The open conformation of a *Pseudomonas* lipase

Joseph D Schrag[1]*, Yunge Li[1], Miroslaw Cygler[1], Dietmar Lang[2], Tanja Burgdorf[2], Hans-Juergen Hecht[2], Rolf Schmid[3], Dietmar Schomburg[2]*, Timothy J Rydel[4†], Joel D Oliver[4], Larry C Strickland[4], C Michelle Dunaway[4], Steven B Larson[5], John Day[5] and Alexander McPherson[5]

Background: The interfacial activation of lipases results primarily from conformational changes in the enzymes which expose the active site and provide a hydrophobic surface for interaction with the lipid substrate. Comparison of the crystallization conditions used and the structures observed for a variety of lipases suggests that the enzyme conformation is dependent on solution conditions. *Pseudomonas cepacia* lipase (PCL) was crystallized in conditions from which the open, active conformation of the enzyme was expected. Its three-dimensional structure was determined independently in three different laboratories and was compared with the previously reported closed conformations of the closely related lipases from *Pseudomonas glumae* (PGL) and *Chromobacterium viscosum* (CVL). These structures provide new insights into the function of this commercially important family of lipases.

Results: The three independent structures of PCL superimpose with only small differences in the mainchain conformations. As expected, the observed conformation reveals a catalytic site exposed to the solvent. Superposition of PCL with the PGL and CVL structures indicates that the rearrangement from the closed to the open conformation involves three loops. The largest movement involves a 40 residue stretch, within which a helical segment moves to afford access to the catalytic site. A hydrophobic cleft that is presumed to be the lipid-binding site is formed around the active site.

Conclusions: The interfacial activation of *Pseudomonas* lipases involves conformational rearrangements of surface loops and appears to conform to models of activation deduced from the structures of fungal and mammalian lipases. Factors controlling the conformational rearrangement are not understood, but a comparison of crystallization conditions and observed conformation suggests that the conformation of the protein is determined by the solution conditions, perhaps by the dielectric constant.

Addresses: [1]Biotechnology Research Institute, NRC of Canada, 6100 Royalmount Ave., Montreal, Quebec, H4P 2R2, Canada, [2]Dept. of Molecular Structure Research, GBF, Mascheroder Weg 1, 38124 Braunschweig, Germany, [3]Institute for Technical Biochemistry, University of Stuttgart, Allmanding 31, 70569 Stuttgart, Germany, [4]The Procter & Gamble Company, Miami Valley Laboratories, Cincinnati, Ohio 45253-8707, USA and [5]Department of Biochemistry, University of California, Riverside, California 92521, USA.

[†]Present address: Molecular Structure Corporation, 3200 Research Forest Drive, The Woodlands, Texas 77381-4238, USA.

*Corresponding author.
E-mail: Joe.Schrag@nrc.ca
D.Schomburg@Uni-Koeln.DE

Key words: crystallography, α/β-hydrolase fold, interfacial activation, lipase, three-dimensional structure

Received: 9 October 1996
Revisions requested: 25 October 1996
Revisions received: 9 December 1996
Accepted: 9 December 1996

Electronic identifier: 0969-2126-005-00187

Structure 15 February 1997, 5:187–202

Introduction

The many lipase structures reported in the last few years have helped to explain many of the long known properties of these enzymes. The structures of lipases from diverse sources, ranging from microbes (including fungi, yeast and bacteria) to mammalian enzymes, in the large part all conform to the α/β-hydrolase fold [1], a recently discovered structural motif common to a wide variety of hydrolases [2]. All of these structures clearly showed these lipases to be serine hydrolases with catalytic triads resembling those of serine proteases.

The defining characteristic of lipases for many years was the phenomenon termed 'interfacial activation' [3,4]. This term is used to describe the greatly increased enzymatic activity of lipases which is observed when the substrate concentration exceeds the critical micelle concentration (CMC). A number of different explanations of this phenomenon have been proposed [4]. Some focus on properties of the substrate and these can be grouped together as proponents of the 'substrate' theory. Other explanations of interfacial activation, proponents of the 'enzyme' theory, centered on properties of the lipase. In particular, it was suggested that lipases undergo a conformational change at the interface.

The first three-dimensional structures of lipases, in which the active sites were occluded by surface loops, clearly indicated the necessity of a conformational change to expose the catalytic site [5–7]. Subsequently, the structures of lipase–inhibitor complexes demonstrated the nature of the structural changes that make the active sites

accessible [8–11]. In addition to providing access to the active site, the structural rearrangements also change the surface properties of the enzymes and in some cases form the oxyanion hole. In each case described, the movement of the so-called 'lid' exposes a large hydrophobic surface area surrounding the active site. This movement results in an amphipathic molecule which could be properly oriented for interaction of the active site with a lipid interface. These structures appeared to present a consolidated view of lipase function consistent with the enzyme theory of activation and reconciled the solubility of lipases in aqueous solution, their function in a hydrophobic environment and the interfacial activation phenomenon. This structural explanation of interfacial activation gained further support from the structure of cutinase [12], which is not activated by interfaces and does not have surface loops which occlude the active site.

Crystallization of human and horse pancreatic lipase [6,13], *Geotrichum candidum* lipase [7], *Candida rugosa* lipase [14], and *Penicillium camembertii* lipase [15] using polyethyleneglycol (PEG) as a precipitant resulted in the closed conformations of the enzyme. In the cases of *Rhizomucor miehei* lipase and human pancreatic lipase, the open conformations were crystallized from PEG solutions, but only after the formation of enzyme–inhibitor complexes or enzyme–cofactor complexes in the presence of organic solvents or detergents [8,9]. In contrast, the crystallization of *Candida rugosa* lipase from organic solvents resulted in the open conformation, even without prior formation of a complex with an inhibitor [16]. This suggests that the conformations of lipases can be altered by controlling the solvent conditions. The three-dimensional structures of *Pseudomonas glumae* lipase (PGL) and *Chromobacterium viscosum* lipase (CVL) crystallized from PEG solutions were recently reported [17,18]. Consistent with the other lipases crystallized from PEG, the closed conformation was observed. *Pseudomonas* lipases have also been crystallized from n-propanol [19,20]. Based on the behavior of other lipases, *Candida rugosa* lipase in particular, one might expect that these conditions would produce the open conformation of the enzyme. This report describes the open conformation of *Pseudomonas cepacia* lipase (PCL) crystallized from n-propanol as determined independently in three laboratories. These crystal structures and comparison with the closed conformations of the *P. glumae* and *C. viscosum* lipases clearly indicate that conformational changes are important for interfacial activation of these bacterial lipases and that the protein conformation depends strongly on the solution conditions.

Results

The crystallization conditions used by the three independent groups varied slightly. At the Biotechnology Research Institute (BRI) and Gesellschaft für Biotechnologische Forschung (GBF) the crystals were grown at pH 8.5–8.7, whereas at Procter and Gamble (P&G) pH 6.5 was utilized. The GBF crystals were grown at 12°C, those at BRI at 18°C and those at P&G at room temperature (~20°C). Despite these differences, the crystals obtained were all isomorphous with those of *Pseudomonas fluorescens* lipase [19]. They belong to space group C2 and the unit cell dimensions varied only slightly from one group to another. There is one molecule in the asymmetric unit giving a packing density V_M of 2.4 Å^3 Da^{-1}, corresponding to a solvent content of 48% [21].

Comparison of three independent structures

Solution of the high resolution crystal structure of the same protein in three different laboratories is an uncommon occurrence. The three independent models of PCL presented here, therefore, provide a good measure of the precision and reliability of crystallographic methods. The data in Table 1 show that all of the three independently determined structures are of comparable quality in both resolution and in deviations of bond lengths and angles from ideality. After eliminating the loops which are poorly defined in one or more of the three structures (residues 18–26 and 219–222) from the comparison, the root mean square (rms) differences in the positions of mainchain atoms are approximately 0.3 Å. Comparison of the three structures including sidechains results in rms differences of about 0.5 Å for 2236 atoms. These deviations are approximately equal to the coordinate error estimated from the Luzzatti plot [22] for the GBF structure.

Analysis of the mainchain torsion angles shows that 88% of the amino acids are located in the most favored regions of the Ramachandran plot and 10% in the additionally allowed regions. Only residues Ser87 and Leu234 adopt strained conformations. The electron density corresponding to both residues is very well defined. Ser87 is the catalytic serine residue and adopts ϕ,ψ angles of 58° and –133°, comparable to those of the catalytic serine residues of other lipases [23]. In all three structures the ϕ,ψ angles of Leu234 are approximately 60° and –50°; Leu234 is the residue preceding a one residue insertion in PCL, as compared to PGL and CVL.

Global fold

Figure 1 shows a ribbon diagram of PCL along with a schematic of the topology of the protein. The topology is very similar to that of PGL and CVL [17,18]. This is to be expected as PCL exhibits 84% sequence identity with PGL and CVL (the amino acid sequence of PGL is identical to that of CVL). Many of the features of the α/β-hydrolase fold are maintained. The central β sheet in the core of the molecule conforms to strands β3–β8 of the α/β-hydrolase fold [2]. Throughout this report the β strands will be numbered to be consistent with the numbering of the consensus α/β-hydrolase fold, so the first strand is named β3. The active-site Ser87 lies at the C-terminal end of strand

Table 1

Root mean square differences (Å) in atom positions of three independent models.

	P&G	GBF	PGL	CVL
BRI*				
Mainchain only	0.301 (1236)	0.310 (1236)	0.638 (1072)	0.514 (1032)
Sidechains included	0.521 (2236)	0.476 (2236)		
P&G†				
Mainchain only	-	0.305 (1276)	0.611 (1072)	0.555 (1032)
Sidechains included		0.476 (2299)		
GBF‡				
Mainchain only		-	0.669 (1072)	0.511 (1032)

Numbers in parentheses are the number of atoms compared. *BRI = the Biotechnology Research Institute model; †P&G = the Procter & Gamble model; ‡GBF = the Gesellschaft für Biotechnologische Forschung model.

