Crystal Structure of Soybean Lipoxygenase L-1 at 1.4 Å Resolution[†,‡]

Wladek Minor,[§,||] Janusz Steczko,[⊥,#] Boguslaw Stec,[○] Zbyszek Otwinowski,[△] Jeffrey T. Bolin,[§] Rick Walter,[▽,∞] and Bernard Axelrod*[,⊥]

Departments of Biological Sciences and Biochemistry, Purdue University, West Lafayette, Indiana 47907, Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, 30239 Krakow, Poland, Department of Chemistry, Boston College, Chestnut Hill, Massachusetts 02167, Department of Chemistry, University of Texas Southwest Medical Center, Dallas, Texas 75235, and Department of Biochemistry, Cornell University, Ithaca, New York 14853

Received March 8, 1996; Revised Manuscript Received June 17, 1996[⊗]

ABSTRACT: Lipoxygenases, which are widely distributed among plant and animal species, are Fe-containing dioxygenases that act on lipids containing (*Z*,*Z*)-pentadiene moieties in the synthesis of compounds with a variety of functions. Utilizing an improved strategy of data collection, low temperature, and synchrotron radiation of short wavelength, the structure of ferrous soybean lipoxygenase L-1, a single chain protein of 839 amino acid residues, has been determined by X-ray crystallography to a resolution of 1.4 Å. The *R*-factor for the refined model is 19.7%. General features of the protein structure were found to be consistent with the results of prior crystallographic studies at lower (2.6 Å) resolution. In contrast to the prior studies, the binding of a water molecule to the active site Fe was established. The octahedral coordination sphere of the Fe also includes the side chains of His^{499}, His^{504}, His^{690}, and Asn^{694} as well as the terminal carboxylate of Ile^{839}, which binds as a monodentate ligand. Asn^{694} is involved in a number of labile polar interactions with other protein groups, including an amide−aromatic hydrogen bond, and appears to be a weak ligand. Several possible access routes for dioxygen and fatty acids to the internal active site and substrate binding cavity are described. The protein structure restricts access to the Fe site such that the formation of an organo-Fe intermediate seems improbable. Structural restrictions pertinent to other proposed reaction intermediates, such as planar pentadienyl and nonplanar allyl radicals, are also discussed.

Lipoxygenases are nonheme, nonsulfur iron dioxygenases that act on lipid substrates containing one or more (*Z*,*Z*)-1,4-pentadiene moieties. The primary reaction products are hydroperoxides of conjugated (*E*,*Z*)-dienes. Common polyunsaturated fatty acids, such as linoleic, linolenic, and arachidonic acids, are the natural substrates for these enzymes, which are widely distributed among plants and animals. Although plant lipoxygenases contain about 25% more amino acid residues than the animal enzymes (ca. 850 vs 650 residues), the two groups have amino acid sequences that are highly related in certain regions, and the primary reactions they catalyze are essentially the same.

In plants, the immediate products are apparently involved in defense mechanisms against pathogens and may be precursors of metabolic regulators (Siedow, 1991; Gardner, 1991). It has also been suggested that lipoxygenases may serve as seed storage proteins (Kato et al., 1993; Tranbarger et al., 1991). In animals, where the predominant substrate is arachidonic acid, the products are precursors of highly potent physiological effectors involved, *inter alia*, in vasoconstriction, blood clotting, bronchoconstriction, granulocyte chemotaxis, and inflammation (Ford-Hutchinson et al., 1994; Yamamoto, 1992). A possible role of mammalian 15-lipoxygenase[1] in the formation of atherogenic plaques has also been considered (Sigal et al., 1994).

All active lipoxygenases thus far tested are monomeric and contain one nonheme, nonsulfur iron atom per molecule. The metal ion, whose properties and environment have been most intensely examined by physical methods for the soybean seed isozyme L-1, is generally believed to be in an octahedral ligand environment. However, there remains some uncertainty concerning the coordination number and the specific geometry of the site (Scarrow et al., 1994; and below).

Under various circumstances, the Fe atom is converted from a high-spin Fe(II) resting state (as isolated) to a

[†] This work was supported in part by Grant MCB 93045465 from the National Science Foundation to B.A. and W.M. and by a grant from the Lucille P. Markey Charitable Trust to Michael G. Rossmann for the development of structural studies at Purdue. J.T.B. was supported in part by Grant GM-52381 from the National Institutes of Health. J.S. was supported in part by an International Research Scholar's award from the Howard Hughes Medical Institute, HHMI 75195-543901. This is Journal Paper 14987 from the Purdue University Agricultural Experiment Station.

[‡] Atomic coordinates have been deposited with the Brookhaven Protein Data Bank under the entry code 1YGE.

* Author to whom correspondence and requests for reprints should be addressed.

[§] Department of Biological Sciences, Purdue University.

[||] Current address: Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, VA 22908.

[⊥] Department of Biochemistry, Purdue University.

[#] Polish Academy of Sciences.

[○] Boston College.

[△] University of Texas Southwest Medical Center.

[▽] Cornell University.

[∞] Current address: Department of Biological Sciences, Purdue University, West Lafayette, IN 47907.

[⊗] Abstract published in *Advance ACS Abstracts,* August 1, 1996.

[1] Abbreviations: L-1, L-2, and L-3, soybean lipoxygenase isozymes 1, 2, and 3; 5-Lox (5-lipoxygenase), 12-Lox (12-lipoxygenase), and 15-Lox (15-lipoxygenase), lipoxygenases that incorporate oxygen at position 5, 12, or 15 of arachidonic acid and related substrates; PDB, Brookhaven Protein Data Bank (Bernstein et al., 1977); PEG, poly(ethylene glycol) (polyoxyethylene); MIR, multiple isomorphous replacement; CCD, charge-coupled device; EXAFS, extended X-ray absorption fine structure; XANES, X-ray absorption near edge structure; CD, circular dichroism; MCD, magnetic circular dichroism; RBP, retinol binding protein; ALBP, adipocyte lipid binding protein.

S0006-2960(96)00576-4 CCC: $12.00 © 1996 American Chemical Society

Table 1: Summary of X-ray Diffraction Measurements

crystal	detector[a]	source[b]	temp (K)	wavelength (Å)	resolution (Å)	R_{sym} (%)	I/σ, last shell
L-1 osmium	SDMW	RRA	277	1.542	3.0	3.9	1.5
L-1 mercury	Raxis-IIc	RRA	293	1.542	3.0	4.9	1.3
L-1 native	Raxis-IIc	RRA	293	1.542	2.6	6.6	1.4
L-1 native	DIP2000	MRA	293	1.542	2.1	6.7	1.7
L-1 native	CCD	A-1	100	0.908	1.4	3.7	4.5

[a] SDMW, multiwire area detector from San Diego Multiwire; Raxis-IIc, Rigaku/Fuji phosphorimaging plate system from Molecular Structures Corp.; DIP 2000, MAC Science/Fuji phosphorimaging plate system from MAC Science; CCD, detector based on charged-coupled device, as described by Tate et al. (1995). [b] RRA, Rigaku RU-200 rotating anode X-ray generator; MRA, MAC Science rotating anode X-ray generator; A-1, beam line A-1 at Cornell High Energy Synchrotron Source.

catalytically active, high-spin Fe(III) state (Slappendel et al., 1982; Cheesbrough & Axelrod, 1983). For example, it was observed some time ago that L-1 is activated by exposure to the product of the reaction, e.g., 13-hydroperoxy-9-*cis*,-11-*trans*-octadecadienoic acid (Haining & Axelrod, 1958). Similar behavior was subsequently noted for all other lipoxygenases tested. Activation induces a hyperchromic shift in the visible spectrum of the L-1 isozyme (Pistorius & Axelrod, 1974). Moreover, the resting state of the L-1 isozyme is EPR silent at liquid helium temperature, whereas the activated enzyme exhibits a signal at $g \approx 6$ (de Groot et al., 1975; Pistorius et al., 1976).

When the primary structures of the soybean L-1 and L-2 isozymes were determined, a region rich in histidine residues was observed (Shibata et al., 1987, 1988). Further investigation showed that His^{499}, His^{504}, and His^{690} (L-1 numbers) are essential for binding of iron (Steczko et al., 1992; Steczko & Axelrod, 1992). Subsequent crystallographic studies at 2.6 Å resolution by Boyington et al. (1993) and by us (Minor et al., 1993) demonstrated that the side chains of these residues are part of the coordination sphere of the iron. Both laboratories also identified the C-terminal carboxylate group of Ile^{839} as a fourth ligand. In our earlier work, we identified Oδ1 of Asn^{694} as a probable ligand, giving a total of five ligands occupying positions consistent with octahedral coordination; binding of water or hydroxide at the sixth position could not be excluded or confirmed. In contrast, Boyington and co-workers inferred from their model that the iron exists as a four-coordinate center with the terminal carboxylate and three histidine ligands in highly distorted octahedral geometry; tetrahedral geometry with one highly distorted ligand was considered a reasonable alternative. Asn^{694} was excluded as a ligand, apparently on the basis of the Fe−ligand distance.

The identification of the ligands is consistent with a wealth of amino acid sequence data. In each of the ~35 sequences of plant and animal lipoxygenases that have been reported to date, the three histidine ligands are conserved. Moreover, all of the sequences terminate in Ile with the exception of a single 5-Lox from rat basophilic leukemia cells (Balcerak et al., 1988), which appears to be truncated at the carboxyl end by two residues such that it ends with a valine residue. By modifying two murine lipoxygenases, Chen et al. (1994) have shown that replacement of the C-terminal isoleucine with valine results in an enzyme having substantial activity. Other replacements result in inactivation.

Here we report the crystal structure of the Fe(II) (resting) form of L-1 lipoxygenase at the resolution of 1.4 Å and low temperature of 100K. This very high resolution is noteworthy for such a large protein (839 amino acids) and allows a more complete and precise description of the protein as well as the iron center. The high-resolution structure is consistent with five certain ligands in octahedral geometry, including the three conserved histidine side chains, the terminal carboxylate, and a water molecule. The Oδ1 atom of Asn^{694} is in the proper direction to serve as a sixth ligand but is sufficiently distant that it appears to interact weakly with the Fe.

EXPERIMENTAL PROCEDURES

Crystals. Detailed procedures for protein purification, crystallization, preparation of Hg and Os derivatives, and data collection at room temperature have been previously described (Steczko et al., 1990; Minor et al., 1993). Crystals were obtained at 21 °C by equilibrating sitting drops of 0.1 M acetate buffer at pH = 5.6 containing 4% (w/v) PEG-3400 and 4−5 mg/mL enzyme against 0.2 M buffer containing 8% PEG-3400. For cryodiffraction experiments, the crystals were transferred to a similarly buffered solution containing 28% (v/v) ethylene glycol and 14% (w/v) PEG-3350 through three solutions with intermediate concentrations of these reagents. Incubation times were approximately 2 min at each stage. After a 1−2 min incubation in the final solution, the crystals were mounted in a nylon loop and placed in a nitrogen gas stream at 100 K to achieve flash freezing.

Lipoxygenase L-1 crystallizes in the monoclinic space group $P2_1$ with one protein molecule (94 kDa) in the asymmetric unit; the solvent content is approximately 48% by volume. At room temperature the unit cell parameters are $a = 95.6$ Å, $b = 94.3$ Å, $c = 50.3$ Å, and $\beta = 91.3°$. The unit cell of the frozen crystals (100 K) is smaller in volume by 2%: $a = 94.9$ Å, $b = 94.0$ Å, $c = 49.9$ Å, and $\beta = 90.1°$. Enzyme recovered from dissolved crystals has normal catalytic activity.

Diffraction Measurements at Lower Resolution and Phasing. For diffraction experiments at room temperature and lower resolution, Cu Kα radiation from rotating anode X-ray generators was used in conjunction with three different systems: an Raxis-II imaging plate system (Molecular Structures Corp.), a DIP-2000 imaging plate system (MAC Science Co.), and a dual multiwire area detector system (San Diego Multiwire Systems).

