Three-Dimensional Structure of the Muscle Fatty-Acid-Binding Protein Isolated from the Desert Locust *Schistocerca gregaria*[†,‡]

Norbert H. Haunerland,[*,§] Bruce L. Jacobson,[‖] Gary Wesenberg,[‖] Ivan Rayment,[‖] and Hazel M. Holden[*,‖]

Department of Biological Sciences, Simon Fraser University, Burnaby, Canada, and Institute for Enzyme Research, Graduate School and Department of Biochemistry, University of Wisconsin, Madison, Wisconsin 53705

Received June 28, 1994; Revised Manuscript Received August 4, 1994[⊗]

ABSTRACT: The three-dimensional structure of the fatty-acid-binding protein isolated from the flight muscle of the desert locust *Schistocerca gregaria* has been solved and refined to a crystallographic *R*-value of 18.5% for all measured X-ray data from 30.0- to 2.2-Å resolution. Crystals employed in the investigation were grown from 2.6 to 2.8 M ammonium sulfate solutions, buffered at pH 7.5 and containing 2–5% 2-methyl-2,4-pentanediol. They belonged to the space group $P2_1$ with unit cell dimensions of a = 61.6 Å, b = 44.8 Å, c = 63.9 Å, and β = 113.6° and two molecules per asymmetric unit. The protein fold consists of ten strands of antiparallel β-pleated sheet that wrap around to form a β-barrel. In addition, there are two small α-helices and six type I, two type II, and two type II′ turns. The two molecules pack in the asymmetric unit as a dimer with a local 2-fold rotational axis. The subunit–subunit interface involves amino acid side chains located in the area of the helix–turn–helix motif and the turn between β-strands E and F. It is this area that has been speculated to form the portal through which fatty acids enter the binding cavity. There are 23 solvent molecules that are conserved between the two independent molecules in the asymmetric unit. Nine of these waters play important structural roles. A three-dimensional comparison between the insect and human muscle fatty-acid-binding proteins shows that their α-carbons superimpose with a root-mean-square deviation of 0.77 Å for 89 structurally equivalent atoms. This high tertiary homology between these two proteins is surprising in light of the fact that they developed independently for more than 500 million years.

Fatty-acid-binding proteins are low molecular weight proteins thought to play a role in fatty-acid transport and metabolism (Veerkamp *et al.*, 1991). Since their discovery by Ockner *et al.* (1972) and Mishkin *et al.* (1972), they have been the subject of intense investigations. It is now known that in vertebrate species there are at least four structurally and functionally distinct fatty-acid-binding proteins isolated from the liver, the intestine, muscles, and adipocytes. These proteins belong to a superfamily of intracellular lipid-binding proteins which also includes P2 myelin protein, cellular retinol and retinoic-acid-binding proteins, and gastrotropin (Matarese *et al.*, 1989). The three-dimensional structures for a number of these vertebrate proteins have been solved to various resolutions, including P2 myelin protein (Jones *et al.*, 1988; Cowan *et al.*, 1993), rat intestinal fatty-acid-binding protein (Sacchettini *et al.*, 1989; Scapin *et al.*, 1992), chicken liver basic fatty-acid-binding protein (Scapin *et al.*, 1990), bovine heart muscle fatty-acid-binding protein (Müller-Fahrnow *et al.*, 1991), murine adipocyte lipid-binding protein (Xu *et al.*, 1992, 1993), and human muscle fatty-acid-binding protein (Zanotti *et al.*, 1992). In addition, the crystal structures of the cellular retinol-binding proteins, I and II, are now known (Cowan *et al.*, 1993; Winter *et al.*, 1993). All of these proteins display a similar molecular architecture with ten strands of antiparallel β-pleated sheet forming an up-and-down β-barrel that envelops the long-chain fatty acid. For an elegant review of the structures of fatty-acid-binding proteins, see Banaszak *et al.* (1994).

Vertebrate species, however, are not the only organisms to possess fatty-acid-binding proteins. These types of lipid-transport molecules have been identified in invertebrates, namely, in the blood fluke *Schistosoma mansoni* (Moser *et al.*, 1991), in insects, specifically in two locust species, *Schistocerca gregaria* (Haunerland & Chisholm, 1990) and *Locusta migratoria* (Van der Horst *et al.*, 1993), and in the tobacco hornworm *Manduca sexta* L. (Smith *et al.*, 1992). The presence of these proteins in insects is not surprising, however, since insects, like any other organism with a circulatory system, require a mechanism by which long-chain fatty acids can be safely and efficiently transported throughout both intra- and extracellular compartments. Furthermore, lipids play important roles in energy storage, especially for those insect species that migrate long distances. Both lepidopteran species, such as *M. sexta*, and orthopteran species, most prominently locusts, are known to fuel extended flight exclusively through the oxidation of fatty acids (Beenakkers *et al.*, 1985). Prior to migration, vast amounts of lipid are taken up and stored in their fat bodies; during flight, these lipids are mobilized and transported to the flight muscles, where energy is retrieved by mitochondrial β-oxidation. Locust flight muscles are among the most active muscle cells known with respect to lipid utilization; these tissues have been shown to metabolize fatty acids at a much higher rate than their vertebrate counterparts (Crabtree & Newsholme, 1975). The high flux of fatty acids required to sustain migratory flight necessitates an efficient fatty-acid-transport mechanism. Thus, one would expect that fatty-acid-binding proteins participate in both fatty-acid uptake by insect midgut and transport to the mitochondria in flight muscle cells.

[†] This research was supported in part by grants from the NIH (HL42322 to H.M.H.) and the Heart and Stroke Foundation of British Columbia and Yukon (N.H.H.). H.M.H. is an Established Investigator of the American Heart Association.

[‡] X-ray coordinates for the *S. gregaria* fatty-acid-binding protein have been deposited in the Brookhaven Protein Data Bank (filename 1FTP).

* To whom correspondence should be addressed.

[§] Simon Fraser University.

[‖] University of Wisconsin.

[⊗] Abstract published in *Advance ACS Abstracts,* September 15, 1994.

