Journal of Biomolecular NMR, 9 (1997) 213–228
ESCOM

J-Bio NMR 424

Solution structure of human intestinal fatty acid binding protein: Implications for ligand entry and exit

Fengli Zhang[a], Christian Lücke[b], Leslie J. Baier[c], James C. Sacchettini[d] and James A. Hamilton[a,*]

[a]*Department of Biophysics, Boston University School of Medicine, 80 East Concord Street, Boston, MA 02118, U.S.A.*
[b]*Johann Wolfgang Goethe-Universität, Marie-Curie-Straße 9, D-60439 Frankfurt am Main, Germany*
[c]*Phoenix Epidemiology and Clinical Research Branch, NIDDK, NIH, Phoenix, AZ 85016, U.S.A.*
[d]*Department of Biochemistry and Biophysics, Texas A+M University, College Station, TX 77845, U.S.A.*

Received 11 October 1996
Accepted 12 December 1996

Keywords: Human intestinal fatty acid binding protein; Isotope enrichment; Multidimensional NMR spectroscopy; Sequential assignments; Solution structure

Summary

The human intestinal fatty acid binding protein (I-FABP) is a small (131 amino acids) protein which binds dietary long-chain fatty acids in the cytosol of enterocytes. Recently, an alanine to threonine substitution at position 54 in I-FABP has been identified which affects fatty acid binding and transport, and is associated with the development of insulin resistance in several populations including Mexican-Americans and Pima Indians. To investigate the molecular basis of the binding properties of I-FABP, the 3D solution structure of the more common form of human I-FABP (Ala^{54}) was studied by multidimensional NMR spectroscopy. Recombinant I-FABP was expressed from *E. coli* in the presence and absence of ^{15}N-enriched media. The sequential assignments for non-delipidated I-FABP were completed by using 2D homonuclear spectra (COSY, TOCSY and NOESY) and 3D heteronuclear spectra (NOESY-HMQC and TOCSY-HMQC). The tertiary structure of human I-FABP was calculated by using the distance geometry program DIANA based on 2519 distance constraints obtained from the NMR data. Subsequent energy minimization was carried out by using the program SYBYL in the presence of distance constraints. The conformation of human I-FABP consists of 10 antiparallel β-strands which form two nearly orthogonal β-sheets of five strands each, and two short α-helices that connect the β-strands A and B. The interior of the protein consists of a water-filled cavity between the two β-sheets. The NMR solution structure of human I-FABP is similar to the crystal structure of rat I-FABP. The NMR results show significant conformational variability of certain backbone segments around the postulated portal region for the entry and exit of fatty acid ligand.

Introduction

Fatty acid binding proteins (FABPs) are members of a family of intracellular lipid binding proteins, which have been isolated from a variety of tissues including muscle, liver, myelin, adipose tissue and the small intestine (Matarese et al., 1989; Veerkamp et al., 1991; Veerkamp and Maatman, 1995). These small, cytosolic proteins are thought to facilitate the transport and trafficking of lipids (Glatz and Van der Vusse, 1990; Kaikaus et al., 1990; Veerkamp et al., 1991; Bass, 1993; Banaszak et al., 1994), although their exact roles remain to be defined clearly. The intestinal fatty acid binding protein (I-FABP) is a highly abundant protein expressed solely in enterocytes of the proximal small intestine, which absorb and secrete dietary nutrients, such as fatty acids. The same tissue

*To whom correspondence should be addressed.

Abbreviations: FABP, fatty acid binding protein; I-FABP, intestinal fatty acid binding protein; H-FABP, heart fatty acid binding protein; L-FABP, liver fatty acid binding protein; ILBP, ileal lipid binding protein; COSY, correlation spectroscopy; TOCSY, total correlation spectroscopy; NOESY, nuclear Overhauser enhancement and exchange spectroscopy; HMQC, heteronuclear multiple-quantum coherence; HSQC, heteronuclear single-quantum coherence; DIANA, distance geometry algorithm for NMR applications; REDAC, redundant dihedral-angle constraint; rmsd, root-mean-square deviation.

contains three other intracellular lipid binding proteins: liver fatty acid binding protein (L-FABP), ileal lipid binding protein (ILBP) and cellular retinol binding protein II (CRBP II). These proteins are less specific and bind other lipids in addition to fatty acids (Veerkamp and Maatman, 1995). I-FABP consists of 131 amino acids and contains a single ligand-binding site that displays a high affinity for both saturated and unsaturated long-chain fatty acids in vitro (Lowe et al., 1987). In addition, human I-FABP has been shown to participate in the intracellular transport of long-chain fatty acids in vivo (Baier et al., 1996).

Studies of human I-FABP may offer a unique opportunity to define the roles of FABP in lipid transport and metabolism. A polymorphism has recently been identified in the gene which encodes human I-FABP (Baier et al., 1995). This A to G single base polymorphism results in an alanine (Ala^{54}) to threonine (Thr^{54}) substitution at residue 54 of I-FABP. Genetic studies with Pima Indians, a population with a high prevalence of insulin resistance and non-insulin dependent diabetes mellitus (NIDDM), have shown that the Thr^{54}-encoding I-FABP genotype (frequency = 0.29) is associated with increased fasting lipid oxidation rates and insulin resistance when compared to the more common Ala^{54}-encoding I-FABP genotype (frequency = 0.71) (Baier et al., 1995). Recombinant Thr^{54} protein has a twofold higher affinity for long-chain fatty acids in vitro as compared to recombinant Ala^{54} protein (Baier et al., 1995). In addition, enterocyte-like cells expressing the Thr^{54} protein transport long-chain fatty acids from their apical to basolateral surfaces at a faster rate than cells expressing the Ala^{54} protein (Baier et al., 1996).

The structure of human I-FABP has not been determined yet. X-ray crystallography has been used to determine the structure of rat I-FABP (Sacchettini et al., 1989, 1992; Scapin et al., 1992), which shares 82% homology with the human protein. The crystal structure has also been determined for several other FABPs, including human heart FABP (Zanotti et al., 1992; Young et al., 1994), bovine heart FABP (Müller-Fahrnow et al., 1991), bovine myelin FABP (Jones et al., 1988) and murine adipocyte FABP (Xu et al., 1992). In contrast, only two FABP solution structures, namely bovine heart FABP (Lücke et al., 1992; Lassen et al., 1993,1995) and ILBP (Lücke et al., 1996), have been determined by NMR spectroscopy. The X-ray crystal structures and the NMR solution structures have shown the same structural motif: 10 antiparallel β-strands (A–J) which form two nearly orthogonal β-sheets of five β-strands each, and two short α-helices (αI and αII) that connect the β-strands A and B. Elements of the secondary structure of rat I-FABP in solution have also been recently elucidated by multidimensional NMR (Hodsdon et al., 1995). Overall, the secondary structure of rat I-FABP, derived from analysis of chemical shift values, was similar to that determined by X-ray crystallography (Sacchettini et al., 1989).

To help understand the molecular basis of fatty acid binding to human I-FABP, and to compare its structure to that of other FABPs, we have analyzed the more common form of I-FABP (Ala^{54}) by 3D NMR spectroscopy. In this paper, we present the sequential assignments and the solution structure of non-delipidated human Ala^{54} I-FABP. Uniformly ^{15}N-enriched human I-FABP and heteronuclear 2D/3D NMR spectroscopy were used to establish the assignments and to obtain distance constraints. The solution structure of human I-FABP was calculated by using a distance geometry program, followed by distance-restrained simulated annealing and subsequent energy minimization.

Materials and Methods

Expression and purification of human I-FABP

The cDNA for human Ala^{54} I-FABP was isolated and ligated into the pET-3d expression vector (Novagen, Madison, WI, U.S.A.) as previously described (Baier et al., 1995). Protein was expressed in transformed BL21 cells (Novagen), and grown in LB media (for unlabeled protein) or M9 medium with $^{15}NH_4Cl$ [^{15}N, 98%+] (Cambridge Isotope Laboratories, Andover, MA, U.S.A.) as the only nitrogen source (for ^{15}N-labeled protein). Protein expression was induced in the transformed cells during cell growth at 37 °C by addition of 0.4 mM isopropyl β-D-thiogalactopyranoside (IPTG) when cells reached an OD_{600} of 0.5. About 3 h after induction, the bacteria were harvested by centrifugation. To purify I-FABP, a French press was used for cell lysis. Ammonium sulfate (final concentration of 60%) was used for protein precipitation. After dialysis to remove excess salt, the protein was applied to a Q-Sepharose column (Pharmacia, Piscataway, NJ, U.S.A.) that was equilibrated with Tris buffer at 4 °C using a flow rate of 200 ml/h. The fractions were analyzed by SDS-polyacrylamide gel electrophoresis. All fractions containing a band at ~15 kDa were combined and concentrated by using a Stir Cell with YM3 membrane (Amicon, Beverly, MA, U.S.A.). The protein was then applied to a G-75 Superdex column by using an FPLC system (Pharmacia) with a flow rate of 1 ml/min to remove small amounts of larger molecular weight impurities. The *E. coli* derived fatty acids which remained bound to the purified protein were not removed. These bound lipids presumably consisted of fatty acids of different chain lengths (ranging from C12 to C20) and different degrees of saturation but with 79% consisting of palmitic and stearic acid (Lowe et al., 1987). Therefore, the non-delipidated Ala^{54} I-FABP used in these studies represents the 'holo'-protein.

NMR experiments

All NMR measurements were performed at 37 °C using a 3–4 mM concentration of purified I-FABP in a

