Fatty-Acid-Binding Protein from the Flight Muscle of *Locusta migratoria*: Evolutionary Variations in Fatty Acid Binding[†]

Christian Lücke,[‡] Ye Qiao,[§] Herman T. B. van Moerkerk,[||] Jacques H. Veerkamp,[||] and James A. Hamilton*[,§]

Max Planck Research Unit for Enzymology of Protein Folding, Weinbergweg 22, D-06120 Halle (Saale), Germany, Department of Physiology and Biophysics, Boston University School of Medicine, 715 Albany Street (W302), Boston, Massachusetts 02118-2526, and Department of Biochemistry, University of Nijmegen, Geert Groteplein 30, 6525 GA Nijmegen, The Netherlands

Received February 3, 2006; Revised Manuscript Received March 14, 2006

ABSTRACT: Intracellular lipid-binding proteins have evolved from a common ancestral gene with the appearance of mitochondrial oxidation, to guarantee, for example, transport of fatty acids through the aqueous cytosol to their site of utilization. The mammalian forms of these lipid carriers are structurally well-characterized and have been categorized, on the basis of sequence similarities and several typical ligand-binding features, into four subfamilies. Only a single complex structure of an invertebrate fatty-acid-binding protein (FABP) has been reported to date, which reveals a unique ligand-binding arrangement yet unknown in vertebrate FABPs. In the present study, the structure of a second invertebrate FABP (locust muscle) complexed with a fatty acid has been determined on the basis of intermolecular NOE connectivities between the protein and the uniformly ^{13}C-enriched oleate ligand. The resulting ligand conformation, although resembling the closely related mammalian heart- and adipocyte-type FABPs, is characterized by certain binding features that differ significantly from the typical hairpin-turn ligand shapes of the latter forms. This is primarily due to an alanine-to-leucine substitution in locust FABPs that produces a steric hindrance for ligand binding. A comparison with an FABP from tobacco hornworm larvae furthermore demonstrates that certain amino acid substitutions that appear to be specific for invertebrates decidedly influence the binding arrangement inside the protein cavity. Hence, as a result of these evolutionary variations, invertebrate FABPs may display a much greater diversity in intracellular lipid binding than observed for the mammalian transport proteins, thus possibly providing new insights for the design of modified lipid carriers.

The intracellular lipid-binding proteins (i-LBPs)[1] began to evolve from a common ancestral gene nearly 1 billion years ago (*1*), prior to the vertebrate/invertebrate split that took place approximately 650 million years ago. They comprise a family of cytosolic 14−16 kDa proteins of very similar three-dimensional fold (*2*, *3*). The primary function of these lipid carriers has been presumed to be the efficient transport of lipids that are poorly soluble in water (e.g., fatty acids, bile acids, or retinoids) through the cytosol to the different cellular compartments for utilization or storage. For example, the fatty acid transport to mitochondria for β oxidation became a major requirement for energy production in eukaryotes (*4*). Numerous other functional roles have been suggested for FABPs based on biophysical data, physiological disorders, and more recently various studies using null mutations for specific FABPs in mice (*4*−*7*); however, some of these associations or functions, such as, for example, (i) protection of polyunsaturated fatty acids, (ii) regulation of the postprandial lipid metabolism, or (iii) influences on insulin resistance and atherosclerosis, may not be relevant to insects.

The flight muscle tissues of migratory birds and insects contain unusually high levels of fatty-acid-binding protein (FABP). Approximately 18% of the total protein content in the flight muscle of desert locust (*Schistocerca gregaria*), for example, is an FABP referred to as *Sg*-FABP (*8*). The migratory locust (*Locusta migratoria*) carries a homologous protein, *Lm*-FABP, whose primary structure differs from *Sg*-FABP in only three amino acid positions (*9*). Both *Sg*- and *Lm*-FABP are closely related to the mammalian i-LBP *subfamily IV* (*3*, *10*), in particular, to the heart (H-) and adipocyte (A-) FABP forms with 44 and 43% sequence identity (65 and 64% sequence similarity), respectively. It

[†] This work has been supported by NIH Grant HL26335 to J.A.H.

* To whom correspondence should be addressed: Department of Physiology and Biophysics, Boston University School of Medicine, 715 Albany Street, W302, Boston, MA 02118-2526. Telephone: (617) 638-5048. Fax: (617) 638-4041. E-mail: jhamilt@bu.edu.

[‡] Max Planck Research Unit for Enzymology of Protein Folding.

[§] Boston University School of Medicine.

[||] University of Nijmegen.

[1] Abbreviations: A-FABP, adipocyte-type FABP; B-FABP, brain-type FABP; EM, energy minimization; FABP, fatty-acid-binding protein; H-FABP, heart-type FABP; HSQC, heteronuclear single-quantum correlation; i-LBP, intracellular lipid-binding protein; L-FABP, liver-type FABP; *Lm*-FABP, *Locusta migratoria* FABP; MD, molecular dynamics; *Ms*-FABP, *Manduca sexta* FABP; NOESY, nuclear Overhauser and exchange spectroscopy; *Sg*-FABP, *Schistocerca gregaria* FABP; TOCSY, total correlation spectroscopy.

10.1021/bi060224f CCC: $33.50
Published on Web 04/28/2006

was consequently proposed that the unusually high FABP content in migratory insects such as locust ensures efficient fatty acid utilization for sustained flight activity (*11*). Developmental changes in FABP expression, concentration, and intracellular distribution are further indications for an important role of this lipid carrier in the energy production (*8*, *12*). The cellular H-FABP levels have also been linked to energy consumption in vertebrates: in the flight muscle tissues of migratory birds, such as the barnacle goose (*Branta leucopsis*) and the western sandpiper (*Calidris mauri*), H-FABP concentrations vary with the migration season and miscellaneous developmental stages (*13*−*15*). Moreover, the extreme fatigue exhibited by H-FABP knockout mice after prolonged physical activity (*16*) is yet another strong evidence that these types of FABPs play a crucial role as intracellular lipid transporters to enhance fatty acid metabolism inside muscle cells.

The i-LBPs have been studied extensively in the last 20 years. A comparison of the sequential similarities and ligand-binding features resulted in a division of mammalian i-LBPs into four subfamilies (*10*). *Subfamily I* contains the intracellular retinoid-binding proteins; *subfamily II* contains the liver-type (L-) FABPs and ileal lipid-binding protein; and *subfamily III* contains the intestinal FABP. *Subfamily IV*, finally, consists of a larger variety of different FABP types; they all exhibit (i) a typical U-shaped bend in the ligand conformation, with the fatty acid carboxylate group coordinated to the side chains of one tyrosine and two arginine residues, and (ii) a short 3_{10}-helical loop at the N terminus of the protein. Moreover, this latter subfamily also exhibits a rather large cluster of internal water molecules, forming an extensive hydrogen-bond network inside the protein cavity that increases both protein stability and ligand-binding affinity (*17*).

The amino acid sequences of FABPs from invertebrates show about 25−47% sequence similarity compared to those from vertebrates (*1*). To date, three-dimensional structures of invertebrate FABP are known only for *Sg*-FABP and an FABP form (*Ms*-FABP) isolated from the midgut of tobacco hornworm (*Manduca sexta*) larvae (*18*, *19*). Interestingly, the reported orientation of the fatty acid ligand inside the binding cavity of *Ms*-FABP is unlike any of the conformational arrangements found in mammalian FABP. The crystal structure of the locust muscle protein (i.e., *Sg*-FABP), on the other hand, represents merely the apo form. In light of the fact that most FABP types had evolved prior to the separation of vertebrates from invertebrates, the question remains, how fatty acid binding by insect FABPs differs from the characteristic binding features found in mammalian FABPs.

We previously characterized the three-dimensional fold of *Lm*-FABP in solution using high-resolution nuclear magnetic resonance (NMR) spectroscopy (*3*). In the present study, we describe the conformation of the oleate ligand bound to *Lm*-FABP, on the basis of intermolecular distance restraints obtained from NOE contacts between the uniformly ^{13}C-enriched ligand and the protein. The conformational features of the resulting complex structure were compared with those of other invertebrate and vertebrate FABPs such as *Ms*-FABP, human H-FABP, and murine A-FABP. We find that certain amino acid substitutions that are unique to invertebrate FABPs clearly distinguish their ligand-binding arrangements from the corresponding mammalian proteins, thus suggesting a possibly greater conformational diversity for lipid molecules bound to the invertebrate FABP forms.

