Catalysis at the Interface: The Anatomy of a Conformational Change in a Triglyceride Lipase†,‡

Urszula Derewenda,§ Andrzej M. Brzozowski,∥,⊥ David M. Lawson,∥ and Zygmunt S. Derewenda*,§

MRC Group in Protein Structure and Function, Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7, and Department of Chemistry, University of York, Heslington, York, YO1 5DD U.K.

Received July 19, 1991; Revised Manuscript Received November 19, 1991

ABSTRACT: The crystal structure of an extracellular triglyceride lipase (from a fungus *Rhizomucor miehei*) inhibited irreversibly by diethyl *p*-nitrophenyl phosphate (E600) was solved by X-ray crystallographic methods and refined to a resolution of 2.65 Å. The crystals are isomorphous with those of *n*-hexylphosphonate ethyl ester/lipase complex [Brzozowski, A. M., Derewenda, U., Derewenda, Z. S., Dodson, G. G., Lawson, D. M., Turkenburg, J. P., Bjorkling, F., Huge-Jensen, B., Patkar, S. A., & Thim, L. (1991) *Nature 351*, 491–494], where the conformational change was originally observed. The higher resolution of the present study allowed for a detailed analysis of the stereochemistry of the change observed in the inhibited enzyme. The movement of a 15 amino acid long "lid" (residues 82–96) is a hinge-type rigid-body motion which transports some of the atoms of a short α-helix (residues 85–91) by over 12 Å. There are two hinge regions (residues 83–84 and 91–95) within which pronounced transitions of secondary structure between α and β conformations are caused by dramatic changes of specific conformational dihedral angles (ϕ and ψ). As a result of this change a hydrophobic area of ca. 800 Å^2 (8% of the total molecule surface) becomes exposed. Other triglyceride lipases are also known to have "lids" similar to the one observed in the *R. miehei* enzyme, and it is possible that the general stereochemistry of lipase activation at the oil–water interfaces inferred from the present X-ray study is likely to apply to the entire family of lipases.

Lipolytic enzymes have recently become a subject of intense structural studies. Apart from the molecular basis of their catalytic function, the activation of these enzymes at an oil–water interface is of particular interest. This is a unique property which evolved in response to the problem of handling insoluble substrates, and it distinguishes true lipolytic proteins from esterases which act upon soluble esters. Triglyceride lipases are particularly sensitive to interfacial activation (Sarda & Desnuelle, 1958), but other lipolytic enzymes, most notably phospholipase A2, also share this property. Various hypotheses have been suggested in the past with regard to the molecular mechanism of interfacial activation. In the case of phospholipase A2 recent X-ray crystallographic investigations have shown no conformational change associated with catalysis at the interface; instead, it is suggested that an isostructural and isoenergetic transfer of phospholipid occurs from the substrate aggregate (micelles, monolayers, vesicles, or membranes) to the catalytic site through a well-defined, rigid hydrophobic channel (Scott et al., 1990a,b; White et al., 1990).

To this day crystal structures of three neutral triglyceride lipases have been elucidated by X-ray crystallographic techniques: an extracellular enzyme purified from a fungus *Rhizomucor miehei* (Brady et al., 1990), human pancreatic lipase (Winkler et al., 1990), and most recently another fungal enzyme from *Geotrychum candidum* (Schrag et al., 1991). In spite of the evolutionary gaps which separate the fungal enzymes from the human digestive lipase certain intriguing regularities are observed. All three enzymes contain mixed (though predominantly parallel) β-pleated sheets, although the connectivities vary. GcL[1] and the larger, catalytic N-terminal domain of hPL share the same β-sheet topology found also in wheat serine carboxypeptidase II (Liao & Remigton, 1990), AChE from *Torpedo californica* (Sussman et al., 1991), dienelactone hydrolase (Pathak & Ollis, 1990), and holoalkane dehalogenase (Franken et al., 1991). RmL, the smallest of these enzymes, also contains a largely parallel sheet, but the connectivities between the strands are slightly different (Brady et al., 1990). All three lipases are serine hydrolases (interestingly, so are AChE and wheat serine carboxypeptidase, while dienelactone hydrolase contains a cysteine triad). In RmL and hPL the catalytic centers are made up of identical constellations of residues, namely, Ser/His/Asp triads originally described in chymotrypsin (Blow et al., 1969), but GcL is unusual in that a Glu replaces Asp in a similar triad. A similar triad is found in AChE (Sussman et al., 1991). In all of these enzymes the nucleophilic residue (Ser or Cys) is located in a sharp turn between one of the β-strands and a buried α-helix. Derewenda and Derewenda (1991) postulate that this structural motif (β-ϵSer-α) in which the nucleophilic Ser adopts an unusual secondary structure may be common to all serine-dependent lipases and esterases.

In none of the three lipases are the active centers exposed to solvent (as is the case in proteinases and acetylcholinesterase); they are instead concealed beneath short amphipathic helices. This led to a hypothesis (Brady et al., 1990;

† This research was financed by a grant to the MRC of Canada Group in Protein Structure and Function (U.D. and Z.S.D.) and NOVO-Nordisk (Copenhagen, Denmark) Research Laboratories (Z.S.D., A.M.B., and D.M.L.).

‡ Crystallographic coordinates have been submitted to the Brookhaven Protein Data Bank.

* To whom correspondence should be addressed.

§ University of Alberta.

∥ University of York.

⊥ Permanent address: Department of Crystallography, Institute of Chemistry, University of Lodz, 91 416 Lodz, Poland.

[1] Abbreviations: GcL, *Geotrychum candidum* lipase; RmL, *Rhizomucor miehei* lipase; hPL, human pancreatic lipase; TIM, triosephosphate isomerase; AChE, acetylcholinesterase; E600, diethyl *p*-nitrophenyl phosphate; rms, root mean square; FFT, fast Fourier transform.

0006-2960/92/0431-1532$03.00/0

Table I: Details of the Xentronics Data Collection

resolution limit (Å)	no. of reflections possible	no. of reflections collected	average redundancy	no. of singly measured reflections (%)	$\langle I/\sigma(I)\rangle$	no. of reflections $I > 2\sigma(I)$ (%)	R_{merge} (%)
<4.82	1481	1372	4.6	75 (5.4)	43.4	1283 (93.5)	5.52
<3.82	1424	1369	4.3	70 (5.1)	26.8	1242 (90.7)	7.20
<3.34	1397	1366	3.7	129 (9.4)	13.2	1150 (84.2)	10.51
<3.03	1379	1346	3.3	178 (13.2)	6.7	972 (72.2)	17.96
<2.82	1396	1364	3.0	242 (17.7)	3.9	786 (57.6)	26.07
<2.65	1358	1315	2.5	367 (27.9)	2.4	584 (44.4)	36.81
total	8435	8132	3.6	1061 (13.1)	16.2	6017 (74.0)	10.18

Winkler et al., 1990) that in contrast to phospholipase A2 triglyceride lipases may undergo a conformational change in response to adsorption at the oil–water interface. This proposal has recently been confirmed by an X-ray crystallographic study of RmL inhibited by *n*-hexylphosphonate ethyl ester (Brzozowski et al., 1991). Limited resolution of the data (complete to 3.2 Å, with only 27% between 3.2 and 3.0 Å) resulted in a qualitative description of the conformational change observed upon inhibition: a rigid-body hinged motion consisting of a rotation of the helix of 169° around its axis and a translation (measured with respect to the center of gravity of the "lid") of over 8 Å (Brzozowski et al., 1991). No more detailed analysis was possible due to the poor quality of the crystals.

