Volume 331, number 1,2, 123–128 FEBS 13020 September 1993

The crystal structure of triacylglycerol lipase from *Pseudomonas glumae* reveals a partially redundant catalytic aspartate

M.E.M. Noble[a,*], A. Cleasby[a,**], L.N. Johnson[a], M.R. Egmond[b], L.G.J. Frenken[b]

[a]*Laboratory of Molecular Biophysics, Department of Biochemistry, South Parks Road, Oxford, UK*
[b]*Unilever Research Laboratorium, Olivier van Noortlaan 120, NL-3133 AT Vlaardingen, The Netherlands*

Received 21 July 1993; revised version received 9 August 1993

The family of lipases (triacylglycerol-acyl-hydrolases, EC 3.1.1.3) constitutes an interesting class of enzymes because of their ability to interact with lipid–water interfaces, their wide range of substrate specificities, and their potential industrial applications [1,2]. Here we report the first crystal structure of a bacterial lipase, from *Pseudomonas glumae*. The structure is formed from three domains, the largest of which contains a subset of the α/β hydrolase fold and a calcium site. Asp^{263}, the acidic residue in the catalytic triad, has previously been mutated into an alanine with only a modest reduction in activity [3].

Triacylglycerol acylhydrolase; EC 3.1.1.3; Crystal structure; Density modification; Catalytic triad; Hydrolase fold

1. INTRODUCTION

Recently, X-ray crystallographically determined structures of lipases from *Rhizomucor meihei* (RML, [4]), *Geotrichum candidum* (GCL, [5]), and human pancreas (HPL, [6]) have shed much light on the structural basis of the biochemical peculiarities of this class of enzyme. Despite diverse amino acid sequences, the crystal structures of the lipases resemble an α/β fold which is observed in a broad family of hydrolase enzymes [7,8]. Common to this family is a nucleophile–histidine–acidic residue catalytic triad, which is observed in the lipases as either a Ser-His-Asp triad (RML, HPL), or a Ser-His-Glu triad (GCL). Whereas GCL is clearly a member of the α/β hydrolase family, HPL shows similarity over only a subset of its secondary structural elements, and RML is even less closely related [7].

Differences are also seen in the mode of interfacial activation [9,10]; when the catalytic serine of RML is covalently modified, a relatively simple helical lid movement occurs, whereas in the presence of bile salts, activation of HPL involves both interaction with an accessory protein (co-lipase), and a more complicated conformational change [10]. Defining the under-lying principles of lipase activity requires detailed investigation of a range of lipase families. One such family is formed by the bacterial lipases, particularly those from *Pseudomonas* species [11–13], several of which share considerable sequence homology. The member of this lipase family with the most industrially desirable properties [2], derived from *P. glumae*, has been chosen as the subject of more detailed study. This has led to the functional characterisation of the protein [3,14] as well as its crystallisation [15], and now structure determination.

*Corresponding author. Fax: (44) (865) 510454.

***Present address.* Glaxo Group Research, Greenford Rd., Middlesex, England, UK.

2. EXPERIMENTAL

Crystals of *P glumae* lipase were grown at 4°C by the hanging drop technique as described previously [15], with 27–29% polyethylene glycol 8000 as precipitant, and a droplet containing 10–20 mg ml^{-1} protein in 0.1 M Tris, pH 9.0, in the presence of 10% v/v acetone and trace amounts of n-dodecyl-β-D-glucopyranoside. Four copies of the mature lipase (M_r = 33,100 Da) were expected to form the asymmetric unit, giving rise to a Matthews volume of 3.0 $Å^3Da^{-1}$. Heavy atom derivatives were prepared by transferring crystals from the mother liquor in which they grew to an equivalent mother liquor to which the heavy atom compound had been added. A concentration of 5 mM and a soaking time of 1 week was used in all cases except for K_2PtCl_4, for which the higher concentration of 50 mM, and the longer soaking time of 5 weeks was found to be necessary to give significant binding. All data were collected on a Xentronics area detector with a Rigaku rotating anode source, and processed with the Xengen package [16]. Heavy atom structures were solved by a combination of difference Patterson and difference Fourier techniques using programs of the CCP4 suite [17]. Heavy atom refinement and phase calculation were performed using the program MLPHARE. Statistics of the data used, and of the phasing are given in Table I. The initial MIR map was very unclear, reflected by the low mean figure of merit to 3.0 Å resolution.