MATERIALS AND METHODS

Expression and Purification. Lm-FABP was expressed and purified as previously described (*9*). The recombinantly expressed protein (15 064 Da molecular weight) included an additional methionine residue at the N terminus (referred to as Met0). The endogenous ligand was removed by slowly eluting the protein with 20 mM phosphate buffer (pH 7.4, 0.05% sodium azide) at 37 °C through a Lipidex 5000 column (*20*). Next, 4.1 μL of pure, uniformly ^{13}C-labeled oleic acid were dissolved in 180 μL of KOH (0.1 M) and subsequently added stepwise, in small aliquots and under stirring, to the dilute (15 mL) protein solution, to avoid extreme changes in the local pH. This mixture was kept at room temperature for 24 h before washing with additional buffer in centrifugal concentrators for the removal of excess lipid.

NMR Measurements. The NMR sample with a protein concentration of 2 mM was prepared in 20 mM phosphate buffer (90:10 H_2O/D_2O, v/v) at pH 5.5. A Bruker DMX 500 MHz spectrometer with a 5 mm inverse triple-resonance probe was used to carry out all NMR experiments. The NMR data were collected at 35 °C in a phase-sensitive mode, implementing time-proportional phase incrementation (TPPI) for quadrature detection. One-dimensional (1D) ^{13}C spectra were acquired with carbon−proton decoupling but without carbon−carbon decoupling. Two-dimensional ^{1}H/^{1}H total correlation spectroscopy (TOCSY) (80 or 6.4 ms spinlock time) and ^{1}H/^{1}H nuclear Overhauser and exchange spectroscopy (NOESY) (150 ms mixing time) spectra were employed for the sequential resonance assignment according to the classical strategy via nuclear Overhauser effect (NOE) connectivities (*21*). ^{1}H/^{13}C heteronuclear single-quantum correlation (HSQC) and ^{13}C-edited 3D NOESY−HSQC (150 ms mixing time; 2048 × 68 × 256 time-domain data sizes in the ω_3 [^{1}H], ω_2 [^{13}C], and ω_1 [^{1}H] dimensions, respectively) experiments made use of pulsed field gradients for coherence selection and artifact suppression, as well as gradient sensitivity enhancement schemes wherever appropriate (*22*, *23*). Quadrature detection in the indirectly detected dimensions was achieved by either the states TPPI (*24*) or the echo/antiecho (*23*) method. All chemical-shift values were referenced to external sodium 2,2-dimethyl-2-silapentane-5-sulfonate (Cambridge Isotope Laboratories, Andover, MA) to ensure consistency among all spectra (*25*).

For processing, the 1D ^{13}C data were multiplied by an exponential function with a line-broadening factor of 3 prior to Fourier transformation. The spectral data were processed on a Silicon Graphics O2 workstation using the XWIN-NMR 2.6 software package (Bruker, Rheinstetten, Germany). In case of the 2D and 3D experiments, a $\pi/2$ phase-shifted squared sine-bell function was used for apodization in all dimensions. In the 3D spectra, forward linear prediction, to extend the time-domain data, and zero-filling were applied in the indirectly detected dimensions. A polynomial baseline correction was employed wherever necessary. The final matrices consisted of 2048 × 2048 and 2048 × 128 × 512 real data points in the 2D and 3D spectra, respectively. Peak

picking and data analysis of the transformed spectra were performed using AURELIA 2.5.9 (Bruker, Rheinstetten, Germany).

Structure Calculation. For the structure elucidation of the *Lm*-FABP/oleate complex, the X-ray coordinates of *Sg*-FABP (PDB code 1FTP, chain B) were used. All following steps were performed with the INSIGHT 2000 software package (Accelrys, Inc., San Diego, CA), which includes the DISCOVER module for energy minimization (EM) and molecular dynamics (MD) calculations. First, charges were introduced for all Asp^-, Glu^-, Arg^+, Lys^+, and His^+ residues, with the exception of the internal His96 imidazole ring, whose chemical-shift values resembled the noncharged and nontitratable His93 in H-FABP (*17*). Second, the side chains of residues 45, 75, and 85 were replaced corresponding to the *Lm*-FABP sequence and subsequently energy-minimized according to the EM protocol described below, with all other atom coordinates of the protein kept fixed.

Next, the oleate molecule was placed manually into the protein cavity, by searching for a position of minimal potential energy in the region where the most pronounced chemical-shift changes had been observed in the NMR spectra after relipidation of *Lm*-FABP with oleate. Thereby, the ligand was docked in two alternative ways: (i) in a U-shaped oleate conformation with a *cis* double bond, analogous to the binding arrangement in H-FABP (*26*), and (ii) in a more stretched oleate conformation with a *trans* double bond, where the methyl group extends in the direction of the turn between helices αI and αII similar to the bound ligand in *Ms*-FABP. Because the subsequent restrained-EM and restrained-MD calculations always produced very similar binding arrangements, with a *cis* double bond in the bound oleate, we concluded that this conformation is energetically preferred. Presented here are therefore only the results obtained when starting from the U-shaped oleate with a *cis* double bond conformation.

The structure calculations were performed with DISCOVER in several steps, employing the consistent valence force field (*27*), with the dielectric constant set equal to *r* (distance in angstroms). The applied EM protocol consisted of several steps. First, two cycles of 10 000 iterations each were calculated using the steepest descent algorithm until a maximum derivative of 0.5 kcal/Å (cycle 1) or 1.0 kcal/Å (cycle 2) was reached. Next, 15 000 iterations were performed using the conjugate-gradient method up to a maximum derivative of 0.01 kcal/Å. For restrained-EM calculations, the experimentally obtained intermolecular NOE data were included with a force constant of 20 kcal $Å^{-2}$ mol^{-1} as upper limit distance restraints, which had been categorized into short (≤3 Å), medium (≤4.5 Å), or long (≤6 Å) distance limits according to strong, medium, or weak NOE signal intensities, respectively. To save computer time and avoid distortions of the X-ray structure, only the atom coordinates of the oleate molecule and the side chains of the residues located inside the binding cavity were minimized, whereas the rest of the protein was kept fixed. The energy-minimized complex structure was subsequently applied to a 100 ps restrained-MD run (i.e., with NOE restraints), performing 100 000 iterations in 1 fs steps, with the temperature set to 35 °C. Every 4000 steps (i.e., 4 ps), a conformer was written to the disk. Finally, each stored MD conformer was subjected once more to a restrained EM according to the same protocol as described above.

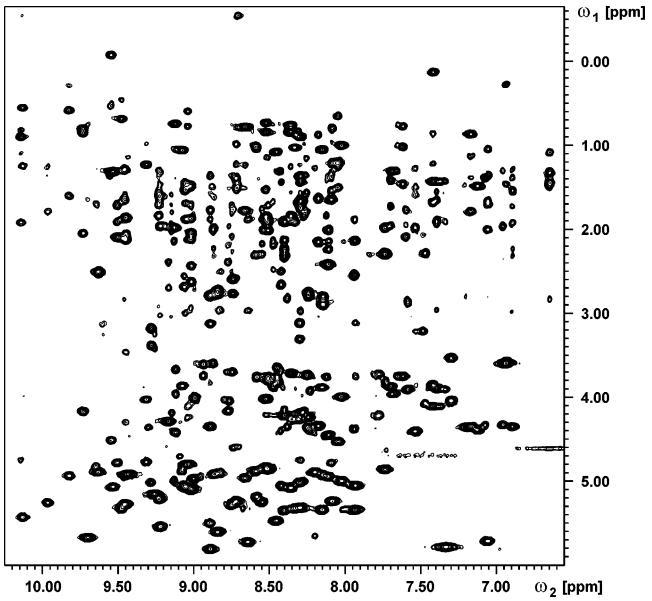

FIGURE 1: Amide region of a homonuclear $^1H/^1H$ TOCSY spectrum collected with the $^{13}C_{18}$-oleate/*Lm*-FABP complex sample at 35 °C (500.13 MHz proton resonance frequency). The mostly β-sheet structure of the protein provides a high signal dispersion.