In search of better quality crystals of an "activated" enzyme we have crystallized another complex of RmL using a commercially available serine proteinase inhibitor, diethyl *p*-nitrophenyl phosphate (E600). Upon inhibition *p*-nitrophenol is cleaved off and a covalent bond is formed between the hydroxyl oxygen of the catalytic serine and the phosphorus atom of the diethyl phosphate. The stereochemical details of the complex in the immediate vicinity of the catalytic center are identical to those observed in *n*-hexylphosphonate complex (Brzozowski et al., 1991). Consequently the present work does not result in any additional insight into the catalytic mechanism, the details of which can be largely inferred from numerous studies of serine proteinases. Enhanced resolution did result, however, in a much more detailed description of the conformational change in the enzyme. The three known structures of lipases indicate that a conformational change in these enzymes in response to the oil–water interface is an essential part of their function. Our report, which constitutes the first detailed stereochemical description of an activated triglyceride lipase, focuses on this phenomenon.

Materials and Methods

Purified RmL (Huge-Jensen et al., 1987; Boel et al., 1988) was donated by NOVO-Nordisk Research Laboratories (Copenhagen, Denmark). Tris-HCl solution (20 mM, pH 8.05) was saturated with E600 (Sigma) by agitating the mixture for several hours, following which the E600 phase was discarded. Lyophilized protein was then dissolved to a concentration of ca. 20 mg/mL, and the mixture was incubated overnight at 37 °C. Precipitated protein was separated by centrifuging at 15000*g*. The remaining solution was used for crystallization, which was carried out using a hanging drop technique. Poly(ethylene glycol) (MW 600) was used as a precipitant. Best crystals grew over the range of precipitant concentration of 8–12%. They were isomorphous with those observed for the *n*-hexylphosphonate complex (space group $C222_1$, $a = 48.6$ Å, $b = 93.9$ Å, $c = 122.1$ Å), but were found to be thicker and easier to handle. Data were collected using a Xentronics (Siemens) area detector mounted on a conventional source and used as described elsewhere (Derewenda & Helliwell, 1989). The final data set (see Table I for all the details of the crystallographic data), processed using the XENGEN suite of programs (Howard et al., 1987), included a total of 8132 unique reflections (96.4% complete to 2.65-Å resolution). The overall merging *R* factor ($R_m = \sum|I - \langle I\rangle|\sum I$) was 0.102. In comparison with the earlier study (Brzozowski et al., 1991) almost 3000 (ca. 60% of the earlier data set) more unique reflections, half of which exceed $2\sigma(I)$ level, were measured. Consequently the reflections/parameters ratio, which is critical in crystallographic refinement, increased significantly from 0.63 to 0.89 (the atomic parameters of water molecules are included in the latter figure). The crystallographic least-squares refinement of the atomic model was initiated with the *n*-hexylphosphonate complex coordinates (protein atoms only), for which the starting conventional crystallographic *R* factor ($\sum|F_o - F_c|/\sum F_c$) was 0.255. The PROLSQ restrained refinement (Hendrickson, 1985) and FFT was used. The E&S PS330 version of FRODO (Jones, 1978) was used for all electron density map fitting and structural analyses. Unless otherwise specified, all calculations have been carried out using the CCP4 suite of crystallographic programs (Science and Engineering Research Council, U.K. Daresbury Laboratory) either in its VMS or Irix implementations. The solvent-accessible surfaces were calculated using the DSSP program (Kabsch & Sander, 1983).

All structural comparisons with the native enzyme were carried out using a highly refined set of atomic coordinates of the latter (crystallographic *R* factor of 0.129 for all reflections for which $I > 3\sigma(I)$ in the 7.5–1.9-Å resolution range, with interatomic distances deviating by 0.015 Å from target values). A full structural analysis of the native enzyme at 1.9-Å resolution is currently being prepared for publication.

Results

Crystallographic Refinement and Model Quality. Five initial cycles of alternating *x*, *y*, *z* and *B* factor restrained refinement followed by three cycles of combined positional and thermal parameters refinement resulted in an *R* factor of 0.211. The FFT matrix contributions were scaled by a factor of 0.2 with respect to the stereochemical restraints imposed by PROLSQ in order to maintain good geometry of the model. As the initial *B* factors were those taken from the native model (the limited resolution of the *n*-hexylphosphonate complex study did not allow for temperature factor refinement), considerable bias toward these initial values of *B* factors was observed. All temperature factors were reset to 20 Å^2, and refinement was resumed with an *R* factor of 0.239. Seven cycles of simultaneous *x*, *y*, *z* and *B* factor refinement brought it down to 0.194, at which stage a $2F_o - F_c$ and $F_o - F_c$ electron density maps were calculated and analyzed. The highest peak of positive density coincided with the expected position of diethyl phosphate, and a molecular model of the latter was built into the density and included in subsequent refinement. Solvent molecules were now included, provided that the associated electron density exceeded the rms density at least by a factor of 2. The solvent structure analyzed in detail in the native enzyme (unpublished data) provided useful guidelines

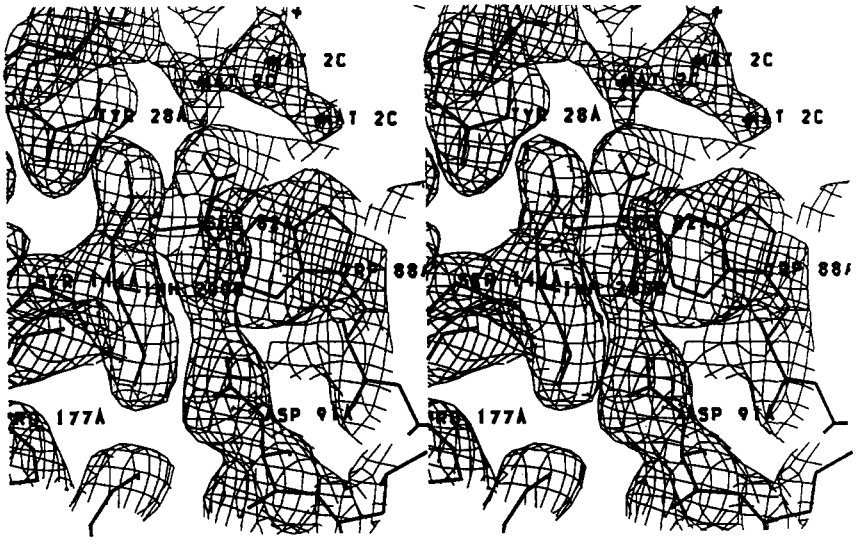


FIGURE 1: Representative final $2F_o - F_c$ electron density showing the position of the bound diethyl phosphate and the neighboring residues, including Ser-82. Note the well-ordered and clearly resolved water molecules [the coefficients of the Fourier sum for a $2F_o - F_c$ electron density map are $(2F_o - F_c) \exp(i\phi_c)$, where F_o is an observed structure factor magnitude and F_c and ϕ_c are the amplitude and phase, respectively, of a structure factor calculated from a refined model].