The phase problem was initially solved to 2.6 Å resolution with data from the Raxis-II and multiwire detector systems (Table 1). Osmium and mercury derivatives (Minor et al., 1993) provided the basis for phase estimation at 3.0 Å resolution by the multiple isomorphous replacement (MIR) method as implemented in MLPHARE (Otwinowski, 1991). Anomalous scattering information for both derivatives was utilized in this process. Phase improvement with extension

Table 2: Resolution-Dependent Statistics for the 1.4 Å Native L-1 Diffraction Data

resolution shell (Å)[a]	av intensity	av error	fraction of reflections with given no. of observations (%)							linear *R*-factor[b]	square *R*-factor[c]
			0	1	2	3	4	≥5	≥1		
40.00−3.80	222 744	5256	3.0	6.8	13.4	8.7	20.5	37.5	97.0	0.026	0.029
3.80−3.02	144 076	3564	1.0	7.6	15.3	12.4	25.4	38.2	99.0	0.031	0.031
3.02−2.63	57 378	1478	1.0	4.1	10.3	12.6	15.5	56.5	99.0	0.050	0.063
2.63−2.39	36 580	863	1.6	2.5	6.2	8.6	13.3	67.9	98.4	0.057	0.062
2.39−2.22	28 033	699	1.8	1.7	4.8	6.8	13.9	70.9	98.2	0.066	0.072
2.22−2.09	21 958	564	2.1	1.5	4.8	8.7	16.1	66.8	97.9	0.072	0.076
2.09−1.99	15 229	547	3.3	4.6	16.7	36.1	39.4	0.0	96.7	0.073	0.079
1.99−1.90	10 530	482	4.0	3.9	17.7	34.5	39.7	0.0	96.0	0.083	0.087
1.76−1.71	4 214	223	4.4	4.5	20.5	33.0	37.7	0.0	95.6	0.105	0.103
1.62−1.58	2 261	188	4.5	4.2	21.2	34.7	35.4	0.0	95.5	0.155	0.146
1.51−1.48	1 447	184	5.4	3.7	22.9	36.2	31.8	0.0	94.6	0.225	0.203
1.42−1.40	890	198	8.4	7.4	26.5	34.4	23.3	0.0	91.6	0.343	0.345
all data	29 259	819	4.0	4.2	17.4	27.5	29.4	17.6	96.0	0.044	0.036

[a] The shells are a sample from a set of 20 shells containing approximately equal numbers of possible reflections. [b] Linear *R*-factor $= \sum_h\sum_i|(I_{hi} - \langle I_h\rangle)|/\sum_h\sum_i I_{hi}$. [c] Square *R*-factor $= \sum_h\sum_i(I_{hi} - \langle I_h\rangle)^2/\sum_h\sum_i I_{hi}^2$.

to 2.6 Å was performed by the use of a solvent flattening program that incorporated a maximum entropy weighting scheme to minimize the difference between calculated (density-modified) and observed amplitudes (Otwinowski, unpublished). This procedure resulted in phases of sufficient quality to allow the unequivocal identification of the Fe atom in an anomalous difference map (Minor et al., 1993). For the large COOH-terminal domain, the standard electron density map was very clear, and both main chain and side chain density features were readily interpretable. In contrast, the density for the NH_2-terminal domain was comparatively poor and difficult to interpret. Model building by the use of the programs O (Jones et al., 1991) and CHAIN (Sack, 1988) was initiated at this stage. However, the resolution of the data and maps was significantly extended before the initial model building was completed.

Extension of Resolution and Preliminary Refinement. The first approach to extend the resolution was to measure diffraction data using the prototype of the DIP-2000 imaging plate scanner. This instrument has an image scanning time approximately one-third that of the Raxis-II system, allowing the use of much smaller oscillation angles, 0.6° vs 1.5°. This appears to be a major factor in a substantial improvement in the extent and quality of the higher resolution data. A nearly complete data set (89%) to 2.15 Å resolution was collected from one crystal. With this new data set, phase extension was repeated *de novo*; i.e., the MIR phase set was extended to 2.15 Å resolution by the use of the same maximum entropy-based solvent flattening program. In this case, 2000 cycles of phase improvement/extension were carried out with automatic assignment of a new molecular envelope every 40 cycles.

The resulting electron density map, although only slightly better in terms of overall backbone connectivity, provided significantly improved detail in many previously ambiguous regions. A full model was constructed, which specified an *R*-factor of 40%. Nevertheless, the density and model remained less reliable in the NH_2-terminal domain, perhaps reflecting higher conformational mobility of and within this domain, as described below.

The consistently less reliable density for the NH_2-terminal domain obtained from different sets of data measured at room temperature prompted us to seek further improvement by performing diffraction measurements at cryogenic temperatures using synchrotron radiation. In fact, a dramatic extension of the data to 1.4 Å resolution was achieved by flash freezing the crystals and collecting data at 100 K, using synchrotron radiation and a new CCD detector (Tate et al., 1995), and using the enhanced HKL software package (Otwinowski, 1993; Minor, 1993) for analyzing the diffraction images. This high-resolution data set was measured using X-rays of 0.908 Å wavelength provided by the recently improved A-1 wiggler station at the Cornell High Energy Synchrotron Source. The active area of the CCD detector was 50 × 50 mm, and its effective pixel size was 50 × 50 μm.

In this experiment, all of the data were obtained from one crystal with a nominal mosaicity of 0.28° in a period of roughly 14 h. However, because of the large dynamic range of the intensities detected between 40 and 1.4 Å, the data were measured in three passes. The first pass, with 40 s exposures and 0.6° oscillation angle per exposure, was designed to obtain high quality data in the resolution range 2.1−1.4 Å. The second pass, with 10 s exposures and 1.5° oscillation angle, was intended to maximize the signal to noise ratio for the resolution shell 3.0−1.9 Å. The last pass, with exposure for 5 s and oscillation range 5°, was used to collect data in the resolution shell between 50 and 3.0 Å. The crystal-to-detector distances were 32, 50, and 75 mm, respectively, for the three passes; the center of the detector was always set to a 2θ angle of 0°. By increasing the distance during measurement of the lower resolution data, a reduction in background and improvement in the spatial resolution of reflections from neighboring planes was achieved. This multipass data collection strategy resulted in high data redundancy and completeness (Table 2).

All data sets were processed, scaled, and postrefined using the HKL package (Denzo, Scalepack, and XdisplayF). Partial reflections were added to determine the full intensity of reflections measured on consecutive images. In the resolution range 40−1.4 Å, 96% of the 169 235 theoretically possible reflections were measured, including 92% of the data between 1.42 and 1.40 Å.

Model Improvement and Further Refinement. The model built initially at 2.6 Å and improved at 2.15 Å resolution was refined at the latter resolution by the use of the program X-PLOR (Brünger et al., 1987). All reflections with $F/\sigma > 2$ in the resolution range 10−2.15 Å were used for the refinement. During this stage, each refinement cycle included simulated annealing (Brünger et al., 1990), refinement

of positions via 150 cycles of Powell minimization, refinement of *B*-factors, and map inspection and manual rebuilding. Simulated annealing by molecular dynamics at 1.4 Å resolution usually consisted of a 0.5 ps trajectory at 2000 K followed by a slow-cooling protocol in which the temperature was gradually quenched to 300 K in 50 K steps. After the refinement converged with $R = 25.4\%$, the model was used to obtain phases for the 1.4 Å resolution data set. At this point, the model included 6667 protein atoms, representing all of the possible residues and no water molecules.

Because of a change in unit cell parameters, rigid body refinement (using X-PLOR) was necessary to adapt the model to the 1.4 Å data. This was accomplished in three steps: at first, the entire molecule was treated as a rigid body; next, the two domains were fitted independently; and finally, the individual helices and β-strands were treated independently. During the rigid body refinement of the entire molecule, the high-resolution limit of the data was increased in 0.5 Å steps between 5.5 and 3.5 Å; the low-resolution limit was always 15 Å. The domains were refined first against 15−3.5 Å resolution data. The high-resolution limit was changed to 3.0 Å for all subsequent refinements with the domains or smaller fragments.

Following the rigid body refinements, atomic coordinates and isotropic *B*-values were refined against the 1.4 Å data using X-PLOR. An initial simulated annealing protocol was followed by 150 cycles of positional refinement and refinement of temperature factors. The result was a model with an *R*-factor of 25% at 1.4 Å resolution and very high quality electron density maps. Further rebuilding, the addition of waters, and X-PLOR refinement has yielded the current model, which has an *R*-factor of 19.7% based on data between 10 and 1.4 Å resolution with $F/\sigma \geq 2$. An estimate for *R*-free (Brünger, 1992) was obtained at this stage by application of the simulated annealing refinement protocol described above. Based on a test set including 10% of the reflections, *R*-free was 24.3%; for comparison, the *R*-factor for the reflections in the working set was 19.3%. Simulated annealing protocols (Hodel, 1992) were also used at this stage to prepare omit maps in the vicinity of the Fe and its ligands. The omitted region was defined by two overlapping 8 Å spheres surrounding Asn^{694} and the Fe atom; atoms in a 3 Å thick boundary shell surrounding the omitted region were positionally restrained.

With respect to the current model and the procedures used to date, the refinement is effectively converged. However, further refinement is possible and will be performed so as to incorporate multiple conformations for protein side chains, variable occupancies for solvent species, and, if warranted, anisotropic disorder for the Fe atom and ligands (see below).

A summary of refinement and model statistics is presented in Table 3. The model includes the entire protein structure (6667 atoms) and 1377 water molecules. The latter were assigned to well-defined positive peaks in ($|F_o| - |F_c|$) difference electron density maps if the peaks were within hydrogen-bonding distance of appropriate protein atoms or other protein-bound solvent molecules. All atoms in the model were assigned full occupancy throughout the refinement. The quality of the structure and the progress of the refinement were monitored frequently with the aid of the program PROCHECK (Laskowski et al., 1993). Very tight stereochemical restraints were enforced on the protein structure in the final stages of refinement to facilitate the use of the model in future experiments. However, Fe−ligand bond lengths were not explicitly restrained.

Fitting of Hypothetical Substrate Models to the Active Site. Atomic structures for linoleic acid and arachidonic acid were constructed and used within the O program to manually explore the fit of possible conformations of substrates and proposed intermediates to the active site. Nonbonded contacts between substrate and protein atoms as well as the torsion angles of the substrate were monitored. The following geometric parameters were used for the substrate models: C−C single-bond lengths, 1.53 Å; C=C double-bond lengths, 1.33 Å; C−O bond lengths, 1.29 Å; $CH_3-CH_2-CH_2-$ bond angles, 110°; $-CH_2-CH_2-CH_2-$ bond angles, 112°; $-C=C-C-$ bond angles, 126°; and C−C−O bond angles, 120°.

Visualization and Analysis of Internal Cavities. A variety of programs were used to display and analyze the large internal cavities of L-1, including GRASP (Nicholls et al., 1991), VOIDOO (Kleywegt & Jones, 1994), and tools provided in the QUANTA package (version 4, Biosym/MSI). The final analysis and visualization of molecular surfaces, as defined by Connolly (1983), were performed with QUANTA. Protein C, N, O, and S atoms were assigned radii of 1.90, 1.82, 1.65, and 2.0 Å, respectively; the probe sphere was assigned a radius of 1.4 Å.

RESULTS AND DISCUSSION

Strategy of the Diffraction Measurements. The improvement in resolution obtained by using frozen crystals and intense synchrotron radiation of short wavelength is not often so dramatic as reported here. A contributing factor may be the strategy of data collection used in this case. Three passes with different counting times, detector distances, and oscillation ranges were used to improve data quality in different resolution ranges. This procedure allowed exploitation of the dynamic range of the detector without sacrificing the signal-to-noise ratio for high-resolution data or overloading the low-resolution data. As a result, the data set has a high level of precision: the overall R_{sym} is 3.6%. In addition, the data set has good overall completeness, 96%, and average multiplicity, more than six observations per reflection (1.1×10^6 observations of 1.6×10^5 reflections). In the shell of highest resolution, the completeness is 91% and the average multiplicity is three observations per reflection. Additional statistics as a function of resolution are presented in Table 2.

