoint[c] (M)	ΔG_{H_2O} (kcal mol^{-1})	midpoint (M)
WT	5.22 ± 0.33	1.36	5.59 ± 0.59	1.41
I23C	6.13 ± 0.40	1.26	7.49 ± 0.42	1.41
S53C	5.80 ± 0.27	1.21	7.09 ± 0.77	1.46
V60C	6.02 ± 0.29	1.18	6.18 ± 0.38	1.21
L72C	5.08 ± 0.30	1.37	5.47 ± 0.35	1.49
L89C	4.62 ± 0.18	1.0	3.90 ± 0.47	1.14
A104C	7.76 ± 0.90	1.46	7.10 ● 0.73	1.54

[a] Fluorescence experiments (Ex = 290 nm, Em = 322 nm for L89C and Em = 327 nm for others) using 1 μM protein in 20 mM potassium phosphate, 0.1 mM EDTA, and 1 mM DTT, pH 7.2, or 20 mM sodium pyrophosphate/1 mM DTT, pH 9.6, at 20 °C. Data were fit to a two-state model as previously described (Ropson et al. 1990). [b] ΔG_{H_2O} is the apparent free energy difference between native and unfolded forms of protein extrapolated to 0 M GdnHCl. [c] The concentration of GdnHCl present at the midpoint of the transition.

On the basis of either ΔG_{H_2O} or the denaturation midpoint, Table II also shows that the mutants and wild-type IFABP were slightly more stable at pH 9.6 than at pH 7.2 but that there are slight changes in the stability relative to the wild-type.

Thiol Group Reactivity with DTNB. For the thiol reactivity studies of unfolded protein at pH 9.6, described below, urea was used instead of GdnHCl because it was difficult to maintain the high pH values in the presence of the high salt associated with GdnHCl. Since 8 M urea was sufficient to unfold wild-type protein (Ropson et al., 1990), 9.5 M urea was used in these experiments. Table III shows rate constants for the reaction of DTNB with native and unfolded mutant proteins at pH 9.6. When denatured in 9.5 M urea/20 mM sodium pyrophosphate, pH 9.6, all six mutants exhibited a single phase with rate constants ranging from 0.0075 to 0.07 μM^{-1} s^{-1}. A control experiment for the denatured proteins using glutathione (Glu-Cys-Gly) under the same conditions (9.5 M urea) gave a rate constant of 0.014 μM^{-1} s^{-1}. The kinetics of native mutants reacting with DTNB at pH 9.6 were remarkably different from those of the unfolded mutants. Thus, L89C was unreactive toward DTNB. Of the remaining five reactive mutants, the reactivity of cysteines in the mutants decreased in the order I23C > L72C ≫ S53C > V60C ~ A104C. The fastest rate was 0.42 μM^{-1} s^{-1} (I23C) and the slowest 0.003 μM^{-1} s^{-1} (V60C and A104C). The rate constants observed for I23C and L72C were greater than those observed for the unfolded proteins. Data for I23C, V60C, L72C, and A104C, however, could only be fit using a two-exponential equation. For these reactions, the total amplitudes were proportional to protein concentration, and the ratio of the amplitudes of the second to the first phase differed somewhat for different mutants.

Table III: Reactivity of Cysteines in IFABP Mutants with DTNB at pH 9.6

	unfolded[a] k (μM^{-1} s^{-1})	apo native[b]			holonative[c]		
		k_1 (μM^{-1} s^{-1})	k_2 (μM^{-1} s^{-1})	A_2/A_1[d]	k_1 (μM^{-1} s^{-1})	k_2 (μM^{-1} s^{-1})	A_2/A_1
I23C	0.070	0.42	0.045	0.58	0.053	0.0063	0.10
S53C	0.039	0.011	*e*		0.0075	*e*	
V60C	0.014	0.0033	0.00069	0.88	0.00013	*e*	
L72C	0.0075	0.22	0.014	1.6	0.00050	0.00016	2.3
L89C	0.057	<0.0001	*e*		nd[f]	nd	
A104C	0.031	0.0025	0.00036	0.53	<0.0001	*e*	
glutathione	0.014	0.024	*e*				

[a] 2 μM protein reacted with 100 μM DTNB in 9.5 M urea and 10 mM sodium pyrophosphate, pH 9.6 at 20 °C. Errors for rate constants were within ± 2%. [b] Reactions were carried out in the same buffer as above in the absence of urea. [c] A 2-fold molar excess of oleate was added to 4 μM protein in 20 mM sodium pyrophosphate, pH 9.6, prior to mixing with an equal volume of 128 μM DTNB at 20 °C. [d] Ratio of the amplitudes of the second to the first phase. [e] Not detectable. [f] Not determined.