β5 in the strand-turn-helix motif described previously [2,24]. The mainchain takes a sharp turn at Gly110, located at the C-terminal end of strand β6. Strand β7 is the longest strand in the β sheet and its N-terminal half (residues 202–206) forms hydrogen bonds to strand β6 in a parallel manner and to residues 195–199 in an antiparallel pattern. The C-terminal half of strand β7 (residues 206–211) hydrogen bonds to both strands β6 and β8 in a pattern typical of parallel β strands. This hydrogen-bonding pattern is disrupted by a bulge at Leu278 in strand β8. The catalytic acid, Asp264, is part of a loop which follows strand β7 and the triad histidine, His286, is located in a loop following strand β8. Additionally, residues 22–23 and 27–28 form a short two-stranded antiparallel β sheet which lies between strands β3 and β4; residues 214–220 and 223–228, connected by a β-hairpin loop, also form a two-stranded antiparallel β sheet following strand β7.

There are 11 α helices, of which four pack against the central β sheet. The helices which correspond to the conserved helices A–F of the α,β-hydrolase fold as defined by Ollis *et al.* [2] are: α1 = A, α2 = B, α3 = C, α7 = D, α10 = E and α11 = F.

Active-site cleft

The active site is comprised of Ser87, His286 and Asp264, which form a number of hydrogen bonds. The sidechain of the active-site histidine, His286, is hydrogen bonded to Oγ of Ser87 and to Oδ1 of Asp264. In addition to the hydrogen bond to His286, the sidechain Asp264 hydrogen bonds to one water molecule, the sidechain of Glu289 and the mainchain atoms of Leu266 and Val267. The mainchain carbonyl oxygen of Asp264 hydrogen bonds to the amide nitrogen of His286. The carbonyl oxygen atom of His286 is hydrogen bonded to His86, which in turn hydrogen bonds to Tyr29. Glu289 makes a weak hydrogen bond with a distance of 3.5 Å to His286 and a strong hydrogen bond of 2.73 Å to Asp264. Protein engineering [25] and modeling studies [18] demonstrate that this arrangement of residues allows the functional substitution of Asp264 by Glu289.

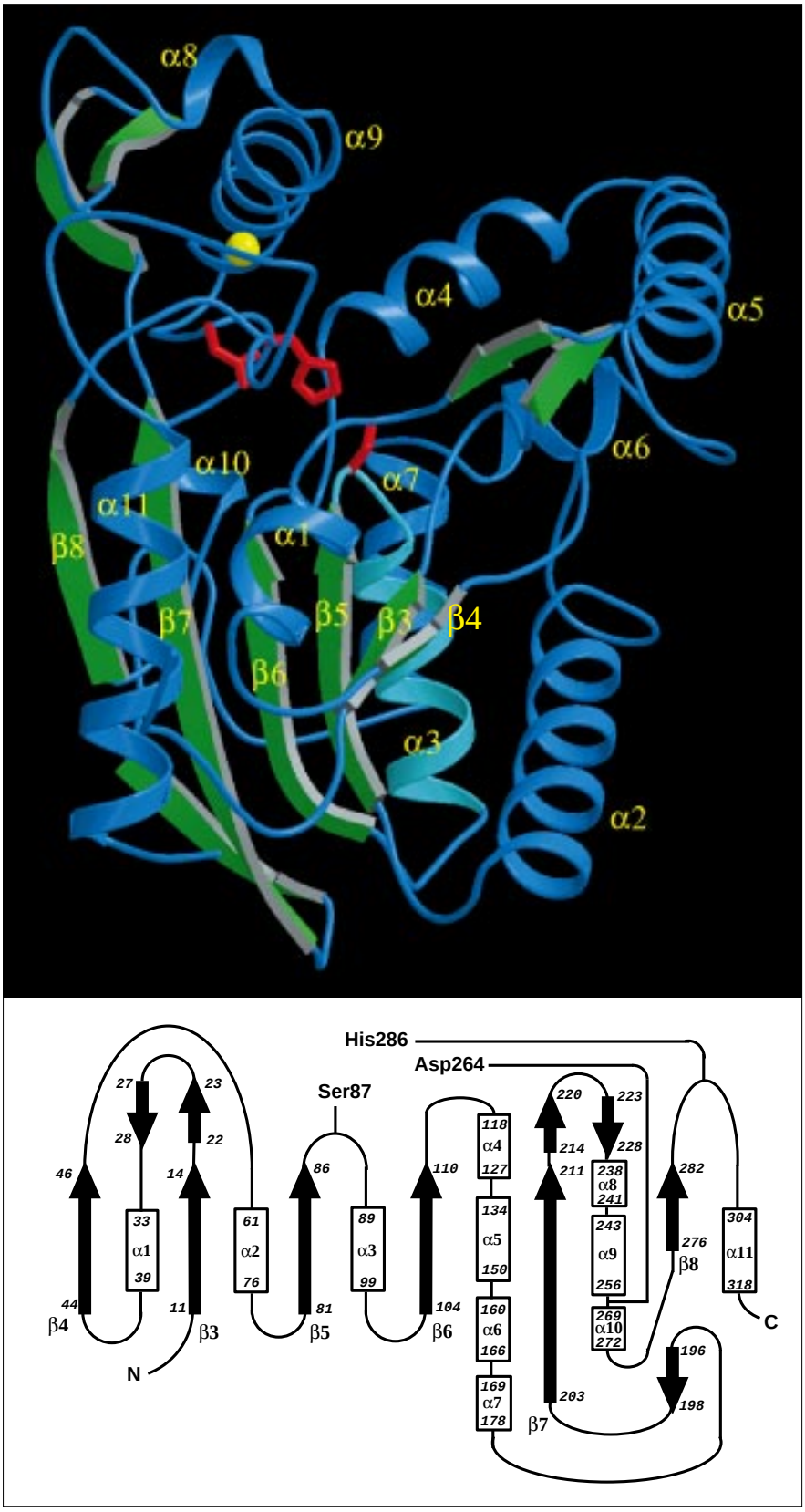
The active site serine lies at the bottom of a cleft in the protein and is exposed to the solvent (Fig. 2). The opening of the cleft is ovoid and is 10 Å × 25 Å across. The cleft is about 15 Å deep. The projection of Leu17 into the cleft above the active-site Ser87 gives the cleft a boomerang shape and divides the active-site cleft into two branches. The floor of the cleft is U-shaped and at the bottom, near Ser87, is about 4 Å wide.

The walls of the cleft surrounding the active site are formed primarily by hydrophobic residues. The cleft is bordered by helices α4, α5, and α9 and by residues 17–29. On one side of Leu17, the base of the cleft is formed by Pro113 and Leu167 and the walls are formed by residues from helices α4, α5, and α9. On the other side of Leu17, the cleft is bounded by residues from loop 17–29 and helices α5 and α9. One can easily imagine that one of the fatty acyl chains of a triglyceride would bind in each arm of the cleft. The direction of the catalytic serine residue suggests that the scissile acyl chain binds in the arm of the cleft formed by helices α4, α5 and α9, which is also the deeper of the two clefts. Superposition of PCL with lipase–inhibitor complexes of *R. miehei*, *C. rugosa* and pancreatic lipases also suggests this placement of the scissile acyl chain.

Oxyanion hole

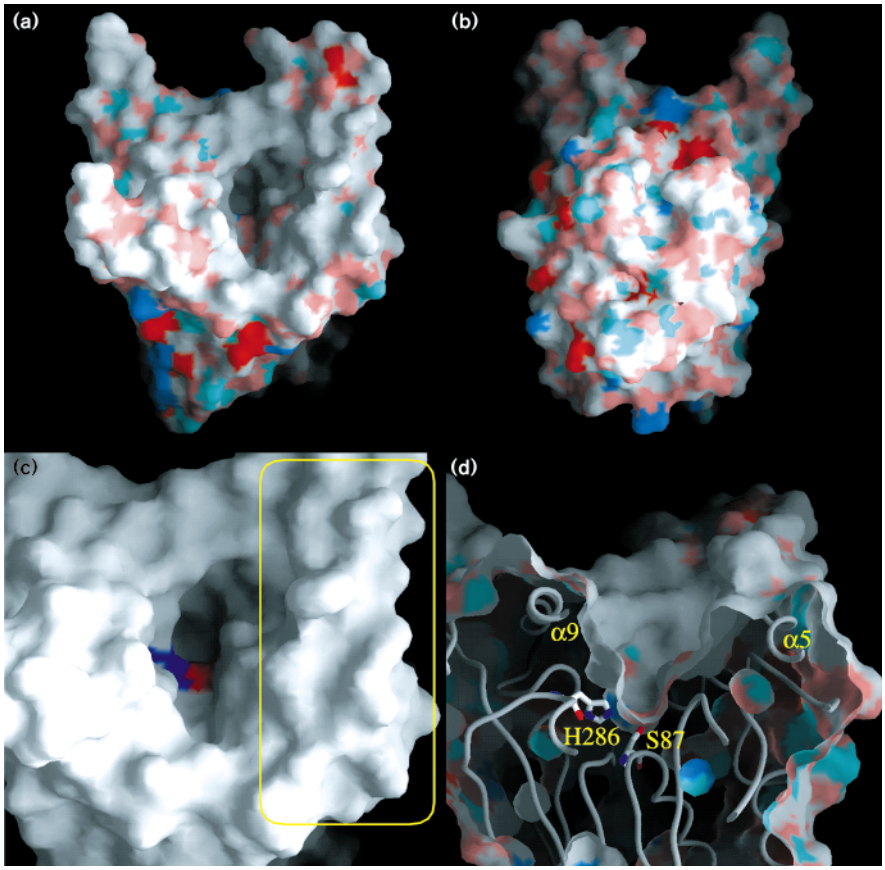
Superposition of PCL with other inhibited lipase structures (e.g. *R. miehei* [8], *Candida antarctica* lipase B [11] and *C. rugosa* [10]) reveals that the amide nitrogen atoms of Gln88 and Leu17 form the oxyanion hole at the C-terminal end of strand β3 and the loop after strand β4. The GBF model includes a water molecule in the oxyanion position. Arg61 is located at the center of a hydrogen-bonding network which stabilizes the region around the oxyanion hole. The guanidinium group of this arginine residue hydrogen bonds to the mainchain carbonyl oxygen atoms of His15, Gly16 and Ser50 and to the sidechain of Asp55. The mainchain amide of Arg61 hydrogen bonds to the carbonyl oxygen of Ser54. Hydrogen bonds between

Figure 1



The overall fold of PCL. (a) The fold of PCL is shown in a ribbon representation with β strands represented as arrows (green) and helices as coils (blue). The position of a Ca^{2+} ion is indicated by a yellow sphere. The sidechains of the catalytic triad residues are shown in red. (Figure produced using the program MOLSCRIPT [43]). (b) Schematic topology diagram of PCL. Secondary structure assignments are according to Kabsch and Sander [44]; helices are shown as rectangles and β strands as arrows. The active-site residues are indicated.