Quality of Electron Density Maps and Current Model. The final ($2|F_o| - |F_c|$) map is of very high quality throughout most of the structure (Figure 1). Most phenylalanine and tryptophan residues as well as some proline residues have a low density hole through the center of the ring as expected for maps at this resolution. However, four short segments have significantly less interpretable density. The electron density for residues 1−6 is especially weak and fragmented, suggesting that this is a mobile segment with many conformations. For residues 22−30, which form part of an extended loop, the density indicates the course of the polypeptide and the location of the side chains, but it is not fully continuous. Two additional loops, comprising residues 70−73 and 116−120, are represented by continuous but less than definitive density. All four segments are included in the current model without any attempt to model multiple

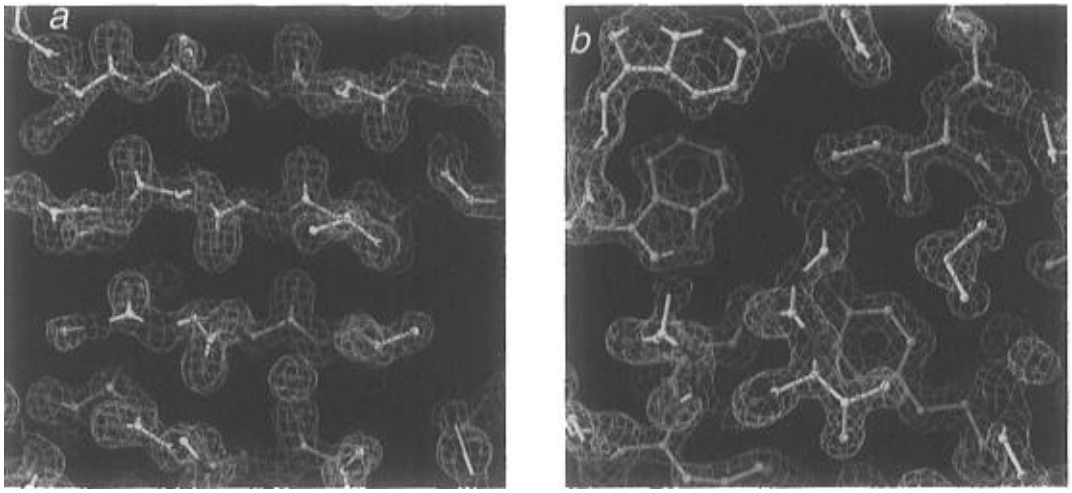

FIGURE 1: Portions of the current 1.4 Å resolution $2|F_o| - |F_c|$ map corresponding to (a) several β strands in the C-domain (residues 404−408, 430−436, 446−452, and 464−469) and (b) a hydrophobic cluster in the C-domain. The map was contoured at 1.3 σ, where σ is the standard deviation of the map.

Table 3: Summary of Refinement Statistics

resolution range	10.0−1.4
standard *R*-factor	19.7%
protein (non-H) atoms	6500
solvent oxygen atoms	1377
geometric deviations[a]	
bond lengths	0.011 Å
bond angles	1.7°
improper dihedral angles	1.5°
average *B*-factors	
main chain of N-domain	34.6 Å^2
main chain of C-domain	16.7 Å^2
all atoms of N-domain	34.0 Å^2
all atoms of C-domain	17.2 Å^2
water oxygen atoms	39.7 Å^2

[a] Root mean square deviations from the standard values defined by Engh & Huber (1991).

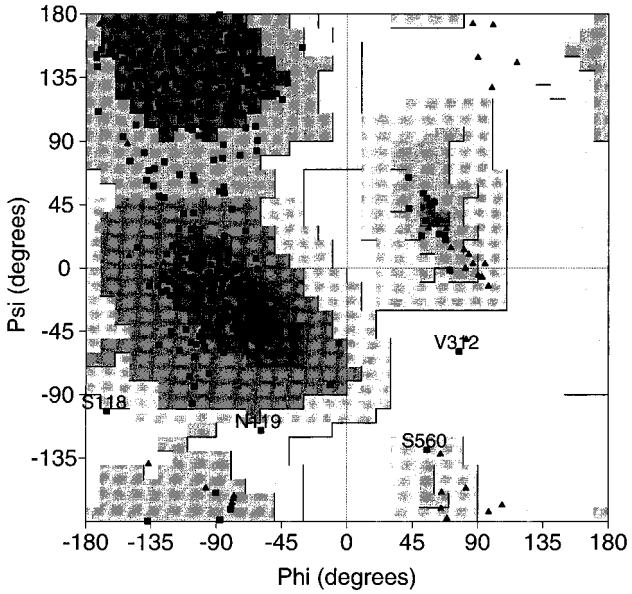

FIGURE 2: Ramachandran plot (Ramachandran & Sasisekharan, 1968) of the refined L-1 structure created by the program PROCHECK (Laskowski et al., 1993). Glycine and non-glycine $\psi-\phi$ values are designated by triangles and squares, respectively. Disallowed, generously allowed, favorable, and most favorable regions are indicated by progressively darker shading.

conformations. The coordinates for residues 1−6 are reliable only as indicators of the volume this segment is likely to occupy.

The model is in excellent agreement with expected stereochemical parameters as illustrated by data presented in Table 3 as well as a Ramachandran plot (Ramachandran & Sasisekharan, 1968) of the main chain dihedral angles ϕ and ψ, shown in Figure 2. One *cis* peptide, Phe^{365}−Pro^{366}, has been identified. Two residues have unusual backbone conformations that are stabilized by hydrogen bonding: Val^{312}, $\phi = +77°$, $\psi = -59°$; and Ser^{560}, $\phi = +55°$, $\psi = -129°$. Overall, the side chain torsional angles are very tightly distributed around the most common rotamers. The density indicates that a number of side chains have multiple conformations, but these have not yet been modeled. A Luzzati plot (Luzzati, 1952) suggests that the average positional error for the model is 0.20 Å (not shown).

General Architecture. The observed fold is fully consistent with the description provided by Boyington et al. (1993) on the basis of their work at 2.6 Å resolution. Thus soybean lipoxygenase L-1 is an ellipsoid of dimensions 90 Å × 65 Å × 60 Å and may be conveniently divided into two spatially isolated domains (Figure 3). The smaller N-domain, which comprises the NH_2-terminal residues 1−146, is a barrel-like β sandwich located at one end of the larger C-domain and relatively loosely associated with it. The C-domain, comprising residues 147−839, has predominately helical secondary structure and includes the catalytic Fe-binding site. A possible alternative division of the structure into five domains is considered below.

The N-Domain. The β sandwich of the N-domain is constructed from two four-stranded antiparallel sheets. At present, the function of the N-domain remains obscure and the question of its dispensability remains open. Mammalian lipoxygenases are substantially smaller proteins and lack approximately 140 NH_2-terminal residues found in L-1, which suggests the N-domain may be dispensable for L-1 and related plant-derived lipoxygenases. However, the generation of a truncated L-1 wherein residues 2−140 were replaced by five residues yielded an inactive enzyme (Steczko et al., 1991). Limited proteolysis of L-1 generates two tightly associated major fragments (Ramachandran et al., 1992) including a 30 kDa product that spans the N-domain plus a portion of the C-domain through residue 274. The other major fragment (60 kDa) begins at

FIGURE 3: Ribbon diagram of the three-dimensional structure of L-1. The smaller N-domain, a single β sandwich, is shown in green. The C-domain is in blue and contains the active site iron, which is represented by the magenta sphere. The two larger β structures in the C-domain are in orange. The yellow fragments near the Fe indicate two segments of the π helix within longer helices; these segments provide three histidine side chains that are Fe ligands. The drawing was prepared with MOLSCRIPT (Kraulis, 1991).

residue 318 and presumably includes the remainder of the domain. The cleaved enzyme was active, and in some circumstances slightly hyperactive, but attempts to resolve, or resolve and reconstitute, the fragments resulted in inactive preparations.

The topological identity and three-dimensional structural similarity of the N-domain and the C-terminal colipase binding domain of mammalian pancreatic lipases were recognized previously (Boyington et al., 1993) and are quite striking despite relatively low sequence similarity. For example, a superposition of the Cα atoms of the N-domain with those of the colipase binding domain of horse pancreatic lipase (Bourne et al., 1994; PDB entry 1HPL) was performed using tools provided in the program O (Jones et al., 1991). An alignment of 81 Cα's was achieved with a root-mean-square difference of 1.6 Å (alignment threshold 3.8 Å, minimum segment length three residues). Residues from all β strands are included in the eight aligned segments, but only 10% of the residues are chemically identical and a majority of these are located within the hydrophobic core of the domain. Although the surface of the N-domain corresponding to the lipase's colipase binding surface is exposed, there is only one identical residue (L-1 Tyr^{102}, lipase Tyr^{403}) among those specifically implicated in colipase binding (Bourne et al., 1994; van Tilbeurgh et al., 1992). Comparison of the structures thus provides little insight concerning the functional significance of their similarity.

The gross structure of the N-domain is also superficially similar to a number of proteins/domains that bind hydrophobic molecules such as retinol and fatty acids (Banaszak et al., 1994). However, the interior of the N-domain is tightly packed with aromatic and aliphatic side chains so that fatty acid binding within the domain is most unlikely. The resemblance is strongest to proteins/domains such as retinol binding protein, RBP (Cowan et al. , 1990), that have an eight-stranded barrel followed by a helix: the N-domain has a similar size, includes eight β strands, and is followed by α helix 1, which spans residues 158 through 171 and is assigned to the C-domain in the division described above. The resemblance does not extend to the topological order of the β strands, which is ABCDEFGH in RBP and GFADCBEH in L-1.

Higher Mobility/Disorder of the N-Domain. The N-domain makes a loose contact with the C-domain through a solvent-filled crevice and is significantly less well ordered, as indicated by comparison of refined temperature factors. Thus the average *B*-value for all main chain atoms in the N-domain is 34 $Å^2$ as opposed to 17 $Å^2$ for the C-domain. This discrepancy is not strongly biased by the N-domain's disordered loops (see above); the average (and median) value for the N-domain is 33 $Å^2$ if 20 residues with main chain average *B* greater than 60 $Å^2$ are omitted from the calculation.

Examination of the spatial distribution of the *B*-values suggests that the N-domain moves as a unit such that residues with lower main chain *B*-values occur near the domain−domain contact and those with larger *B*-values occur farther away. The pattern of *B*-values (see Figure 4A) suggests it is more likely that the barrel-like N-domain rocks across the surface of the C-domain as opposed to swinging away from it via a motion articulated by the short segment connecting the N-domain to the rest of the molecule. This conclusion is supported by correlation coefficients[2] derived from the *B*-values of Cα's and their distances from a reference Cα: when the reference Cα is near the middle of one of the strands closest to the C-domain, the correlation is higher than when it is in the connecting segment. For example, the correlation coefficients were 0.95 or above when Cα's 109, 129, or 138−143 served as the reference point but were 0.90 or less when the reference point was any Cα in the connecting segment. Residues 109, 129, and 141 are central residues within the three β strands closest to the C-domain, and each is at least 14 Å from $Cα^{145}$, the first residue in the connecting segment. The highest correlation, 0.96, was obtained when $Cα^{141}$ served as the reference point; Figure 4B illustrates the correlation for this case.

The C-Domain. As noted above, the C-domain (residues 147−839) is rich in helical structure. It contains 20 helices that include at least six 3_{10}, α, or π hydrogen bonds as well as several shorter, helical segments of one turn or less. Residues 473 through 518 span the longest helix, helix 9,[3] which is 65 Å long, lies at the center of the domain, and is approximately parallel and adjacent to the longest intradomain Cα−Cα chord ($Cα^{157}$−$Cα^{719}$, 86 Å). Among the other helices, the longer ones are more or less parallel or antiparallel to helix 9 so that the core of the domain may be described as a somewhat disorganized multihelix bundle. The active site lies within this bundle. Two antiparallel β sheets, with three and five strands, are also present in the C-domain. Both sheets lie at or near the surface of the domain and are separated from the active site by the bundle of helices.

Several helices contain notable deviations from the features of a typical α helix in terms of geometry and/or hydrogen bonding. For example, as noted previously (Boyington et

[2] Cα atoms in the highly mobile loops of the N-domain were excluded by using only Cα atoms with *B*-values less than 40 $Å^2$. The correlation coefficient was $C = \sum_i(\langle B\rangle - B_i)(\langle d\rangle - d_i)/(\sum_i(\langle B\rangle - B_i)^2(\langle d\rangle - d_i)^2)^{1/2}$, where B_i is the *B*-factor for $Cα_i$, $\langle B\rangle$ is the average *B* for the set, d_i is the distance from $Cα_i$ to the reference Cα atom, and $\langle d\rangle$ is the average distance for the set.

[3] For the sake of comparisons with the structure described by Boyington et al. (1993), the labels assigned to elements of secondary structure correspond to those found in that paper as well as in PDB entry 2SBL.

