The Crystal Structure of a Novel, Inactive, Lysine 49 PLA2 from Agkistrodon acutus Venom

AN ULTRAHIGH RESOLUTION, AB INITIO STRUCTURE DETERMINATION*

Qun Liu‡‡, Qingqiu Huang‡§, Maikun Teng§, Charles M. Weeks, Christian Jelsch**, Rongguang Zhang‡, and Liwen Niu‡§ §§

From the ‡Key Laboratory of Structural Biology, Chinese Academy of Sciences, the §Department of Molecular and Cell Biology, School of Life Sciences, University of Science and Technology of China, 96 Jinzhai Road, Hefei, Anhui 230026, China, the §Hauptman-Woodward Medical Research Institute, Buffalo, New York 14203-1196, the **Laboratoire de Cristallographie et Modélisation des Matériaux Minéraux et Biologiques, CNRS-Faculté des Sciences BP 239, 54506 Vandoeuvre-les-Nancy Cedex, France, and the §§Structural Biology Center, Argonne National Laboratory, Argonne, Illinois 60439

The crystal structure of acutohaemolysin, a lysine 49 phospholipase A2 protein with 1010 non-hydrogen protein atoms and 232 water molecules, has been determined ab initio using the program SnB at an ultrahigh resolution of 0.8 Å. The lack of catalytic activity appears to be related to the presence of Phe102, which prevents the access of substrate to the active site. The substitution of tryptophan for leucine at residue 10 interferes with dimer formation and may be responsible for the additional loss of hemolytic activity. The ultrahigh resolution of the experimental diffraction data permits alternative conformations to be modeled for disordered residues, many hydrogen atoms to be located, the proto-

Phospholipase A2 (PLA2, EC 3.1.1.4) is an enzyme that hydrolyzes the sn-2 fatty aryl bond of phospholipids to produce free fatty acid and lysophospholipids (1). Although PLA2s from different species share similar structures, they show diverse pharmacological and toxicological functions, including myotoxicity, anticoagulant activity, hemolyticity, and lipolyticity (2, 3). Accurate atomic resolution crystal structures are required to answer questions concerning the possible structural basis for these biological functions. According to conventional wisdom, the catalytic activity of PLA2s depends on the binding of a calcium ion and on the presence of residue Asp49 (4). (The conventional numbering system of Renetseder et al. (5) is used throughout this report.) The substitution of other residues for Asp49 eliminates the binding of calcium ion and is thought to abolish catalytic activity (6). However, the side-chain Nζ atom of lysine residues can play a role similar to that of a calcium ion, thereby enabling Lys49 PLA2s to display noticeable cata-

Received for publication, May 18, 2003, and in revised form, July 9, 2003

Published, JBC Papers in Press, July 19, 2003, DOI 10.1074/jbc.M305210200

* This work was supported by research grants from the National Natural Science Foundation of China (Grants 39870108, 30025012, and 30121001), the 973 Plan of the Ministry of Science and Technology of China (Grant G1999075603), the Chinese Academy of Sciences (Grants STZ-2-07 and STZ98-2-12) (to L. N. and M. T.), and by National Institutes of Health Grant EB-002057 (to C. M. W.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1MC2) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

§§ To whom correspondence should be addressed. Tel./Fax: 86-551-360-3046; E-mail: lwniu@ustc.edu.cn.

† The abbreviations used are: PLA2, phospholipase A2; r.m.s., root mean square.

‡‡ This work was supported by research grants from the National Natural Science Foundation of China, the Chinese Academy of Sciences, the Department of Molecular and Cell Biology, School of Life Sciences, University of Science and Technology of China, the Hauptman-Woodward Medical Research Institute, Buffalo, New York, the Laboratoire de Cristallographie et Modélisation des Matériaux Minéraux et Biologiques, CNRS-Faculté des Sciences, Vandoeuvre-les-Nancy Cedex, France, and the Structural Biology Center, Argonne National Laboratory, Argonne, Illinois 60439.

When ultrahigh resolution (e.g. better than 0.9 Å) crystallographic diffraction data are available, the deformation of the...
atomic electron density due to chemical bonding can be observed (16). Following a refinement using a spherical atom model, residual maps may reveal additional information about chemical bonding and electron transfer between atoms. At subatomic resolutions (\(d < 0.5 \, \text{Å}\)), the experimental molecular charge density can be precisely refined using a multipolar electron density model (17). Multipolar density can be decomposed into three contributions: the core spherical density, the valence spherical density, and the valence non-spherical density. Such information is important for deriving the overall electrostatic potential distribution in protein structures (18–22), for elucidating the mechanisms of biological processes involving electron transfers between catalytic atoms, and for determining the oxidation states of active metal ions in metalloproteins (23).

It is often asserted that charge density analyses are restricted to diffraction data from crystals of small molecules (24). Recently, however, it has been demonstrated that charge density refinement for biological macromolecules can also be carried out if the diffraction resolution is high enough and if the atomic thermal motion is sufficiently low. For example, the application of charge density refinement to the structure of a scorpion protein at 0.96-Å resolution showed visible bonding density between some main-chain atoms (21). The best example of charge density analysis applied to biological macromolecules, however, is provided by exhaustive refinement, at the ultrahigh resolution of 0.54 Å, of the valence electron distribution in crambin, a small protein with 46 residues (22). Herein we report the results of a charge density analysis applied to acutohaemolysin.

