Protein Science (1998), 7:1332–1339. Cambridge University Press. Printed in the USA.

The three-dimensional structure of a helix-less variant of intestinal fatty acid-binding protein

RUTH A. STEELE, DANIEL A. EMMERT, JEFF KAO, MICHAEL E. HODSDON, CARL FRIEDEN, AND DAVID P. CISTOLA

Departments of Biochemistry & Molecular Biophysics and Chemistry, Washington University School of Medicine, St. Louis, Missouri 63110

(RECEIVED November 19, 1997; ACCEPTED March 4, 1998)

Abstract

Intestinal fatty acid-binding protein (I-FABP) is a cytosolic 15.1-kDa protein that appears to function in the intracellular transport and metabolic trafficking of fatty acids. It binds a single molecule of long-chain fatty acid in an enclosed cavity surrounded by two five-stranded antiparallel β-sheets and a helix-turn-helix domain. To investigate the role of the helical domain, we engineered a variant of I-FABP by deleting 17 contiguous residues and inserting a Ser-Gly linker (Kim K et al., 1996, *Biochemistry 35*:7553–7558). This variant, termed Δ17-SG, was remarkably stable, exhibited a high β-sheet content and was able to bind fatty acids with some features characteristic of the wild-type protein. In the present study, we determined the structure of the Δ17-SG/palmitate complex at atomic resolution using triple-resonance 3D NMR methods. Sequence-specific ^{1}H, ^{13}C, and ^{15}N resonance assignments were established at pH 7.2 and 25 °C and used to define the consensus $^{1}H/^{13}C$ chemical shift-derived secondary structure. Subsequently, an iterative protocol was used to identify 2,544 NOE-derived interproton distance restraints and to calculate its tertiary structure using a unique distance geometry/simulated annealing algorithm. In spite of the sizable deletion, the Δ17-SG structure exhibits a backbone conformation that is nearly superimposable with the β-sheet domain of the wild-type protein. The selective deletion of the α-helical domain creates a very large opening that connects the interior ligand-binding cavity with exterior solvent. Unlike wild-type I-FABP, fatty acid dissociation from Δ17-SG is structurally and kinetically unimpeded, and a protein conformational transition is not required. The Δ17-SG variant of I-FABP is the only wild-type or engineered member of the intracellular lipid-binding protein family whose structure lacks α-helices. Thus, Δ17-SG I-FABP constitutes a unique model system for investigating the role of the helical domain in ligand-protein recognition, protein stability and folding, lipid transfer mechanisms, and cellular function.

Keywords: intestinal fatty acid binding protein; NMR spectroscopy; protein structure

Intestinal fatty acid-binding protein or I-FABP is a water-soluble 15.1-kDa protein that binds a single molecule of long-chain fatty acid. It is abundantly expressed in the polarized epithelial cells of the small intestinal mucosa and is thought to function in the transcytoplasmic transport and trafficking of fatty acids absorbed from the diet. It belongs to a family of intracellular lipid-binding proteins that bind fatty acids, retinoids, sterols, and other amphiphilic lipids (Banaszak et al., 1994, Veerkamp & Maatman, 1995; Glatz & van der Vusse, 1996).

Members of the iLBP family vary considerably in amino acid sequence and lipid-binding properties, but exhibit the same β-clam fold. This fold consists of two five-stranded, antiparallel β-sheets that surround an interior cavity containing the bound ligand and ordered water molecules (Sacchettini & Gordon, 1993; Banaszak et al., 1994). It also contains a helix-turn-helix domain between β-strands A and B (Fig. 1). The helical domain covers one end of the binding cavity and is thought to participate in the entry, exit, and transfer of lipids (Sacchettini et al., 1989; Hodsdon et al., 1996; Hsu & Storch, 1996; Hodsdon & Cistola, 1997a, 1997b).

To investigate the role of the helical domain, we previously engineered and characterized a non-native variant of I-FABP, designated Δ17-SG (Cistola et al., 1996; Kim et al., 1996). The 17 residues forming the helix-turn-helix domain were deleted and replaced by a two-residue linker, Ser-Gly. An initial characteriza-

Reprint requests to: Dr. David P. Cistola, Washington University School of Medicine, 660 South Euclid Avenue, Campus Box 8231, St. Louis, Missouri 63110; e-mail: cistola@cosine.wustl.edu.

Abbreviations: CSI, chemical shift index; Δ17-SG, the helix-less variant of I-FABP made by deleting residues 15–31 and replacing them with a Ser-Gly linker; HNCO, correlation of amide proton and nitrogen with carbonyl carbon of previous residue; HNCACB, correlation of amide proton/nitrogen with C_α and C_β of same and previous residue; HCACO, intraresidue correlation of H_α and C_α with carbonyl carbon; HCACON, correlation of H_α and C_α, via carbonyl carbon, with amide ^{15}N of next residue; I-FABP, intestinal fatty acid binding protein; iLBP, intracellular lipid-binding protein; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser and exchange spectroscopy; RMSD, root-mean-square deviation; $RMSD_{av}$, root-mean-square deviation from the mean coordinates; TOCSY, total correlation spectroscopy.