0006-2960/94/0433-12378$04.50/0

Table 1: Intensity Statistics for the Native X-ray Data Set

	resolution range (Å)							
	overall	30–5.57	–3.97	–3.25	–2.82	–2.53	–2.31	–2.14
no. of measurements	23500	1821	3112	3587	3839	3942	3803	3172
no. of independent reflections[a]	16510	1049 (585)	1812 (985)	2292 (1116)	2615 (1105)	2890 (992)	3075 (718)	2777 (392)
% of the theoretical no. of reflections	92	98	99	99	97	96	91	77
average intensity		1562	1850	1070	436	233	143	89
σ		67	71	49	27	21	21	23
R-factor[b] (%)	3.5	2.0	2.0	2.6	4.0	6.3	10.1	17.1

[a] This is the number of reduced observations. Shown in parentheses is the number of independent measurements for which there were duplicate or symmetry-related observations. [b] R-factor = $\sum|I - \bar{I}|/\sum I \times 100$.

Recently, the three-dimensional structure of a fatty-acid-binding protein from the midgut of *M. sexta* L. has been solved and refined to 1.75-Å resolution (Benning *et al.*, 1992). As expected, the overall molecular fold of the insect molecule is very similar to other lipid-binding proteins such as the P2 myelin protein and the adipocyte lipid-binding protein. However, while the carboxylic acid moiety of the fatty-acid ligand binds in the same general vicinity within the β-barrel as seen in the vertebrate proteins, the positioning of the hydrocarbon chain after C6 is completely different (Benning *et al.*, 1992).

The fatty-acid-binding protein isolated from the adult desert locust *S. gregaria* has been shown to be the most abundant soluble muscle protein, comprising more than 18% of the total cytosolic protein (Haunerland & Chisholm, 1990; Haunerland *et al.*, 1993). The primary structure, known from the cDNA, indicates that the protein has significant homology with vertebrate fatty-acid-binding proteins over its entire 133 amino acid residue sequence (Price *et al.*, 1992). Interestingly, this insect protein demonstrates the greatest sequence homology of 41% identity to the human heart muscle fatty-acid-binding protein despite its evolutionary distance (Price *et al.*, 1992). The *S. gregaria* protein displays much lower homology with the fatty-acid-binding protein from *M. sexta* L. midgut. Here we describe the crystallization, structural determination, and least-squares refinement of fatty-acid-binding protein from *S. gregaria* muscle and compare its molecular motif with those of the fatty-acid-binding proteins isolated from *M. sexta* L. midgut and human muscle.

MATERIALS AND METHODS

Crystallization. The fatty-acid-binding protein from *S. gregaria* was isolated and purified according to published procedures (Haunerland & Chisholm, 1990). For crystallization experiments, the lyophilized protein was dissolved in 10 mM HEPES and 5 mM NaN_3, pH 7.0, to a concentration of 20 mg/mL. A search for crystallization conditions was conducted by the hanging drop method of vapor diffusion. Small clusters of crystals were observed growing at 2.7–3.0 M ammonium sulfate buffered at pH 7.5 with 50 mM Na^+,K^+/PO_4. The addition of 2–5% 2-methyl-2,4-pentanediol yielded thin diamond-shaped crystals.

By precession photography, the crystals were shown to belong to the space group $P2_1$ with unit cell dimensions of a = 61.6 Å, b = 44.8 Å, c = 63.9 Å, and β = 113.6° and two molecules per asymmetric unit. The Matthew's coefficient (V_m) for these crystals was 2.7 Å^3/Da, thus corresponding to a solvent content of approximately 54% (Matthews, 1968). The crystals diffracted to a nominal resolution of 2.2 Å.

X-ray Data Collection and Processing. For X-ray data collection, one crystal was mounted in a quartz capillary tube. A native X-ray data set was collected to 2.2-Å resolution at 4 °C with a Siemens X1000D area detector system. The X-ray source was nickel-filtered Cu Kα radiation from a Rigaku RU200 X-ray generator operated at 50 kV and 50 mA. These X-ray data were subsequently processed with the data reduction software package XDS (Kabsch, 1988a,b) and internally scaled according to the algorithm of Fox and Holmes (1966) as implemented by Dr. Phil Evans. Relevant X-ray data collection statistics may be found in Table 1. The native X-ray data set was 92% complete to 2.2-Å resolution.

Structural Determination and Least-Squares Refinement. Initial protein phases were determined by molecular replacement techniques with the software package AMORE (Navaza, 1987). The three-dimensional X-ray coordinates for the P2 myelin protein, kindly provided by Dr. T. Alwyn Jones, served as the search model. The long-chain fatty acid was removed from the model, and nonconserved amino acid residues were truncated to alanines. The cross-rotation functions and translational searches were calculated with X-ray data from 8.0- to 4.0-Å resolution. Two solutions corresponding to the two molecules in the asymmetric unit were obtained and are as follows: (i) α = 114.01°, β = 123.22°, γ = 355.61°, a = 0.3980, b = 0.0000, and c = 0.0848; (ii) α = 62.55°, β = 159.59°, γ = 361.03°, a = 0.1682, b = 0.3972, and c = 0.3216. Following rigid body refinement, the R-factor was 47.6%.

Upon inspection of the initial electron density map, it was obvious that there were significant differences in several of the surface loops between the insect protein and P2 myelin and that the course of the polypeptide chain through these regions was not entirely clear. In order to expedite the refinement of the insect model, the electron density for the two molecules in the asymmetric unit was averaged according to the algorithm of Bricogne (1976) and an "averaged" model for the insect protein was built into the resulting electron density. This averaged model was placed back into the unit cell, and an initial cycle of least-squares refinement was conducted with the program package TNT (Tronrud *et al.*, 1987). The starting R-factor for all measured X-ray data from 30.0- to 2.2-Å resolution was 33.9%. Additional cycles of manual model building and least-squares refinement reduced the R-factor to 18.5%. Small peaks of electron density located within 3.5 Å of potential hydrogen-bonding groups were modeled as solvent molecules. A total of 140 solvent molecules was included in the refinement. The average temperature factor for the solvent was 45.5 Å^2 with 24 of the water molecules having temperature factors below 30 Å^2. The electron density associated with some of the solvent molecules was rather bulky, suggesting that they may be ordered sulfate or phosphate ions. At this resolution, however, they have been modeled into the electron density as waters. Relevant least-squares refinement statistics may be found in Table 2. A plot of the mean main-chain temperature factors for both molecules in the asymmetric unit is given in Figure 1, and a Ramachandran plot of all non-glycinyl residues is shown in Figure 2. The average temperature factor for all backbone atoms was 34.1