Figure 2



The active-site cleft of PCL. **(a)** The molecular surface colored according to atom type: charged oxygen atoms are red, polar oxygen atoms and carbonyl oxygens are pink, nitrogen atoms are blue and carbon atoms are white. **(b)** The molecule in (a) rotated by 180° on a vertical axis to show the side of the molecule opposite the active-site cleft; atoms are colored as in (a). **(c)** The cleft as viewed from above; helix α5 (lid) is shown boxed. Helix α9 forms the left wall of the cleft and helix α4 forms the upper wall in this view. The exposed surface of Ser87 is shown in red and that of His286 is shown in blue. **(d)** Cutaway view of the active-site cleft roughly perpendicular to the view in (a). The active-site Ser87 and His286 are shown along with a backbone ribbon. Helices α9 and α5 are visible at the left and right of the cleft, respectively. (Figure produced using the program GRASP [45].)

the carbonyl oxygen of Gly16 and the amide nitrogen of Gly19 and between the amide nitrogen of Thr18 and Gly51 also help to stabilize this region. The sidechain of His15 hydrogen bonds to either Leu49 or Ala47.

The position of the amide nitrogen atom of Leu17 in the BRI structure differs in position by 0.7 Å from that in the GBF and P&G structures and the plane of the bond suggests that the direction of the hydrogen in the BRI model would be less favorable for hydrogen bonding to the oxyanion. This difference probably results from the disorder seen for residues 18–27 in the BRI structure.

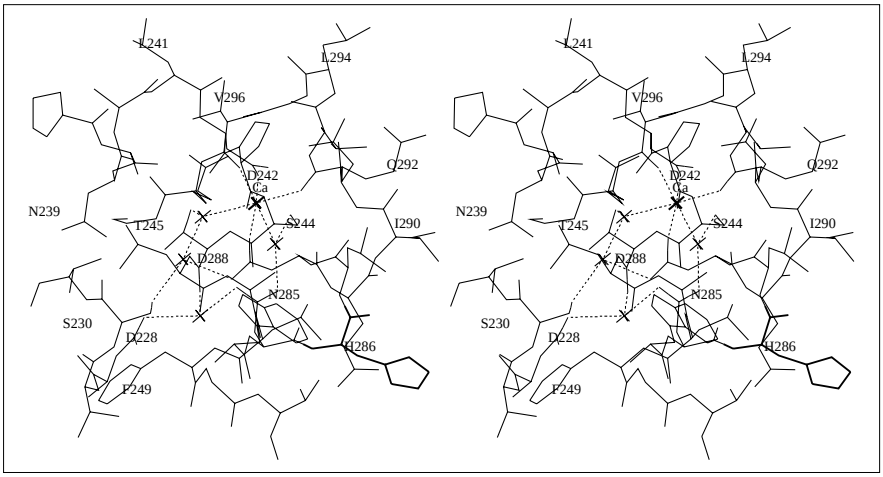
Ca^{2+} site

As reported by Noble *et al.* [17], PCL contains a bound Ca^{2+} ion. The Ca^{2+} ion is six-coordinated, making contacts with four oxygen atoms of the protein and with two water molecules, as reported for CVL by Lang *et al.* [18]. The four protein atoms which coordinate the Ca^{2+} ion include the carbonyl oxygen atoms of Gln292 and Val296 and the Oδ2 sidechain oxygen atoms of Asp242 and Asp288 (Fig. 3). The peptide bond between Gln292 and Leu293 adopts a *cis* conformation; the *cis* conformation was also observed in CVL. This allows a hydrogen bond to form between the amide hydrogen of Leu293 and the sidechain of Asp242. The Ca^{2+} is 14.5 Å from the active-site serine, so it is difficult to envision anything beyond a structural role for this ion. Nonetheless, this ion bridges helix α9 (which forms part of the wall of the cleft) to the loop which positions the catalytic His286 in the cleft and, therefore, could play an important stabilizing role. The Ca^{2+} site in PCL is not similar in location to the Ca^{2+} site in pancreatic lipase. After superposition of the central β sheets of these lipases, the Ca^{2+} ions are nearly 30 Å apart and are approximately equidistant from the catalytic serine residue, but in opposite directions.

Common solvent molecules

There are 60 water molecules common (closer than 1 Å) to all three structures. There are an additional 13 water molecules which are common to the BRI and GBF structures, one which is common to the BRI and P&G structures and 28 common to the P&G and GBF structures. Most of these molecules form the first hydration layer and are distributed around the surface of the molecule. The back side of the protein, opposite the binding cleft, has many solvent molecules. Only 12 of the common water molecules are buried inside the lipase. In addition, two water molecules which

Figure 3



Stereo view of the Ca^{2+}-binding site in PCL. The Ca^{2+} ion is six-coordinated by four protein atoms and two water molecules. Note the *cis* peptide bond which orients the carbonyl group of Gln292 towards the Ca^{2+} ion. Internal water molecules in the region of the Ca^{2+}-binding site are shown as crosses. The catalytic His286 is shown in bold lines; dotted lines represent hydrogen bonds or bonds to the Ca^{2+} ion.

are common to BRI and GBF and one solvent molecule modeled only in the GBF structure are internal.

Most of the internal water molecules are in locations which clearly suggest a role in stabilizing structural elements. All of these internal water molecules are present in the CVL structure as well [18] and their positions correspond to cavities observed in PGL. Two of the internal water molecules coordinate the Ca^{2+} ion (Fig. 3). Two additional common water molecules provide a hydrogen-bonding network which stabilizes the positions of helix α9 and the loop containing the catalytic His286 (Fig. 3). A water molecule modeled only in the GBF structure lies just below the protein surface and helps to stabilize the calcium-binding loops. Common internal water molecules are also located on either side of the central β sheet and help to stabilize secondary structure elements. On one side of the β sheet a water molecule hydrogen bonds to the sidechain of Asp36 and to the carbonyl oxygen of Val305, thereby bridging helices α1 and α11. Other common water molecules on this side of the β sheet fill cavities between helix α11 and strands β6 and β7 and help to anchor the loop which contains the catalytic histidine (Fig. 4a). On the opposite side of the β sheet water molecules help to stabilize the region at the C-terminal end of strands β6 and β7 (Fig. 4b). This region includes the loop containing the catalytic acid, Asp264, and loops which form part of the walls of the substrate-binding cleft.

Finally, common water molecules help to anchor loops which appear to be important in the activation of the lipase (discussed below). One of these water molecules is located at the C-terminal ends of strands β3 and β4, anchoring loop 17–28. Loop 50–53 is stabilized at either end by water molecules: one bridges the backbone atoms of Gln53 and Gly51 to the the sidechain of Gln88 and the other anchors backbone atoms of Asp55 and Ala57 to the N-terminal turn of helix α2.

Discussion

Comparison of PCL with other *Pseudomonas* lipases

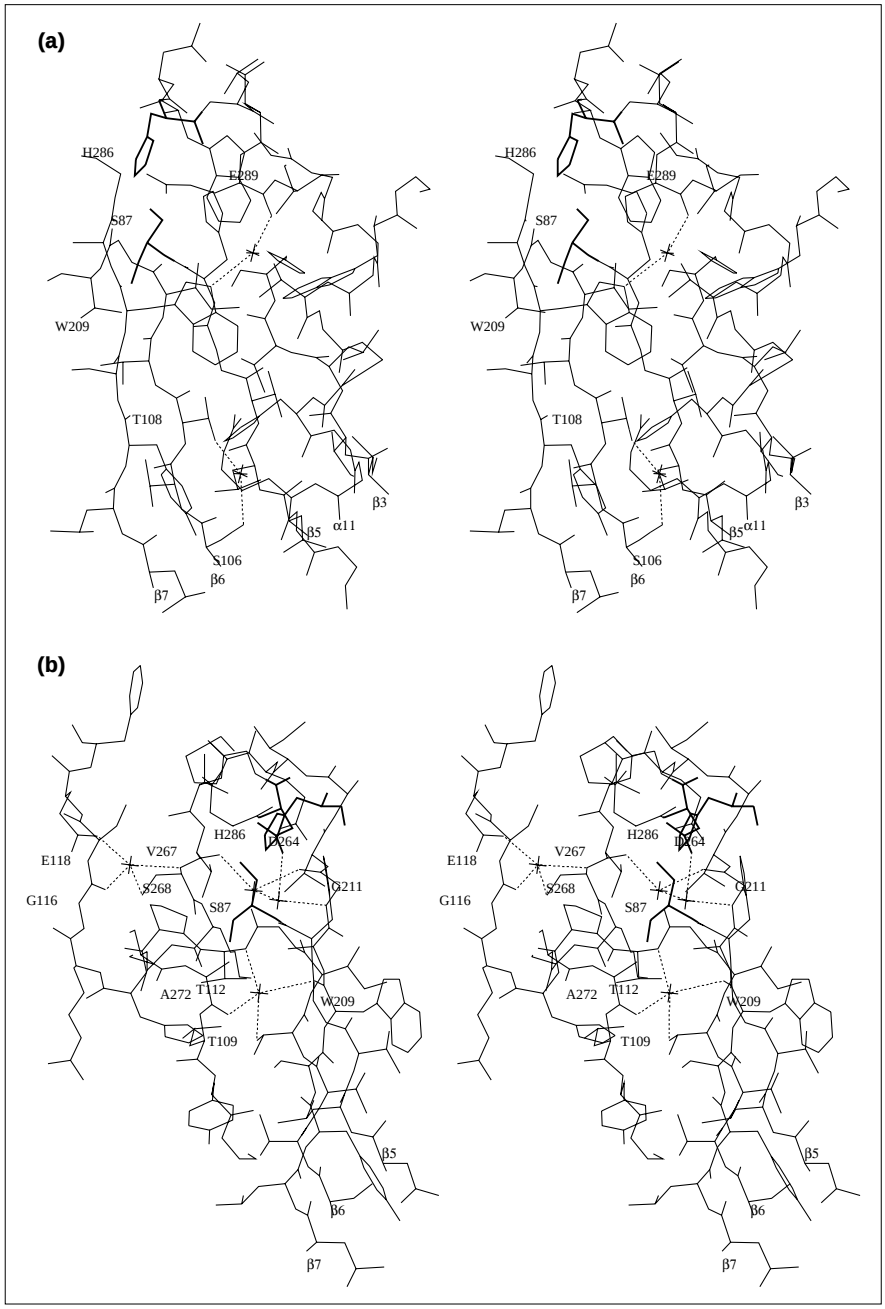
A comparison of the amino acid sequence of either PGL or CVL with that of PCL indicates that 84% of the residues are identical [26]. The lipases of these strains belong to the same subgroup of lipase producing strains from the family of *Pseudomonadaceae*. Of the 50 amino acids which differ in these lipases, 24 are substituted conservatively.

