Crystal Structure of Recombinant Murine Adipocyte Lipid-Binding Protein†,‡

Zhaohui Xu,§ David A. Bernlohr,∥ and Leonard J. Banaszak*,§

Department of Biochemistry, University of Minnesota Medical School, Minneapolis, Minnesota 55455, and Department of Biochemistry, College of Biological Sciences, University of Minnesota, St. Paul, Minnesota 55108

Received August 6, 1991; Revised Manuscript Received October 23, 1991

ABSTRACT: Adipocyte lipid-binding protein (ALBP) is the adipocyte member of an intracellular hydrophobic ligand-binding protein family. ALBP is phosphorylated by the insulin receptor kinase upon insulin stimulation. The crystal structure of recombinant murine ALBP has been determined and refined to 2.5 Å. The final *R* factor for the model is 0.18 with good canonical properties. Crystalline ALBP has a conformation which is essentially identical to that of intestinal fatty acid binding protein and myelin P2 protein. Although the crystal structure is of the apo- form, a cavity resembling that in other family members is present. It contains a number of bound and implied unbound water molecules and shows no large obvious portal to the external milieu. The cavity of ALBP, which by homology is the ligand-binding site, is formed by both polar and hydrophobic residues among which is tyrosine 19. Y19 is phosphorylated by the insulin receptor kinase as described in the accompanying paper [Buelt, M. K., Xu, Z., Banaszak, L. J., & Bernlohr, D. A. (1992) *Biochemistry* (following paper in this issue)]. By comparing ALBP with the earlier structural results on intestinal fatty acid binding protein, it is now possible to delineate conserved amino acids which help form the binding site in this family.

The intracellular trafficking of hydrophobic compounds appears to involve a family of proteins now referred to as hydrophobic ligand-binding proteins. Different cell types express high levels of one or more of this gene family [for reviews, see Sweetser et al. (1987) and Matarese et al. (1990)]. The members of this family may have relatively major differences in their amino acid sequences as well as in the specificity for binding different hydrophobic ligands. One of these proteins is adipocyte lipid binding protein, ALBP;[1] ALBP is also known as 422, aP2, or p15 in the biochemical literature (Bernlohr et al., 1985; Hunt et al., 1986; Hresko et al., 1988). It is a 14.6-kDa protein expressed exclusively in adipose cells or cultured adipogenic cell lines (Bernlohr et al., 1984). In vitro assays showed that ALBP binds both long chain fatty acid and retinoic acid with 1 mol of ligand bound per mole of protein (Matarese & Bernlohr, 1988). Although the precise role of ALBP in adipocyte lipid metabolism is not known, it is likely that the protein serves as an intracellular fatty acid carrier facilitating the uptake and utilization of these hydrophobic lipids (Waggoner & Bernlohr, 1990).

Other intracellular hydrophobic ligand-binding proteins have been purified from a variety of tissue sources. These proteins are considered to form a multigene family based on their gene structures and amino acid sequence similarity (Matarese et al., 1990). Crystal structures have been obtained for two members of the intracellular lipid-binding protein family: myelin P2 at 2.7-Å resolution (Jones et al., 1988) and both the apo- and holo- forms of rat intestinal fatty acid binding protein, IFABP, at 2.0-Å resolution (Sacchettini et al., 1989a,b).

Despite the fact that P2 and IFABP have only 29% amino acid identities and somewhat different ligand-binding selectivity, they have highly similar conformations. The tertiary structure of both proteins is described as a "β-clam" and consists of two nearly orthogonal β-sheets formed by 10 antiparallel β-strands. The two sheets actually surround an internal ligand-binding site. In the case of P2, the hydrophobic binding site accommodates retinol, retinoic acid, and oleic acid (Uyemura et al., 1984), while in IFABP the binding site is mainly for long-chain fatty acids (Lowe et al., 1987).

The overall conformation of these intracellular proteins also resembles extracellular hydrophobic ligand-binding proteins. The conformational similarity is necessarily limited in that the extracellular family has members with 160 or more amino acids. Nonetheless, the two segments of antiparallel β-sheets are present. In cases where structures are known, the ligand-binding sites are contained between the two sheets. This family includes serum retinol-binding protein (Newcomer et al., 1984), β-lactoglobulin (Sawyer et al., 1985; Papiz et al., 1986; Monaco et al., 1987), and insecticyanin (Holden et al., 1987; Huber et al., 1987a,b). The similarities between the intra- and extracellular proteins suggest that the antiparallel β-clam structure serves as an important motif for hydrophobic ligand binding and transport.

Lipid metabolism in adipocytes is largely controlled via the molecular actions of insulin. Insulin, by binding to cell surface receptors, generates a series of regulatory events among which is the control of triacylglycerol synthesis (Flatt, 1970). Interestingly, ALBP has been shown to be a substrate of the activated insulin receptor kinase in 3T3-L1 adipocytes (Bernier et al., 1987). The phosphorylation is exclusively on tyrosine 19 (Hresko et al., 1988; Chinander & Bernlohr, 1989). Fatty acid binding has a profound effect on the phosphorylation reaction by lowering the K_m (Hresko et al., 1990; Buelt et al., 1991a). In the accompanying paper, we demonstrate that phosphorylation of ALBP blocks ligand binding (Buelt et al.,

† This study was supported by research grants from the NIH to L.J.B. (GM 13925) and to D.A.B. from the NSF (NSF/DMB 8552942) and the American Diabetes Association, Minnesota Affliate.

‡ The coordinates for the final structure have been deposited in the Protein Data Bank at the Brookhaven National Laboratory.

* Author to whom correspondence should be addressed: Department of Biochemistry, University of Minnesota, 435 Delaware St. SE, Minneapolis, MN 55455.

§ University of Minnesota Medical School.

∥ College of Biological Sciences, University of Minnesota.

[1] Abbrevations: ALBP, adipocyte lipid-binding protein; IFABP, intestinal fatty acid binding protein; P2, bovine myelin lipid binding protein; PC, Patterson correlation.

0006-2960/92/0431-3484$03.00/0

Table I: Progress of Refinement

	no. of reflections[a,b]	R factor	B (Å²)	solvent included	RMSD bond length (Å)	RMSD bond angle (deg)	RMSD planarity (deg)
(1) starting model	2976	0.458	15.0		0.065	4.12	9.015
(2) rigid-body refinement	2976	0.456	15.0		0.065	4.12	9.012
(3) simulated annealing	4579	0.235	group		0.019	3.17	1.506
(4) model rebuilt using $(2F_o - F_c)$ and $(F_o - F_c)$ electron density maps							
(5) simulated annealing	4579	0.220	indiv		0.018	3.77	1.408
(6) model rebuilt using $(2F_o - F_c)$ and $(F_o - F_c)$ electron density maps, H_2O included							
(7) simulated annealing	4579	0.197	indiv	31	0.018	3.73	1.366
(8) model rebuilt using $(2F_o - F_c)$ and $(F_o - F_c)$ electron density maps, H_2O included							
(9) simulated annealing	4579	0.172	indiv	88	0.016	3.47	1.139
(10) model rebuilt using $(2F_o - F_c)$ and $(F_o - F_c)$ electron density maps, H_2O included							
(11) simulated annealing	4773	0.183	indiv	69	0.017	3.46	1.070

[a] $F > 2\sigma(F)$, except for step 11, where all data were included. [b] Steps 1 and 2 included data from 10 to 3 Å. For all subsequent steps, the data were from 8 to 2.5 Å.