FIGURE 2: A schematic representation of the native RmL lipase molecule showing the location of the surface loop containing the "lid" fragment. The ribbon shows all of the molecule with the exception of the loop, which is shown in a thick line connecting the individual Cα atoms; three of these atoms within the "lid" region are labeled.

for the identification of ordered water molecules in the complex. A total of 170 water molecules were identified at this stage and included in subsequent refinement. Eleven cycles of refinement reduced the R factor to 0.16 [only reflections for which $I > \sigma(I)$ were used in this round], and another set of maps were calculated. Solvent structure was reviewed, and the current total of 237 water molecules were included. The protein model was carefully analyzed for any possible remaining errors. Three Arg residues were found to show static disorder with Arg-30 in three possible conformations, Arg-80 in two conformations, and Arg-86 in two main conformations. These alternative conformations have been treated accordingly in subsequent refinement. The last round of crystallographic refinement consisted of three x, y, z and five x, y, z, B cycles.

The final crystallographic R factor for the refined model of the RmL–E600 complex was 0.146 for all data between 10 and 2.65 Å for which $I > \sigma(I)$ and 0.162 for all data within the same resolution limits (the final set of coordinates has been deposited in the Brookhaven Protein Data Bank). The stereochemical parameters of the atomic model of RmL in complex with E600 are listed in Table II. All the parameters show satisfactory values indicating proper convergence of refinement; the relatively high number of water molecules identified in spite of the resolution limit of 2.65 Å is due to the high quality and contrast of the electron density map (see Figure 1). This is frequently observed when the native protein model is derived from high-resolution refinement (as in the present study).

Overall Comparison with the Native Molecule. The RmL lipase is a single-chain (269 residues) protein molecule with three disulfide bridges (Cys-29–Cys-268, Cys-40–Cys-43, Cys-235–Cys-244); four out of a total of 13 prolines (34, 209, 229, 250) are in the cis conformation. The central core of the molecule is formed by a mixed (though predominantly parallel) eight-stranded β-sheet. All interstrand connections are right-handed, and four α-helices have also been identified. The catalytic center in RmL is formed by a serine proteinase type triad of Ser-144, Asp-203, and His-257. A long surface loop (residues 80–109) connects the C-terminal end of strand 3 with the N-terminal end of a long kinked helix which leads to strand 4. This "loop", in principle, can be classified as a large Ω loop, although it defies a classical definition (Leszczynski & Rose, 1986) in that it exhibits well-defined secondary structure in its helical fragment. It is a 17-residue-long (82–96) fragment of this surface loop which obscures the entrance to the active site in the native enzyme (Figure 2).

Table II: Stereochemical Parameters for the Refined Atomic Model

	standard deviation	σ
(1–2) bond distances	0.019 Å	0.020 Å
(1–3) angle distances	0.069 Å	0.040 Å
(1–4) distances	0.087 Å	0.060 Å
planarity	0.015 Å	0.020 Å
chiral volumes	0.159 Å^3	0.120 Å^3
nonbonded contacts		
single torsion	0.234 Å	0.500 Å
multiple torsion	0.325 Å	0.500 Å
conformation torsion angles		
planar (0°, 180°)	2.63°	3.000°
staggered (±60°, 180°)	23.77°	5.000°
orthonormal (±90°)	33.70°	15.000°
isotropic temperature factor		
main-chain bond	0.79 Å^2	1.000 Å^2
main-chain bond	1.36 Å^2	1.500 Å^2
side-chain bond	1.60 Å^2	1.500 Å^2
side-chain angle	2.57 Å^2	2.000 Å^2

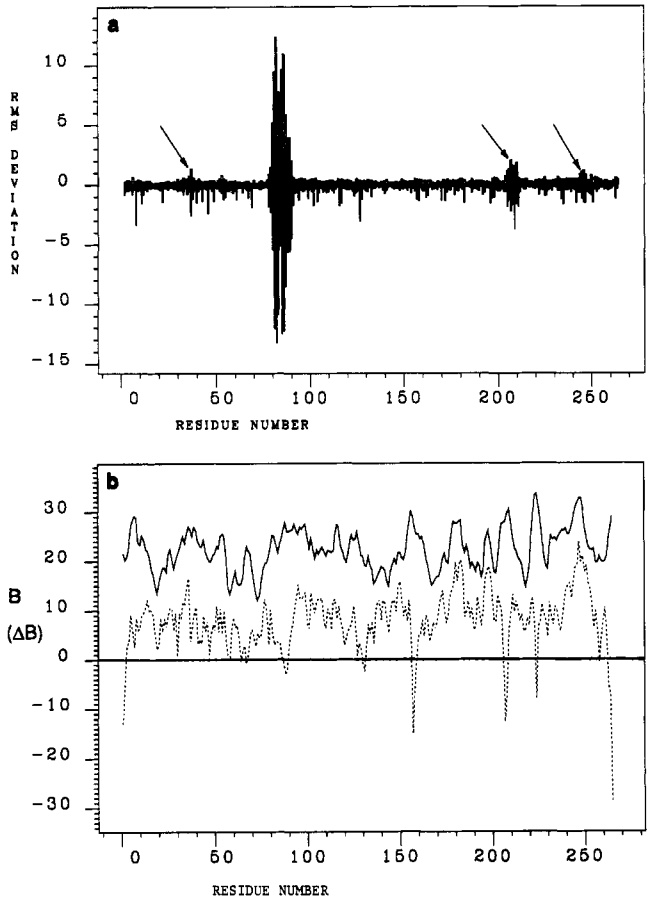


FIGURE 3: Overall comparison of the native and inhibited molecules of RmL. (a) Rms difference (Å) between the native and inhibited molecules of RmL lipase. Averages have been calculated for both the main chain (upper histogram) and side chain [lower (negative) histogram] within each residue and are shown against a residue number. Arrows denote regions of intermolecular contacts excluded from the least-squares fitting. (b) Mean main-chain isotropic temperature factors (averaged per residue) of the inhibited complex, and the difference plot (dotted line) obtained by subtracting the mean main-chain temperature factor of the native structure from the analogous value in the complex.

When the native and E600-inhibited structures of lipase are superimposed by least-squares optimization of all main-chain atoms, four regions of localized conformational change can be identified. They include three short segments (residues 40–41, 209–216, and 247–253) which are involved in intermolecular contacts in either one or the other of the crystal forms (see the section on crystal packing) and the region of the "lid" (residues 82–96) where the change is most profound. Least-squares optimization based on main-chain atoms excluding all above-mentioned regions (Figure 3a) resulted in an rms deviation between the two structures of 0.34 Å (932 main-chain atoms were used). There is, however, a striking similarity in the structure of the central part of the lid, the α-helix 85–92 (Ile-Arg-Asn-Trp-Ile-Ala-Asp-Leu), between the native and inhibited forms of the enzyme: when they are superposed, the rms deviation for all main-chain atoms is 0.31 Å.