Superposition of PGL and PCL shows that 257 Cα atoms (80% of the total) superimpose with an rms difference in position of only 0.39 Å. The crystal structure of PGL, determined at 3 Å resolution [17], has four molecules in the asymmetric unit. Superposition of these four independent molecules shows five loops whose conformations are somewhat variable and may reflect flexibility in the loops and the influence of crystal packing forces. These loops comprise residues 17–28 (PGL numbering), residues 132–135, residues 150–157, residues 214–228 and residues 232–235. Each of these loops is on the surface of the protein. Pairwise comparisons of the four independent molecules give rms positional differences for backbone atoms of loops 132–135 and 232–235 of about 1 Å and 1.5 Å, respectively. For three of the molecules, the positional differences in loop 215–224 are also about 1.5 Å; in the fourth molecule the position of this loop is shifted by about 4 Å. The positions of the backbone atoms of loop 17–29 vary from one molecule to another by more than 2 Å and loop 150–156 varies by more than 3 Å.

The amino acid sequence of CVL is identical to that of PGL and the crystal structure of CVL provides a fifth independently determined view of the molecule. The structural

Figure 4

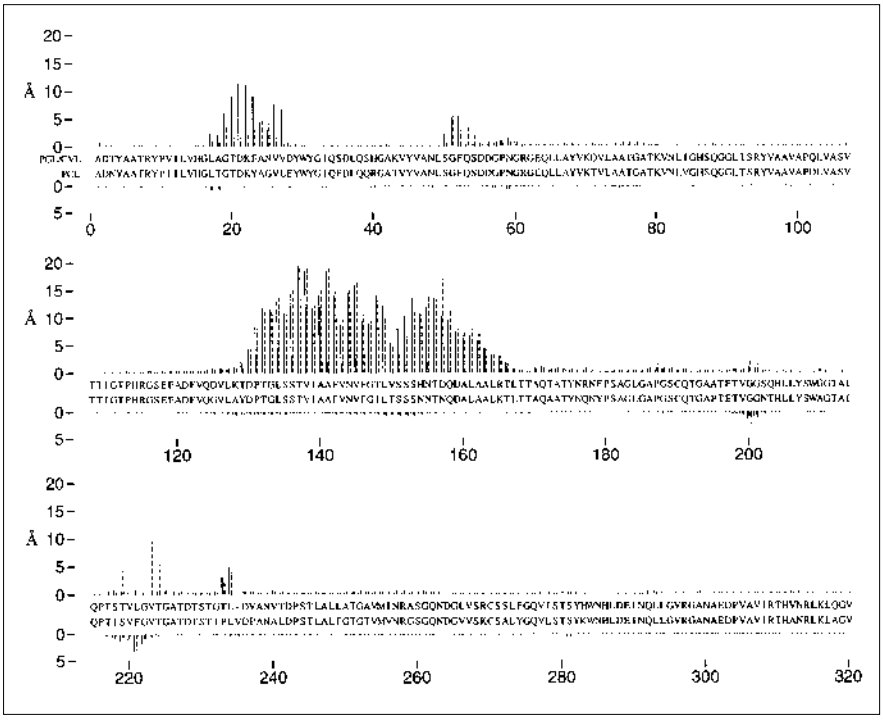
Common internal water molecules. Catalytic triad residues are marked with bold lines and hydrogen bonds by dashed lines; water molecules are shown as crosses.
(a) Structural water molecules found on the concave side of the central β sheet.
(b) Structural water molecules found on the convex side of the central β sheet.



differences between CVL and PGL have been described previously [18] and, interestingly, these differences involve the same regions of the protein that differ among the PGL monomers. The magnitude of the positional shifts of some of these loops in CVL are significantly larger than the variation observed in PGL. Loops 17–28 and 115–224 of CVL differ in position from PGL by as much as 10 Å, residues 150–156 differ by as much as 6 Å, and residues 132–135 differ by about 2 Å. In addition, residues 49–54 differ in position by as much as 4 Å and helix 137–147 (137–149 in PGL) is shifted by an average of about 3 Å, although the active site remains occluded. As residues 220–222 are missing in the CVL structure, possibly a result of proteolytic cleavage [18], the flexibility of loop 214–228 is easily understood. The other structural differences between PGL and CVL are less easily explained, but may be related to pH differences in the crystallization conditions [18].

Comparison of PCL with PGL or CVL emphasizes the flexibility of these same regions of the protein (Fig. 5).

Figure 5



Positional differences of Cα atoms between *Pseudomonas* lipase structures. The differences between PCL and PGL/CVL are shown above the sequences of PCL and PGL/CVL: differences between PCL and molecule A of PGL are represented by solid lines and differences between PCL and CVL are represented by dashed lines. Below the sequences are the pairwise differences in Cα positions of the three independent models of PCL. Under each residue are three lines representing the differences between the GBF and P&G models, the GBF and BRI models, and the BRI and P&G models, respectively, from left to right. The P&G model does not contain Ala1 and residues 19–27 were omitted from comparisons with the BRI model as these residues were disordered in the BRI crystal structure.

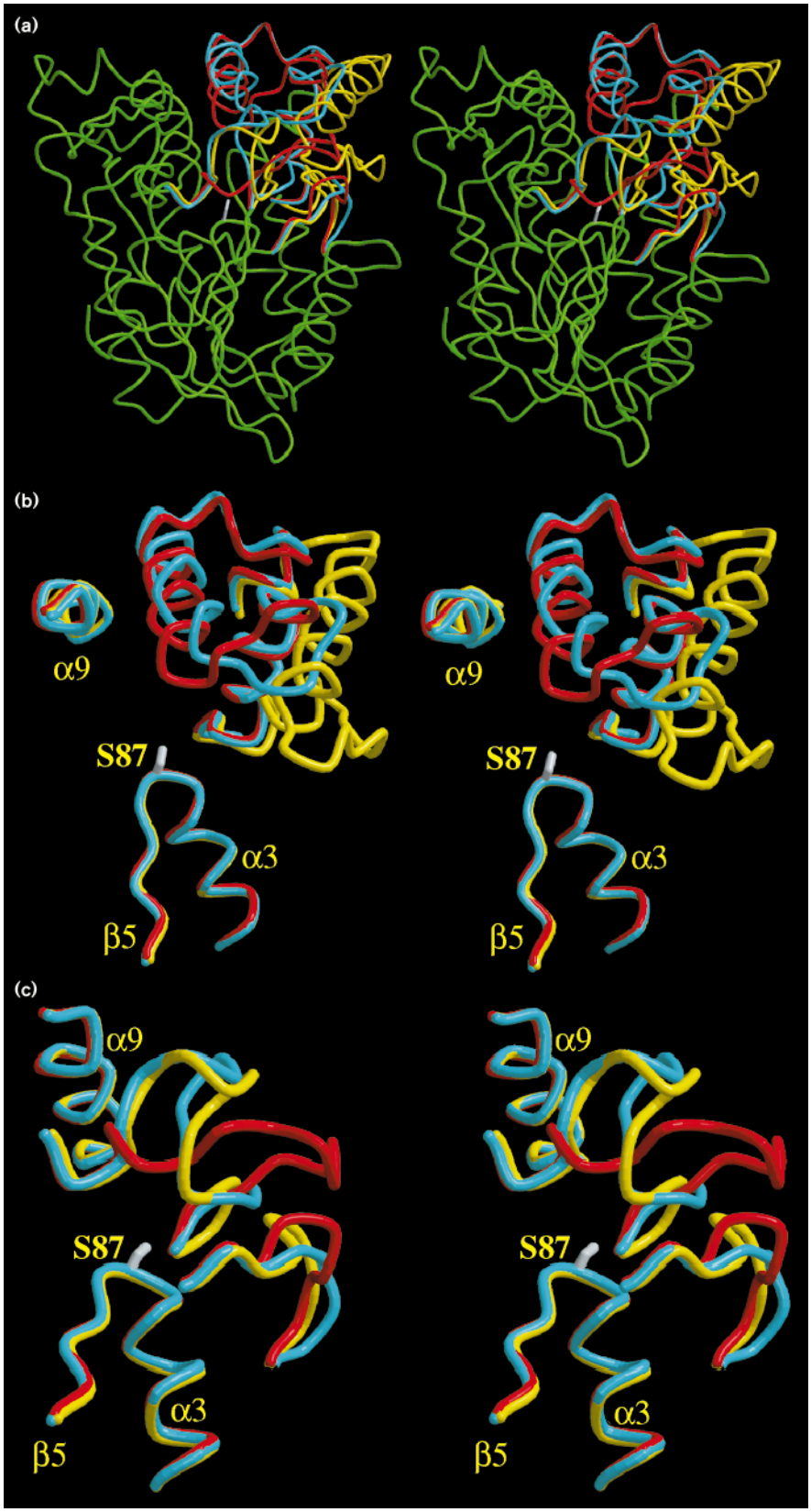
Superposition of CVL [18] and PCL shows that 258 Cα atoms superimpose with an rms difference in position sof only 0.5 Å. The major structural differences involve three adjacent loops in the region of the active site. The largest differences between PCL and CVL involve a large omega loop encompassing residues 128–166. The N-terminal hinge point for the movement of this loop is marked by sharp differences in the ϕ,ψ angles of Thr129, which shift from ϕ = –127°, ψ = –36° in the CVL structure to ϕ = –78°, ψ = –21° in the PCL structure. Differences in the ϕ,ψ angles at Thr166 mark this residue as the C-terminal hinge point. Some plasticity in this omega loop is apparent, as the differences between CVL and PCL involve conformational changes in addition to large positional shifts. The largest positional shift is that of helix α5 (residues 137–147 in CVL, 137–149 in PGL), which is displaced in PCL (relative to CVL) by as much as 20 Å, with an average displacement of about 14 Å (Figs 5,6). In addition to this positional shift, helix α5 is extended in PCL at both the N and C termini. Residues 134–136, having no regular secondary structure in CVL, form the N-terminal turn of helix α5 in PCL. Helix α5 extends to residue 150 in PCL, almost one full turn longer than that of CVL. The conformations of the residues following helix α5 are very different in PCL and PGL or CVL. The positional differences of these residues range from about 4 Å for residue 151 to about 17 Å for Asn157. Residues 156–162 form helix α6 in CVL, whereas the comparable helix in PCL involves residues 160–166. At residue Thr166 the rms deviation once again drops below 2 Å.