**EXPERIMENTAL PROCEDURES**

**Crystalization and Diffraction Data Collection—Acutohaemolysin** was purified and crystallized using previously reported procedures (25). Crystals with a size of \(0.8 \times 0.8 \times 1.0 \, \text{mm}\) were soaked in a 0.05 M sodium citrate solution (pH 5.6) containing 18% (v/v) isopropanol, 18% (v/v) polyethylene glycol 4000, and 5% (v/v) glycerol and then subjected to cryogenic flash freezing. Diffraction data were collected from cryoprotected crystals, under a nitrogen stream at a temperature of 100 K, at beamline 19-ID of the Advanced Photon Source (Argonne). The wavelength of the incident x-rays was set to 0.92 Å to meet the requirements for ultrahigh resolution data collection. A total of 600 frames was recorded from one crystal with different exposure time to cover all of the reciprocal space at both low and high resolution. Data processing was performed using the program HKL2000 (26), and the statistics of data collection and processing are summarized in Table I. The cryo-protected crystals of acutohaemolysin belong to space group C2 with unit cell dimensions \(a = 44.743 \, \text{Å}, b = 59.100 \, \text{Å}, c = 45.319 \, \text{Å}\), and \(\beta = 117.421^\circ\). These dimensions are slightly shorter than those observed under room temperature conditions. The agreement factor \(R_{\text{merge}}\) and the data completeness were 5.5% and 76.7%, respectively, for all resolutions up to 4.0 Å. The proper experimental deconvolution requires very accurate data between 20 and 1.0 Å. The \(R_{\text{merge}}\) values examined following completion of data collection and processing are summarized in Table I. The crystallographic \(R\) factor was calculated using the LEVY/EVAL programs in the SnB package (27), and then 107,100 triplet structure variants were generated using the 10,710 strongest \(E\) values of the resolution range 8–0.8 Å were calculated using the LEVY/EVAL programs in the SnB package (27), and then 107,100 triplet structure variants were generated using the 10,710 strongest \(E\) values. Next, trial structures consisting of 300 randomly positioned atoms (31, 32) were created, and each of them was subjected to 1,071 cycles of dual-space Shake-and-Bake refinement that included parameter shift optimization of the minimal function \(R_{\text{min}}\) and selection of 300 new peaks in each cycle (12). Three figures of merit (i.e., \(R_{\text{min}}\), a crystallographic \(R\) factor, and a correlation coefficient between \(E_{\text{obs}}\) and \(E_{\text{calc}}\) (28)) were calculated for each trial, and a histogram of the \(R_{\text{min}}\) values examined following completion of refinement for each of the 107,100 trials suggested a solution. After 140 cycles, a trace of the \(R_{\text{min}}\) values versus SnB cycle number showed a sudden drop (characteristic of a solution) resulting in a decrease of \(R_{\text{min}}\) from an initial value of 0.48 to a stable value of 0.439. Trial 16 was refined for 107 more dual space cycles while selecting the 964 strongest peaks in each cycle, and a final 107th cycles of Fourier refinement were performed while slowly increasing the number of reflections phased from 10,710 to 43,457 (i.e., all reflections with \(E > 0.75\)). The final \(R_{\text{min}}\), crystallographic \(R\) factor, and correlation coefficient values were 0.443, 0.32, and 0.66, respectively.

**Structural Model Building and Refinement**—To avoid lower ranking peaks that might be spurious, only the strongest 500 peaks were then input to unconstrained free-atom refinement using SHELXL (29). The highest 17 peaks were assigned to be sulfur atoms, and the remaining peaks were assigned as oxygen atoms. Refinement continued until there was no further decrease in the crystallographic \(R\) factor, which converged to a value of 0.25 after 100 cycles. Next, the main chain was built automatically by the program ARP/wARP (30) using diffusion coefficients derived from 20 and 1.0 Å, i.e., \(C_{\text{g}}(20) = 1.3\) and \(C_{\text{g}}(1.0) = 1.1\) Å. A total of 600 frames was used in the protein electron density map derived from the crystal structure. Hydrophilic residues at the molecular surface, it was very easy to recognize and build almost all side chains based on both \(2F_{\text{calc}} - F_{\text{obs}}\) and \(F_{\text{calc}} - F_{\text{obs}}\) maps. Hydrophilic residues with low electron density were assigned according to a cDNA clone from the same species (31). The crystallographic \(R\) factor of the model built in this way dropped to a value of 22% without including any solvent molecules.

Further refinement with the SHELXL package was performed with cross-validation using 5% of the observed reflections that were selected at random. Using data in the resolution range 10–1.2 Å, the model was first refined to convergence with an \(R\) factor of 22.2% and an \(R_{\text{free}}\) factor of 24.0%. Then, water molecules were added gradually. Using the instruction TRUN 1.3 0.02, the upper resolution limit was gradually expanded to 1.0 Å. Both the electron density map and the model were inspected carefully, and the model was adjusted to fit the density using the program O (32). Thereafter, all reflections were used, and, after addition of more water, the resolution was extended to 0.8 Å yielding \(R\) and \(R_{\text{free}}\) values of 18.91% and 20.08%, respectively.

Subsequent refinements were carried out by introducing anisotropic thermal parameters for all atoms except for the water molecules. This adjustment gave a sharp decrease in both \(R\) and \(R_{\text{free}}\) (13.98% and 14.96%, respectively). At this point, alternative conformations could be distinguished by using \(2F_{\text{calc}} - F_{\text{obs}}\) and \(F_{\text{calc}} - F_{\text{obs}}\) maps as well as clues in the diagnostic table output by SHELXL. For these modeled conformations, the initial \(B\) factors of all side-chain atoms were reset to an isotropic value of 10 Å², the occupancies of the main conformers were fixed at 0.66, and atoms with isotropic and anisotropic \(B\) factors were refined alternatively. Altogether, alternative conformations for nine residues were modeled. At this stage, the \(R\) value was 10.72%, and the \(R_{\text{free}}\) value was 12.63%. Distance restraints were applied during all refinements; the allowed standard deviations from the distance targets was of 0.02 Å for covalent bonds and 0.04 Å for bonds linked by two bonds. Near the end of refinement, the addition of riding hydrogen atoms made stereochemical-requirements further reduced \(R\) and \(R_{\text{free}}\) to 9.21% and 12.1%, respectively.