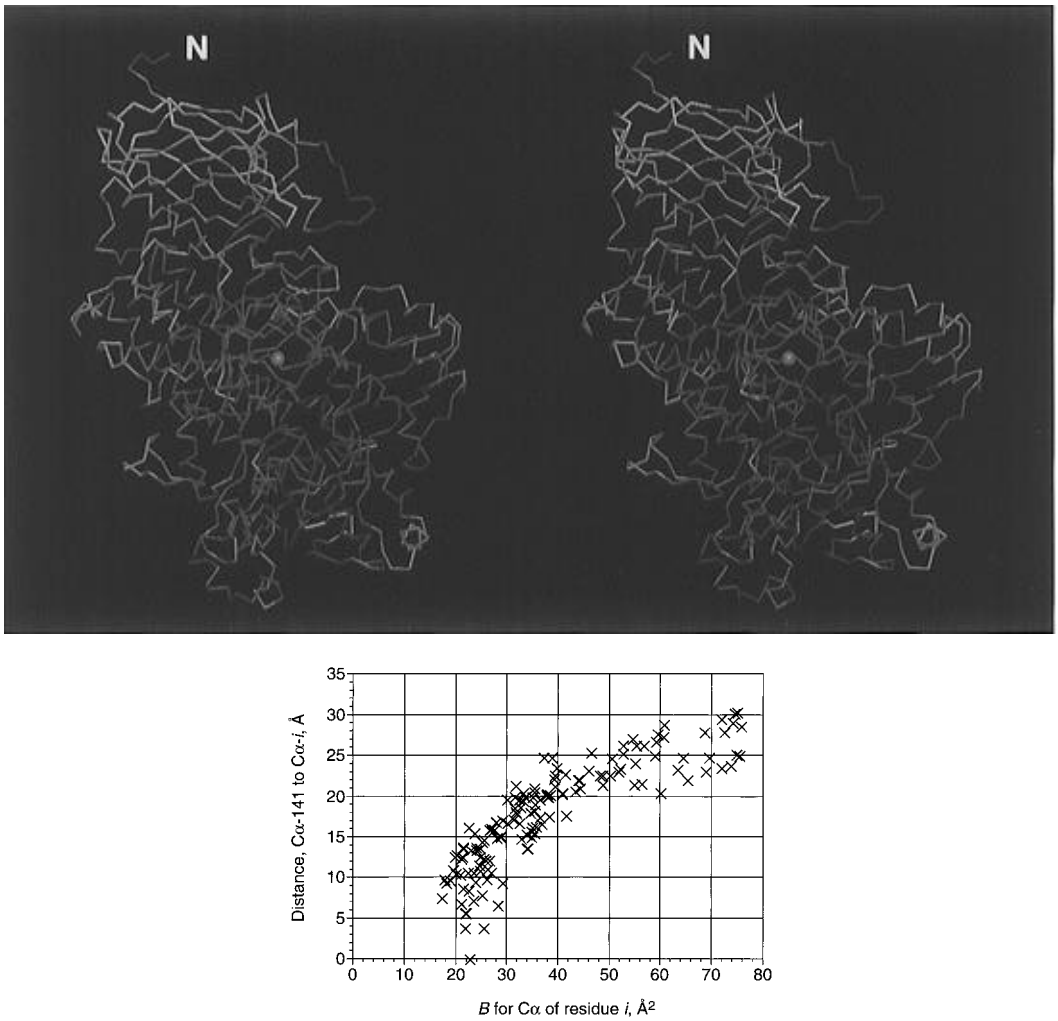

FIGURE 4: (a, top) Stereo Cα trace of L-1 showing the spatial distribution of the temperature factors. The colors are assigned according to the *B*-values of the backbone atoms as follows: colors from violet to light blue correspond to *B*-values between 10 and 35 Å^2; colors from yellow to red indicate *B*-values between 35 and 80 Å^2. The green sphere represents the iron atom. In the N-domain, residues near the domain−domain contact exhibit lower *B*-values, whereas residues having higher values tend to be distant from that region. (b, bottom) Correlation between the *B*-values of N-domain residues and their distance from Cα^{141}.

al., 1993), two very long helices, helix 9 (residues 473−518) and helix 18 (residues 671−702), possess internal π-helical segments that contribute side chains to the primary coordination shell of the active site Fe atom. In the case of helix 9, the π structure includes residues 494 through 504 and ligands His^{499} and His^{504}. For helix 18, residues 684 through 690 form the π structure, and His^{690} is a ligand.

The two π segments possess features in common that may be representative of the structural transitions necessary to accommodate π helices within longer helical segments. One shared feature is an extremely negative (for helices) ϕ torsion angle for the residue immediately prior to the first of the backbone amide groups involved in π-type hydrogen bonding. For helix 9, the amide group of Ser^{498} is the first hydrogen bond donor in the π helix, and the ϕ angle for Met^{497} is −105°. The same ϕ value is found for Ser^{687}, which precedes the first π hydrogen bond donor in helix 18, Ala^{688}. A second common feature is the presence of a residue with an extremely negative ϕ angle and an hydroxylated side chain three residues before the last π hydrogen bond donor: Thr^{503}, $\phi = -108°$, in helix 9; and once again Ser^{687}, $\phi = -105°$, in helix 18. In both cases, the side chain hydroxyl group of residue *i* apparently forms hydrogen bonds with the carbonyl group of residue *i−4* and with the amide group of residue *i+1*, thus intervening between groups that otherwise would be linked by π hydrogen bonds. Moreover, the carbonyl group of residue *i−1* is turned away from the axis of the helix and does not participate in helical hydrogen bonding, whereas the amide group of residue *i* is involved in an α-type hydrogen bond with the carbonyl group of residue *i−4*. These consistent disruptions suggest that the threonine or serine residue stabilizes and perhaps initiates the transition from the π form to a tighter helix.

Helices 2 and 11 provide additional examples of interesting variations in helical structure. Helix 2, residues 254−278, joins 3_{10}, α, and π segments within a single continuous helix. It begins with one turn of 3_{10} helix that merges into two turns of α helix. The helix then widens and bends to accommodate nearly a full turn of π structure and a proline residue before resuming α structure. The helix is capped at its carboxyl-terminal end by a hydrogen bonded turn that includes a single residue in left-handed α conformation (Lys^{277}). Helix 11, which borders the substrate binding site, exhibits an unusual discontinuity. It begins at residue 534 in regular α structure but opens at residues 545 and 546 to incorporate a dramatic interruption in the form of a reverse turn spanning residues 547−551. Remarkably, residues 552 and 553 then align with the first part of the helix, allowing

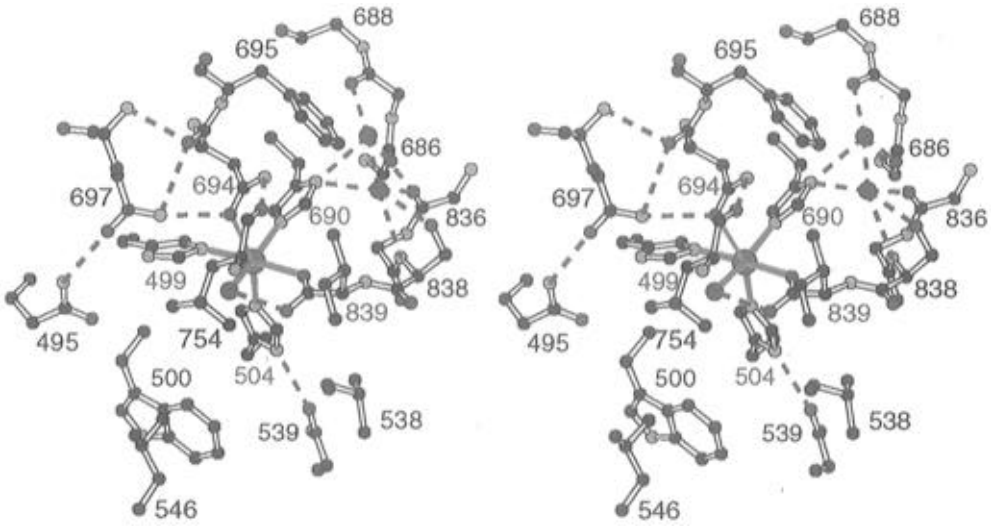

FIGURE 5: Stereoview illustrating the vicinity of the active site Fe atom. The hydrogen bond network (red dashed lines) and other nonbonded contacts in the vicinity of the active site are illustrated. The nearest portion of the probable fatty acid binding pocket is within the space bordered by residues Gln[495], Trp[500], His[504], Ile[538], and Leu[546]. Waters molecules are represented by small red spheres, and the iron is illustrated by a larger magenta sphere. Fe ligand vectors are in green, and the residue numbers of the protein ligands are labeled in red; other residue numbers are in black. Carbon, nitrogen, oxygen, and sulfur protein atoms are in gray, blue, red, and yellow, respectively. The figure was generated by MOLSCRIPT (Kraulis, 1991).

their amide groups to form hydrogen bonds with the carbonyl groups of residues 545 and 546. The polypeptide chain returns to α-helical structure from this point to its termination at residue 555. Thus the helix is effectively extended past the interruption by an additional turn of helix involving residues 551−555.

An Alternative Assignment of Domains. Division of the L-1 structure into two domains is supported by the spatial separation of the N-domain from the remainder of the structure. Nevertheless, alternative assignments of domains are possible and warrant consideration. In one of these, the structure is divided into five domains such that four smaller domains of 90−170 residues associate on the surface of one larger, C-terminal domain. In this case, the first domain spans residues 1−170 so that it includes the N-terminal β sandwich and helix 1 and is thus analogous to the structure of RBP, as described above. The second domain includes residues 171−250, which form a contiguous but loosely organized unit; the hydrogen-bonded secondary structure is limited to turns and short, isolated β ladders. Domain three, residues 251 to ca. 340, is organized into a three-stranded antiparallel β sheet and three well-defined helices (2, 3, and 4). Domain four begins and ends approximately at residues 341 and 470 and comprises the five-stranded β sheet and two associated helices (7 and 8). The fifth and final domain begins approximately at residue 471 near the N-terminus of core helix 9, includes the remainder of the polypeptide chain (368 residues), and holds the catalytic site. It is exclusively α helical and encompasses more than ten helices organized about the two long, core helices (9 and 18). Recognition of this alternative assignment of domains prompts consideration of the possible functional roles the peripheral domains may play. For example, motions of the domains could facilitate binding, transport, and release of substrates and products. Indeed, likely entrances/exits to the internal cavities implicated in substrate binding and passage of the substrate to the binding site (see below) are located at interfaces between the catalytic domain 5 and peripheral domains 3 and 4.

Table 4: Coordination Geometry of the Active Site Fe

	distances (Å)	
defining atoms	this work	2SBL[a]
Fe−His[499] Nε2	2.23	2.27
Fe−His[504] Nε2	2.26	2.22
Fe−His[690] Nε2	2.21	2.25
Fe−Ile[839] O1	2.40	2.16
Fe−Asn[694] Oδ1	3.05	3.16
Fe−Wat[842] O	2.56	
	angles (deg)	
defining atoms	this work	2SBL[a]
His[504] Nε2−Fe−His[499] Nε2	95	92
His[504] Nε2−Fe−His[690] Nε2	106	100
His[504] Nε2−Fe−Ile[839] O1	92	107
His[504] Nε2−Fe−Wat[842] O	97	
His[499] Nε2−Fe−His[690] Nε2	102	105
His[499] Nε2−Fe−Wat[842] O	86	
Ile[839] O1−Fe−His[690] Nε2	86	82
Ile[839] O1−Fe−Wat[842] O	82	
His[499] Nε2−Fe−Ile[839] O1[b]	167	159
His[690] Nε2−Fe−Wat[842] O[b]	154	
Asn[694] Oδ1−Fe−His[499] Nε2	75	68
Asn[694] Oδ1−Fe−His[690] Nε2	92	96
Asn[694] Oδ1−Fe−Ile[839] O1	79	92
Asn[694] Oδ1−Fe−Wat[842] O	66	
Asn[694] Oδ1−Fe−His[504] Nε2[b]	161	157

[a] Values derived from PDB entry 2SBL. [b] These angles should be 180° for ideal octahedral coordination; all other angles should be 90°.