Table 2: Refinement Statistics

resolution limits (Å)	30.0–2.2
R-factor (%)[a]	18.5
no. of reflections used	16724
no. of protein atoms	2106
no. of solvent molecules	140
weighted root-mean-square deviations from ideality	
bond length (Å)	0.018
bond angle (deg)	2.116
planarity (trigonal) (Å)	0.007
planarity (other planes) (Å)	0.007
torsional angle (deg)[b]	19.938

[a] R-factor = $\Sigma|F_o - F_c|/\Sigma|F_o|$. [b] The torsional angles were not restrained during the refinement.

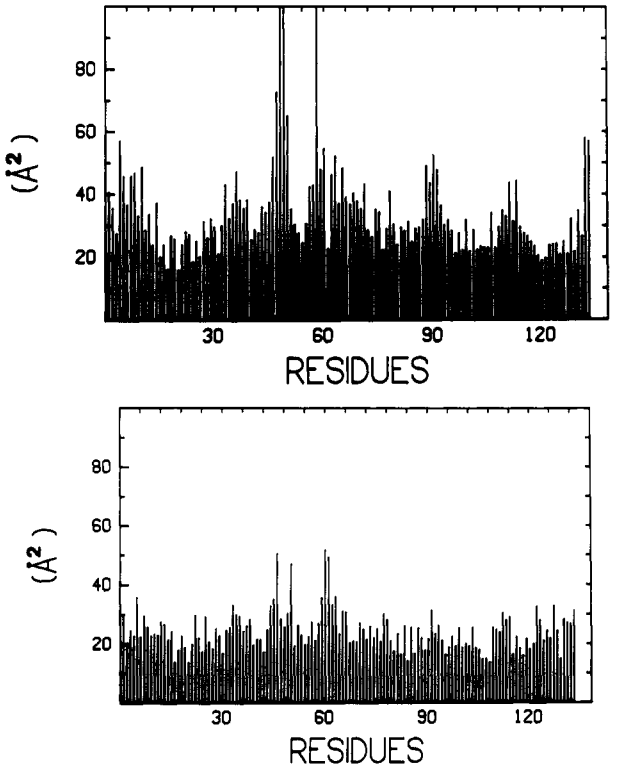

FIGURE 1: Plot of the mean main-chain temperature factors versus amino acid residue number. Overall, the temperature factors for molecule I in the asymmetric unit (a, top) are higher than those for molecule II (b, bottom).

and 25.4 Å², respectively, for molecules I and II in the asymmetric unit. A representative portion of the electron density map is displayed in Figure 3. The electron density was well-ordered except for the following side chains: in molecule I, lysines 8, 32, 50, 52, 56, 60, 67, 81, 83, 112, 131, isoleucines 29 and 59, Gln 88, Asn 92, and Gln 133; in molecule II, lysines 22, 50, 60, 67, 81, 83, 93, 99, 112, 120, 131, Leu 35, Glu 42, and Ile 105. Also, according to the amino acid sequence based on the cDNA sequence, residue 92 is an asparagine (Price *et al.*, 1992). The electron density for residue 92 appears more like that for a serine but has been built into the model as an asparagine. X-ray coordinates for the *S. gregaria* fatty-acid-binding protein have been deposited in the Brookhaven Protein Data Bank (Bernstein *et al.*, 1977).

RESULTS

As can be seen in Figure 1, the backbone atoms for molecule I in the asymmetric unit have higher overall temperature factors than for those of molecule II. There are 12 protein–protein contacts within 3.5 Å between molecule I and symmetry-related molecules in the crystalline lattice. Specifically, these crystalline packing interactions occur at the N-terminus and at the regions defined by Gly 25 to Arg 31, Asp 78 to Val 82, and Gly 100 to His 102. As expected from the overall lower temperature factors, the interactions between molecule II and its symmetry-related partners are more extensive with 25 protein–protein contacts within 3.5 Å. For molecule II, crystalline packing interactions between neighboring molecules occur at the N-terminus and at regions delineated by Thr 15 to Glu 30, Ile 45 to Asp 49, and Asp 78 to Val 125. In light of the differences in temperature factors and for the sake of simplicity, the following discussion will refer only to molecule II unless otherwise indicated. It should be noted, however, that the α-carbons for the two molecules in the asymmetric unit superimpose with a root-mean-square deviation of 0.45 Å according to the algorithm of Rossmann and Argos (1975).

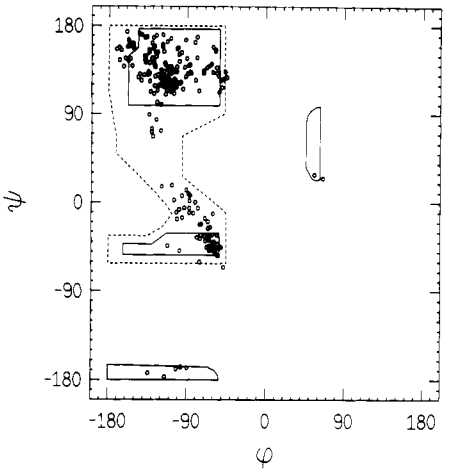

FIGURE 2: ϕ,ψ plot of all non-glycinyl main-chain dihedral angles. Fully allowed ϕ,ψ values are enclosed by solid lines; those only partially allowed are enclosed by dashed lines.

