J. Mol. Biol. (1996) **264**, 585–602

JMB

The NMR Solution Structure of Intestinal Fatty Acid-binding Protein Complexed with Palmitate: Application of a Novel Distance Geometry Algorithm

Michael E. Hodsdon, Jay W. Ponder and David P. Cistola*

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

The three-dimensional solution structure of rat intestinal fatty acid-binding protein (I-FABP) complexed with palmitate has been determined using multidimensional triple-resonance NMR methods. The structure is based on 3889 conformational restraints derived mostly from 3-D ^{13}C- and ^{15}N-resolved nuclear Overhauser (NOESY) experiments. The 3-D NOESY data for this 15.4 kDa complex contained an average of nine possible interpretations per cross-peak. To circumvent this ambiguity, an eight-stage iterative procedure was employed to gradually interpret and introduce unambiguous distance restraints during subsequent rounds of structure calculations. The first stage of this procedure relied critically upon an initial structural model based on the consensus ^{1}H/^{13}C chemical shift-derived secondary structure and a set of symmetry-checked restraints derived from the 3-D ^{13}C-resolved NOESY spectrum. The structures were calculated using DISTGEOM, a program that implements a novel distance geometry algorithm with pairwise Gaussian metrization. A central feature of this algorithm is the use of an iteratively optimized Gaussian distribution for the selection of trial distances, which overcomes the tendency of metrization to produce crushed structures. In addition, this algorithm randomly selects pairwise elements of the distance matrix, which results in an improved sampling of conformational space for a given computational effort. The final family of 20 distance geometry/simulated annealing structures exhibited an average pairwise C^{α} root-mean-square deviation of 0.98 Å, and their stereochemical quality, as assessed by PROCHECK, was comparable to that of 2.5 Å X-ray crystal structures. The NMR structure was compared with the X-ray crystal structure of the same ligand/protein complex and was found to be essentially identical within the precision of the results. The NMR structure was also compared with that of the palmitate complex with bovine heart FABP, which shares 30% sequence identity with rat I-FABP. The overall folds were the same, but differences were noted with respect to the presence or absence of apparent conformational heterogeneity and the location and conformation of the bound fatty acid.

Keywords: intestinal fatty acid-binding protein; NMR spectroscopy; protein structure; distance geometry; lipid transport and trafficking

**Corresponding author*

Abbreviations used: iLBPs, intracellular lipid-binding proteins; I-FABP, intestinal fatty acid-binding protein; H-FABP, heart-muscle fatty acid-binding protein; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser and exchange spectroscopy; HSQC, heteronuclear single-quantum correlation spectroscopy; HMQC, heteronuclear multiple-quantum correlation spectroscopy; TINKER, a program package that includes the distance geometry program DISTGEOM employed in the present study; RMSD, root-mean-square deviation; DDM, difference distance matrix; $O(N)^2/O(N)^3$, to the order of the number of atoms squared or cubed; 3-D, three-dimensional.

0022–2836/96/480585–18 $25.00/0

Introduction

The intracellular lipid-binding proteins (iLBPs) are a family of soluble, predominantly cytoplasmic proteins that bind polar lipids such as fatty acids, retinoids and sterols (Ockner *et al*., 1972; Banaszak *et al*., 1994; Veerkamp & Maatman, 1995). The iLBPs appear to function in the uptake, transport, trafficking and sequestration of cellular lipids (Sweetser *et al*., 1987; Matarese *et al*., 1989; Glatz and van der Vusse, 1989, 1990; Kaikaus *et al*., 1990; Veerkamp *et al*., 1991; Bass, 1993). Some iLBPs may also be directly or indirectly involved in the regulation of gene expression, cell growth and differentiation in conjunction with the retinoid-X, retinoic acid and peroxisome-proliferator-activated nuclear receptors (Kaikaus *et al*., 1990; Bass, 1993; Veerkamp & Maatman, 1995).

An individual mammalian organism expresses at least nine distinct iLBP genes and 13 distinct proteins, ranging in amino acid sequence identity from approximately 20 to 80% (Banaszak *et al*., 1994; Veerkamp & Maatman, 1995). Some cell types, such as adipocytes, hepatocytes, and cardiac and skeletal myocytes, appear to express a single, dominant iLBP variant. However, the small intestinal enterocyte abundantly expresses four distinct iLBP variants. These four proteins are termed intestinal fatty acid-binding protein, cellular retinol-binding protein-II, ileal lipid-binding protein and liver fatty acid-binding protein and have different ligand specificities, stoichiometries and affinities (Glatz *et al*., 1993; Sacchettini & Gordon, 1993; Banaszak *et al*., 1994). The molecular basis and functional significance of these differences remain unclear.

We are using the four iLBPs expressed in the small intestine as a model system for investigating the basis of molecular recognition in the iLBP family. Starting with I-FABP, we are developing a database of sequence-specific ^{1}H, ^{13}C and ^{15}N NMR resonance assignments, which provide the basis for characterizing and comparing the structural and dynamical properties of these proteins in solution at atomic resolution. We recently reported the resonance assignments and chemical shift-derived secondary structure for the 15.4 kDa complex between rat I-FABP and palmitate (Hodsdon *et al*., 1995). The assignments and secondary structure were determined entirely on the basis of chemical shifts and scalar couplings with no reliance upon NOE data, which can be difficult to interpret for macromolecular complexes of this size.

In the present study, we employed the resonance assignments and chemical shift-derived secondary structure, along with 3-D ^{13}C- and ^{15}N-resolved NOESY experiments, to solve the tertiary structure of holo I-FABP in solution. The results permitted a detailed assessment of the protein conformation in solution at atomic resolution and were compared with the X-ray crystal structure for the same complex (Sacchettini *et al*., 1989). The results were also compared with related protein complexes, including the NMR structure of palmitate complexed with bovine heart fatty acid-binding protein (Lassen *et al*., 1995), which shares 30% sequence identity with rat I-FABP. Finally, this study provided an opportunity to evaluate and improve protocols for the determination of NMR structures of larger proteins, particularly with respect to the sampling and convergence properties of distance geometry.

Results and Discussion

Resonance assignments and NOESY data

The ^{1}H, ^{13}C, and ^{15}N resonance assignments for *Escherichia coli*-derived rat I-FABP complexed with palmitate at pH 7.2 and 37°C were previously reported (Hodsdon *et al*., 1995). The current database includes complete backbone and side-chain aliphatic ^{1}H and ^{13}C assignments for 127 of 131 residues and backbone amide $^{1}H/^{15}N$ assignments for 119 residues. The remaining 12 amide protons were not observed, even in gradient-enhanced experiments without presaturation, because of rapid exchange with solvent under these sample conditions. In the present study, a number of aromatic ring ^{1}H and fatty acid $^{1}H/^{13}C$ assignments were established.

Most of the distance restraints used to determine the tertiary structure of the protein were derived from 3-D ^{13}C- and ^{15}N-resolved NMR spectra collected on samples of uniformly [88%-^{13}C, 99%-^{15}N]-enriched I-FABP complexed with perdeuterated palmitate (Hodsdon *et al*., 1995). Three individual ^{15}N planes from the 3-D ^{15}N-resolved NOESY spectrum are shown in Figure 1(A). These planes illustrate NOE correlations involving the backbone amides of G65, I23, A69 and F62. Figure 1(B) illustrates three individual ^{13}C planes from the 3-D ^{13}C-resolved NOESY spectrum, illustrating NOE correlations involving the side-chain carbon atoms / protons of L89δ, I108δ, I103α and T79α. Distance restraints involving the bound palmitate were derived from 2-D ^{13}C-NOESY-HSQC spectra collected for samples of unenriched I-FABP complexed with selectively and uniformly ^{13}C-enriched palmitate. These spectra were similar to those shown in Figures 5 and 6 of Cistola *et al*. (1996).

Derivation and description of the final set of structural restraints

A combined total of 5313 NOE cross-peaks were identified in the 3-D ^{13}C- and ^{15}N-resolved NOESY spectra. The chemical shift of the scalar-coupled ^{13}C or ^{15}N nucleus, provided by the third frequency dimension of the spectrum, established the interpretation of one of the two dipolar-coupled protons represented by each NOESY cross-peak. However, the interpretation of the ^{1}H chemical shift for the remaining member of the proton pair was

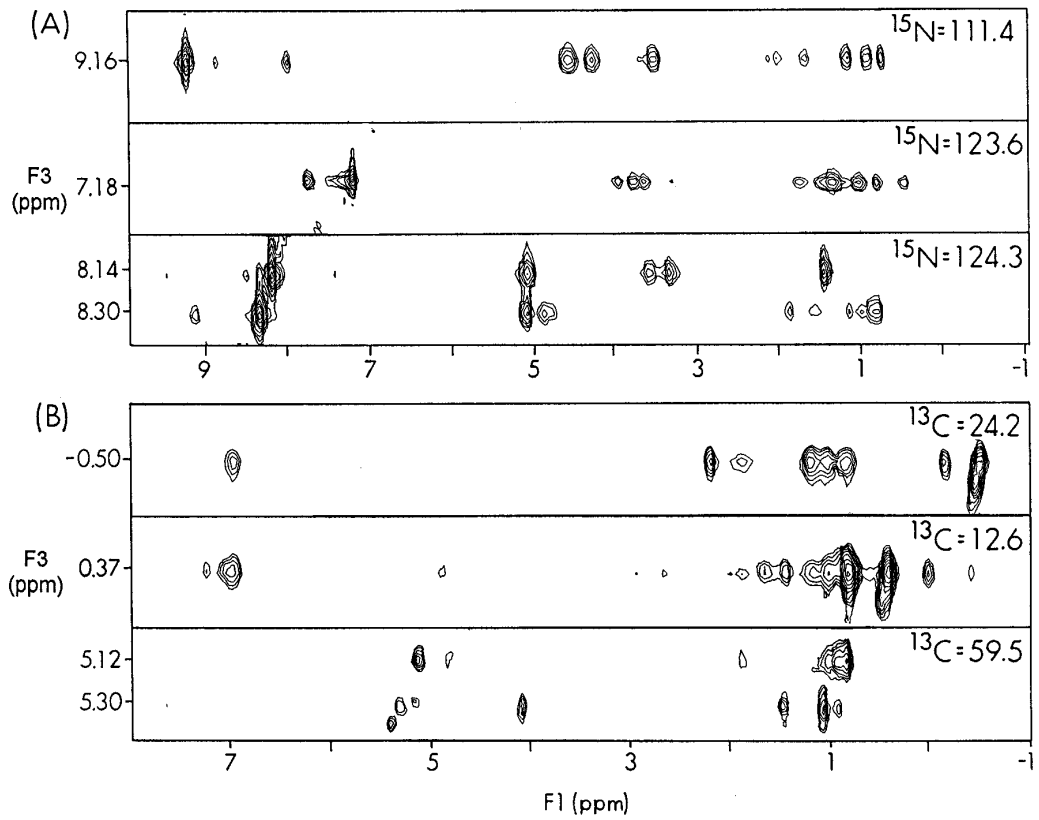

Figure 1. Selected strips from F1 / F3 (^{1}H / ^{1}H) planes of the (A) 3-D ^{15}N-NOESY-HMQC and (B) 3-D ^{13}C-NOESY-HSQC spectra of uniformly ^{13}C / ^{15}N-enriched I-FABP complexed with 3 mM palmitate (pH 7.2) at 37°C. The chemical shift of the *F*2 dimension is given in the upper right-hand corner of each plane. (A) NOE correlations involving the backbone amide ^{1}H / ^{15}N atoms for residues G65 (9.16 ppm), I23 (7.18 ppm), A69 (8.14 ppm) and F62 (8.30 ppm). (B) Correlations involving side-chain and backbone ^{1}H / ^{13}C atoms for L89δ (–0.50 ppm), I108δ (0.37 ppm), I103α (5.12 ppm) and T79α (5.30 ppm).

ambiguous in nearly all cases. This ambiguity resulted from proton chemical shift degeneracy rather than insufficient resolution. A complete search of our ^{1}H assignment database revealed an average of nine possible interpretations for each partially assigned 3-D NOESY cross-peak. Resolving the ambiguities in interpreting or ''assigning'' the NOESY cross-peaks represented a significant challenge in the determination of the structure of this 15.4 kDa complex. Four-dimensional CN- and CC-NOESY experiments (Kay *et al*., 1990; Clore *et al*., 1991) can be used to reduce such ambiguity; however, initial attempts at collecting 4-D data for this fatty acid-protein complex were unsuccessful. Alternatively, a non-NMR-derived structural model, such as an X-ray crystal structure or homology model, could have been explicitly or implicitly used to assign ambiguous NOE cross-peaks. However, this approach can introduce bias into the final family of structures by weighting the structures toward the model. Therefore, we employed an iterative procedure based solely on the use of experimental NMR data in order to achieve an independently determined structure of holo-I-FABP.