RESULTS AND DISCUSSION

Structural Characterization of Lm-FABP. We have previously reported the NMR resonance assignment of a recombinant *Lm*-FABP sample that included a mixed population of bound fatty acid ligands, which were endogeneous to the bacterial expression system used for the protein production (*28*). In the present study, the heterogeneous fatty acid population was exchanged by delipidation and subsequent relipidation with uniformly ^{13}C-labeled oleate, leading to better resolved spectral data with fewer multiple spin systems (Figure 1). As illustrated in Figure 2A, the resulting spectra of the oleate complex displayed, compared to the previous NMR experiments with a mixed ligand population, significant chemical-shift changes (i.e., >0.10 ppm) exclusively for residues that line the so-called portal region of the protein (Table 1). These data indicate that the lipid-binding site is located inside the large protein cavity near the fatty acid portal (Figure 2B), as in the case of the mammalian FABPs of i-LBP *subfamily IV*; this was subsequently confirmed by intermolecular NOE connectivities between the fatty acid ligand and the protein (Figure 2C). Consequently, the ^1H resonances of all 134 protein residues (including Met0) were assigned for the *Lm*-FABP/oleate complex, with the exception of 1 missing and 8 incomplete side-chain assignments. This sequence-specific resonance assignment has been deposited at the BioMagResBank database (http://www.bmrb.wisc.edu) under accession number BMRB-6931.

Even though the three-dimensional structure of *Lm*-FABP has not been reported to date, we had previously characterized its structural features based on high-resolution NMR data (*28*). The secondary-structure elements indicated in Figure 3 were derived from NOE connectivities typical for α-helix and β-sheet conformations (*21*). A total of 10 antiparallel β strands were identified spanning residues Gly6−Thr15 (βA), Val40−Leu46 (βB), Phe51−Lys56 (βC),

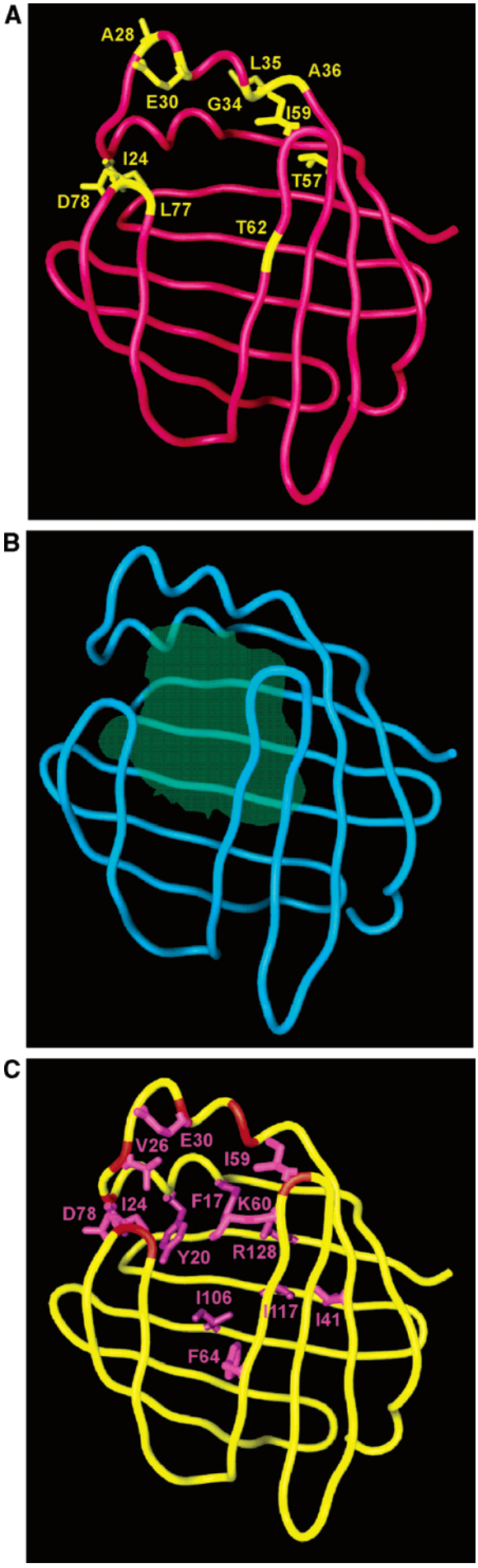

Table 1: Residues that Show Chemical-Shift Changes Larger than 0.10 ppm between (i) *Lm*-FABP with a Heterogeneous Fatty Acid Population and (ii) *Lm*-FABP in Complex Only with Oleate[a]

location	shifted backbone proton resonances	shifted side-chain proton resonances
helix αI		Ile24
helix αII	Ala28, Glu30, Gly34, Leu35, Ala36	Ala28, Glu30, Leu35
turn βC−βD		Thr57, Ile59
β-strand D	Thr62	
turn βE−βF	Leu77, Asp78	Asp78

[a] Differences in backbone and side-chain resonances are listed in separate columns. Most of these residues are located within structure elements that surround the fatty acid portal region.

Asn61−Phe66 (βD), Glu70−Thr76 (βE), Arg80−Asp89 (βF), Lys93−Lys99 (βG), Pro103−Phe110 (βH), Cys115−Lys120 (βI), and Val125−Ala132 (βJ). The α helices comprised residues Asn16−Ile24 (αI) and Ala28−Ala36 (αII). Finally, *Lm*-FABP exhibited a helical loop at the N terminus (Val1−Ala5) that is a typical feature of the mammalian i-LBP *subfamily IV*, which includes H- and A-FABP among others (*3*, *10*). Two other NMR spectral features, which are characteristic for mammalian H- and B-FABP, provided additional evidence for a very high similarity in the structural and functional properties of *Lm*-FABP and H-FABP (*28*): (1) proton spin-system heterogeneities of residues located in the fatty acid portal region and (2) slow exchange of labile protons belonging to Thr76 $O^{\gamma 1}$ and His96 $N^{\epsilon 2}$. The former feature is indicative of distinct ligand conformational states in the presence of several competing fatty acid types (*29*). The slowed labile proton exchange, on the other hand, is due to the intricate hydrogen-bond network inside the lipid-binding cavity (*17*).

The tertiary fold of *Lm*-FABP should be basically identical to *Sg*-FABP, because the amino acid sequences of these two locust muscle proteins differ in only three positions (i.e., 98% sequence identity): Ile45/Val45 at the end of β-strand B, Glu75/Asp75 at the end of β-strand E, and Thr85/Ile85 in the middle of β-strand F. Only the first, conservative substitution (I45V) involves a side chain that is located in the protein interior, whereas the side chains of the other two substituted residues reside on the protein surface. To confirm that the three-dimensional conformation of *Lm*-FABP resembles *Sg*-FABP, we searched for long-range NOE contacts between specific side-chain protons inside *Lm*-FABP that correspond to proton−proton distances of not more than 5 Å in the crystal structure of *Sg*-FABP. For example, Ile41 H^{δ} (β-strand B) displayed NOE connectivities to Cys115 H^{β} and H^{γ} as well as Ile117 H^{δ} (both β-strand I), and Ile86 H^{δ}

FIGURE 2: Ribbon representations of locust muscle FABP. (A) Chemical-shift differences observed between nondelipidated *Lm*-FABP (with a mixed fatty acid population) and *Lm*-FABP in complex only with oleate. Colored in yellow are backbone segments (worm) and side chains (rods) of residues that showed significant (>0.10 ppm) chemical-shift differences between the two holo forms, thus indicating that these regions are affected by the presence of the ligand. (B) The large internal protein cavity near the fatty acid portal, where the ligand can be incorporated, is shaded in green. (C) Colored in red and magenta are backbone segments and side chains, respectively, of residues showing intermolecular NOE connectivities with the bound oleate ligand. [This figure was produced with GRASP (*40*) using the X-ray coordinates of *Sg*-FABP (PDB code 1FTP).]