A ribbon representation of molecule II is given in Figure 4. Those amino acid residues participating in standard secondary structural elements are listed in Table 3. Like other members of its class, the molecule is characterized by ten β-strands (A–J) forming two layers that are nearly orthogonal to one another and consisting of six and four β-strands each. All of the β-strands adopt standard dihedral angles except for β-strand A, formed by Lys 8 to Thr 15, which contains a distinctive kink at residue Asp 12 (ϕ = –112.3, ψ = –43.2). While the separation between α-carbons in the β-sheet is typically 4.5–5.0 Å, there is a noticeable gap between β-strands D and E where the separation between α-carbons approaches 12 Å. The molecule also contains six type I turns, two type II turns, and one type II′ turn. The turn delineated by Asp 47 to Lys 50, although built as a type II′ turn, is not well-defined in the present electron density map and most likely adopts a variety of conformations. Additionally, the loop connecting β-strands G and H does not adopt the standard ϕ,ψ angles for classical reverse turns. A list of the dihedral angles for the reverse turns is given in Table 4.

Since the *S. gregaria* protein crystallized with two molecules per asymmetric unit, it is instructive to examine the packing motif. Intermolecular contacts between molecules in the crystalline lattice may provide clues as to the regions on the protein responsible for interactions with other molecules. In the case of the locust protein, the two molecules packed in the asymmetric unit with a local 2-fold rotational axis as displayed in Figure 5. There are 62 contacts equal to or below 4.0 Å between these two molecules at their subunit–subunit interface. This interface is formed by Gly 25 to Glu 30, Asp 78 to Arg

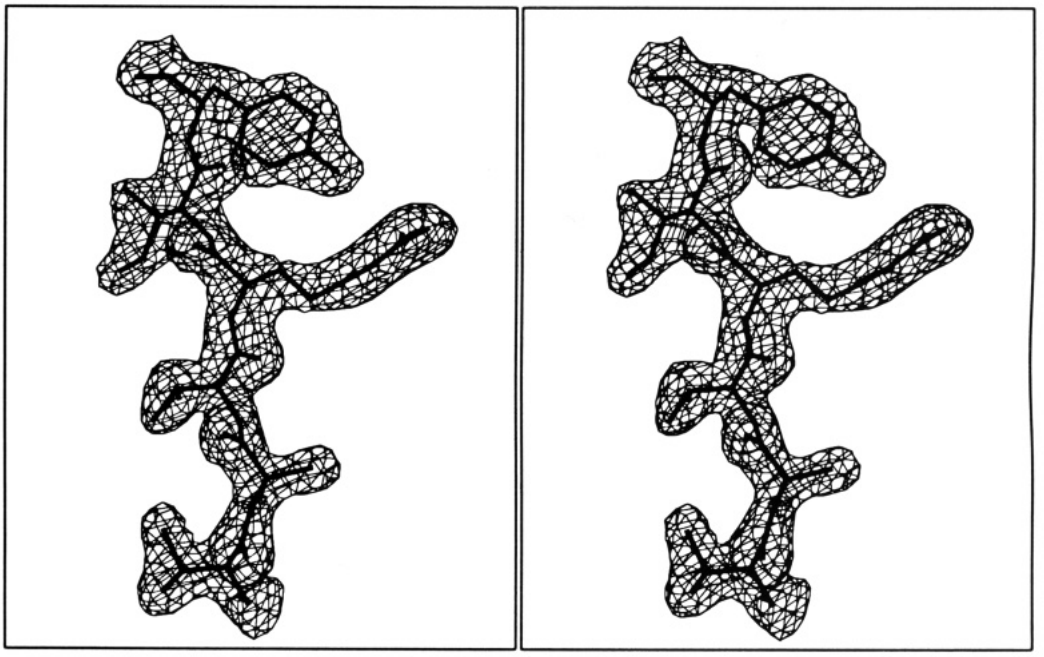

FIGURE 3: Representative portion of the electron density map calculated to 2.2-Å resolution. The electron density shown was calculated with coefficients of the form $(2F_o - F_c)$ and contoured at 1σ. This region corresponds to residues Val 125, Ala 126, Thr 127, Arg 128, and Ile 129 in molecule II.

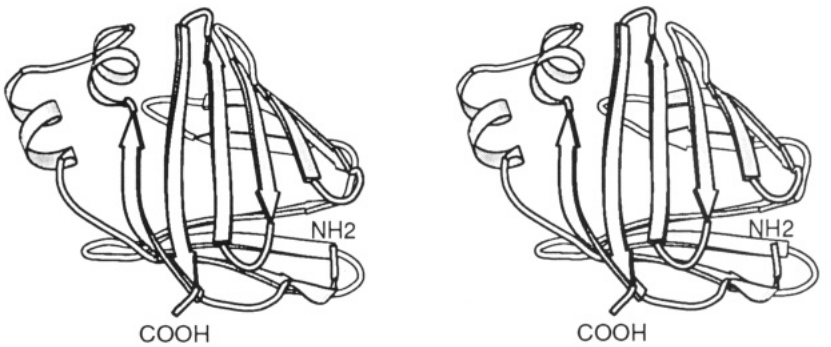

FIGURE 4: Ribbon representation of the fatty-acid-binding protein. This figure was prepared with the software package MOLSCRIPT (Kraulis, 1991).

Table 3: List of Secondary Structural Elements

amino acid residues	type of structure	amino acid residues	type of structure
Lys 2–Ala 5	~type I turn	Lys 67–Glu 70	type II turn
Phe 4–Ile 7	type II turn	Phe 72–Glu 75	β-sheet (E)
Lys 8–Thr 15	β-sheet (A)	Thr 76–Gly 79	type I turn
Phe 17–Met 21	α-helix	Arg 80–Asp 89	β-sheet (F)
Lys 22–Gly 25	type I turn	Gly 90–Lys 93	~type I turn
Ala 28–Gly 34	α-helix	Leu 94–Gln 98	β-sheet (G)
Val 40–Leu 46	β-sheet (B)	Thr 104–Phe 110	β-sheet (H)
Asp 47–Lys 50	type II′ turn	Ser 111–Gln 114	type I turn
Phe 51–Lys 56	β-sheet (C)	Cys 115–Lys 120	β-sheet (I)
Thr 57–Lys 60	type I turn	Leu 121–Leu 124	type II′ turn
Asn 61–Phe 66	β-sheet (D)	Val 125–Ala 132	β-sheet (J)