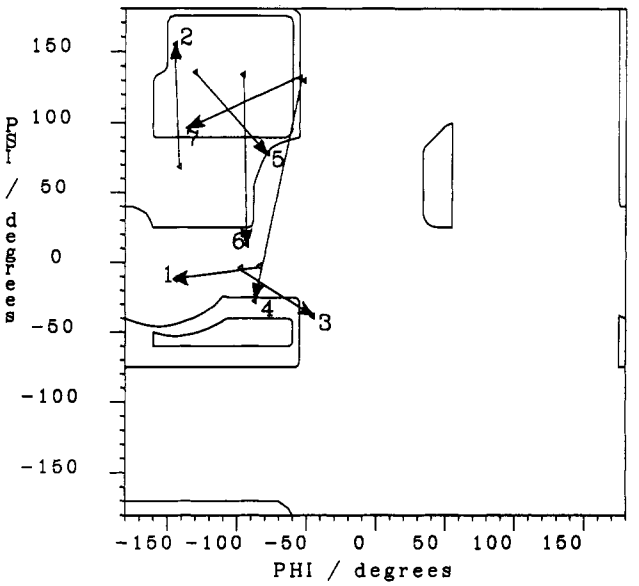
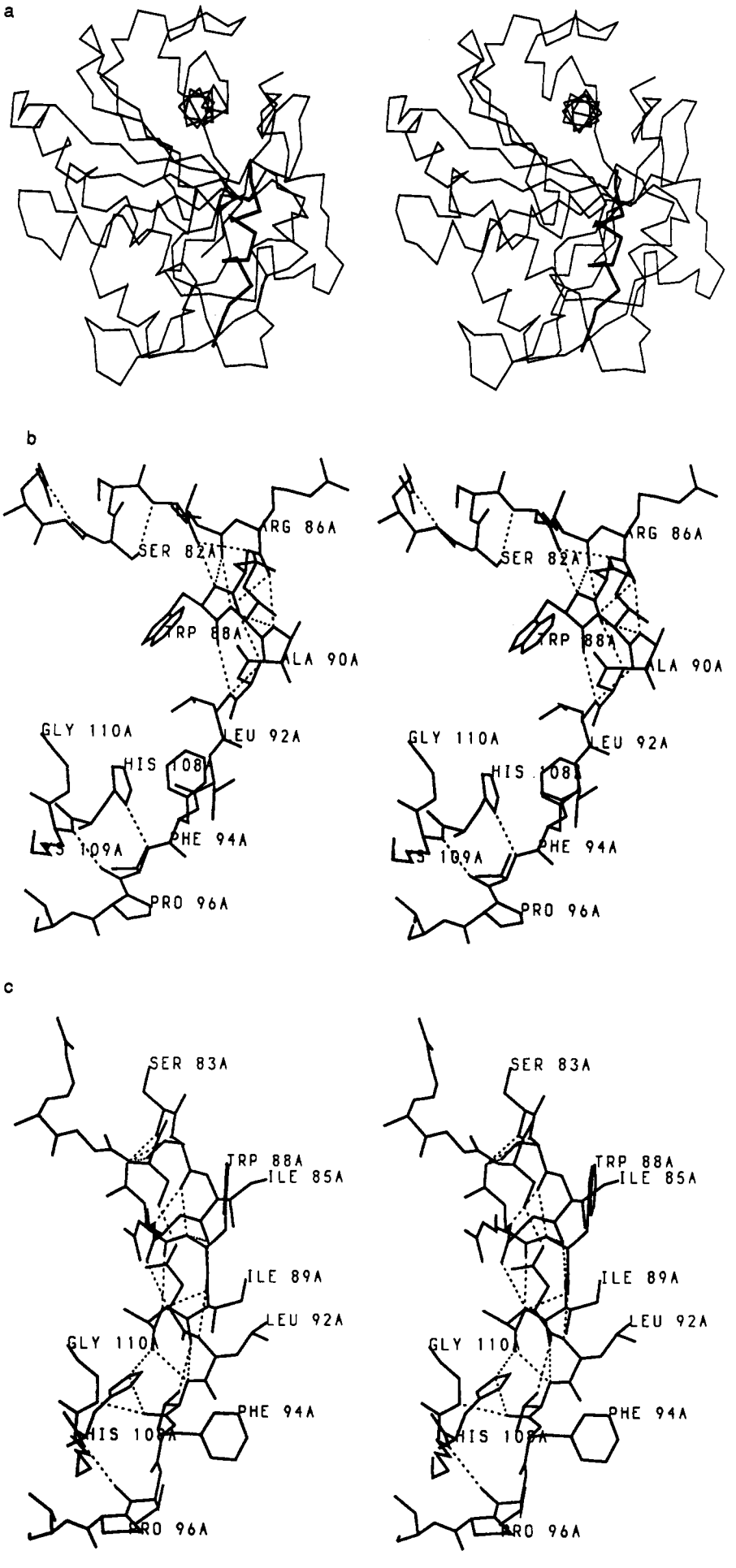


FIGURE 4: A Ramachandran plot showing conformational transitions of residues involved in the interfacial activation mechanism. The arrows are directed from the point indicating the secondary structure in the native enzyme toward that observed in the inhibited molecule. (1) Ser-83; (2) Ser-84; (3) Asp-91; (4) Leu-92; (5) Thr-93; (6) Phe-94; (7) Val-95.

Although the precision of the refinement of isotropic temperature factors is severely affected by the limited resolution of the data, it would still be possible to detect major differences between the native and inhibited structures. Figure 3b illustrates that there are no such differences, apart from the overall increase in thermal disorder upon inhibition; this is consistent with the observed poorer quality of the crystals of the complex.

Conformational Changes in the "Lid" and Hinge Regions. In order to identify and characterize the hinge regions, we have analyzed the changes in the conformational torsion angles (ϕ and ψ) versus the residue number (Figure 4). There are two hinge regions. At the N-terminal end of the "lid" Ser-83 and Ser-84 each undergo a conformational change. Ser-83, in spite of a change in the ϕ angle of 60°, remains within the γ_R region of the Ramachandran plot [the secondary structure definition in this paper is based on the notation introduced by Efimov; the δ region of conformational space is below β in the Ramachandran plot, the γ_R region is to the left of the classical α-helix (Efimov, 1986)] while Ser-84 changes its conformation from δ to β. The latter change is accomplished by a change of ψ in Ser-84. At the C-terminal end of the lid the conformational transition is initiated at residue 91 and affects four subsequent amino acids (Asp-91, Leu-92, Thr-93, and Phe-94), all of which undergo changes in their conformations: $\delta \rightarrow \alpha$, $\beta \rightarrow \alpha$, $\beta \rightarrow \delta$, and $\beta \rightarrow \gamma_R$, respectively. Val-95 manages to retain its β secondary conformation. Again, all the changes can be traced to individual dihedral angles.


a
b
ARG 86A
SER 82A
TRP 88A
ALA 90A
GLY 110A
LEU 92A
HIS 108A
PHE 94A
PRO 96A
c
SER 83A
TRP 88A
ILE 85A
ILE 89A
LEU 92A
GLY 110A
PHE 94A
HIS 108A
PRO 96A

