THE JOURNAL OF BIOLOGICAL CHEMISTRY
© 1993 by The American Society for Biochemistry and Molecular Biology, Inc.

Vol. 268, No. 17, Issue of June 15, pp. 12843-12847, 1993
Printed in U.S.A.

Insights into Interfacial Activation from an Open Structure of *Candida rugosa* Lipase*

(Received for publication, January 8, 1993, and in revised form, February 18, 1993)

Paweł Grochulski‡§, Yunge Li‡, Joseph D. Schrag‡, François Bouthillier‡, Penny Smith¶, David Harrison¶, Byron Rubin¶‖, and Miroslaw Cygler‡‖

From the ‡Biotechnology Research Institute, National Research Council of Canada, Montréal, Québec H4P 2R2, Canada and ¶Eastman Kodak Co., Rochester, New York 14650

The structure of the *Candida rugosa* lipase determined at 2.06-Å resolution reveals a conformation with a solvent-accessible active site. Comparison with the crystal structure of the homologous lipase from *Geotrichum candidum*, in which the active site is covered by surface loops and is inaccessible from the solvent, shows that the largest structural differences occur in the vicinity of the active site. Three loops in this region differ significantly in conformation, and the interfacial activation of these lipases is likely to be associated with conformational rearrangements of these loops. The "open" structure provides a new image of the substrate binding region and active site access, which is different from that inferred from the structure of the "closed" form of the *G. candidum* lipase.

The presence of a water/lipid interface dramatically enhances the hydrolytic activity of lipases (1, 2). The activation phenomenon has been suggested to be associated with a conformational change in the enzyme (3, 4). The need for rearrangement is supported by the structures of three lipases reported to date. In all of them the Ser-His-Asp/Glu catalytic triads are occluded by a polypeptide flap (lid) and are not exposed to the solvent (5–8). The topological location of the flap varies among the three lipases, and its length and complexity increases with the size of the molecule. The first confirmation of structural rearrangement was provided by the crystal structures of the small *Rhizomucor miehei* lipase complexed with inhibitors (9, 10). They revealed a large displacement of a single loop that occludes the active site in the native structure. This rearrangement opened access to the active site serine and exposed a number of hydrophobic residues, creating a significantly larger hydrophobic area for interaction with the interface and substrate binding. Cutinase, which also hydrolyzes triacylglycerols, does not show the interfacial activation displayed by other lipases and its active site was found to be accessible to solvent (11). The lack of interfacial activation was attributed to the absence of a flap (11). *Candida rugosa* lipase (CRL)[1] belongs to the same enzyme family as *Geotrichum candidum* lipase (GCL) and acetylcholinesterases (7, 12). The 2.06-Å resolution structure of CRL reported here is the first example of a native interface-activatable lipase in an "open" form, with an accessible active site. Taken together with GCL in its "closed" form (7, 13), the two structures provide insight into the conformational rearrangement associated with the interfacial activation of these higher molecular weight lipases and identify the regions likely to be involved in substrate binding. This new information modifies our previous notions, which were inferred from the GCL structure alone (7) and from its comparison to acetylcholinesterase (13).

The fungus *C. rugosa* (formerly *cylindracea*) produces several closely related lipase isoforms, five of which have been cloned and sequenced (14, 15).[2] All genes code for 534-amino acid proteins with M_r of approximately 60,000. The isoforms show ~40% amino acid sequence identity to GCL and ~25% to acetylcholinesterase (7, 12). Conserved residues include the catalytic triad, disulfide forming cysteines, and some salt bridges.