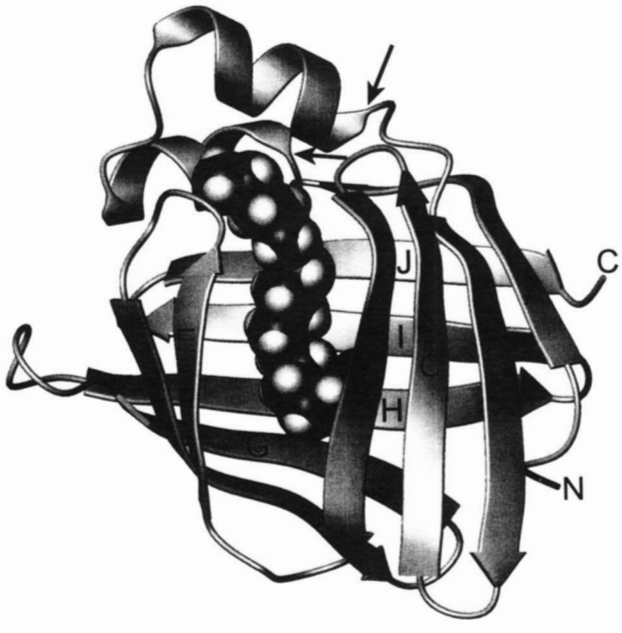

Fig. 1. Solution-state NMR structure of wild-type rat intestinal fatty acid binding protein complexed with palmitate (Hodsdon et al., 1996). The X-ray crystal structure of this complex is essentially identical (Sacchettini et al., 1989). The arrows indicate residues at the beginning and end of the deletion site in the Δ17-SG variant. The bound palmitate is shown in CPK form with its carboxy terminus behind strand D. This figure was created with MOLMOL (Koradi et al., 1996).

tion of Δ17-SG indicated that it is remarkably stable and exhibits a high β-sheet content. The refolding kinetics for Δ17-SG are nearly identical to those of the wild-type protein and imply that the helices are not a nucleation site for protein folding. The binding of fatty acids to Δ17-SG exhibits some features similar to the wild-type protein and others that are notably different. The buried pairwise interactions involving the carboxyl terminus of the fatty acid in the wild-type complex appear to be preserved in Δ17-SG, but the interactions involving the methyl terminus are absent. The kinetic progress curve for oleate association with wild-type I-FABP shows a rate-limiting step that is absent for Δ17-SG. This rate-limiting step is thought to represent an order–disorder transition involving the helical domain that allows the fatty acid access to the internal binding cavity (Cistola et al., 1996; Hodsdon & Cistola, 1997a, 1997b).

In the present study, we determined the three-dimensional structure of holo Δ17-SG in solution using triple-resonance 3D NMR methods. The backbone conformations of the wild-type and helix-less variants of I-FABP, determined using nearly identical methods and sample conditions, can now be compared at atomic resolution.

Results

Assignments and secondary structure

Sequence-specific ^{1}H, ^{13}C, and ^{15}N NMR resonance assignments were established at pH 7.2, a value previously used in the assignment of the apo and holo forms of wild-type I-FABP (Hodsdon et al., 1995, 1996; Hodsdon & Cistola, 1997b). At this pH, backbone amide protons that are not involved in well-structured regions of the protein may exchange rapidly with the protons of bulk water and become invisible in amide proton-detected NMR spectra. This exchange property has both advantages and disadvantages in the NMR analysis. Because of missing amide correlations, the resonance assignment process is more difficult at physiological pH compared with lower pH values. This difficulty can be circumvented in part by relying more heavily on ^{1}H and ^{13}C backbone/side-chain correlations (Hodsdon et al., 1995). An important advantage is that the pattern of missing amide resonances at pH 7.2 can provide physiologically relevant information about locally disordered regions of the protein backbone (Hodsdon et al., 1996; Hodsdon & Cistola, 1997a, 1997b).

For the Δ17-SG variant of I-FABP, backbone and side-chain aliphatic $^{1}H/^{13}C$ assignments were established for 102 of its 116 residues. For backbone $^{1}H/^{15}N$ correlations, 90 of the possible 116 were assigned; the remaining 26 were missing, even in gradient- and sensitivity-enhanced spectra. Therefore, these 26 residues have rapid amide or amino proton exchange rates and are likely to reside in less-structured regions of the protein backbone. Eleven of the missing 26 are in contiguous residues that straddle the deletion site (Fig. 2). An additional eight are located in turns and were also missing in gradient-enhanced spectra of the wild-type protein (Hodsdon et al., 1995, 1996; Hodsdon & Cistola, 1997b). The current resonance assignment database for Δ17-SG is provided as Supplementary Material.

Figure 2 displays the consensus chemical shift indices (CSI) for Δ17-SG and wild-type I-FABP and the resulting solution-state secondary structure. Chemical shift values are a direct output of the resonance assignment process and their interpretation is considerably less ambiguous than NOE patterns, upon which many NMR secondary structures are based. The $^{1}H/^{13}C$ consensus CSI algorithm of Wishart and Sykes emphasizes repeating patterns of chemical shifts to identify regular elements of a secondary structure with over 90% accuracy (Wishart et al., 1992; Wishart & Sykes, 1994a, 1994b). The chemical shift-derived secondary structure also serves as an important restraint on the interpretation of otherwise ambiguous NOEs used in the initial stage of tertiary structure determination (Hodsdon et al., 1996).

With the obvious exception of the helical region, the consensus chemical shift indices and missing amide profile for Δ17-SG indicate that its β-strand topology is nearly identical to that of wild-type I-FABP (Fig. 2). One notable difference is that the distal portion of β-strand A, adjacent to the deletion site, is missing and presumably unstructured in Δ17-SG.