Various iterative approaches have been employed to assign NOESY data (e.g. Kraulis *et al*., 1989; Clore *et al*., 1990; Forman-Kay *et al*., 1991; Güntert *et al*., 1993; Meadows *et al*., 1994). Generally, a subset of unambiguously assigned NOE cross-peaks is used to calculate an initial ensemble of NMR structures, which is subsequently employed in an iterative manner to interpret the remaining NOE cross-peaks. This approach is successful if there is a sufficient number of unambiguous NOEs to calculate an initial set of structures with the proper fold and if the initial structures adequately sample the conformational space consistent with the full data set (Nilges, 1995). However, the 3-D NOESY data for this 15.4 kDa complex contained very few cross-peaks that could be assigned unambiguously, too few to calculate an initial ensemble of structures. Therefore, the first step of our iterative procedure relied critically on the development of an initial NMR-based structural model derived from the combined use of the ^{1}H / ^{13}C consensus chemical shift index and a subset of symmetry-checked, ''pseudo-4-D'' restraints derived from the 3-D ^{13}C-resolved NOESY experiment. Knowledge of the chemical shift-derived secondary structure of I-FABP, determined unambiguously without the use of NOEs (Hodsdon *et al*.,

Table 1. Final set of conformational restraints

Type	Number
Total distance restraints	3889 (30 restraints/residue)
Intra-protein distance restraints	
Intraresidue	708
Sequential ($i, i \pm 1$)	965
Medium-range ($i, i \pm 2, i \pm 3, i \pm 4$)	536
Long-Range	1622
Intra-ligand distance restraints	5
Ligand-protein distance restraints	53
Torsional restraints	
Helical Φ and Ψ restraints[a]	32

[a] Derived from the $^{1}H/^{13}C$ chemical shift-derived secondary structure (Hodsdon *et al*., 1995).

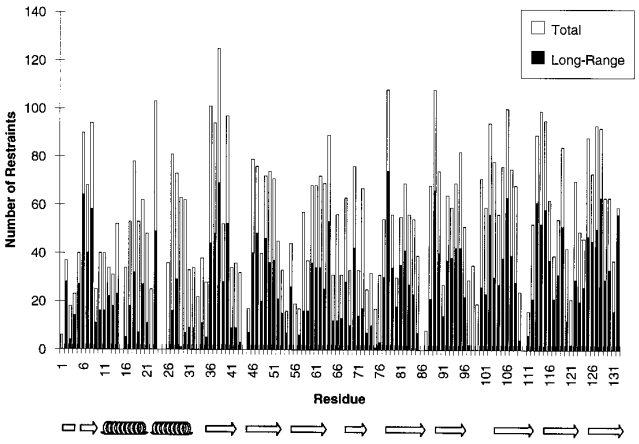

Figure 2. The distribution of the total (open bars) and long-range (filled bars) distance restraints used for the calculation of the final ensemble of NMR structures. Long-range distance restraints are defined as $|i-j| > 4$, where i and j are residue numbers in the sequence. In this Figure, each inter-residue restraint is allotted twice, once for each of the dipolar-coupled partners. The chemical shift-derived secondary structure of I-FABP (Hodsdon *et al*., 1995) is shown below the plot.

1995), dramatically restricted the number of possible tertiary folds the protein could adopt. In addition, the symmetry check eliminated most of the incorrect interpretations of a subset of the cross-peaks from the 3-D ^{13}C-resolved NOESY spectrum. When used together, this information established the molecular topology of I-FABP in solution and provided an initial NMR-derived model for assigning additional NOESY data. From this initial interpretation, 1679 interproton distance restraints were derived and served as input for the calculation of an ensemble of somewhat imprecise, but accurate tertiary structures that established the global fold of I-FABP. This ensemble was then used to reinterpret the NOESY data, and further rounds of structure calculations and NOE assignments were performed in an iterative manner, as detailed in Materials and Methods.

The final set of structural restraints is described in Table 1, and the distribution of restraints along the protein sequence is plotted in Figure 2. The total number of restraints, 3889 or 30 per residue, is among the highest for proteins whose structures have been determined by NMR. Also, the number of long-range restraints, important for determining tertiary structure, represent a substantial fraction of the total. The distance restraints were interpreted in a conservative manner by setting the upper bound on the distance between the two dipolar-coupled protons to 5.0 Å and their lower bound distance to 1.86 Å. Because of potential complications arising from incomplete relaxation, spin-diffusion and the influence of the directly attached ^{13}C and ^{15}N nuclei, the NOE cross-peak intensities were considered unsuitable in this case for any further lowering of the upper bound or subclassification into bins such as strong, medium and weak.

The conformational restraints involving the bound palmitate included 53 ligand-protein and five intra-ligand distance restraints (Table 1). The ligand-protein NOE restraints involved the 2–4 and 14–16 positions of the fatty acid. The lack of restraints involving the middle of the hydrocarbon chain resulted from chemical shift degeneracy of the carbons and protons and the inability to assign NOE correlations in this region. All of the observed intra-ligand NOEs were short-range, and no interactions were observed between the two ends of the bound fatty acid. The existence of a salt-bridge interaction between the guanidinium group of Arg106 and the carboxylate group of palmitate, a feature of the 2.0 Å crystal structure (Sacchettini *et al*., 1989), has also been independently observed by NMR (Cistola *et al*., 1989, 1996; Jakoby *et al*., 1993). Therefore, an additional restraint, requiring that the distance between one of the guanido protons of Arg106 and one of the palmitate carboxylate oxygen atoms be less than 1.8 Å, was included in the structure calculations.

Computational strategy and rationale

The final set of 3889 inter-proton distance restraints was used as input for a computational procedure based on distance geometry and simulated annealing. It has been previously recognized that some classical implementations of distance geometry provide an insufficient sampling of the conformational space consistent with the data and generate families of overly extended structures with RMSD values that are too low (Metzler *et al*., 1989; Havel, 1990). The technique of metrization was developed to improve the sampling properties of distance geometry and provide a more realistic assessment of the precision of an ensemble of structures (Havel, 1990; Kuszewski *et al*., 1992). However, metrization is rarely used in practice, perhaps because it is computationally more expensive and can produce other kinds of artifacts (Havel, 1990, 1991). Our experience with this 15.4 kDa complex indicated that some implementations of distance geometry, such as XPLOR version 3.1 and an earlier version of DISTGEOM,

had difficulty producing properly embedded structures when metrization was employed. The problem originated from the uniform distribution used to select the trial distances, and the severity of the problem appeared to increase with molecular mass. The trial distances appeared to be systematically underestimated, giving rise to protein structures which appeared crushed and tangled after embedding (see Materials and Methods). Simulated annealing failed to refine these structures to convergence, generally producing final structures with improper local geometry and poor agreement with the experimental restraints.

To overcome these problems, a distance geometry algorithm employing an improved metrization procedure was developed and applied to the NMR structure determination for I-FABP. This algorithm is implemented in the current version of DISTGEOM/TINKER developed by J. W. Ponder (unpublished results). This distance geometry algorithm has three distinguishing features. First, it employs an iteratively optimized Gaussian distribution, rather than a uniform distribution, for the selection of trial distances during metrization. Although the use of a uniform distribution may be warranted in the case of an unconstrained polypeptide chain up to 60 residues (Havel, 1990), a Gaussian distribution centered at a fractional distance of 0.6 to 0.7 appears to be more appropriate for larger, more compact molecules such as globular proteins (see Figure 1 of Oshiro *et al*., 1991). In DISTGEOM, the Gaussian distribution is iteratively optimized in terms of its mean and standard deviation by comparing properties of the embedded structures with those estimated directly from the distance matrix (J.W.P., unpublished results). Second, the DISTGEOM algorithm chooses pairwise elements of the distance matrix, rather than atoms in the molecule, as in ''random atom'' metrization procedures (Havel, 1990; Kuszewski *et al*., 1992). The choice of pairwise elements of the distance matrix affords a greater degree of randomness in the order in which the trial distances are chosen during metrization and results in a better sampling of the conformational space for a given computational effort. Finally, to make these calculations feasible, DISTGEOM employs a more computationally efficient algorithm for the application of the metrization step. The new method is formally $O(N^2)$ instead of $O(N^3)$, as compared with other current algorithms. We refer to this distance geometry method as pairwise Gaussian metrization.

In the simulated annealing procedure following distance geometry, we deliberately employed a simple penalty function containing terms only for the experimental NMR restraints and standard local covalent bond geometries. Thus, it did not include the 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. We chose this conservative approach so that the NMR structures presented here would be a more direct representation of the experimental restraints. Energy minimization involving additional terms may be incorporated into a future stage of refinement, but for now, the results are not complicated or biased by such procedures.

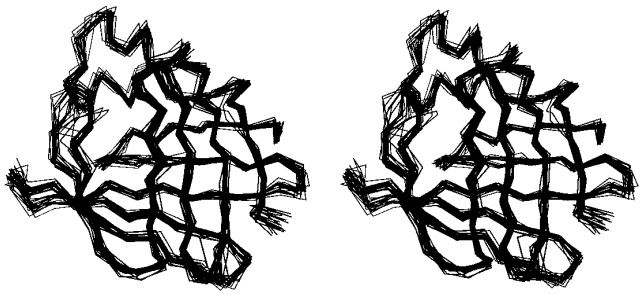

Figure 3. Stereo diagram of 20 superposed backbone C^α traces representing the final ensemble of NMR structures for I-FABP complexed with palmitate. These structures were calculated using the distance geometry/simulated annealing protocol detailed in Materials and Methods. The average pairwise C^α RMSD of this family is 0.98 Å. For clarity, the ligand is not shown in this Figure, even though it is present in the structures represented here (see also Figures 4 and 5). This Figure, as well as Figures 4, 5 and 10, were generated using MOLSCRIPT (Kraulis, 1991).

In calculating the structure of this fatty acid-protein complex, the ligand atoms were present throughout the procedure and thus, were embedded and refined along with the protein. This approach prevented the protein from sampling and occupying conformational space that should have been occupied by ligand (and *vice versa*) and permitted an assessment of the precision of the structures in the ligand-binding cavity.

Description of the final structures

The final family of 20 embedded and refined structures is presented as a stereoplot in Figure 3. The structures are represented by superposed traces through the backbone C^α coordinates of the protein. The traces correspond to 20 separate distance geometry/simulated annealing calculations using identical input data. Therefore, the superposed structures in Figure 3 provide a qualitative and quantitative assessment of the precision of the data and of the conformational space consistent with the data. The conformation of the protein backbone appears to be well determined. The average pairwise root-mean-square deviation for the family is slightly less than 1 Å, a realistic value given the large number of restraints combined with the current lack of stereospecific assignments (Havel, 1991). Note that this value corresponds to the average pairwise C^α RMSD, as opposed to the RMSD about the mean atomic coordinates (''average structure''). No attempt was made to calculate an average structure at this stage.