MATERIALS AND METHODS

The protein used for crystallization was purified from the crude *C. rugosa* lipase obtained from Sigma or Kodak (16). Purification of crude *C. rugosa* lipase by gel filtration, ion exchange chromatography, and chromatofocusing separated multiple isoforms, some of which are likely to be products of heterogeneous glycosylation. Crystals were obtained by the vapor diffusion method independently at BRI and Kodak laboratories (16) using 2-methyl-2,4-pentanediol as a precipitant. Identical orthorhombic crystals (space group $C222_1$; cell dimensions a = 64.9, b = 97.5, c = 175.6 Å) were obtained in both cases. The x-ray data to 2.06-Å resolution used in the refinement were collected on the R-AXIS area detector from a crystal of isoform pI = 4.4 (Sigma). The structure was determined by the molecular replacement method using program AUTOMR (17). Although the molecular replacement solution using GCL as a model was clear, the refinement proved to be difficult due to significant local differences in some parts of the structure. Systematic omission from the rotation function calculations of a sliding window of 40 residues of GCL allowed us to exclude from the starting model those parts of the GCL structure that were the most likely to differ from CRL. This procedure helped to minimize model bias and to select the best starting model for the refinement. The starting model selected in this way contained 65% of all GCL residues. This model was slowly extended through many rounds of molecular dynamics refinement (X-PLOR; Ref. 18) and model refitting (FRODO; Ref. 19). The tracing of the polypeptide chain in the region 62–92 (flap) and 122–129 became apparent only in the advanced stages of the refinement, when the *R*-factor dropped below 0.23. The current *R*-factor, including 237 water molecules and 3 sugar residues, is 0.142 for 21,869 reflections (8–2.06-Å resolution)

* This is publication 33715 in a series from the National Research Council of Canada. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates (code 1CRL) have been deposited in the Protein Data Bank, Brookhaven National Laboratory, Upton, NY with a release date of January 1, 1994.

§ Permanent address: Inst. of Physics, Technical University of Łódź, 93-005 Łódź, Poland.

‖ To whom correspondence should be addressed.

[1] The abbreviations used are: CRL, *C. rugosa* lipase; GCL, *G. candidum* lipase.

[2] L. Alberghina, personal communication.

with $I \geq \sigma(I)$. The root-mean-square deviation in bond lengths and bond angles from their ideal values are 0.012 Å and 2.8°, respectively, and the average B factor is 19.2 Å^2. The electron density map is in excellent agreement with the amino acid sequence of Kawaguchi *et al.* (14).

RESULTS AND DISCUSSION

Description of the Structure—CRL is a single-domain molecule (Fig. 1) and belongs to the family of α/β hydrolase fold proteins (20). In this paper, the nomenclature for the secondary structural elements follows that recommended as a standard for describing cholinesterases and the lipase members of the lipase/esterase family of homologous proteins (12, 21). An explanation of this nomenclature is given in the legend to Fig. 1. The β-strands are numbered in the same way as in Ref. 20, and the correspondence between the names of the conserved helices of the α/β hydrolase fold (20) are A = $\alpha_{3,4}$, B = $\alpha_{4,5}$, C = $\alpha_{5,6}$, D = $\alpha^4_{6,7}$, E = $\alpha^4_{7,8}$, and F = $\alpha^2_{8,9}$. As inferred from sequence alignments (7, 12), the catalytic triad is formed by Ser-209, His-449, and Glu-341 (Fig. 2). Ser-209 is embedded in the characteristic super-secondary structural motif, strand-turn-helix, found in all other lipases (5–7), cutinase (11), and acetylcholinesterase (22).

Most of the major structural features of CRL are the same as in GCL (7, 13). Four hundred fifty structurally equivalent Cα atoms, 84% of all residues, superimpose with a root-mean-square deviation of 0.99 Å. Despite the overall similarity, important differences distinguish CRL and GCL and provide insight into the conformational rearrangement upon interfacial activation. There are 84 pairs of Cα atoms that differ in their positions by more than 2.5 Å after superposition of the two structures (Fig. 3). The differences in the C-terminal and the nearby 250–273 region, distant from the active site, are probably unrelated to the catalytic function of the protein. Large conformational differences which appear to be catalytically important involve three loops in the vicinity of the active site: $L_{b3,2}$ (residues 62–92, referred to in the text as a flap, Fig. 4), $L^1_{3,4}$ (residues 122–129), and $L^4_{6,7}$ (residues 294–305).

The loop that differs most in the two proteins is the flap. This loop in CRL is 7 residues shorter than the corresponding loop in GCL and assumes a very different orientation relative