1992). To assess the structure/function relationship of ligand binding and phosphorylation, a crystal structure study has been undertaken. Murine ALBP has been cloned and expressed in *Escherichia coli*, purified to homogeneity, and crystallized (Xu et al., 1991). This crystalline form of ALBP contains no ligand. In the text which follows, the refined crystal structure of ALBP to 2.5-Å resolution is reported.

Materials and Methods

Crystals and X-ray Data Collection. Detailed information concerning protein purification, crystallization, and X-ray data collection can be found in a previous report (Xu et al., 1991) and will be mentioned here in a summary form. Recombinant murine apo-ALBP crystallizes in the orthorhombic space group $P2_12_12_1$ with the following unit cell dimensions: a = 34.4 Å, b = 54.8 Å, and c = 76.3 Å. The asymmetric unit contains one molecule with a molecular weight of 14 500. The entire diffraction data set was collected on one crystal. In the resolution range ∞–2.5 Å, 5115 of the 5227 theoretically possible reflections were measured. Unless otherwise noted, the diffraction data with intensities greater than 2σ were used for structure determination and refinement. As can be seen in Table I, this included about 96% of the measured data.

Molecular Replacement. The tertiary structure of crystalline ALBP was solved by using the molecular replacement method incorporated into the XPLOR computer program (Brunger et al., 1987). The refined crystal structure of myelin P2 protein without solvent and fatty acid was used as the probe structure throughout the molecular replacement studies. We are indebted to Dr. A. Jones and his colleagues for permission to use their refined P2 coordinates before publication.

(*1*) *Rotation Search.* The rotation search was carried out using the Patterson search procedures in XPLOR. The probe Patterson maps were computed from structure factors calculated by placing the P2 coordinates into an orthorhombic cell with 100-Å edges. One thousand highest Patterson vectors in the range of 5–15 Å were selected and rotated using the pseudoorthogonal Eulerian angles (θ_+, θ_2, θ_-) as defined by Lattman(1985). The angular search interval for θ_2 was set to 2.5°; intervals for θ_+ and θ_- are functions of θ_2. The rotation search was restricted to the asymmetric unit θ_- = 0–180°, θ_2 = 0–90°, and θ_+ = 0–720° for the $P2_12_12_1$ space group (Rao et al., 1980). XPLOR produces a sorted list of the correlation results simplifying final interpretation (Brunger, 1990).

(*2*) *Patterson Correlation Refinement.* To select which of the orientations determined from the rotation search was the correct solution, a Patterson correlation refinement of the peak list of the rotation function was performed. This was carried out by minimization against a target function defined by Brunger (1990) and as implemented in XPLOR. The search model, P2, was optimized for each of the selected peaks of the rotation function.

(*3*) *Translation Search.* A translation search was done by using the P2 probe molecule oriented by the rotation function studies and refined by the Patterson correlation method. The translation search employed the standard linear correlation coefficient between the normalized observed structure factors and the normalized calculated structure factors (Fujinaga & Read, 1987; Brunger, 1990). X-ray diffraction data from 10–3-Å resolution were used. Search was made in the range x = 0–0.5, y = 0–0.5, and z = 0–0.5, with the sampling interval 0.0125 of the unit cell length.

Structure Refinement. The refinement of the structure was based on an energy function approach (Brunger et al., 1987): arbitrary combinations of empirical and effective energy terms describing crystallographic data as implemented in XPLOR. Molecular model building was done on an IRIS Workstation (Silicon Graphics) with the software TOM, a version of FRODO (Jones, 1978).

The initial model of ALBP was built by simply putting the amino acid sequence of ALBP into the molecular structure of myelin P2 protein. After a 20-step rigid-body refinement of the positions and orientations of the molecule, crystallographic refinement with simulated annealing was carried out using a slow-cooling protocol (Brunger et al., 1989, 1990). Temperature factor refinement of grouped atoms, one for backbone and one for side-chain atoms for each residue, was initiated after the R factor dropped to 0.249.

The new coordinates were checked and adjusted against a $(2|F_o| - |F_c|)$ and a $(|F_o| - |F_c|)$ electron density map, where $|F_o|$ and $|F_c|$ are the observed and calculated structure factor amplitudes. Phases are calculated from the crystal coordinates. The Fourier maps were calculated on a grid corresponding to one-third of the high-resolution limit of the input diffraction data. All residues were inspected on the graphics system at several stages of the refinement. The adjustments were made on the basis of following criteria: (a) that an atom was located in low electron density in the $(2|F_o| - |F_c|)$ map or negative electron density in the $(|F_o| - |F_c|)$ map; (b) that the parameters for the ϕ,ψ angles placed the residue outside the acceptable regions in the Ramachandran diagram. Iterative refinement and model adjustment against a new electron density map was

Table II: Molecular Orientation and Position

(A) Molecular Orientation in the Crystal Cell

peak index	before PC refinement θ_1	θ_2	θ_3	after PC refinement θ_1	θ_2	θ_3	rotation function	Patterson correlation
1	148.3	47.5	354.0	145.4	46.1	351.5	10.918	0.1398
2	139.6	47.5	351.8	145.4	46.1	351.5	9.830	0.1397
3	150.6	47.5	2.7	145.9	44.5	350.1	9.736	0.1265
6	154.6	50.0	340.6	145.8	46.1	351.1	9.167	0.1392
7	135.5	47.5	6.9	145.4	46.2	351.5	8.983	0.1398
90	150.5	40.0	345.5	145.5	46.1	351.4	7.453	0.1396
98	126.7	40.0	6.7	146.5	46.5	352.8	7.403	0.1286

(B) Position of Molecular Center in the Crystal Cell

TF_{max}	R factor (%)	x	y	z	Peak Height[a] correct	noise	S/N
0.419	47.0	8.609	23.272	10.497	10.7	8.3	1.3

[a] Arbitrary units of correlation function.

carried out until the R factor appeared unaffected. Isotropic temperature factors for individual atoms were then included in the refinement.

The next stage of the crystallographic study included the location of solvent molecules. They were identified as well-defined positive peaks in the electron density maps within hydrogen-bonding distance of appropriate protein atoms or another solvent atom. Solvent atoms were assigned as water molecules and refined as oxygen atoms. Those that refined to positions too close to other atoms, ended up located in low electron density, or had associated temperature factors greater than 50 Å^2 were removed from the coordinate list in the subsequent stage. The occupancy for all atoms, including solvent molecules, was kept at 1.0 throughout the refinement. Detailed progress of the crystallographic refinement is given in Table I.