Effect of Fatty Acid Binding on the Reactivity of Cysteine Mutants. As shown in Table III and by the above data, all the mutants except L89C were reactive with DTNB. Since some of the cysteine residues are inside the cavity that is the binding pocket for fatty acid, the effect of a 2-fold excess of oleate added to mutants was tested at pH 9.6. As expected, since these residues are pointing into the binding pocket, oleate almost completely protected V60C and A104C from reacting. It also dramatically reduced the reactivity of L72C from 0.22 to <0.0005 μM^{-1} s^{-1}, suggesting that this cysteine was also affected by the bound oleate. The effect of oleate binding on the reactivity of S53C and I23C was considerably less with rate decreases of 32% and 87%, respectively.

pH Dependence of the Reaction Rate. The reaction rate of cysteine with DTNB depends on the pK_a of the thiol group. In the above experiments, pH 9.6 was chosen because it was expected that the thiol group would be deprotonated. In order to determine the pK_a of the thiol group, we measured the DTNB reaction rate as a function of pH for two mutants (L72C and A104C) which were quite different in their reactivities at pH 9.6. The data for both these mutants fit a simple titration curve with a pK_a value of 8.7. The same value was obtained for either kinetic phase of the reaction. Other mutants were tested at pH 8.5 and 7.2. The reaction rate of S53C was pH-independent, indicating an abnormal pK_a value for this cysteine residue. All other mutants, however, showed pH dependencies consistent with pK_a values between 8 and 9.2.

Reactivity of Cysteine Mutants with 4-PDS. 4-PDS has a similar structure to DTNB, but it does not have nitro or carboxyl groups and is therefore not negatively charged at pH values above 7.2 as is DTNB. To examine the possible effect of a charge on the reaction rate, we measured reaction rates for the native mutants with 4-PDS for comparison to the DTNB reaction (Table IV) at pH 9.6. In general, the native protein reaction rates were 4–10-fold faster when using 4-PDS. Mutants which showed biphasic kinetics with DTNB also showed biphasic kinetics with 4-PDS. The reaction of I23C with 4-PDS, however, was too fast to measure, being completed within the dead time of stopped-flow (<2 ms). L89C, which did not react with DTNB, reacted slowly with 4-PDS in a biphasic fashion. The amplitude of the second phase was one-half to one-fifth that of the first phase for mutants showing biphasic kinetics. To compare the difference in reactivity of a thiol group toward DTNB and 4-PDS, we measured the rate constants of unfolded protein and of glutathione using either DTNB or 4-PDS at pH 9.6. Tables III and IV show that the results are similar for both DTNB and 4-PDS.

Table IV: Reactivity of Cysteine Mutants with 4-PDS at pH 9.6

	unfolded[a]	apo native[b]		
	k (μM^{-1} s^{-1})	k_1 (μM^{-1} s^{-1})	k_2 (μM^{-1} s^{-1})	A_2/A_1[c]
I23C	0.045	>4		
S53C	0.043	0.016	*e*	
V60C	0.042	0.013	0.0036	0.42
L72C	0.026	2.0	0.093	0.19
L89C	0.034	0.00084	0.000017	0.50
A104C	0.036	0.025	0.0026	0.38
glutathione[d]	0.025	0.043	*e*	

[a] 4 μM protein reacted with 102 μM 4-PDS in 9.3 M urea and 10 mM sodium pyrophosphate, pH 9.6 at 20 °C. Errors for rate constants were within ±2%. [b] 4 μM protein reacted with 90 μM 4-PDS in 10 mM sodium pyrophosphate pH 9.6 buffer at 20 °C. Reaction using I23C was completed within the dead time of the stopped-flow. [c] Ratio of the amplitudes of the second to the first phase. [d] 4 μM glutathione reacted with 172 μM 4-PDS in 10 mM sodium pyrophosphate buffer, pH 9.6 at 20 °C. [e] Not detectable.