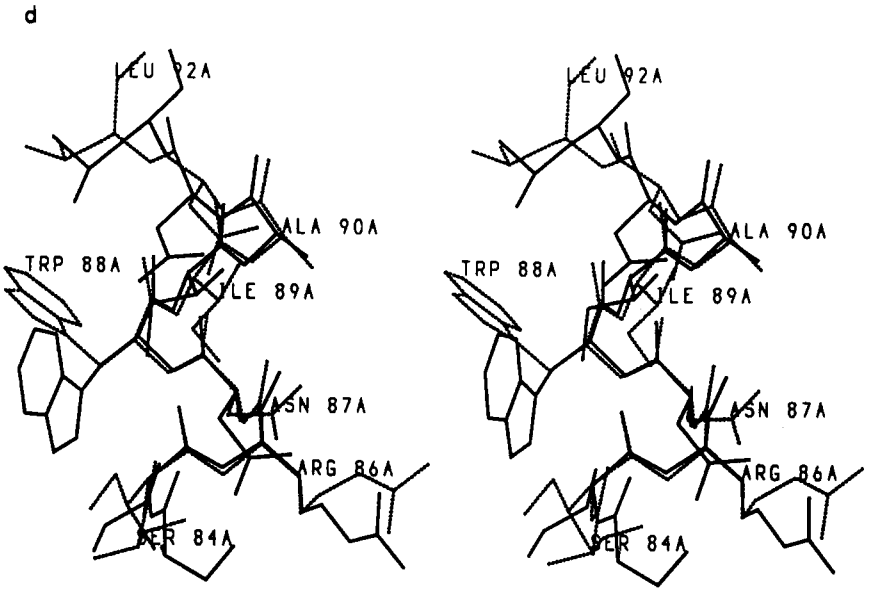


FIGURE 5: A structural comparison of the "lid" fragments in the native and activated species: (a) A schematic view of the RmL molecule (only the Cα positions are shown) showing the positions of the "lid" residues in the native (thin line) and inhibited (thick line) molecules. (b) The "lid" in native RmL; hydrogen bonds are shown as broken lines. (c) The "lid" in the inhibited form; the view is kept the same in both (a) and (b). (d) A comparison of the least-squares fitted helices of the "lid".

Among the side chains of the "lid" amino acids only Trp-88 undergoes a significant change: the χ_1 and χ_2 angles are 179° and 79°, respectively, in the native form and −74° and −125° in the inhibited species.

Hydrogen Bonding. In the native enzyme the lid residues are involved in a number of hydrogen bonds (Figure 5), although in the segment between Ser-82 and Phe-94 all of these bonds are internal. In the native enzyme the hydroxyl of Ser-82 is hydrogen bonded to the amide of Ser-84. From this point the main-chain atoms form the usual 1←5 α-helical bonds with Leu-92 assuming the 3_{10}-helix conformation frequently observed in amino acids and not involved in any other contacts. The C-terminal fragment of this region (93–96) is stabilized by the interactions of the main-chain atoms of Val-95: N to Nδ1 of His-108 and O to N of Lys-109.

Upon activation, changes in this network are observed. Ser-82 (involved in catalysis) is stabilized in its new conformation by an interaction of its hydroxyl with Oδ1 of Asp-91. The polar side chain of Asn-87, now buried beneath the "lid", forms a number of new interactions: Oδ1 binds the amide nitrogens of Thr-93 and Phe-94, while Nδ2 forms a bond with the carbonyl of Tyr-60. The hydroxyl group of Ser-84 becomes H-bonded to Oδ1 of Asp-61. The helical central part of the "lid" retains its structural integrity, with the carbonyl oxygen of Ile-89 forming an additional H-bond with the hydroxyl of Thr-93. Leu-92 becomes incorporated into the helix, while Thr-93 assumes the 3_{10} conformation to become the last amino acid in the helix. The imidazole of His-108 changes its H-bonding partner from the amide nitrogen of Val-95 to the carbonyl oxygen of Ala-90, with Nε2 involved in place of Nδ1. The change from a hydrogen bond acceptor (in the native molecule) to a donor (in the activated species) is easily accomplished since only one hydrogen bond is formed by the imidazole in each structure.

It is interesting to note that in the active enzyme the "lid" gains three additional interactions with the main domain of the protein molecule.

In the E600 complex the "lid" occupies a new position on the surface of the molecule, some 8 Å away from the native location. This deep surface depression extending over 10 Å into the molecule is filled in the native enzyme by 18 water molecules, half of which are directly hydrogen bonded to the polar protein groups. During the conformational change all but three of these solvent molecules are expelled. In the E600 complex these three molecules become buried and mediate the polar contacts formed primarily by Asn-87.

Changes in the Exposed Surface. The moving "lid" brings about a profound change in the nature of the exposed surface of the active lipase molecule. We have quantified this change in the E600 complex by analyzing the changes in the exposed surfaces of the amino acids. A plot of the difference solvent-accessible surface (Figure 6a) reveals that there are 12 distinctly hydrophobic amino acids (Ile-85, Trp-88, Ile-89, Leu-92, Phe-94, Val-205, Leu-208, Phe-213, Val-254, Leu-255, Leu-258, Leu-267) which become significantly more exposed upon activation. The total newly exposed area taking into account these residues alone amounts to 732 Å^2 or 7% of the total surface of 10 809 Å^2. This is accompanied by a significant loss of predominantly polar surface. The polar residues of the lid (Ser-84, Asn-87, Asp-91, and Thr-93) account for 329 Å^2, while the tetrapeptide 58–61 (Tyr-Asp-Thr-Asn), which interacts with the "lid" in its new position, loses 124 Å^2 of the surface exposed to solvent in the active enzyme. All these changes occur around the area of the active site (Figure 6b).

Crystal Packing. The availability of atomic coordinates of RmL crystallized in two distinct crystal forms makes it possible to assess the possible influence of crystal packing on the molecular structure of the protein. As was already pointed out, a comparison of the native and inhibited structures revealed three regions (Figure 3) which have been affected in either form by intermolecular interactions. All of these changes involve external loops, which occasionally move by 1–2 Å, but these shifts seem highly localized and of no further significance.

A detailed analysis of the inhibited form reveals that the predominant intermolecular interaction involves the large hydrophobic area exposed by the conformational change, including the "lid". Figure 7 shows a pair of molecules related by crystallographic symmetry (a 2-fold axis) involved in this interaction. The hydrophobic surfaces are completely buried in this crystal dimer, and no water molecules are observed at the interface. The indole groups of two symmetry-related