TABLE 1
CHEMICAL SHIFT VALUES OF THE ^{1}H AND ^{15}N RESONANCES ASSIGNED FOR HUMAN I-FABP

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	^{15}N	Others	Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	^{15}N	Others
A1	n.a.	n.a.	n.a.	n.a.		L30	7.20	4.10	1.75, 1.66	122.6	$C^{\gamma}H$ 1.45
F2	8.22	4.40	3.12, 1.47	n.a.	$C^{\delta}H$ 7.22						$C^{\delta}H$ 0.79, 0.61
					$C^{\epsilon}H$ 6.88	A31	8.48	4.14	1.55	124.0	
					$C^{\zeta}H$ 6.56	A32	7.41	4.11	1.33	120.2	
D3	8.35	4.77	2.61, 2.49	120.9		H33	7.52	4.90	3.31, 2.87	121.1	$C^{\delta 2}H$ 7.18
S4	8.20	4.21	3.25, 2.57	123.3							$C^{\epsilon 1}H$ 8.10
T5	8.05	5.09	3.92	120.0	$C^{\gamma}H$ 1.14	D34	7.27	3.99	2.94, 2.84	121.4	
W6	9.41	5.22	3.10, 2.85	131.9	$C^{\delta}H$ 6.79	N35	8.92	4.23	2.98, 2.70	121.9	$N^{\delta}H$ 7.32, 6.71
					$N^{\epsilon 1}H$ 9.82						N^{δ} 115.7
					$C^{\epsilon 3}H$ 7.61	L36	6.85	4.32	1.78, 1.65	122.3	$C^{\gamma}H$ 1.62
					$C^{\zeta 3}H$ 7.00						$C^{\delta}H$ 1.09, 0.86
					$C^{\eta 2}H$ 6.87	K37	9.37	5.58	1.85, 1.77	131.1	$C^{\gamma}H$ 1.57, 1.35
					$C^{\zeta 2}H$ 7.23						$C^{\delta}H$ 1.64, –
					$N^{\epsilon 1}$ 130.3						$C^{\epsilon}H$ 2.89, –
K7	9.91	5.43	1.83, –	126.7	$C^{\gamma}H$ 1.46, –	L38	9.30	5.56	1.56, 1.12	126.4	$C^{\gamma}H$ 1.48
					$C^{\delta}H$ 1.74, 1.65						$C^{\delta}H$ 0.85, 0.36
					$C^{\epsilon}H$ 2.95, –	T39	9.12	5.21	4.01	124.6	$C^{\gamma}H$ 1.13
V8	8.76	3.29	1.70	131.6	$C^{\gamma}H$ 0.87, 0.22	I40	8.58	4.93	2.18	132.0	$C^{\gamma}H_2$ 1.70, 0.93
D9	9.65	4.91	2.52, 2.43	131.3							$C^{\gamma}H_3$ 0.75
R10	7.76	4.72	1.79, 1.58	116.3	$C^{\gamma}H$ 1.28, –						$C^{\delta}H$ 0.54
					$C^{\delta}H$ 3.14, –	T41	8.69	4.59	3.86	125.0	$C^{\gamma}H$ 1.04
					$N^{\epsilon}H$ 7.21	Q42	8.97	4.15	1.63, 1.55	130.7	$C^{\gamma}H$ 1.73, 1.70
					N^{ϵ} 87.7						$N^{\epsilon}H$ 7.51, 6.30
S11	8.51	5.26	3.73, 3.68	117.0							N^{ϵ} 110.8
E12	9.41	4.71	1.79, –	126.3	$C^{\gamma}H$ 2.14, 2.03	E43	8.81	4.35	1.83, 1.67	132.6	$C^{\gamma}H$ 1.96, –
N13	8.27	4.41	3.14, 2.94	123.2	$N^{\delta}H$ 7.00, 6.89	G44	8.88	3.90, 3.53		120.7	
					N^{δ} 114.0	N45	8.74	4.80	3.10, 2.90	128.9	$N^{\delta}H$ 7.76, 6.90
Y14	8.81	5.29	2.10, –	128.2	$C^{\delta}H$ 6.89						N^{δ} 115.8
					$C^{\epsilon}H$ 6.54	K46	7.88	4.88	1.83, 1.66	124.1	$C^{\gamma}H$ 1.26, 1.09
D15	8.53	4.08	2.60, –	122.5							$C^{\delta}H$ 1.52, –
K16	7.94	4.01	1.83, 1.70	123.0	$C^{\gamma}H$ 1.56, 1.50						$C^{\epsilon}H$ 2.78, –
					$C^{\delta}H$ 1.73, –	F47	9.13	4.81	1.50, 1.07	130.0	$C^{\delta}H$ 6.89
					$C^{\epsilon}H$ 3.01, 2.98						$C^{\epsilon}H$ 6.56
F17	7.69	3.95	3.17, 3.07	124.3	$C^{\delta}H$ 6.91						$C^{\zeta}H$ 6.78
					$C^{\epsilon}H$ 6.83	T48	8.35	4.77	3.82	118.5	$C^{\gamma}H$ 0.98
					$C^{\zeta}H$ 6.79	V49	9.61	4.73	2.05	131.3	$C^{\gamma}H$ 0.84, –
M18	8.16	3.23	2.10, –	121.4	$C^{\gamma}H$ 1.59, 1.28	K50	9.14	4.94	1.94, 1.78	130.7	$C^{\gamma}H$ 1.43, 1.30
E19	8.34	3.77	1.85, –	123.6	$C^{\gamma}H$ 2.14, 2.01						$C^{\delta}H$ 1.63, –
K20	7.98	4.05	2.03, 1.88	126.3	$C^{\gamma}H$ 1.43, 1.29						$C^{\epsilon}H$ 2.90, –
					$C^{\delta}H$ 1.65, 1.56	E51	9.05	4.96	2.07, 1.93	133.3	$C^{\gamma}H$ 2.29, 2.03
					$C^{\epsilon}H$ 2.93, –	S52	8.98	5.50	3.90, 3.82	125.3	
M21	7.50	3.62	1.73, 1.47	118.4	$C^{\gamma}H$ 1.24, 1.07	S53	9.09	4.98	4.04, 3.84	124.7	
G22	7.67	4.00, 3.61		111.2		A54	9.12	4.05	1.02	123.2	
V23	7.37	3.66	1.37	124.5	$C^{\gamma}H$ 0.81, 0.63	F55	8.09	4.46	3.24, 2.66	115.0	$C^{\delta}H$ 7.13
N24	8.81	4.35	3.08, 2.57	131.4	$N^{\delta}H$ 7.51, 6.88						$C^{\epsilon}H$ 7.28
					N^{δ} 116.1						$C^{\zeta}H$ 7.06
I25	8.51	3.69	1.80	126.1	$C^{\gamma}H_2$ 1.40, 1.18	R56	7.44	4.67	1.70, –	116.2	$C^{\gamma}H$ 1.22, –
					$C^{\gamma}H_3$ 0.86						$C^{\delta}H$ 2.95, –
					$C^{\delta}H$ 0.67						$N^{\epsilon}H$ 6.66
V26	7.57	3.69	2.06	124.3	$C^{\gamma}H$ 0.98, 0.91						N^{ϵ} 86.7
K27	7.60	4.10	1.75, 1.47	122.0	$C^{\gamma}H$ 1.56, 1.38	N57	8.34	5.91	2.75, –	121.7	$N^{\delta}H$ 7.50, 6.88
					$C^{\delta}H$ 1.50, –						N^{δ} 116.6
					$C^{\epsilon}H$ 2.83, –	I58	8.93	4.88	1.67	120.6	$C^{\gamma}H_2$ 1.24, 1.08
R28	8.46	3.84	1.82, 1.71	122.4	$C^{\gamma}H$ 1.47, 1.26						$C^{\gamma}H_3$ 0.79
					$C^{\delta}H$ 3.11, 3.05						$C^{\delta}H$ 0.48
					$N^{\epsilon}H$ 7.21	E59	8.46	5.28	1.93, –	126.7	$C^{\gamma}H$ 2.04, –
					N^{ϵ} 87.7	V60	9.42	4.44	2.21	130.9	$C^{\gamma}H$ 1.19, –
K29	7.69	4.04	1.86, –	122.2	$C^{\gamma}H$ 1.40, –	V61	8.12	5.10	1.85	128.9	$C^{\gamma}H$ 0.92, 0.79
					$C^{\delta}H$ 1.60, –						
					$C^{\epsilon}H$ 2.91, –						

TABLE 1 (continued)

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	^{15}N	Others	Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	^{15}N	Others
F62	8.44	4.85	2.77, 1.53	125.1	$C^{\delta}H$ 6.46 $C^{\epsilon}H$ 6.89 $C^{\zeta}H$ 6.43	K94	8.84	5.02	1.55, –	124.1	$C^{\gamma}H$ 1.26, – $C^{\delta}H$ 1.56, – $C^{\epsilon}H$ 2.92, 2.83
E63	9.50	5.31	1.94, 1.82	123.8	$C^{\gamma}H$ 2.18, –	R95	8.70	4.75	1.73, –	126.3	$C^{\gamma}H$ 1.29, –
L64	9.12	4.81	1.07, 1.03	130.0	$C^{\gamma}H$ 1.51 $C^{\delta}H$ 0.90, –						$C^{\delta}H$ 2.90, 2.82 $N^{\epsilon}H$ 9.05
G65	9.23	4.25, 3.62		112.1							N^{ϵ} 87.8
V66	7.95	4.34	2.34	125.8	$C^{\gamma}H$ 1.10, 0.88	T96	8.48	4.05	4.28	120.6	$C^{\gamma}H$ 1.09
T67	8.67	4.31	3.87	132.8	$C^{\gamma}H$ 1.07	D97	8.89	4.30	2.85, 2.70	124.3	
F68	9.51	5.11	3.50, 3.38	129.1	$C^{\delta}H$ 7.61 $C^{\epsilon}H$ 6.99 $C^{\zeta}H$ 6.87	N98 G99	8.53 7.74	4.72 4.09, 3.86	3.08, 2.54	119.9 111.8	$N^{\delta}H$ 7.50, 5.51 N^{δ} 110.0
N69	8.33	5.61	2.74, –	120.0	$N^{\delta}H$ 7.35, 6.71 N^{δ} 114.9	N100	8.36	4.68	3.00, 2.62	122.8	$N^{\delta}H$ 6.67, 6.50 N^{δ} 110.9
Y70	8.89	4.65	3.00, 2.09	126.4	$C^{\delta}H$ 6.77 $C^{\epsilon}H$ 6.55	E101 L102	8.63 8.94	4.78 5.14	1.98, – 1.81, 1.64	123.2 128.8	$C^{\gamma}H$ 2.34, 2.16 $C^{\gamma}H$ 1.45
N71	7.72	5.28	2.23, –	127.7	$N^{\delta}H$ 7.70, 6.83 N^{δ} 116.7	N103	8.53	5.60	2.83, 2.53	127.5	$C^{\delta}H$ 0.91, 0.76 $N^{\delta}H$ 7.43, 7.34 N^{δ} 116.8
L72	9.00	3.95	2.03, –	125.2	$C^{\gamma}H$ 1.83 $C^{\delta}H$ 0.93, 0.66	T104	8.97	5.58	4.03	120.7	$C^{\gamma}H$ 0.74
A73	8.82	3.69	0.79	124.7		V105	8.11	5.40	1.96	130.3	$C^{\gamma}H$ 0.94, 0.80
D74	7.14	4.31	2.92, 1.99	117.1		R106	9.37	5.05	1.33, 1.18	128.0	$C^{\gamma}H$ 1.63, 1.56
G75	7.85	4.29, 3.41		112.1							$C^{\delta}H$ 2.93, –
T76	7.75	3.73	3.84	123.2	$C^{\gamma}H$ 0.94						$N^{\epsilon}H$ 6.11
E77	8.63	4.63	2.08, 1.93	133.1	$C^{\gamma}H$ 2.02, 1.83						N^{ϵ} 86.4
L78	9.20	5.26	1.70, 1.23	128.5	$C^{\gamma}H$ 1.53 $C^{\delta}H$ 0.60, 0.48	E107 I108	8.53 8.77	4.85 4.79	1.99, 1.83 1.64	124.3 128.3	$C^{\gamma}H$ 2.07, – $C^{\gamma}H_2$ 1.07, –
R79	8.50	5.36	1.51, –	122.6	$C^{\gamma}H$ 1.23, – $C^{\delta}H$ 3.09, 3.03 $N^{\epsilon}H$ 7.21 N^{ϵ} 87.7	I109	8.88	4.21	1.69	133.5	$C^{\gamma}H_3$ 0.73 $C^{\delta}H$ 0.48 $C^{\gamma}H_2$ 1.32, 0.95 $C^{\gamma}H_3$ 0.82
G80	8.81	4.04, 3.81		118.3							$C^{\delta}H$ 0.73
T81	6.74	4.54	3.69	109.6	$C^{\gamma}H$ 0.86	G110	9.00	3.99, 3.63		122.7	
W82	9.07	5.29	2.52, 2.43	124.0	$C^{\delta}H$ 4.89 $N^{\epsilon 1}H$ 10.37 $C^{\epsilon 3}H$ 7.00 $C^{\zeta 3}H$ 6.87 $C^{\eta 2}H$ 6.51 $C^{\zeta 2}H$ 6.71 $N^{\epsilon 1}$ 134.5	D111 E112 L113 V114 Q115	8.66 7.90 7.87 9.26 9.34	4.73 5.12 4.33 4.51 5.70	2.86, 2.76 1.98, – –0.85, 0.66 2.00 1.94, 1.91	127.8 122.1 127.4 132.0 133.6	$C^{\gamma}H$ 2.45, 2.20 $C^{\gamma}H$ 0.74 $C^{\delta}H$ –0.52, 0.02 $C^{\gamma}H$ 0.93, 0.79 $C^{\gamma}H$ 2.18, 2.16 $N^{\epsilon}H$ 7.49, 6.45
S83	9.02	4.77	3.82, 3.69	118.9							N^{ϵ} 110.9
L84	8.82	5.17	1.85, 1.35	129.5	$C^{\gamma}H$ 1.50 $C^{\delta}H$ 0.95, 0.88	T116 Y117	8.72 9.29	4.91 5.16	3.88 2.49, 1.81	124.2 129.6	$C^{\gamma}H$ 1.13 $C^{\delta}H$ 6.41
E85	8.93	4.58	1.81, 1.78	131.4	$C^{\gamma}H$ 2.09, 1.93						$C^{\epsilon}H$ 6.50
G86	8.94	3.97, 3.64		122.4		V118	9.14	5.31	2.03	124.1	$C^{\gamma}H$ 0.92, 0.89
N87	7.73	4.22	2.19, –	128.5	$N^{\delta}H$ 8.05, 6.81 N^{δ} 114.8	Y119	8.99	4.84	3.02, 2.56	131.4	$C^{\delta}H$ 6.88 $C^{\epsilon}H$ 6.78
K88	7.99	4.96	1.89, –	122.1	$C^{\gamma}H$ 1.64, 1.29 $C^{\delta}H$ 1.62, 1.56 $C^{\epsilon}H$ 2.93, –	E120 G121 V122	9.48 8.28 8.29	3.37 4.05, 3.50 4.16	1.29, 0.75 2.34	129.7 105.7 126.9	$C^{\gamma}H$ 1.88, 1.42 $C^{\gamma}H$ 0.94 0.88
L89	8.49	4.99	1.07, 0.76	127.1	$C^{\gamma}H$ 0.94 $C^{\delta}H$ –0.33, –0.48	E123 A124	8.48 9.22	5.67 5.07	1.99, 1.94 1.33	130.6 129.8	$C^{\gamma}H$ 2.16, –
I90	9.57	4.65	1.93	128.5	$C^{\gamma}H_2$ 1.46, 1.32 $C^{\gamma}H_3$ 0.86 $C^{\delta}H$ 0.84	K125	8.58	5.47	1.44, 1.41	119.3	$C^{\gamma}H$ 1.25, 0.94 $C^{\delta}H$ 1.52, 1.37 $C^{\epsilon}H$ 2.77, –
G91	9.78	4.72, 2.29		125.2		R126	9.14	4.73	1.79, 1.58	123.2	$C^{\gamma}H$ 1.96, 1.65
K92	7.70	4.70	1.38, –	131.0	$C^{\gamma}H$ 1.18, – $C^{\delta}H$ 1.49, – $C^{\epsilon}H$ 2.82, –						$C^{\delta}H$ 2.90, – $N^{\epsilon}H$ 6.05 N^{ϵ} 86.4
F93	8.39	5.03	2.86, 2.60	126.2	$C^{\delta}H$ 6.95 $C^{\epsilon}H$ 6.78 $C^{\zeta}H$ 6.56	I127	8.69	4.95	1.59	126.7	$C^{\gamma}H_2$ 1.32, – $C^{\gamma}H_3$ 0.85 $C^{\delta}H$ 0.72