80, and Asp 101 to His 102. There are three specific hydrogen-bonding interactions within the interface: the backbone amide nitrogen of Ala 28 (molecule I) lies within 2.9 Å of $O^{\delta 1}$ of Glu 30 (molecule II), N^{ϵ} of Arg 80 (molecule I) lies within 3.1 Å of $O^{\gamma 1}$ of Asp 101 (molecule II), and the backbone amide nitrogen of Asp 101 (molecule I) lies within 2.5 Å of $O^{\gamma 1}$ of Asp 101 (molecule II). The helix–turn–helix motif is located at the subunit–subunit interface. A systematic study of oligomeric proteins has shown that subunit–subunit interface surface areas range from 670 Å^2 for superoxide dismutase to over 10 000 Å^2 for catalase (Janin *et al.*, 1988). For the fatty-acid-binding protein dimer described here, the interface is approximately 900 Å^2 using a search probe radius of 1.4 Å (Lee & Richards, 1971).

When the α-carbons for the two locust fatty-acid-binding proteins in the asymmetric unit are superimposed, there are 23 solvent molecules that likewise superimpose within 1 Å. Most of these waters are located at the surface with only three being buried in the binding pocket as can be seen in Figure 6. Nine of the waters form hydrogen bonds to backbone carbonyl oxygens or amide nitrogens and two interact with surface side-chain atoms. The other nine waters located at the surface play more structural roles. Four serve as bridges between β-strands A and B, B and C, G and H, and H and I. There is an additional water located at the N-terminus of β-strand H that hydrogen bonds to residues in β-strand I. Of the remaining four waters, two are located in the gap between β-strands D and E. These waters stabilize the type II turn defined by Lys 67 to Glu 70 where one bridges the backbone amide nitrogen of Lys 67 to the carbonyl oxygen of Glu 70 and the other anchors the type II turn to β-strand F. Finally, two waters are located in the nonclassical reverse turn connecting β-strands G and H.

A comparison between the fatty-acid-binding proteins isolated from *S. gregaria* (molecule II) and *M. sexta* L. is given in Figure 7. The α-carbons for these two proteins

Table 4: Dihedral Angles for Reverse Turns

		molecule I				molecule II			
amino acid	type	ϕ_2	ψ_2	ϕ_3	ψ_3	ϕ_2	ψ_2	ϕ_3	ψ_3
Lys 2–Ala 5	~I	−64.9	−19.6	−86.9	−6.4	−60.4	−10.6	−105.4	2.5
Phe 4–Ile 7	II	−55.6	137.6	87.8	−11.6	−62.8	143.0	79.3	−2.9
Lys 22–Gly 25	I	−55.5	−42.1	−88.0	8.4	−58.6	−40.0	−90.9	14.1
Asp 47–Lys 50	II′	98.5	−148.3	−85.1	8.1	70.9	−112.6	−95.3	6.7
Thr 57–Lys 60	I	−74.6	−59.1	−49.7	−67.0	−63.8	−33.7	−98.8	−6.8
Lys 67–Glu 70	II	−45.7	127.7	82.5	−2.2	−55.6	131.7	77.7	−1.9
Thr 76–Gly 79	I	−54.0	−32.4	−89.1	4.7	−52.4	−44.2	−81.7	0.3
Glu 90–Lys 93	~I	−64.5	−47.4	−108.9	17.2	−73.6	−22.5	−119.8	16.4
Ser 111–Gln 114	I	−71.7	−25.3	−97.8	−15.8	−62.2	−32.5	−101.7	−10.4
Leu 121–Leu 124	II′	60.1	−122.9	−92.6	−4.8	60.6	−117.2	−95.1	−14.6

superimpose with a root-mean-square deviation of 1.3 Å for 91 structurally equivalent atoms. Of the 23 conserved water molecules described above, five are conserved between these two insect proteins as shown in Figure 7. Three of these waters bridge together β-strands A and B, B and C, and H and I. Another conserved water forms a hydrogen bond with a backbone amide nitrogen. The fifth water, located between β-strands D and E, is likewise conserved in adipocyte lipid-binding protein, cellular retinol-binding protein II, and intestinal fatty-acid-binding protein (Banaszak *et al.*, 1994).

Figure 8 shows a comparison between the fatty-acid-binding proteins from locust (molecule II) and human muscle. The α-carbons for these two proteins superimpose with a root-mean-square deviation of 0.77 Å for 89 structurally equivalent atoms. As shown in Figure 8, there are two water molecules that superimpose within 1.0 Å between these two proteins. One of these hydrogen bonds to the backbone amide nitrogen of Gly 27 (Gly 26 in the human muscle protein) which is located in the loop region connecting the two α-helices. The other water is again located in the gap between β-strands D and E where it stabilizes the reverse turn and bridges this turn to β-strand F.

Since the protein used for the crystallization experiments contained equimolar amounts of endogenous fatty acids (Haunerland & Chisholm, 1990), it was expected that electron density corresponding to the ligand would be obvious and that it would be located within the β-barrel. After least-squares refinement of the model, however, it became clear that some of the fatty-acid ligand had been removed during the crystallization process. The addition of 2-methyl-2,4-pentanediol to the crystallization medium most likely removed the fatty acid from the protein. Consequently, the structure described here is of the apoprotein. It should be noted, however, that there may not have been complete removal of the fatty-acid ligand from the protein and that some of the waters built into the binding pocket may correspond to portions of the lipid.

In the case of the fatty-acid-binding protein from *M. sexta* L., one of the carboxylate oxygens of the fatty acid was shown to interact with the hydroxyl group of Tyr 129 and the δ-guanidino group of Arg 127, while the other carboxylate oxygen was shown to lie within hydrogen-bonding distance of Gln 39 and a sulfate group (Benning *et al.*, 1992). The sulfate may have been an artifact resulting from crystallization of the protein out of ammonium sulfate at pH 4.5. Removal of the sulfate group could possibly allow the amino group of Lys 105 to move into position to form a hydrogen bond with the carboxylate moiety of the fat. In the *S. gregaria* fatty-acid-binding protein, Tyr 130 and Arg 128 adopt orientations similar to those observed in the *M. sexta* L. protein. However, in the *S. gregaria* protein, Gln 39 and Lys 105 have been replaced by Ile 41 and Arg 108, respectively.