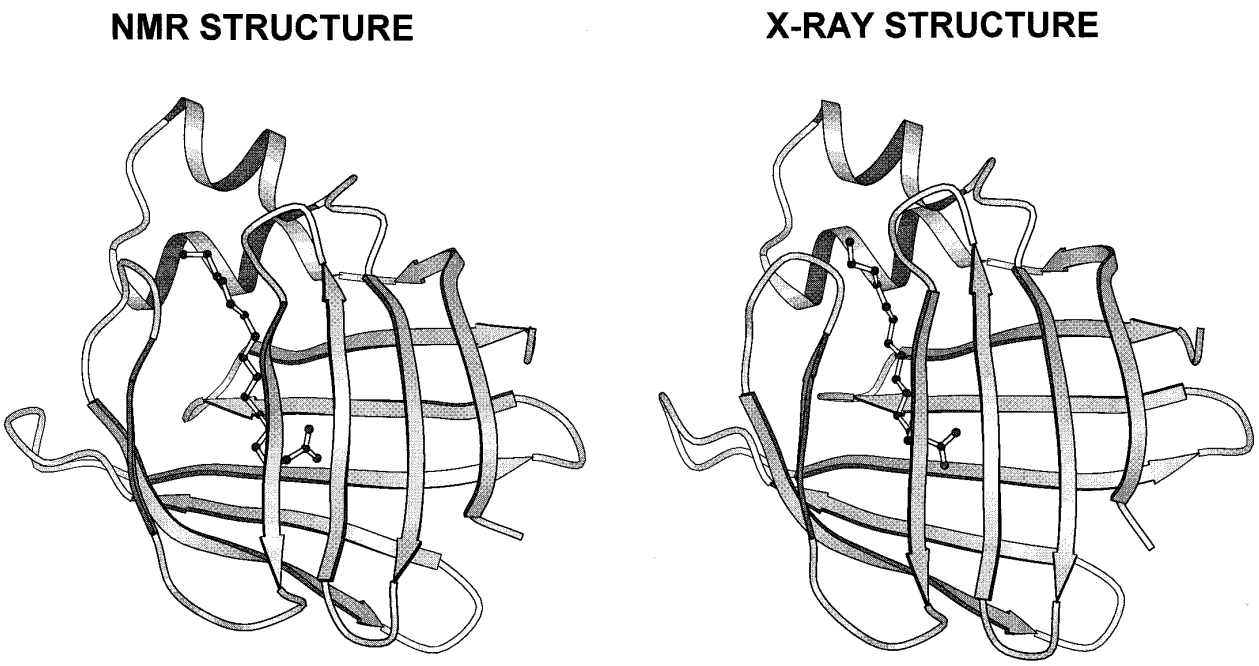

Figure 4. Ribbon diagrams representing a single selected structure from the NMR ensemble in Figure 3 (left) and the X-ray crystal structure (Sacchettini *et al*. 1989). In the NMR structure, the positions of the secondary structure elements were defined by the consensus $^{1}H/^{13}C$ chemical shift index for holo I-FABP (Hodsdon *et al*., 1995). The bound fatty acid is shown in ball and stick format.

Also, we saw no clear justification in this case for excluding any regions of the protein that appeared more variable, such as the N terminus, so the reported RMSD value corresponds to the entire protein sequence.

A single member of the family of 20 was chosen and displayed as a ribbon diagram in order to further illustrate the general features of the NMR structure of holo-I-FABP (Figure 4, left). The structure consists of two orthogonally packed β-sheets, each containing five β-strands. Two α-helices extend from the turn between the first and second β-strands and pack against one edge of the β-sheet domain. This structural motif has been described as a β-clam and has been found for all members of the lipid-binding family whose tertiary structure has been determined to date (Banaszak *et al*., 1994).

The location and conformation of the bound palmitate in the family of NMR structures is shown in Figures 4 and 5. The ligand is bound in an interior cavity between the two β-sheets, and its overall conformation is very similar to that observed in the X-ray crystal structure. At the current level of refinement, the heterogeneity in the local structure of the bound palmitate is slightly larger than that observed in the protein side-chains (Figure 5). One reason is that the protein backbone partially anchors the side-chains and effectively restricts their conformations. However, the palmitate is provided the extra freedom of rotation and translation relative to the protein and is limited only by the experimental restraints. Also, as indicated above, our current data set currently contains ligand-protein restraints involving the two ends, but not the middle, of the fatty acid hydrocarbon chain. Nevertheless, the general

Figure 5. The ensemble of 20 final NMR structures emphasizing the bound fatty acid and several selected side-chains. The position of the protein backbone is indicated by a single C^{α} trace taken from Figure 3. The bound palmitate is located near the center of the Figure with its carboxyl end pointing toward the lower right and its methyl end toward the upper left. Starting at the top of the Figure and proceeding clockwise around the structure, the displayed side-chains are Lys27, Phe128, Arg106, Leu64, and Tyr70.

Table 2. Stereochemical quality of structures as assessed by PROCHECK version 3.0

	Average values for family of 20 structures	Values for single structure	Reference values for 2.5 Å X-ray structures
Ramachandran plot statistics			
Residues in most favored regions [A, B, L]	68(58.1%)	73(62.4%)	76.6 ± 10.0%
Residues in additionally allowed regions [a, b, l, p]	41(35.0%)	36(30.8%)	
Residues in generously allowed regions [~a, ~b, ~l, ~p]	6(5.1%)	6(5.1%)	
Residues in disallowed regions	2(1.7%)	2(1.7%)	
Total number of non-glycine and non-proline residues	117(100%)	117(100%)	
Main-chain statistics			
Standard deviation of ω angle (degrees)	0.7	0.3	6.0 ± 3.0
Number of bad contacts per 100 residues	6.4	6.1	10.5 ± 10.0
Standard deviation of ''Zeta-angle'' (degrees)	5.2	5.2	3.1 ± 1.6
Standard deviation of H-bond energy (kcal/mol)	0.9	0.9	0.9 ± 0.2
Overall PROCHECK G-factor	–0.6	–0.6	–0.6 ± 0.3
Side-chain statistics (standard deviations in degrees)			
Chi-1 *gauche* minus	27.9	33.9	22.7 ± 6.5
Chi-1 *trans*	25.9	24.5	22.7 ± 5.3
Chi-1 *gauche* plus	26.9	24.8	21.3 ± 4.9
Chi-1 ''pooled''	27.3	26.5	22.0 ± 4.8
Chi-2 *trans*	27.4	23.9	23.1 ± 5.0

location and orientation of the bound fatty acid is well established.

The proper use of metrization was found to be important for obtaining converged, refined structures in the presence of the bound palmitate. In the embedded structures obtained using uniform metrization, the ligand was frequently located on the exterior of the protein. These structures could not be refined to convergence using simulated annealing. In order to satisfy the experimental restraints to the palmitate, the protein structure contorted itself to present interior side-chains to its exterior, resulting in high penalty functions and grossly distorted structures (see Materials and Methods). This artifact was avoided in the current study by the use of pairwise Gaussian metrization.

The stereochemical quality of the family of NMR structures and the single selected structure was analyzed using PROCHECK version 3.0 (Laskowski *et al*., 1993). The PROCHECK suite of programs was originally designed to compare newly determined X-ray crystal structures to previously determined structures of similar resolution. Many of its criteria for stereochemical quality are dependent on the degree of resolution of the crystal structure model and on the method of refinement. Protein solution structures determined by NMR spectroscopy are not characterized by a parameter analogous to the resolution of an X-ray structure. As well, the method of refinement used in this study differed from the typical refinement of crystal structures in that only the experimental NMR restraints and local covalent bond geometries were used. Nevertheless, the inclusion of an analysis of stereochemical quality has become standard in the description of newly determined protein structures, and PROCHECK was considered to be a generally suitable method for assessing structural quality.†

Stereochemical quality parameters for the family of NMR structures and for the single selected structure are summarized in Table 2. Statistics are provided for the subdivision of backbone ϕ/ψ dihedral angles into variously favored regions of a Ramachandran plot. On average, 93.1% of the residues were found in allowed regions of the Ramachandran plot and 58.1% of the residues were in the most favored regions. This latter value is slightly lower than the typical range observed for crystal structures of 2.5 Å resolution. However, as indicated above, the family of NMR structures were refined with a limited penalty function lacking a general torsional energy term. The inclusion of such a term would be expected to localize backbone torsions to more favorable values.

Other main- and side-chain parameters determined from the PROCHECK analysis are indicated in Table 2. Overall, the stereochemical quality of the family of NMR structures and the single selected structure best matched the typical ranges observed for 2.5 Å resolution crystal structures. This is exemplified by the overall PROCHECK *G*-factor of –0.6. The only significant deviations were found for the ω angle defining the planarity of the peptide bond and the Zeta-angle, a virtual torsion defining the C^{α} chirality. During refinement of the NMR structures, peptide bond planarity was regulated as a component of the penalty function resulting in a narrow distribution of the ω angle for the family. Flexible restrictions on bond angles in the penalty function resulted in broader distributions than has typically been found in comparable crystallographic structures, hence a slightly higher standard deviation of the C^{α} chirality defined by the Zeta-angle.

† After this study was completed, we learned of a new version of the PROCHECK software designed for the analysis of NMR structures. Information can be obtained at the following web site: http://www.biochem.ucl.ac.uk/bsm/

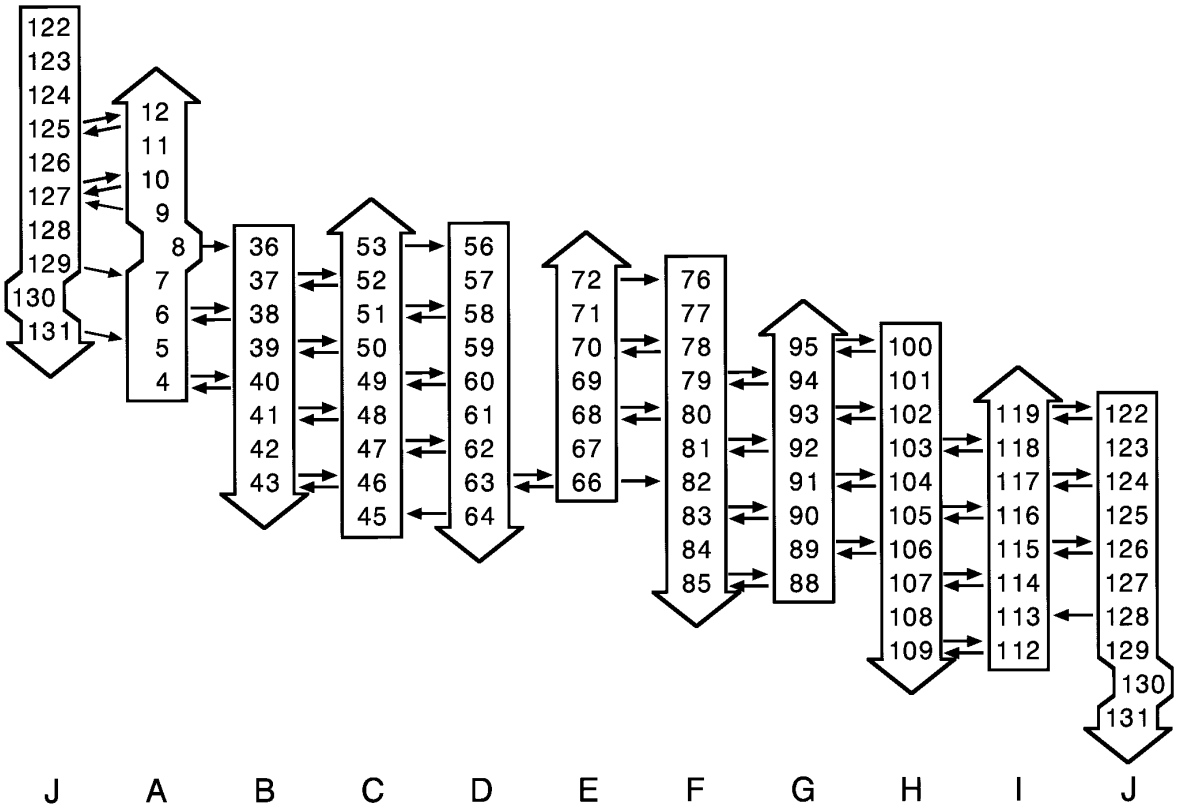

Figure 6. Network of β-sheet hydrogen bonds identified from an analysis of the final ensemble of NMR structures. Each arrow represents a single hydrogen-bond pointing from the backbone amide group of one residue to the backbone carbonyl of another. The presence of an arrow indicates that at least 50% of the members of the ensemble exhibited a hydrogen bond defined by the criteria in the text.