TABLE 1 (continued)

Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	^{15}N	Others	Residue	NH	$C^{\alpha}H$	$C^{\beta}H$	^{15}N	Others
F128	10.02	5.22	3.23, 3.09	130.2	$C^{\delta}H$ 7.11 $C^{\epsilon}H$ 6.97 $C^{\zeta}H$ 7.07	K130	8.99	4.12	1.47, 1.33	127.1	$C^{\gamma}H$ 1.82, – $C^{\delta}H$ 1.29, – $C^{\epsilon}H$ 2.81, 2.72
K129	8.89	5.29	2.05, 1.89	122.4	$C^{\gamma}H$ 1.50, – $C^{\delta}H$ 1.73, – $C^{\epsilon}H$ 2.99, –	D131	8.09	4.47	2.34, 2.16	133.9	

All NMR measurements with I-FABP were performed at 37 °C in a buffer solution consisting of 20 mM KH_2PO_4, 0.05% NaN_3 and 5% D_2O at pH 6.5. The ^1H chemical shift values refer to external sodium 3-(trimethylsilyl)[2,2,3,3-$^2H_4$]propionate. The ^{15}N chemical shift values refer to external $^{15}NH_4Cl$. Resonances that have not been assigned are marked n.a.

buffer solution consisting of 20 mM KH_2PO_4, 0.05% NaN_3 and 5% D_2O at pH 6.5. For experiments in D_2O, the above buffer was replaced by a perdeuterated buffer (20 mM KD_2PO_4, 0.05% NaN_3) at 4 °C. This perdeuterated buffer was prepared from the protonated buffer (95% H_2O, 5% D_2O) by lyophilizing and redissolving in D_2O twice without readjusting the pH. HSQC spectra were collected within 15 min after changing the temperature of the sample from 4 to 37 °C. The 66 NH protons that were observed were considered to be involved in hydrogen bonds. The ^1H chemical shift values refer to external sodium 3-(trimethylsilyl)[2,2,3,3-$^2H_4$]propionate (Cambridge Isotope Laboratories). The ^{15}N chemical shift values refer to external 2.9 M $^{15}NH_4Cl$ in 1 M HCl.

The NMR data were collected on a Bruker DMX 500 MHz spectrometer with an Aspect Station computer using

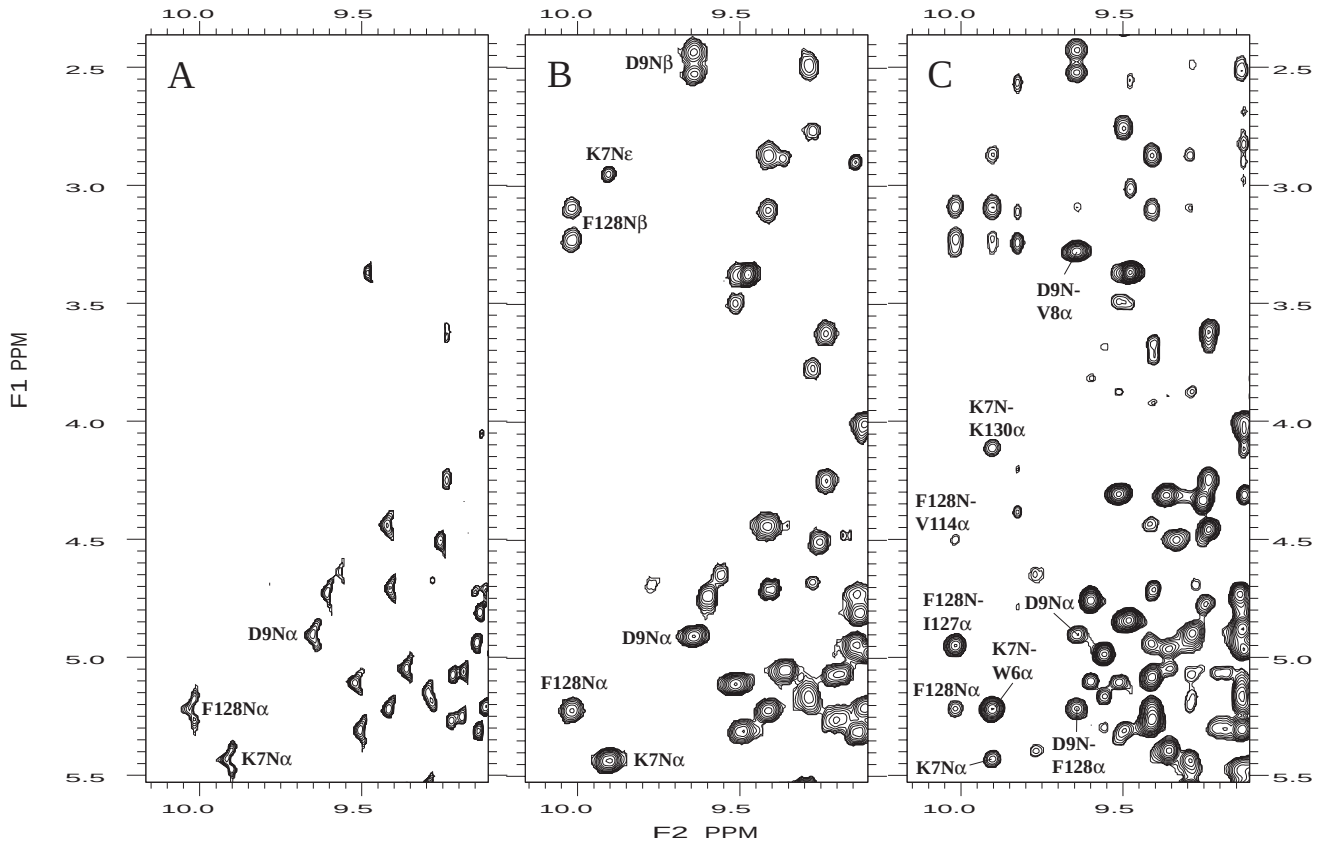

Fig. 1. Homonuclear 2D spectra of human I-FABP: COSY, TOCSY and NOESY, as shown in panels (A), (B) and (C), respectively. ^1H frequency 500 MHz, 20 mM potassium phosphate in H_2O with 5% D_2O, T = 37 °C, pH = 6.5, 512 increments. (A) Contour plot of part of a COSY-type spectrum obtained from a TOCSY experiment with very short (3 ms) spin-lock time. The assignments for the NH-$C^{\alpha}H$ connectivities of the spin systems F128, K7 and D9 are indicated as examples. (B) Contour plot of the same region as in panel (A) from a TOCSY spectrum with 80 ms spin-lock time. The assignments for the NH connectivities to the side-chain are again indicated for the spin systems F128, K7 and D9. (C) Contour plot of the same region as in panel (A) from a NOESY spectrum with a mixing time of 200 ms. Note that shorter mixing times were also used (data not shown). The strong NOE cross peaks between the NH resonances of F128, K7 and D9 and the $C^{\alpha}H$ resonances of the corresponding preceding amino acids, which were used for the sequential assignments, are indicated. The weaker NOE cross peaks from intraresidue NH-$C^{\alpha}H$ connectivities, as well as the NOE cross peaks from long-range interstrand NH-$C^{\alpha}H$ connectivities, are also indicated.

a 5 mm inverse triple resonance probe. The spectra were recorded in a phase-sensitive mode with time-proportional phase incrementation (TPPI). Quadrature detection was used with the carrier placed in the center of the spectrum on the water resonance. Selective presaturation was used for water suppression during the relaxation delay. In the NOESY, 2D HMQC-NOESY and 3D NOESY-HMQC experiments, water saturation was also applied during the mixing time. Decoupling of the ^{15}N nuclei during acquisition was achieved by a GARP sequence (Shaka et al., 1985).