DISCUSSION

Several studies have investigated the properties of cysteine residues inserted into proteins by site-directed mutagenesis. Martensson et al. (1993), for example, have introduced cysteine residues into carbonic anhydrase II to map the properties of a folding intermediate by measuring the accessibility of the thiol group to alkylation. Ballery et al. (1990, 1993) introduced cysteine residues into yeast phosphoglycerate kinase also in order to study the folding process by thiol group accessibility. These studies have been directed toward investigating the reactivity of the thiol group during refolding. Wynn and Richards (1993) have introduced a specific cysteine into *E. coli* thioredoxin which they then modified with a series of straight-chain aliphatic thiosulfonates. Such derivatives can then be used to study amino acid packing. In our studies, we are initially interested in the reactivity of inserted cysteine residues in the native and unfolded forms of IFABP.

Location of Cysteine Residues. Wild-type rat intestinal fatty acid binding protein does not contain cysteine residues. We can therefore create any number of mutants that contain only a single cysteine residue. Figure 2 (top) is a stereoview of the apoprotein backbone obtained from the known crystal structure (Scapin et al., 1992) where the location of all the cysteine residues of all the mutants we have prepared is shown. The mutations are distributed throughout the structure and located either in the middle of β-strands (mutants V60C and A104C), at or near turns between strands (mutants S53C, L72C, and L89C), or between the two small helical stretches (mutant I23C). In terms of side-chain size, the replacements have been generally conservative, with cysteine replacing either serine, alanine, leucine, valine, or isoleucine. As shown by the figure, there is a large cavity in the structure. This cavity contains the fatty acid (not shown). The cysteine residues at positions 60 and 104 are located deep in the pocket into which the fatty acid binds while Cys72 is located near the entrance to that pocket. Although no energy minimization has been