Comparison to the X-ray crystal structure

The NMR and X-ray crystal structures of I-FABP complexed with palmitate are displayed side-by-side as ribbon diagrams in Figure 4. In terms of both local and global structure, they appear to be very similar. A more detailed and quantitative comparison of the two structural models was provided by an analysis of hydrogen-bonding, root-mean-square-deviations (RMSD), and difference distances matrices (DDM).

Backbone hydrogen bonds were located in the family of NMR structures using the criteria of less than 2.5 Å separation between the amide hydrogen and the carbonyl oxygen and bond angles for N-H-O and H-O-C greater than 90°. Hydrogen bonds identified between a pair of residues in over 50% of the structures in the NMR family were considered valid. Given the strict 2.5 Å cutoff and the 1 Å precision of the ensemble, this criterion was quite stringent. Note that no ''hydrogen bond restraints'' were used in the calculation and refinement of these NMR structures. Rather, the final chemical shift and NOE-derived structures were used to directly identify the positions of hydrogen bonds according to the criteria outlined above. Diagrammed in Figure 6 are the 67 inter-strand hydrogen bonds located within the β-sheet region of the protein. The hydrogen bond topology was nearly identical to that seen in the X-ray crystal structure. Few backbone hydrogen bonds were identified between β-strands D and E, but some interstrand NOEs involving side-chains were observed. In the crystal structure, contacts between these strands are provided by bridging solvent molecules and some side-chain interactions. The present NMR study provides no information on solvent interactions with the protein, although at least one ordered water molecule adjacent to the fatty acid C-2 position has been identified by ^{19}F-^{1}H NMR (Cistola & Hall, 1995). Although not shown in Figure 6, intrahelical backbone hydrogen bonds were also observed and exhibited the expected $i / i + 4$ pattern for a regular α-helix. The positions of side-chains are not yet defined with sufficient precision to justify the assignment of hydrogen bonds involving side-chain moieties.

A quantitative comparison of the family of NMR structures and the X-ray crystal structure was provided by an analysis of the RMSDs of atomic coordinates and DDMs of internal coordinates. The top panel of Figure 7 provides the distribution of average C^{α} deviations along the protein sequence for the family of superposed NMR structures shown in Figure 3. The most-ordered regions of the backbone display average deviations of

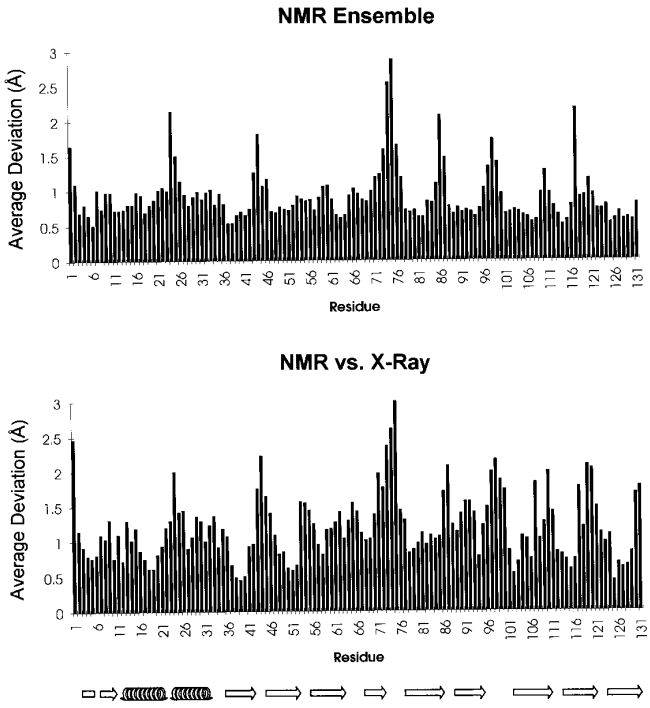

Figure 7. Distributions of the average C^{α} deviations along the sequence of the protein for NMR ensemble (top) and for the NMR ensemble *versus* the X-ray structure (bottom). The *y*-axis represents the pairwise average deviation of the C^{α} coordinates calculated for every possible pair of superposed structures. In the top panel, the average deviations between the C^{α} coordinates for every pair of structures in the final NMR ensemble is plotted. In the bottom panel, the average of the deviations between each NMR structure and the crystal structure are shown. The chemical shift-derived secondary structure is shown at the bottom of the Figure.

approximately 0.8 Å with the average for the entire protein being slightly less than 1 Å. Isolated regions of significantly increased values can be seen for a subset of the turns between secondary structure elements. A comparison of the results in Figures 2 and 7 shows a correlation between regions of increased C^{α} deviations and regions lacking long-range distance restraints. A similar pattern is observed in the bottom panel in Figure 7, which shows the distribution of the average deviations between the ensemble of NMR structures and the X-ray structure. For the entire protein, the average pairwise C^{α} RMSD of the NMR structure ensemble from the X-ray crystal structure is 1.3 Å. Again, there is a correlation of regions with increased deviation between the NMR and X-ray structure with regions lacking long-range NMR restraints. Hence, those areas of NMR structure that differ most from the X-ray structure are less well defined within the ensemble of NMR structures.

Distance and difference-distance matrices have been used to identify and analyze elements of protein structure (Kundrot & Richards, 1987). Distance matrices were calculated for a single structure by determining the intramolecular distances between all possible pairs of C^{α} positions. Two protein structures were then compared by generating a difference distance matrix, calculated by arithmetic subtraction of their individual distance matrices. Analyzing the heterogeneity in the family of NMR structures by this method provides more information than obtained from average deviation and RMSD values. Whereas an increased average deviation simply indicates a region of local structural heterogeneity, a DDM analysis of the same region is capable of describing the heterogeneity relative to the rest of the protein structure. Hence, the presence of structural domains varying in position relative to the global structure but fixed internally can be identified.

Distance matrices and DDMs describing the family of NMR structures and their comparison to the X-ray structure are provided in Figure 8. The close contacts between individual elements of secondary structure are clearly visible in the average distance matrix for the family of NMR structures (lower-left, first matrix) and the X-ray structure distance matrix (lower-left, second matrix). The largest average differences between C^{α} distances within the NMR family (upper-right, first matrix) were found to correspond to the same set of turns between secondary structure which gave increased average deviations. The DDM indicates that the heterogeneity in these structural regions is purely local as the increased difference distance values appeared approximately uniform relative to the rest of the protein. The differences between the family of NMR structures and the X-ray structure (upper-right, second matrix) are seen to correspond primarily to the poorly determined turn regions of the NMR structures. Otherwise, the distance matrices for the X-ray structure and the average of the NMR family show no systematic differences.

The above comparisons indicate that, within the precision of the data, the NMR and X-ray structures of I-FABP complexed with palmitate appear to be essentially identical.

Comparison to the structures of homologous proteins

X-ray crystal structures containing defined fatty acids have been determined for four other members of the iLBP family besides I-FABP. These include the myelin P2 protein (Jones *et al*., 1988), heart/muscle fatty acid-binding protein (Müller-Fahrnow *et al*., 1991; Zanotti *et al*., 1992; Young *et al*., 1994), adipocyte lipid-binding protein (Xu *et al*., 1993), and the insect midgut fatty acid-binding protein 2 (Benning *et al*., 1992). All five proteins display the same overall β-clam topology and bind a single molecule of fatty acid in an interior cavity between the two β-sheets. However, the location and conformation of the bound fatty acid differs. In some cases, the hydrocarbon chain of the fatty acid adopts a U-shaped conformation, whereas in others, it is more extended. In addition, the polar end of the fatty acid interacts with the

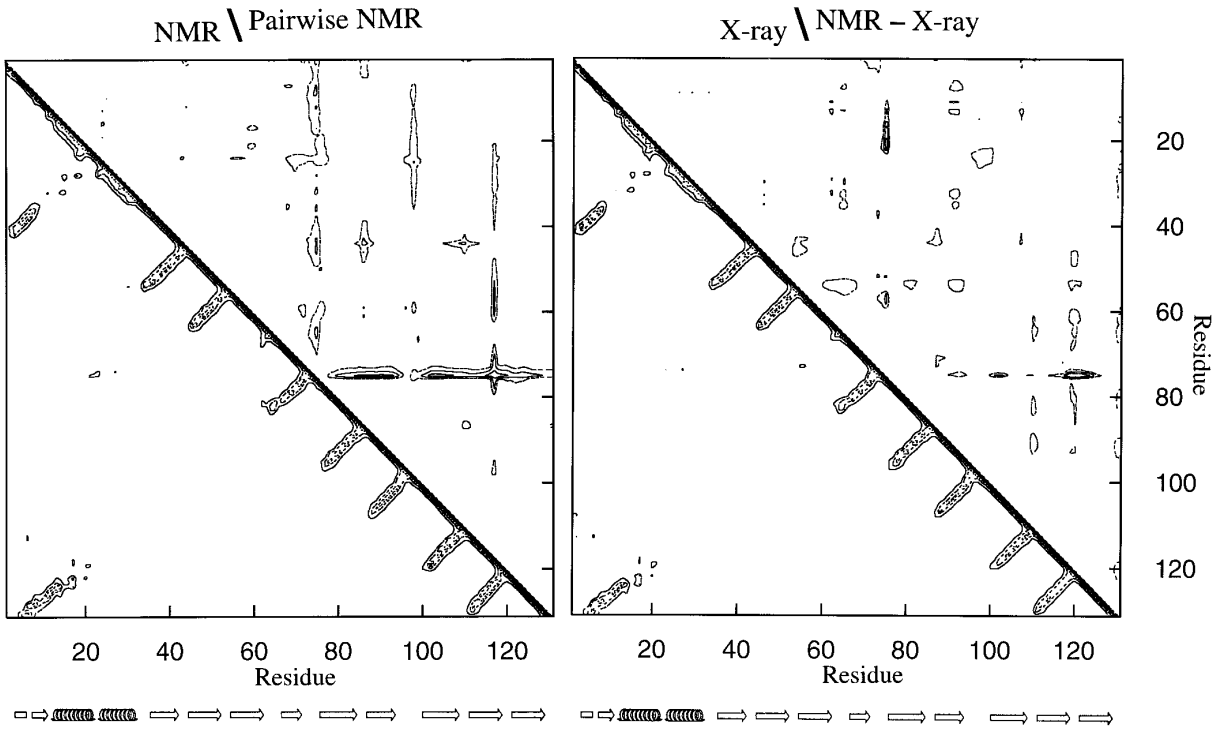

Figure 8. Distance matrices and difference distance matrices illustrating comparisons within the NMR family and between the NMR and X-ray crystal structures. In the left panel, the average C^{α} distance matrix for the family of NMR structures (lower left, NMR) is compared to the matrix of the average difference between the C^{α} distance matrices for every possible pair of structures in the family (upper right, Pairwise NMR). In the right panel, the C^{α} distance matrix for the X-ray structure (lower left, X-ray) is compared to the matrix of the average difference between the family of NMR structures and the X-ray structure (upper right, NMR—X-ray). The matrices are contoured with five lines at 2, 3, 4, 6, and 8 Å, except the Pairwise NMR matrix which was contoured at 1, 1.5, 2, 3, and 4 Å. The solution secondary structure is illustrated at the bottom of each matrix for reference.

protein in one of two distinct motifs (Banaszak *et al*., 1994). The first, described as the P2 motif, is adopted by the myelin, adipocyte, heart and insect midgut proteins. In these structures, the fatty acid carboxyl group interacts with Arg126, Tyr128 and Arg106, although the interaction with Arg106 is indirect and occurs through a bridging water molecule. The second motif is unique to I-FABP, and is characterized by a direct ion-pair/hydrogen bond interaction between the fatty acid carboxylate and Arg106, with no participation from residues 126 and 128 (Sacchettini *et al*., 1989). The NMR results restrain the fatty acid to interact with Arg106 and not residues 126 or 128 and provide additional evidence for the unique binding motif exhibited by I-FABP.