NOE-derived tertiary structure

A total of 2,544 NOE-derived distance restraints or an average of 22 per residue were used as input for the final round of structure calculations. A summary of restraint statistics is provided in Table 1, and the distribution of NOEs per residue is shown in Figure 3. For the final ensemble, 23 structures were calculated and 3 were discarded based on penalty function values greater than two standard deviations above the mean. The average penalty function value for the 20 final structures was 0.13 ± 0.04. Values less than 10 are considered acceptable (Hodsdon et al., 1996). The low penalty value and the restraint violation statistics listed in Table 1 indicate that the structures agree well with the experimental restraints.

The stereochemical quality of the final ensemble was analyzed using PROCHECK-NMR (Laskowski et al., 1996), and the perti-

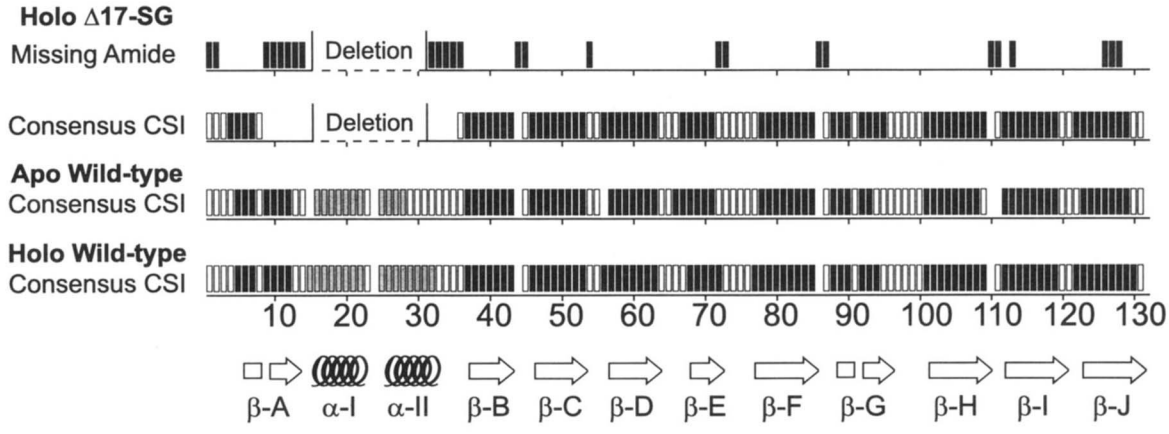

Fig. 2. Bar graph representation of chemical shift indices (CSI) for holo I-FABP (Δ17-SG). The consensus CSI plots for wild-type apo and holo I-FABP are shown for comparison. Black bars represent a chemical shift index of 1, white bars an index of 0, and light gray bars an index of −1. Gaps in the CSI plot represent residues for which there is no chemical shift information. Residues for which amide proton chemical shifts were missing are shown as black bars in the top row of the figure. The secondary structure for holo wild-type I-FABP and its nomenclature are shown at the bottom of the figure. The residues deleted from the wild-type protein for the Δ17-SG variant are also indicated.

nent structural statistics are shown in Table 2A. All of the structures have acceptable covalent geometry, as shown by small deviations from ideal bond lengths and angles. Excluding the ill-defined unassigned loop containing the deletion site, an average of 96.5% of residues were found in allowed regions of the Ramachandran plot. Approximately 40% of these are in the most favorable regions. Note that this ensemble of structures was refined using a simple penalty function without dihedral angle restraints or torsional energy terms that may otherwise restrict backbone angles to more favorable ϕ/ψ values. It is our practice to report first-generation NMR structures that are based solely on experimental NMR restraints and standard local covalent bond geometries (Hodsdon et al., 1996). At this stage, we deliberately avoid the use of general terms for van der Waals attractive energies, torsional energies, and electrostatic energies that are often used in the refinement of X-ray and NMR structures. This conservative approach is employed so that the first-generation structures will be a direct representation of the experimental NMR restraints, uncomplicated and unbiased by energy-based refinement procedures.

The precision of the ensemble, as defined by various atomic RMS differences, is summarized in Table 2B. The backbone heavy atom $RMSD_{av}$, excluding unassigned and unrestrained residues 9–20,[1] is 0.86 and 0.68 Å for β-strand residues, as defined by the consensus CSI. The value of the $RMSD_{av}$ for all heavy atoms, including side chains, is 1.34 Å. Stereospecific assignments have not yet been established for this protein. Therefore, the side chains are less well defined than the backbone because of the large upper bound corrections used for prochiral side-chain atoms (Hodsdon et al., 1996).

Table 1. *Final restraint statistics for I-FABP (Δ17-SG)*[a]

Distance restraints	
Total	2,544 (21.9/residue or 24.5/residue[b])
Intraresidue	590
Sequential (i, $i \pm 1$)	598
Medium range ($\|i - j\| \leq 3$)	270
Long range ($\|i - j\| > 3$)	1,086
Restraint violations	
Upper bounds	
Average violation (Å)	0.08 ± 0.00
Largest violation (Å)	0.08
No. of violations	1 of 50,880[c] (0.002%)
Lower bounds	
Average violation (Å)	0.07 ± 0.05
Largest violation (Å)	−0.18
No. of violations	62 of 50,880[c] (0.12%)

[a]Analyzed using AQUA v1.2.1 (Laskowski et al., 1996).

[b]The latter number excludes residues 9–20 in the unassigned, unrestrained deletion-site loop.

[c]The product of the number of restraints and the number of structures in the final ensemble.