$^{1}H,^{1}H$ total correlation spectroscopy (TOCSY) (Griesinger et al., 1988), nuclear Overhauser enhancement and exchange spectroscopy (NOESY) (Jeener et al., 1979), $^{1}H,^{15}N$ 2D HMQC-TOCSY (Lerner and Bax, 1986) and 2D HMQC-NOESY spectra (Shon and Opella, 1989) were obtained with 512 t_1 increments and 2048 t_2 data points. $^{1}H,^{1}H$ double-quantum filtered COSY (DQF-COSY) spectra (Piantini et al., 1980) were obtained with 1024 t_1 increments and 2048 t_2 data points. $^{1}H,^{15}N$ heteronuclear single-quantum coherence (HSQC) spectra (Bodenhausen and Ruben, 1980) were obtained with 512 t_1 increments and 1024 t_2 data points; the ^{1}H carrier frequency was set to the center of the amide region. $^{1}H,^{15}N$ 3D TOCSY-HMQC spectra (Marion et al., 1989b; Zuiderweg and Fesik, 1989) and 3D NOESY-HMQC spectra (Marion et al., 1989a) were obtained with 256, 64 and 1024 data points for $t_1(^{1}H)$, $t_2(^{15}N)$ and $t_3(^{1}H)$, respectively. Spin-lock times for TOCSY experiments were 80 ms for both 2D and 3D TOCSY-type experiments. COSY-type information was obtained by running a TOCSY experiment with a short (3 ms) spin-lock time. The mixing time for 3D NOESY-HMQC and 2D HMQC-NOESY experiments was set to 150 ms; several mixing times, ranging from 50 to 200 ms, were used for $^{1}H,^{1}H$ NOESY experiments.

The data sizes after Fourier transformation were $2048(^{1}H) \times 2048(^{1}H)$ for 2D $^{1}H,^{1}H$ homonuclear experiments (TOCSY, NOESY, etc.) and $1024(^{15}N) \times 2048(^{1}H)$ for 2D $^{1}H,^{15}N$ heteronuclear experiments (2D HMQC-TOCSY and 2D HMQC-NOESY). For $^{1}H,^{15}N$ 3D TOCSY-HMQC and 3D NOESY-HMQC, the data were Fourier transformed to $512(^{1}H) \times 64(^{15}N) \times 2048(^{1}H)$, and then the upper half of t_3 was discarded (all peaks are present on the lower half of t_3 only) to a final size of $512(^{1}H) \times 64(^{15}N) \times 1024(^{1}H)$.

The NMR data were processed with the Bruker

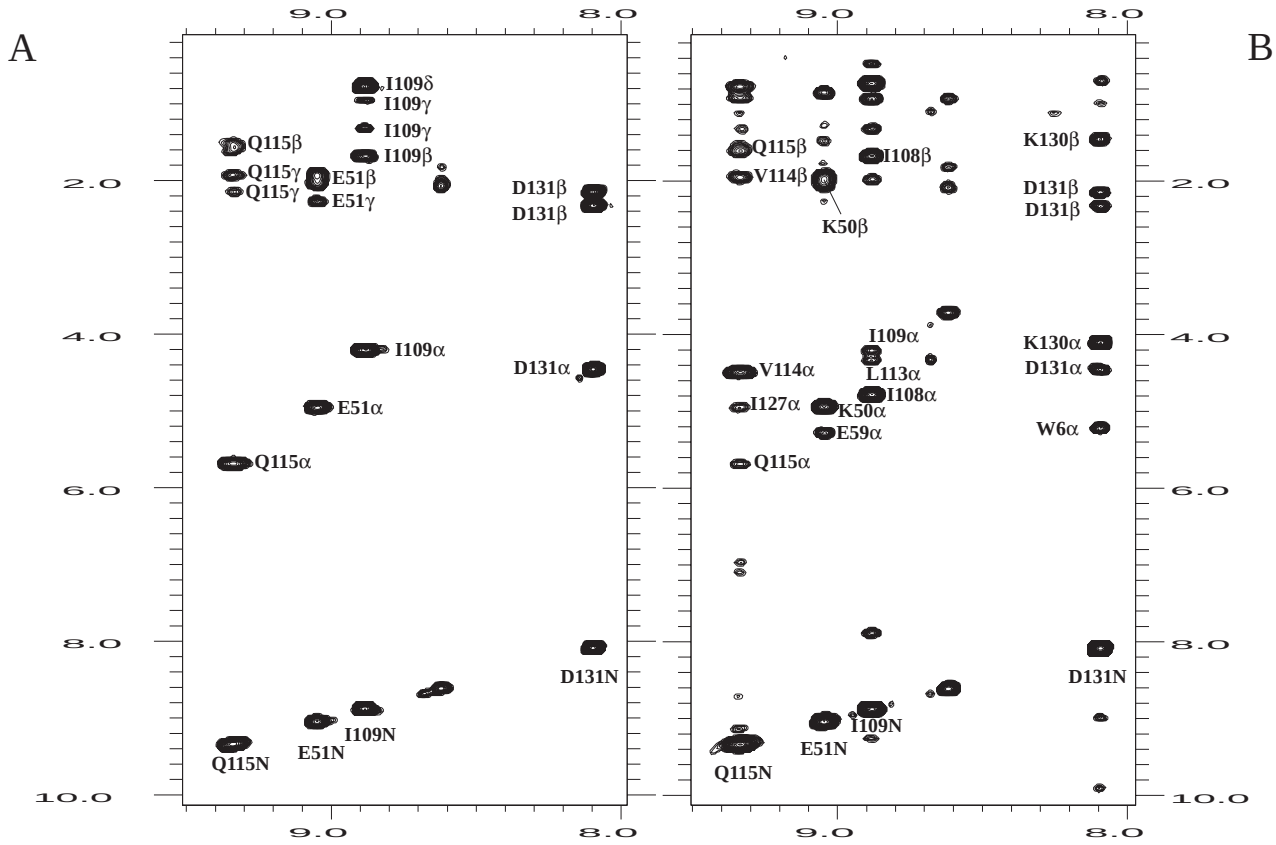

Fig. 2. 2D slices from 3D TOCSY-HMQC and NOESY-HMQC spectra of ^{15}N-labeled I-FABP. The experimental conditions were similar to those of the homonuclear 2D studies. Data were collected with $1024(^{1}H)$, $64(^{15}N)$ and $256(^{1}H)$ points for t_1, t_2 and t_3, respectively. (A) A contour plot of a 2D slice from the 3D TOCSY-HMQC with a spin-lock time of 80 ms. The assignments of the NH connectivities to $C^{\alpha}H$ and to all side-chain protons of the spin systems Q115, E51, I109 and D131 are indicated. (B) A contour plot of the same region as in panel (A) from the 3D NOESY-HMQC with a mixing time of 150 ms. The strong sequential NH-$C^{\alpha}H$, intraresidue NH-$C^{\alpha}H$ and long-range interstrand NH-$C^{\alpha}H$ connectivities are again indicated.

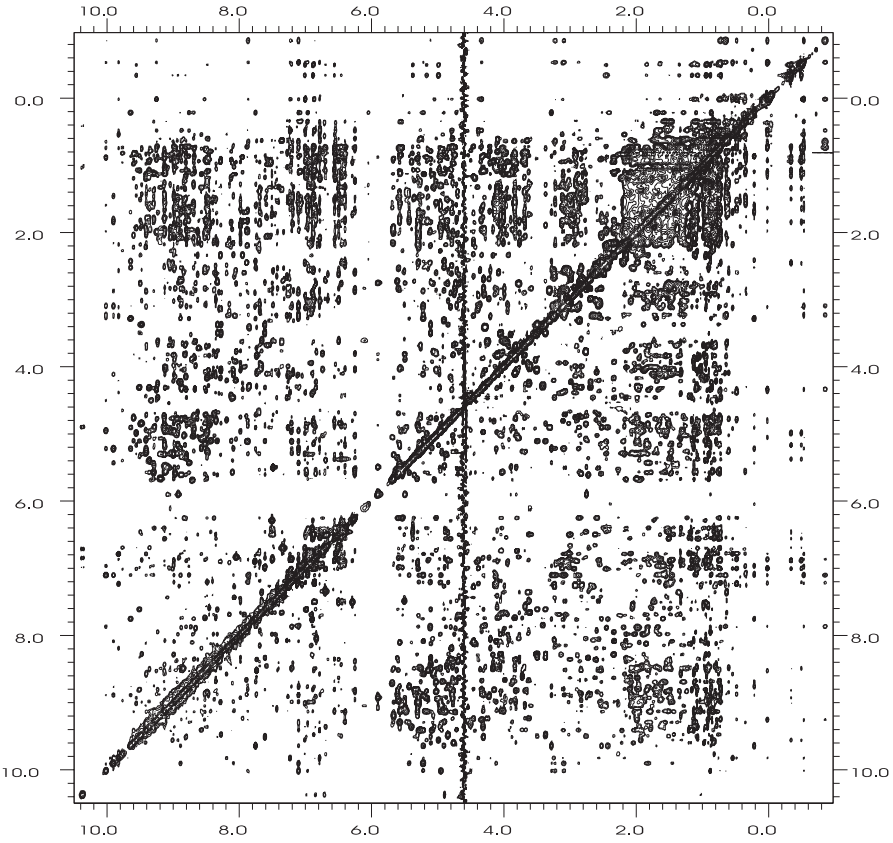

Fig. 3. Contour plot of the entire ^1H,^1H NOESY spectrum. ^1H frequency 500 MHz, 20 mM potassium phosphate in H_2O with 5% D_2O, T = 37 °C, pH = 6.5, mixing time 200 ms, 512 increments. The amide region of human I-FABP is well dispersed due to the high percentage of β-sheet structure, as seen in the NOESY spectrum for example. The strong NH-$C^\alpha H$ cross peaks in the fingerprint region are an indication of β-sheet structure. The strong cross peaks in the $C^\alpha H$-$C^\alpha H$ region are characteristic for antiparallel β-sheet structure.

UXNMR program. Peak picking and assignments were carried out by using NMRI and TRIAD (Tripos, St. Louis, MO, U.S.A.). The distance geometry calculations were performed using the DIANA 2.1 program from Peter Günthert (ETH, Zürich, Switzerland). Simulated annealing and subsequent energy minimization were carried out using SYBYL 6.2 (Tripos) as described previously for ILBP (Lücke et al., 1996).

Results

General strategy

The sequence-specific resonance assignments (Wüthrich, 1986) of human I-FABP were carried out by using homonuclear 2D DQF-COSY, TOCSY and NOESY, and heteronuclear 2D HMQC-TOCSY, 2D HMQC-NOESY, 3D TOCSY-HMQC and 3D NOESY-HMQC spectra. We have recently made complete ^1H sequential assignments for a closely related protein, porcine ILBP, based entirely on homonuclear 2D NMR spectra. This was possible because of extraordinary resolution of the ^1H homonuclear 2D NMR spectra obtained for ILBP. In the case of human I-FABP, we found significant overlaps in some regions of the ^1H homonuclear 2D NMR spectra. One important difference, in addition to the fact that I-FABP has four more residues than ILBP, was that the present studies were carried out at pH 6.5 because I-FABP turned out to be unstable at lower pH (strong precipitation), whereas ILBP remained fairly stable in acidic buffer and was studied at pH 5.0. The heteronuclear 3D TOCSY-HMQC and 3D NOESY-HMQC spectra displayed excellent separation in the amide proton region. Although the homonuclear 2D experiments had some overlapping peaks, they provided essential information about $C^\alpha H$ and side-chain protons. Moreover, the resolution in $t_1(^1H)$ for homonuclear 2D experiments was slightly better than that for heteronuclear 3D experiments because of the time limitations on 3D data collection.