Results

Molecular Replacement. From the initial rotation search, the 101 highest peaks were chosen for further study. These are shown in Figure 1A. The highest peak of the rotation function had a value 4.8 times the standard deviation above the mean and 1.8 times the standard deviation above the next highest peak. This orientation was consistently the highest peak for diffraction data within the resolution ranges of 10–5, 7–5, and 7–3 Å. Apart from peak number 1, six strong peaks emerged after PC refinement, as can be seen in Figure 1B. These peaks all corresponded to approximately the same orientation as peak number 1. Three of them were initially away from that orientation and converged to it during the PC refinement. The final results from the rotation function studies are given in Table IIA.

A translation search as implemented in XPLOR was used to find the molecular position of the now oriented P2 probe in the ALBP unit cell. The results are summarized in Table IIB. Only a single position emerged at $x = 0.250$, $y = 0.425$, $z = 0.138$ with a correlation coefficient of 0.419. The initial R factor for the P2 coordinates in the determined molecular orientation and position was 0.470 including X-ray data in the resolution range of 10–3 Å. A rigid-body refinement of orientation and position reduced the starting R factor to only 0.456, probably attesting to the efficacy of the Patterson refinement in XPLOR.

Refined Structure of apo-ALBP. The refined ALBP structure has an R factor 0.183 when all observed X-ray data (4773 reflections) between 8.0 and 2.5 Å are included. The rms deviation of bond lengths, bond angles, and planarity from ideality is 0.017 Å, 3.46°, and 1.07°, respectively. An estimate of the upper limit of error in atomic coordinates is obtained

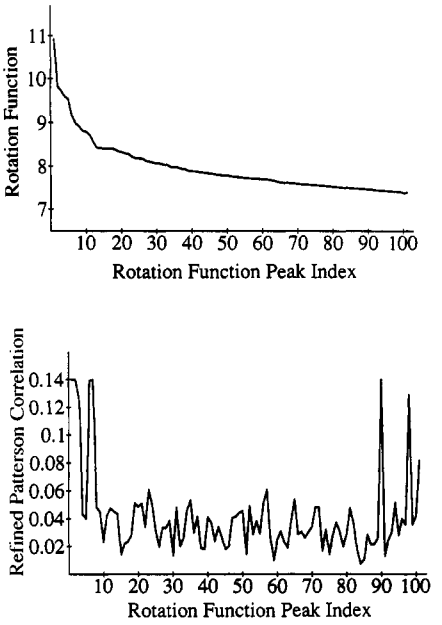

FIGURE 1: Rotation function results—P2 into crystalline ALBP. (A, top) Plot of the 101 best solutions to the rotation function each peak numbered in the horizontal direction. The correlation between the Patterson's of the probe molecule and the measured ALBP X-ray results are shown in the vertical direction and are given in arbitrary units. (B, bottom) Description of the rotation studies after Patterson correlation refinement. The peak numbers plotted in both panels A and B are the same.

by the method of Luzzati (1952). Figure 2A,B summarizes the overall refined model. The plot presented in Figure 2A suggests that the upper limit for the mean error of the refined ALBP coordinates is around 0.34 Å. The mean temperature factors for main-chain and side-chain atoms are plotted in Figure 2B.

Careful examination of $(2|F_o| - |F_c|)$ and $(|F_o| - |F_c|)$ maps at each refinement step led to the conclusion that no bound ligand was present. There was no continuous positive electron density present near the ligand-binding site as identified in both P2 (Jones et al., 1988) and IFABP (Sacchettini et al., 1989a). The absence of bound fatty acid in crystalline ALBP is consistent with the chemical modification experiment which indicates ALBP purified from *E. coli* is devoid of fatty acid (Xu et al., 1991). The final refined coordinate list includes 1017 protein atoms and 69 water molecules.

Structural Properties of Crystalline ALBP. A Ramachandran plot of the main chain dihedral angles ϕ and ψ is shown in Figure 3. In the refined model, 13 residues have

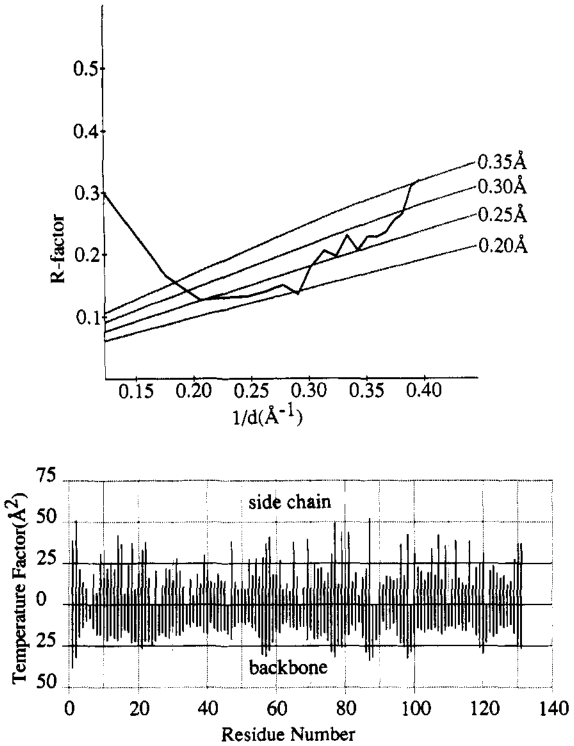

FIGURE 2: ALBP refinement results. (A, top) Theoretical estimates of the rms positional errors in atomic coordinates according to Luzzati (1952) are shown superimposed on the curve for the ALBP diffraction data. The coordinate error estimated from this plot is 0.25 Å with an upper limit of about 0.35 Å. (B, bottom) Mean values of the main-chain and side-chain temperature factors are plotted versus the residue number. The temperature factors are those obtained from the final refinement cycles.

Table III: Structural Properties of ALBP

```
                       1         2         3         4         5
ALBP[1]       CDAFVGTWKLVSSENFDDYMKEVGVGFATRKVAGMAKPNMIISVNGDLVTIRSESTFKN
conserved[2]     * ****    **    *   *     **          *   *   *    * * *
positive φ[3]      *        *        *                     *
cav contact[4]         *     *  **  * *   *  ******* *          ***** ***
sec str[5]         ββββAββββ αααIαααα   αααIIαααα   βββBβββ  βββCβββ

              6         7         8         9         10
ALBP          TEISFKLGVEFDEITADDRKVKSIITLDGGALVQVQKWDGKSTTIKRKRD
conserved         * **  *    **        *     *    *  *      *
positive φ           *         *          *          *
cav contact   ***         ***** *              * *      * * *
sec str       ββDββ     ββEβ     ββββFββββ  ββββGβββ  ββββHβββββ

              11        12       13
ALBP          GDKLVVECVMKGVTSTRVYERA
conserved                 *   *
positive φ    *         **
cav contact        * *        * *
sec str         ββββIβββ  ββββJββββ
```