An improvement in the MIR electron density was achieved by application of solvent flattening and histogram matching techniques according to the program SQUASH [18]. This gave rise to clearly defined subunit boundaries. No clear indication of molecular symmetry was apparent from self-rotation calculations, or from inspection of the solvent flattened map. The first non-crystallographic symmetry operator was determined by carrying out a real space density correlation search using subunit masks and centres determined from the solvent flattened map. A version of the operator improvement program 'a-rt-improve' [19] was used for this purpose, so modified as to allow a full search of rotational space on an Eularian angle search grid.

A clear peak identified the transformation relating molecules A and B within the asymmetric unit. Phase improvement was carried out using this operator for partial averaging, combined with solvent flattening.

In the resulting 3 0 Å map, helices could be identified in the two averaged, as well as the two unaveraged molecules By matching these helices, the remaining transformations could be determined, and phase improvement performed using averaging over all four molecules. This was begun at 5 0 Å resolution, and extended to 3 0 Å, at incremental steps in resolution of approximately one reciprocal lattice unit. At each resolution, five cycles of averaging, with phase combination by SIGMAA [20] were performed. All averaging operations were carried out with the 'a' suite of programs [19] This protocol gave a very clear map (Fig. 1) in which the model could be built with the program 'O' [21].

Structure refinement was carried out using 'O' for manual rebuilding, and X-PLOR for automatic least squares and simulated annealing. A test set of 5% of the data was isolated at the start of refinement, and excluded from X-PLOR refinement. This set was used to monitor R_{free} [22] at different stages of refinement. For the first two rounds of simulated annealing, a single molecule was used, with strict non-crystallographic symmetry. In subsequent refinement, tight non-crystallographic symmetry restraints were applied to all protein atoms except for loops 18–29, 132–135, 150–160, 216–225, 216–225, 232 -236. For these loops, the map clearly indicated differences between the four molecules At the end of refinement, individual atomic B-factors were refined. These were restrained by both 1-2 and 1-3 bonding restraints, as well as the non-crystallographic symmetry restraint. Before B-factor refinement was begun, the conventional R value was 23.0%, and the free R value was 29.4%. After the final refinement, including B-factor refinement, the conventional R value was 15.9%, and the free R value was 22.6%. The small difference between R and R_{free} indicates that there has been little over-fitting of the data, and justifies the refinement of individual atomic temperature factors despite limited resolution. The deviation from ideal bond lengths of the final model was 0.011 Å, and the deviation from ideal bond angles was 2.5°. No residue has disallowed main-chain torsion angles, and the pooled side chain deviation around the preferred Chi-1 angles is only 14.4° compared to a typical value of 25.8° for other structures of 3.0 Å resolution (according to the PROCHECK package [23]).

3. RESULTS AND DISCUSSION

The structure of *P. glumae* lipase (PGL) was solved by the technique of multiple isomorphous replacement (MIR), with non-crystallographic symmetry averaging used to produce an interpretable map despite poor initial phases. The four molecules which form the crystallographic asymmetric unit are not related by proper symmetry. The free-R value [22] was used to identify the appropriate weights for geometric and non-crystallographic terms in X-PLOR [24] refinement, with the result that the final R_{free} is only 6.7% higher than the conventional R-factor.

The structure of PGL is composed of three domains (Fig. 2). The largest domain contains a six stranded β

Table I

Crystallographic information

Spacegroup $P2_12_12_1$, a = 158.16, b = 158.64, c = 63.36, $\alpha = \beta = \gamma = 90.0$

Data collected·

Dataset	Native	$NaAuCl_4$	$KAuCl_4$	$Hg(CH_3CO_2)$[a]		K_2PtCl_4
				(a)	(b)	
Number of crystals	2	3	2	3	1	1
Total observations	92,181	49,554	20,157	44,657	60,449	32,170
Merging R-factor[b]	4.6	5.0	6.2	5.1	7.2	4.8
Number of reflections	25,924	19,394	14,850	19,846	15,814	22,280
Completeness to 3.54 Å	92	75	62	78	64	82
Completeness 3.54–3.0 Å	59	33	18	32	23	46
Phasing statistics.						
MFID[c]	–	0.11	0.11	0 21	0.18	0.11
Number of sites	–	3	3	4	4	2
R-factor[d]	–	0.74	0.77	0.76	0.75	0.80
Phasing power	–	1.5	1.4	1 4	1.4	1.2

Mean figure of merit to 3.0 Å (All 25,882 phased reflections) = 0.53

[a] Two sets of data collected at different times which showed poor merging, and slightly different site occupancies.