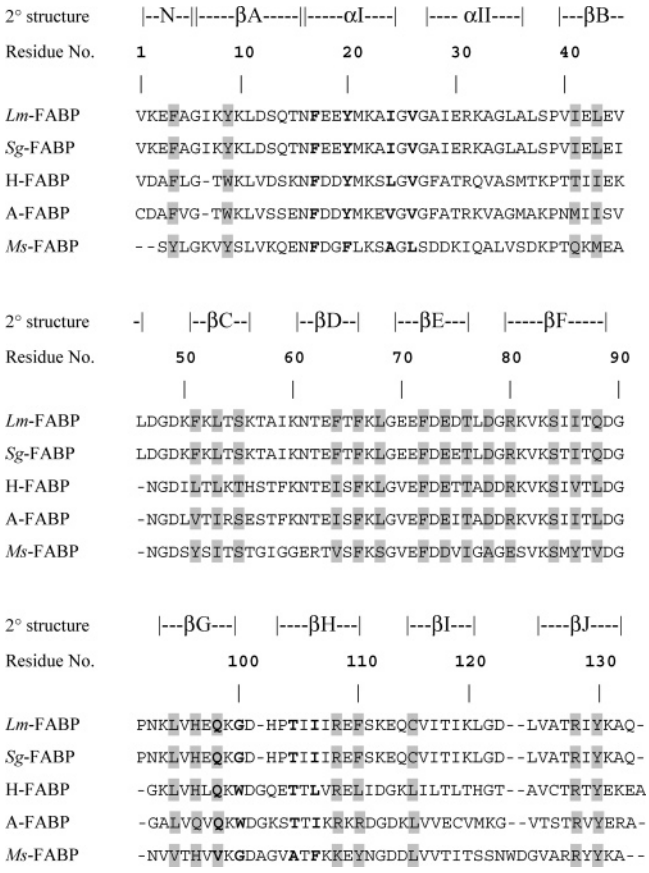

FIGURE 3: Amino acid sequence comparison of *Lm*-FABP with *Sg*-FABP, human H-FABP, murine A-FABP, and *Ms*-FABP from tobacco hornworm. The secondary-structure elements (2°) are indicated in the top row; the amino acid numbering refers to *Lm*-FABP. Shaded in gray are residues that play a role in the intricate hydrogen-bond network inside the H-FABP cavity. Letters in bold represent the residues that line the hydrophobic tail of the palmitate ligand in *Ms*-FABP.

(β-strand F) showed NOE contacts with the ring protons of Phe64 and Phe66 (both β-strand D) as well as Leu53 $H^{\delta 1}$ (β-strand C). On the basis of these and other NOE data, it could be concluded that *Lm*-FABP and *Sg*-FABP have the same three-dimensional structure (*28*).

Three-Dimensional Structure of the Lm-FABP/Oleate Complex. In comparison with the mammalian H-FABP forms, the two locust muscle FABPs exhibit three insertions and two deletions in their amino acid sequences (Figure 3). The insertions occur at position 7 (Ile7 at the beginning of β-strand A), position 46 (Leu46 at the end of β-strand B), and position 91 (Pro91 in the turn between β-strands F and G). The deletions pertain to a residue in the turn between β-strands G and H (Gly99 in human H-FABP) and the C-terminal residue (Ala132 in human H-FABP). These sequential differences have the following effects on the protein structure: (i) a bulge is produced in β-strand A because of the insertion of Ile7 (*Lm*-FABP); (ii) the tight β-turns βB−βC, βF−βG, and βG−βH are disrupted because of the insertions of Leu46 and Pro91 (both *Lm*-FABP) and the deletion of Gly99 (H-FABP); and (iii) the C terminus is shortened because of the deletion of Ala132 (H-FABP). Aside from these minor distinctions, the backbone folds of the FABPs derived from the muscle tissues of vertebrates and invertebrates are very similar. Moreover, a comparison of those residues showing fatty acid contacts (i.e., $\leq$4.5 Å distance) in H-FABP with the corresponding residues in locust muscle FABP has already established that most of these amino acids are conserved (*9*). Hence, the fatty acid ligand should be bound to locust muscle FABP in a manner similar to mammalian H-FABPs, i.e., in a bent conformation in the upper part of the protein cavity, on top of a network of hydrogen-bonded water molecules that stabilize both the protein structure and the ligand binding (*17*).

As expected, the 1D ^{13}C NMR spectrum of $^{13}C_{18}$-oleate bound to *Lm*-FABP showed only one carbonyl signal at 185 ppm, thus indicating a single ligand-binding site (Figure 4A). In addition, the methylene resonances of several other oleate carbon units (C2, C3, C7, C8, C11, C17, and C18) could be assigned on the basis of chemical-shift values and intramolecular NOEs (Figure 4B); these chemical-shift values have also been deposited at the BioMagResBank database under accession number BMRB-6931. (The C9 and C10 moieties, which form the carbon−carbon double bond, could not be unambiguously assigned because of very weak signal intensities; it appears, however, that the ^{13}C and ^{1}H resonances of both groups overlap around 132.7 and 5.26 ppm, respectively.) The ^{13}C-edited 3D NOESY−HSQC spectrum furthermore provided intermolecular NOEs between the ^{13}C-labeled oleate ligand and various protons of the nonlabeled *Lm*-FABP (Figure 4C). These NOE data were converted into 81 intermolecular distance restraints between the protein and the ligand (see Table S1 in the Supporting Information) and were subsequently employed in restrained-EM and restrained-MD calculations to obtain the structure of the complex. Those *Lm*-FABP residues showing NOE connectivities with the oleate ligand are located mainly in the upper part of the protein cavity, as was shown in Figure 2C.

The *Lm*-FABP/oleate complex was calculated on the basis of the X-ray structure of *Sg*-FABP, which had been previously modified in those three positions where the amino acid sequences of both locust FABPs differed. After the ligand was docked, only the protein side chains inside the binding pocket as well as the bound oleate were subjected to minimization, whereas the rest of the protein molecule remained fixed in the conformation that corresponds to crystallized *Sg*-FABP and is supported by the NOE data of *Lm*-FABP. The resulting structure ensemble (Figure 5A), which has been deposited at the Protein Data Bank under PDB code 2FLJ, presents 25 energy-minimized conformers that derive from a 100 ps restrained-MD simulation. These conformers were collected every 4 ps during the MD run, starting 4 ps after the beginning of the calculation. The bound ligand shows a very well-defined orientation at the buried carboxylate end, while the methyl tail in the portal region is partially defrayed. This result was not unexpected, because the carboxylate group is apparently coordinated by the side chains of Arg108, Arg128, and Tyr130 (Figure 5B), analogous to the mammalian FABPs of *subfamily IV*. Moreover, it had been demonstrated previously by another group applying ^{13}C NMR relaxation measurements that, in the case of palmitate bound to rat intestinal FABP, the immersed carboxylate moiety of the ligand displayed a lower mobility than the noncoordinated methyl end near the entry portal (*30*). This is a general consequence of the close interactions typically found between the fatty acid carboxlate group and basic or polar amino acid side chains in lipid-binding proteins, as exemplified also by serum albumin (*31*).