**Experimental Charge Density Refinement**—Charge density refinement for acutohaemolysin was carried out using the program MOPRO (33), an extension of the program MOLLY (16) for biological macromolecules. This program is able to refine large structures with the use of conjugate gradient iterations, and the sparse matrix principle was applied to the least squares method. A common difficulty in charge density refinement is the separation of the anisotropic mean square displacements of the atoms from the static molecular electron distribution. The proper experimental deconvolution requires very accurate diffraction data to ultrahigh resolution (e.g., 0.4–0.6 Å). The contribution of the bonding electrons is negligible to the diffraction signal beyond 0.8 Å (23), but scattering by the core electrons, located close to the atomic nuclei, extends to 0.5 Å or better depending on the size of the atomic thermal factors. Therefore, in small molecules, charge density analyses as well as accurate atomic positional and thermal parameters are obtained by a high order refinement where only reflections in the upper resolution ranges are weighted.

In the case of acutohaemolysin, the starting structure for charge density studies was the final model obtained from the SHELXL refinement. Only the atoms with a \(B\) factor lower than 9 Å² were allowed to vary in the high order MOPRO refinement, and the stereo-chemical restraints that were applied were similar to those used in SHELXL. The program MOPRO enables the application of reflection weights that depend on \(I_{\text{obs}}\), on its uncertainty \(\sigma_{I}\), and on the resolution \(d\). The \(W_{\text{min}}\) weighting scheme results in reflections at high resolution that are overweighted, as in Equation 1.

\[
W_{\text{min}} = \frac{2}{\sigma_{I}^2 + 0.001I_{\text{obs}}^{-1/2}d^{-2}} \quad (\text{Eq. 1})
\]
Structural refinement and validation

Resolution limits (Å) 20.0–0.80
R-factor (for data with F > 4σp) (%) 9.21
R-factor (for all data) (%) 10.52
Goodness of fit
Number of amino acid residues 122
Number of protein atoms 1010
Number of solvent water molecules 232
Number of isopropanol molecules 2
Uncertainties and r.m.s. deviations Full-matrix inversion
Bond distances (Å) 0.0183
Bond angles (°) 1.1789
ω dihedrals angles (°) 0.7973
Overall average G-factor
Anisotropy Mean Sigma Number of atoms
Protein atoms 0.443 0.153 1010
Solvent molecules 0.384 0.130 232
r.m.s. Z scores (should be close to 1.0)
Bond lengths 0.832
Bond angles 1.214
Peptide bond ω angle restraints 1.012
Side-chain planarity 1.208
Improper dihedral distribution 0.874
Inside/outside distribution 1.088
Volumes of all atoms
Ramachandran plot (for non-Gly and non-Pro)
Residues in most favored regions (%) 91.6
Residues in additional allowed regions (%) 8.4

* Completeness is the ratio of the number of observed reflections to the number of possible reflections.
* The values for the data in the highest resolution shell (0.83–0.80 Å) are shown in parentheses.
* Goodness of fit = \[ \sum_{hkl} (F_{obs}(hkl) - F_{calc}(hkl))^2 / \sum_{hkl} F_{obs}^2(hkl) \] (n − p) where n is the number of reflections, p is the number of refined parameters, and F_{obs} and F_{calc} are the observed and calculated amplitudes, respectively, of the structure factors for reflection hkl, and w is a weighting scheme.
* The G-factor and the Ramachandran plot were generated by program PROCHECK (34).
* All r.m.s. Z scores are from WHATCHECK (35) with the exception of the volume, which was calculated by PROVE (36).

Using this weighting scheme, the charge density could be partially retrieved for regions of the protein with low thermal motion. When a spherical atom model was used, a residual map revealed the non-modulated deformation of the electron density after high order refinement.

Coordinates and anisotropic displacement parameters were not refined simultaneously but, rather, in successive refinement cycles. The Scale factor and solvent correction parameters were regularly refined using a different weighting scheme,

\[ W_{ij} = (2σ^2_i + 0.001σ^2_j)^{1/2} \] (Eq. 2)

which was independent of the resolution. Refinement using all diffraction data and both weighting schemes (until convergence of the positional and thermal parameters was achieved) resulted in a final R-factor of 15%. This relatively high value was due to overweighting of the upper resolution reflections during the refinement.

RESULTS AND DISCUSSION

Model Quality and Stereochemical Parameters—The structure of acutohaemolysin contains 1010 non-hydrogen protein atoms, 2 isopropanol molecules, and 232 water molecules, and it was refined to a crystallographic R factor of 9.21% (for data with F > 4σp) or 10.52% (for all data). Fourier maps clearly revealed the atomicity of the electron density, and non-hydrogen protein atoms (i.e. C, N, and O) could be easily differenti-ated, because the heavier atoms have a larger spherical radius of density. Moreover, peaks accounting for 428 hydrogen atoms were observed on difference Fourier maps, thereby indicating that the refined model is of high quality.