[1]This numbering corresponds to that of helix-less I-FABP (Δ17-SG). The equivalent residues in the wild-type I-FABP are 9–14 and 32–35.

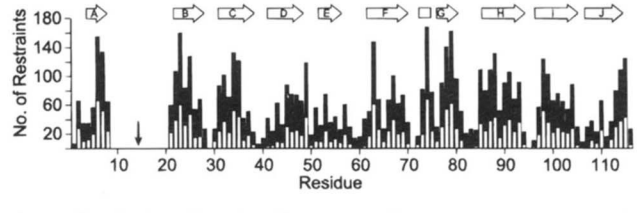

Fig. 3. Distribution of total and long-range distance restraints used for the calculation of the final ensemble of structures for I-FABP (Δ17-SG). The full height of each bar represents the total number of restraints for each residue. The height of the unfilled portion represents the number of long range restraints, defined as $|i - j| > 3$, where i and j are residue numbers. In this figure, interresidue restraints are counted twice, once for each residue. The arrow indicates the site of the deletion of the helices from I-FABP.

Table 2. *Structural and stereochemical statistics for I-FABP (Δ17-SG)*

A. Structural statistics[a]	
RMSD from ideal bond lengths (Å)	0.017 ± 0.002
RMSD from ideal bond angles (deg)	3.94 ± 0.49
	Ramachandran plot statistics[b]
Residues in allowed regions	100(96.5%)[c]
Most favored regions	42(40.4%)
Additionally allowed regions	47(45.1%)
Generously allowed regions	11(11.0%)
Residues in disallowed regions	4 (3.5%)

B. Atomic RMS differences[d]

	Average backbone heavy atom RMSD		Average all heavy atom RMSD	
	To mean	Pairwise	To mean	Pairwise
All residues	1.66 ± 0.51	2.41 ± 0.68	2.21 ± 0.45	3.20 ± 0.62
All residues excl. 9–20[b]	0.86 ± 0.12	1.25 ± 0.17	1.34 ± 0.14	1.94 ± 0.19
β-sheet[e]	0.68 ± 0.13	0.99 ± 0.16	1.19 ± 0.16	1.72 ± 0.21

[a]Analyzed using PROCHECK-NMR version 3.4.4 (Laskowski et al., 1996).
[b]For all residues excluding residues 9–20, which are in the unrestrained deletion-site loop.
[c]Based on a total of 104 residues.
[d]Analyzed using MOLMOL version 3.4 (Koradi et al., 1996).
[e]Calculated for residues 4–7, 22–28, 31–38, 41–48, 52–56, 63–70, 73–80, 86–94, 97–104, and 107–114, which are defined as β-sheet residues by the consensus CSI.

Collectively, the agreement of the structures with the experimental restraints, the number of restraints per residue, the stereochemical quality of the structures, and the precision of the ensemble indicate that the final structure is of sufficient quality to compare its backbone conformation to that of the wild-type protein.

A stereo diagram of the structure ensemble is shown in Figure 4. Except for the deletion-site loop, the protein adopts an all β-sheet structure. The first antiparallel β-sheet consists of strands A–E, the vertically oriented strands at the front face of the molecule. The second β-sheet, consisting of strands F–J, is represented by the horizontally oriented strands at the back face of the molecule. The ligand-binding cavity between the two β-sheets is preserved. Overall, the β-sheet domain is well defined and closely resembles that of the wild-type protein. The unassigned, unrestrained resi-

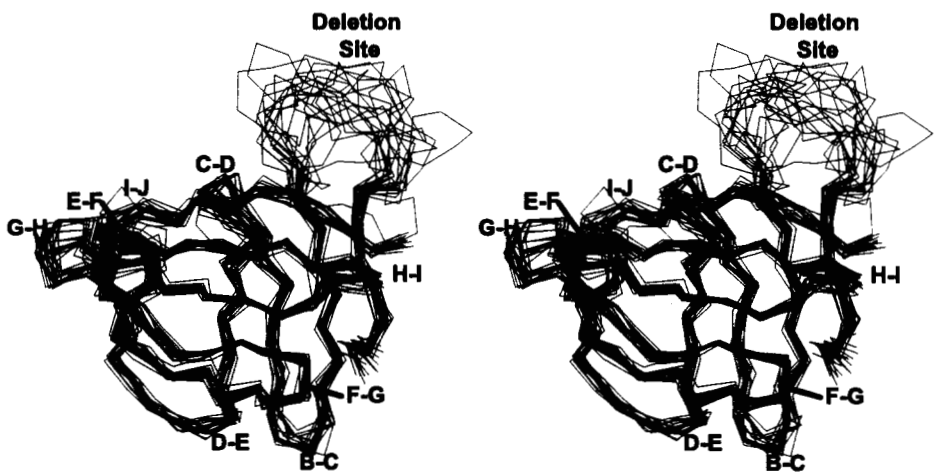

Fig. 4. Stereo diagram of 20 superimposed backbone C_α traces for I-FABP (Δ17-SG) complexed with palmitate. The turns between β-strands are labeled according to the flanking strands. The large ill-defined loop labeled "deletion site" is where the helix-turn-helix motif was deleted from the wild-type I-FABP. The structures were overlaid, omitting residues 9–20, which comprise this disordered loop. The complexed palmitate, although present in the sample, was not included in the calculations and is not shown.

dues 9–20[1] form an ill-defined loop that spans the deletion site and connects β-strands A and B. This loop is likely to be relatively unstructured in solution. The Δ17-SG I-FABP structure also contains a large opening that connects the interior ligand binding cavity with exterior solvent, as discussed below.