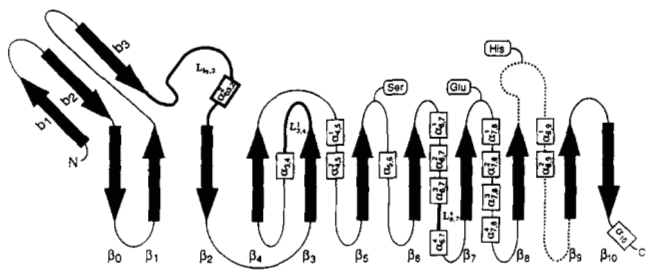


FIG. 1. **Topology and stereo view of the *C. rugosa* lipase polypeptide chain.** The β-strands of the large β-sheet are marked by β_i where i ranges from 0 to 10. The strands were numbered in a manner consistent with the numbering of the strands in the α/β hydrolase fold (20). Since the large β-sheet in CRL is larger than the conserved fold, the first strand is numbered 0. CRL also has a small β-sheet whose strands are designated by b_j, where j ranges from 1 to 3. The loops between β-strands are marked as $L^k_{i,j}$, where the subscripts denote the numbers or names of β-strands connected by the loop. The superscript refers to the consecutive part of the loop bordered by the secondary structural elements. When no superscript is used, $L_{i,j}$ refers to the entire length of the polypeptide between the indicated β-strands. α-Helices are designated $\alpha^k_{i,j}$, where the subscripts i and j refer to the loop in which the helix is embedded and the superscript k refers to the sequential number of the helix within the loop. The three loops near the active site which differ between CRL and GCL ($L_{b3,2}$, $L^1_{3,4}$, $L^4_{6,7}$) are marked in *bold lines*. The *dashed line* of $L_{8,9}$ indicates a rare left-handed connection between adjacent parallel β-strands. Assignment of the secondary structural elements is as follows: b_1 = 2–6, b_2 = 10–14, β_0 = 15–17, β_1 = 20–26, b_3 = 50–53, $\alpha_{b3,2}$ = 73–83, β_2 = 100–104, β_3 = 115–120, $\alpha_{3,4}$ = 136–145, β_4 = 150–154, $\alpha^1_{4,5}$ = 167–172, $\alpha^2_{4,5}$ = 177–191, β_5 = 202–208, $\alpha_{5,6}$ = 210–220, β_6 = 235–240, $\alpha^1_{6,7}$ = 253–266, $\alpha^2_{6,7}$ = 274–279, $\alpha^3_{6,7}$ = 283–290, $\alpha^4_{6,7}$ = 318–324, β_7 = 332–339, $\alpha^1_{7,8}$ = 343–347, $\alpha^2_{7,8}$ = 355–365, $\alpha^3_{7,8}$ = 371–380, $\alpha^4_{7,8}$ = 403–425, β_8 = 430–437, $\alpha^1_{8,9}$ = 452–456, $\alpha^2_{8,9}$ = 464–475, β_9 = 502–506, β_{10} = 509–513, α_{10} = 519–525.

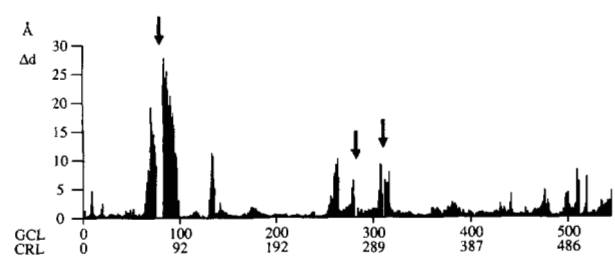


FIG. 3. **The plot of the distance between the structurally corresponding Cα atoms of GCL and CRL as a function of residue number.** The two structures were superimposed by minimizing the root-mean-square difference in position between the equivalent atoms. This superposition placed 450 Cα atoms in equivalent positions with an root-mean-square deviation of 0.99 Å. This plot is based on this superposition of the two lipase molecules. The *arrows* indicate deletions in the CRL sequence.