The stereochemical parameters of the model were analyzed by the use of programs PROCHECK (34), WHATCHECK (35), and PROVE (36), and the results are shown in Table I. The main-chain dihedral angles of all non-glycine and non-proline residues lie in the preferred and allowed regions (91.6% and 8.4%, respectively) of the Ramachandran diagram (37), and the standard deviations of the bond lengths and bond angles (0.018 Å and 2.3°, respectively) are reasonable. The overall average G factor of 0.03 computed by PROCHECK is very close to its theoretic value of zero, and this is another indication of the excellent overall quality of the model. All root-mean-square (r.m.s.) Z scores are close to 1.0 except for those of the bond angles and side-chain planarity. The poor r.m.s. Z scores for these two quantities might be the result of the relatively low quality of the standard parameters used by WHATCHECK.

In addition, the standard deviations of the stereochemical parameters were also estimated directly from the experimental data by inversion of the full normal matrix (38), and the pa-
The amino acid residues of acutohaemolysin and four other Lys\(^{49}\) PLA\(_{2}\)s. PDB accession codes are given in parentheses. Nine residues (Tyr\(^{25}\), Asn\(^{27}\), Leu\(^{39}\), Lys\(^{76}\), Gln\(^{89}\), Met\(^{92}\), Phe\(^{102}\), Phe\(^{117}\), and Glu\(^{132}\)) plus one insertion (Ser\(^{131}\)) for which acutohaemolysin differs from all other sequences are shown in red. This alignment was performed using the program CLUSTALW (39).

Calcium-binding Loop—The calcium-binding loop, illustrated in Fig. 4, is a highly conserved feature of the PLA\(_{2}\) enzymes, including the Lys\(^{49}\) variants that do not have a calcium ion present in this region. The existence of a consensus sequence Tyr\(^{25}\)-Gly-Cys-Asn-Cys-Gly-X-X-Arg-Gly\(^{35}\) implies that this region has a similar function in all members of the family. Compared with the neighboring region containing Cys\(^{27}\)-Asn-Cys-Gly\(^{30}\), the electron density of the peptide segment Val\(^{31}\)-Gly-Gly\(^{35}\) appears to be quite poor, indicating that these residues have relatively high thermal motion. In fact, the average temperature factor for Val\(^{31}\)-Gly-Gly\(^{33}\) is 16 Å\(^2\), double the value for the segment Cys\(^{27}\)-Asn-Cys-Gly\(^{30}\). Fig. 4 also shows that a strong electron density peak marks the position of a water molecule in the calcium-binding loop. This structurally well ordered water molecule (named WAT2010) forms two hydrogen bond and four electrostatic interactions in the range 2.76–3.38 Å with the carbonyl oxygen atoms of four residues in the loop as well as the N\(_{\alpha}\) atoms of Lys\(^{49}\) and Lys\(^{122}\) (see Table II for the geometric parameters of hydrogen bonds). Therefore, these six atoms, including the carbonyl oxygen atoms of Val\(^{31}\) and Gly\(^{33}\), stabilize the water molecule, and it is reasonable to infer that the segment Val\(^{31}\)-Gly-Gly\(^{33}\) has an impact on the behavior of the water molecule. The electron density accounting for the two hydrogen atoms of this water molecule could be clearly seen on the maps.

The calcium-binding loop is one of most conservative regions in PLA\(_{2}\)s family. For the enzymatic activity of Asp\(^{49}\) PLA\(_{2}\)s, Scott et al. (41) proposed that the backbone amide N-H of residue Gly\(^{30}\) plays a role in stabilizing the oxyanion of the substrate tetrahedral intermediate (41). On the other hand, for the enzymatic activity of Lys\(^{49}\) PLA\(_{2}\)s, Lee et al. (9) demonstrated that the hyperpolarization of the peptide bond between residues 29 and 30, induced by the hydrogen interaction between the main-chain carbonyl of Cys\(^{29}\) and Lys\(^{122}\) via a water molecule, may increase the affinity for the enzyme to the reaction product (fatty acid). In acutohaemolysin, the poor density appearance and high average temperature factor of peptide segment Val\(^{31}\)-Gly-Gly\(^{33}\) indicate relatively high dynamics.
The hydrogen interactions between carbonyl oxygen atoms of Val31 and Gly33 to the structurally well ordered water molecule, which has electrostatic interactions with the N\textsubscript{ε}/H9256 atoms of Lys49 and Lys122, suggest coordination between Val31-Gly-Gly33, the water molecule (WAT2000), and the peptide bond of Cys29-Gly30. Therefore, it could be postulated that this loop may be of importance in stabilizing the substrate tetrahedral intermediate and in releasing the reaction products.

**Catalytic Site**—PLA\textsubscript{2}s use the conserved catalytic triad consisting of His\textsuperscript{48}, Tyr\textsuperscript{52}, and Asp\textsuperscript{99} (see Figs. 2 and 5) to hydrolyze substrates into free fatty acids and lysophospholipids (1). It is believed that His\textsuperscript{48} and Asp\textsuperscript{99}, together with a nearby structurally conserved water molecule, are involved in the nucleophilic attack of the water molecule at the sn-2 position of the phospholipid substrate. The success of this process requires the stabilization of a tetrahedral transition-state intermediate (41). For Asp\textsuperscript{49} PLA\textsubscript{2}s containing calcium ion, the essential factors stabilizing the tetrahedral intermediate are the interactions of calcium ion with the catalytic water molecule, with the carboxyl oxygen atom of Asp\textsuperscript{49}, and with the main-chain carbonyl oxygen atoms of the calcium-binding loop. Due to the substitution of lysine for Asp\textsuperscript{49}, Lys\textsuperscript{49} PLA\textsubscript{2}s are unable to bind calcium, but some of them have, nevertheless, been demonstrated to have at least limited catalytic activity (7). Furthermore, the crystal structures of two Lys\textsuperscript{49} PLA\textsubscript{2}s complexed with their substrates reveal that these proteins have the ability to bind substrates to their active sites (8, 42). It has been proposed that Lys\textsuperscript{49} PLA\textsubscript{2}s possess interrupted catalytic activity, because the residue Lys\textsuperscript{122} hyperpolarizes the peptide bond between residues Cys\textsuperscript{29} and Gly\textsuperscript{30}, resulting in failure to release the reaction product (9).