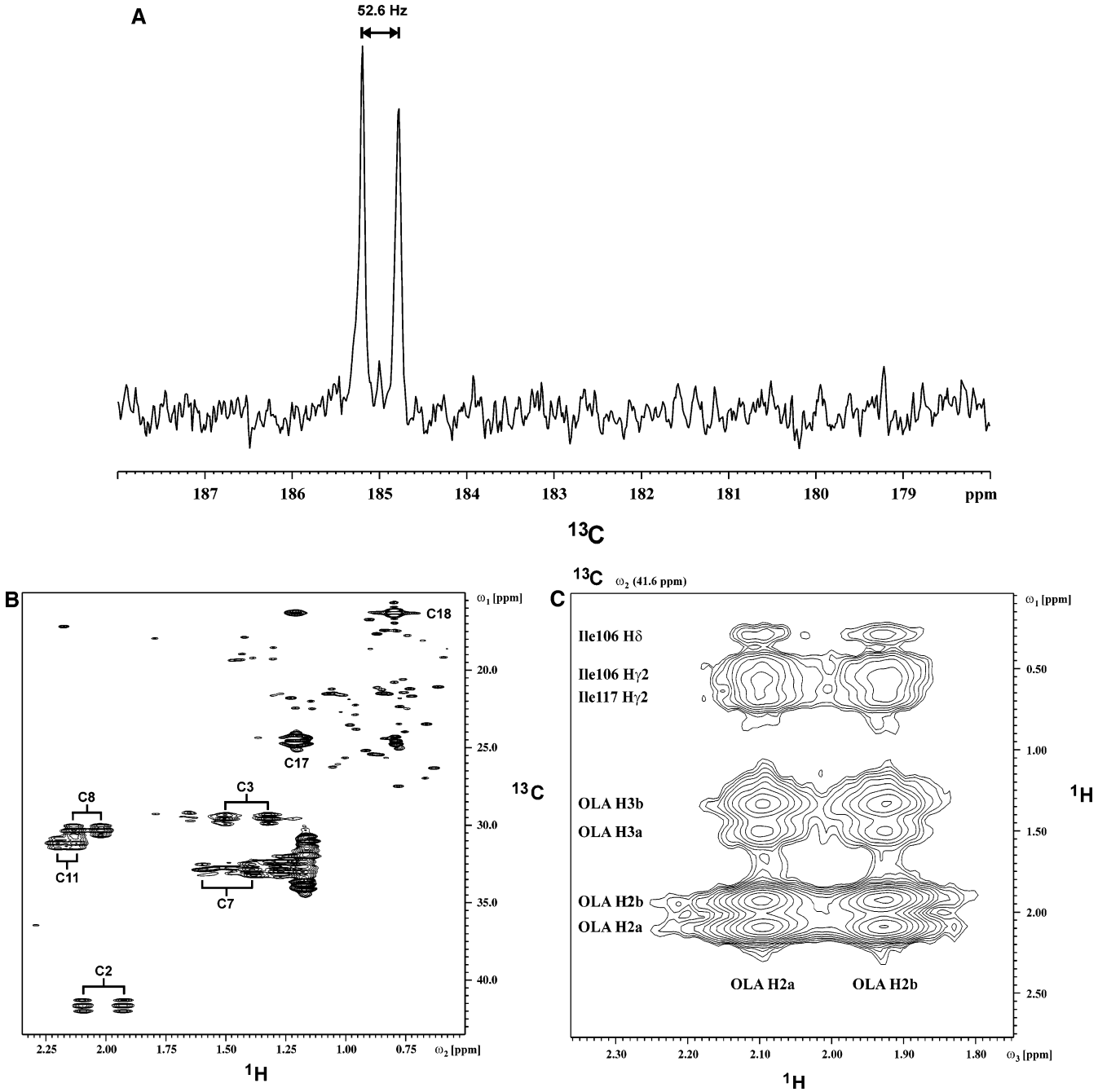

FIGURE 4: NMR spectra collected with the $^{13}C_{18}$-oleate/*Lm*-FABP complex. (A) Carbonyl region of a 1D ^{13}C NMR spectrum (1125 scans) obtained without carbon−carbon decoupling. The doublet at 185 ppm represents the C1 carbon resonance, which couples with the adjacent C2 atom ($^3J_{C1-C2} = 52.6$ Hz), thus indicating that only a single binding site is occupied. Nonbound oleate, which under identical conditions shows a very weak doublet at 182 ppm in aqueous solution, is not detected here. (B) Aliphatic region of a 2D $^1H/^{13}C$ HSQC spectrum that was acquired without carbon−carbon decoupling. The resonances of the assigned methyl and methylene groups are marked. The weak signals in the upper right corner derive from the methyl groups of the nonlabeled protein. (C) ^{13}C slice of the 3D NOESY−HSQC spectrum at the carbon frequency of the oleate (OLA) C2 atom at 41.6 ppm. Both C2 methylene protons (H2a at 2.10 ppm and H2b at 1.93 ppm) show NOE connectivities with each other, with the adjacent C3 methylene protons (H3a and H3b), and with several methyl protons belonging to *Lm*-FABP residues Ile106 and Ile117.

The ligand conformation observed in the *Lm*-FABP/oleate complex is nevertheless unusual. The hydrophobic tail points toward the entry portal, whereas in most mammalian FABPs of i-LBP *subfamily IV*, such as H-FABP, brain (B-) FABP, and epidermal (E-) FABP, this end curls back into the cavity in a U-shaped conformation (Figure 6A). The only known exception is murine A-FABP, where just the very long-chain fatty acid arachidonate displays a U-shape in the bound form (*32*), while the long-chain fatty acids palmitate, stearate, and oleate bind with the hydrophobic tail of the ligand always extending toward the open portal (*33*, *34*). These differences between various FABPs in the orientation of the fatty acid methyl terminus can be explained by the presence of a phenylalanine ring (Figure 6B) that is highly conserved in mammalian FABPs. In the case of H-FABP, the phenyl ring has been proposed to act essentially as a lid at the entry portal (*35*), which interacts with the hydrophobic tail of the bound ligand and thus stabilizes the complex. In A-FABP, however, this phenyl ring points away from the binding cavity, thus allowing the hydrophobic tail of the bound ligand to orient