The structural rearrangement of this large omega loop is correlated with the structural differences observed in the regions Leu17 to Val26 and Ser50 to Ser54 (Figs 5,6). The positions of the mainchain atoms of residues 17–26 of PCL are more similar to CVL than PGL, but still differ by as much as 4 Å with an average rms deviation observed for the Cα positions of these residues of 3 Å. The conformation of this loop is similar in CVL and PCL, but its position is shifted. The amide nitrogen of Leu17, which forms part of the oxyanion hole, is shifted by 0.74 Å but maintains the same orientation as in the CVL structure. The ϕ,ψ angles of residues 17–20 suggest a lower conformational energy for this oxyanion loop in PCL. Differences in the mainchain conformations of residues 18–20 direct some hydrophobic sidechains of these residues towards the active site, contributing to the hydrophobicity of the active-site cleft. The movement of this loop closes a hydrophilic trough which was observed in the CVL structure [18]. Several water molecules located in this trough in CVL are excluded by mainchain atoms in PCL.

In addition, the mainchain conformation of residues 50–54 of PCL is different from either of the conformations observed in PGL or CVL, their positions differing by as much as 4–6 Å. The conformational rearrangement of this

Figure 6

Stereo views of the positional and conformational differences of important loops. **(a)** Cα trace of PCL with the omega loop, the oxyanion loop, and loop 47–55 highlighted in yellow. The same loops of PGL and CVL are shown in red and blue, respectively. The sidechain of Ser87 is shown in gray. **(b)** The large omega loop encompassing helix α5. The nucleophile elbow, the sidechain of Ser87, and helix α9 (which forms one wall of the binding cleft) are included for reference. The color scheme is as in (a). **(c)** Residues 16–31 encompassing the oxyanion loop (center) and residues 47–55 (bottom right) are shown. The color scheme is as in (a). (Figures produced using the programs MOLSCRIPT [43] and Raster3D [46].)



loop directs the hydrophobic sidechain of Phe52 towards the active-site cleft. The position of this loop is stabilized by hydrogen bonds to Arg61, which also provides hydrogen bonds to the oxyanion loop 17–26. The guanidinium group of Arg61 is shifted in position by ~2.8 Å and the hydrogen-bonding pattern differs between PCL and CVL. In PCL, Arg61 forms a number of hydrogen bonds: NH1 hydrogen bonds to the carbonyl oxygen atoms of Ala50 and Gly16; NH2 hydrogen bonds to the carbonyl oxygen of His15 and to the sidechain oxygen atom of Thr92; and Nε hydrogen bonds to the sidechain of Asp55. In contrast, the pattern observed in CVL is: NH1 hydrogen bonds to the carbonyl oxygens of Gly51 and Leu17; NH2 is bridged to the carbonyl oxygen of His15 by a water molecule; and Nε hydrogen bonds to the carbonyl oxygen of Gly51.

Two of the water molecules common to all three structures of PCL and CVL appear to stabilize secondary structures which provide an anchor for these loop movements. One of these water molecules bridges strands β3 and β4 and is located near Leu17 and Leu49 (the apparent pivot points for two of the loops which move). The other water molecule appears to stabilize Gly51 and Gln53 which seem to be the pivot points of the mobile loop region 50–54.

Other regions of PCL show displacements of residues by more than 2 Å from their equivalent positions in PGL or CVL, but these displacements do not appear to be related to the catalytic function of the enzyme. Residues 233–234 of PCL are displaced by about 3–4 Å, probably due to the fact that these residues immediately precede the one residue insertion in the sequence of PCL (PCL numbering: PCL 1–234 = PGL/CVL 1–234; PCL 235 has no PGL/CVL equivalent; PCL 236–320 = PGL/CVL 235–319). The electron density for residues 233–234 and the inserted Val235 is very clear and the positions of these residues are not in question. Leu234 makes hydrophobic contacts with a symme-try related molecule. Residues 219, 223 and 224 of PCL differ from their positions in CVL by more than 4 Å, but this is undoubtedly related to the three-residue deletion of residues 220–222 in CVL. The Cα position of Gly200 differs from its counterpart in PGL/CVL by 2.5 Å. This residue is a part of a Gly-Gly sequence which is likely to be very flexible and is near a symmetry-related molecule. The difference in conformation may result from flexibility in the mainchain of this loop or from crystal packing forces. None of these loops are adjacent to the active site and the differences in these regions are, therefore, not likely to be of any catalytic significance.

The conformational differences between PCL and PGL/CVL can not be explained in terms of amino acid sequence differences. The sequences around loop 50–53 are identical (Fig. 5) and the differences between PGL and CVL, with identical sequences, in the conformation of loop 17–28 is greater than the difference between CVL and PCL. Three of the five sequence differences between PCL and PGL/CVL in this loop are conservative substitutions (Phe/Tyr, Val/Leu, and Asp/Glu). Only seven sequence differences are found in the large omega loop which displays the largest conformational differences. Two of these are the first two residues of the loop and two others are conservative substitutions. The only sequence differences in helix α5, which occludes the active site in PGL/CVL, are Thr/Ile148 and Val/Thr150. Only small differences in the contacts of these residues are observed and these small differences result from changes in the polypeptide backbone conformation in this region and not from the sidechain differences.

Other regions of the protein (e.g. residues 249–283) have invariant conformations despite a higher frequency of amino acid substitutions than that observed in the large omega loop.

Molecular surface

Although the sequences of PGL/CVL and PCL are not identical, comparison of their molecular surfaces is instructive. As many of the sequence differences between PGL/CVL and PCL result from conservative substitutions, the comparison of surface areas probably provides a good description of the differences between the open and closed conformations rather than merely representing differences between two different lipases. The total surface area of PCL is about 800 $Å^2$ larger than that of CVL. In addition, the molecular surface of PCL is more hydrophobic than that of PGL/CVL. The molecular surface area contributed by polar and charged residues in PCL is 1000 $Å^2$ less than in CVL, whereas the molecular surface from hydrophobic and aromatic residues is 1700 $Å^2$ greater in PCL than in PGL/CVL. The same trend is seen if the comparison is based on atom type rather than residue type. The exposed surface of carbon atoms is 1000 $Å^2$ greater in PCL, whereas the combined area of nitrogen and oxygen atoms is decreased by 300 $Å^2$. These comparisons clearly indicate that the open conformation of PCL exposes a large hydrophobic surface which is not accessible in the closed conformation of PGL/CVL. This parallels the observations for *R. miehei* lipase [8], pancreatic lipase [9] and *C. rugosa* lipase [14].

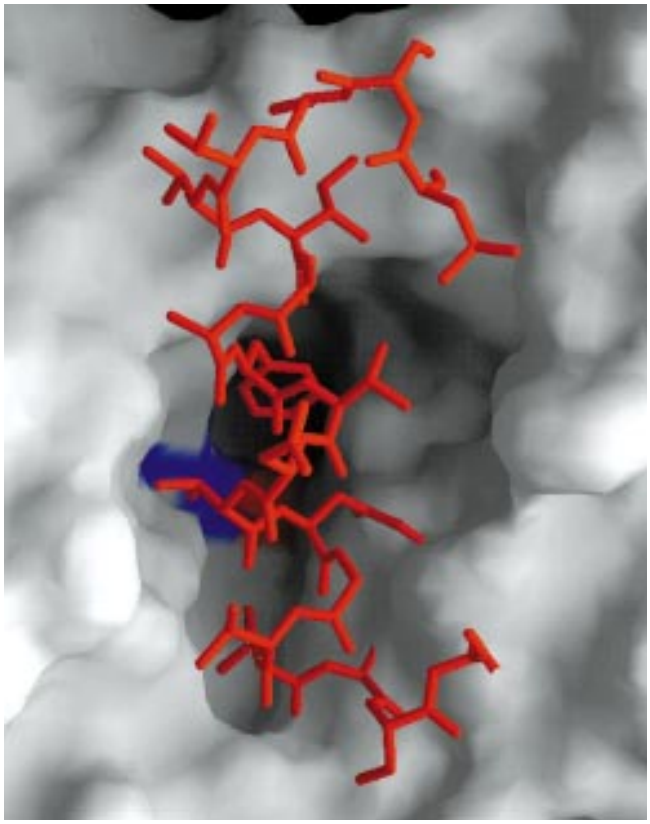
The increased hydrophobic surface in the open conformation of PCL is localized mainly to the cleft surrounding the active site. Of all the residues which line the cleft, only three differ between PGL/CVL and PCL: Tyr23 of PGL/CVL is a conservatively substituted phenylalanine in PCL; Leu266 in PGL/CVL is replaced by valine in PCL; and Thr150 in PCL is a valine residue in PGL/CVL. The nature and shape of the cleft are, therefore, well conserved.

The movement of the lid in PCL also masks a number of hydrophilic groups resulting in a decrease in the exposed

hydrophilic surface relative to PGL/CVL. Many of the residues along helix α5 have either mainchain or sidechain atoms, or both, masked by contacts with other parts of the molecule. Included in this group are Ser135 (carbonyl oxygen), Thr137 (Oγ1), Ala140 (carbonyl oxygen), Asn144 (carbonyl oxygen and Oδ1), and Gly147 (carbonyl oxygen). Other residues with a drop in exposed hydrophilic surface are Ser151, Ser152, Ser153, and Gln158; loop 50–54 also shows a large decrease in exposed surface. Many of the polar mainchain atoms of these residues become buried as a result of the movement of the lid.