An NMR solution structure has also been established for bovine heart fatty acid-binding protein (H-FABP) complexed with palmitate. This protein shares 30% sequence identity with rat I-FABP. The H-FABP structure was based initially on 1071 restraints derived from 2-D ^{1}H spectra of an unenriched sample (Lücke *et al*., 1992; Lassen *et al*., 1993) and subsequently, 2371 restraints derived from 2-D ^{1}H and 3-D ^{15}N-resolved spectra of a uniformly ^{15}N-enriched sample (Lassen *et al*., 1995). The overall fold was the same as observed here for I-FABP, but some differences were noted in the conformation of the bound ligand and several other regions of the protein. Unlike I-FABP, the bound palmitate in the NMR structure of H-FABP adopts a U-shaped conformation as described above. Also, apparent differences in the protein backbone were observed in α-helix II and the turns between β strands E and F, and G and H. In the ensemble of H-FABP NMR structures, the backbone conformation of helix II appeared to be more variable compared with the rest of the protein (Figure 5 of Lassen *et al*., 1995). In addition, chemical shift heterogeneity was observed in the NMR spectra for each isoform of bovine H-FABP. For the pI 4.9 isoform, multiple peaks were observed at positions 22 to 36, 53 to 60 and 119 to 122 (Lücke *et al*., 1992). Based on these findings, it was concluded that helix II and adjacent portions of the molecule maintain four distinct backbone conformations with lifetimes of at least 10 ms (Lassen *et al*., 1993). Multiple peaks were also observed for the pI 5.1 isoform at positions 2, 3, 6, 30 to 34, 36, 53 to 60, 75 to 76, 88 to 90, 100, 101 and 122 (Lassen *et al*., 1995).

The NMR results for I-FABP complexed with palmitate did not exhibit evidence for such chemical shift or conformational heterogeneity. Only one set of resonances was observed at each

amino acid position in the protein (Hodsdon *et al*., 1995). Also, helix II was quite well defined in the ensemble of structures (Figure 3), and the consensus $^{1}H/^{13}C$ chemical shift index was consistent with a defined second helix (Hodsdon *et al*., 1995). The reasons for these apparent differences in the NMR results for holo I-FABP and holo H-FABP are currently unknown, but could have significance with respect to the mechanisms of ligand entry and transfer for both proteins (see below).

Functional implications

In the small intestinal enterocyte, I-FABP is thought to facilitate the flux of fatty acids from the plasma membrane primarily to the smooth endoplasmic reticulum, where they are enzymatically converted to triacylglycerols for export *via* the chylomicron pathway. The detailed mechanism by which I-FABP acquires its ligand, transports it to the organelle membrane and releases it remains unclear. The 1.2 Å X-ray structure of apo I-FABP revealed no opening large enough to admit a long-chain fatty acid molecule (Scapin *et al*., 1992). If the X-ray structure is representative of a closed, ligand-inaccessible state in solution, then some type of conformational change must occur in order to admit the ligand. We recently observed kinetic evidence for such a change by comparing the ligand-binding properties of a helix-less variant of I-FABP with those of the wild-type protein (Cistola *et al*., 1996; Kim *et al*., 1996). The results indicated that the α-helical region of wild-type I-FABP was directly or indirectly involved in a conformational change from a closed to an open, ligand-accessible state.

More recently, the NMR solution structure of the apo form of I-FABP has been determined, as detailed elsewhere (Hodsdon & Cistola, 1996a). Unlike the holo NMR structure presented here, the apo NMR structure exhibited discrete disorder in the backbone conformation, as evidenced by ^{1}H and ^{13}C chemical shifts, amide ^{1}H exchange rates and interproton NOEs. In addition, apo I-FABP exhibited increased backbone mobility in the same region of the molecule involving residues 26 to 35, 54 to 59 and 73 to 75, as evidenced by comparative amide ^{15}N relaxation and ^{1}H saturation transfer rates (Hodsdon & Cistola, 1996b). This region of the molecule is thought to comprise the ligand entry portal. The discrete disorder and increased mobility in the backbone conformation of apo I-FABP implies that a manifold of open and closed states exists in solution at equilibrium. We proposed that the binding of ligand stabilizes a series of cooperative interactions, analogous to a C-terminal helix capping box, that shift the equilibrium toward the closed state (Hodsdon & Cistola, 1996a).

With regard to ligand release and transfer, Storch and colleagues have established evidence that some members of the iLBP family, specifically the heart and adipocyte proteins, transfer ligand to acceptor membranes *via* a collisional mechanism (Kim & Storch, 1992; Wootan *et al*., 1993). This collisional mechanism appears to be mediated by interactions between the proteins and negatively charged lipids in the membrane (Sunderland & Storch, 1993; Wootan & Storch, 1994; Herr *et al*., 1995). Lysine residues at positions 22 and 59 of heart FABP, as well as one introduced by mutagenesis at position 28, were found to substantially affect the rate of transfer (Herr *et al*., 1996). Similar transfer studies have not yet been performed with I-FABP, but it contains lysine and arginine residues at positions 16, 20, 27, 28 and 29 that could mediate a similar collisional mechanism.

Finally, the gene for I-FABP has been shown to exhibit polymorphism in humans (Polymeropoulos *et al*., 1990; Prochazka *et al*., 1993; Humphreys *et al*., 1994). One polymorphism, observed in the Pima Indian population of Arizona, occurs at codon 54 and gives rise to alanine- and threonine-encoding alleles, with frequencies of 0.71 and 0.29, respectively (Baier *et al*., 1995). The Pima population has a very high prevalence of type II diabetes, and insulin resistance is both a hallmark and a major risk factor for that disease. Individuals homozygous or heterozygous for the threonine allele exhibit an increased rate of fat oxidation and insulin resistance, and these metabolic factors could be related to a difference in the fatty acid binding properties of the Ala- and Thr-containing proteins (Baier *et al*., 1995). In other mammalian iLBPs, the corresponding residue at position 54 is typically Thr, Asn or Ser, and there are no other known examples of Ala at this position (Banaszak *et al*., 1994; Veerkamp & Maatman, 1995). Rat I-FABP, which is 85% identical with human, contains Asn at position 54. In the structures of holo rat I-FABP, the side-chain of Asn54 forms a hydrogen bond with the backbone carbonyl of residue 33, which is adjacent to the C-terminal end of helix II. Presumably, threonine could form a similar side-chain-to-backbone hydrogen bond, but Ala could not. The role of this hydrogen bond is currently unclear, but it could participate in the regulation of the order-disorder transition observed for apo I-FABP in solution that appears to affect the entry and exit of ligand (Hodsdon & Cistola, 1996a,b). Studies addressing the cooperative interactions between residues in this region of I-FABP are underway.

Materials and Methods

Sample preparation

The protocols used for the biosynthesis, purification and delipidation of uniformly ^{13}C- and ^{15}N-enriched I-FABP have been described elsewhere (Hodsdon *et al*., 1995), and similar methods were used for the preparation of unenriched I-FABP (Jakoby *et al*., 1993). The final protein concentration was 3 mM with a buffer composition of 20 mM potassium phosphate, 50 mM potassium chloride, 0.05% sodium azide (pH 7.2), suspended either in 20% $^{2}H_2O$ (H_2O sample) or 99.996% $^{2}H_2O$ ($^{2}H_2O$ sample). The isotope-enriched protein was

Table 3. Parameters used for NMR experiments

A. *Aquisition parameters*	Nucleus			No. of complex points			Spectral width (kHz)			
Experiment	F1	F2	F3	F1	F2	F3	F1	F2	F3	Scans
^{15}N-NOESY-HMQC	^{1}H	^{15}N	^{1}H	192	40	256	6.50	2.20	6.50	16
^{13}C-NOESY-HSQC	^{1}H	^{13}C	^{1}H	140	28	192	4.80	2.80	4.80	32
[^{1}H,^{13}C]-HSQC	^{13}C	^{1}H	—	256	512	—	7.60	6.50	—	32
[^{1}H,^{13}C]-HSQC-TOCSY	^{13}C	^{1}H	—	200	512	—	5.00	6.50	—	128
[^{1}H,^{1}H]-TOCSY-HSQC	^{1}H	^{1}H	—	186	512	—	3.00	3.00	—	64
[^{1}H,^{13}C]-NOESY-HSQC	^{13}C	^{1}H	—	300	512	—	5.00	5.00	—	128
[^{1}H,^{1}H]-NOESY	^{1}H	^{1}H	—	500	512	—	6.50	6.50	—	32

B. *Processing parameters*	Exponential factor			Gaussian factor			No. of processed real points		
Experiment	F1	F2	F3	F1	F2	F3	F1	F2	F3
^{15}N-NOESY-HMQC	n	–3	–3	n	0.030	0.025	256	64	512
^{13}C-NOESY-HSQC	5	–5	–5	n	0.020	0.010	256	64	512
[^{1}H,^{13}C]-HSQC	–2	–5	—	0.043	0.020	—	512	1024	—
[^{1}H,^{13}C]-HSQC-TOCSY	–4	7	—	n	0.040	—	512	1024	—
[^{1}H,^{1}H]-TOCSY-HSQC	5	n	—	n	0.075	—	512	1024	—
[^{1}H,^{13}C]-NOESY-HSQC	5	n	—	n	n	—	512	1024	—
[^{1}H,^{1}H]-NOESY	–7	–5	—	0.060	0.080	—	1024	1024	—

n, not used

complexed with a stoichiometric amount of perdeuterated palmitate, and the unenriched protein, with either [U-^{13}C]-, [2-^{13}C]- or [16-^{13}C]palmitate using the protocol of Cistola *et al*. (1989). The possibilities of incomplete delipidation or incomplete complex formation could be ruled out using 2-D ^{1}H TOCSY and ^{1}H/^{15}N HSQC spectra, since most of the backbone resonances for apo I-FABP were not superimposable with those of holo I-FABP (Hodsdon & Cistola, 1996b). The ^{13}C-enriched fatty acids were obtained from Cambridge Isotope Laboratories; their isotopic and chemical purity was established using ^{1}H and ^{13}C NMR spectra of samples dissolved in deuterated chloroform.