Although all of the required catalytic apparatus (including His\textsuperscript{48}, Tyr\textsuperscript{52}, Tyr\textsuperscript{73}, and Asp\textsuperscript{99}, and the catalytic water molecule WAT2010) is conserved in the acutohaemolysin structure, no measurable activity is exhibited in standard assays (25). When the catalytic triad of acutohaemolysin is superimposed with the corresponding residues of four structurally homologous Lys\textsuperscript{49} PLA\textsubscript{2}s, at least one of which possesses catalytic activity (9), it is apparent that all the triads have nearly identical conformations and are anchored tightly to the scaffold composed of two \textalpha helices (Fig. 5). Therefore, it is clear that the local conformation of the triad in acutohaemolysin is not responsible for the lack of catalytic ability, and the structural basis for this functional deficiency needs to be investigated further.

By sequence alignment (see Fig. 1), Phe\textsuperscript{102} is detected as a unique residue that exists only in acutohaemolysin, whereas other Lys\textsuperscript{49} PLA\textsubscript{2}s possess a short side-chain residue (e.g. Val) at this site. As shown in Fig. 5, by comparison of the local conformation near the catalytic triad of acutohaemolysin with that of piratoxin II, a catalytically active Lys\textsuperscript{49} PLA\textsubscript{2} complexed with a 13-carbon fatty acid molecule from Bothrops pirajai venom (9), it is apparent that the benzene ring of Phe\textsuperscript{102} may interfere with the binding of potential substrates. If this fatty acid molecule were placed at the same position in the acutohaemolysin structure, there would be an impossibly short contact distance (about 1.3 Å) between it and the benzene ring of Phe\textsuperscript{102}. Therefore, the docking of a similar or larger fatty acid molecule into the hydrophobic channel of acutohaemolysin seems to be prohibited. Because all other structural components in or near the active site are the same, and because the catalytic process requires a space to allow substrate entry and product release, it can be postulated that Phe\textsuperscript{102} is responsible for the loss of catalytic activity. Clearly, this hypothesis should...
is the hydrogen-bond donor for the Gly31 water molecule. The carbonyls of Cys27 and Asn28 also receive hydrogens from Lys49 and Lys122. The structurally ordered (and highly conserved) water molecule 2010 has multiple interactions with neighboring atoms, and peaks on the $F_{\text{obs}} - F_{\text{calc}}$ map indicate that this water is the hydrogen-bond donor for the Gly31 and Gly33 carbonyl oxygen atoms and receives hydrogens from Lys49 and Lys102. The carbonyls of Cys27 and Asn28 also have electrostatic interactions with the water molecule.

be validated experimentally by the site-directed mutation of F102V in acutohaemolysin.

The acutohaemolysin structure provides a high resolution look at the hydrogen-bond network in the vicinity of the active site. Peaks of electron density corresponding to the hydrogen atoms riding on the C$\delta_3$, C$\epsilon_1$, and N$\epsilon_2$ atoms of His$^{48}$ as well as the O$\delta$ atom of Tyr$^{52}$ have been readily observed (Fig. 6). The two hydrogen-bond distances involving the O$\delta$ atom of Asp$^{99}$ are 2.59 and 2.77 Å, indicating the strength of these interactions. Because the dissociation $pK_a$ value of the N$\epsilon_2$H$^+$ group in a free histidine residue is about 6.0, the proton could easily bind to or release from the N$\epsilon_2$ atom under neutral conditions. Acutohaemolysin was crystallized at pH 5.6 (a little lower than the dissociation $pK_a$ value) so that the environment was appropriate for the protonation of both the N$\delta_1$ and the N$\epsilon_2$ atoms. However, it should be remembered that the $pK_a$ values of amino acid residues located at the surface of proteins are known to be highly dependent on the local chemical environment (43). The protonation of the N$\epsilon_2$ atom is favored by its role as hydrogen-bond donor to the negatively charged carboxylate of Asp$^{99}$. On the other hand, N$\delta_1$ seems to be non-protonated and is the receptor in the hydrogen bond involving WAT2010.

Table II
Parameters for the hydrogen bonds illustrated in Figs. 4 and 6

| Atoms          | Distance X → Y | Distance X → H | Distance H → Y | Angle X → H → Y | Fig. |
|---------------|----------------|----------------|----------------|-----------------|-----|
| H–2010W       | 2.80           | 1.98           | 0.99           | 138.6           | 4   |
| H–2010W       | 2.76           | 1.84           | 1.15           | 133.4           | 4   |
| H–99OD1       | 2.77           | 1.74           | 1.05           | 165.8           | 6   |
| H–99OD1       | 2.59           | 1.82           | 0.87           | 146.4           | 6   |
| Average       | 2.73           | 1.85           | 1.02           | 146.1           |     |