[1] The N-terminal Met is missing in the recombinant protein and therefore in the crystal structure. [2] Conserved in ALBP, IFABP, and P2. [3] Positive ϕ value for peptide torsional angles in ALBP and IFABP. [4] Residues which contain one or more atoms in contact with the hydrophobic ligand-binding cavity. [5] The secondary structural elements found in ALBP are marked with β, indicating strands of the sheets, and α, for helical residues. The definitions of α- and β-structure are according to Kabsch and Sander (1983).

positive ϕ angles, 9 of which belong to glycine residues. There are 11 glycine residues in ALBP, all associated with good quality electron density. Seven of the nine glycines with positive ϕ angles are conserved in three fatty acid binding proteins whose structures are known (ALBP, IFABP, and P2), and these are listed in Table III. G46 and G67 are conserved in all hydrophobic ligand binding proteins. All conserved glycine residues with positive ϕ angles in ALBP also have positive ϕ angles in IFABP (Sacchettini et al., 1989a). As seen in Figure 4, all the glycine residues with positive ϕ angles are located in the turns in the polypeptide chain with the exception of G6. G6 has a positive ϕ but is not in a turn. Clearly a consensus amino acid sequence for hydrophobic

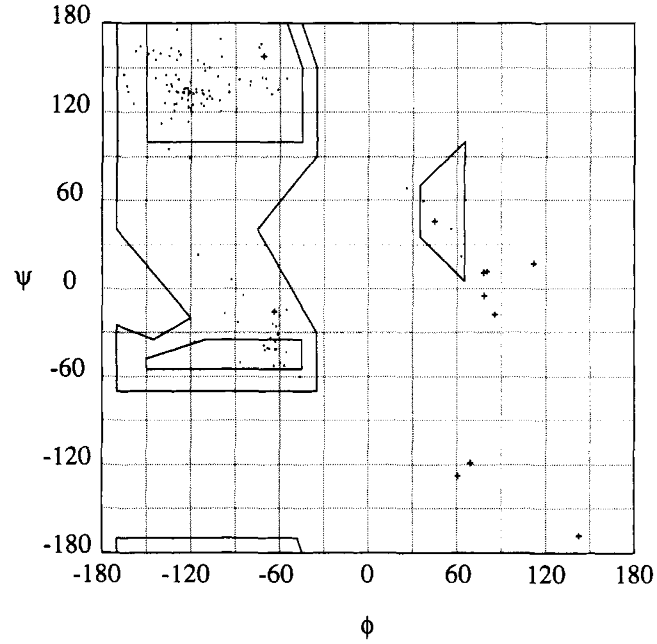

FIGURE 3: Ramachandran plot of the crystallographic model of ALBP. The main-chain torsion angle ϕ (N–C_α bond) is plotted versus ψ (C_α–C bond). The following symbols are used: (·) non-glycine residues; (+) glycine residues. The enclosed areas of the plot show sterically allowed angles.

ligand-binding proteins must include the glycines.

There are four non-glycine residues with positive ϕ angles. Among them, Asn15 is conserved in all intracellular hydrophobic ligand-binding proteins. Three residues, N15, D77, and K120, also have positive ϕ angles in the IFABP structure (Sacchettini et al., 1989a). The location of these residues is also shown in Figure 4. K120 together with G121 forms one of the three turns that have two residues with positive ϕ values. The other two are 88–89 and 98–99.

The surface and interior properties of ALBP molecule can be determined from the coordinates by assessing the solvent accessibility using Lee and Richards (1971) algorithm as implemented in XPLOR. In the structure of ALBP, it is useful to remember that a relatively large cavity is to be found in the center of the molecule and that this cavity must contain some disordered water particularly in the apo- form of the structure. A solvent molecule released in the internal cavity would find "cavity" accessible atoms, but these will not appear in a typical accessibility study. Using a solvent probe with a radius of 1.6 Å, the variation in solvent accessibility is shown in Figure 5A,B.

The accessibilities for the main-chain atoms show that most of the β-strands and the two α-helices contain buried atoms while the turns are generally more accessible to solvent. This is most striking when looking at the backbone atoms in the C-terminal half of the molecule as can be seen in Figure 5A. The four sharp accessibility spikes all appear at sharp turns between β-strands.

The values for side-chain accessibility shown in Figure 5B resemble a noise curve which is fairly typical of periodical secondary structures. With exceptions noted below, those side chains with large accessible surface area are mainly charged or neutral polar amino acids, while the probe unaccessible area is composed of the side chains of hydrophobic and also some polar amino acid residues. The latter might be taken to be typical of these cavity proteins. Three arginine residues, R78, R106, and R126, have a relatively low value of accessible surface. They appear to be located in the interior of the

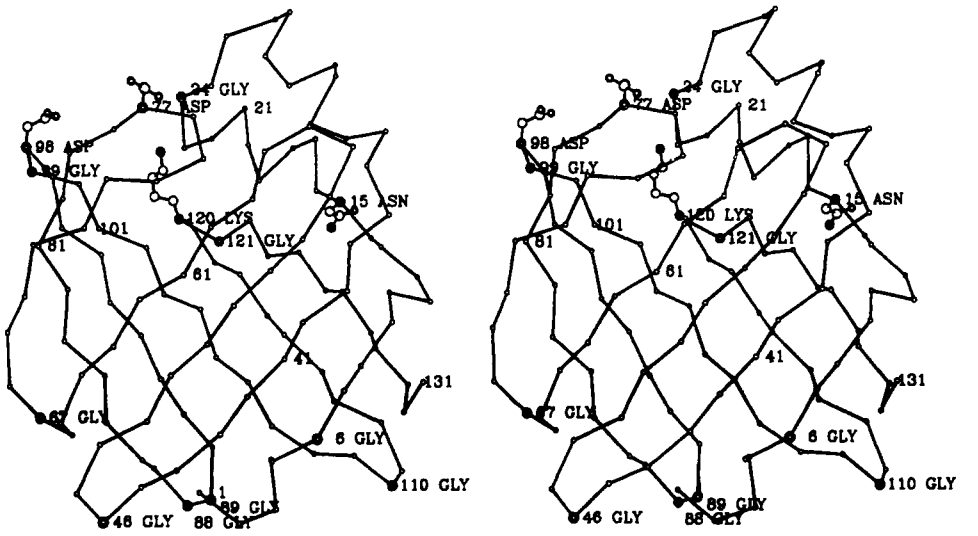

FIGURE 4: Stereodiagram of ALBP. The stereodiagram depicts a C_α model of ALBP. Residues with positive ϕ value for peptide torsional angles are also shown in the form of ball and stick representations. To the right of each C_α atom is the amino acid name and sequence number.