Coordination of the Active Site Fe Atom. Geometric features of the Fe site are presented graphically in Figure 5 and quantified in Table 4. Under the conditions of this study, the active site Fe has five ligands at distances less than 2.6 Å: the side chains of His[499], His[504], and His[690]; one oxygen atom (O1) from the carboxylate group of the protein's terminal residue, Ile[839]; and one water molecule (number 842). A sixth potential ligand, the side chain Oδ1 atom of Asn[694], is positioned in an appropriate direction to act as a sixth ligand but is 3.05 Å distant from the Fe atom. Thus Asn[694] is weakly associated with the Fe but completes the coordination sphere. The binding of the terminal carboxylate

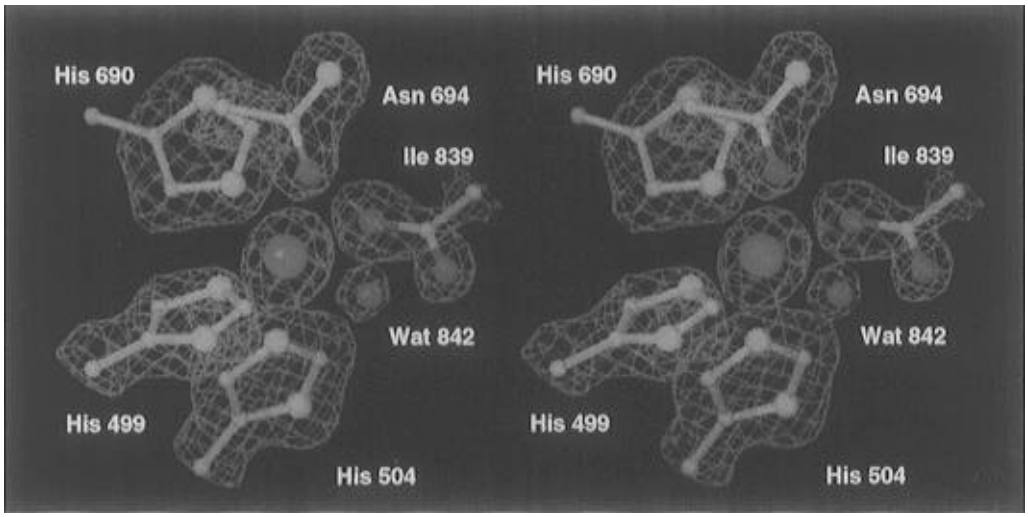

FIGURE 6: $|F_o| - |F_c|$ density from a simulated annealing omit map (Hodel et al., 1992) calculated in the vicinity of the Fe atom and its ligands. The omitted region consisted of two overlapping spheres of 8 Å radius surrounding Wat^{842} and Asn^{694}. The map was calculated using data in the range of 10−1.4 Å resolution and is contoured at 2.5σ, where σ is the standard deviation of the map. The side chains of residues 499, 504, 690, and 694 as well as the atoms C, O1, and O2 of Ile 839 are illustrated. Atom colors are as follows: C, orange; N, cyan;, O, red; Fe, magenta.

is unequivocally monodentate: the second carboxylate oxygen atom (O2), which forms a hydrogen bond with the water ligand, is 3.55 Å distant from the Fe.

Relative to our previous report, there are two significant changes: a water molecule has been located *trans* to His^{690} and adjacent to the active site cavity; the relatively weak association of Asn^{694} has been established. Relative to the structure of Boyington et al. (1993), the most significant difference is the observation of the water ligand. Its absence in the Boyington 2SBL model[4] seems likely to be a consequence of the differences in experimental conditions, which include temperature (100 K in this work vs room temperature), pH (5.6 in this work vs 7.0), and the identity and concentration of cosolutes: 8−14% PEG-3350, 28% ethylene glycol, and 0.1 M sodium acetate in this work vs 4.6 M sodium formate, 1.0 M ammonium acetate, 600 mM LiCl, and 20 mM MES. The water ligand was clearly visible in the first maps calculated with 1.4 Å data and was among the first waters added to the model. It has a refined *B*-factor of 27 $Å^2$, which is comparable to the *B*-factor for the Fe atom, 22 $Å^2$. However, the water ligand was not observed in early stages of the current study, which depended on lower resolution data measured at room temperature and phases based on an unrefined and incomplete model. A simulated annealing omit map confirming the presence of the water ligand and illustrating the orientation of Asn^{694} is shown in Figure 6. Additional information pertaining to the presence of a water ligand and to the possible stronger binding of Asn^{694} in other circumstances is considered below.

The average Fe−ligand distance in this crystal structure of ferrous L-1, 2.33 Å (excluding Asn^{694}), as well as each of the individual distances, is larger than the first shell distance of 2.13 Å determined for a five-coordinate model with three N and two O ligands in a recent EXAFS study of L-1 by Scarrow et al. (1994); prior EXAFS studies of L-1 yielded shorter average values between 2.05 and 2.09 Å for models with 6 ± 1 ligands (Navaratnam et al., 1988; Van der Heijdt et al., 1992). The average distance in the four-coordinate, Boyington 2SBL model is 2.22 Å. Based on a survey of relevant small molecule crystal structures (Scarrow et al., 1994), the average Fe(II)−ligand distance is expected to be within a range of 2.1−2.2 Å. The difference between the other measurements or expected values and the larger average value observed in this crystal structure may be within the random error of the analysis. However, the consistently longer individual distances suggest a systematic error that may, for example, be related to the refinement method; note that no specific restraints were applied to the Fe−ligand distances whereas bonds, angles, and nonbonded contacts among protein atoms were strongly restrained. As noted above, alternative refinement schemes or methods will be tested in future studies.

The Fe-binding site has distorted octahedral geometry but is much less distorted than indicated by previously reported data. For example, the ligand−Fe−ligand angles for the four strong protein ligands deviate from the expected values for an octahedron by an average of 9°. For these ligands, Boyington et al. (1993) originally reported data that gave an average of 14°, whereas the value for their 2SBL model is 12°. For our model, inclusion of the water ligand and Asn^{694} increases the value to 11°. Boyington et al. (1993) also noted that the Fe−Nϵ2 bond for His^{499} was 33° out of the plane of the imidazole ring in their model and discussed the possibility of sp^3 hybridization for this nitrogen; the bonds and rings were described as coplanar for His^{504} and His^{690}. In our model, the Fe−Nϵ2 bonds are 11°, 6°, and 1° out of the plane of the ring for His^{499}, His^{504}, and His^{690}, respectively; the corresponding values for the Boyington 2SBL model are 18°, 6°, and 3°.

In our model, the locations of His^{504} and Asn^{694} may be considered axial by geometric criteria. The Fe atom is displaced by 0.32 Å toward Nϵ2 of His^{504} from the least-squares plane defined by the bonded atoms of the other four ligands, which themselves deviate from the plane by a root-mean-square value of 0.17 Å. Thus, if Asn^{694} is discounted

[4] Some structural data reported in Boyington et al. (1993) are inconsistent with the associated PDB entry, 2SBL. Descriptions ascribed to "the (Boyington) 2SBL model" are derived from the PDB coordinates. When it is necessary to refer to conflicting data appearing in Boyington et al. (1993), an explicit reference to that publication has been made.

as a ligand, the overall coordination could be described as square pyramidal with His^{504} as the axial ligand.

The results of spectroscopic studies are generally in agreement with the geometry observed in our crystal structure and are certainly more consistent with this structure than with the four coordinate Boyington 2SBL model. Two independent examinations of XANES data, coupled to the EXAFS studies cited above, produced preferred models with five- (Scarrow et al., 1994) or six-coordinate Fe in the ferrous form of the enzyme (Van der Heijdt et al., 1992). CD and MCD data were interpreted in terms of a six-coordinate rhombically distorted octahedral site (Whittaker & Solomon, 1988). Mössbauer spectra were also consistent with octahedral symmetry and led by extrapolation to a proposal of a six-coordinate site (Dunham et al., 1990). It is important to note that many of these measurements, as well as the diffraction measurements reported here, were made with frozen samples in the presence of low molecular weight alcohols used as glassing agents. The X-ray absorption studies of Scarrow et al. (1994) suggest that such reagents may affect the coordination number and Fe−ligand distances.

The Water Ligand. Wat^{842} binds to the Fe in the equatorial plane of the complex. Its position is *trans* to His^{690} and is at the boundary of a large internal cavity that is presumably the site of fatty acid binding. As illustrated in Figure 5, binding of the water ligand is stabilized by a hydrogen bond with the nonligated oxygen atom (O2) of the terminal carboxylate group. For the ferrous enzyme, evidence relevant to the presence or absence of a water/hydroxide ligand from other methods is limited. EXAFS studies suggest 5−7 ligands (Van der Heijdt et al., 1992; Scarrow et al., 1994) and are thus consistent with models that include a bound water. In contrast, Whittaker and Solomon (1988) did not observe changes in the MCD spectrum of ferrous L-1 in the presence of KN_3, NaF, or NaOCN and concluded that binding of the Fe by exogenous (nonprotein) ligands did not occur. In the case of the ferric enzyme, the evidence is somewhat more compelling in that the EXAFS studies strongly favor a six-coordinate site and suggest that an hydroxide ion is a ligand. In addition, Nelson (1988) reported that features of EPR spectra obtained from ferric L-1 were broadened in the presence of $H_2{}^{17}O$. Although the EPR spectrum was perturbed by the addition of ethanol or cyanide, broadening was not eliminated.

Asn^{694}. As noted above, Oδ1 of Asn^{694} is too distant from the Fe (3.05 Å) to be considered a strong ligand. This result is consistent with the observation that substitution of the corresponding asparagine (Asn^{713}) in the *E. coli* expressed soybean L-3 isozyme with histidine, alanine, or serine resulted in proteins that maintained the ability to bind Fe (Kramer et al., 1994). Similar results were obtained for replacements of Asn^{554} in *E. coli* expressed human 5-Lox with glutamine, aspartate, or alanine (Hammarberg et al., 1995). MCD studies have suggested that the ferrous and ferric forms of L-1 have very similar Fe coordination (Zhang et al., 1991). In contrast, changes in the Fe coordination as a result of oxidation and/or substrate binding have been implied by several X-ray absorption studies (Scarrow et al., 1994; Van der Heijdt et al., 1992, 1995). Thus it remains possible that Asn^{694} acts as a ligand in forms of the enzymes important for catalysis. Indeed, the alanine- and serine-substituted L-3 proteins described above were catalytically inactive, and the 5-Lox variants described above all possessed marginal or unmeasurable activity.

In this context it is interesting that the orientation of the Asn^{694} carboxamide group is involved in four polar or weakly polar interactions: (i) the interaction between O$\delta1^{694}$ and the Fe atom; (ii) a possible hydrogen bond between O$\delta1^{694}$ and N$\epsilon2$ of Gln^{697} (3.1 Å, distorted geometry); (iii) a hydrogen bond between N$\delta2^{694}$ and the backbone carbonyl group of Leu^{754} (2.8 Å, good geometry); and (iv) an amino−aromatic interaction (Burley & Petsko, 1986; Levitt & Perutz, 1988) involving an N$\delta2^{694}$ proton and the side chain of Phe^{695} (3.1 Å from Nδ2 to the center of the ring, good geometry). Sequence alignments indicate that the latter interaction is highly conserved among lipoxygenases with an asparagine residue at the equivalent position: all but one also have a phenylalanine residue at the position corresponding to Phe^{695}.

Each of these interactions is clearly visible in Figure 5. The dominant interaction is likely to be iii, and it appears that only the weaker interactions ii and iv are in conflict with the formation of an Fe−O$\delta1^{694}$ bond: adjusting the side chain torsion angles of Asn^{694} so as to move O$\delta1^{694}$ closer to the Fe by at least 0.6 Å does not significantly lengthen the N-to-O distance of interaction iii nor does it require poor hydrogen-bonding geometry. Such an adjustment appears to eliminate interactions ii and iv, but the latter may be rescued by a small adjustment of Phe^{695}. Moreover, the proton that mediates interaction ii also participates in a second, distorted hydrogen bond with the carbonyl group of Asn^{694}, and this hydrogen bond may be strengthened if the conformation of Asn^{694} is adjusted; i.e., the energetic cost of sacrificing the interaction of N$\epsilon2^{797}$ with O$\delta1^{694}$ may be recovered by an improved interaction with the carbonyl group. Thus it is structurally feasible that oxidation of the Fe, or other changes in the structure of the active site that occur on substrate binding, might strengthen the ligation of O$\delta1^{694}$ to the Fe.

Observations concerned with the behavior of the Fe atom during refinement and the appearance of the final electron density maps may also be relevant to the possible formation of a stronger Fe−O$\delta1^{694}$ bond. Simulated annealing was performed at various stages of the refinement both along the productive path and in control experiments. It was generally observed that the ligands would return to the same positions within a few hundredths of an angstrom but that the Fe position would vary by tenths of angstroms along the axis of N$\epsilon2^{504}$ and O$\delta1^{694}$. In fact, the final model has an extreme Fe−O$\delta1^{694}$ separation. Furthermore, the final electron maps show features consistent with an anisotropic fluctuation of the position of the Fe. The ($2|F_o| - |F_c|$) map has negative densities of small volume, at three times the standard deviation of the map, within the ($|F_o| - |F_c|$) density of the Fe atom and along the His^{690}−Wat^{842} axis. In addition, there are positive features associated with both the Fe atom and O$\delta1^{694}$ along the vector between them. Taken together, these observations are consistent with disorder in the Fe−O$\delta1^{694}$ interaction and suggest that the Fe and O$\delta1^{694}$ are more closely associated in a significant fraction of the molecules in the crystal.