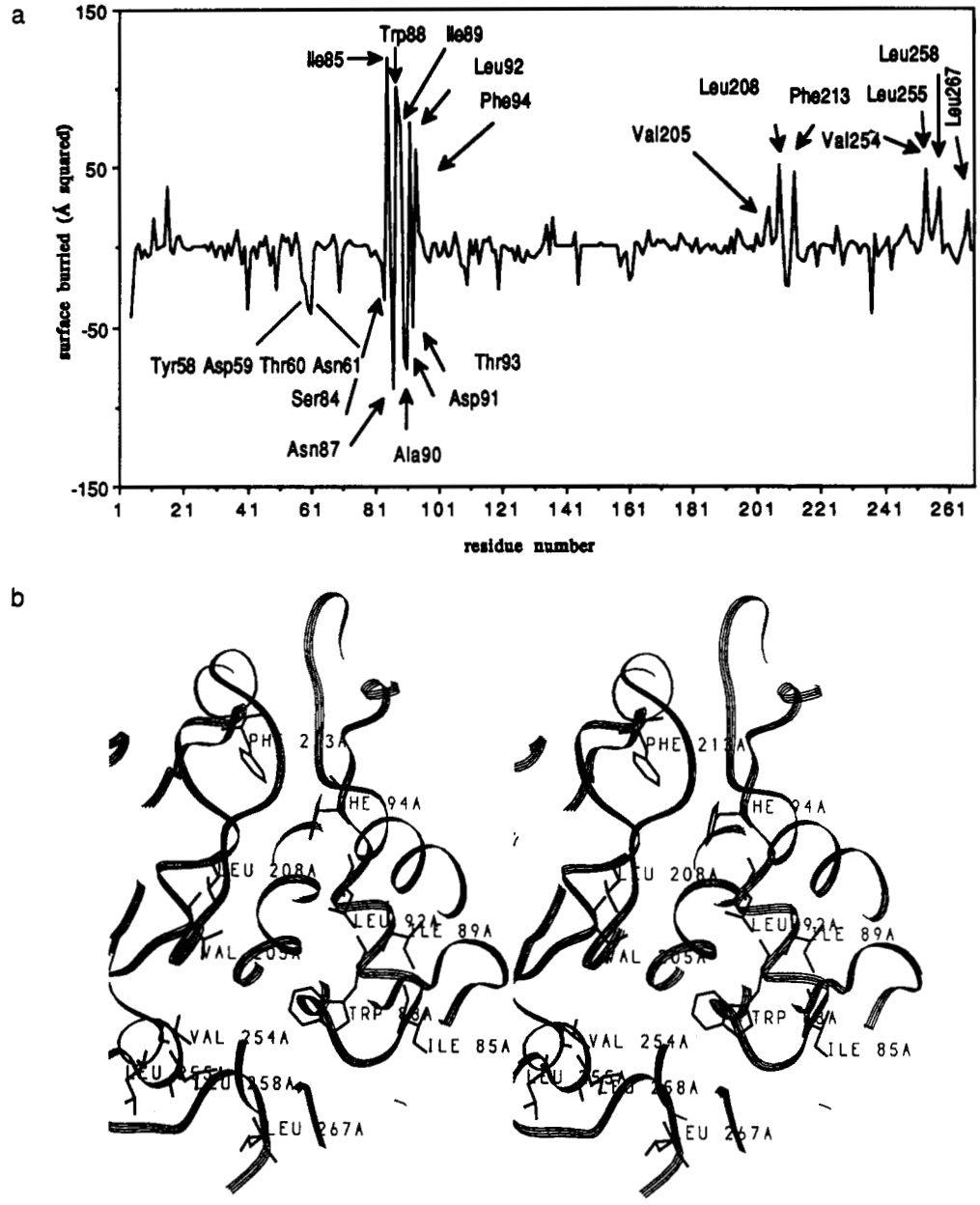


FIGURE 6: (a) A plot of the difference surface area versus the residue number. For each residue the solvent-accessible area in the inhibited form was subtracted from that in the native enzyme. The residues believed to play the key role are indicated with arrows. (b) The location of the 12 hydrophobic residues which become significantly more exposed in the inhibited enzyme. For clarity a thin slab at the surface of the molecule (around the entrance to the active site) is shown. The catalytic serine is located at the tip of the helix seen in the center.

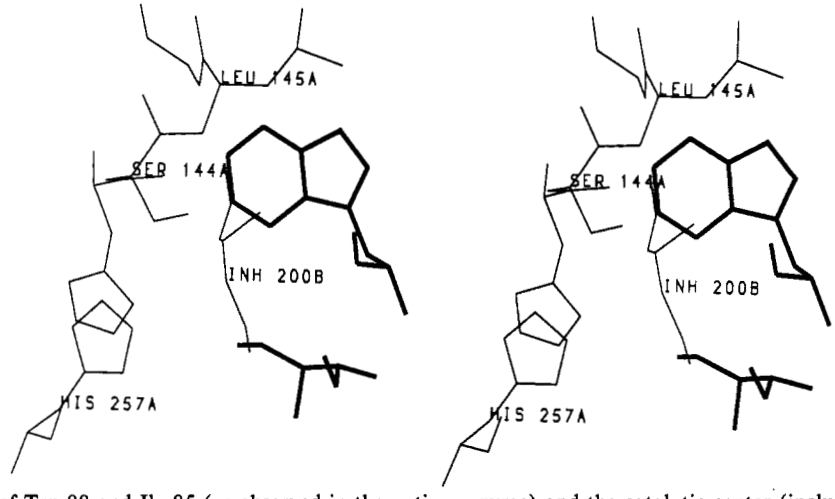


FIGURE 7: Superposition of Trp-88 and Ile-85 (as observed in the native enzyme) and the catalytic center (including the diethyl phosphate inhibitor) of the complex. The two models are superposed as described in the text.

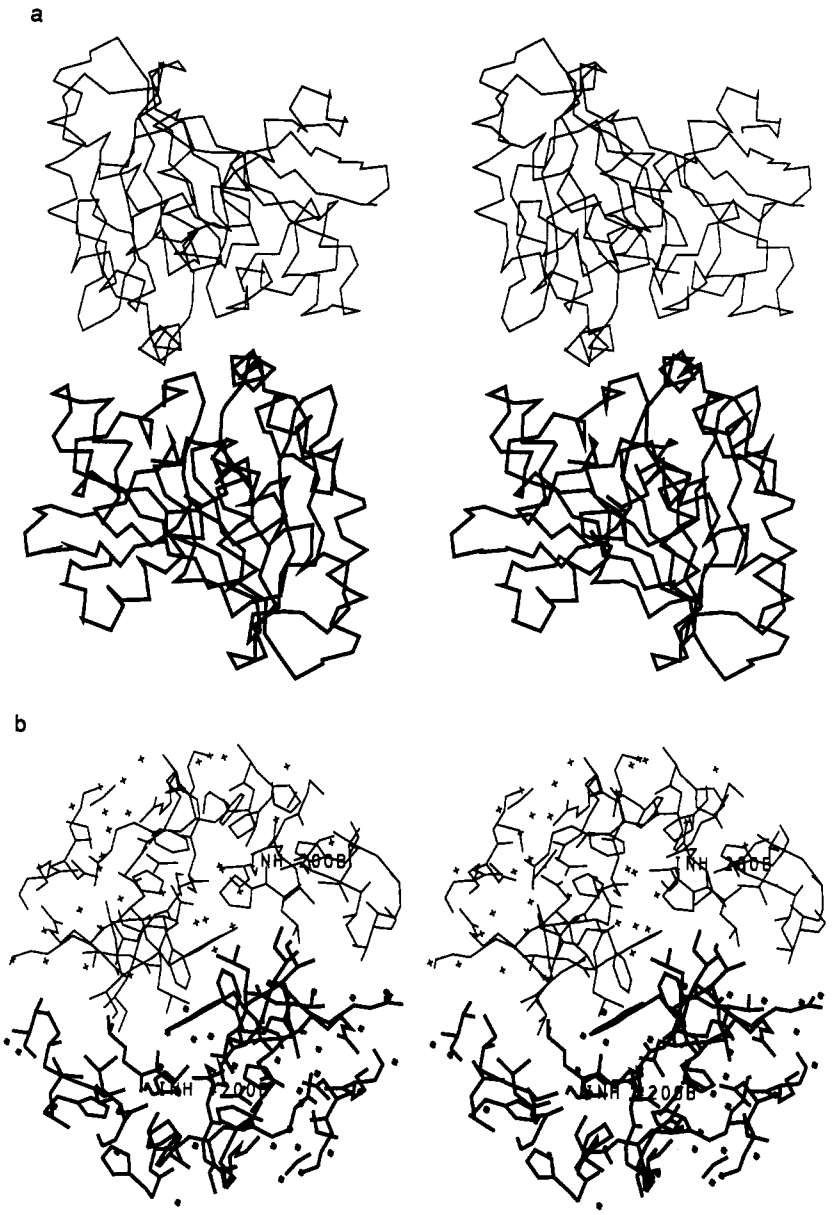


FIGURE 8: The main intermolecular interaction within the crystal of the complex: (a) A schematic view (Cα atoms only are shown) of two RmL molecules related by a 2-fold axis. The two short helices in the middle of the diagram are the "lids". (b) An expanded view of the central region of (a) showing the interactions between the "lids" in the crystal dimer; the stars correspond primarily to single protein atoms cut from their covalent partners by the sphere border. The positions of the two symmetry-related inhibitor molecules are labeled.