The combination of heteronuclear 3D and homonuclear 2D experiments made it possible to obtain complete ^1H and ^{15}N assignments for the structure calculation of human I-FABP. First, correlation-type experiments were used to group the proton (and nitrogen) resonances and to assign them to different amino acid residue types (spin systems). Heteronuclear 3D TOCSY-HMQC and homonuclear TOCSY experiments provided scalar couplings between all protons of the same spin system. The homonuclear 2D DQF-COSY and the TOCSY with very short

spin-lock time provided through-bond connectivities for protons three bonds (or less) apart. Next, dipolar coupling interactions in heteronuclear 3D NOESY-HMQC and homonuclear 2D NOESY experiments were used to connect the spin systems sequentially. For structural calculations, the heteronuclear 3D NOESY-HMQC and homonuclear 2D NOESY data were converted into proton distance constraints. These data were used as input for distance geometry calculations to determine the tertiary structure of the protein. Finally, the structures were refined by simulated annealing and energy minimization in the presence of NOE distance constraints.

Sequential resonance assignments

The ^{1}H and ^{15}N sequential resonance assignments were completed for 130 of the 131 amino acid residues in human I-FABP, except the assignment for ^{15}N of residue Phe^{2}. The N-terminal residue (Ala^{1}) appeared to display a high degree of heterogeneity, because no NMR signal was observed for Ala^{1}. The ^{1}H and ^{15}N assignments of human I-FABP are listed in Table 1.

Figure 1 shows contour plots of part of the fingerprint region of the homonuclear 2D spectra. The assignments for spin systems F128, K7 and D9 were labeled as examples. In the TOCSY spectrum with very short spin-lock time (panel A), where only COSY-type connectivities are detected, the vicinal NH-C$^{\alpha}$H cross peaks have been marked. In the TOCSY spectrum with long spin-lock time (panel B), cross peaks between NH and side-chain protons of the same spin system are observable in addition to the vicinal NH-C$^{\alpha}$H cross peaks. As illustrated, the cross peaks F128 NH-C$^{\beta}$H, K7 NH-C$^{\epsilon}$H and D9 NH-C$^{\beta}$H are visible in this spectral region. In the NOESY spectrum (panel C), the vicinal NH(i)-C$^{\alpha}$H(i) connectivities for F128, K7 and D9 show relatively weak cross peaks, while the NH(i)-C$^{\alpha}$H(i–1) connectivities for these

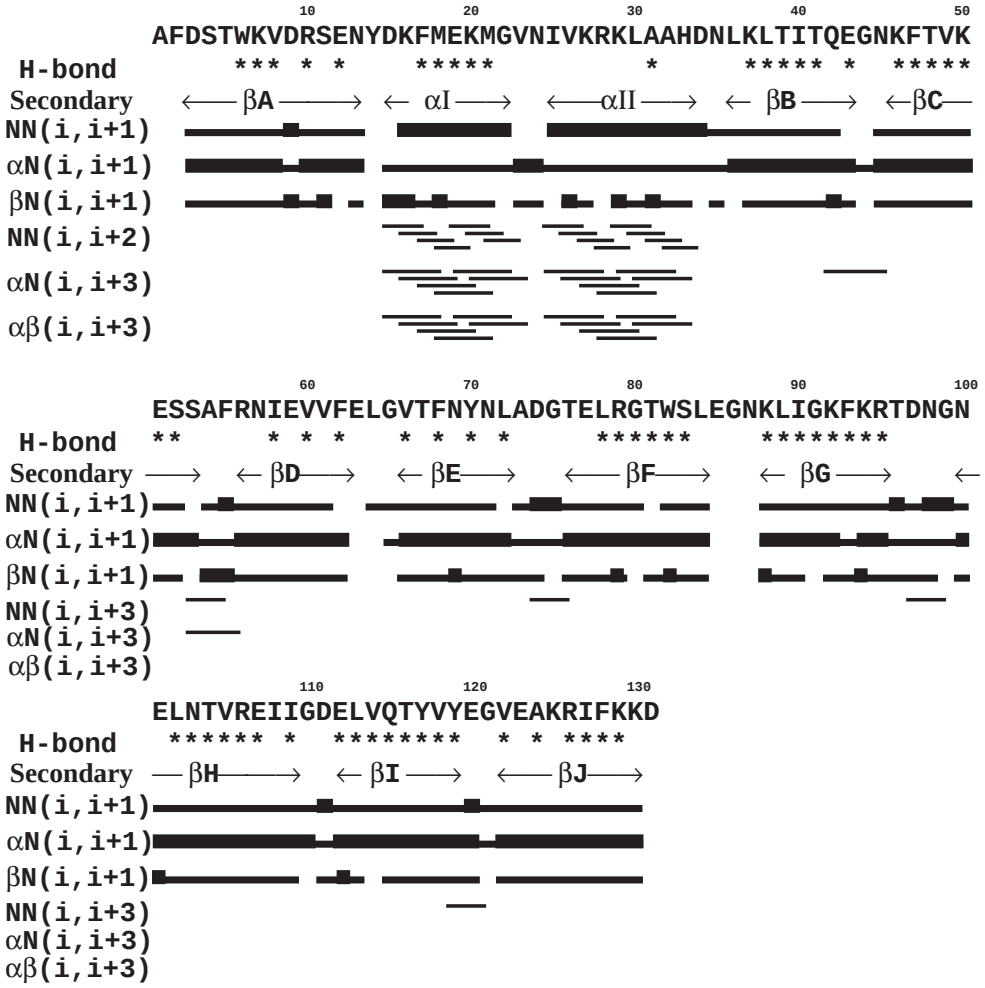

Fig. 4. Amino acid sequence and survey of strong and weak NOEs, used to establish the sequence-specific resonance assignments and to identify elements of regular secondary structure in I-FABP. Differences in the NOE intensities of the sequential $d_{\alpha N}$ (indicating β-strand structures), d_{NN} (indicating helical and turn structures) and $d_{\beta N}$ connectivities are represented by block height, with medium and weak NOEs both represented by shorter blocks. The $d_{\alpha N}(i,i+3)$, $d_{\alpha\beta}(i,i+3)$ and $d_{NN}(i,i+2)$ NOE connectivities, which are characteristic for segments of α-helical structure, are displayed by lines that start and end at the positions of the interacting residues. The slow-exchanging amide protons (hydrogen bond) and the secondary structure elements predicted by the NOE patterns are also indicated in the figure.

TABLE 2
STRUCTURAL STATISTICS FOR HUMAN I-FABP

Number of atoms	
Total	2121
Hydrogen	1079
Non-hydrogen	1042
Effective distance constraints	
Total	2519
Intraresidue	241
Interresidue sequential ($\|i-j\|=1$)	751
Interresidue medium-range ($2 \le \|i-j\| \le 4$)	358
Interresidue long-range ($\|i-j\| \ge 5$)	1169
Number of residual constraint violations	
Above 0.5 Å	0.0 ± 0.0
Between 0.5 and 0.4 Å	0.4 ± 0.7
Between 0.4 and 0.3 Å	0.3 ± 0.7
Between 0.3 and 0.2 Å	3.7 ± 2.5
Cartesian coordinate rmsd between the backbone heavy atoms (Å)	
Overall	1.09
β-Sheets	0.79
α-Helix I	1.15
α-Helix II	1.64
Portal[a]	1.67
Non-portal[a]	0.97
Energies (kcal/mol)[b]	
E_{total}	−10.6
E_{bond}	48.9
E_{angle}	355.6
$E_{torsion}$	428.9
E_{vdW}	−844.0

[a] The rmsd calculation for the portal region includes residues 24–33, 43–46 and 72–76. The rest of the residues (3–23, 34–42, 47–71 and 77–130) are included in the rmsd calculation for the non-portal region.

[b] All data are obtained from the 10 best DIANA conformations, except that the energies are calculated for the best DIANA conformation after simulated annealing with subsequent energy minimization in the presence of distance constraints.

residues display very strong cross peaks – a feature typical for β-strand structures. The cross peaks F128NH-V114C$^{\alpha}$H, K7NH-K130C$^{\alpha}$H and D9NH-F128C$^{\alpha}$H reflect the connectivities between neighboring β-strands in the antiparallel β-sheet structure. Figure 2 shows the same ^{15}N slice from both the 3D TOCSY-HMQC and the 3D NOESY-HMQC spectra. The connectivities for residues Q115, E51, I109 and D131 are indicated. The resonances at 8.65 ppm belong to an adjacent ^{15}N slice.

Secondary structure

The complete secondary structure of human I-FABP was determined from NOE distance constraints obtained from heteronuclear 3D NOESY-HMQC (Fig. 2B) and homonuclear NOESY (Fig. 3) spectra. The NOESY spectrum shows relatively good chemical shift dispersion. Overlaps in the NH region of the homonuclear spectra could be resolved in the 3D NOESY-HMQC spectra.

The large number of strong NH-C$^{\alpha}$H cross peaks as illustrated in Figs. 1C, 2B and 3 indicates that human I-FABP has a high content of β-strand structure. Furthermore, the abundance of strong C$^{\alpha}$H-C$^{\alpha}$H cross peaks shown in Fig. 3 demonstrates that these β-strands comprise an antiparallel β-sheet structure. A survey of sequential NOE connectivities within the human I-FABP sequence (Fig. 4) further indicates the presence and the location of regular secondary structure elements. Strong $d_{\alpha N}(i,i+1)$ connectivities and weak (or absent) $d_{NN}(i,i+1)$ connectivities, which are predominant throughout the human I-FABP amino acid sequence, are indications for β-strands. Strong $d_{NN}(i,i+1)$ and $d_{\alpha\beta}(i,i+3)$ as well as weaker $d_{\alpha N}(i,i+3)$ and $d_{NN}(i,i+2)$ connectivities were used to identify two short helices.

Figure 4 summarizes the NOE connectivities for the sequence-specific resonance assignments and identifies elements of regular secondary structure in I-FABP.

Figure 5 shows diagonal plots presenting the backbone NOE (panel A) and all NOE (panel B) connectivities throughout the I-FABP sequence. Helical structure elements are indicated by connectivity patterns parallel and adjacent to the diagonal, while antiparallel β-sheet structures are represented by connectivity patterns orthogonal to the diagonal. The two short α-helical segments already detected in Fig. 4 can also be observed in these diagonal plots. In addition, nine connectivity patterns representing antiparallel β-sheet structures formed by 10 β-strands become apparent in this type of representation. The distribution of backbone NOEs (Fig. 5A) indicates that there is a 'gap' between β-strands D and E. However, when all the NOE connectivities are plotted, as shown in Fig. 5B, the gap is not so evident, which suggests that the gap lacks long-range NOE connectivities between backbone protons, but displays multiple long-range NOEs involving side-chain protons.

Tertiary structure

A total of 3170 upper distance limit constraints were used as input for distance geometry calculations with the program DIANA. Of these, 3038 were NOE distance constraints obtained from the heteronuclear 3D NOESY-HMQC and homonuclear 2D NOESY data, and 132 were derived from 66 hydrogen bonds observed with D_2O exchange experiments. Among these 66 slowly exchanging NH protons, 60 are located in the β-sheets and six are in the helices. The hydrogen bonds in the β-sheets are assigned based on multiple NOE connectivities for antiparallel β-sheets. The hydrogen bonds in the helices are assigned to the α-type based on the observation of NOE connectivities for $d_{\alpha N}(i,i+4)$ that are unique for α-type helices. The NOE constraints were classified into three different upper distance limit categories, 2.5, 4.5 and 6.0 Å, corresponding to strong, medium and weak cross peaks. The DIANA program regarded 651 of these upper distance limit constraints as irrelevant, because they did

not restrict the distance between the two respective atoms. The remaining 2519 nontrivial upper distance limit constraints consisted of 241 intraresidue, 751 sequential, 358 medium-range (two to four residues apart) and 1169 long-range (five or more residues apart) distance constraints, as shown in Table 2.