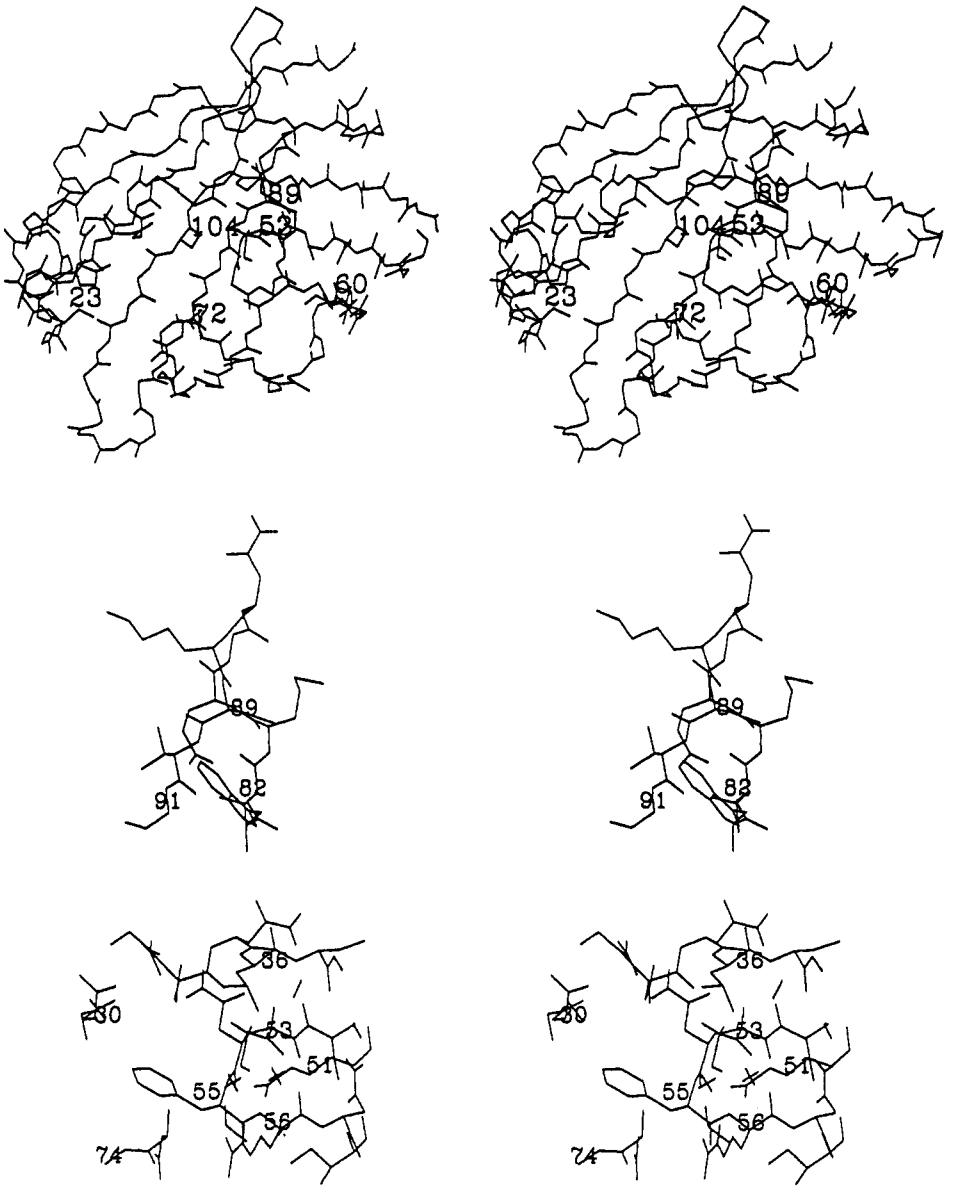

FIGURE 2: Stereoviews of apo-IFABP based on the available crystal structure at 1.2 Å (Scapin et al., 1992). (Top) The complete structure with all six mutants is illustrated. Only the backbone is shown except for the cysteine residues. Normal residues were replaced with cysteine using the molecular modeling program Insight II. There has been no attempt to energy-minimize the structure after substituting cysteine for the native amino acid. (Middle) The region around Cys89 with residues 82–91 is shown. Note the proximity of Trp82 to Cys89, probably accounting for the decreased fluorescence of the L89C mutant. (Bottom) The region within 10 Å of the S atom of Cys53 is shown. Labeled residues are Leu30, Leu36, Glu51, Cys53, Phe55, Arg56, and Asp74. The X's represent ordered water. All views are in the same orientation.

performed, none of these three residues appears to be easily accessible to solvent. Cys89, located near a turn between two β-strands, is buried within a hydrophobic core surrounded by aromatic residues. A stereoview of this region of the protein is shown in Figure 2 (middle). It is this portion of the protein that we have suggested may be involved in the early folding steps (Ropson & Frieden, 1992). Close by this cysteine side chain is Trp82, and this may account for the decreased fluorescence of the native form of L89C relative to the wild-type or other mutants since the thiol group is an effective quencher of indole fluorescence (Permyakov, 1993). This cysteine is clearly not solvent-accessible since it does not react with DTNB. Cys23 is located on the surface of the protein while Cys53 appears to be partially buried. This latter cysteine has an abnormally low pK_a value (discussed below). For the most part, as shown in Figure 1 and Table II, none of these amino acid substitutions markedly affected the stability of the protein. The most destabilizing change occurred for the L89C mutant, a mutant in which the thiol group was unreactive toward DTNB. One interesting observation was that all the proteins were slightly more stable at pH 9.6 than at pH 7.2. While the fluorescence of L89C was dramatically changed relative to wild-type, some of the other mutants showed small changes. The cysteine of A104C, a mutant in which the fluorescence is decreased 15%, is near Trp82 but not as close as Cys89. Fluorescence is slightly increased in the S53C mutant, but this cysteine group is closer to Trp6 than to Trp82.

Reactivity of Cysteine Residues. Certainly the most dramatic result was the great difference in reactivity of the different cysteine residues toward chemically modifying reagents. We had expected that most of the cysteine substitutions would be unreactive because they appeared to be buried in the structure. Instead, only one, Cys89, was unreactive. Those located deep within the fatty acid binding pocket (Cys104 and Cys60) reacted with the modifying reagents only slowly compared to the unfolded protein and