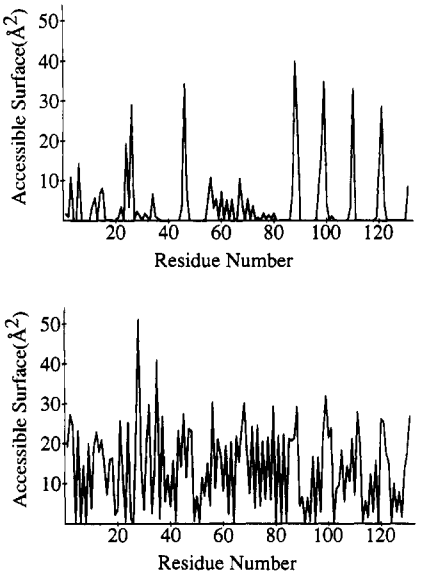

FIGURE 5: Accessibility studies of ALBP atoms. (A, top) Plot of residue number vs averaged accessible surface area (Å^2) for the main-chain atoms in the crystal structure of ALBP. (B, bottom) Accessible surface area (Å^2) for the side-chain atoms averaged for that residue. The calculation is based on the Lee and Richards (1971) algorithm as implemented in XPLOR.

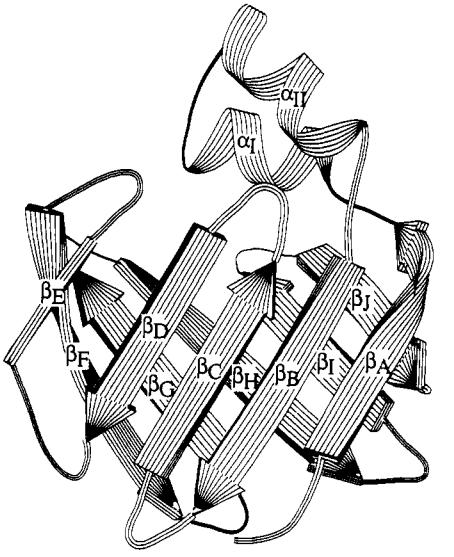

FIGURE 6: Secondary structural components of ALBP. The secondary structural components of crystalline ALBP are drawn in cartoon fashion, arrows for β-structure and coils for α-helices. The two α-helices are near the N-terminal and are labeled αI and αII. β-strands are labeled from βA to βJ.

protein. R78 is involved in a salt bridge with E76, while the positive charges on R106 and R126 are largely shielded by solvent molecules. These internal arginines have relatively low temperature factors and are thought to be important for fatty acid binding (Buelt & Bernlohr, 1990; Sacchettini et al., 1989a; Jones et al., 1988).

Other interesting features of the accessibility relationships is the fact that the two most accessible side chains as seen in Figure 5B are hydrophobic: A28 and M35. In IFABP, A28 is homologous with a valine, V26. M35 is homologous to a histidine in IFABP and a leucine in P2. The location of M35 and A28 in ALBP means that there is a small hydrophobic patch on the solvent side of the protein. Later discussion will note that this region is located on two interconnected helices which may form a portal to the interior cavity and that the first helix contains Y19, a side chain which is phosphorylated by the insulin receptor kinase.

Secondary Structure of ALBP. The secondary structural elements were assigned to ALBP on the basis of the Kabsch and Sander (1983) algorithm. Figure 6 is a ribbon diagram showing the idealized elements of secondary structure. The framework of the molecule is antiparallel β-sheets, formed by 10 β-strands wrapped together to form a flattened cone. The up-down β-barrel has also been observed in several other proteins and is considered a structural motif by Richardson and Richardson (1989).

The strands A–J comprise residues 6–14, 39–45, 48–54, 60–64, 70–73, 79–87, 90–97, 100–109, 112–119, and 122–130. Tight turns are found in positions 46–47, 88–89, 98–99, 110–111, and 120–121. A bulged turn occurs at 74–78. The loop between strands A and B consists of a helix–turn–helix motif. The loops between C and D, and D and E, do not fit a category. Main-chain hydrogen bonds occur regularly between neighboring strands except strands D and E. The spaces between the two strands are largely filled by side chains. As will be discussed in a later section, the gap between βD and βE might result from a favored protein folding pathway.

A right-handed twist is observed as a characteristic feature for the β-sheets. Additionally, strand βA has a sharp kink in its middle caused by β-bulge at position 11 (Figure 7A). This