Discussion

A helix-less variant of rat I-FABP was engineered by deleting 17 contiguous residues spanning the helix-turn-helix segment and inserting a two-residue linker, Ser-Gly. Initially, the overall structure content of Δ17-SG was assessed by comparing its CD and NMR spectra with those of the wild-type protein (Cistola et al., 1996; Kim et al., 1996). In the current study, we established sequence-specific NMR resonance assignments and determined the three-dimensional structure of Δ17-SG in solution. Its backbone conformation is now defined at atomic resolution and can be compared with that of wild-type I-FABP.

A superposition of the average C_α traces of the wild-type and Δ17-SG structures is shown in Figure 5. The structure of Δ17-SG exhibits topological features remarkably similar to those of wild-type I-FABP. These features include the two orthogonally related, five-stranded, antiparallel β-sheets. They also include the widened "gap" between β-strands D and E. Like the wild-type protein, this gap does not constitute an actual opening into the interior and is largely occupied by side chains. Moreover, Δ17-SG maintains a ligand binding cavity between the two β-sheets. We previously demonstrated that Δ17-SG binds one molecule of palmitate or oleate with some features characteristic of the wild-type protein (Cistola et al., 1996).

Although the Δ17-SG and wild-type structures are similar overall, there are two notable differences. The first difference is the ill-defined loop flanking the deletion site of Δ17-SG. This loop replaces the helix-turn-helix segment in the wild-type structure. Because resonance assignments could not be established for the loop, it was not possible to identify chemical shift or NOE-based restraints that would further define its conformation. Likewise, it was not possible to characterize the dynamic properties of the loop using backbone relaxation and quantitative saturation transfer experiments. However, the selective absence of amide resonances, even in gradient- and sensitivity-enhanced HSQC experiments, implies that the backbone amide protons in the loop are in rapid exchange with bulk water (lifetimes ≤ ms) and not involved in hydrogen bonds. Because of the lack of conformational restraints, the backbone traces of the deletion-site loop in Figure 4 should not be taken as literal representations of the conformation of this segment of the molecule.

The loop, which spans residues 9–20,[1] was longer than anticipated. The Δ17-SG variant was engineered to contain only a small linker to connect the newly formed ends of β-strands A and B. In the wild-type structure, β-strand A is kinked and effectively forms two segments (Figs. 1, 2). The first segment of β-strand A hydrogen bonds with strand B, and the second, with strand J. The interstrand hydrogen bonds between strands A and J are absent in the Δ17-SG structure. Thus, the distal half of β-strand A in the wild-type structure corresponds to the proximal part of the disordered loop in Δ17-SG. Residues 17–20 correspond to 32–35 in the wild-type protein. These residues have long-range interactions that assist in the C-terminal capping of helix II (see below). The absence of these interactions in the helix-less variant may also contribute to the length of the disordered loop. It may be possible to eliminate most or all of the unstructured loop without introducing other significant perturbations. For example, a construct containing a deletion of wild-type residues 9–35 and an insertion of S-G (Δ25-SG) could conceivably yield a second-generation helix-less variant that is more compact and stable than Δ17-SG.

The second, and most striking, difference between the structures of Δ17-SG and wild-type I-FABP is the large opening that connects the interior binding cavity with exterior solvent. Figure 6 displays solvent-accessible surface maps of two NMR structures selected from the ensembles for holo wild-type and holo helix-less I-FABP. The structures are oriented in an identical fashion such that the viewer is looking directly toward the helical domain, colored magenta in the left panel of Figure 6. In the wild-type structure, the interior cavity is covered by the helical domain and is inaccessible to solvent. The cavity is hidden from view in the left panel of Figure 6. In contrast, the interior cavity is solvent-accessible and clearly visible in the Δ17-SG structure. The yellow portion in the middle of the structure represents the β-sheet surface at the "bottom" (far end) of the interior cavity. The ill-defined

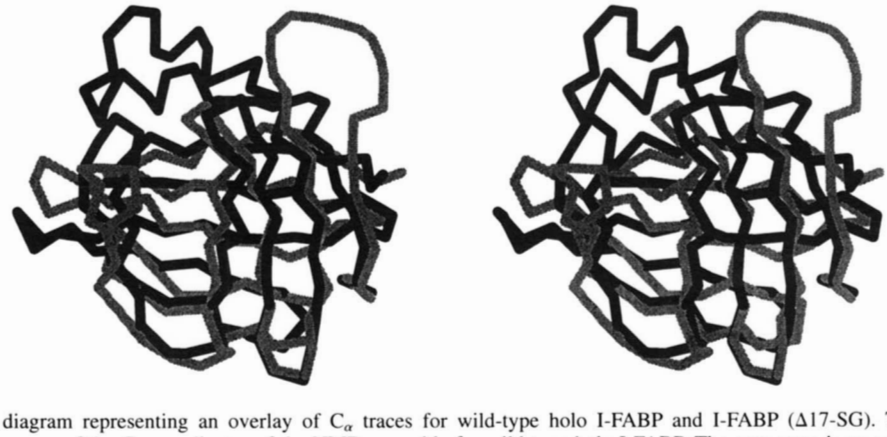

Fig. 5. Stereo diagram representing an overlay of C_α traces for wild-type holo I-FABP and I-FABP (Δ17-SG). The black trace represents an average of the C_α coordinates of the NMR ensemble for wild-type holo I-FABP. The gray trace is an average of the C_α coordinates of the NMR ensemble for I-FABP (Δ17-SG). The structures are superposed using the C_α atoms for residues that are common to both structures (1–8 and 21–116 in Δ17-SG or 1–8 and 36–131 in wild type).