[b] Mean merging R ($= \frac{\Sigma_h \Sigma_i |I_h - I_{h,i}|}{\Sigma_h \Sigma_i |I_{h,i}|} \times 100\%$) for individual blocks of data used in the final dataset.

[c] Mean Fractional Isomorphous Difference $= \frac{\Sigma_h |FPH_h - FP_h|}{\Sigma_h |FP_h|} \times 100\%$.

[d] Figures quoted for R-factor and Phasing power relate to acentric reflections only.

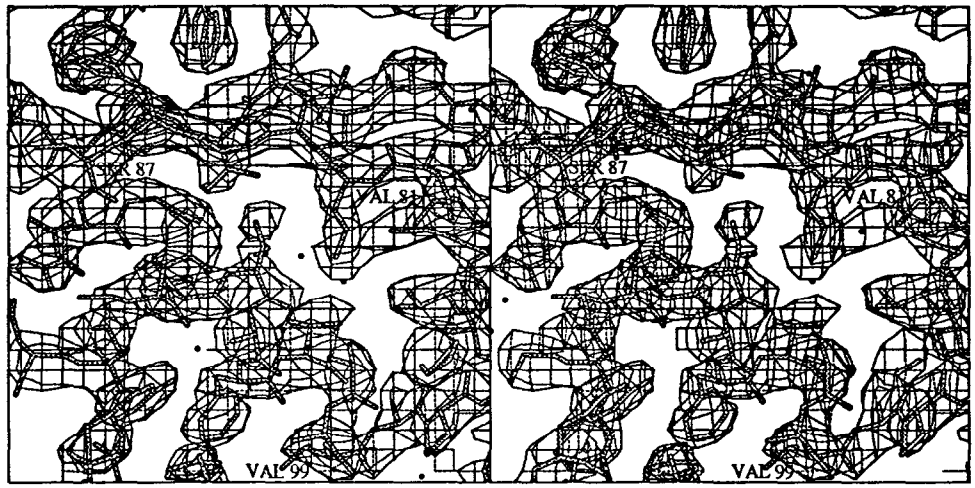
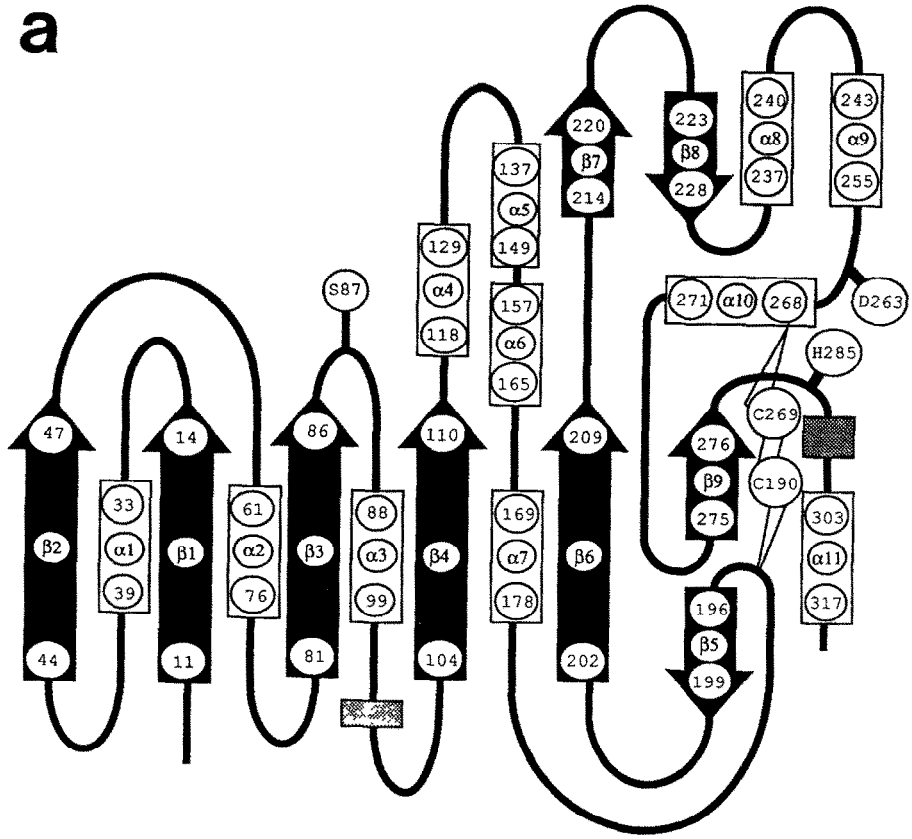