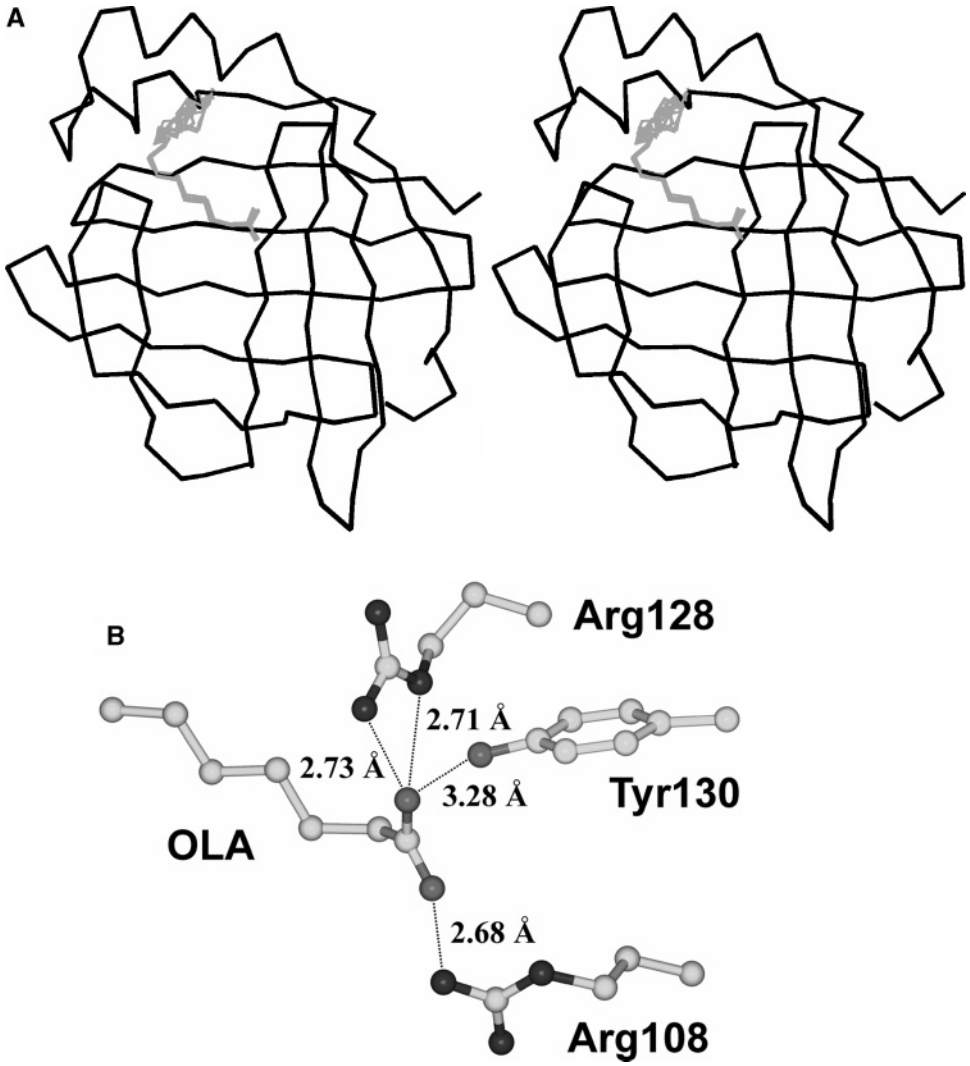

FIGURE 5: (A) Stereo representation of *Lm*-FABP (black C^{α} trace) in complex with oleate (gray carbon chain). The carboxylate group of the ligand is well-defined inside the protein cavity, while the hydrophobic tail displays a higher conformational variability near the fatty acid entry portal. (B) Analogous to mammalian FABPs of i-LBP *subfamily IV*, the oleate (OLA) carboxylate group in *Lm*-FABP shows a binding geometry that suggests coordinative interactions with the side chains of the highly conserved residues Arg108, Arg128, and Tyr130. Heavy atom pairs that show a hydrogen-bond geometry are indicated by dotted lines marked with the corresponding distance values.

itself toward the entry portal. In the A-FABP/arachidonate complex, on the other hand, the tail of the ligand curls back into the cavity even though the entry portal remains in the "open" state. Presumably, as in the case of serum albumin (*31*, *36*), the methyl terminus of this very long-chain fatty acid would be too exposed to the bulk solvent, thus making such a protected position of the hydrophobic tail energetically more favorable. In *Lm*-FABP, finally, the phenylalanine is replaced by Ile59, whose side chain has a similar orientation as Phe57 in A-FABP, i.e., with the portal open. As a consequence, similar to A-FABP, the methyl end of the oleate ligand in *Lm*-FABP occupies the space where the phenyl ring of Phe57 is located in H-FABP (Figure 6B).

A sequence comparison (see Figure 3) of all of those residues, which in human H-FABP are in direct contact with the internal water cluster or the ligand carboxyl group (*17*), shows a high degree of similarity between *Lm*-FABP and H-FABP (23 of 26 residues are conserved; >60% identity). An analogous comparison between *Lm*-FABP and A-FABP yields the same result. Nevertheless, despite the above-mentioned similarity in the oleate-binding conformation between *Lm*-FABP and A-FABP, there is a significant difference between the locust and mammalian protein forms. In *Lm*-FABP, the central section of the fatty acid carbon chain is aligned slightly (>1 Å) further into the back of the binding cavity compared to the ligand positions in H- and A-FABP (Figure 6B). This effect had been previously predicted (*18*), because the rather large Leu77 side chain, which is conserved in the βE−βF turn of the locust FABP forms, apparently diverts the ligand as a result of steric hindrance, in contrast to the mammalian FABPs of *subfamily IV* where the corresponding residue is a highly conserved alanine. As a result of this displacement, the oleate ligand in *Lm*-FABP differs clearly from the typical U-shaped turn conformation observed in the mammalian counterparts.

Comparison between Vertebrate and Invertebrate FABPs. Hence, it appears that, although they separated from vertebrates only after the evolutionary development of the four i-LBP subfamilies typically found in mammalian organisms (*1*), the FABP forms of invertebrate species feature certain amino acid substitutions that produce a more diverse collection of binding arrangements than observed in vertebrates. This is also evident from the palmitate-binding scenario reported for *Ms*-FABP by Benning and co-workers (*19*),

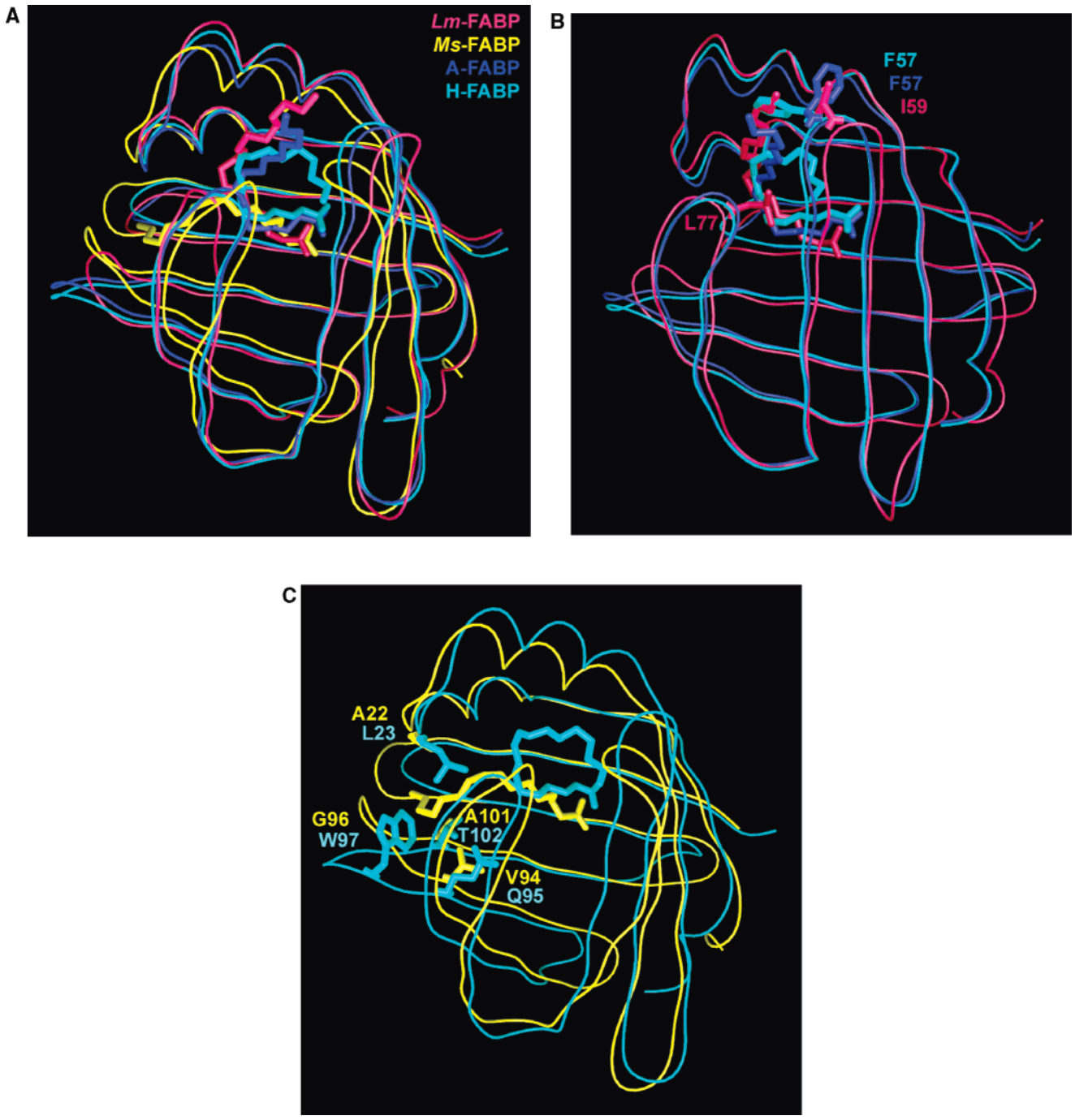

FIGURE 6: (A) Superposition of the calculated *Lm*-FABP/oleate complex with the X-ray structures of *Ms*-FABP (with palmitate), murine A-FABP (with oleate), and human H-FABP (with oleate). The overall three-dimensional fold (backbone worm) is very similar for both the vertebrate and invertebrate proteins, but the ligand-binding scenarios (rods) differ notably. The invertebrate FABP from tobacco hornworm (*Ms*-FABP; yellow) displays a highly unusual extended orientation of the ligand, whereas the bent oleate conformation in *Lm*-FABP (red) shows some resemblances to the mammalian A-FABP (blue) and H-FABP (cyan) forms. (B) Structural comparison of *Lm*-FABP with human H-FABP and murine A-FABP employing the same color code as in A. The backbone folds are very similar for all three proteins. A major difference is observed in the turn $\beta C-\beta D$ that (i) exhibits an isoleucine residue (I59) in *Lm*-FABP instead of the highly conserved phenylalanine found in mammalian FABPs and (ii) is pushed out further in *Lm*-FABP. As a consequence, similar to A-FABP, the ligand entry portal in *Lm*-FABP is open, thus allowing the oleate methyl group to protrude from the binding cavity, in contrast to H-FABP where the phenyl ring lid is closed. (C) Structural comparison of *Ms*-FABP with human H-FABP employing the same color code as in A. In the vertebrate protein, the oleate methyl end curls back in a U-shaped conformation, whereas the palmitate in *Ms*-FABP stretches nearly linearly through the protein pocket. This extended ligand conformation in *Ms*-FABP is apparently favored by several amino acid substitutions, in particular the less bulky side chains in positions A22 and G96. [This figure was produced with GRASP (*40*).]