Trp-88 are stacked in an antiparallel fashion but do not form a π-complex.

Discussion

Two questions must be considered before conclusions can be drawn from the present study: (1) is the conformational change absolutely essential for the expression of the catalytic activity of the enzyme, and (2) can the observed new conformation be taken as the one characteristic of the activated enzyme adsorbed at the oil–water interface?

The first question is resolved conclusively by a comparison of the environs of the catalytic site in the native and inhibited molecules (Figure 8). In the former case two specific amino acids, Ile-85 and Trp-88, block the position which the inhibitor occupies in the complex. This observation alone indicates that a conformational change is essential for the activation of the enzyme.

The issue of whether the crystal structure represents exactly that of a molecule adsorbed at an oil–water interface is more difficult to answer. Clearly, the involvement of the newly exposed surfaces in a crystal interaction may have altered some details of the molecular architecture. In particular, the conformation of the side chain of Trp-88 in the crystal structure can be influenced by the interactions of the two symmetry-related indole groups. However, the conformation of the "lid" is stabilized by a set of new hydrogen bonds, and it appears from the analysis of the coordinates that there are only two stable extreme positions of the "lid", with probably a plethora of possible metastable intermediates. The structural details of the inhibited molecule as well as the affinity of this molecule toward hydrophobic surfaces (see the crystal-forming dimer interaction) is fully consistent with the view that the observed changes are representative of those occurring in the heterogeneous medium of an oil–water mixture. Further experimental solution studies may resolve this issue conclusively.

The very low rms difference between the atomic coordinates of the core of the molecule in the native and complexed states (0.34 Å) is indicative of relative rigidity of the molecule and compares well with analogous values calculated for other proteins undergoing similar changes [e.g., in TIM the rms

deviation of the α-carbons between the native and complexed forms is 0.42 Å (Joseph et al., 1990)]. It can also be argued that the value of 0.34 Å can be taken as the upper estimate of the coordinate error in the present study, as there is no other chemical reason for the atomic coordinates of the internal cores of the two molecules to differ.

The rms difference calculated for main-chain atoms of the "lid" fragments before and after inhibition is even lower (0.31 Å). This is typical of a rigid body type motion. Such motions have been frequently described in multidomain proteins as bending domain motions [for review see Bennet and Huber (1984)]. A rigid-body displacement of a small structural element within a protein molecule is less common, but not unknown. In TIM an 11-residue loop closes over the active site upon substrate binding, and while it retains its secondary structure intact, it is transported over a distance of 7 Å by the action of flexible hinge regions (Joseph et al., 1990). However, the individual residues within those hinges remain within the same allowed conformational regions throughout the cycle, whereas in RmL we observe transitions along the $\beta \leftarrow \delta \rightarrow \alpha$ trough in the Ramachandran plot, which require more energy. However, the particularly high-energy barriers of forbidden Ramachandran regions and, specifically, a high-energy ridge of $\phi = 0$ are avoided, as all amino acids remain within the $\phi < 0$ half of the Ramachandran plot. Figure 4 illustrates only the simplest possible routes of conformational changes, along straight lines joining the two conformational states. The actual transition may be more complex. However, we have carefully examined the two structures, and we conclude that the pathways close to those shown in Figure 4 are quite probable as they do not involve any major collisions during the conformational change.

The considerable increase in the exposed hydrophobic area upon inhibition (or rather activation) is unlikely to be coincidental. The 12 hydrophobic residues clustered around the entrance to the catalytic site (see Results) create a patch of hydrophobic surface which in terms of its area (ca. 750 Å^2) is slightly higher than the average buried accessible enzyme surface in proteinase–inhibitor complexes and very similar to the average buried antibody surface in four analyzed antibody–antigen complexes (Janin & Chothia, 1990). The energy which could be gained by burying this surface in apolar medium is usually estimated at approximately 25 cal·Å^{-2}·mol^{-1} (Chothia, 1974; Eisenberg & McLachan, 1958), or approximately 19 kcal·mol^{-1} in total for the RmL transition. The burying of polar residues, as long as their hydrogen-bonding potential is saturated (as seems to be the case here), is likely to result in very small loss of energy (Roseman, 1988). For an enzyme adsorbed to the interface the loss of translational and rotational entropy is probably similar in magnitude to the entropy lost on dimer formation (Finkelstein & Janin, 1989) which amounts to approximately 15 kcal·mol^{-1}. Thus, according to this rather qualitative argument, the balance of forces is consistent with RmL being stabilized at the hydrophobic surface. It must be remembered, however, that the energy estimates continue to be debated. For example, a recent report (Honig, 1991) indicates that accounting for size and curvature of molecules results in a hydrophobic effect estimate equivalent in magnitude to the surface tension at the hydrocarbon–water interface (72 cal·Å^{-2}·mol^{-1}), considerably higher than the hitherto used estimates.

The X-ray structures of complexes in a crystalline state fail to answer directly any questions relating to the dynamics of interfacial activation or the sequence of events at the interface. It can be easily hypothesized that the enzyme "opens up" in response to a hydrophobic surface and proceeds to hydrolyze the triglycerides within the micelle. On the other hand, lipid molecules are oriented so that their polar heads, rather than the apolar fatty acid chains, point to the surface; also, it is the polar head of a triglyceride which is bound and immobilized by the enzyme, along with one fatty acid chain (Brzozowski et al., 1991). Solution studies coupled with site-directed mutagenesis may shed more light on the phenomenon.

Although the present study involves a relatively small extracellular fungal protein, the observations reported here may have a direct bearing on our understanding of the structure/function relationships in mammalian systems. The crystallographic structure of hPL reveals the presence of a long loop with a helical apex obscuring the active site (Winkler et al., 1990). Interestingly, the position of this short helix with respect to the active center is virtually identical to that in RmL, though the differences in the tertiary structure rule out any homology [atomic coordinates of hPL have not been deposited, and our comparisons have been based on hPL coordinates derived from the stereo figures of Winkler et al. (1990) as described by Derewenda and Cambillau (1991)]. It is also established that both hepatic and lipoprotein lipases exhibit sequence homology with the pancreatic enzyme, which warrants a conclusion that a common tertiary motif characterizes this family of enzymes (Datta et al., 1988; Presson et al., 1989; Derewenda & Cambillau, 1991). Many lipases are regulated at the molecular level by various mechanisms. Pancreatic lipase is inactivated by bile salts but regains activity in the presence of its cofactor—colipase. Lipoprotein lipase requires apolipoprotein CII for full activity, whereas hepatic lipase has no partner. Outside this family, bile salt activated lipase is inactive in the absence of bile salts. It is therefore possible that the interactions modulating the enzymes involve the "lids", whose altered amino acid sequences reflect different requirements imposed by cofactors. Further structural studies will undoubtedly resolve this fundamental question.

Acknowledgments

The experimental part of the work (crystal growth and data collection) was completed in the laboratories of the Protein Structure Group, Department of Chemistry, University of York, U.K., prior to Z.S.D. and U.D. moving to Edmonton. That stage was financed by NOVO-Nordisk and the Bridge lipase project of the European Economic Community. We thank Dr. Anita Sielecki (University of Alberta, Edmonton) for help with the DSSP program and Drs. R. Read, L. Smillie, and M. Michalak (University of Alberta, Edmonton) for critical reading of the original manuscript. We thank the anonymous reviewers of *Biochemistry* for helpful suggestions which resulted in significant improvements with respect to the original text.