Using the DIANA program, first a total of 50 initial conformations were calculated. Subsequently, the redundant dihedral-angle constraint (REDAC) strategy was applied. There was no significant improvement of the target function value after four cycles of repeated application of the REDAC procedure. The backbone structures of the 10 best conformations calculated with the DIANA program are shown superimposed in Fig. 6. In general, all structures exhibit good covalent geometry. A Ramachandran plot comprising all residues of the 10 structures is shown in Fig. 7A. The number within each data point corresponds to the model number from which that point is derived. The majority of the dihedral torsion angles lie within energetically favorable regions. Among the dihedral torsion angle values determined from the 10 best conformations, 75.6% are in most favored regions, 22.7% are in additional allowed regions and 1.7% are in generally allowed regions; no residue is in disallowed regions (Laskowski et al., 1993). Figure 7B shows the average rmsd values for the backbone heavy atoms (N, C, C^{α} and O) of the 10 best conformations as a function of residue number. The N-terminal residue (Ala^{1}) was not plotted because no resonance assignments were obtained for Ala^{1}.

The average rmsd between all backbone heavy atoms (N, C, C^{α} and O) of the 10 best conformations is 1.09 Å. In addition, Fig. 7B shows that the protein structure is overall well defined. The 10 β-strands are very well defined, with some segments displaying rmsd values smaller than 0.50 Å. The average rmsd between the backbone heavy atoms (N, C, C^{α} and O) of the 10 β-strands is 0.79 Å. As expected, the C- and N-terminal regions, the two α-helices, and the turns that connect the β-strands are less well defined compared to the 10 β-strands. Among these less defined regions, the helix αII and the turn connecting β-strands E and F show larger average rmsd values. The structure statistics, including the number of constraint violations, energies and rmsd values, are listed in Table 2.

Discussion

Structural features of human I-FABP

The DIANA conformation with the smallest target function, after simulated annealing and energy minimization in the presence of distance constraints, yielded a solution structure of I-FABP very similar to that of other fatty acid binding proteins, as shown in Fig. 8. The 10

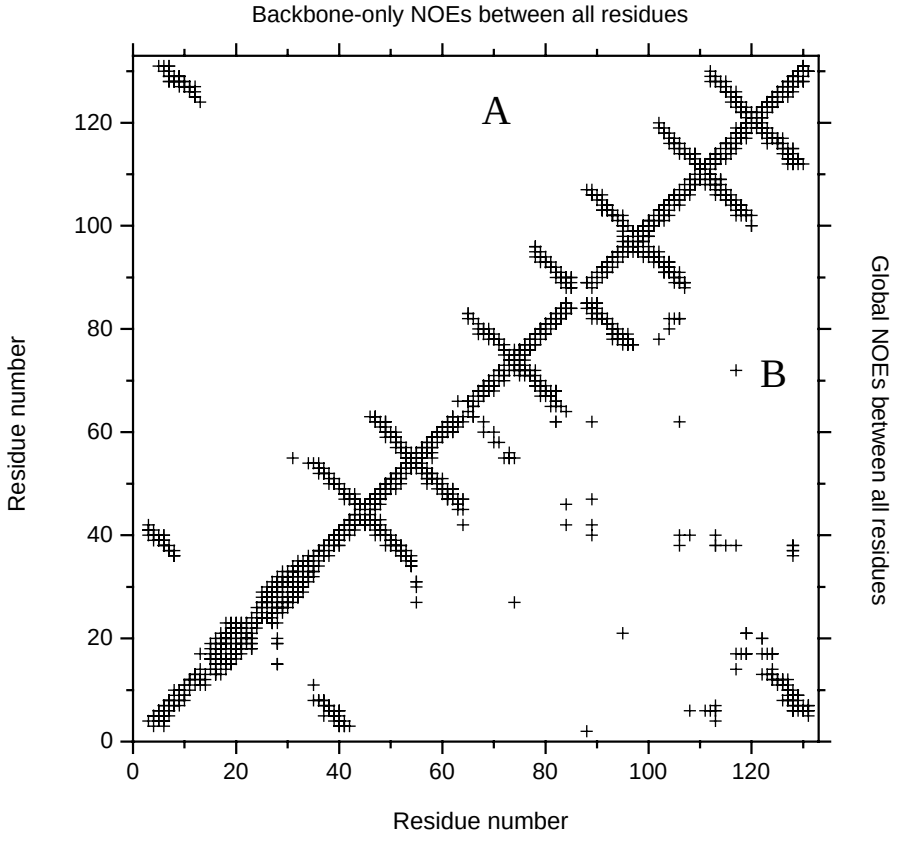

Fig. 5. Diagonal plots presenting the backbone NOE (A) and all NOE (B) connectivities throughout the I-FABP sequence.

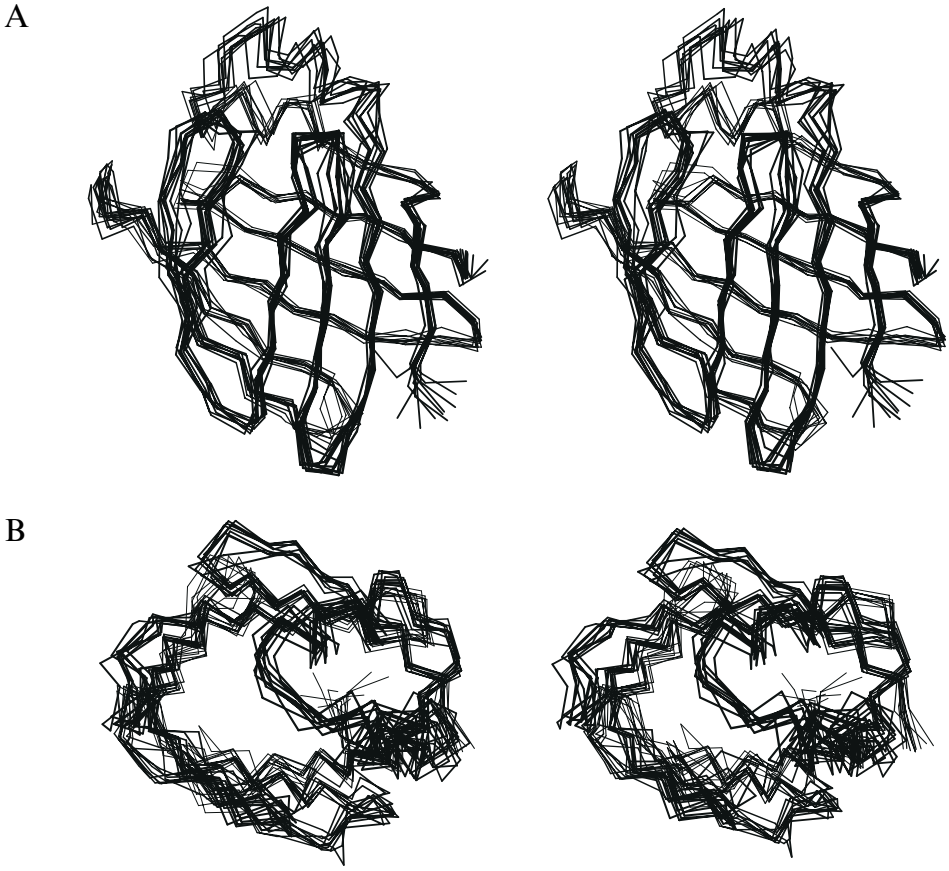

Fig. 6. Tertiary structure of human I-FABP obtained from distance geometry calculations with 2519 distance constraints that were derived from heteronuclear 3D NOESY-HMQC and homonuclear 2D NOESY data. The backbone structures of the 10 best DIANA conformations are shown superimposed as viewed from the side (A) and from the top (B). The putative binding cavity is evident in the top view. The average rmsd between the backbone heavy atoms (N, C, C^{α} and O) of these 10 best conformations, excluding the N-terminal residue, is 1.09 Å. (Produced with MOLSCRIPT; Kraulis, 1991.)

antiparallel β-strands (A–J), which are arranged in two nearly orthogonal β-sheets, form a β-clam structure closed on one side by two short α-helices (αI and αII). The opposite side of the β-clam is closed by side chains of mostly hydrophobic amino acids. There is a kink in the middle of β-strand A (at Asp^{9}) caused by a β-bulge conformation. The first half of β-strand A is part of the first β-sheet, while the second half of β-strand A is part of the second β-sheet. The same is true for β-strand F, which also serves as a vital link between the two β-sheets that form the β-clam structure. The hydrogen-bonding network within the β-sheet structure is continuous, except for a so-called gap between β-strands D and E. Despite the lack of main-chain hydrogen bonding (as shown in Fig. 5A), the gap is filled with the side chains of amino acids from β-strands D and E. There are 28 NOE connectivities between β-strands D and E, most of which involve at least one side-chain proton. It is believed that the ordered water molecules also present in the gap region help stabilize the protein structure. Xu et al. (1992) suggested that the gap is the result of a favored folding pathway, in which the intermediate folding steps lead to an open β-sheet, and the β-sheet then closes to a β-barrel but cannot form main-chain hydrogen bonds between β-strands D and E.

Figure 9 shows the backbone of human I-FABP with the side chains from different residue types highlighted. Most hydrophobic side chains (Fig. 9A) are located in the interior of the protein. The side chains for charged residues (Fig. 9B) are located primarily on the surface of the protein. The polar side chains (Fig. 9C) are more evenly distributed over the entire protein. A cavity is formed in the interior of I-FABP because of the loose packing of the protein. Even though in the NMR structure the internal cavity of human I-FABP is less well defined compared to the rat I-FABP X-ray structure, the overall features of the cavity are very similar. In the DIANA structure with the least violations of NOE distance constraints after simulated annealing and energy minimization (Fig. 8), the cavity for human I-FABP is lined by 25 amino acids (15 hydrophobic, 4 charged and 6 polar residues); all except for Ile^{58} and Asn^{71} are in the same position as in rat I-FABP. Especially notable is the fact that Phe^{55}, which was suggested to be important for fatty acid entry, as well as residues Arg^{106} and Arg^{126}, key amino acids for fatty

acid binding, are almost unchanged in their alignment compared to rat I-FABP. However, in rat I-FABP the cavity is lined by a total of 31 amino acids. This difference is probably due to slight variations in the orientation of side chains close to the cavity in the NMR conformations, which can cause large differences in the actual arrangement of side chains lining the cavity. For this same reason, the cavity of the human I-FABP is defined slightly different in all 10 conformations obtained from the NMR constraints.