Figure 7 shows a superposition of the closed conformation of PGL and the molecular surface of PCL. It is immediately obvious from the superposition that helix 137–149 occludes the binding cleft. It is also apparent that the residues of helix α5 display very good surface complimentarity with the wall of the cleft which is formed by helix α9. This complimentarity should be expected as only residues 148 and 150 differ between PGL/CVL and PCL. Residue 148 lies on the surface leaving ample space for the substitution of an isoleucine sidechain for a threonine sidechain. Residue 150 varies from threonine in PCL to valine in PGL/CVL. As the sizes of these sidechains are similar, no big difference in the packing should be expected.

Figure 7



Comparison of the open and closed conformations. Helix α5 of PGL is superimposed onto the molecular surface of PCL. The active site is colored as in Figure 2c. (Figure produced using the program GRASP [45].)

Crystal contacts in the lid

Recently, the three-dimensional structures of *Humicola lanuginosa* lipase and *Rhizopus delemar* lipase, both closely related to the *Rhizomucor* lipase, were described [15,27]. The lid of the *Humicola* lipase was disordered, suggesting that it was mobile. This mobility was attributed to the absence of any crystal interactions with residues of the lid. The asymmetric unit of the *R. delemar* lipase crystals had one molecule whose lid was in the closed conformation and one whose lid was in a position intermediate between closed and open. Based on these observations, the investigators suggested that the crystal-packing forces stabilize one of many conformations available to the protein in solution and that the two conformation explanation of interfacial activation is an oversimplification. However, for the lids of the *Rhizopus*, *Humicola*, and *Rhizomucor* lipases, a single loop moves and both the length of this loop and the observed range of its motion are small compared to those of other lipases. In contrast, the conformational differences observed between PGL, CVL and PCL involve three loops and positional changes of 10–20 Å. Although this suggests that the differences observed are not simply the result of stabilization of one particular conformation by crystal-packing interactions, observation of the crystal-packing interactions of the lid is, nonetheless, of interest.

The lid of PCL is involved in crystal-packing interactions which are almost entirely van der Waals contacts. The largest contact area involves interactions of helix α4 with helix α9 and the two antiparallel β strands encompassing residues 214–228 of the symmetry-related molecule. The only hydrogen bonds in this interaction are provided by several bridging water molecules. Helix α5 contacts its equivalent from a second symmetry-related molecule, again involving only van der Waals contacts. The helices from the two molecules cross one another so that the sidechains of residues Val145 and Phe142 interdigitate with their symmetry equivalents. Phe146 and Leu149 also contact their symmetry mates. A small contact area between Pro131 and a third symmetry-related molecule is also observed.

The crystals of PGL have four molecules per asymmetric unit, so there are contacts between independent molecules as well as contacts with symmetry-related molecules. Each of the four molecules shows different contacts with neighboring molecules. Some, but not all of the molecules have contacts involving residues in the lid: these involve primarily residues 21–28, 130–135, and 150–155. In some cases there are also contacts involving residues 50–52. After superimposing each of the four independent molecules, one sees that the rms difference in position for these loops ranges from 1–4 Å. Residues 230–234 are frequently involved in

intermolecular interactions and show a similar rms deviation in position of 1–4 Å.

The crystals of CVL have one molecule per asymmetric unit. The loop regions at either end of helix α5 (the lid), residues 130–135 and residues 150–158, are involved in crystal packing contacts. The contacts involving residues 130–135 are largely van der Waals interactions, whereas residues 153–158 make numerous hydrogen bonds with their counterparts in the symmetry-related molecule. Asp144 in the middle of helix α5 interacts with residues 235–239 of another molecule, but the hydrogen-bond distances of 3.5 Å suggest only weak hydrogen-bonding interactions. The Gln53 sidechain also hydrogen bonds to the mainchain carbonyl atoms of a symmetry-related molecule.

The rms deviations of only 1–4 Å among the four molecules in the PGL crystals suggest that crystal-packing forces alone are not likely to cause the large (>10 Å) differences in conformation observed when comparing PCL to PGL/CVL. This observation also suggests that these differences may reveal important information regarding the activation of the lipase. One interpretation of the PGL, CVL and PCL structures is that they represent sequential stages in the transition between the closed and open conformations of the lipase. Based on the change in orientation of the oxyanion loop and the partial movement of helix α5, Lang *et al.* [18] have previously suggested that the CVL structure represents an intermediate in the transition between the closed and open conformations. Movement of helix α5 (of the large omega loop) to expose the active site together with changes in the oxyanion loop and loop 50–54 to expose hydrophobic sidechains in the active-site cleft result in the open conformation, as exemplified by the PCL structure. Whether these structures represent sequential stages or not must remain speculative, but the activation of these lipases must certainly involve concerted conformational changes in each of these three adjacent loops. The comparison of PCL with PGL clearly demonstrates that the open position of helix α5 can not be attained without adjustments in the conformation of the oxyanion loop. Movement of loop 50–54 from its position in CVL is essential in order for helix α5 and residues 151–157 (near the C terminus of the omega loop) to attain the position observed in PCL. The differences observed in the four molecules of PGL suggest that a fraction of the differences observed between PGL and CVL or PCL are related to crystal-packing forces and other similar conformations are probably accessible in solution. However, comparison of the three crystal structures clearly demonstrates that the conformations of these three loops are interrelated and the magnitude of the differences suggests that they are likely to control the activation of these lipases.

The crystal structures of PGL and CVL together with the three independent determinations of the PCL structure, reported here, provide eight independent representations of these lipases, which share at least 80% sequence identity. This large number of independent models, available for no other lipases, obtained under different crystallization conditions provides a unique opportunity to examine the effects of crystallization conditions on the conformation of the protein. The closed conformation seen in PGL was obtained in crystals grown from high concentrations of PEG, although some organic solvent and detergent were present [17]. Reduction of the PEG concentration, also in the presence of organic solvent and detergent, produced a slightly different closed conformation in CVL [18]. Finally, crystallization of PCL from a high concentration of organic solvent resulted in an open conformation. In conjunction with the arguments presented above, suggesting that crystal-packing interactions contribute at most only a small portion of these conformational differences, these results strongly suggest a relationship between the solution conditions and the conformation observed. In aqueous solution, the closed conformation is preferred, whereas in a more hydrophobic environment the open conformation, with its hydrophobic cleft surrounding an exposed active site, is preferred. As both the open and closed conformations of the bacterial lipases have been obtained at both pH 6.5 and pH 8–9, we conclude that pH is not likely to be the principal factor which determines the conformation. The influence of organic solvents and detergents in producing open conformations might suggest that the dielectric constant is a more likely trigger for the movement of the lid.

Despite the advances in understanding lipase function provided by the many structural studies in recent years, there are still many questions left unanswered. The molecular basis of substrate specificity of the various lipases is still largely unknown. Some progress towards understanding stereoselectivities has been made [28], but much remains to be investigated. The interfacial activation appears to result largely from conformational changes. However, the substrate selectivities and stereoselectivities observed are probably the result of a complex interplay of enzyme and substrate, perhaps involving structural changes in both the enzyme and the substrate. The determination of structures of the lipases in solution would help to determine the flexibility and conformations of the lids. Knowledge of the structure of the lipase bound to a lipid interface would also help to define its interaction with the lipid, possibly providing clues towards understanding the molecular bases of substrate and stereoselectivities. Studies of this nature would help to define the relative importance of structural changes in the enzyme versus structural changes in the substrate to lipase activity.

Biological implications

For many years the defining characteristic of lipases was their greatly enhanced activity when the substrate is presented at an oil–water interface, a phenomenon termed

'interfacial activation'. We have only begun to understand this phenomenon in recent years. The three independently determined crystal structures of the open conformation of *Pseudomonas cepacia* lipase described in this report together with the closed conformations reported for the closely related *Pseudomonas glumae* [17] and *Chromobacterium viscosum* [18] lipases provide eight independent three-dimensional structures of bacterial lipases. Comparison of these structures and the crystallization conditions used clearly demonstrates that the conformations of multiple loops surrounding the active site are interrelated and change in response to the solvent conditions. In aqueous conditions the closed conformation is favored and in the presence of organic solvents the open conformation is preferred. These structures indicate that a large part, if not all, of the interfacial activation of *Pseudomonas* lipases results from a conformational change in the enzyme. Similar activation mechanisms were deduced for *Rhizomucor miehei* lipase [8], pancreatic lipase [9], and *Candida rugosa* lipase [10,16]. In contrast, *Candida antarctica* lipase B [11] and cutinase [12] do not have occluded active sites and are not interfacially activated. Conformational changes, thus, appear to be crucial in interfacial activation of lipases. In the closed conformations of interfacially activated lipases, the active sites are occluded by α helices. Major positional displacements of the surface loops encompassing these helices result in exposure of the catalytic serine residues, which are surrounded by a hydrophobic cleft in the enzyme. Similar to the pancreatic lipase [9] and to the postulated rearrangement of the *Geotrichum candidum* lipase [16], multiple loops in the *Pseudomonas* lipases undergo positional shifts. As observed for pancreatic lipase and *C. rugosa* lipase [14], the positional movements of the occluding loops of the *Pseudomonas* lipases are accompanied by conformational rearrangements of some residues in the loops. The parameters which most influence the conformations of lipases remain uncertain, but the dielectric constant of the solution is one strong possibility.

Materials and methods

Determination of the BRI model

P. cepacia lipase was purchased from Genzyme (Cambridge, MA) and used without further purification. The lyophilized enzyme was dissolved in water to a protein concentration of 16–20 mg ml^{-1} and dialyzed against water overnight at 4°C. The dialyzed protein was crystallized from 50% n-propanol in 50 mM Tris buffer, pH 8.5 by either hanging-drop or sitting-drop vapor diffusion at 18°C. These conditions were adapted from the crystallization conditions reported for *P. fluorescens* lipase by Larson *et al.* [19]. The crystals had unit cell dimensions of a = 91.5, b = 47.3, c = 85.2 Å and β = 121.25° and belonged to space group C2.

All diffraction data were collected on a RAXIS-IIC image plate area detector mounted on a Rigaku RU-300 rotating-anode X-ray generator operated at 50 kV and 100 mA. Copper Kα radiation was monochromated by a graphite crystal. The crystal to detector distance was 99.5 mm. Oscillation images were collected using an angle of 1.8° and a data collection time of 36 min per frame. All data used in the structure determination were collected from a single crystal. Data were reduced and merged using the RAXIS software. A total of 24 516 observations with $I > \sigma I$ were reduced to 14 497 unique reflections to a resolution of 2.1 Å with an R_{merge} of 6.03%. Table 2 summarizes the quality of the data.