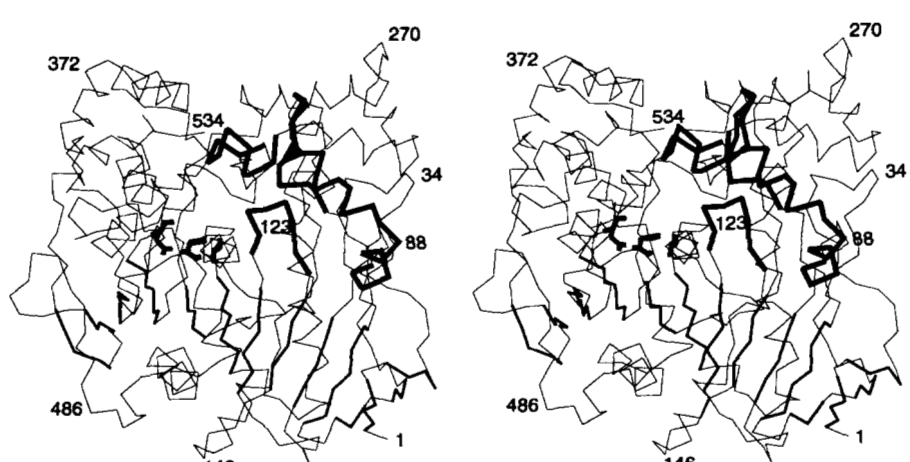


FIG. 2. **Stereo view of the Cα tracing of CRL showing the putative substrate binding depression.** Ser-209, Glu-341, and His-449, the 3 residues of the catalytic triad, are shown in full. β-Strands are shown as *medium lines*. Loops believed to be catalytically important that differ significantly in conformation from their counterparts in GCL are shown in *thick lines*.

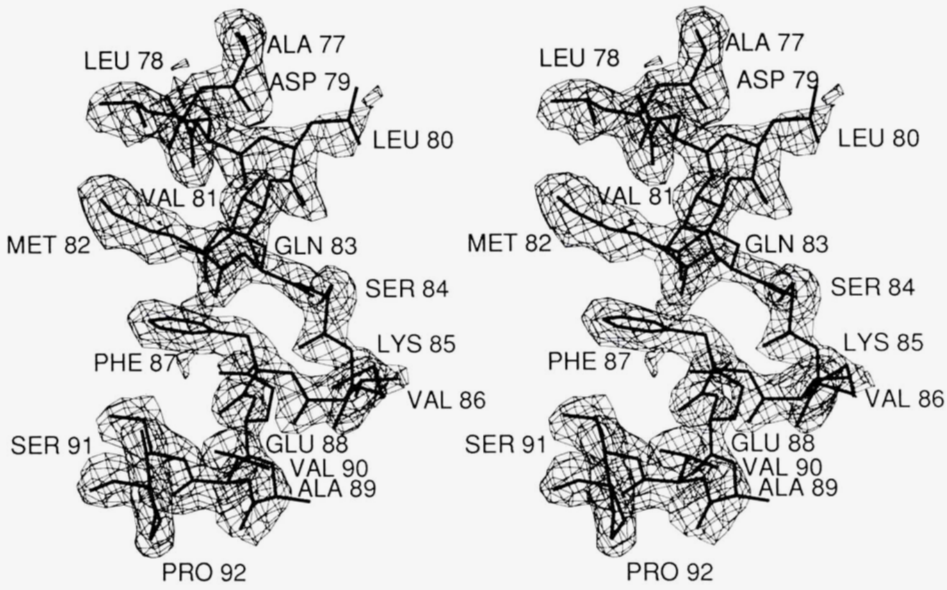
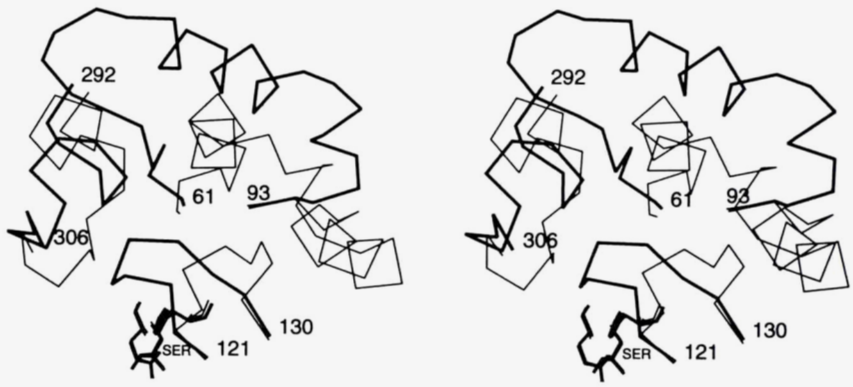


FIG. 4. **$3F_o - 2F_c$ electron density showing the quality of the map in the region of the flap.** The map is contoured at the 1σ level.

FIG. 5. **Stereo plot showing the differences in the conformation and position of the loops near the active site which differ between CRL and GCL.** The two structures were superimposed as described in the legend to Fig. 3. The loops of CRL are shown in *thick lines* and GCL is shown in *thin lines*.