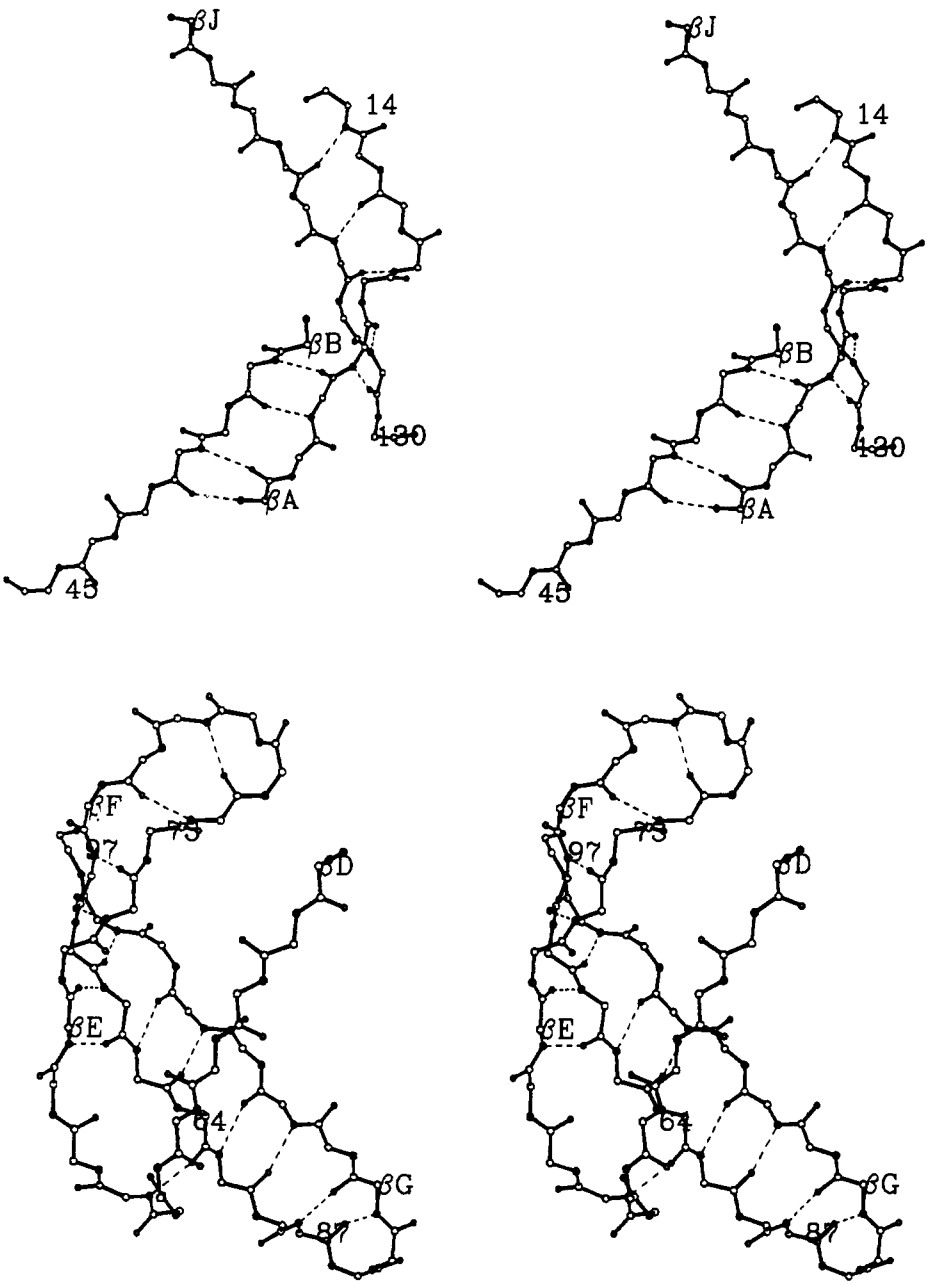

FIGURE 7: β-Strand bending leading to the formation of continuous β-sheets in ALBP. The stereodiagram shows the polypeptide backbone with interstrand hydrogen bonds indicated by dash lines. The name of each β-strand is placed near the first residue, and the residue number of the last residue in each strand in also shown. Panel A (top) contains βA, βB, and βJ. Panel B (bottom) contains βD to βG.

introduces a 70° change in direction. The net effect is that half of the strand runs nearly perpendicular to the other half and permits βA to hydrogen bond to both βB and βJ at the C-terminal. At the leftmost end of the molecules as shown in Figure 6, βF is also severely distorted due to the presence of a β-bulge at position 80–81. As can be seen in Figure 7B, the bend in βF (residues 76–88) allows hydrogen bonding between both βE and βG, forming the connection between the two sheets. The two β-strand bends in βA and βF are the key factors in the formation of the nearly orthogonal β-sheets characteristic of the clam-like structure.

The only region in ALBP that does not have β-structure is found between residues 15 and 35. Here the molecule contains two interconnected α-helices (αI and αII). αI contains eight residues, F16–V23, and terminates with a Gly–Val–Gly turn. The helix labeled αII consists of residues 27–35; it runs antiparallel to αI at an angle of about 30°. The helix–turn–helix motif effectively hinders the solvent and perhaps ligand access to the protein core. These secondary structural elements are held in place by several hydrogen bonds. For instance, a water-mediated hydrogen bond occurs between the indole nitrogen of W97 and the carbonyl oxygen of V23. This interaction apparently stabilizes the helices in the position close to the sheet.

The Ligand-Binding Cavity. The striking similarity between the overall chain fold of ALBP, P2, and IFABP, coupled with the knowledge that ALBP has a hydrophobic binding site with a high affinity constant for oleic acid (Buelt & Bernlohr, 1990), has prompted the obvious suggestion that the three molecules have similar binding sites. Indeed, there is a large empty volume in the interior of the apo-ALBP molecule as is shown in Figure 8A. The volume of this cavity is calculated to be 609 Å^3 (Levitt & Banaszak, 1992). One molecule of oleic acid, estimated to be ~285 Å^3, could be accommodated into the cavity without any difficulty, and work is in progress on preparing crystals of ALBP with bound oleic acid.

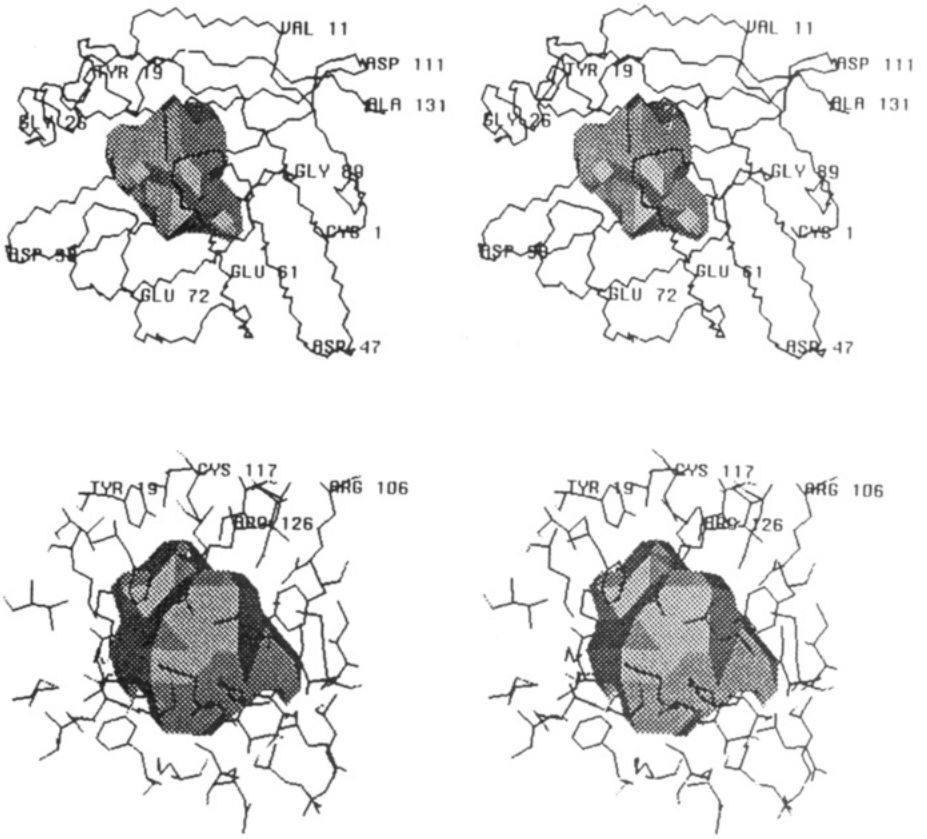

FIGURE 8: Ligand-binding cavity in ALBP. The stereodiagram shown in panel A (top) shows the main-chain atoms of ALBP with the solid surface depicting the cavity region as calculated by the method of Levitt and Banaszak (1992). The α-carbon atoms of a few residues are labeled to permit identification of the position in the amino acid sequence. The bottom leftmost corner of each label is positioned next to the C_α atom. Panel B (bottom) contains a stereodiagram of the cavity in solid surface and all side chains in contact with the cavity surface. Important residues which are discussed further in the accompanying text are labeled.

All of the side chains in contact with the cavity are shown in the stereodiagram contained in Figure 8B. Only a few are labeled, but a listing of all cavity contacting residues can be found in Table III. The location of the cavity corresponds roughly to the position found in IFABP (Sacchettini et al., 1989a) and that reported for P2 (Jones et al., 1988). Two arginines, R106 and R126, are inside the cavity as can be seen in Figure 8B. Presumably one or both of these arginines will be responsible for interacting with the carboxylate of a bound fatty acid as they are in both IFABP and P2. Y19 is also shown in Figure 8B and is a cavity contacting residue. Interestingly, Y19 is phosphorylated by the insulin receptor kinase despite its very buried nature in the apoprotein. A detailed discussion concerning the possible mechanism of the phosphorylation reaction can be found in the accompanying paper (Buelt et al., 1992). As also can be seen from the data in Figure 5B, Y19 is a solvent inaccessible residue.