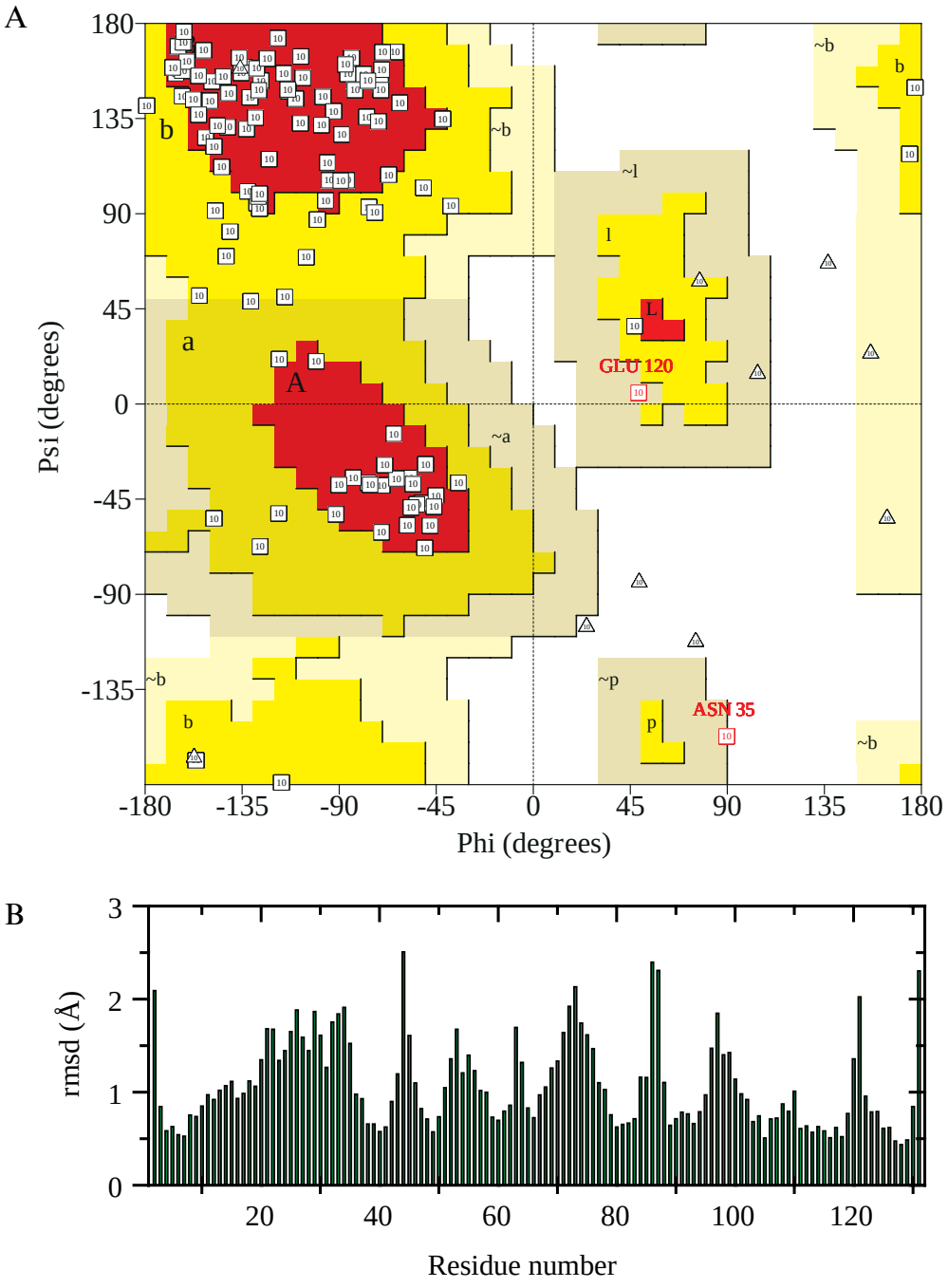

Fig. 7. Structural features of the 10 best DIANA conformations. (A) The Ramachandran plot shows the phi–psi torsion angles for all residues (except N- and C-terminal residues) in the 10 structures. The number within each data point corresponds to the model number from which that point is derived. The most favored regions (A, B, L), additional allowed regions (a, b, l, p) and generally allowed regions (~a, ~b, ~l, ~p) are based on an analysis of 118 protein X-ray structures with a resolution of at least 2.0 Å and an R-factor no greater than 20% (Laskowski et al., 1993). For a total of 10×131 residues in the 10 NMR conformations, 900 torsion angle values are in most favored regions, 270 are in additional allowed regions, 20 are in generally allowed regions and none are found in disallowed regions, while 20 residues are terminal residues and 100 are glycine residues shown as triangles. (Note that there is no proline in human I-FABP.) (B) Plot of the average rmsd per residue between the backbone heavy atoms (N, C, C^{α} and O) of the 10 best NMR conformations. The β-sheet structures are better defined with an average rmsd of 0.79 Å, compared to the overall rmsd for the entire protein structure of 1.09 Å (excluding the N-terminal residue).

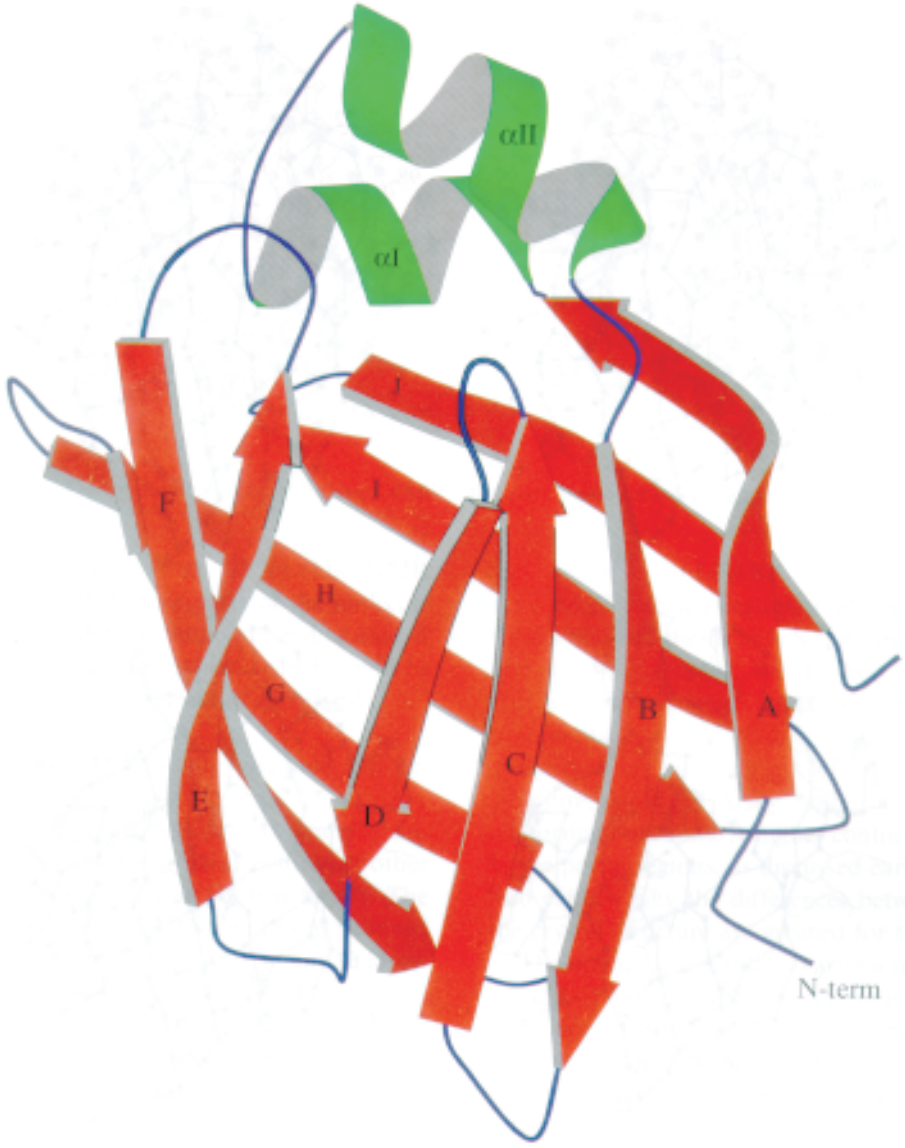

Fig. 8. A ribbon drawing showing the DIANA structure of human I-FABP with the least violations of the NOE distance constraints (lowest target function) after simulated annealing and subsequent energy minimization.

Overall, the solution structure of human I-FABP is well defined except for certain turns and certain backbone segments, as shown in Fig. 6 with the 10 superimposed NMR structures. Based on the packing density of atoms on the surface of the rat I-FABP, it was suggested that the fatty acids enter and leave the binding cavity through a portal region, which is formed by helix αII, the turn between β-strands C and D, and the turn between β-strands E and F (Sacchettini et al., 1992). These are precisely the regions that show higher variations in the NMR structures. From our NMR results, as shown in Fig. 7B, larger rmsd values (e.g. >1 Å) could be interpreted to be due either to a lack of NOE constraints that define the protein structure sufficiently or to a higher conformational variability inherent to that part of the structure. The residues from Met^{18} to Asn^{35}, which cover the C-terminal end of helix αI and all of helix αII, for example show rmsd values >1 Å between the best 10 NMR conformations. The same is true for the segments from Ser^{52} to Phe^{55} (turn between β-strands C and D) and from Tyr^{70} to Thr^{76} (turn between β-strands E and F). Since there are 37 and 18 NOE connectivities from helices αI and αII, respectively, to the β-barrel, these protein segments appear to actually display a higher conformational variability. Comparing the two helices, the differences in the number of slow-exchanging backbone amide protons (see Fig. 4) and the rmsd values (see Table 2) both seem to indicate a higher conformational variability for helix αII, possibly due to its involvement in the fatty acid entry portal. Similar results were observed in the other two NMR solution structures of FABP determined to date, bovine heart FABP (Lücke et al., 1992; Lassen et al.,

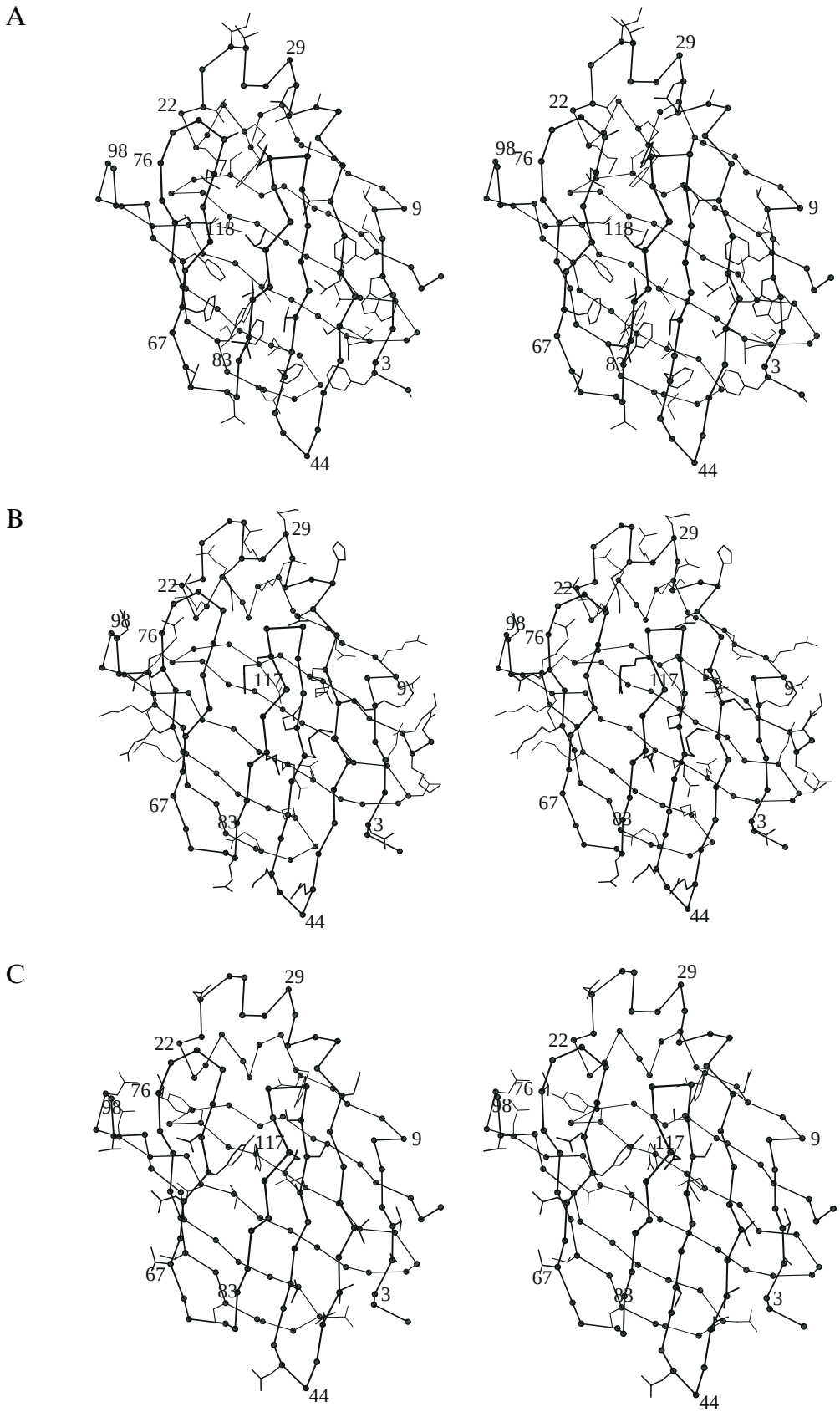

Fig. 9. Stereodiagrams of the backbone of human I-FABP with different side-chain types highlighted: (A) the hydrophobic side chains of all alanine, isoleucine, leucine, methionine, phenylalanine, tryptophan and valine residues; (B) the charged side chains of all arginine, aspartic acid, glutamic acid, histidine and lysine residues; (C) the polar side chains of all asparagine, cysteine, glutamine, serine, threonine and tyrosine residues.