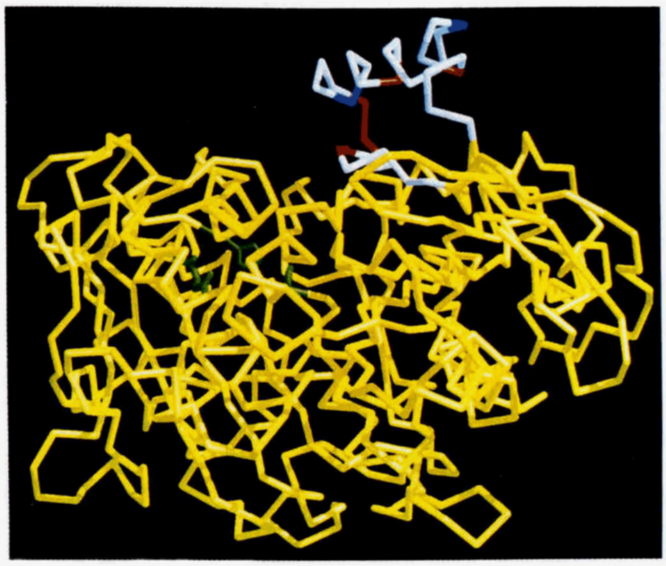


FIG. 6. **Cα trace of CRL showing the orientation of the flap relative to the rest of the molecule.** The residues of the flap are color-coded as described in the Fig. 7 legend, emphasizing the amphipathic character of the flap. Triad residues are shown in full (*green*).

to the remainder of the molecule. The positions of the tips of the two flaps differ by as much as 25 Å (Figs. 3 and 5). Strikingly, in the open conformation observed for CRL, the flap extends nearly perpendicular to the protein surface (Fig. 6) and forms one wall of a large depression that surrounds the active site (Fig. 7*a*). The active site Ser-209 lies at the bottom of this depression with its O_γ atom exposed to the solvent. The depression surrounding Ser-209 shows more hydrophobic character than any other solvent-exposed surface of the molecule. The face of the flap directed toward the active site is hydrophobic, composed mainly of aliphatic side chains. The face opposite the active site is hydrophilic in character and is stabilized through interactions with the protein surface. Similar stabilization has been observed in the *R. miehei* lipase-inhibitor complexes (9, 10). In GCL, the flap packs against the bulk of the molecule with its hydrophobic face buried and its hydrophilic face forming one exterior surface of the enzyme. In this position, the flap completely covers the active site (Fig. 7*b*). Removal of the flap from GCL exposes a surface that displays a hydrophobic nature similar to the depression of CRL.

The flap of CRL interacts with two areas of a symmetry-related molecule; Leu-73 and Leu-80 contact a hydrophobic patch formed by residues 439 to 459 whereas Val-86 and Val-90 contact their 2-fold related counterparts. These van der Waals interactions involve only six atom-atom distances