Fig. 5. A stereo view of the superimposed catalytic triads of acutohaemolysin (red) and four other Lys$^{49}$ PLA2s (i.e. 1CLP, blue; 1QLL, purple; 1PPA, green; and 1GOD, cyan) illustrates structural conservation in this region. A fatty acid molecule co-crystalized with piratoxin II (1QLL) suggests that the structural basis for the low catalytic activity of acutohaemolysin is interference with substrate approach to the triad by the side chain of Phe$^{102}$.

either dynamic or static. Dynamic disorder can be due to thermal vibrations or to rotations around single bonds, whereas static disorder can result from slight displacements of parts of protein molecules in the crystal lattice or local differences in solvation. Relatively low resolution x-ray diffraction experiments give an average positional distribution for disordered atoms or residues. At ultrahigh resolution, however, it is possible to decipher the disorder and to view alternative conformations. In the case of acutohaemolysin, such conformations have been modeled for several residues, including Trp$^{10}$, Val$^{18}$, Leu$^{24}$, Arg$^{43}$, Met$^{92}$, Arg$^{107}$, Glu$^{108}$, Asp$^{111}$, and Ser$^{116}$. All of these residues are located at the molecular surface, and some of these cases of disorder may have possible functional significance.

In some Lys$^{49}$ PLA2s, the N-terminal $\alpha$ helix and an antiparallel $\beta$ sheet in the region of residues 74–85 (noted as the $\beta$ wing in Fig. 2) play a role in the formation of a molecular dimer that, in turn, has been considered essential for the presence of hemolytic activity (23, 42). On the other hand, acutohaemolysin exists as a monomer in solution, and it has no hemolytic activity. Whereas other Lys$^{49}$ PLA2s typically possess Leu or Phe residues at position 10 at the end of the N-terminal $\alpha$ helix, this residue is a tryptophan in acutohaemolysin (Fig. 1). The large indole group of Trp$^{10}$ shows two preferred conformations differing by a rotation of about 180° around the C$\beta$–C$\gamma$ bond (Fig.
7), and a series of other conformers with varying torsion angles about this bond may, in fact, exist. The electron density in this region is poor due to the disorder as well as the large atomic thermal motion (B(eq) = 20–40 Å²). Nevertheless, it is clear that the Trp¹⁰ side chain, which occupies a volume of about 300 Å³, extends into the surrounding solvent, and its flexibility may prevent molecular aggregation and dimer formation from occurring. Thus, the presence of Trp¹⁰ may be responsible for the absence of hemolytic activity in acutohaemolysin.

Although acutohaemolysin exists as a monomer in solution, it packs in the crystal as a dimer composed of two molecules related by a crystallographic 2-fold axis parallel to the b axis of the unit cell. This dimer is distinctly different from the hemolytically active dimer observed in other Lys⁴⁹ PLA₂s. One of the interactions that appears to stabilize the crystallographic dimer is a pair of salt bridges between the disordered Arg¹⁰⁷ and Glu¹⁰⁸ residues. The two positively charged Nζ atoms of Arg¹⁰⁷ in one monomer are hydrogen-bonded to the two negatively charged Oe atoms of Glu¹⁰⁸ in the other (symmetry-related) half of the dimer with the shortest contact distances being 2.75 and 2.98 Å. In addition, the Arg-Glu salt-bridge pairs are further stabilized by longer range electrostatic interactions with N...O distances shorter than 4 Å (44). The definite contribution of salt bridges and longer range electrostatic interactions in the crystal packing formation could be supported by the analysis of several electrostatic variants of cutinase surface residues in different crystal forms (45).

Solvent Molecules—X-ray diffraction experiments performed at liquid nitrogen temperatures permit more solvent molecules to be located than would be possible using data measured at higher temperature. In all, 210 fully occupied water molecules, 22 half-occupied water molecules, and 2 isopropanol molecules were located on the acutohaemolysin difference electron-density maps. All of these solvent molecules were visible as electron density peaks contoured at 4σ, and they possess equivalent isotropic thermal parameters (B(eq)) lower than 40 Å². Each water molecule forms at least one hydrogen bond with adjacent protein atoms or other solvent atoms.

The thermal anisotropy of both protein and solvent atoms in acutohaemolysin was analyzed statistically (see Table I) using PARVATI, a program that has been implemented to study this property for structures determined at a resolution higher than 1.4 Å (46). Anisotropy is defined as the square axial ratio of an ellipsoid and can be determined from the anisotropic displacement parameters. For the protein atoms, the distribution of anisotropy values has a typical Gaussian shape, and the average anisotropy of 0.44 is almost the same as the expected value of 0.45 statistically computed from 67 crystal structures with a resolution higher than 1.4 Å. On the other hand, the anisotropy distribution for the solvent molecules has two peaks. One of these peaks (at ~0.44) matches the results for the protein atoms, but the other (at ~0.25) indicates a larger anisotropy. Such a bimodal distribution implies that the water molecules in the refined structure of acutohaemolysin can be divided into two hydration shells. Water molecules in the inner shell (anisotropy peak at 0.44) have thermal parameters similar to those of the protein atoms and could be considered to be an integral part of the protein molecule. Inner-shell waters exhibit an average hydrogen-bond length of 2.70 Å, whereas the corresponding value for waters in the outer shell is 3.13 Å. These results are similar to observations made for the hydration shells around other protein molecules (47).