where the hydrophobic tail of the ligand extends into a section of the protein cavity that is not occupied in any of the known vertebrate FABP structures (Figure 6C). Similar to *Lm*-FABP, specific amino acid replacements in *Ms*-FABP also produce a rather unique binding conformation.

In comparison with other i-LBP amino acid sequences, *Ms*-FABP shows the greatest similarity with human H-FABP of *subfamily IV* (33% identity or 60% similarity) but is not much more distant to human L-FABP of *subfamily II* (29% identity or 57% similarity). In fact, there is a conflict in the

literature whether on an evolutionary basis *Ms*-FABP is more closely related to H-FABP or L-FABP (*1*, *37*). With regard to the internal water network found in H-FABP (Figure 3), *Ms*-FABP certainly shows a lower degree of conservation (17 of 26 residues conserved; 27% identity) than *Lm*-FABP, in particular for the respective cavity side chains in the turn βE−βF. However, when the conformation of the bound ligand in *Ms*-FABP is compared with that in H-FABP and L-FABP, the fatty acid carboxylate end (at least up to C6) is rather similar in *Ms*-FABP and human H-FABP, while the positions and orientations of the two fatty acid ligands in the cavity of rat L-FABP are quite different (*38*). Hence, here we compare briefly the binding features of *Ms*-FABP and H-FABP to demonstrate the drastic effects of certain amino acid replacements on the ligand-binding arrangement.

The hydrophobic tail of the palmitate ligand in *Ms*-FABP is lined by residues Phe15, Phe18, Ala22, Leu24, Val94, Gly96, Ala101, and Phe103 (Figure 3). In the H-FABP sequence, only one of these residues is conserved (Phe16), while three are replaced with polar side chains (Tyr19, Gln95, and Thr102), two with smaller hydrophobic side chains (Val25 and Leu104), and two with much more bulky side chains (Leu23 and Trp97). The combination of these substitutions, in particular Ala22 (*Ms*-FABP) → Leu23 (H-FABP) and Gly96 (*Ms*-FABP) → Trp97 (H-FABP) for steric reasons, presumably prevent fatty acid binding in this part of the H-FABP cavity (Figure 6C). Interestingly, Phe103 of *Ms*-FABP also occurs in human B-FABP, where it plays a role in the interaction with the π-electron system of polyunsaturated fatty acid ligands; aside from this single exception, however, all of the other respective residues in B-FABP are identical to H-FABP.

CONCLUSIONS

The X-ray structure of human H-FABP (*26*) has shown that a water molecule cluster exists inside the binding cavity, forming an intricate hydrogen-bond network as also evidenced by several slow-exchanging protons observed with NMR (*17*). This network is created mainly by two factors: (i) several polar side chains in the center section of the ligand-binding cavity and (ii) a patch of hydrophobic residues at the bottom of the cavity. Most of these residues defining the internal water structure in H-FABP are also present in both locust flight muscle FABPs (Figure 3). This close evolutionary relationship between these lipid carriers suggests that there should be marked structural (i.e., very stable protein conformation) and functional (i.e., high fatty-acid-binding affinity) similarities. In fact, the structure of the locust FABPs as derived from NMR and X-ray shows a strong resemblance to the mammalian H-FABP in both overall fold and ligand-binding properties: first, the 3_{10}-helical loop at the N terminus, which is a hallmark of mammalian FABPs belonging to i-LBP *subfamily IV*, is also present in the locust FABPs. Second, the coordination of the fatty acid carboxylate group by Arg108, Arg128, and Tyr130 of *Lm*-FABP is analogous to *subfamily IV* as well (*10*). Third, the binding affinity of oleic acid to *Lm*-FABP (47 μM) is of the same order as that of H-FABP (43 μM), suggesting very similar overall binding characteristics (*9*). Finally, the spin-system heterogeneities observed in the portal region of nondelipidated *Lm*-FABP had previously been reported only for mammalian H-FABP and B-FABP (*29*, *39*).

Nevertheless, despite this high degree of similarity, the locust protein exhibits several features in the ligand-binding arrangement that are atypical for vertebrate FABPs. For example, in the case of *Lm*-FABP, the side chain of Leu77 apparently forces the fatty acid into a conformation that is significantly different from the typical U-shaped bend found in mammalian FABPs of i-LBP *subfamily IV*. Furthermore, residue Ile59 replaces the highly conserved phenyl ring lid at the fatty acid entry portal, thus leading to an orientation of the ligand tail that is similar to mammalian A-FABP, where the portal lid is, rather uncommon for *subfamily IV* members, in an open state. Despite these differences in the ligand-binding arrangement, however, the oleic acid binding affinity of *Lm*-FABP is as high as that of the highly specialized mammalian H-FABP.

The only other hitherto known invertebrate FABP structure (*Ms*-FABP) shows a completely novel fatty acid ligand orientation, which is due to a number of very specific amino acid substitutions. It therefore seems that, even though *Lm*-FABP does show a fairly close resemblance to the FABP types of the mammalian i-LBP *subfamily IV*, the invertebrate FABPs may have produced a greater structural (and possibly also functional) diversity that still needs to be explored further. It is conceivable that other invertebrate species have developed yet unknown binding scenarios that may have optimized the transport of fatty acids or other lipids to warrant particular physiological functions. Hence, such new conformational features might be of special interest to protein engineers who intend to design improved binding characteristics for this widespread family of lipid carriers.

ACKNOWLEDGMENT

The authors thank Primož Pristovšek (National Institute of Chemistry, Ljubljana, Slovenia) for expert advice regarding the computational techniques.

SUPPORTING INFORMATION AVAILABLE

Supplementary Table S1, the upper distance limits used as intermolecular distance restraints in the EM and MD calculations of the *Lm*-FABP/oleate complex. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Schaap, F. G., van der Vusse, G. J., and Glatz, J. F. C. (2002) Evolution of the family of intracellular lipid binding proteins in vertebrates, *Mol. Cell. Biochem. 239*, 69−77.
2. Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S., and Jones, T. A. (1994) Lipid-binding proteins: A family of fatty acid and retinoid transport proteins, *Adv. Protein Chem. 45*, 89−151.
3. Lücke, C., Gutiérrez-González, L., and Hamilton, J. (2003) Intracellular lipid binding proteins: Evolution, structure, and ligand binding, in *Cellular Proteins and Their Fatty Acids in Health and Disease* (Duttaroy, A. K., and Spener, F., Eds.), pp 95−118, Wiley-VCH, Weinheim, Germany.
4. Zimmerman, A. W., and Veerkamp, J. H. (2002) New insights into the structure and function of fatty acid-binding proteins, *Cell. Mol. Life Sci. 59*, 1096−1116.
5. Ek, B. A., Cistola, D. P., Hamilton, J. A., Kaduce, T. L., and Spector, A. A. (1997) Fatty acid binding proteins reduce 15-lipoxygenase-induced oxygenation of linoleic acid and arachidonic acid, *Biochim. Biophys. Acta 1346*, 75−85.
6. Zhang, F., Lücke, C., Baier, L. J., Sacchettini, J. C., and Hamilton, J. A. (2003) Solution structure of human intestinal fatty acid