The structure was solved by molecular replacement using the coordinates of PGL (Protein Data Bank [29] code 1tah) as a starting model. The rotation and translation functions were calculated using data from 10–4 Å and 10–3 Å, respectively, in X-PLOR 3.1 [30]. Calculation of the rotation function followed by PC refinement of the top 190 rotation peaks gave a clear solution with a correlation coefficient of 0.281, whereas the next highest peak had a correlation coefficient of 0.077. The translation search also gave a clear solution and the rotated and translated PGL coordinates gave an R factor of 48% and a free R factor of 48%. Refinement of the rotated and translated PGL coordinates in X-PLOR with 8–3 Å data consisted of rigid-body minimization, conjugate gradient minimization, simulated annealing using a slow cooling protocol followed by conjugate gradient minimization, and B factor optimization again followed by conjugate gradient minimization. A $3F_o–2F_c$ map was calculated and those loops in the model which did not fit the density were removed. The largest loops removed included residues 129–166, 217–239, and 17–28. Three small loops, 35–37, 50–52, and 75–76, were also removed from the model. Removal of these loops reduced the R factor from 40% (free R factor 42%) to 35% (free R factor 40%). Another refinement cycle omitting the rigid-body refinement was done and subsequently the model was extended to include as many of the missing loops as possible using the program O [31] according to $3F_o–2F_c$ and $F_o–F_c$ density maps. At the same time the amino acid sequence was corrected. The density map was clear enough to distinguish sequence differences, not only between PGL and PCL, but also between the two PCL amino acid sequences, as reported by Gilbert [26]. The density clearly fits the sequence of *Pseudomonas* M-12-33 [26] with the exception of residue 283: the density clearly indicates that this residue is lysine rather than leucine. The calculated molecular

Table 2

Data collection and refinement summary.

Data	BRI	P&G	GBF
No. of reflections	14 497	14 416	17 688
No. of observations	24 516	35 247	33 550
R_{merge} (%)	6.0	3.9	9.5
Resolution (Å)	2.1	2.2	2.0
Completeness (%)	79.4	91.8	83.3
Refinement program	X-PLOR SHELXL-93	X-PLOR PROLSQ	X-PLOR
Resolution range	8–2.1	7–2.2	8–2.0
No. of reflections	12 694	13 749	-
R factor (all data)*	17.5	15.2	18.8
Free R factor†	22.0	20.3	-
R factor $F > 2\sigma F$	15.8	14.0	-
Deviations from ideality			
bond length (Å)	0.008	0.013	0.02
bond angles (°)	1.50	1.87	1.97
dihedral angles (°)	23.04	25.36	22.37
improper angles (°)	1.23	5.24	1.59
B factors‡			
B value model	isotropic	isotropic	isotropic
mainchain	19.22	16.95	11.3
sidechain	20.09	19.13	12.2
all protein	19.61	18.24	11.7
solvent	29.22	31.94	28.2

*R factor = $\Sigma || F_{obs} | - | F_{calc} || / \Sigma | F_{obs} |$. †The free R factor was calculated using 10% of the data. ‡Values given are mean values.

weight of the published sequence is 33 131 Da and the mass measured by ion spray mass spectrometry is 33 140 ± 3 Da, consistent with the expected 15 Da mass difference between leucine and lysine at residue 283. The resolution was extended to 2.1 Å in subsequent refinement cycles which were followed by manual refitting of the model. The R factor dropped to 16.3% for data between 8 and 2.1 Å resolution with $I > 1.7\sigma I$.

Residues 19–27 are poorly ordered and only the general direction of the chain in this region is suggested by fragmented density. These residues are included in the final model as alanine or glycine with occupancies of 0.2 merely to indicate a tentative location for this loop. Lys22 is included as lysine as density which could correspond to its sidechain is observed. Accordingly, a slightly higher occupancy of 0.4 has been assigned to this residue. Additionally, residues 219–223 are poorly ordered and, though the direction of the mainchain is clear, no sidechains are visible for these residues. They have been included in the final model as glycine residues with occupancies of 0.5. The final model also includes 91 water molecules and a Ca^{+2} ion. The final refinement statistics are listed in Table 2.

Determination of the P&G model

The Amano A PCL was obtained from the Amano International Enzyme Co. of Japan. The commercial enzyme powder contained approximately 50% glycine by weight, which was removed using a Pharmacia PD-10 column. Prior to crystallization, the enzyme was lyophilized and redissolved in double-distilled water to a protein concentration of 16 mg ml^{-1}. Diffraction quality crystals were obtained by hanging-drop vapor diffusion in Linbro plates at room temperature using the crystallization conditions reported for *P. fluorescens* lipase by Larson *et al.* [19] as a guide. The droplets were made by mixing 5 μl of the enzyme stock with 5 μl of the reservoir solution which contained 21–24% n-propanol, 0.2 M sodium citrate, and 0.1 M MES, pH 6.5 and were suspended over a 1 ml reservoir. The crystals obtained had cell dimensions of a = 90.5, b = 47.0, and c = 85.0 Å with β = 121.5° and were of space group C2.

All diffraction data were collected by the screenless rotation technique using a MAR Research 180 mm diameter imaging plate detector system mounted on a Rigaku RU-200 rotating-anode X-ray generator. The generator was powered to 50 kV and 180 mA and the CuKα radiation was monochromated with a graphite crystal. The native lipase intensity data were collected from a single crystal using a crystal to detector distance of 100 mm. Each of 125 frames of data were exposed for 2 min with a 1° crystal rotation (φ) per frame. These data were reduced and merged using the XDS software package [32] and the data quality statistics are summarized in Table 2. A total of 35 247 observations reduced to 14 416 unique reflections to a resolution of 2.1 Å with a merging R factor on intensities of 3.9%.

The multiple isomorphous replacement (MIR) method was initially used in the structure determination. The heavy-atom salts were dissolved in the aforementioned crystal growth reservoir solution. The mercury derivative was prepared by soaking a crystal overnight in 2 mM K_2HgI_4 in reservoir solution and the platinum derivative was made by soaking a crystal in 10 mM K_2PtCl_6 in reservoir solution for 13 days. Both soaks were conducted at 4°C. The *N*-iodosuccinimide anhydride derivative was soaked for 24 h in the crystal reservoir solution, made 100 mM with the iodo compound. This solution was then removed from the crystal and the crystal was washed four times with fresh reservoir solution. All derivative X-ray intensity data were collected similarly to the native lipase data. Phase calculations were conducted using the HEAVY program package [33].

The mercury derivative was used to obtain the initial single isomorphous replacement phases, which allowed the identification of the three Pt sites using difference Fourier techniques. The phases resulting from these two derivatives (mean figure of merit, <m> = 0.6, 12–2.8 Å resolution) were used to identify the sites in the iodinated derivative. Calculation of phases using all three derivatives including anomalous data yielded <m> = 0.72 for 12–2.8 Å data.

Shortly after the calculation of the three derivative MIR map was completed, the coordinates of the PGL became publicly available from the Protein Data Bank (code 1tah). Molecular replacement calculations were immediately executed using the MERLOT program [34]. Crowther rotation function and the Crowther and Blow translation function calculations were conducted using 8–3.2 Å intensity data. The Crowther rotation function calculations gave a clear solution using either the PGL coordinates (11.3σ) or PCL coordinates generated from the PGL coordinates (11.4σ). The corresponding Crowther and Blow translation function for the best rotation function model had a peak height of 13.1σ. This resulting molecular replacement structure solution was refined using the X-PLOR package [30] with 7–2.8 Å data. The refinement protocol included rigid-body refinement and 2–3 iterations of crystallographic simulated-annealing refinement followed by B factor refinement. Examination of the resulting $2F_o–F_c$ density map, using the FRODO program [35] on an Evans and Sutherland PS390 A+ graphics display, revealed that residues 129–165 did not fit the map and should not be included in the next round of X-PLOR refinement. Several rounds of model building using FRODO and refinement in X-PLOR followed. These culminated in a 7–2.5 Å structure which included all protein atoms for residues 2–320 and had an R factor = 0.184 (90% data) and a free R factor = 0.263 (10% of the data). The MIR map and a phase-combined map produced from the SQUASH program [36] were also used in addition to the $2F_o–F_c$ map in building in the missing parts of the structure. From this point forward the structure was refined using SHELXL-93 [37]. Further iterations of refinement and model-building, including the addition of the Ca^{+2} ion and water molecules, and extension of the resolution to 2.2 Å were required to reach the final PCL structure. The final structure has 118 water molecules and an R factor of 14.3% for 7–2.2 Å data. The details regarding the final structure and refinement are summarized in Table 2.

Determination of the GBF model

Commercial lipase (triacylglycerol lipase, EC 3.1.1.3) of *P. cepacia* (Amano Pharmaceutical Co.) was purified by a single hydrophobic interaction chromatography step and used for crystallization. Crystallization was achieved by the sitting-drop vapor diffusion method using CRYSCHEM crystallization plates with a Biomek automated laboratory station (Beckman, Mannheim, Germany). The optimized conditions were: a reservoir solution consisting of 29–30% (v/v) 1-propanol in 100 mM Tris/HCl buffer pH 8.5–8.7. The sitting drop, before mixing with reservoir, consisted of 22 mg ml^{-1} lipase in 10 mM Tris/HCl buffer, pH 7.5. The crystallizations were done at 12°C. These conditions were adapted from the crystallization conditions reported for *P. cepacia* by Bornscheuer *et al.* [20]. The monoclinic C2 crystals have unit cell dimensions a = 91.3 Å, b = 47.3 Å, c = 85.4 Å and β = 121.4°.

Native data to 1.9 Å resolution were collected using synchrotron radiation from the wiggler beam line No. 6 at the DORIS storage ring, HASYLAB, DESY, Hamburg. The data were recorded at a wavelength of 1.0 Å at 10°C on a MAR imaging plate area detector. This data set was collected at a crystal to detector distance of 20 cm using oscillations of 1° per exposure. The digitized images were evaluated using the DENZO program package [38] to obtain integrated intensities. The merged data set contained 17 688 unique reflections within the resolution range of 2.0 to 10.0 Å and an R_{merge} of 9.5%. The completeness of the data set was 83%. Table 2 summarizes the quality of the data.