Hydrogen Atoms—Ultrahigh resolution diffraction data are of good quality if a high percentage of the hydrogen atoms can be recognized on difference electron density maps. In the case of acutohaemolysin, a total of 428 hydrogens, including most main-chain and many side-chain atoms (about 38% of the possible 1132 hydrogen atoms in the structure), could be discerned on F(obs) – F(calc) difference maps contoured at 2.1σ (1σ = 0.09/e Å²). All of these peaks had a distance of 1.0 ± 0.3 Å to the nearest non-hydrogen atom. There were 89 and 88 peaks (72% of all possible) observed for the electron density of the hydrogen.
atoms riding on the main-chain C\textsubscript{\text{a}} and N atoms, respectively. Although a relatively low percentage (28\%) of the side-chain hydrogen atoms was observed, the peaks of electron density for the hydrogen atoms of some residues in hydrophobic regions could be found easily. The geometry of the hydrogen bonds listed in Table II gives an indication of the good quality of the hydrogen positions.

Experimental Deformation Density Analysis—The ultrahigh resolution (0.8 Å) and high quality ($R_{\text{merge}} = 5.5\%$) of the acutohaemolysin diffraction data as well as the low experimental temperature factors estimated from a Wilson plot (5.6 Å\textsuperscript{2}) and observed as the average for the refined model parameters (6.5 Å\textsuperscript{2}) support the application of charge density analysis. For some hydrophobic residues in the core region of the structure, the atomic temperature factors are even below 4.0 Å\textsuperscript{2}, low enough to allow partial differentiation of the deformation density features from the thermal movement. Following MOPRO refinement of the coordinates and thermal parameters of the protein atoms with $B_{\text{eq}} < 9$ Å\textsuperscript{2}, a residual map displayed unambiguously the covalent bonding densities in the regions with low $B$-factors. These regions include the catalytic residue His\textsuperscript{49} (Fig. 8A) and residue Tyr\textsuperscript{64} (Fig. 8C). Planes through several peptide bonds show residual peaks corresponding to valence electrons (e.g. peptide bond Phe\textsuperscript{92}-Ala\textsuperscript{93} shown in Fig. 8B). Some bonding densities are even visible for some residues in the difference map after the refinement with the SHELXL program.

Disulfide bonds are formed by the oxidation of two thiol groups of cysteine residues. The formation of such intra-molecular or inter-molecular bonds is important to fold and stabilize functional structures. Besides, disulfide bonds are sensitive to temperature or reducing environments and may easily dissociate or reform by increasing or decreasing the temperature. S–S bonds are weak covalent interactions, and the corresponding bonding density peaks appear lower compared with other bonds (C–S, C–C, and C–N, etc.) in a charge density analysis of L-cystine performed at 0.45-Å resolution (44). After high order refinement, the map also shows residual electron density between the two sulfur atoms that form the Cys\textsuperscript{43}-Cys\textsuperscript{95} disulfide bond (Fig. 8D).

The presence of seven disulfide bridges in the protein (Fig. 2) and the relatively low amount of bulk solvent volume in the unit cell contribute to the overall low thermal motion in the PLA2 structure. Because a significant number of the protein atoms (39\%) have low $B$ factors in the range of 4–6 Å\textsuperscript{2}, the crystals reported here seem to have the potential to diffract to resolutions even higher than 0.8 Å (22), if the crystallization conditions and the data collection techniques are optimized. A

![Fig. 8](image-url)  
Residual maps that were obtained after high order refinement and contoured at $\pm 0.05$ e/Å\textsuperscript{3} reveal peaks of electron density between the bonded atoms. The solid blue lines show the electron density accumulation, and the dashed red lines show the electron density depletion (i.e. blue, positive density; red, negative density). These residual maps illustrate the catalytic residue His\textsuperscript{49} (A), the dipeptide Phe\textsuperscript{92}-Ala\textsuperscript{93} (B), residue Tyr\textsuperscript{64} (C), and the disulfide bond between Cys\textsuperscript{43} and Cys\textsuperscript{95} (D).
charge density analysis performed using such data would provide additional information about the protonation and electronic states of the residues in acutohaemolysin.

Acknowledgments—We are grateful to Dr. Thomas R. Schneider (Department of Structural Chemistry, Göttingen University, Göttingen, Germany), who kindly provided us with the new version of SHELX97 to carry out the macromolecular refinement, to Dr. Quan Hao (MacCHESS, Cornell University, New York) for the help in the preparation of the manuscript, and to Drs. Jane Griffin and Robert Blessing (Hauptman-Woodward Institute) for critical reading of the manuscript.