binding protein with a naturally-occurring single amino acid substitution (A54T) that is associated with altered lipid metabolism, *Biochemistry 42*, 7339−7347.
7. Boord, J. B., Maeda, K., Makowski, L., Babaev, V. R., Fazio, S., Linton, M. F., and Hotamisligil, G. S. (2004) Combined adipocyte-macrophage fatty acid-binding protein deficiency improves metabolism, atherosclerosis, and survival in apolipoprotein E-deficient mice, *Circulation 110*, 1492−1498.
8. Haunerland, N. H., Andolfatto, P., Chisholm, J. M., Wang, Z., and Chen, X. (1992) Fatty-acid-binding protein in locust flight muscle. Developmental changes of expression, concentration and intracellular distribution, *Eur. J. Biochem. 210*, 1045−1051.
9. Maatman, R. G. H. J., Degano, M., van Moerkerk, H. T. B., van Marrewijk, W. J. A., van der Horst, D. J., Sacchettini, J. C., and Veerkamp, J. H. (1994) Primary structure and binding characteristics of locust and human muscle fatty-acid-binding proteins, *Eur. J. Biochem. 221*, 801−810.
10. Hanhoff, T., Lücke, C., and Spener, F. (2002) Insights into binding of fatty acids by fatty acid binding proteins, *Mol. Cell. Biochem. 239*, 45−54.
11. Haunerland, N. H. (1994) Fatty acid binding protein in locust and mammalian muscle. Comparison of structure, function and regulation, *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol. 109*, 199−208.
12. Haunerland, N. H., Chen, X., Andolfatto, P., Chisholm, J. M., and Wang, Z. (1993) Developmental changes of FABP concentration, expression, and intracellular distribution in locust flight muscle, *Mol. Cell. Biochem. 123*, 153−158.
13. Guglielmo, C. G., Haunerland, N. H., and Williams, T. D. (1998) Fatty acid binding protein, a major protein in the flight muscle of migrating western sandpipers, *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol. 119*, 549−555.
14. Pelsers, M. M., Butler, P. J., Bishop, C. M., and Glatz, J. F. C. (1999) Fatty acid binding protein in heart and skeletal muscles of the migratory barnacle goose throughout development, *Am. J. Physiol. 276*, R637−R643.
15. Guglielmo, C. G., Haunerland, N. H., Hochachka, P. W., and Williams, T. D. (2002) Seasonal dynamics of flight muscle fatty acid binding protein and catabolic enzymes in a migratory shorebird, *Am. J. Physiol. Regul. Integr. Comput. Physiol. 282*, R1405−R1413.
16. Binas, B., Danneberg, H., McWhir, J., Mullins, L., and Clark, A. J. (1999) Requirement for the heart-type fatty acid binding protein in cardiac fatty acid utilization, *FASEB J. 13*, 805−812.
17. Lücke, C., Huang, S., Rademacher, M., and Rüterjans, H. (2002) New insights into intracellular lipid binding proteins: The role of buried water, *Protein Sci. 11*, 2382−2392.
18. Haunerland, N. H., Jacobson, B. L., Wesenberg, G., Reyment, I., and Holden, H. M. (1994) Three-dimensional structure of the muscle fatty-acid-binding protein isolated from the desert locust *Schistocerca gregaria*, *Biochemistry 33*, 12378−12385.
19. Benning, M. M., Smith, A. F., Wells, M. A., and 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.
20. Glatz, J. F. C., and Veerkamp, J. H. (1983) Removal of fatty acids from serum albumin by Lipidex 1000 chromatography, *J. Biochem. Biophys. Methods 8*, 57−61.
21. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
22. Kay, L. E., Keifer, P., and Saarinen, T. (1992) Pure absorption gradient enhanced heteronuclear single quantum correlation spectroscopy with improved sensitivity, *J. Am. Chem. Soc. 114*, 10663−10665.
23. Schleucher, J., Sattler, M., and Griesinger, C. (1993) Coherence selection by gradients without signal attenuation: Application to the three-dimensional HNCO experiment, *Angew. Chem., Int. Ed. Engl. 32*, 1489−1491.
24. Marion, D., Ikura, M., Tschudin, R., and Bax, A. (1989) Rapid recording of 2D NMR spectra without phase cycling. Application to the study of hydrogen exchange in proteins, *J. Magn. Reson. 85*, 393−399.
25. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) ^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR, *J. Biomol. NMR 6*, 135−140.
26. Young, A. C., Scapin, G., Kromminga, A., Patel, S. B., Veerkamp, J. H., and 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.
27. Dauber-Osguthorpe, P., Roberts, V. A., Osguthorpe, D. J., Wolff, D. J., Genest, M., and Hagler, A. T. (1988) Structure and energetics of ligand binding to proteins: *E. coli* dihydrofolate reductase trimethoprin, a drug-receptor system, *Proteins 4*, 31−47.
28. Lücke, C., Kizilbash, N., van Moerkerk, H., Veerkamp, J., and Hamilton, J. (2003). Letter to the editor: NMR assignment and structural characterization of the fatty acid binding protein from the flight muscle of *Locusta migratoria*, *J. Biomol. NMR 25*, 355−356.
29. Lücke, C., Rademacher, M., Zimmerman, A. W., van Moerkerk, H. T. B., Veerkamp, J. H., and Rüterjans, H. (2001) Spin-system heterogeneities indicate a selected-fit mechanism in fatty acid binding to heart-type fatty acid binding protein (H-FABP), *Biochem. J. 354*, 259−266.
30. Zhu, L., Kurian, E., Prendergast, F. G., and Kemple, M. D. (1999) Dynamics of palmitic acid complexed with rat intestinal fatty acid binding protein, *Biochemistry 38*, 1554−1561.
31. Hamilton, J. A. (2004) Fatty acid interactions with proteins: What X-ray crystal and NMR solution structures tell us, *Prog. Lipid Res. 43*, 177−199.
32. LaLonde, J. M., Levenson, M. A., Roe, J. J., Bernlohr, D. A., and Banaszak, L. J. (1994) Adipocyte lipid-binding protein complexed with arachidonic acid. Titration calorimetry and X-ray crystallographic studies, *J. Biol. Chem. 269*, 25339−25347.
33. Xu, Z., Bernlohr, D. A., and 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.
34. LaLonde, J. M., Bernlohr, D. A., and Banaszak, L. J. (1994) X-ray crystallographic structures of adipocyte lipid-binding protein complexed with palmitate and hexadecanesulfonic acid. Properties of cavity binding sites, *Biochemistry 33*, 4885−4895.
35. Eads, J., Sacchettini, J. C., Kromminga, A., and Gordon, J. I. (1993) *Escherichia coli*-derived rat intestinal fatty acid binding protein with bound myristate at 1.5 Å resolution and I-FABPArg106 → Gln with bound oleate at 1.74 Å resolution, *J. Biol. Chem. 268*, 26375−26385.
36. Choi, J. K., Ho, J., Curry, S., Qin, D., Bittman, R., and Hamilton, J. A. (2002) Interactions of very long-chain saturated fatty acids with serum albumin, *J. Lipid Res. 43*, 1000−1010.
37. Haunerland, N. H., and Spener, F. (2004) Properties and physiological significance of fatty acid binding proteins, *Adv. Mol. Cell Biol. 33*, 99−123.
38. Thompson, J., Winter, N., Terwey, D., Bratt, J., and Banaszak, L. (1997) The crystal structure of the liver fatty acid-binding protein. A complex with two bound oleates, *J. Biol. Chem. 272*, 7140−7150.
39. Rademacher, M., Zimmerman, A. W., Rüterjans, H., Veerkamp, J. H., and Lücke, C. (2002) Solution structure of fatty acid-binding protein from human brain, *Mol. Cell. Biochem. 239*, 61−68.
40. Nicholls, A., Sharp, K. A., and Honig, B. (1991) Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins 11*, 281−296.

BI060224F