References

Bennett, W. S., & Huber, R. (1984) *CRC Crit. Rev. Biochem. 15*, 291–384.

Blow, D. M., Birktoft, J. J., & Hartley, B. S. (1969) *Nature 221*, 337–339.

Boel, E., Huge-Jensen, B., Christiansen, M., Thim, L., & Fiil, N. P. (1988) *Lipids 23*, 701–706.

Brady, L., Brzozowski, A. M., Derewenda, Z. S., Dodson, E., Dodson, G., Tolley, S., Turkenburg, J. P., Christiansen, L., Huge-Jensen, B., Norskov, L., Thim, L., & Menge, U. (1990) *Nature* **343**, 767–770.

Brzozowski, A. M., Derewenda, U., Derewenda, Z. S., Dodson, G. G., Lawson, D., & Turkenburg, J. P. (1991) *Nature 351*, 491–494.

Chothia, C. (1974) *Nature 248*, 338-339.
Datta, S., Luo, C.-C., Li, W.-H., VanTuinen, P., Ledbetter, D. H., Brown, M. A., Chen, S.-H., Liu, S.-W., & Chan, L. (1988) *J. Biol. Chem. 263*, 1107-1110.
Derewenda, Z. S., & Helliwell, J. R. (1989) *J. Appl. Crystallogr. 22*, 123-137.
Derewenda, Z. S., & Cambillau, C. (1991) *J. Biol. Chem. 266*, 23112-23119.
Derewenda, Z. S., & Derewenda, U. (1991) *Biochem. Cell. Biol.* (in press).
Efimov, A. V. (1986) *Mol. Biol. (Moscow) 20*, 250-260.
Eisenberg, D., & McLachlan, A. D. (1985) *Nature 319*, 199-203.
Finkelstein, A., & Janin, J. (1989) *Protein Eng. 3*, 1-4.
Franken, S. M., Roseboom, H. J., Kalk, K. H., & Dijkstra, B. W. (1991) *EMBO J. 10*, 1297-1302.
Janin, J., & Chothia, C. (1990) *J. Biol. Chem. 265*, 16027-16030.
Jones, A. (1978) *J. Appl. Crystallogr. 11*, 268-272.
Joseph, D., Petsko, G. A., & Karplus, M. (1990) *Science 249*, 1425-1428.
Hendrickson, W. A. (1985) *Methods Enzymol. 115*, 252-270.
Honig, B. (1991) 15th International Congress of Biochemistry, Jerusalem, Collected Abstracts, p 123.
Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H., & Salemme, F. (1987) *J. Appl. Crystallogr. 20*, 383-387.
Huge-Jensen, B., Gailuzzo Rubano, D., & Jensen, R. G. (1987) *Lipids 22*, 559-565.
Kabsch, W., & Sander, C. (1983) *Biopolymers 22*, 2577-2637.
Leszczynski, J. F., & Rose, G. D. (1986) *Science 234*, 849-855.
Liao, D. I., & Remington, S. J. (1990) *J. Biol. Chem. 256*, 6528-6531.
Pathak, D., & Ollis, D. (1990) *J. Mol. Biol. 214*, 497-525.
Persson, B., Bengtsson-Olivecrona, G., Enerback, S., Olivecrona, T., & Jornval, H. (1989) *Eur. J. Biochem. 179*, 39-45.
Sarda, L., & Desnuelle, P. (1958) *Biochim. Biophys. Acta 30*, 513-521.
Schrag, J. D., Li, Y., Wu, S., & Cygler, M. (1991) *Nature 351*, 761-764.
Scott, D. L., Otwinowski, Z., Gelb, M., & Sigler, P. B. (1990a) *Science 250*, 1563-1566.
Scott, D. L., White, S. P., Otwinowski, Z., Yuan, W., Gelb, M. G., & Sigler, P. B. (1990b) *Science 250*, 1541-1546.
Sussman, J. L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L., & Silman, I. (1991) *Science 253*, 872-879.
White, S. P., Scott, L. S., Otwinowski, Z., Gelb, M., & Sigler, P. B. (1990) *Science 250*, 1560-1563.
Winkler, F. K., D'Arcy, A., & Hunziker, W. (1990) *Nature 343*, 771-774.

Crystal Structure at 1.5-Å Resolution of d(CGCICICG), an Octanucleotide Containing Inosine, and Its Comparison with d(CGCG) and d(CGCGCG) Structures†,‡

Vinod D. Kumar,*,§ Robert W. Harrison,§ Lawrence C. Andrews,∥ and Irene T. Weber§

Macromolecular Structure Laboratory, National Cancer Institute, Frederick Cancer Research and Development Center, Advanced Biosciences Laboratories Basic Research Program, P.O. Box B, Frederick, Maryland 21702-1201, and Department of Pharmacology, Bluemle Life Sciences Building, Thomas Jefferson University, 233 South 10th Street, Philadelphia, Pennsylvania 19107

Received July 8, 1991; Revised Manuscript Received October 28, 1991

ABSTRACT: The octadeoxyribonucleotide d(CGCICICG) has been crystallized in space group $P6_522$ with unit cell dimensions of $a = b = 31.0$ Å and $c = 43.7$ Å, and X-ray diffraction data have been collected to 1.5-Å resolution. Precession photographs and the self-Patterson function indicate that 12 base pairs of Z-conformation DNA stack along the c-axis, and the double helices pack in a hexagonal array similar to that seen in other crystals of Z-DNA. The structure has been solved by both Patterson deconvolution and molecular replacement methods and refined in space group $P6_5$ to an R factor of 0.225 using 2503 unique reflections greater than $3.0\sigma(F)$. Comparison of the molecules within the hexagonal lattice with highly refined crystal structures of other Z-DNA reveals only minor conformational differences, most notably in the pucker of the deoxyribose of the purine residues. The DNA has multiple occupancy of C:I and C:G base pairs, and C:I base pairs adopt a conformation similar to that of C:G base pairs.

Inosine is a purine nucleoside whose neutral base forms stable base pairs with all four conventional bases, and the strength of the base pairing is approximately equal in each case (Martin et al., 1985). Inosine occurs naturally in the wobble position of the anticodon of some t-RNA's, where it appears to pair with adenosine in addition to cytidine and uridine, the nucleosides which pair with guanosine in that position. Poly(rI) and poly(dI) form stable helices with poly(rC) and poly(dC) (Felsenfeld & Miles, 1967) and serve as templates for the incorporation of cytosine into products of DNA and RNA polymerases (Hall et al., 1985). Oligonucleotides containing inosines have been used extensively as probes to screen human cDNA or genomic DNA libraries (Takahashi et al., 1985).

† This research was sponsored in part by the National Cancer Institute, DHHS, under Contract NO1-CO-74101 with ABL.

‡ The coordinates have been deposited in the Brookhaven Protein Data Bank (cgcicicg_dna.pdb).

* Author to whom correspondence should be addressed.

§ Present address: Department of Pharmacology, Bluemle Life Sciences Building, Thomas Jefferson University, 233 S. 10th St., Philadelphia, PA 19107.

∥ Present address: International Centre for Diffraction Data, 1601 Parklane, Swarthmore, PA 19081.

0006-2960/92/0431-1541$03.00/0