With respect to the hydrophobic tail of the fatty-acid ligand, its positioning in the *M. sexta* L. protein was found to be quite different from that observed in all vertebrate fatty-acid-binding proteins (Benning *et al.*, 1992). In the *M. sexta* L. protein, residue 32 is a leucine. This position in the vertebrate proteins is typically a glycine or an alanine. Because of this bulky side chain, the fatty-acid ligand in the *M. sexta* L. protein curls off at C6 in a direction opposite to that observed in the binding proteins from vertebrate sources. In the *S. gregaria* protein, this structurally equivalent residue is Gly 34, suggesting that the fatty acid may bind in a manner more like that observed in the vertebrate proteins. On the other hand, in the heart muscle protein, for example, residue 75 is an alanine (Müller-Fahrnow *et al.*, 1991). In the *S. gregaria* protein, the structurally equivalent residue is Leu 77. The bulky side chain of Leu 77 would collide with C8 of the fatty acid if it were to bind to the locust protein in exactly the same manner as observed for the vertebrate proteins.

DISCUSSION

The tertiary structure of locust flight muscle fatty-acid-binding protein follows the same overall three-dimensional motif that is characteristic for intracellular lipid-binding proteins. As expected from the primary structures, the protein is similar in its tertiary structure to the fatty-acid-binding protein from vertebrate muscle. In contrast, there are larger structural differences visible between locust fatty-acid-binding protein and the midgut fatty-acid-binding protein from another insect, *M. sexta* L. The structural conservation of muscle fatty-acid-binding proteins is evident when their α-carbon backbones are compared. The root-mean-square deviation between the locust and the human muscle fatty-acid-binding proteins for 89 structurally equivalent residues is 0.77 Å, as opposed to 1.3 Å calculated between the *M. sexta* L. midgut fatty-acid-binding protein and the locust molecule.

Although the present study is of the apoprotein, the binding cavity and the orientation of the amino acids implicated in fatty-acid binding suggest that the ligand probably binds in a manner similar to that observed in the vertebrate muscle fatty-acid-binding protein. Specifically, the carboxylate head group of the ligand should follow the "P2 motif" found in muscle fatty-acid-binding protein, P2 myelin, and adipocyte lipid-binding protein (Banaszak *et al.*, 1994). The hydrophobic tail is most likely not bound in the conformation observed in the *M. sexta* L. protein (Benning *et al.*, 1992). Rather, its orientation may resemble that found in vertebrate fatty-acid-binding proteins, although subtle differences must exist since the bulky side chain of Leu 77 would sterically interfere with C8 of the hydrocarbon chain. Details concerning protein–lipid interactions in the *S. gregaria* system, however, must await the results from the crystallization experiments of the holoprotein that are currently in progress.