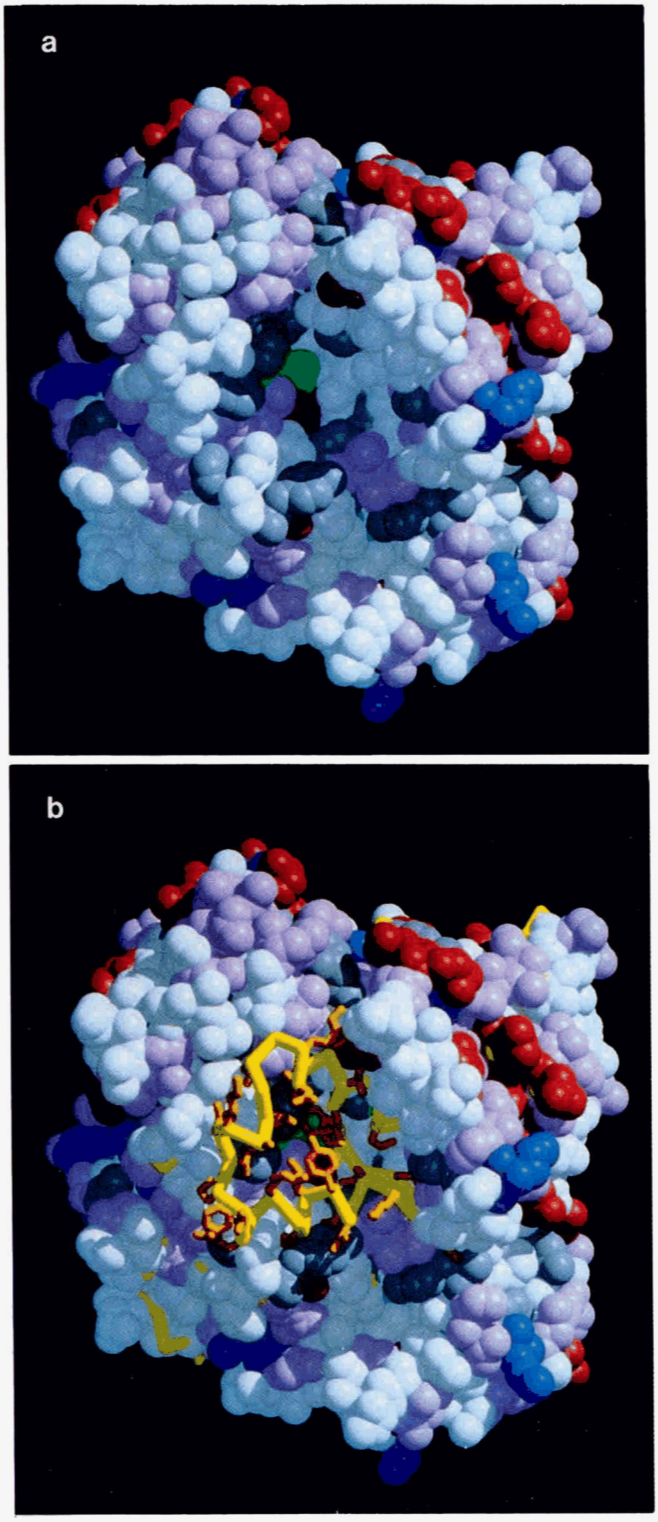


FIG. 7. **Van der Waals representation of CRL in an orientation similar to that in Fig. 2.** *a*, the catalytic Ser-209 is clearly visible at the bottom of the depression. The area surrounding this serine shows a distinct alternating pattern of hydrophobic and polar, uncharged residues. The residues are color-coded as follows: *dark blue*, Arg, Lys; *light blue*, His; *red*, Asp, Glu; *gray*, Tyr, Trp, Phe; *pink*, Thr, Ser, Asn, Gln; *green*, catalytic triad residues Ser-209, Glu-341, and His-449; *white*, Ile, Leu, Val, Ala, GLy, Pro, Met, Cys. *b*, composite of van der Waals representation of CRL and Cα trace of GCL after superposition of their structures. The two structures were superimposed as described in the Fig. 3 legend. CRL is shown in the same van der Waals representation seen in *a*. The tracing of GCL (*yellow*) is visible only in the regions where the two structures have significantly different conformations. The side chains of the flap and $L^1_{3,4}$ of GCL are also shown (*orange*). The catalytic triad of CRL and GCL superimpose with a root-mean-square difference of only 0.26 Å. One can clearly see that the flap and underlying loop of GCL completely occlude the substrate binding depression and the catalytic site. Most of the other differences are in the positions of surface side chains.

shorter than 4.0 Å. In contrast, the intramolecular contacts between residues on the hydrophilic face of the flap and the adjacent protein surface are more extensive. This suggests that the observed structure of CRL is an intrinsically stable conformation of the enzyme and that the position of the flap is not determined by the intermolecular forces involved in forming the crystal.

The second conformational difference leading to a more open, accessible active site in CRL is the orientation of the $L^1_{3,4}$ loop. Despite a high degree of sequence identity in this loop, the positions observed in CRL and GCL are very different (Fig. 5). In CRL, this loop extends toward the center of the protein in a conformation closely resembling that of the corresponding loop in native (22) and complexed forms (23) of acetylcholinesterase. The counterpart in GCL extends in the opposite direction, partially occluding the active site, and is tucked beneath the flap (7, 13) (Figs. 3 and 7*b*). The geometry of $L^1_{3,4}$ and the active site of CRL suggest that the NH groups of Ala-210 and Gly-124 will form the oxyanion hole. The corresponding atoms, NH of Ala-218 and Ala-132, in GCL were also proposed to contribute to the oxyanion hole (7), so the position of $L^1_{3,4}$ appears to be very important to catalysis.