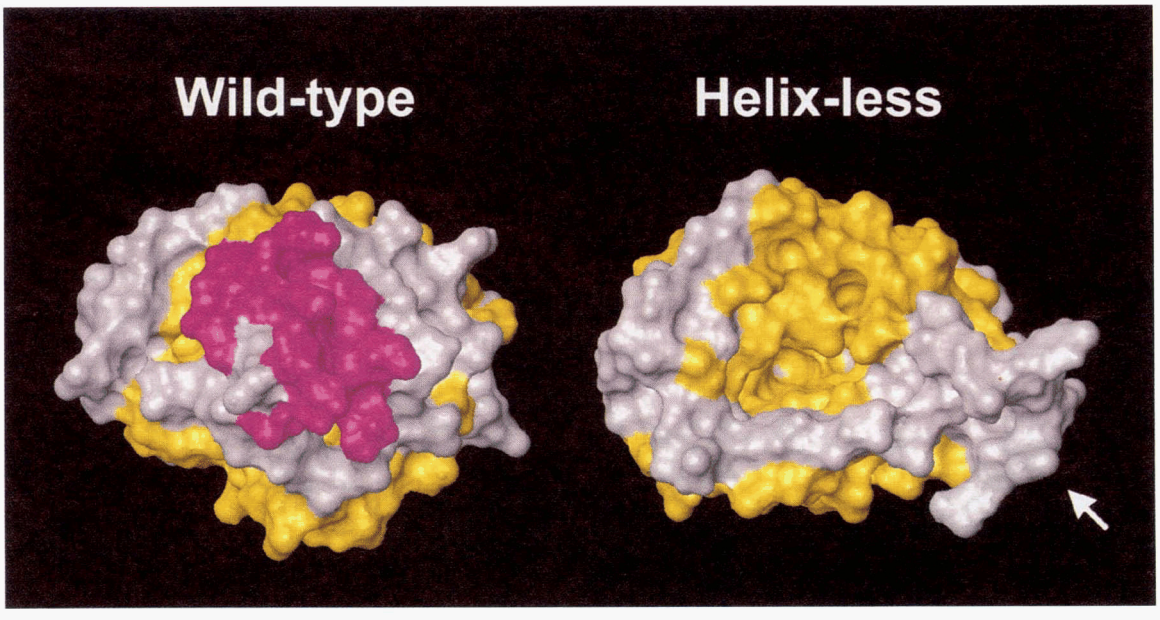

Fig. 6. Solvent-accessible surface representations of wild-type and helix-less I-FABP NMR structures. The view is from directly above the helices and proposed portal entrance. The surface is colored by residue type; yellow, β-sheet residues; magenta, helical residues; gray, all other residues. This figure was made with the program MOLMOL v2.4, using a Lee and Richards contact surface (Lee & Richards, 1971) with a solvent molecule of radius 1.4 Å. The arrow indicates the deletion-site loop.

loop, designated by the arrow, is not nearly large enough to cover the opening to the cavity, even if this loop were to fold down on the β-sheet domain.

The overall shape of the wild-type structure has been described as a clam shell. In contrast, the overall shape of the Δ17-SG structure is more like a tea cup, with one open end.

The open-ended, solvent-accessible cavity in Δ17-SG helps explain its unique fatty acid-binding properties (Cistola et al., 1996). Isotope-directed 2D NMR studies using palmitate selectively ^{13}C-enriched at each end of the molecule indicated that the interactions involving the carboxyl terminus of the fatty acid are preserved in the helix-less variant. In contrast, the interactions at the methyl terminus, adjacent to the helical domain in wild-type I-FABP, are absent in Δ17-SG. Moreover, kinetic progress curves for oleate association exhibit a rate-limiting plateau for wild-type I-FABP that is absent for Δ17-SG. This plateau was interpreted as a conformational transition, involving the helical domain, that is necessary for fatty acid entry into the interior cavity. Because of the large opening, the entry and exit of fatty acid from Δ17-SG is unimpeded, and a protein conformational transition is not required.

A structural and dynamical characterization of the apo and holo forms of wild-type I-FABP in solution has led to a revised mechanism of fatty acid entry and exit known as the "dynamic portal hypothesis" (Hodsdon et al., 1996; Hodsdon & Cistola, 1997a, 1997b). According to this hypothesis, the region encompassed by the distal portion of helix II and the C–D and E–F turns regulates the entry and exit of fatty acid by undergoing a backbone order–disorder transition in solution. The fatty acid enters through this locally disordered region, the dynamic portal. Fatty acid binding shifts the equilibrium toward the ordered state and closes the dynamic portal by stabilizing a series of interactions resembling a helix-capping box (Hodsdon & Cistola, 1997a). In the absence of fatty acid, these interactions are destabilized and helix II becomes decoupled from the C–D turn and frayed at its C-terminal end. The release of fatty acid may be catalyzed by any process that destabilizes the capping box interactions, such as collisional interactions between the fatty acid/I-FABP complex and target membranes (Hsu & Storch, 1996).