Table III lists all the 41 amino acids containing one or more atoms in contact with the internal binding site. Included in this list is one glycine, 18 nonpolar residues, and 22 polar amino acids. Of somewhat more interest is the nature of the side chains that are in contact with ligand-binding site. Analysis of this list shows that of the 41 amino acids indicated in Table III in contact with the ligand cavity, 9 do not have side-chain atoms defining the ligand site; these include K58, N59, K37, E54, E61, M35, V32, I73, and G34. Of the 32 side chains in contact with the cavity, less than half, 15, are hydrophobic. The remaining 17 are polar residues including 4 arginines, 1 aspartic acid, and 1 glutamic acid.

The $(|F_o| - |F_c|)$ map computed with the calculated phases near the completion of the refinement had eight isolated peaks of electron density with values above 3σ which lie on the inside surface of the cavity. These peaks were interpreted to be structured water molecules. In addition, three water molecules are found slightly within and four water molecules are found just outside of the calculated cavity. In all cases, the water positions are able to make hydrogen bonds to one or more atoms belonging to ALBP. In apo-ALBP, the ligand-binding cavity is estimated to be about 600 Å^3. Allowing 12 Å^3 per water molecule, the ligand-binding cavity in crystalline ALBP would contain about 50 water molecules of which 15 appear to be ordered in the 2.5-Å electron density map.

The algorithm used to study the cavity when carried out at intervals of 1 Å indicates a very small opening from the exterior to the ligand-binding cavity. This opening occurs between strands D and E, formed by main-chain atoms of N59, E61, T74, and A75 and side-chain atoms of K58, T60, I62, E72, I73, and A75. Recall that this is the place where normal interstrand hydrogen bonding was replaced by side-chain interactions. Since the opening is less than that which would be required as a portal for fatty acid entry, the functional significance of this small opening remains unknown.

Discussion

Structure Comparison with IFABP. The ALBP structure was compared with the refined structures of apo-IFABP [R = 0.185 for 7–1.96-Å resolution data; the rms deviation of bond length from ideal values was 0.016 Å (Sacchettini et al., 1989b)] and of holo-IFABP [R = 0.178 for 7–2-Å resolution data; the rms deviation of bond length from ideal values was 0.009 Å (Sacchettini et al., 1989a)]. The structures of apo- and holo-IFABP have been compared extensively by Sacchettini et al. (1989b), and generally only minor conformational changes accompany fatty acid binding. In IFABP, the

largest difference occurs at residues 73 and 110. The rms deviations for the main-chain atoms of these two residues are 1.63 and 1.85 Å, respectively. Residue 73 is located at the opening on the molecule surface which was thought to be a potential portal for the fatty acid. The difference at residue 110 is due to the flip of the peptide bond. In holo-IFABP, the turn between βH and βI is a type II turn, while in apo-IFABP, it is a type I turn. Taking into account these two regions in apo-ALBP, the similarity to the apo-form of IFABP is greater.

The cavity volume in IFABP is about 450 Å^3 or two-thirds of that associated with ALBP. Not unexpectedly, the number of amino acids containing atoms in contact with the cavity is slightly less, 36 as opposed to 41 residues in ALBP. Of these 36, 33 have side-chain atoms near the cavity, and, of these 33, 25 are in structurally homologous positions in ALBP. On the basis of this relatively poor sampling, one could argue that these 25 locations are fundamental to defining the hydrophobic ligand-binding site. However, when the identities in these 25 locations are defined for both IFABP and ALBP, there are only six positions. These six positions in ALBP are M20, F57, A75, D76, R106, and R126. R106 and R126 are conserved for interacting with the carboxylate of fatty acids. In cellular retinol-binding protein, only R126 remains an identity (Demmer et al., 1986). The roles the other residues (M20, F57, A75, and D76) play is unclear. Clustered together on one side of the molecule, they may be near the ligand-binding portal.

One other comparison between ALBP and IFABP can be made. Of the 33 amino acids in contact with the cavity in IFABP, 20 are hydrophobic and 13 are polar. The cavity in ALBP, as noted above, is somewhat more "polar", having 15 hydrophobic and 17 internal polar residues.

The conclusions drawn from the above comparisons are the following: (1) Even though the basic clam structure is very similar for various members of this family, the nature of internal side-chain packing can cause variations in the volume of the cavity binding site. (2) The close similarity in main-chain structure produces the obvious homology in conformational positions which are in contact with the internal binding site. However, unlike most enzyme systems, these cavity side chains are not highly conserved! (3) Although the cavity was designed for hydrophobic ligands, it contains a large proportion of polar side chains. This accounts for the fact that the apo-structures do not appear to undergo major conformational changes upon binding large ligands. Rather the cavity region in the absence of hydrophobic ligand, contains both bound and disordered water molecules. The bound waters of course interact with the internal polar atoms. (4) In ALBP and IFABP, the internal cavity does not appear to have any large obvious portal to the external milieu. One is forced to conclude that binding is accompanied by transient conformational changes permitting the hydrophobic ligand to enter/exit the binding site.

Folding Mechanism. The large separation and lack of hydrogen bonding between the strands βD and βE remains a puzzle. It is present in ALBP as well in the earlier crystal structures of IFABP and P2 (Sacchettini et al., 1989a; Jones et al., 1988). One hypothesis that could be used to explain the lack of hydrogen-bonded sheet formation between these two elements is that the gap is the result of a favored folding pathway; that is, intermediates consist of more than one β-sheet which then fold together to form the clam arrangement. Using pseudo-torsional angles about the lines connecting the α-carbon atoms (in a C_α model), we have attempted to "unfold" the model focusing especially on rotations around the conserved glycine residues: G6, G24, G46, G67, G89, G110, and G121. Clearly such hypothetical adjustments have many degrees of freedom, and we are not sure that all possibilities were tried. In no case was it possible to unfold the native conformation into a single β-sheet. However, torsional rotations around the pseudo-G67 angle produce two segments of β-sheet. One is four-membered and consists of βA, βB, βC, and βD. The other would contain six members including βE through βJ. βD and βE become one continuous albeit bent β-strand which connects the two sheet segments. The suggestion, therefore, is that the clam-like structure forms by the folding of a more open structure containing two β-sheet segments.