It must be noted that amido oxygen atoms are weak ligands for Fe and other transition metals and that Fe ligation by an amido oxygen would be unusual. In fact, a search for small molecule examples of Fe−amido oxygen bonds using the Cambridge Structure Database System (Allen & Kennard,

1993) revealed only three structurally distinct compounds. Nevertheless, four examples of bonds between transition metals and amido oxygen atoms have been observed in protein crystal structures reported within the past year. These include a glutamine to Mn bond in isopenicillin *N*-synthase (Roach et al., 1995), asparagine to Zn bonds in purple acid phosphatase (Sträter et al., 1995) and calcineurin (Griffith et al., 1995), and a bond between an asparagine and a unknown metal (Mn, Fe, or Co) in protein phosphatase 1 (Goldberg et al., 1995).

The Extended Fe Binding Site. The interactions that contribute to the binding of the water ligand and the positioning of Asn^{694} are part of a larger network of hydrogen bonds and noncovalent interactions that may play important roles in the structure of the active site in states along the reaction pathway. Many of these are illustrated in Figure 5. Only one of the histidine ligands, the axial ligand His^{504}, is positioned by a protein−protein hydrogen bond. Moreover, it appears that the location of His^{504} may be a critical anchor in an otherwise relatively plastic active site. First, the hydrogen bond links Nδ2^{504} with the Oδ1 atom of a conserved residue, Asn^{539}. Second, Asn^{539} is in turn hydrogen bonded via an Nδ proton to the side chain Oϵ1 atom of the conserved residue Glu^{508}, which is further hydrogen bonded via its Oϵ2 atom to the backbone amide group of Gly^{218}. It is possible that the entire chain of hydrogen bonds is important for positioning His^{504}; for human 5-Lox, replacement of the residue corresponding to Glu^{508} (of L-1) with a glutamine residue resulted in an inactive enzyme (Ishii et al., 1992).

His^{690}, which is *trans* to the water ligand and the substrate binding cavity, forms hydrogen bonds via its Nδ1 proton with one or both of two water molecules that are part of a chain of waters leading to the surface of the protein. Contrary to Figure 4 of Boyington et al. (1993),[5] the Nδ1 atom of His^{499} is not involved in a hydrogen bond with Gln^{495} or any other protein residue or water molecule. The closest potential acceptor is indeed Oϵ1^{495}, but it is 3.6 Å distant from Nδ1^{504} and 2 Å (34°) out of the plane of the imidazole ring. It is interesting that Oϵ1^{495}, which borders the fatty acid binding cavity, is an unsatisfied hydrogen bond acceptor. Perhaps rearrangement of the active site during catalysis causes it to accept a hydrogen bond from His^{499} or to interact with the substrate.

The apparent requirement for an isoleucine or valine terminal residue (839 in L-1) was discussed above. Although this preference may be explained by interactions between the side chain and the substrate, it may also reflect a role for a β-branched side chain in functional Fe coordination. Note that nonbonded contacts involving the two Cγ atoms may influence the orientation of the terminal carboxylate as well as the conformations of Phe^{695} and Leu^{754}.

Cavities and Substrate Access Channels. Boyington et al. (1993) identified two major cavities in the C-domain. Cavity I is a conical invagination that leads from a ~10 Å wide opening at the surface of the protein (~20 Å from the Fe) to a small pocket defined by the backbone carbonyl group of Ser^{498} and the side chains of Thr^{503}, Leu^{693}, and Val^{693}. At this point, 8 Å distant from the Fe atom, the pocket is just large enough for a single water molecule. In the view of Figure 5, the end of cavity I lies approximately on a line that ends at the Fe atom and bisects the His^{499}−His^{690} edge of the coordination polyhedron. Cavity II is a 40 Å long, narrow, and winding cavity that extends in two directions from a bend near the Fe atom. The more remote end of the cavity was defined by the side chains of residues Met^{341} and Leu^{480}, approximately 30 Å from the Fe; the closer end was not well defined. In the view of Figure 5, the water ligand is above the bend in the cavity, and the remote end is to the left.

Boyington et al. (1993) noted that many of the residues[6] lining cavity I have hydrophobic side chains and that some are conserved. They concluded that the cavity "presents an ideal path for the movement of molecular oxygen from the outside into one of the two unoccupied coordination sites of the iron". Both points warrant further discussion.

First, it is misleading to characterize this cavity as hydrophobic. In our current model, at least 22 water molecules are found in cavity I and within 20 Å of the Fe atom. All but five of these waters form hydrogen bonds with the protein; nine distinct backbone carbonyl or amide groups and 10 different side chains are involved, and there are 25 water−protein hydrogen bonds.

Second, the "unoccupied" Fe coordination sites are not empty in our model but rather are occupied by the side chain of Asn^{694} and the water ligand. Moreover, a direct path to these sites is blocked by a 4.4 Å contact between the Cγ2 atoms of Thr^{503} and Val^{693} (>6 Å from the Fe) as well as the side chains of His^{499}, His^{690}, and Asn^{694}. Movement of an oxygen molecule to either site from cavity I requires displacement of the existing occupant as well as a substantial rearrangement of the protein structure that must (i) remove the restriction formed by the side chains of Thr^{503} and Val^{693} and (ii) either disturb a strongly associated ligand (His^{499} or His^{690}) or drastically alter the protein structure near Asn^{694}. Contrary to the conclusions of Boyington et al. (1993), a minor reorientation of the Asn^{694} side chain is not sufficient to provide access to the Fe. The key restriction in this case is a 4.1 Å contact between the Fe and Cβ^{694}, which is unaffected by changes in the side chain torsion angles and unlikely to be increased by alteration of the backbone torsion angles or gross displacement because Asn^{694} is within helix 18. Furthermore, reorientation of Asn^{694} so as to free a site for O_2 to bind, as proposed by Boyington et al. (1993), seems unlikely. Not only must this reorientation disrupt four polar interactions involving Oδ1^{694} and Nδ2^{694} (see above), it must move Oδ1^{694} within a limited volume by approximately 3 Å against the resistance of several nonbonded contacts. It appears that such an event would require a substantial rearrangement of additional residues and/or a gross movement of helix 18.

Given the above, it seems unlikely that cavity I is ideally or uniquely suited for O_2 delivery. A likely alternative would be a path that utilizes cavity II, which provides immediate access to the Fe through the water binding site. A third

[5] Figure 4 of Boyington et al. (1993) is inconsistent with the results of this work and with PDB entry 2SBL and appears to be in error. The figure implies either that one atom of the side chain carboxamide group of Gln^{495} is both a hydrogen bond donor and an acceptor or that the side chain His^{499} is in the imidazolate (doubly deprotonated) state.

[6] Residues that surround cavity I include Cys^{357}, Val^{358}, Ile^{359}, Arg^{360}, Tyr^{409}, Ile^{412}, Tyr^{493}, Met^{497}, Ser^{498}, Leu^{501}, Asn^{502}, Thr^{503},Val^{570}, Asn^{573}, Trp^{574}, Val^{575}, Asp^{578}, Gln^{579}, Leu^{581}, Asp^{584}, Lys^{587}, Arg^{588}, Tyr^{610}, Trp^{684}, Leu^{689}, and Val^{693}.

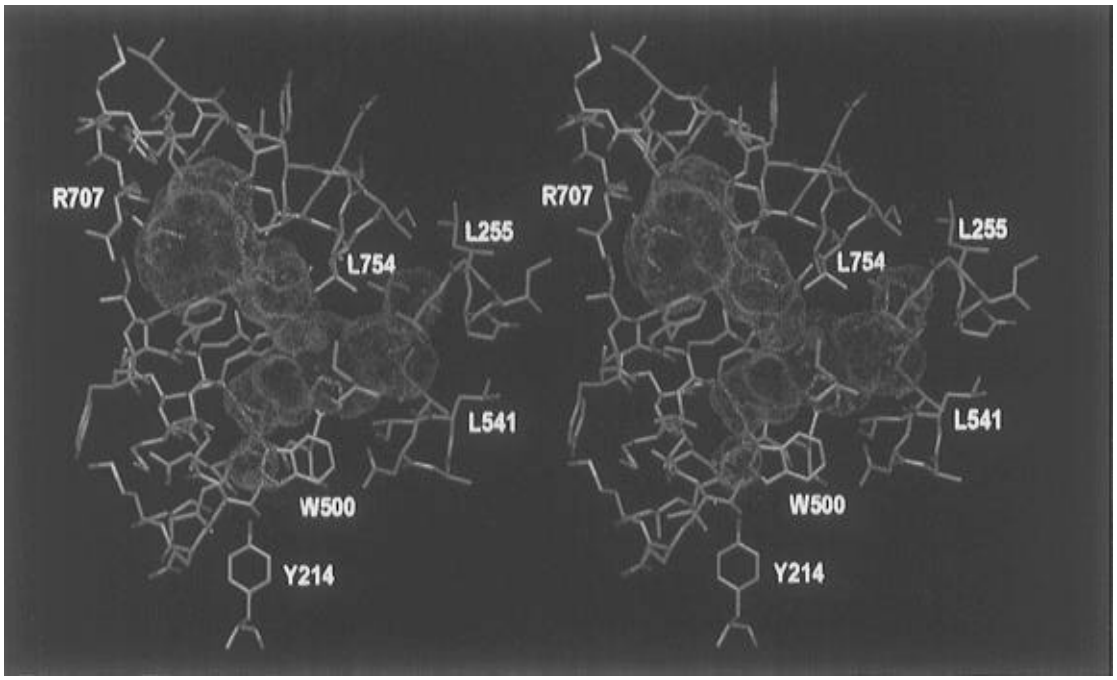

FIGURE 7: Shape, volume, and location of cavity IIa (the putative fatty acid binding pocket) illustrated as a molecular surface (Connolly, 1983) in green dots superimposed on the refined model. The figure was prepared with the QUANTA package as described in the text using a probe radius of 1.4 Å and radii of 1.90, 1.82, 1.65, and 2.00 Å for C, N, O, and S atoms, respectively. A few residues are numbered to mark the location of features described in the text. With the exception of Leu[546] and Phe[557], residues in front of the cavity were deleted to allow a clear view. The iron atom is represented by a purple sphere. Water molecules are represented by smaller, red spheres. Carbon, nitrogen, oxygen, and sulfur protein atoms are in gray, blue, red, and yellow, respectively.

possibility follows a chain of water molecules that leads from the vicinity of the side chain of His[690] to the surface of the protein along a winding and narrow 20 Å pathway that passes between Pro[834] and Ile[839] and reaches the surface in the vicinity of His[248] and Arg[533]. Leu[255], Asn[534], Ala[758], and Glu[761] are also along the path. Approach of O_2 via this route seems less likely because it is substantially more confined and convoluted than the cavity I path and would similarly require rearrangement of residues in the Fe binding site to reach the Fe or fatty acid substrate.

Boyington et al. (1993) describe cavity II as a narrow passage with sharp bends near the Fe atom and at the constriction formed by the side chains of Arg[707] and Val[354]. Because the passage is completely blocked by a 3.5 Å contact between these side chains, it is more appropriate to define two cavities, IIa[7] and IIb,[8] with IIa close to the Fe and serving as the probable binding cavity for the fatty acid substrate. Both IIa and IIb are sandwiched between two layers of helices, with helices 9 and 11 on one side and helices 2, 6, 18, and 21 on the other. A molecular surface for cavity IIa is illustrated in Figure 7.

Cavity II is wholly internal, and thus the point(s) of entry and exit for the fatty acid is (are) difficult to identify. Boyington et al. (1993) proposed that the substrate enters cavity IIb after a rearrangement of the side chains of Met[341] and Leu[480] at the distal end of the cavity, which in our model would open a 6.5 Å wide portal. Rearrangement of the side chains of Arg[707] and Val[354] would be required to allow the fatty acid to pass from cavity IIb into IIa. Although subtle changes at this point should be sufficient to open a channel for water, a larger opening may be required for fatty acid substrates containing one or more (*Z*,*Z*)-1,4-pentadiene units. If so, the side chain of Arg[707] must be displaced significantly, which would require the disruption of a hydrogen bond (perhaps a buried salt bridge) it forms with Asp[490]. Moreover, the Arg[707] side chain would be displaced into IIb or IIa, either further blocking the access channel (IIb) or reducing the free volume in the proposed binding cavity (IIa).