NMR spectroscopy

All spectra were collected at 37°C using a Varian Unity-500 NMR spectrometer equipped with a Nalorac 5 mm triple-resonance probe. Quadrature detection in the indirectly detected dimensions was achieved using hypercomplex data collection. Processing of the time-domain spectra was performed on a Sun SPARC-2 workstation using VNMR version 4.2 (Varian Associates). For spectral analysis, the processed and phased frequency-domain spectra were imported into NMR COMPASS version 2.5 (Molecular Simulations, Inc.) running on a Silicon Graphics INDY/R4000 workstation. Table 3 summarizes the acquisition and processing parameters for each of the NMR experiments described below. Unless otherwise stated, GARP-1 (Shaka *et al*,. 1985) or WALTZ-16 (Shaka *et al*., 1983) schemes were used for broad-band decoupling. For solvent suppression, all experiments employed low-power presaturation (<25 μW) of the water resonance during both the NOESY mixing time and the relaxation delay. The ^{1}H, ^{13}C and ^{15}N chemical shifts were referenced as previously described (Hodsdon *et al*., 1995).

Three-dimensional ^{15}N-NOESY-HMQC and ^{13}C-NOESY-HSQC experiments were performed using the H_2O and $^{2}H_2O$ samples, respectively. The ^{15}N-NOESY-HMQC spectrum was acquired with a mixing time of 150 ms and a 1.0 second relaxation delay. A delay of 4.5 ms ($1/2J_{NH}$) was utilized to establish antiphase magnetization prior to the generation of ^{1}H-^{15}N multiple-quantum coherence. A NOESY mixing time of 200 ms and a 0.9 second relaxation delay were used for the acquisition of the ^{13}C-NOESY-HSQC spectrum, with a 1.7 ms ($1/4J_{CH}$) delay during the INEPT subsequences. During the t_1 and t_3 periods, MPF10 (Fujiwara *et al*., 1993) was employed for broad-band decoupling of the 15 kHz ^{13}C spectral window.

Two-dimensional [^{1}H,^{13}C]-HSQC, [^{1}H,^{13}C]-HSQC-TOCSY, [^{1}H,^{1}H]-TOCSY-HSQC, and [^{1}H,^{13}C]-NOESY-HSQC spectra were collected for the samples of I-FABP complexed with [U-^{13}C]- , [2-^{13}C]- and [16-^{13}C]palmitate. A delay value of 1.7 ms ($1/4\ J_{CH}$) was used during the INEPT subsequences. The isotropic mixing times were arrayed at 15, 30 and 45 ms in the TOCSY experiments, and 150 ms mixing times were used in the NOESY experiments.

Analysis and assignment of the 3-D NOESY spectra

Cross-peaks in the 3-D ^{15}N- and ^{13}C-resolved NOESY spectra were picked manually in a computer-assisted manner by visual inspection of each ^{1}H-^{1}H spectral plane. A final list of 5313 peaks was obtained, with 3499 and 1814 derived from the ^{13}C- and ^{15}N-resolved spectra, respectively. The assignment of these 5313 cross-peaks was complicated by the severe degeneracy of ^{1}H chemical shifts that is not atypical for a 15 kDa protein, and an iterative protocol was used to circumvent this problem, as detailed below.

First, an in-house program called SEARCH was used to systematically search the database of I-FABP ^{1}H assignments at a tolerance of $\pm$0.04 ppm. This tolerance was empirically calibrated by averaging the chemical shift differences between a number of well-resolved intraresidue peaks along the indirectly detected proton dimension (*F*1); examples included the chemical shifts for α-β and β-α correlations in corresponding planes of the ^{13}C-resolved NOESY spectrum. The database search generated a list of all possible interproton distance restraints consistent with each observed NOE cross-peak. A total of 47,415 interpretations of 5313 cross-peaks, or an average of nine interpretations per cross-peak, was found. Very few of the individual cross-peaks had a single, unambiguous interpretation. Second, a subset of the ambiguous NOEs were tentatively assigned by searching all of the possible interpretations for symmetry-related cross-peaks on different ^{13}C or ^{15}N planes. This was performed automatically using the

in-house program SYMMETRY. For cross-peaks that passed the symmetry check, the ambiguity in the assignment was greatly reduced or eliminated. Third, the symmetry-checked restraints were used along with the $^{1}H/^{13}C$ consensus chemical shift-derived secondary structure of I-FABP (Hodsdon *et al*., 1995) to generate an initial NMR-based structural model of I-FABP. This model resembled a two-dimensional topological diagram and established that the β-strands were related in an antiparallel, meandering fashion. The precise register of the β-strands was not a part of this initial model and intentionally left vague. The two α-helices were considered as an extension of the turn between the first and second β-strands. Fourth, a single interpretation of some NOESY cross-peaks was then established using this structural model and a simple set of rules. If a single unambiguous intra-residue interpretation existed, it was accepted and other possible interpretations ignored. If no intra-residue correlation was found, then distance restraints consistent with the initial structural model were considered. Only unambiguous restraints were selected. For a large number of cross-peaks, no unambiguous interpretation was found, and they were set aside at this stage. This initial stage of 3-D NOESY cross-peak interpretation and assignment yielded a total of 1680 interproton distance restraints: 276 intraresidue, 432 sequential, and 972 long-range.

The initial set of 1680 restraints was used as input for the distance geometry and simulated annealing calculations described below. Twenty structures were generated using an upper bound of 7.0 Å, plus any prochiral corrections, for the experimental distance restraints. Because the precise register of neighboring β-strands was not a feature of the initial structural model, the interpretation of inter-strand NOEs represented a possible source of error. The increased upper bound was intended to accommodate some of the possible error in the initial interpretation of the NOESY cross-peaks. The average pairwise C^{α} RMSD for the 20 structures derived from these restraints was 2.5 Å.

This initial ensemble of I-FABP NMR structures acted as a powerful tool for the second manual interpretation of the NOESY data. The initial set of 1680 distance restraints were discarded, and the entire list of 5313 NOESY cross-peaks were manually reinterpreted using the new ensemble of structures. The ambiguities in the assignment of the ^{1}H chemical shift were resolved by considering the approximate global fold of the protein, not just the specific inter-proton distances in the structures. That is, only those cross-peaks that had a single interpretation less than 7 Å (corresponding to the closest distance between the proton pairs in the full ensemble of structures) were accepted at this stage and interpreted according to the rules stated above. This second manual interpretation gave rise to 1858 distance restraints (386 intraresidue, 540 sequential, and 932 long-range). As this set of restraints was based on a more detailed three-dimensional structural model, the upper bound was not artificially lengthened as before and a value of 5 Å was used. A second set of 20 structures was similarly generated with an average pairwise C^{α} RMSD of 2.4 Å.

Distance geometry calculations are useful for detecting inconsistencies or inaccuracies in the experimental restraints (Havel, 1991), and the frequently violated restraints were considered likely to be inaccurate. Therefore, restraints that were violated in five or more of the ensemble of 20 structures were discarded, and a subsequent family of structures was calculated. During later cycles of structure determination, however, these initially discarded restraints were not prevented from being reselected as valid interpretations of NOESY cross-peaks.

After two cycles of deletions and structure calculations, the set of 1858 restraints described above was reduced to 1798 and the average pairwise RMSD of the subsequent structures was 2.6 Å. At this stage, a two-dimensional homonuclear NOESY spectrum of unenriched holo I-FABP was then analyzed to identify 242 aromatic ring restraints. The aromatic ring protons were assigned based upon the intensity of the intraresidue NOE correlation to the β protons and confirmed by a set of NOESY correlations which agreed with the family of NMR structures generated thus far. The NOEs involving aromatic protons were analyzed within the context of the current NMR structures and set aside if ambiguous. The addition of the aromatic ring restraints led to a total of 2040 distance restraints and an ensemble of 20 structures with an average pairwise RMSD of 1.9 Å.

Further NOESY cross-peaks were assigned in an automated procedure using the in-house program DISTFIL and filtered for duplicate restraints using the program UNIQUE. All possible interpretations of the NOESY cross-peaks were considered according to how frequently they were satisfied in the current family of NMR structures. In the first automated step, distance restraints were added to the current set of 2040 if they were satisfied in all 20 structures derived from the set. This resulted in the addition of 808 restraints and a new family of 20 structures with an average pairwise RMSD of 1.5 Å. A requirement of satisfying 18 out of these 20 structures added 363 more restraints. Finally, restraints were added according to a criterion of 16 out of 20 structures in three further cycles of additions. At each step of the automated procedure, restraints were deleted if poorly satisfied in the family of derived structures using the procedure described above. This set of distance restraints consisted of 672 intra-residue, 906 sequential, and 2004 long-range restraints, including 58 involving the bound palmitate, for a total of 3582. During this process the average pairwise RMSD of the generated structures was reduced to 1.2 Å.

The final modification to the list of distance restraints was to eliminate the correction applied to some of the restraints involving prochiral resonances. Restraints involving methylene, geminal methyl, or aromatic ring protons required a correction to their upper bound in order to account for a correlation to either of their pro-chiral atoms. The 5 Å upper bounds for these restraints were increased by 1.76, 4.66, and 4.30 Å, respectively. This method of treating prochiral atoms in the absence of stereospecific assignments corresponds to the ''third method'' defined by Kuntz *et al*. (1989). It differs from the pseudoatom approach in that full atom representations and real atomic van der Waals radii are used. For distance geometry calculations, a separate restraint was entered into the bounds matrix for each of the pro-chiral resonances. In subsequent simulated annealing calculations, a weighting in the penalty function equal to the inverse of its degeneracy was used. For example, a single NOESY correlation involving a pair of methylene protons was considered as two separate distance restraints with upper bounds of 6.76 Å and each was weighted by 1/2 in the penalty function. In some cases, cross-peaks of comparable intensity were identified to both resolved pro-chiral resonances. In these cases, the prochiral correction was removed, since it was unlikely that one of these cross-peaks resulted solely

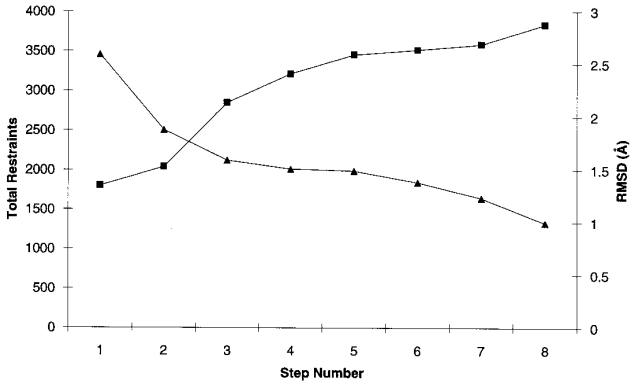

Figure 9. The progress of the iterative procedure for NOE cross-peak assignment and structure calculation. The *y*-axis displays the total number of interproton distance restraints (squares) and the average pairwise C^{α} RMSD values for the ensemble of 20 structures (triangles) at each major step of the iterative procedure.

from spin-diffusion. It was also necessary to replace the previous single restraint with a pair of restraints involving both of the specific pro-chiral resonances. A total of 249 restraints were modified as such, resulting in a final set of 3889 restraints (Table 1).

The progress of the iterative protocol for NOE interpretation and structure calculation is plotted in Figure 9. With each major step of the iteration, the number of distance restraints increased and the average pairwise C^{α} RMSD for the family decreased. Convergence in the total number of restraints was achieved after about six to seven iterations. Note that the apparently discontinuous increase in the number of restraints in step 8 resulted from the change discussed at the end of the previous paragraph. Further iterations may have resulted in a decreased RMSD without an increased total number of restraints. Such an artifactual decrease in RMSD would result from a reshuffling of restraint interpretations rather than a meaningful increase in the number of restraints. Therefore, we optimized the iterative procedure against the total number of restraints rather than the RMSD.