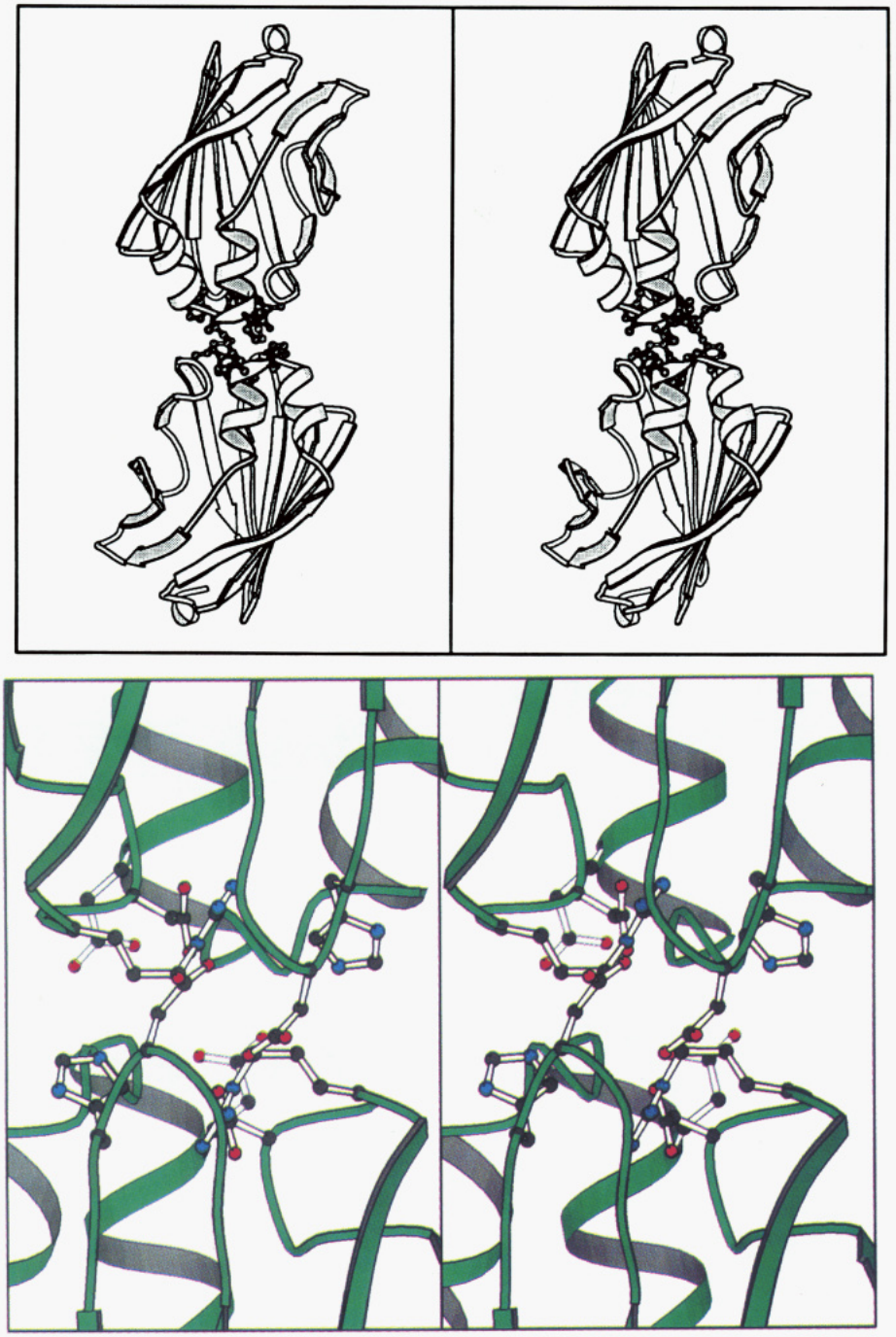

FIGURE 5: Molecule–molecule interactions. (a, top) The two fatty-acid-binding proteins located in the asymmetric unit pack with a 2-fold rotational axis located nearly perpendicular to the plane of the page. (b, bottom) A close-up view of the molecule–molecule interface is shown. The model has been rotated from the orientation in (a) by 180° about the vertical axis. The molecule–molecule interface region is formed by residues Gly 25, Val 26, Gly 27, Ala 28, Ile 29, Glu 30, Asp 78, Gly 79, Arg 80, Asp 101, and His 102. For the sake of clarity, only polar residues are shown as ball-and-stick models.

The high structural homology between the muscle fatty-acid-binding proteins is rather surprising considering the evolutionary distance between locusts and humans. Vertebrates and invertebrates are commonly believed to have branched off a common ancestor in the Cambrian period. The fact that, in spite of more than 500 million years of independent development, the structures of human and locust muscle fatty-acid-binding protein are so similar implies that many details of the structure, including the binding cavity and surface properties, may be required for the function of the protein in muscle cells. The biochemical roles of fatty-acid-binding proteins are not entirely clear. While it is generally assumed that these proteins facilitate fatty-acid transport, evidence has been presented that they influence fatty acid metabolism in more direct ways (Veerkamp *et al.*, 1991). Moreover, the existence of a multigene family of tissue-specific fatty-acid-binding proteins strongly suggests that the different proteins have specialized metabolic functions. While some differences in binding affinities exist, the binding cavity is very similar in most of these proteins. It may be more instructive to consider the differences in surface properties between the various tissue-specific proteins. Not surprisingly, the high degree of structural similarity between human and locust muscle fatty-acid-binding protein includes many surface areas.

In this regard, it is interesting to consider the intermolecular interactions between the two fatty-acid-binding protein

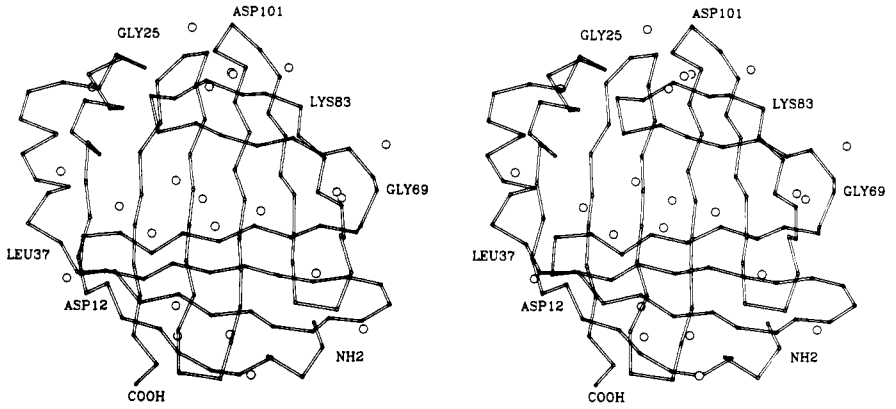

FIGURE 6: Location of the conserved water molecules. There are 23 conserved water molecules between the two fatty-acid-binding proteins in the asymmetric unit and are indicated by the open circles.

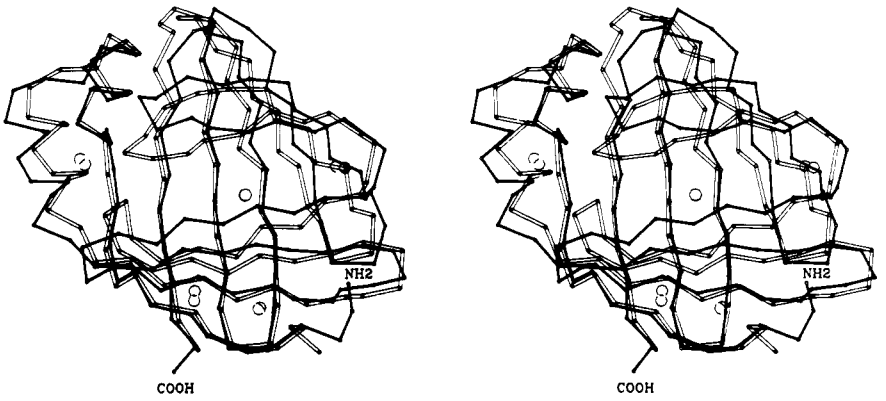

FIGURE 7: Superposition of the α-carbons for the insect fatty-acid-binding proteins isolated from *M. sexta* L. and *S. gregaria*. The models for the *M. sexta* L. and *S. gregaria* proteins are displayed in open and filled bonds, respectively. There are five conserved water molecules between these two proteins as indicated by the open circles. The orientation is the same as in Figure 6.