The third conformational difference observed in the proximity of the active site is of the $L^4_{6,7}$ loop, which abuts the $L^1_{3,4}$ loop (Fig. 5). The conformation of the former segment in CRL accommodates the shifted $L^1_{3,4}$ loop, but the differences observed between CRL and GCL may also result from a 2-amino acid deletion in CRL in the $L^4_{6,7}$ loop.

GCL and CRL show many similarities in their kinetics and substrate specificity (24, 25), and both are activated by the presence of an interface. Taken together with the amino acid sequence similarity and the overall similarity in their three-dimensional structures, a similar mechanism of activation is expected. The reported increase in the rate of CRL-catalyzed esterification of 2-hydroxy acids in solvents with a low dielectric constant (26) further suggests that the open conformation of CRL observed in the crystal may have been induced by the crystallization conditions (35% 2-methyl-2,4-pentanediol). These findings and the fact that the large conformational and positional differences in the loops are concentrated mainly in the region of the active site strongly suggest that the CRL and GCL structures represent conformational states near the opposite ends of the activation pathway. The GCL structure represents the inactive, closed state; CRL, with an accessible active site, represents either an active state or an intermediate one, close to the active conformation.

Model for Interfacial Activation—Inspection of the GCL structure (7) and the observation of an internal cavity extending from the active site serine to two surface helices, one from the flap and one from loop $L_{6,7}$ (13), suggested that the direction of substrate entry into the active site was along a line approximately parallel to the β-strands of the large β-sheet (see Ref. 7). This notion was strengthened by the comparison with acetylcholinesterase, whose active site gorge (22) corresponds in position to the cavity in GCL (13). This suggested that the flap and loop $L_{6,7}$ would move away from one another to allow entry of the substrate. The emergence of the structure of the open conformation of CRL presents a

substantially different view of the structural changes responsible for the interfacial activation of GCL and CRL. The most extensive reorientation in the proposed transition from the closed to the open conformation involves unwinding and swinging of the flap around its disulfide-bridged base, *toward* the $L_{6,7}$ loop. The hydrophobic face of the flap, which is buried in the closed conformation, becomes exposed and forms one wall of the binding cleft in the open conformation. The location of the resulting depression suggests that substrate approaches the catalytic site from a direction roughly perpendicular to strands β_1 to β_5. The open conformation of the flap will likely be additionally stabilized by hydrophobic interactions with the lipid layer, which are mimicked in the crystal by intermolecular contacts. In conjunction with the movement of the flap, transition from the closed to the open states involves a change in the conformation of the $L^1_{3,4}$ loop. Such a movement is expected in the opening of GCL, in particular, in order to fully expose the active site and to move the NH of Ala-132 into the proper position for the putative oxyanion hole. The formation of the oxyanion hole by movement of $L^1_{3,4}$ in GCL would be similar to the formation of part of the oxyanion hole in *R. miehei* lipase by movement of the flap (9). If loop $L^1_{3,4}$ moves to the conformation observed in CRL, a concomitant conformational change of the loop $L^4_{6,7}$ is required. Transition of CRL from the open to the closed conformation may not require the reverse movement of $L^1_{3,4}$ and $L^4_{6,7}$. Instead, this transition may involve movement of the flap alone.

Substrate Binding Depression—The presumed substrate binding depression is apparent in Fig. 2 and 7*a*. The floor of the depression is formed by the C-terminal ends of strands β_4 and β_5, the catalytic Ser-209, and residues 121–129 of the $L^1_{3,4}$ loop. One wall is formed by the flap ($L_{b3,2}$), and residues 296–299 of the $L^4_{6,7}$ loop, two of the loops whose positions differ between CRL and GCL. The other wall of the depression is formed by residues 341–350 ($\alpha^1_{7,8}$), residues 442–459 ($L^1_{8,9}$, $\alpha^1_{8,9}$), and residues 397–400 ($L^4_{7,8}$). The surface of the cleft shows a pattern of interchanging hydrophobic/aromatic and hydrophilic patches extending from the active site (Fig. 7*a*). The hydrophilic areas are almost entirely uncharged polar residues. One may speculate that the fatty acyl chains of bound triacylglycerol extend along the hydrophobic patches. There are only two surface polar residues in the vicinity of Ser-209: Glu-208 and Ser-450. They could play a role in hydrogen bonding of the carbonyl groups of the glycerol head of the substrate. Correlation of sequence differences in the cleft of the various lipase isoforms with differences in substrate specificity should further define substrate binding residues and provide a basis for mutagenesis studies aimed at defining the structural determinants of substrate specificity.