Structure calculations

All calculations were performed using the TINKER molecular modeling package running on a DEC 2100-A500 computer. This package contains the program DISTGEOM, which has the option of implementing distance geometry with pairwise Gaussian metrization. The features of this algorithm are outlined above in Results and Discussion and will be described in detail elsewhere. Prior to embedding, pairwise elements of the distance matrix were randomly chosen for the selection of trial distances. The choice of each trial distance within the bounds was weighted by a Gaussian distribution centered at approximately 0.6 to 0.7. The mean and standard deviation of the distribution was initially estimated based upon the size of the structure and the amount of information available in the restraints. Triangle inequality bounds smoothing was performed after each distance selection until 5% of the distances were metrized. The remainder of the trial distances were selected without metrization. After each structure of the ensemble was embedded, the Gaussian trial distribution was optimized to produce structures with the proper scaling of local distances. The mean of the Gaussian distribution typically reached a final value of 0.61 after two to three iterations. Hence, the problem of producing embedded structures which are improperly compressed or expanded was minimized.

After embedding, the distance geometry structures were regularized in DISTGEOM using the technique of majorization described by Havel (1991) followed by a brief minimization of the penalty function. Further refinement of the structures employed simulated annealing. The structures were initially equilibrated at 200°C for 1000 steps of molecular dynamics at a time-step of 0.04 ps, followed by 10,000 dynamics steps of cooling to 0°C at a time-step of 0.02 ps. The final refined structures typically had penalty function values less than 1 (see below). Structures with values higher than 10 were consistently observed to have distorted helices, conformations inconsistent with the experimental chemical shift data.

The penalty function for the molecular mechanics calculations described above consisted of a series of atom-based energy terms to enforce local bond geometry, prevent non-bonded contacts and implement the experimental restraints. Based upon the consensus chemical shift-derived secondary structure of I-FABP (Hodsdon *et al*., 1995), backbone Φ/Ψ dihedral restraints were implemented for the two α-helices located from residues 15 to 22 and 25 to 32. A flat-well parabolic penalty function restrained the Φ torsional angles between $-80°$ and $-40°$ and Ψ between $-60°$ and $-20°$ weighted such that a 10° deviation in either direction outside of these ranges resulted in a unit change of the penalty function. The dihedral restraints helped to prevent the helices from becoming trapped in local minima. Dihedral restraints were not enforced for the β-strand segments; they were not necessary for obtaining converged structures. A similar flat-well parabolic penalty function restrained the NOE distances to values between their upper and lower bounds such that a 1 Å violation resulted in a unit contribution to the penalty function. Except for those involving prochiral groups, distance restraints were applied with an upper bound of 5.0 Å and a lower bound of 1.86 Å.

For the final ensemble, a total of 23 structures were calculated and three were discarded based on high penalty function values. The penalty values for the final 20 were 0.55 ($\pm$0.37) (mean $\pm$ standard deviation) and the three obvious outliers had values of 19.5, 13.1 and 19.2. To determine whether the calculation of a larger number of structures would influence the final result, we subsequently extended the final ensemble to a total of 69 structures. Again, only a small fraction (6 of 69) exhibited elevated penalty function values. The average pairwise C^{α} RMSD for the larger group of structures was 1.1 Å, only slightly larger than the value of 0.98 Å determined for the original 20 structures. (Note that RMSD values scale with the number of structures in the ensemble.) Since the DG/SA algorithm had a convergence rate of $\sim$90% in this case, the original number of structures was sufficient to achieve adequate sampling. This example illustrates the power of properly implemented embedding algorithms for obtaining good starting structures for subsequent refinement and for facilitating convergence.

Figure 10 illustrates the influence of the type of distance geometry metrization algorithm on the I-FABP structures generated. The structure in Figure 10(A) was calculated using a uniform trial distribution and four-atom metrization (Havel, 1990, 1991; Kuszewski *et al*., 1992). Note the crushed and tangled appearance,

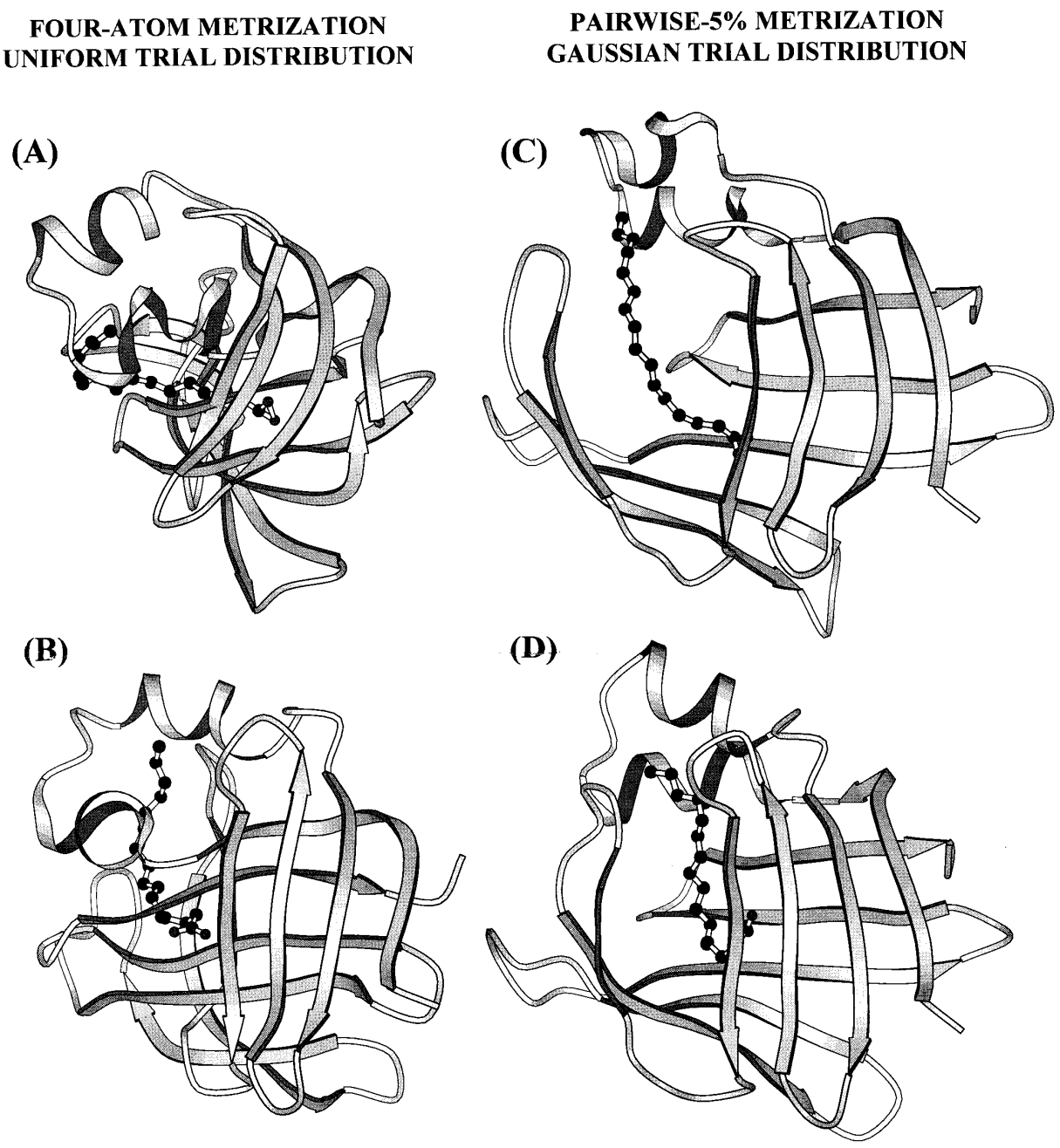

Figure 10. Ribbon diagrams illustrating the influence of the distance geometry metrization algorithm on the calculated structures both prior to ((A) and (C)) and after ((B) and (D)) simulated annealing refinement. All of the calculations were performed using DISTGEOM, which has the option of altering the trial distribution and the type of metrization employed. The structures in (A) and (B) were calculated using four-atom metrization with a uniform trial distribution; this is the type of algorithm implemented in X-PLOR version 3.1. The structures in (C) and (D) were calculated using pairwise-5% metrization with an optimized Gaussian trial distribution. This latter algorithm was used in the current study. For reference, the correct positions of secondary structure elements, as defined by $^{1}H/^{13}C$ chemical shifts, are indicated in each ribbon diagram; this is not meant to imply that regions of the calculated structures shown here actually adopt ϕ/ψ angles consistent with regular α-helix and β-sheet, especially in the distorted structures shown on the left. All of the calculations used the final set of distance restraints listed in Table 3. The structures shown are examples from an ensemble of structures obtained for each set of calculations; other members of the ensemble showed the same general characteristics as shown here.

which resulted from improper scaling of local distances. Subsequent refinement of this structure using simulated annealing (Figure 10(B)) was unable to produce an untangled structure; for example, note the three C-terminal ''β-strands'' protrude through the gap between the fourth and fifth β-strands. In contrast, the structure in Figure 10(C) was calculated using pairwise Gaussian metrization, as employed in the current study.

Although slightly expanded, this structure has the correct topology. Subsequent refinement with simulated annealing produced the structure shown in Figure 10(D). The crushed, tangled structures in Figure 10(A) and (B) were not atypical of the full ensemble calculated using the uniform trial distribution and four-atom metrization; nearly all of the structures in this ensemble were severely distorted.

Availability of software and molecular coordinates

Information about the DISTGEOM program and the TINKER molecular modeling package and instructions for downloading may be obtained at the following web site: http://dasher.wustl.edu (128.252.162.151). The molecular coordinates for the final ensemble of 20 NMR structures and the conformational restraints used to calculate them have been deposited in the Brookhaven Protein Data Bank (ID codes 1URE and R1UREMR, respectively).

Acknowledgements

This work was supported by grants from the National Science Foundation (MCB-9205665 to D.P.C.), the National Institutes of Health (GM-24483 to J.W.P.), the American Digestive Health Foundation, and institutional start-up funds. The Unity-500 spectrometer was supported in part by the Markey Center for Research in the Molecular Biology of Disease at Washington University. D.P.C. gratefully acknowledges a Johnson & Johnson/Merck Research Scholar Award from of the American Digestive Health Foundation. The authors are indebted to Drs David Wishart and Brian Sykes for providing their program for calculating chemical shift indices. We also thank Drs Chang-guo Tang for NMR advice and assistance and James Toner for the biosynthesis and purification of the samples.

References

Baier, L., J., Sacchettini, J. C., Knowler, W. C., Eads, J., Paolisso, G., Tataranni, P. A., Mochizuki, H., Bennett, P., Bogardus, C. & Prochazka, M. (1995). An amino acid substitution in the human intestinal fatty acid binding protein is associated with increased fatty acid binding, increased fat oxidation, and insulin resistance. *J. Clin. Invest.* **95**, 1281–1287.

Banaszak, L. J., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S. & Jones, T. A. (1994). Lipid-binding proteins: a family of fatty acid and retinoid transport proteins. *Advan. Protein Chem.* **45**, 89–151.

Bass, N. M. (1993). Cellular binding proteins for fatty acids and retinoids: similar or specialized functions? *Mol. Cell. Biochem.* **123**, 191–202.

Benning, M. M., Smith, A. F., Wells, M. A. & Holden, H. M. (1992). Crystallization, structure determination and least squares refinement to 1.75 Å resolution of the fatty-acid-binding protein isolated from *Manduca sexta L*. *J. Mol. Biol.* **228**, 208–219.

Cistola, D. P. & Hall, K. B. (1995). Probing internal water molecules in proteins using two-dimensional ^{19}F-^{1}H NMR. *J. Biomol. NMR,* **5**, 415–419.

Cistola, D. P., Sacchettini, J. C., Banaszak, L. J., Walsh, M. T. & Gordon, J. I. (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.

Cistola, D. P., 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.

Clore, G. M., Appella, E., Yamada, M., Matsushima, K. & Gronenborn, A. M. (1990). Three-dimensional structure of interleukin 8 in solution. *Biochemistry,* **29**, 1689–1696.