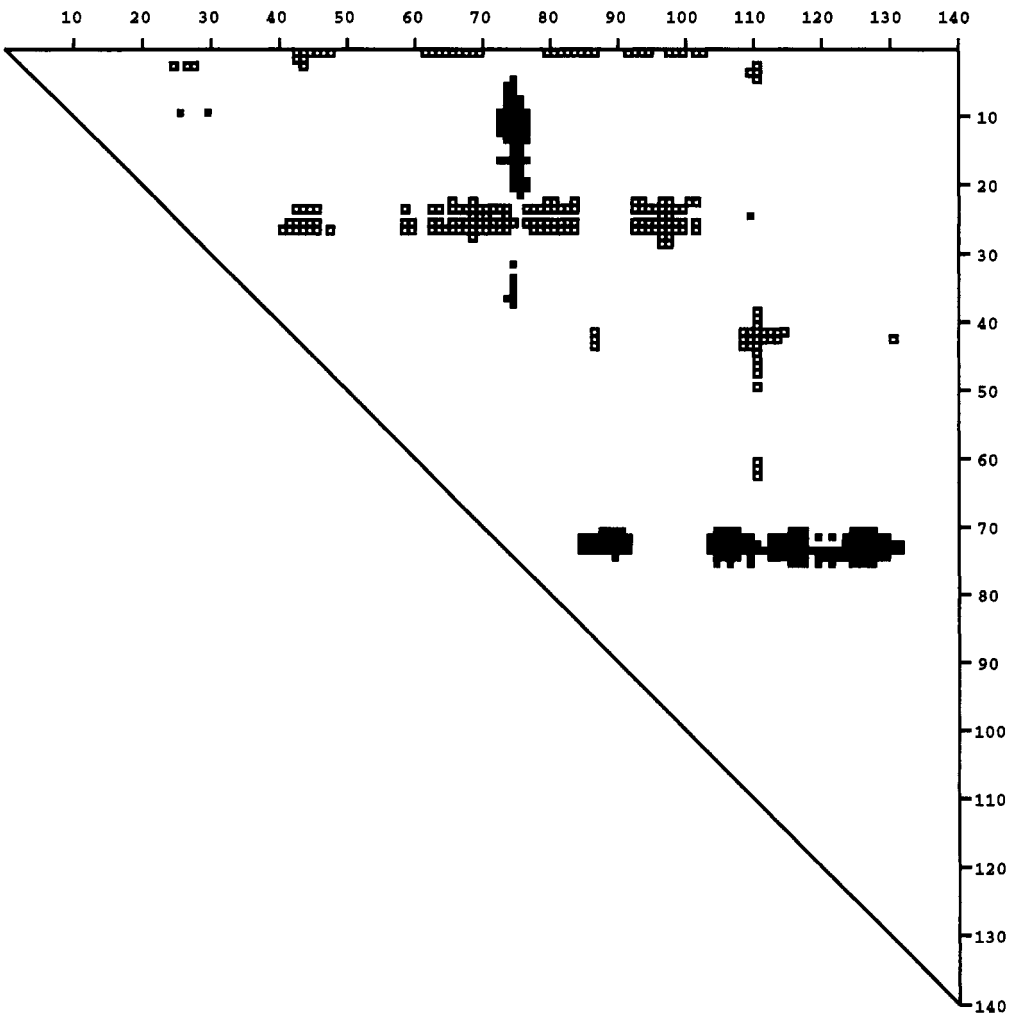

FIGURE 3: Difference distance map between the apo- and holo-forms of IFABP. The solid squares represent a threshold of 1.0 Å of main-chain α-carbons moving apart (i.e., only those changes greater than 1.0 Å are shown) whereas the open squares represent those α-carbons that have moved closer together as a consequence of ligand binding. See Materials and Methods for details.

were further protected from reaction by the presence of oleate. The low reactivity in the absence of oleate may be a consequence either of slow diffusion of DTNB or 4-PDS into the binding pocket or of an environment not conducive to reaction. Cys104, however, appears to have a relatively normal pK_a value of 8.7, suggesting a reasonably normal microenvironment. On the other hand, the reaction rate for DTNB with Cys53, located within a turn between two β-strands, is about the same in the folded protein as in the unfolded protein and is pH-independent from 7.2 to 9.6. Isoelectric focusing experiments (data not shown) and the reaction rate with sulfhydryl reagents are consistent with a low pK_a value. The reason for this low pK_a is not clear. The region around Cys53 is shown in a stereoview in Figure 2 (bottom). While a positively charged arginine residue (56) is nearby, there is a glutamate residue (51) just as close. The carboxyl of Glu51 is hydrogen-bonded to the guanidino moiety of the arginine via Tyr70 (Scapin et al., 1992). Located within 10 Å of the S atom are two leucine residues as well as Phe55 and Asp74. The reactivity of this cysteine with DTNB or 4-PDS indicates that the group is partially buried although two ordered water molecules are close by. A Connolly surface solvent accessibility calculation using Insight II indicates that this residue is not accessible to the bulk solvent. The results imply that there may be some conformational flexibility in this region of the protein.