Acknowledgments—We thank Marc Desrochers for assistance in computations, and especially for help in preparing the figures, and the Molecular Structure Corp. (The Woodlands, TX) for help in data collection.

REFERENCES

1. Brockman, H. L., Law, J. H., and Kézdy, F. J. (1973) *J. Biol. Chem.* **248,** 4965–4970
2. Macrae, A. R. (1983) in *Microbial Enzymes and Technology* (Fogarty, W. M., ed) pp. 225–250, Applied Science, London
3. Desnuelle, P., Sarda, L., and Ailhaud, G. (1960) *Biochim. Biophys. Acta* **37,** 570–571
4. Verger, R. (1980) *Methods Enzymol.* **64,** 340–392
5. Winkler, F. K., D'Arcy, A., and Hunziker, W. (1990) *Nature* **343,** 771–774
6. 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., and Menge, U. (1990) *Nature* **343,** 767–770
7. Schrag, J. D., Li, Y., Wu, S., and Cygler, M. (1991) *Nature* **351,** 671–674
8. Van Tilbeurgh, H., Sarda, L., Verger, R., and Cambillau, C. (1992) *Nature* **359,** 159–162
9. 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., and Thim, L. (1991) *Nature* **351,** 491–494
10. Derewenda, U., Brzozowski, A. M., Lawson, D. M., and Derewenda, Z. S. (1992) *Biochemistry* **31,** 1532–1540
11. Martinez, C., De Geus, P., Lauwereys, M., Matthyssens, G., and Cambillau, C. (1992) *Nature* **356,** 615–618
12. Cygler, M., Schrag, J. D., Sussman, J. L., Harel, M., Silman, I., Gentry, M. K., and Doctor, B. P. (1993) *Protein Sci.* **2,** 366–382
13. Schrag, J. D., and Cygler, M. (1993) *J. Mol. Biol.* **220,** in press
14. Kawaguchi, Y., Honda, H., Toniguchi-Morimura, J., and Iwasaki, S. (1989) *Nature* **341,** 164–166
15. Longhi, S., Fusetti, F., Grandori, R., Lotti, M., Vanoni, M., and Alberghina, L. (1992) *Biochim. Biophys. Acta* **1131,** 227–232
16. Rubin, B., Jamison, P., and Harrison, D. (1991) in *Lipases: Structure, Mechanism and Genetic Engineering* (Alberghina, L., Schmid, R. D., and Verger, R., eds) pp. 63–66, VCH Verlagsgesellschaft mbH, Weinheim, Germany
17. Matsuura, Y. (1991) *J. Appl. Crystallogr.* **24,** 1063–1066
18. Brünger, A. T., Kuriyan, J., and Karplus, M. (1987) *Science* **235,** 458–460
19. Jones, T. A. (1978) *J. Appl. Crystallogr.* **11,** 268–272
20. Ollis, D., Cheah, E., Cygler, M., Dijkstra, B. Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J. D., Sussman, J. L., Verscheuren, K. H. G., and Goldman, A. (1992) *Protein Eng.* **5,** 197–211
21. Massoulié, J., Sussman, J. L., Doctor, B. P., Soreq, H., Velan, B., Cygler, M., Rotundo, R., Shafferman, A., Silman, I., and Taylor, P. (1992) in *Multidisciplinary Approaches to Cholinesterase Functions* (Shafferman, A., and Velan, B., eds) pp. 285–288, Plenum Press, New York
22. Sussman, J. L., Harel, M., Frolow, F., Oefner, C., Goldman, A., Toker, L., and Silman, I. (1991) *Science* **253,** 872–879
23. Sussman, J. L., Harel, M., and Silman, I. (1992) in *Multidisciplinary Approaches to Cholinesterase Functions* (Shafferman, A., and Velan, B., eds) pp. 95–107, Plenum Press, New York
24. Sonnet, P. E., and Gazzillo, J. A. (1991) *J. Am. Oil Chem. Soc.* **68,** 11–15
25. Rangheard, M.-S., Langrand, G., Triantaphylides, C., and Baratti, J. (1989) *Biochim. Biophys. Acta* **1004,** 20–28
26. Parida, S., and Dordick, J. S. (1991) *J. Am. Chem. Soc.* **113,** 2253–2259