Clore, G. M., Kay, L. E., Bax, A. & Gronenborn, A. M. (1991). Four-dimensional ^{13}C/^{13}C-edited nuclear Overhauser enhancement spectroscopy of a protein in solution: application to interleukin 1β. *Biochemistry,* **30**, 12–18.

Fujiwara, T., Anai, T., Kurihara, N. & Nagayama, K. (1993). Frequency-switched composite pulses for decoupling carbon-13 spins over ultrabroad bandwidths. *J. Magn. Reson. ser. A,* **104**, 103–105.

Forman-Kay, J. D., Clore, G. M., Wingfield, P. T. & Gronenborn, A. M. (1991). High-resolution three-dimensional structure of reduced recombinant human thioredoxin in solution. *Biochemistry,* **30**, 2685–2698.

Glatz, J. F. C. & van der Vusse, G. J. (1989). Intracellular transport of lipids. *Mol. Cell. Biochem.* **88**, 37–44.

Glatz, J. F. C. & van der Vusse, G. J. (1990). Cellular fatty acid-binding proteins: current concepts and future directions. *Mol. Cell. Biochem.* **98**, 247–251.

Glatz, J. F. C., Vork, M. M., Cistola, D. P. & van der Vusse, G. J. (1993). *Prostaglandins Leukotrienes and Essential Fatty Acids,* **48**, 33–41.

Güntert, P., Berndt, K. D. & Wüthrich, K. (1993). The program ANSO for computer-supported collection of NOE upper distance constraints as input for protein structure determination. *J. Biomol. NMR,* **3**, 601–606.

Havel, T. F. (1990). The sampling properties of some distance geometry algorithms applied to unconstrained polypeptide chains: a study of 1830 independently computed conformations. *Biopolymers,* **29**, 1565–1585.

Havel, T. F. (1991). An evaluation of computational strategies for use in the determination of protein structure from distance constraints obtained by nuclear magnetic resonance. *Prog. Biophys. Mol. Biol.* **56**, 43–78.

Herr, F., Matarese, V., Bernlohr, D. A. & Storch, J. (1995). Surface lysine residues modulate the collisional transfer of fatty acid from adipocyte fatty acid binding protein to membranes. *Biochemistry,* **34**, 11840–11845.

Herr, F., 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, M. E. & Cistola, D. P. (1996a). Discrete backbone disorder in the NMR structure of apo intestinal fatty acid-binding protein in solution: implications for the mechanism of ligand entry. *Biochemistry*. In the press.

Hodsdon, M. E. & Cistola, D. P. (1996b). Ligand binding alters the backbone mobility of intestinal fatty acid-binding protein as monitored by ^{15}N NMR relaxation and ^{1}H exchange. *Biochemistry*. In the press.

Hodsdon, M. E., Toner, J. J. & Cistola, D. P. (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.

Humphreys, P., McCarthy, M., Tuomilehto, J., Tuomilehto-Wolf, E., Stratton, I., Morgan, R., Rees, A., Owens, D., Stengård, J., Nissinen, A., Hitman, G., Turner, R. C. & O'Rahlilly, S. (1994). Chromosome 4q locus associated with insulin resistance in pima indians. Studies in three european NIDDM populations. *Diabetes,* **43**, 800–804.

Jakoby, M. G., Miller, K. R., Toner, J. J., Bauman, A., Cheng, L., Li, E. & Cistola, D. P. (1993). Ligand-protein electrostatic interactions govern the specificity of retinol- and fatty acid-binding proteins. *Biochemistry,* **32**, 872–878.

Jones, T. A., Bergfors, T., Sedzik, J. & Unge, T. (1988). The three-dimensional structure of P2 myelin protein. *EMBO J.* **7**, 1597–1604.

Kaikaus, R. M., Bass, N. M. & Ockner, R. K. (1990). Functions of fatty acid binding proteins. *Experientia,* **46**, 617–630.

Kay, L. E., Clore, G. M., Bax, A. & Gronenborn, A. M. (1990). Four-dimensional heteronuclear triple-resonance NMR spectroscopy of interleukin-1β in solution. *Science,* **249**, 411–414.

Kim, H.-K. & Storch, J. (1992). Mechanism of free fatty acid transfer from rat heart fatty acid-binding protein to phospholipid membranes. Evidence for a collisional process. *J. Biol. Chem.* **267**, 20051–20056.

Kim, K., Cistola, D. P. & Frieden, C. (1996). Intestinal fatty acid-binding protein: the structure and stability of a helix-less variant. *Biochemistry,* **35**, 7553–7558.

Kraulis, P. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.* **24**, 946–950.

Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Pettersson, G., Knowles, J. & Gronenborn, A. M. (1989). Determination of the three-dimensional solution structure of the C-terminal domain of cellobiohydrolase I from *Trichoderma reesei*. A study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing. *Biochemistry,* **28**, 7241–7257.

Kuszewski, J., Nilges, M. & Brünger, A. T. (1992). Sampling and efficiency of metric matrix distance geometry: a novel partial metrization algorithm. *J. Biomol. NMR,* **2**, 33–56.

Kundrot, C. E. & Richards, F. M. (1987). Crystal structure of hen egg-white lysozyme at a hydrostatic pressure of 1000 atmospheres. *J. Mol. Biol.* **193**, 157–170.

Kuntz, I. D., Thomason, J. F. & Oshiro, C. M. (1989). Distance geometry. *Methods Enzymol.* **177**, 159–204.

Laskowski, R., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283–291.

Lassen, D., Lücke, C., Kromminga, A., Lezius, A., Spener, F. & Rüterjans, H. (1993). Solution structure of bovine heart fatty acid-binding protein (H-FABP$_c$). *Mol. Cell. Biochem.* **123**, 15–22.

Lassen, D., Lücke, C., Kveder, M., Mesgarzadeh, A., Schmidt, J. M., Specht, B., Lezius, A., Spener, F. & Rüterjans, H. (1995). Three-dimensional structure of bovine heart fatty-acid-binding protein with bound palmitic acid, determined by multidimensional NMR spectroscopy. *Eur. J. Biochem.* **230**, 266–280.

Lücke, C., Lassen, D., Kreienkamp, H.-J., Spener, F. & Rüterjans, H. (1992). Sequence-specific ^{1}H-NMR assignment and determination of the secondary structure of bovine heart fatty-acid-binding protein. *Eur. J. Biochem.* **210**, 901–910.

Matarese, V., Stone, R. L., Waggoner, D. W. & Bernlohr, D. A. (1989). Intracellular fatty acid trafficking and the role of cytosolic lipid binding proteins. *Prog. Lipid Res.* **28**, 245–272.

Meadows, R. P., Olejniczak, E. T. & Fesik, S. W. (1994). A computer-based protocol for semiautomated assignments and 3D structure determination of proteins. *J. Biomol. NMR,* **4**, 79–96.

Metzler, W. J., Hare, D. R. & Pardi, A. (1989). Limited sampling of conformational space by the distance geometry algorithm: implications for structures generated from NMR data. *Biochemistry,* **28**, 7052–7059.

Müller-Fahrnow, A., Egner, U., Jones, T. A., Rüdel, H., Spener, F. & Saenger, W. (1991). Three-dimensional structure of fatty acid-binding protein from bovine heart. *Eur. J. Biochem.* **199**, 271–276.

Nilges, M. (1995). Calculation of protein structures with ambiguous distance restraints. Automated assignment of ambiguous NOE cross-peaks and disulphide connectivities. *J. Mol. Biol.* **245**, 645–660.

Ockner, R. A., Manning, J. A., Poppenhausen, R. B. & Ho, W. K. L. (1972). A binding protein for fatty acids in cytosol of intestinal mucosa, liver, myocardium, and other tissues. *Science,* **177**, 56–58.

Oshiro, C. M., Thomason, J. & Kuntz, I. D. (1991). Effects of limited input distance constraints upon the distance geometry algorithm. *Biopolymers,* **31**, 1049–1064.

Polymeropoulos, M. H., Rath, D. S., Xiao, H. & Merril, C. R. (1990). Trinucleotide repeat polymorphism at the human intestinal fatty acid binding protein gene (FABP2). *Nucl. Acids Res.* **18**, 7198.

Prochazka, M., Lillioja, S., Tait, J. F., Knowler, W. C., Mott, D. M., Spraul, M., Bennett, P. H. & Bogardus, C. (1993). Linkage of chromosomal markers on 4q with a putative gene determining maximal insulin action in pima indians. *Diabetes,* **42**, 514–519.

Sacchettini, J. C. & Gordon, J. I. (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, J. C., Gordon, J. I. & Banaszak, L. J. (1989). Crystal structure of rat intestinal fatty-acid-binding protein. Refinement and analysis of the *Escherichia coli*-derived protein with bound palmitate. *J. Mol. Biol.* **208**, 327–339.

Scapin, G., Gordon, J. I. & Sacchettini, J. C. (1992). Refinement of the structure of recombinant rat intestinal fatty acid-binding protein at 1.2-Å resolution. *J. Biol. Chem.* **267**, 4253–4269.

Shaka, A. J., Keeler, J., Frenkiel, T. & Freeman, R. (1983). An improved sequence for broadband decoupling: WALTZ-16. *J. Magn. Reson.* **52**, 335–338.

Shaka, A. J., Barker, P. B. & Freeman, R. (1985). Computer-optimized decoupling scheme for wideband applications and low-level operation. *J. Magn. Reson.* **64**, 547–552.

Sunderland, J. E. & Storch, J. (1993). Effect of phospholipid headgroup composition on the transfer of fluorescent long-chain free fatty acids between membranes. *Biochim. Biophys. Acta,* **1168**, 307–314.

Sweetser, D. A., Heuckeroth, R. O. & Gordon, J. I. (1987). The metabolic significance of mammalian fatty-acid-binding proteins: abundant proteins in search of a function. *Annu. Rev. Nutr.* **7**, 337–359.

Veerkamp, J. H. & Maatman, R. G. H. J. (1995). Cytoplasmic fatty acid-binding proteins: their structure and genes. *Prog. Lipid Res.* **34**, 17–52.

Veerkamp, J. H., Peters, R. A. & Maatman, R. G. H. J. (1991). Structural and functional features of different types of cytoplasmic fatty acid-binding proteins. *Biochim. Biophys. Acta*, **1081**, 1–24.

Wootan, M. G. & Storch, J. (1994). Regulation of fluorescent fatty acid transfer from adipocyte and heart fatty acid binding protein by acceptor membrane lipid composition and structure. *J. Biol. Chem.* **269**, 10517–10523.

Wootan, M. G., Bernlohr, D. A. & Storch, J. (1993). Mechanism of fluorescent fatty acid transfer from adipocyte fatty acid binding protein to membranes. *Biochemistry*, **32**, 8622–8627.

Xu, Z., Bernlohr, D. A. & Banaszak, L. J. (1993). The adipocyte lipid binding protein at 1.6 Å resolution. Crystal structures of the apoprotein and with bound saturated and unsaturated fatty acids. *J. Biol. Chem.* **268**, 7874–7884.

Young, A. C. M., Scapin, G., Kromminga, A., Patel, S. B., Veerkamp, J. H. & Sacchettini, J. C. (1994). Structural studies on human muscle fatty acid binding protein at 1.4 Å resolution: binding interactions with three C18 fatty acids. *Structure*, **2**, 523–534.

Zanotti, G., Scapin, G., Spadon, P., Veerkamp, J. H. & Sacchettini, J. C. (1992). Three-dimensional structure of recombinant human muscle fatty acid-binding protein. *J. Biol. Chem.* **267**, 18541–18550.

Edited by P. E. Wright

(Received 19 June 1996; received in revised form 17 September 1996; accepted 24 September 1996)

Supplementary material, comprising three Tables, is available from JMB Online.