Comparison of the results for Δ17-SG and wild-type I-FABP constitutes a test of the dynamic portal hypothesis. Removal of the proposed structural barrier to fatty acid entry (the helix-turn-helix domain) correlates with the loss of a kinetic barrier to fatty acid binding (the rate-limiting step in fatty acid association). However, it is not necessary to delete or disrupt the entire helical domain to eliminate these barriers. Kinetic progress curves for the triple mutant L30K/N54G/F55G are linear like those for Δ17-SG (J. Monsey & D.P. Cistola, unpubl. obs.). This triple mutant was designed to decouple the specific side-chain interactions between helix II and the C–D turn and shift the dynamic portal equilibrium toward the disordered or open state(s).

The results from the present study firmly establish that Δ17-SG is a helix-less, all-β-sheet variant of I-FABP and define its three-dimensional structure in solution. The Δ17-SG structure will provide a useful reference point for the design and engineering of additional lipid-binding protein variants. It is now clear that the helical domain is not required to stabilize the β-sheet domain of I-FABP; the same may be true of other lipid-binding proteins. Nor does the helical domain serve as a nucleation site for protein folding (Kim et al., 1996). Rather, it appears to play a regulatory role in ligand entry, exit, and transfer (Cistola et al., 1996; Herr et al., 1996; Hsu & Storch, 1996). Helix-less I-FABP, with its defined structure, provides a unique and powerful model system for molecular and cellular studies aimed at elucidating the mechanisms of lipid binding, transfer, and trafficking.

Materials and methods

Sample preparation

The Δ17-SG variant of rat I-FABP (accession number M35992) was engineered by replacing residues E15 through G31 of the wild-type protein with a S-G linker (Kim et al., 1996). Uniformly $^{13}C/^{15}N$-enriched I-FABP (Δ17-SG) was bio-synthesized and purified using protocols essentially identical to those used for wild-type I-FABP (Hodsdon et al., 1995). The final yield of purified isotope-enriched protein was approximately 90 mg, about half that of a comparable preparation of wild-type protein. The purified, delipidated protein was concentrated to 2 mM in a buffer containing 20 mM potassium phosphate, 50 mM KCl, and 0.05% sodium azide at pH 7.2. The concentrated sample remained optically clear with no visible sign of aggregation. This sample was complexed with a stoichiometric amount of perdeuterated palmitate (Cistola et al., 1989) and divided into two equal aliquots of 1.1 mL each. Both aliquots were lyophilized; one was redissolved in 1.1 mL of 80% H_2O/20% D_2O, and the other, 99.996% D_2O. Each sample was equally divided and aliquoted into two Wilmad 545-PPT NMR tubes, yielding a total of two "H_2O" and two "D_2O" samples.

Samples of unenriched protein complexed with ^{13}C-enriched palmitate were also prepared (Cistola et al., 1996).

NMR spectroscopy

The NMR spectra were acquired, processed, and referenced as described previously (Hodsdon et al., 1995, 1996; Hodsdon & Cistola, 1997b). All spectra were acquired at 25 °C. A summary of the acquisition and processing parameters is provided as Supplementary Material.

Resonance assignments and chemical shift-derived secondary structure

The ^1H, ^{13}C, and ^{15}N NMR resonances of the protein backbone and side-chain atoms were assigned using a strategy similar to that employed for wild-type I-FABP (Hodsdon et al., 1995). Six three-dimensional experiments were used in a concerted fashion: HNCO, g-HNCACB, ^{15}N-TOCSY, HCACO, HCACON, and HCCH-TOCSY. The g-HNCACB (Wittekind & Mueller, 1993; Muhandriam & Kay, 1994) data set was substituted for the HNCA spectrum used previously. This substitution enhanced the assignment process by providing C_β chemical shifts and correlations that assisted in the initial identification of amino acid type (Grzesiek & Bax, 1993). The resonance assignment process was also facilitated by referring to the chemical shift database for wild-type I-FABP (Hodsdon et al., 1995, 1996).

NOE-derived distance restraints and structure calculations

Interproton distance restraints were obtained largely from 3D ^{15}N- and ^{13}C-resolved NOESY spectra. All restraints were interpreted conservatively with an upper bound of 5.0 Å (plus prochiral corrections, if appropriate) to achieve increased accuracy at the expense of precision (Hodsdon et al., 1996). As observed for wild-type I-FABP, a systematic search of the assignment database revealed that nearly all NOE peaks had multiple possible interpretations or assignments for one proton in each pair. Therefore, an iterative procedure was required to establish the identity of the protons involved in pairwise NOE-based distance restraints (Hodsdon et al., 1996). The first step of this procedure relied critically on the consensus $^1H/^{13}C$ chemical shift-derived secondary structure as a starting model and a set of unambiguous distance restraints derived from a symmetry check of the 3D ^{13}C- and ^{15}N-resolved NOESY cross peaks.

The first set of 1,314 interproton distance restraints was sufficient to calculate a somewhat imprecise, but accurate set of structures that established the overall fold of the protein. This initial set of structures was then used to manually reassign the entire NOE peak list and rule out all interpretations that were clearly inconsistent with the overall fold. This second set of restraints served as input for a second round of structure calculations. Subsequent cycles of cross-peak assignments were performed automatically using in-house programs (Hodsdon et al., 1996). Distance restraints involving aromatic protons were identified from 2D ^1H NOESY data collected using an unenriched sample and added to the middle stages of the iterative procedure. No hydrogen bond restraints were inferred or assumed at any stage of the calculations.