The most surprising observation is that two residues, Cys23 and Cys72, are much more reactive in the native protein than in the unfolded protein. One, Cys23, is located on the surface of the protein, but both the crystal structure of the holoprotein and the partial oleate protection indicate that this residue is located near the terminal methyl group of the fatty acid. It is surrounded by both positively (Lys27 and Arg95) and negatively (Asp74) charged residues. Electrostatic interactions between the positively charged residues and the negatively charged DTNB could explain the high reactivity, but this is unlikely because the uncharged (at pH 9.6) 4-PDS reacts with Cys23 even more rapidly. Indeed, the rate of the reaction with 4-PDS was too fast to be measured under our conditions, suggesting a second-order rate constant of over 40 μM^{-1} s^{-1}, close, perhaps, to the rate of a diffusion-controlled reaction. It is also possible that there is some specific interaction between the modifying reagent and the protein.

The high reactivity of Cys72 is quite surprising because this position appears in the apoprotein to be reasonably well protected from accessibility to solvent (or modifying reagent) by several residues including Asp74, Thr76, Met21, and Arg95. Furthermore, oleate binding significantly inhibits the DTNB reaction, suggesting that the residue is within the binding pocket. Examination of the holoprotein structure shows that Cys72 is indeed in close proximity to the carbon backbone of the fatty acid.

It is quite possible that the ready reactivity of Cys72 is related to motions of the protein. Difference distance maps are a sensitive probe of structural differences between the apo- and holoproteins. Figure 3 shows a C_α distance difference calculation between the apo- and holoproteins (G. Rose, personal communication). It is clear from this representation that only two regions of the protein are primarily altered by fatty acid binding and these regions are in the neighborhood of residues 72–76 and 23–27. As pointed out by Scapin et al. (1992), these regions, in addition to Phe55, may represent the portal for the entrace of fatty acid into the internal cavity of the protein. These two regions and three others (41–47, 83–89, and 107–112) are also characterized by the highest values of the temperature factor (Scapin et al., 1992). Since access to Cys72 appears from the crystal structure to be limited, it seems likely that the high reactivity of this residue is due to conformational flexibility of the protein in the region that represents the entrance of the fatty acid to the interior of the protein. These conformational changes must be rapid (milliseconds or less) since the second-order rate constant for the 4-PDS reaction is at most only an order of magnitude lower than that for a diffusion-controlled reaction. A related observation has been made by Richiera et al. (1992), who found that acrylodan, normally a sulfhydryl modifying reagent, covalently modified Lys27. This residue also is near the entrance for the fatty acid and is one whose position is altered by fatty acid binding (Figure 3).

One puzzling result is the presence of a second kinetic phase in reactions with modifying reagents using native proteins (except for Cys53). The ratio of amplitudes of the second to the first phase using DTNB differed from mutant to mutant, ranging, in the absence of oleate, from 0.53 to 1.6 (Table III). It was first suspected that the two phases might be a consequence of a side reaction between a thiol group and oxygen dissolved in buffers. However, using the V60C mutant, the same two phases were observed when solutions were degassed before use. Incomplete delipidation during purification might also give two phases which would represent reactions of thiol groups in apo- and ligand-bound forms. To test this hypothesis, purified A104C was again run over the Lipidex column. Two phases and no noticeable change in reaction rates were observed. The possibility that structural heterogeneity may exist in these cysteines was also considered. Crystallographic studies on wild-type IFABP (Scapin et al., 1992) identified about 20 residues showing structural heterogeneity, none of them, however, in positions that we replaced. Two-dimensional ^1H NMR spectra (D. Cistola, personal communication) of the A104C and V60C mutants showed two peaks of equal intensity for the V60C mutant that might be assigned to the hydrogens of the C_β of the cysteine, but only one peak was observed for Cys104. It was not clear whether the two peaks in V60C spectra came from hydrogens in two rigid C_β configurations or they represented two slowly exchangeable conformational forms.