A comparison of the newly obtained crystal structure of ALBP with that of IFABP makes it possible to generalize further about the conformational properties of hydrophobic ligand-binding proteins. The volumes of the ligand-binding cavities are related to about 25 structurally related amino acids. However, this number will be subject to change when comparisons include P2 and other members of this family. As mentioned earlier, ALBP is phosphorylated by insulin receptor kinase. The structure presented here is the first direct visualization of an insulin receptor kinase substrate. An analysis of the crystal structure in terms of the phosphorylation site is given in the accompanying paper. The crystal structure of ALBP with bound oleic acid is presently under study.

ACKNOWLEDGMENTS

We are indebted to Professor David Levitt for the use of his program "POCKET". This has been used to determine the ligand-binding cavity in ALBP and IFABP and to analyze the nature of the side chains involved in this part of the structure. We are also indebted to Professor Axel Brunger of Yale University for the use of XPLOR, which was used to both solve and refine the crystal structure of ALBP, and to Professor T. Alwyn Jones and his colleagues for permission to use their refined P2 coordinates before publication. We are grateful to Ed Hoeffner and David Schuller for their continued effort in maintaining the Dietrich Laboratories computing facilities. Finally, we thank the Minnesota Supercomputer Institute for supporting this project by providing supercomputer time.

REFERENCES

Bernier, M., Laird, D. M., & Lane, M. D. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 1844–1848.

Bernlohr, D. A., Angus, C. W., Lane, M. D., Bolanowski, M. A., & Kelly, T. J., Jr. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 5468–5472.

Bernlohr, D. A., Doering, T. L., Kelly, T. J., Jr., & Lane, M. D. (1985) *Biochem. Biophys. Res. Commun. 132*, 850–855.

Brunger, A. T. (1990) *Acta Crystallogr. Sect. A. 46*, 46–57.

Brunger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 223*, 458–460.

Brunger, A. T., Karplus, M., & Petsko, G. A. (1989) *Acta Crystallogr. Sect. A 45*, 50–61.

Brunger, A. T., Krukowski, A., & Erickson, J. (1990) *Acta Crystallogr. Sect. A 46*, 585–593.

Buelt, M. K., & Bernlohr, D. A. (1990) *Biochemistry 29*, 7408–7413.

Buelt, M. K., Shekels, L. L., Jarvis, B. W., & Bernlohr, D. A. (1991) *J. Biol. Chem. 266*, 12266–12271.

Buelt, M. K., Xu, Z., Banaszak, L. J., & Bernlohr, D. A. (1992) *Biochemistry* (following paper in this issue).

Chinander, L. L., & Bernlohr, D. A. (1989) *J. Biol. Chem. 264*, 19564–19572.

Demmer, L. A., Birkenmeier, E. H., Sweetser, D. A., Levin, M. S., Zollman, S., Sparkes, R. S., Mohandas, T., Lusis,

A. J., & Gordon, J. I. (1986) *J. Biol. Chem. 262*, 2458–2467.
Flatt, J. P. (1970) in *Adipose Tissue: Regulation and Metabolic Functions* (Jeanrenaud, B., & Hepp, D., Eds.) pp 93–101, Georg Thieme Verlag, Germany.
Fujinaga, M., & Read, R. J. (1987) *J. Appl. Crystallogr. 20*, 517–521.
Holden, H. M., Rypniewski, W. R., Law, J. H., & Rayment, I. (1987) *EMBO J. 6*, 1565–1570.
Hresko, R. C., Bernier, M., Hoffman, R. D., Flores-Riveros, J. R., Liao, K., Laird, D. M., & Lane, M. D. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 8835–8839.
Huber, R., Schneider, M., Epp, O., Mayr, I., Messerschmidt, A., Pflugrath, J., & Kayser, H. (1987a) *J. Mol. Biol. 195*, 423–434.
Huber, R., Schneider, M., Mayr, I., Muller, R., Deutzmann, R., Suter, F., Zuber, H., Falk, H., & Kayser, H. (1987b) *J. Mol. Biol. 198*, 499–515.
Hunt, C. R., Ro, J. H.-S., Dobson, D. E., Min, H. Y., & Spiegelman, B. M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 3786–3790.
Jones, T. A. (1978) *J. Appl. Crystallogr. 11*, 268–272.
Jones, T. A., Bergfors, T., Sedzik, J., & Unge, T. (1988) *EMBO J. 7*, 1597–1604.
Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577–2637.
Lattman, E. E. (1985) *Methods Enzymol. 115*, 55–77.
Lee, B., & Richards, F. M. (1971) *J. Mol. Biol. 55*, 379–400.
Levitt, D. G., & Banaszak, L. J. (1992) *J. Comput. Graph.* (in press).
Lowe, J. B., Sacchettini, J. C., Laposata, M., McQuillan, J. J., & Gordon, J. I. (1987) *J. Biol. Chem. 262*, 5931–5937.
Luzzati, V. (1952) *Acta Crystallogr. 5*, 808–810.
Matarese, V., & Bernlohr, D. A. (1988) *J. Biol. Chem. 263*, 14544–14551.
Matarese, V., Stone, R. L., Waggoner, D. W., & Bernlohr, D. A. (1990) *Prog. Lipid Res. 28*, 245–272.
Monaco, H. L., Zanotti, G., Spadon, P., Bolognesi, M., Sawyer, L., & Eliopoulos, E. E. (1987) *J. Mol. Biol. 197*, 695–706.
Newcomer, M. E., Jones, T. A., Aqvist, J., Sundelin, J., Eriksson, U., Rask, L., & Peterson, P. A. (1984) *EMBO J. 3*, 1451–1454.
Papiz, M. Z., Sawyer, L., Eliopoulos, E. E., North, A. C. T., Findley, J. B. C., Sivaprasadaro, R., Jones, T. A., Newcomer, M. E., & Kraulis, P. J. (1986) *Nature 324*, 383–385.
Rao, S. N., Jih, J.-H., & Hartsuck, J. A. (1980) *Acta Crystallogr. Sect. A 36*, 878–884.
Richardson, J. S., & Richardson, D. C. (1989) in *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G. D., Ed.) pp 1–98, Plenum Publishing Corp., New York.
Sacchettini, J. C., Gordon, J. I., & Banaszak, L. J. (1989a) *J. Mol. Biol. 208*, 327–339.
Sacchettini, J. C., Gordon, J. I., & Banaszak, L. J. (1989b) *Proc. Natl. Acad. Sci. U.S.A. 86*, 7736–7740.
Sawyer, L., Papiz, M. Z., North, A. C. T., & Eliopoulos, E. E. (1985) *Biochem. Soc. Trans. 13*, 265–266.
Sweetser, D. A., Heuckeroth, R. O., & Gordon, J. I. (1987) *Annu. Rev. Nutr. 7*, 337–359.
Uyemura, K., Yoshimura, K., Suzuki, M., & Kitamura, K. (1984) *Neurochem. Res. 9*, 1509–1514.
Waggoner, D. W., & Bernlohr, D. A. (1990) *J. Biol. Chem. 265*, 11417–11420.
Xu, Z., Buelt, M. K., Banaszak, L. J., & Bernlohr, D. A. (1991) *J. Biol. Chem. 266*, 14367–14370.