Attempts to include the bound palmitate molecule in the structure calculations were unsuccessful. Because of the 17-residue deletion, there were no observable conformational restraints involving the methyl terminus and middle of the fatty acid chain (Cistola et al., 1996). The robust sampling properties of the distance geometry/simulated annealing algorithm led the relatively unrestrained fatty acid hydrocarbon chain to sample a large region of conformational space, resulting in distortions to the protein fold and poor agreement with the experimental restraints. Therefore, the palmitate molecule was omitted from the structure calculations, even though it was present in the NMR samples used for structure determination. For wild-type I-FABP, it was possible to define the conformation of the bound palmitate, as shown in Figure 5 of Hodsdon et al. (1996). Using selectively ^{13}C-enriched palmitate and unenriched Δ17-SG, we previously observed that the interactions involving the carboxyl terminus of the fatty acid in the wild-type complex were preserved in the Δ17-SG complex (Cistola et al., 1996).

Structure calculations were performed on a Silicon Graphics INDY/R4400 workstation using DISTGEOM,[2] a distance geometry/simulated annealing program, that is part of the TINKER[2] protein modeling package. This program uses a novel pairwise Gaussian metrization embedding algorithm that has excellent sampling and convergence properties. The unique features of this program and the calculation and refinement protocols have been discussed elsewhere (Hodsdon et al., 1996). For each round of calculations, approximately 20 structures were generated. Outlier structures with penalty function values greater than two standard deviations above the mean were infrequently observed. Because this embedding algorithm has excellent convergence properties, there is no need to calculate and discard a large number of structures (Hodsdon et al., 1996).

Supplementary material in electronic appendix

Three tables are provided containing the chemical shift resonance assignments for holo Δ17-SG I-FABP and a summary of the ac-

[2]Information about the DISTGEOM program and the TINKER molecular modeling package and instructions for downloading can be obtained at the following web site: http://dasher.wustl.edu/ (128.252.162.151).

quisition and processing parameters for the NMR experiments (suppl.doc).

Acknowledgments

This work was supported by National Institutes of Health grants DK48046 (to D.P.C.) and DK13332 (to C.F.). The authors thank Drs. Chang-guo Tang for NMR assistance and James Toner for the biosynthesis and purification of the $^{13}C/^{15}N$-enriched protein. Dr. Keeyhuk Kim engineered Δ17-SG and performed some of its initial characterization (Cistola et al., 1996, Kim et al., 1996). The DISTGEOM/TINKER software was generously provided by Dr. Jay Ponder.

The molecular coordinates for the ensemble of NMR structures for Δ17-SG I-FABP have been deposited in the Brookhaven Protein Data Bank with ID number 1a57.

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.

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.

Cistola DP, Sacchettini JC, Banaszak LJ, Walsh MT, Gordon JI. 1989. Fatty acid interactions with rat intestinal and liver fatty acid-binding proteins expressed in *Escherichia coli*. A comparative ^{13}C NMR study. *J Biol Chem 264*:2700–2710.

Glatz JF, van der Vusse GJ. 1996. Cellular fatty acid-binding proteins: Their function and physiological significance. *Prog Lipid Res 35*:243–282.

Grzesiek S, Bax A. 1993. Amino acid type determination in the sequential assignment procedure of uniformly $^{13}C/^{15}N$-enriched proteins. *J Biomol NMR 3*:185–204.

Herr FM, Aronson J, Storch J. 1996. Role of portal region lysine residues in electrostatic interactions between heart fatty acid binding protein and phospholipid membranes. *Biochemistry 35*:1296–1303.

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.

Hodsdon ME, Toner JJ, Cistola DP. 1995. ^{1}H, ^{13}C, and ^{15}N assignments and chemical shift-derived secondary structure of intestinal fatty acid-binding protein. *J Biomol NMR 6*:198–210.

Hsu KT, Storch J. 1996. Fatty acid transfer from liver and intestinal fatty acid-binding proteins to membranes occurs by different mechanisms. *J Biol Chem 271*:13317–13323.

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.

Koradi R, Billeter M, Wuthrich K. 1996. MOLMOL: A program for display and analysis of macromolecular structures. *J Mol Graph 14*:51–55, 29–32.

Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. 1996. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *J Biomol NMR 8*:477–486.

Lee B, Richards FM. 1971. The interpretation of protein structures: Estimation of static accessibility. *J Mol Biol 55*:379–400.

Muhandriam DR, Kay LE. 1994. Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity. *J Magn Reson B 103*:203–216.

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. *J Biol Chem 268*:18399–18402.

Sacchettini JC, Gordon JI, Banaszak LJ. 1989. Refined apoprotein structure of rat intestinal fatty acid binding protein produced in *Escherichia coli*. *Proc Natl Acad Sci USA 86*:7736–7740.

Veerkamp JH, Maatman RG. 1995. Cytoplasmic fatty acid-binding proteins: Their structure and genes. *Prog Lipid Res 34*:17–52.

Wishart DS, Sykes BD. 1994a. The ^{13}C chemical-shift index: A simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J Biomol NMR 4*:171–180.

Wishart DS, Sykes BD. 1994b. Chemical shifts as a tool for structure determination. *Methods Enzymol 239*:363–392.

Wishart DS, Sykes BD, Richards FM. 1992. The chemical shift index: A fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry 31*:1647–1651.

Wittekind M, Mueller L. 1993. HNCACB, a high sensitivity 3D NMR experiment to correlate amide-proton and nitrogen resonances with the alpha- and beta-carbon resonances in proteins. *J Magn Reson B 101*:201–205.