CONCLUSIONS

It is well-known that cysteine residues in proteins differ considerably in their reactivity toward alkylating or modifying reagents. By placing these residues in specific regions of a protein whose structure is known, it is possible to obtain some idea of the factors involved in such reactivity. The results, however, were surprising since some residues which would appear to be inaccessible to solvent were reactive toward either DTNB and/or 4-PDS. The most striking observation is that residues that are in the region of the entrace of the fatty acid to the interior of the protein react more rapidly in the native protein than in the unfolded protein. We conclude that this portion of the protein is undergoing rapid conformational change.

ACKNOWLEDGMENT

We thank Dr. George Rose for helpful discussions and for providing the analysis and its description of the information shown in Figure 3, Dr. James Sacchettini for the X-ray coordinates, and Dr. David Cistola for the NMR studies.

REFERENCES

Ballery, N., Minard, P., Desmadril, M., Betton, J.-M., Perahia, D., Mouawad, L., Hall, L., & Yon, J. M. (1990) *Protein Eng. 3*, 199–204.

Ballery, N., Desmadril, M., Minard, P., & Yon, J. M. (1993) *Biochemistry 32*, 708–714.

Benning, M. M., Smith, A. F., Wells, M. A., & Holden, H. M. (1992) *J. Mol. Biol. 228*, 208–219.

Bolen, D. W., & Santoro, M. M. (1988) *Biochemistry 27*, 8069–8074.

Gordon, J. I., Sacchettini, J. C., Ropson, I. J., Frieden, C., Li, E., Rubin, D. C., Roth, K. A., & Cistola, D. P. (1991) *Curr. Opin. Lipidol. 2*, 125–137.

Johnson, W. C., Jr. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 145–166.

Jones, P. D., Carne, A., Bass, N. M., & Grigor, M. R. (1988) *Biochem. J. 251*, 919–925.

Jones, T. A., Bergfors, T., Sedzik, J., & Unge, T. (1988) *EMBO J. 7*, 1597–1604.

Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 88*, 488–492.

Lowe, J. B., Sacchettini, J. C., Laposata, M., McQuillan, J. J., & Gordon, J. I. (1987) *J. Biol. Chem. 262*, 5931–5937.

Martensson, L.-G., & Jonsson, B.-H. (1993) *Biochemistry 32*, 224–231.

Matarese, V., Stone, R. L., Waggoner, D. W., & Bernlohr, D. A. (1989) *Prog. Lipid Res. 28*, 245–272.

Permyakov, E. A. (1993) *Luminescent Properties of Proteins*, p 48, CRC Press, Boca Raton, FL.

Richieri, G. V., Ogata, R. T., & Kleinfeld, A. M. (1992) *J. Biol. Chem. 267*, 23495–23501.

Ropson, I. J., & Frieden, C. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 7222–7226.

Ropson, I. J., Gordon, J. I., & Frieden, C. (1990) *Biochemistry 29*, 9591–9599.

Sacchettini, J. C., Stockhausen, D., Li, E., Banaszak, L. J., & Gordon, J. I. (1987) *J. Biol. Chem. 262*, 15756–15758.

Sacchettini, J. C., Banaszak, L. J, & Gordon, J. I. (1990) *Mol. Cell. Biochem. 98*, 81–93.

Sacchettini, J. C., Scapin, G., Gopaul, D., & Gordon, J. I. (1992) *J. Biol. Chem. 267*, 23534–23545.

Scapin, G., Gordon, J. I., & Sacchettini, J. C. (1992) *J. Biol. Chem. 267*, 4253–4269.

Scapin, G., Spadon, P., Mammi, M., Zanotti, G., & Monaco, H. L. (1990) *Mol. Cell. Biochem. 98*, 95–99.

Sweetser, D. A., Heuckeroth, R. O., & Gordon, J. I. (1987) *Annu. Rev. Nutr. 7*, 337–359.

Winter, N. S., Bratt, J. M., & Banaszak, L. J. (1993) *J. Mol. Biol. 230*, 1247–1259.

Wynn, R., & Richards, F. M. (1993) *Protein. Sci. 2*, 395–403.

Xu, Z. H., Bernlohr, D. A., & Banaszak, L. J. (1993) *J. Biol. Chem. 268*, 7874–7884.

Zanotti, G., Scapin, G., Spadon, P., Veerkamp, J. H., & Sacchettini, J. C. (1992) *J. Biol. Chem. 267*, 18541–18550.