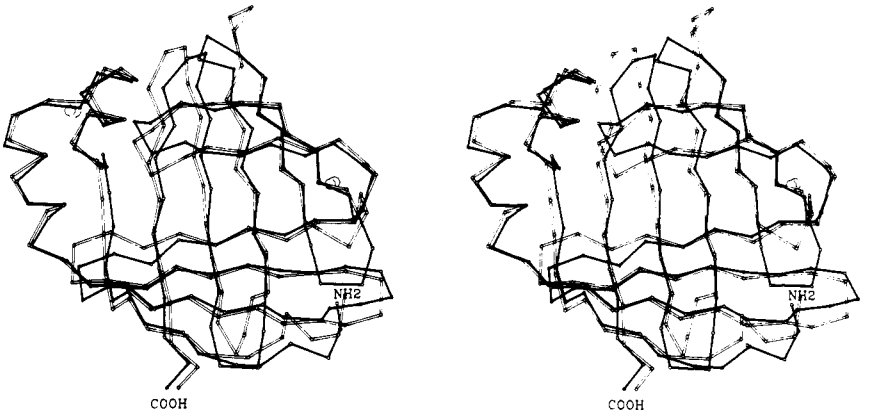

FIGURE 8: Superposition of the α-carbons for *S. gregaria* fatty-acid-binding protein and recombinant human muscle fatty-acid-binding protein (Zanotti *et al.*, 1992). The models for the human and locust proteins are displayed in open and filled bonds, respectively. There are two conserved water molecules between these two proteins as indicated by the open circles. The orientation is the same as in Figure 6. X-ray coordinates for the human muscle protein were obtained from the Brookhaven Protein Data Bank.

molecules located in the asymmetric unit. These interactions involve amino acid side chains located in the area of the helix–turn–helix motif and the turn between β-strands E and F. It is the same area that has been speculated to form the portal that, through dynamic changes, allows entry of fatty acids into the binding cavity (Sacchettini *et al.*, 1989; Banaszak *et al.*, 1994). Such dynamic changes are further supported by the recent findings of Lücke *et al.* (1992), who investigated the solution structure of bovine heart fatty-acid-binding protein. These authors found a high degree of heterogeneity in the second helix of the helix–turn–helix motif, indicating increased conformational flexibility of this part of the protein structure. Interactions between two fatty-acid-binding protein molecules, like those observed in the crystalline lattice, could stabilize the portal area and thus could influence the loading and unloading of lipids.

At present, it is impossible to predict whether the interactions between two protein molecules in the crystal have any physiological significance. Although fatty-acid-binding proteins usually appear to be monomeric proteins, there is some evidence to suggest that the molecules from pig heart, at certain concentrations, exist in different oligomeric states (Fournier *et al.*, 1983). Aggregation was found to influence ligand exchange with other cellular organelles (Fournier & Richard, 1988). Considering the very large concentrations of this protein found in locust flight muscle (Haunerland *et al.*, 1992,

1993), intermolecular interactions between fatty-acid-binding proteins may indeed exist *in vivo*. Further analysis of intermolecular interactions of fatty-acid-binding proteins from muscle and from other tissues could shed light on the tissue-specific physiological functions of these proteins.

ACKNOWLEDGMENT

We thank Drs. Grover Waldrop and Matt Benning for aid in preparing the figures and for critically reading the manuscript.

REFERENCES

Banaszak, L., Winter, N., Xu, Z., Bernlohr, D. A., Cowan, S., & Jones, T. A. (1994) *Adv. Protein Chem. 45*, 90–150.

Beenakkers, A. M. Th., Van der Horst, D. J., & Van Marrewijk, W. J. (1985) in *Comprehensive Insect Biochemistry, Physiology, and Pharmacology* (Kerkut, G. A., & Gilbert, L. I., Eds.) Vol. 10, pp 451–486, Pergamon Press, New York.

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

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rogers, J. R., Kennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.

Bricogne, G. (1976) *Acta Crystallogr., Sect. A 32*, 832–837.

Cowan, S. W., Newcomer, M. E., & Jones, T. A. (1993) *J. Mol. Biol. 230*, 1225–1246.

Crabtree, B., & Newsholme, E. A. (1975) in *Insect Muscle* (Usherwood, P. N. R., Ed.) pp 405–491, Academic Press, London.

Fournier, N. C., & Richard, M. A. (1988) *J. Biol. Chem. 263*, 14471–14479.

Fournier, N. C., Zuker, M., Williams, R., & Smith, I. P. (1983) *Biochemistry 22*, 1863–1872.

Fox, G. C., & Holmes, K. C. (1966) *Acta Crystallogr. 20*, 886–891.

Haunerland, N. H., & Chisholm, J. M. (1990) *Biochim. Biophys. Acta 1047*, 233–238.

Haunerland, N. H., Andolfatto, P., Chisholm, J. M., Wang, Z., & Chen, X. (1992) *Eur. J. Biochem. 210*, 1045–1051.

Haunerland, N. H., Chen, X., Andolfatto, P., Chisholm, J. M., & Wang, Z. (1993) *Mol. Cell. Biochem. 123*, 153–158.

Janin, J., Miller, S., & Chothia, C. (1988) *J. Mol. Biol. 204*, 155–164.

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

Kabsch, W. (1988a) *J. Appl. Crystallogr. 21*, 67–71.

Kabsch, W. (1988b) *J. Appl. Crystallogr. 21*, 916–924.

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

Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379–400.

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

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

Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491–497.

Mishkin, S., Stein, L., Gatmaitan, Z., & Arias, I. M. (1972) *Biochem. Biophys. Res. Commun. 47*, 997–1003.

Moser, D., Tendler, M., Griffiths, G., & Klinkert, M.-Q. (1991) *J. Biol. Chem. 266*, 8447–8454.

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

Navaza, J. (1987) *Acta Crystallogr. A43*, 645–653.

Ockner, R. K., Manning, J. A., Poppenhausen, R. B., & Ho, W. K. L. (1979) *Science 177*, 56–58.

Price, H. M., Ryan, R. O., & Haunerland, N. H. (1992) *Arch. Biochem. Biophys. 297*, 285–290.

Rossmann, M. G., & Argos, P. (1975) *J. Biol. Chem. 250*, 7525–7532.

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

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

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

Smith, A. F., Tsuchida, K., Hanneman, E., Suzuki, T. C., & Wells, M. A. (1992) *J. Biol. Chem. 267*, 380–384.

Tronrud, D. E., Ten Eyck, L. F., & Matthews, B. W. (1987) *Acta Crystallogr., Sect. A 43*, 489–501.

Van der Horst, D. J., van Doorn, J. M., Passier, P. C. C. M., Vork, M. M., & Glatz, J. F. C. (1993) *Mol. Cell. Biochem. 123*, 145–152.

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

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

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

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

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