REFERENCES

1. Ownby, C. L., Selistre de Araujo, H. S., White, S. P., and Fletcher, J. E. (1999) *Toxicol* 37, 411–445
2. Arni, R. K., Fontes, M. R., Barberato, C., Gutierrez, J. M., Diaz, C., and Ward, R. J. (1999) *Arch. Biochem. Biophys.* 366, 177–182
3. Ward, R. J., De Azevedo, W. F., and Arni, R. K. (1998) *Toxicol* 36, 1623–1633
4. Francis, B., Gutierrez, J. M., Leomonte, B., and Kaiser, I. I. (1991) *Arch. Biochem. Biophys.* 284, 352–359
5. Renetseder, R., Brunie, S., Dijkstra, B. W., Drenth, J., and Sigler, P. B. (1985) *J. Biol. Chem.* 260, 11627–11636
6. Gutierrez, J. M., and Leomonte, B. (1995) *Toxicol* 33, 1405–1424
7. Pedersen, J. Z., Leomonte, B., Massoud, R., Gubenek, F., Gutierrez, J. M., and Rufini, S. (1995) *Biochemistry* 34, 4670–4675
8. Shimohigashi, Y., Tani, A., Matsumoto, H., Nakashima, K., Yamaguchi, Y., Oda, N., Takao, Y., Kamiya, H., Kishino, J., Arita, H., and Ohno, M. (1995) *J. Biochem.* 118, 1037–1044
9. Lee, W. H., Da Silva Giotte, M. T., Marangoni, S., Toyama, M. H., Polikarpov, I., and Garnatt, H. C. (2001) *Biochemistry* 40, 28–36
10. Holland, D. R., Clancy, L. L., Muchmore, S. W., Ryde, T. J., Einspahr, H. M., Finzel, B. C., Heinrikson, R. L., and Watenpaugh, K. D. (1999) *J. Biol. Chem.* 265, 17649–17656
11. Ward, R. J., Rodrigues, A. R., Ruggiero, N. J., Arni, R. K., and Casari, G. (1998) *Protein Eng.* 11, 285–284
12. Weeks, C. M., Deftera, G. T., Hauptman, H. A., Thuman, P., and Miller, R. (1994) *Acta Crystallogr.* 50, 210–220
13. Miller, R., Gallo, S. M., Khalak, H. G., and Weeks, C. M. (1994) *J. Appl. Crystallogr.* 27, 613–621
14. Weeks, C. M., and Miller, R. (1999) *J. Appl. Crystallogr.* 32, 120–124
15. Kuhn, P., Knapp, M., Sortis, M. S., Ganshaw, G., Thoene, M., and Bott, R. (1998) *Biochemistry* 37, 13446–13452
16. Jelsch, C., Pichen-Pesme, V., Leomonte, C., and Aubry A. (1998) *Acta Crystallogr.* 54, 1306–1318
17. Hansen, N. K., and Coppens, P. (1978) *Acta Crystallogr.* A 34, 909–921
18. Coppens, P. (1997) *X-ray Charge Densities and Chemical Bonding* Oxford University Press, Oxford, pp. 233–286
19. Stewart, R. F., and Craven, B. M. (1993) *Biophys. J.* 65, 998–1005
20. Spackman, M. A., and Byrom, P. G. (1996) *Acta Crystallogr.* 52, 1023–1035
21. Houssset, D., Benabicha, F., Pichen-Pesme, V., Jelsch, C., Maierbaumer, A., David, S., Fontecilla-Camps, J. C., and Leomonte, C. (2000) *Acta Crystallogr.* 56, 151–160
22. Jelsch, C., Tester, M. M., Lamzin, V., Pichen-Pesme, V., Blessing, R. H., and Leomonte, C. (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97, 2171–2176
23. Leomonte, C., Blessing, R. H., Coppens, P., and Tobard, A. (1986) *J. Am. Chem. Soc.* 108, 6942–6949
24. Coppens, P. (1998) *Acta Crystallogr.* A 54, 779–788
25. Huang, Q. Q., Zhu, X. Y., Li, N., Deng, W. H., Chen, T. B., Rao, P. F., Teng, M. K., and Niu, L. W. (2000) *Acta Crystallogr.* 56, 907–911
26. Owinoziwski, Z., and Minor, W. (1997) *Methods Enzymol.* 276, 307–326
27. Blessing, R. H., Gun, D. Y., and Langs, D. A. (1996) *Acta Crystallogr.* 52, 257–266
28. Fujinaga, M., and Read, R. J. (1997) *J. Appl. Crystallogr.* 20, 517–521
29. Sheldrick, G. M. (1997) *The SHELX-97 Manual*, Göttingen University
30. Perrakis, A., Sixma, T. K., Wilson, K. S., and Lamzin, V. S. (1997) *Acta Crystallogr.* 53, 448–455
31. Fan, C. Y., Qian, Y. C., Yang, S. L., and Geng, Y. (1999) *Genetic Analysis* 15, 15–18
32. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr.* A 47, 109–110
33. Guillot, B., Viry, L., Guillot, R., Leomonte, C., and Jelsch, C. (2001) *J. Appl. Crystallogr.* 34, 214–223
34. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr.* 26, 283–291
35. Hooff, R. W. W., Vriend, G., Sander, C., and Abola, E. E. (1996) *Nature* 379, 272
36. Pontius, J., Richelle, J., and Wodak, S. J. (1996) *Curr. Opin. Struct. Biol.* 6, 307–326
37. Ramachandran, G. N., Ramakrishnan, C., and Sasisekharan, V. (1963) *J. Mol. Biol.* 7, 95–96
38. Dauter, Z., Lamzin V. S., and Wilson, K. S. (1997) *Curr. Opin. Struct. Biol.* 7, 681–688
39. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* 22, 4673–4680
40. Esnouf, R. M. (1997) *J. Mol. Graph. Model.* 15, 112–133
41. Scott, D. L., White, S. P., Otwinowski, Z., Yuan, W., Gelb, M. H., and Sigler, P. B. (1999) *Science* 286, 99–103
42. Yamaguchi, Y., Shimohigashi, Y., Chiwata, T., Tani, A., Chijiwa, T., Lomonte, B., and Ohno, M. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94, 6942–6949
43. Harris, T. K., and Turner, G. J. (2002) *Curr. Opin. Struct. Biol.* 12, 121–136
44. Jelsch, C., Tester, M. M., Lamzin, V., Pichen-Pesme, V., Blessing, R. H., and Leomonte, C. (2000) *Proc. Natl. Acad. Sci. U.S.A.* 97, 2171–2176
45. Kumar, S., and Nussinov, R. (1999) *J. Mol. Biol.* 293, 1241–1255
46. Merritt, E. (1999) *Acta Crystallogr.* 55, 1109–1117
47. Svergun, D. I., Richard, S., Koch, M. H., Sayers, Z., Koprin, S., and Zacek, G. (1998) *Proc. Natl. Acad. Sci. U.S.A.* 95, 2267–2272