Assuming the two barriers can be removed, the substrate must snake along a convoluted path as much as 40 Å in length, and it must displace a large number of protein-associated water molecules as it goes. Displacement of water from cavity IIb may be difficult because the space is relatively confined (cf. Levitt & Park, 1993) and because 11 of the 14 crystallographically ordered waters that fill the cavity have low *B*-values less than 13 Å^2. In that only 3% of the waters have $B \leq 13$ Å^2, the IIb waters are among the least mobile in the crystal structure.

Although the cavity IIb route is feasible, the analysis above suggests that it is not ideal and that alternative routes may exist. One promising alternative is located at the opposite end of cavity IIa, far away from IIb and relatively close to the Fe atom. In this case, access to IIa is barred 8−10 Å from the Fe atom by two contacts between side chain atoms: Thr[259]Oγ1 and Leu[541]Cβ, 4.6 Å; Lys[260]Cβ and Leu[541]Cδ2, 5.3 Å. Sterically unhindered adjustments of the side chains of Thr[259] and Leu[541] appear to be sufficient to open a channel wide enough (>6 Å) for entry of a fatty acid.

[7] Residues that surround cavity IIa include Tyr[214], Leu[255], Glu[256], Gly[258], Thr[259], Leu[496], His[499], Trp[500], His[504], Ile[538], Leu[541], Ala[542], Ile[547], Ile[553], Val[564], Ser[567], and Ile[839], which line the subcavity proximal to the Fe (see Figure 7); Gln[495], Leu[546], Ile[753], and Leu[754], which contribute to both the proximal and distal subcavities; and Val[354], Ser[491], His[494], Ile[552], Ile[553], Thr[556], Phe[557], Gln[697], Gly[701], Ile[704], Arg[707], Thr[709], Ser[747], Leu[748], Val[750], and Ile[751], which line the distal subcavity.

[8] Residues that surround cavity IIb include Trp[340], Met[341], Glu[345], Phe[346], Glu[349], Met[350], Gly[353], Val[354], Asn[355], Val[358], Ile[359], Leu[407], Leu[480], Lys[483], Ala[484], Val[486], Ile[487], Asn[490], Tyr[700], Met[705], Asn[706], Arg[707], Pro[708], and Tyr[734].

The characteristics of the crystallographically ordered water on both sides of the barrier also seem favorable. Two waters are located 5.6 Å apart on opposite sides of the barrier, Wat^{961} (B = 31 Å^2) inside the cavity and Wat^{1469} (B = 32 Å^2) on the outside. The path from Wat^{1469} to the surface is direct and requires displacement of only one ordered water (2063) that is effectively at the surface (5 Å away) and has a *B*-value of 42 Å^2. On the inside, Wat^{961} is hydrogen bonded to the carbonyl oxygen of Leu^{538} but is otherwise surrounded by the nonpolar side chain atoms of Leu^{538}, Leu^{541}, Ala^{542}, Leu^{754}, and Thr^{259}. The distance from Wat^{961} to the Fe's water ligand is 6.3 Å, and only one water molecule intervenes (Wat^{841}, B = 49 Å^2).

Examination of the internal and external molecular surfaces suggests two additional points where a channel might open to connect cavity IIa with the external surface. One of these is marked by Wat^{1828}, which is within 10 Å of the surface and is blocked by the side chain and backbone atoms of Leu^{255} and Ile^{573}. The second is marked by Wat^{2222} (<10 Å from the surface) and is blocked primarily by the side chain of Val^{564}. A branched, solvent-filled channel leads to the surface from the remote side of Val^{564}.

All of the proposed access routes rely on entrances that are slightly smaller than the proposed "portal" of the adipocyte lipid binding protein, ALBP (Banaszak et al., 1994), where, based on an examination of PDB entries 1LIB (Xu et al., 1993) and 1ADL (LaLonde et al., 1994), the restrictive distances between backbone and/or Cβ atoms are larger than 7 Å. If a larger entrance is required, a gross movement of one or more helices would be required; the most likely scenario would be the opening of a gap between helix 2 and helix 11. Any of the proposed entrances might also play a role as exit ports for the ordered and disordered water (Ernst et al., 1995) that occupies the cavities in the absence of fatty acids.

Possible Modes of Substrate Binding and Their Relevance to Chemical Mechanisms. Although the structure of the enzyme is not static and may change upon binding of substrate, it is possible that important features of the structure of the bound substrate, and its interactions with the protein can be predicted from the existing information. For example, it is clear that cavity IIa can accommodate C18 and C20 fatty acids in relatively extended conformations but cannot accommodate U-shaped conformations such as those observed for arachidonic acid bound to ALBP (LaLonde et al., 1994) or for stearate, elaidate, and oleate bound to human muscle fatty acid binding protein (Young et al., 1994).

Viewed in terms of its molecular surface, as in Figure 7, the shape of cavity IIa is complex and has three major branches. However, it is possible to describe cavity IIa in terms of the intersection of two subcavities.[7] The subcavity proximal to the Fe bends by roughly 120° as it passes between Leu^{546} and the Fe atom, has two arms roughly 6−8 Å in length, and has a shape that is approximately 2-fold symmetric about an axis that passes through the Fe atom. One arm leads to the alternative entrance near Leu^{541}, where it ends in a pocket lined primarily with nonpolar side chains, as described above. The other arm passes between Leu^{546} and Gln^{495} to a hydrophobic pocket surrounded by Leu^{496}, Leu^{547}, Ile^{553}, and Trp^{500}. Buried waters are found at the ends of both arms.

One end of the second subcavity intersects the proximal subcavity 4 Å from the Fe atom in an extremely restricted space between the side chains of Gln^{495}, Leu^{546}, and Leu^{754}. From this point, the subcavity bends gently and passes through a narrow opening bordered by the side chains of Gln^{495}, Phe^{557}, Ile^{751}, and Ser^{747} before opening into a larger space that extends to the barrier of Arg^{707}, approximately 12−14 Å from the intersection. A cluster of eight ordered water molecules fills this end of the subcavity, which is bordered by several side chain and backbone atoms capable of forming hydrogen bonds.

Given the flexibility of linoleic acid [9(*Z*),12(*Z*)-octadecadienoic acid] and the protein side chains that border the cavity, it is not possible to define a specific binding mode for the substrate based on the current information. In fact, a unique mode may not be expected: under laboratory conditions, L-1 will convert linoleic acid to both 9- or 13-hydroperoxide products with varying degrees of regio- and stereospecificity (Christopher et al., 1972; Gardner, 1989). L-1 similarly produces different hydroperoxides from 9*Z*,-12*E* and 9*E*,12*Z* substrates (Funk et al., 1987) and exhibits flexible regiospecificity in its reactions with a selection of (*Z*,*Z*)-3,6-dienyl 1-monoadipates (Datcheva et al., 1991; Scheller et al., 1995).

In some of these cases, the stereochemistry of the products is consistent with the existence of binding modes that are reversed in a head-for-tail sense (Gardner, 1989). Given the generally symmetric, bent shape of the proximal subcavity, it is tempting to speculate that the substrate/product variations might be explained by head-to-tail reversed binding modes that utilize only this space and not the distal subcavity. In fact, the bend in the proximal subcavity near the Fe is generally compatible with the ∼90° change of direction that would be imposed on the substrate by a *cis,cis*-pentadiene unit with both the C10−C11 and C11−C12 bonds (i.e., both single bonds involving the central methylene carbon, C11) in the *trans* conformation. However, both arms of the proximal subcavity appear to be somewhat too short to accommodate the longer branch of the substrate if it were to adopt a generally extended conformation. Thus, kinking of the aliphatic chains of the substrate, extension of the substrate through one of the potential openings, and/or rearrangement of the protein would be required for the substrate to bind without utilizing space in the distal subcavity.

It is also necessary to consider whether linoleic acid can bind in a mode that utilizes both subcavities. Although the portal between the subcavities is narrow, manual fitting of models for linoleate and arachidonic acids suggests it is possible for the polar end of the substrate to bind in the water-filled pocket at the end of the distal subcavity while the aliphatic end of the substrate binds in the pocket of the proximal subcavity that terminates near Leu^{541}. In this mode, the substrate has a relatively extended conformation but must gently bend in opposite directions at two points, as in the symbol "∼". One bend is necessary in the vicinity of the Fe atom, where the pentadiene moiety should bind. A second bend would allow the substrate to reach into and conform to the distal subcavity.

Because the substrate's binding mode is uncertain, it is difficult to evaluate structural features of the various steps in the several proposed chemical mechanisms (Nelson & Seitz, 1994). Nevertheless, the environment of the Fe and its water ligand suggest some limitations with respect to the proposed mechanisms. For example, a mechanism proposed

by Corey and Nagata (1987) postulates a σ-organo-Fe intermediate. In terms of the structure of the free enzyme, the Fe's water ligand would be replaced by the C13 atom of the substrate. Although the water site borders on the substrate binding cavity, replacement of the water by C13 seems unlikely because access by a large molecule is restricted by the side chains of Gln^{495}, Gln^{697}, Leu^{754}, and Ile^{839} as well as the C-terminal carboxylate: these protein groups would be in steric conflict with the rest of the pentadiene unit and/or C14. Thus, unless a significant rearrangement of the protein structure occurs, it would appear that direct binding of the substrate to the Fe is prohibited. In contrast, the formation of a peroxy-bridged (Fe−OO−substrate) intermediate, which is a feature of this and other reaction schemes, is not necessarily prohibited; the peroxy group would span the constriction, allowing the substrate to remained in a less restricted space.

The limited space in the substrate binding cavity may also have significance with respect to consideration of the structures of possible fatty acyl radical reaction intermediates. One potential intermediate is the pentadienyl radical wherein the unpaired electron is delocalized over C9 through C13, implying that C8 through C14 are essentially coplanar (cf. Nelson et al., 1990). A recent alternative proposal (Nelson et al., 1994) suggests that the radical may be a Δ^{12}-[9,10,-11]-allyl radical, an isomer of the pentadienyl radical wherein rotation about the C11−C12 bond disallows delocalization of the radical over the C12−C13 π orbital. In this case, the restriction of coplanarity applies separately to the C8 through C11 and C11 through C14 groups, and the substrate must be significantly kinked at C11. The structure of the free enzyme appears to favor such a kinked conformation because of the distances between side chains that border the presumptive fatty acid binding site. Specifically, for the substrate's pentadiene unit to approach the site occupied by the Fe's water ligand, C8 through C14 must fit between the side chains of Gln^{495} and Ile^{538}, which are separated by 10.0 Å. There appears to be insufficient space for a unit that is coplanar from C8 through C14 because in such a structure the distances C9−C13 and C8−C14 are approximately 5.0 Å and 6.0 Å, respectively. Thus some or all of the atoms C8, C9, C13, and C14 would make unacceptably close contacts with Gln^{495} and Ile^{538}. These short contacts could help force a kinked substrate structure, stabilize a kinked radical, and/or promote stereospecific hydrogen abstraction (see below) during radical formation by properly aligning the C11−H bond with the appropriate π system. It should be carefully noted, however, that the L-1 structure does not require formation of a kinked radical, much less a kinked radical of a specific form, because the side chains may reorient on substrate binding and activation of the Fe.

An additional question concerns the identity of an active site base to promote hydrogen abstraction. With linoleic acid as substrate, the predominant reaction involves formal abstraction of the *pro*-S hydrogen atom from C11 and the stereospecific addition of the peroxy group at C13 to form the 13-*S* product. Boyington et al. (1993) suggested that Nϵ2 of His^{499} or the terminal carboxylate might act as proton acceptors. N$\epsilon 2^{499}$ seems an unlikely candidate because it is a ligand to the Fe and must dissociate to accept a proton. Scarrow et al. (1994) argued against the direct involvement of the Fe-bound carboxylate because of its low basicity and suggested an Fe(III)-bound hydroxide, $[Fe(III)-OH]^{2+}$, might serve as the effective base, leading to $[Fe(II)-OH_2]^{2+}$ after acquisition of a hydrogen atom. This proposal is consistent with the presence of a water ligand in the crystal structure as well as recent EXAFS models for the ferric enzyme (Van der Heijdt et al., 1992; Scarrow et al., 1994), which are six coordinate and include a short (1.9 Å) Fe−O bond. It is interesting to note that our attempts to manually fit the substrate models into the proposed binding site suggest it is easier for C11 to approach the site of the water ligand if substrate is kinked in the sense described above.