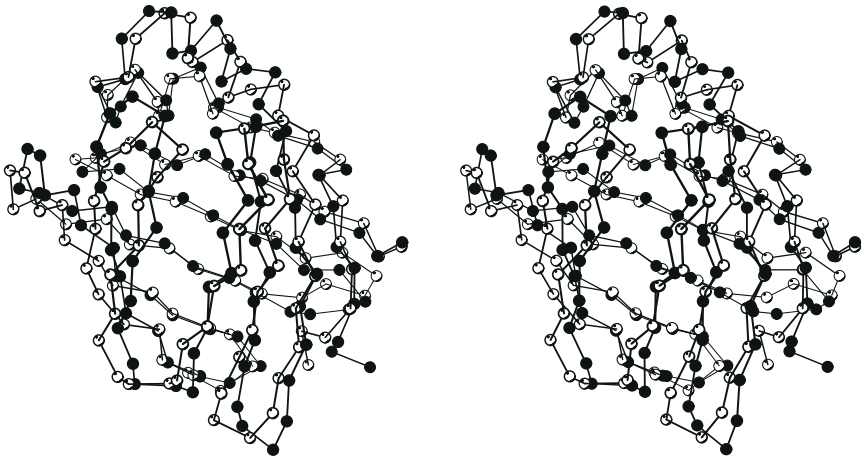

Fig. 10. Superposition of human I-FABP (filled circles) and rat I-FABP (open circles).

1993,1995) and ILBP (Lücke et al., 1996). The NMR results therefore suggest that these parts of the molecule may have some degree of dynamic mobility and/or consist of several different stable conformations, which could be important for the binding and release of fatty acids.

Comparison of human I-FABP solution structure with rat I-FABP X-ray structure

Determination of the NMR solution structure of human I-FABP allows comparison with structures of other members of the same lipid-binding protein family. The crystal structure of human I-FABP has not been reported to date. Rat I-FABP, which has an 82% amino acid sequence identity compared to human I-FABP, has been studied by NMR with ^{13}C- and ^{15}N-labeled protein (Hodsdon et al., 1995). The chemical-shift-derived secondary structure of rat I-FABP is very similar to our results for human I-FABP, but a 3D solution structure for rat I-FABP was not reported. However, the crystal structure of rat I-FABP as both a holo-protein and an apo-protein (with and without palmitic acid, respectively) has been determined to high resolution (Sacchettini et al., 1989, 1992; Scapin et al., 1992). Figure 10 shows the superimposition of human I-FABP (with endogenous *E. coli* fatty acids bound) and rat I-FABP (with palmitic acid bound). The rmsd values between the backbone heavy atoms (N, C, C^{α} and O) of the best human I-FABP conformation and the holo- and apo-forms of the rat I-FABP crystal structure are 1.82 Å and 1.86 Å, respectively. When compared to the ensemble of the 10 best NMR conformations of human I-FABP, the backbone heavy atoms of rat I-FABP (with palmitic acid bound) show an average rmsd of 1.69 Å. Hence, there is a very close correlation between the human and rat I-FABP structures. In contrast, a comparison of the human I-FABP and porcine ILBP (Lücke et al., 1996) NMR structure shows weak correlation. The structures of the best defined regions (most β-strands and some of the turns) of human I-FABP are very similar to rat I-FABP. Those regions that differ from rat I-FABP are the regions that also show larger rmsd values between the 10 best NMR conformations, i.e. they already differ among the 10 best solution structures. Again, the larger rmsd values for certain segments of the protein structure as observed among the 10 best NMR conformations may reflect the dynamics of the solution structure, while the X-ray study only detects one single conformational state. A higher conformational variability in the portal regions, as discussed earlier, could therefore also explain why the differences between the NMR and the crystal structure are greatest for this portal region.

The amino acid residue at position 54 is an alanine in human I-FABP compared to an asparagine in rat I-FABP. They are both located in the turn between β-strands C and D, which is part of the portal region. The alignments of Ala^{54} (human) and Asn^{54} (rat) within the protein structure are very similar. Asn^{54} has a longer side chain, which points towards the end of helix αII. There are six NOE connectivities between Ala^{54} and Asp^{34} in human I-FABP. The Thr^{54} to Ala^{54} substitution in populations such as the Pima Indians will probably affect the interaction pattern between residue 54 and the residue(s) at the end of helix αII. This change could affect the orientation and/or dynamics of helix αII, which in turn might affect the binding and transport of fatty acid ligands.

Biological implications

Knowledge of the solution structure of I-FABP will yield new insights into the mechanism of fatty acid binding, transport and release in enterocytes. The rates at which these events occur appear to have significant physiological consequences. Free fatty acids normally provide an alternative fuel source to glucose for energy during periods of fasting. Biochemical studies have indicated a direct relationship between concentrations of circulating free fatty acids and insulin resistance (Boden et al., 1994). Interestingly, the Ala^{54} to Thr^{54} substitution in human I-

FABP, which affects fatty acid binding as well as fatty acid transport across enterocyte-mimicking cells, has been shown to be associated with insulin resistance in Pima Indians (Baier et al., 1995). A structural comparison between the Ala^{54} and Thr^{54} forms of human I-FABP will provide the molecular basis for the observed physiological differences between individuals with an Ala^{54} or Thr^{54} genotype.

Acknowledgements

This work has been supported by the Welch Foundation and by NIH Grants GM45859 and HL26335 to J.C.S. and J.A.H., respectively, and an NIH cardiovascular training grant to F.Z.

References

Baier, L.J., Sacchettini, J.C., Knowler, W.C., Eads, J., Paolisso, G., Tataranni, P.A., Mochizuki, H., Bennett, P.H., Bogardus, C. and Prochazka, M. (1995) *J. Clin. Invest.*, **95**, 1281–1287.

Baier, L.J., Bogardus, C. and Sacchettini, J.C. (1996) *J. Biol. Chem.*, **271**, 10892–10896.

Banaszak, L.B., Winter, N., Xu, Z., Bernlohr, D.A., Cowan, S. and Jones, T.A. (1994) *Adv. Protein Chem.*, **45**, 89–151.

Bass, N.M. (1993) *Mol. Cell. Biochem.*, **123**, 191–202.

Boden, G., Chen, X., Ruiz, J., White, J.V. and Rossetti, L. (1994) *J. Clin. Invest.*, **93**, 2438–2446.

Bodenhausen, G. and Ruben, D.J. (1980) *Chem. Phys. Lett.*, **69**, 185–191.

Glatz, J.C. and Van der Vusse, G.J. (1990) *Mol. Cell. Biochem.*, **98**, 247–251.

Griesinger, C., Otting, G., Wüthrich, K. and Ernst, R.R. (1988) *J. Am. Chem. Soc.*, **110**, 7870–7872.

Hodsdon, M.E., Toner, J.J. and Cistola, D.P. (1995) *J. Biomol. NMR*, **6**, 198–210.

Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.*, **71**, 4546–4553.

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

Kaikaus, R.M., Bass, N.M. and Ockner, R.K. (1990) *Experientia*, **46**, 617–630.

Kraulis, P.J. (1991) *J. Appl. Crystallogr.*, **24**, 946–950.

Laskowski, R.A., MacArthur, M.W., Moss, D.S. and Thornton, J.M. (1993) *J. Appl. Crystallogr.*, **26**, 283–291.

Lassen, D., Lücke, C., Kromminga, A., Lezius, A., Spener, F. and Rüterjans, H. (1993) *Mol. Cell. Biochem.*, **123**, 15–22.

Lassen, D., Lücke, C., Kveder, M., Mesgarzadeh, A., Schmidt, J.M., Specht, B., Lezius, A., Spener, F. and Rüterjans, H. (1995) *Eur. J. Biochem.*, **230**, 266–280.

Lerner, L. and Bax, A. (1986) *J. Magn. Reson.*, **69**, 375–380.

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

Lücke, C., Lassen, D., Kreienkamp, H.-J., Spener, F. and Rüterjans, H. (1992) *Eur. J. Biochem.*, **210**, 901–910.

Lücke, C., Zhang, F., Rüterjans, H., Hamilton, J.A. and Sacchettini, J.C. (1996) *Structure*, **4**, 785–800.

Marion, D., Driscoll, P.C., Kay, L.E., Wingfield, P.T., Bax, A., Gronenborn, A.M. and Clore, G.M. (1989a) *Biochemistry*, **28**, 6150–6156.

Marion, D., Kay, L.E., Sparks, S.W., Torchia, D.A. and Bax, A. (1989b) *J. Am. Chem. Soc.*, **111**, 1515–1517.

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

Müller-Fahrnow, A., Egner, U., Jones, T.A., Rüdel, H., Spener, F. and Saenger, W. (1991) *Eur. J. Biochem.*, **199**, 271–276.

Piantini, U., Sørensen, O.W. and Ernst, R.R. (1980) *J. Am. Chem. Soc.*, **104**, 6800–6801.

Sacchettini, J.C., Gordon, J.I. and Banaszak, L.J. (1989) *J. Mol. Biol.*, **208**, 327–339.

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

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

Shaka, A.J., Barker, P.B. and Freeman, R. (1985) *J. Magn. Reson.*, **64**, 547–552.

Shon, K. and Opella, S.J. (1989) *J. Magn. Reson.*, **71**, 379–383.

Veerkamp, J., Peeters, R.A. and Maatman, R.G.H.J. (1991) *Biochim. Biophys. Acta*, **1081**, 1–24.

Veerkamp, J. and Maatman, R.G.H.J. (1995) *Prog. Lipid Res.*, **34**, 17–52.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY, U.S.A.

Xu, Z., Bernlohr, D.A. and Banaszak, L.J. (1992) *Biochem. J.*, **31**, 3484–3492.

Young, A.C.M., Scapin, G., Kromminga, A., Patel, S.B., Veerkamp, J.H. and Sacchettini, J.C. (1994) *Structure*, **2**, 523–534.

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

Zuiderweg, E.R.P. and Fesik, S.W. (1989) *Biochemistry*, **28**, 2387–2391.