The phase problem was solved by molecular replacement using the coordinates of *Chromobacterium viscosum* lipase [18] as a starting model. The rotation and translation functions were calculated using data from 8–4 Å with the program AMoRe [39]. The program combines automatically a rotation and translation search. As judged by the correlation coefficient, a convincing solution was obtained and was subjected to a rigid-body refinement protocol using the same software package.

The initial refinement of the *C. viscosum* lipase coordinates with X-PLOR [31] consisted of rigid-body minimization, conjugate gradient minimization, simulated annealing using a slow cooling protocol followed by conjugate gradient minimization using 10–3 Å data. Subsequently, the amino acid sequence of the model was corrected according to the published sequence [40]. Removal of the residues which did not fit into the density of the resulting $2F_o–F_c$ map or collided with symmetry-related molecules, amino acids 130–156 and 217–239, reduced the R factor to 30.9%. In subsequent cycles of manual model fitting with the programs O [31] and FRODO [35] followed by refinement with X-PLOR, the resolution was extended in steps of 0.1 Å. At 2.5 Å the R factor was 27.9% and group B factor optimization was included in the X-PLOR refinement. By inspecting $F_o–F_c$ and 'omit' maps the trace of residues 130–156 and 217–239 could be easily recognized. The model was completely rebuilt and the R factor dropped to 25.9%. The resolution was stepwise extended to 2.0 Å. Additionally, water molecules were added by inspecting $F_o–F_c$ density maps. The final refinement cycle was carried out by restrained parameter least squares [41] with the software package PROLSQ [42]. The R factor was reduced to 18.8% for data between 8 and 2.0 Å resolution.

The final model of the PCL consisted of a polypeptide chain of 320 amino acids. It includes 193 water molecules and a Ca^{2+} ion. Four regions have poorly defined density for the sidechains: residues 75–77, 128–130, 199–201, and 219–222. These regions are located in exterior loops of the molecule. The mean coordinate error of all atoms estimated from Luzzatti plots [22] is 0.2 Å.

Accession numbers

The BRI coordinates have been deposited in the Protein Data Bank under accession code 2lip. This is NRC publication number 39957.

References

1. Rubin, B. (1994). Grease pit chemistry exposed. *Nat. Struct. Biol.* **1**, 568–572.
2. Ollis, D.A., *et al.*, & Goldman, A. (1992). The α/β hydrolase fold. *Protein Eng.* **5**, 197–211.
3. Desnuelle, P. (1972). The lipases. In *The Enzymes*. (Boyer, P.D., ed.), pp. 575–616, Academic Press, Inc., NY, USA.
4. Verger, R. (1980). Enzyme kinetics of lipolysis. *Methods Enzymol.* **64**, 340–392.
5. Brady, L., *et al.*, & Menge, U. (1990). A serine protease triad forms the catalytic centre of a triacylglycerol lipase. *Nature* **343**, 767–770.
6. Winkler, F.K., D'Arcy, A. & Hunziker, W. (1990). Structure of human pancreatic lipase. *Nature* **343**, 771–774.
7. Schrag, J.D., Li, Y., Wu, S. & Cygler, M. (1991). Ser-His-Glu triad forms the catalytic site of the lipase from *Geotrichum candidum*. *Nature* **351**, 761–764.
8. Brzozowski, A.M., *et al.*, & Thim, L. (1991). A model for interfacial activation in lipases from the structure of a fungal lipase–inhibitor complex. *Nature* **351**, 491–494.
9. van Tilbeurgh, H., Egloff, M.-P., Martinez, C., Rugani, N., Verger, R. & Cambillau, C. (1993). Interfacial activation of the lipase–procolipase complex by mixed micelles revealed by X-ray crystallography. *Nature* **362**, 814–818.
10. Grochulski, P., *et al.*, & Cygler, M. (1994). Analogs of reaction intermediates identify a unique substrate binding site in *Candida rugosa* lipase. *Biochemistry* **33**, 3494–3500.
11. Uppenberg, J., *et al.*, & Jones, T.A. (1995). Crystallographic and molecular-modeling studies of lipase B from *Candida antarctica* reveal a stereospecific pocket for secondary alcohols. *Biochemistry* **34**, 16838–16851.
12. Martinez, C., DeGeus, P., Lauwereys, M., Mattyhssens, G. & Cambillau, C. (1992). *Fusarium solani* cutinase is a lipolytic enzyme with a catalytic serine accessible to solvent. *Nature* **356**, 615–618.
13. Bourne, Y., Martinez, C., Kerfelec, B., Lombardo, D., Chapus, C. & Cambillau, C. (1994). Horse pancreatic lipase. The crystal structure refined at 2.3 Å resolution. *J. Mol. Biol.* **238**, 709–732.
14. Grochulski, P., Li, Y., Schrag, J.D. & Cygler, M. (1994). Two conformational states of *Candida rugosa* lipase. *Protein Sci.* **3**, 82–91.
15. Derewenda, U., *et al.*, & Derewenda, Z.S. (1994). An unusual buried polar cluster in a family of fungal lipases. *Nat. Struct. Biol.* **1**, 36–47.
16. Grochulski, P., *et al.*, & Cygler, M. (1993). Insights into interfacial activation from an open structure of *Candida rugosa* lipase. *J. Biol. Chem.* **268**, 12843–12847.
17. Noble, M.E.M., Cleasby, A., Johnson, L.N., Egmond, M.R. & Frenken, L.G.J. (1993). The crystal structure of triacylglycerol lipase from *Pseudomonas glumae* reveals a partially redundant catalytic aspartate. *FEBS Lett.* **331**, 123–128.
18. Lang, D., *et al.*, & Schomburg, D. (1996). Crystal structure of a bacterial lipase from *Chromobacterium viscosum* ATCC 6918 refined at 1.6 Å resolution. *J. Mol. Biol.* **259**, 704–717.
19. Larson, S., Day, J., Greenwood, A., Oliver, J., Rubingh, D. & McPherson, A. (1991). Preliminary investigation of crystals of the neutral lipase from *Pseudomonas fluorescens*. *J. Mol. Biol.* **222**, 21–22.
20. Bornscheuer, U., *et al.*, & Menge, U. (1994). Lipase of *Pseudomonas cepacia* for biotechnological purposes: purification, crystallization and characterization. *Biochim. Biophys. Acta* **1201**, 55–60.
21. Matthews, B.W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497.
22. Luzzatti, P.V. (1952). Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Cryst.* **5**, 802–810.
23. Cygler, M., Schrag, J.D. & Ergan, F. (1992). Advances in structural understanding of lipases. *Biotech. Gen. Eng. Rev.* **10**, 143–184.
24. Derewenda, Z.S. & Derewenda, U. (1991). Relationships among serine hydrolases: evidence for a common structural motif in triacylglyceride lipases and esterases. *Biochem. Cell Biol.* **69**, 842–851.
25. Frenken, L.G., Egmond, M.R., Batenburg, A.M., Bos, J.W., Visser, C. & Verrips, C.T. (1993). Cloning of the *Pseudomonas glumae* lipase gene and determination of the active site residues. *Appl. Environ. Microbiol.* **58**, 3787–3791.
26. Gilbert, J. (1993). *Pseudomonas* lipases: biochemical properties and molecular cloning. *Enzyme Microb. Technol.* **15**, 634–645.
27. Derewenda, U., Swenson, L., Wei, Y. & Derewenda, Z.S. (1994). Conformational lability of lipases in the absence of an oil–water interface: crystallographic studies of enzymes from the fungi *Humicola lanuginosa* and *Rhizopus delemar*. *J. Lipid Res.* **35**, 524–534.
28. Cygler, M., *et al.*, & Gupta, A.K. (1994). A structural basis for the chiral preferences of lipases. *J. Am. Chem. Soc.* **116**, 3180–3186.
29. Bernstein, F.C., *et al.*, & Tasumi, M. (1977). The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.
30. Brünger, A.T. (1992). *X-PLOR, Version 3.1. a System for X-ray Crystallography and NMR*. Yale University, New Haven, CT, USA.
31. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* **47**, 110–119.
32. Kabsch, W. (1988). Evaluation of single-crystal X-ray difraction data from a position-sensitive detector. *J. Appl. Cryst.* **21**, 916–924.
33. Terwilliger, T.C. & Eisenberg, D. (1983). Unbiased three-dimensional refinement of heavy-atom parameters by correlation of origin-removed Patterson functions. *Acta Cryst. A* **39**, 813–817.
34. Fitzgerald, P.M.D. (1988). MERLOT, an integrated package of computer programs for the determination of crystal structures by molecular replacement. *J. Appl. Cryst.* **21**, 273–281.
35. Jones, T.A. (1978). A graphics model building and refinement system for macromolecules. *J. Appl. Cryst.* **11**, 268–272.
36. Zhang, K.Y.J. (1993). SQUASH-combining constraints for macromolecular phase refinement and extension. *Acta Cryst. D* **49**, 213–222.
37. Sheldrick, G. (1993). SHELXL-93 program. Institut fuer Anorganische Chemie, Goettingen, Germany.
38. Gewirth, D. (1994). *The HKL Manual, Edition 3*. Yale University, New Haven, CT, USA.
39. Navaza, J. (1994). AMoRe: an automated package for molecular replacement. *Acta Cryst. A* **50**, 157–163.
40. Jorgensen, S.T., Skov, K.W. & Diderichsen, B. (1991). Cloning, sequence, and expression of a lipase gene from *Pseudomonas cepacia*: lipase production in heterologous hosts requires two *Pseudomonas* genes. *J. Bacteriol.* **173**, 559–567.
41. Hendrickson, W.A. & Konnert, J.H. (1980). Incorporation of stereochemical information into crystallographic refinement. In *Computing in Crystallography*. (Diamond, R., Ramaseshan, S. & Venkatesan, K., eds), pp. 13.01–13.21, Indian Academy of Science, Bangalor, India.
42. Collaborative Computing Project No. 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Cryst. D* **50**, 760–763.

43. Kraulis, P.J. (1991). MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Cryst.* 24, 946–950.
44. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577–2637.
45. Nicholls, A., Sharp, K.A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins* 11, 282–296.
46. Merritt, E.A. & Murphy, M.E.P. (1994). Raster3D version 2.0: a program for photorealistic molecular graphics. *Acta Cryst. D* 50, 869–873.