Scarrow et al. (1994) also discuss several possible schemes for the subsequent steps. Schemes that include a covalent organo-Fe intermediate seem to be disfavored by the crystal structure, as noted above. It also seems unlikely that the mechanism can exploit an Fe-bound hydroxide in the hydrogen abstraction step and subsequently form a peroxy-bridged intermediate. The observed stereochemistry requires that hydrogen atom abstraction and oxygen attack occur on opposite faces of the pentadiene unit, but the environment of the Fe and substrate binding sites does not provide two ligand sites on opposite faces. Thus a mechanism that exploits an Fe-bound hydroxide in the hydrogen abstraction step and an Fe-bound dioxygen in a subsequent step would require the substrate to undergo a large conformational and/or orientational change between the two steps. Such an event seems unlikely within the hindered environment of the active site. Therefore, if the initial hydrogen abstraction is promoted by Fe-bound hydroxide, this interpretation argues against a mechanism wherein the observed variations in regio- and stereospecificity are controlled by the specific location of an Fe-bound dioxygen. On the basis of the structure, it seems more likely that the enzyme engenders specificity by requiring a relatively specific approach of dioxygen through steric interference or other means.

REFERENCES

Allen, F. H., & Kennard, O. (1993) *Chem. Des. Automation News 8*, 31−37.

Balcarek, J. M., Theisen, T. W., Cook, M. N., Varrichio, A., Hwang, S. M., Strohsacker, M. W., & Crooke, S. T. (1988) *J. Biol. Chem. 263*, 13937−13941.

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

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

Bourne, Y., Martinez, C., Kerfelec, B., Lombardo, D., Chapus, C., & Cambillau, C. (1994) *J. Mol. Biol. 238*, 709−732.

Boyington, J. C., Gaffney, B. J., & Amzel, L. M. (1993) *Science 260*, 1482−1486.

Brünger, A. T. (1992) *Nature (London) 355*, 472−475.

Brünger, A. T., Kuriyan, J., & Karplus, M. (1987) *Science 235*, 458−460.

Brünger, A. T., Krukowski, A., & Erickson, J. W. (1990) *Acta Crystallogr., Sect. A 46*, 585−593.

Burley, S. K., & Petsko, G. A. (1986) *FEBS Lett. 203*, 139−143.

Cheesbrough, T. M., & Axelrod, B. (1983) *Biochemistry 22*, 3837−3840.

Chen, X. S., Kurre, U., Jenkins, N. A., Copeland, N. G., & Funk, C. D. (1994) *J. Biol. Chem. 269*, 13979−13987.

Christopher, J. P., Pistorius, E. K., Regnier, F. E., & Axelrod, B. (1972) *Biochim. Biophys. Acta 289*, 82−87.

Connolly, M. L. (1983) *Science 221*, 709−713.

Corey, E. J., & Nagata, R. (1987) *J. Am. Chem. Soc. 109*, 8107−8108.

Cowan, S. W., Newcomer, M. E., & Jones, T. A. (1990) *Proteins: Struct. Funct., Genet. 8*, 44−61.

Datcheva, V. K., Kiss, K., Solomon, L., & Kyler, K. S. (1991) *J. Am. Chem. Soc. 113*, 270−274.
de Groot, J. J., Veldink, G. A., Vliegenthart, J. F., Boldingh, J., Wever, R., & van Gelder, B. F. (1975) *Biochim. Biophys. Acta 377*, 71−79.
Dunham, W. R., Carroll, R. T., Thompson, J. F., Sands, R. H., & Funk, M. O., Jr. (1990) *Eur. J. Biochem. 190*, 611−617.
Engh, R. A., & Huber, R. (1991) *Acta Crystallogr., Sect. A 47*, 392−400.
Ernst, J. A., Clubb, R. T., Zhou, H. X., Gronenborn, A. M., & Clore, G. M. (1995) *Science 267*, 1813−1817.
Ford-Hutchinson, A. W., Gresser, M., & Young, R. N. (1994) *Annu. Rev. Biochem. 63*, 383−417.
Funk, M. O., Jr., Andre, J. C., & Otsuki, T. (1987) *Biochemistry 26*, 6880−6884.
Gardner, H. W. (1989) *Biochim. Biophys. Acta 1001*, 274−281.
Gardner, H. W. (1991) *Biochim. Biophys. Acta 1084*, 221−239.
Goldberg, J., Huang, H., Kwon, Y., Greengard, P., Nairn, A. C., & Kuriyan, J. (1995) *Nature (London) 376*, 745−752.
Griffith, J. P., Kim, J. L., Kim, E. E., Sintchak, M. D., Thomson, J. A., Fitzgibbon, M. J., Fleming, M. A., Caron, P. R., Hsiao, K., & Navia, M. A. (1995) *Cell 82*, 507−522.
Haining, J. L., & Axelrod, B. (1958) *J. Biol. Chem. 232*, 193−202.
Hammarberg, T., Zhang, Y. Y., Lind, B., Radmark, O., & Samuelsson, B. (1995) *Eur. J. Biochem. 230*, 401−407.
Hodel, A., Kim, S. H., & Brünger, A. T. (1992) *Acta Crystallogr., Sect. A 48*, 851−858.
Ishii, S., Noguchi, M., Miyano, M., Matsumoto, T., & Noma, M. (1992) *Biochem. Biophys. Res. Commun. 182*, 1482−1490.
Jones, T. A., Zou, J.-Y., Cowan, S. W., & Kjeldgaard, M. (1991) *Acta Crystallogr., Sect. A 47*, 110−119.
Kato, T., Shirano, Y., Iwamoto, H., & Shibata, D. (1993) *Plant Cell Physiol. 34*, 1063−1072.
Kleywegt, G. J., & Jones, T. A. (1994) *Acta Crystallogr., Sect. D 50*, 178−185.
Kramer, J. A., Johnson, K. R., Dunham, W. R., Sands, R. H., & Funk, M. O., Jr. (1994) *Biochemistry 33*, 15017−15022.
Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
LaLonde, J. M., Bernlohr, D. A., & Banaszak, L. J. (1994) *Biochemistry 33*, 4885−4895.
Laskowski, R., MacArthur, M. W., Moss, D. S., & Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
Levitt, M., & Perutz, M. F. (1988) *J. Mol. Biol. 201*, 751−754.
Levitt, M., & Park, B. H. (1993) *Structure 1*, 223−226.
Luzzati, V. (1952) *Acta Crystallogr. 5*, 802−810.
Minor, W. (1993) XDISPLAYF (computer program), Purdue University, West Lafayette, IN.
Minor, W., Steczko, J., Bolin, J. T., Otwinowski, Z., & Axelrod, B. (1993) *Biochemistry 32*, 6320−6323.
Navaratnam, S., Feiters, M. C., Al-Hakim, M., Allen, J. C., Veldink, G. A., & Vliegenthart, J. F. (1988) *Biochim. Biophys. Acta 956*, 70−76.
Nelson, M. J. (1988) *J. Am. Chem. Soc. 110*, 2985−2986.
Nelson, M. J., & Seitz, S. P. (1994) *Curr. Opin. Struct. Biol. 4*, 878−884.
Nelson, M. J., Seitz, S. P., & Cowling, R. A. (1990) *Biochemistry 29*, 6897−6903.
Nelson, M. J., Cowling, R. A., & Seitz, S. P. (1994) *Biochemistry 33*, 4966−4973.
Nicholls, A., Sharp, K. A., & Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.
Otwinowski, Z. (1991) in *Isomorphous Replacement and Anomalous Scattering* (Wolf, W., Evans, P. R., & Leslie, A. G. W., Eds.) pp 80−86, Science and Engineering Research Council, Daresbury Laboratory, Daresbury, England.
Otwinowski, Z. (1993) in *Proceedings of the CCP4 Study Weekend: Data Collection and Processing* (Sawyer, L., Isaacs, N., & Bailey, S., Eds.) pp 56−62, Science and Engineering Research Council, Daresbury Laboratory, Daresbury, England.
Pistorius, E. K., & Axelrod, B. (1974) *J. Biol. Chem. 249*, 3183−3186.
Pistorius, E. K., Axelrod, B., & Palmer, G. (1976) *J. Biol. Chem. 251*, 7144−7148.
Ramachandran, G. N., & Sasisekharan, V. (1968) *Adv. Protein Chem. 23*, 283−438.
Ramachandran, S., Carroll, R. T., Dunham, W. R., & Funk, M. O., Jr. (1992) *Biochemistry 31*, 7700−7706.
Roach, P. L., Clifton, I. J., Fulop, V., Harlos, K., Barton, G. J., Hajdu, J., Andersson, I., Schofield, C. J., & Baldwin, J. E. (1995) *Nature (London) 375*, 700−704.
Sack, J. S. (1988) *J. Mol. Graphics 6*, 224−225.
Scarrow, R. C., Trimitsis, M. G., Buck, C. P., Grove, G. N., Cowling, R. A., & Nelson, M. J. (1994) *Biochemistry 33*, 15023−15035.
Scheller, G., Jäger, E., Hoffmann, B., Schmitt, M., & Schreier, P. (1995) *J. Agric. Food Chem. 43*, 1768−1774.
Shibata, D., Steczko, J., Dixon, J. E., Hermodson, M., Yazdanparast, R., & Axelrod, B. (1987) *J. Biol. Chem. 262*, 10080−10085.
Shibata, D., Steczko, J., Dixon, J. E., Andrews, P. C., Hermodson, M., & Axelrod, B. (1988) *J. Biol. Chem. 263*, 6816−6821.
Siedow, J. N. (1991) *Annu. Rev. Plant Physiol. Plant Mol. Biol. 42*, 145−188.
Sigal, E., Laughton, C. W., & Mulkins, M. A. (1994) *Ann. N.Y. Acad. Sci. 714*, 211−224.
Slappendel, S., Malmström, B. G., Petersson, L., Ehrenberg, A., Veldink, G. A., & Vliegenthart, J. F. (1982) *Biochem. Biophys. Res. Commun. 108*, 673−677.
Steczko, J., & Axelrod, B. (1992) *Biochem. Biophys. Res. Commun. 186*, 686−689.
Steczko, J., Muchmore, C. R., Smith, J. L., & Axelrod, B. (1990) *J. Biol. Chem. 265*, 11352−11534.
Steczko, J., Donoho, G. A., Dixon, J. E., Sugimoto, T., & Axelrod, B. (1991) *Protein Expr. Purif. 2*, 221−227.
Steczko, J., Donoho, G. P., Clemens, J. C., Dixon, J. E., & Axelrod, B. (1992) *Biochemistry 31*, 4053−4057.
Sträter, N., Klabunde, T., Tucker, P., Witzel, H., & Krebs, B. (1995) *Science 268*, 1489−1492.
Tate, M. W., Eikenberry, E. F., Barna, S. L., Wall, M. E., Lowrance, J. L., & Gruner, S. M. (1995) *J. Appl. Crystallogr. 28*, 196−206.
Tranberger, T. J., Franceschi, V. R., Hildebrand, D. F., & Grimes, H. D. (1991) *Plant Cell 3*, 973−987.
Van der Heijdt, L. M., Feiters, M. C., Navaratnam, S., Nolting, H. F., Hermes, C., Veldink, G. A., & Vliegenthart, J. F. (1992) *Eur. J. Biochem. 207*, 793−802.
Van der Heijdt, L. M., Schilstra, M. J., Feiters, M. C., Nolting, H. F., Hermes, C., Veldink, G. A., & Vliegenthart, J. F. (1995) *Eur. J. Biochem. 231*, 186−191.
van Tilbeurgh, H., Sarda, L., Verger, R., & Cambillau, C. (1992) *Nature (London) 359*, 159−162.
Whittaker, J. W., & Solomon, E. I. (1988) *J. Am. Chem. Soc. 110*, 5329−5339.
Xu, Z., Bernlohr, D. A., & Banaszak, L. J. (1993) *J. Biol. Chem. 268*, 7874−7884.
Yamamoto, S. (1992) *Biochim. Biophys. Acta 1128*, 117−131.
Young, A. C. M., Scapin, G., Kromminga, A., Patel, S. B., Veerkamp, J. H., & Sacchettini, J. C. (1994) *Structure 2*, 523−534.
Zhang, Y., Gebhard, M. S., & Solomon, E. I. (1991) *J. Am. Chem. Soc. 113*, 5162−5175.

BI960576U