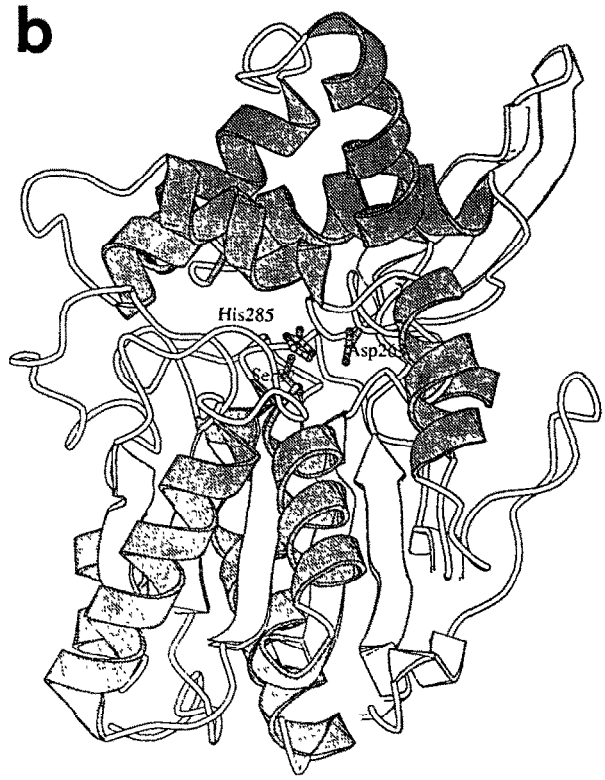
Fig. 1. A stereo diagram of the averaged MIR electron density map showing the region of chain around the catalytic serine, Ser^{87}. Ser^{87} is in a loop that connects a β strand (β3) with an α helix (α3).

sheet flanked on one face by three α helices, and on the other by two α helices. This sheet has considerably fewer strands than the equivalent sheets at the core of RML (8 strands), GCL (11 strands), and HPL (9 strands). Topological equivalents can, however, be identified with the core strands and helices of the α/β hydrolase fold, such that β1 of PGL is equivalent to β3 in the nomenclature of Ollis et al. [8]. The β sheet is principally parallel, with the exception of one of the edge strands, which is formed from a parallel and an anti-parallel section. Each of the other domains is formed by secondary structural elements contiguous in sequence. The first of these domains (residues 118–165) consists of three α helices, one of which may represent a moveable lid to the active site (discussed below). The other (residues 214–271) consists of four α helices and a β hairpin which extends from the body of the protein into solvent.

The identity of the nucleophile of PGL has been suggested by the recognition of a conserved Gly-X-Ser-X-Gly motif (Gly^{85}-His^{86}-Ser^{87}-Gln^{88}-Gly^{89}), and supported by studies which showed that PGL mutant $Ser^{87} \rightarrow$ Ala had no detectable activity [3]. In the crystal structure, this motif is found to form the short loop connecting a β strand (β3) with an α helix (α3). This situation is analogous to the nucleophile elbow in other members of the hydrolase family [7]. The main chain torsion angles of Ser^{87} are $\phi = 60$, $\psi = -128$, compared to values of $\phi = 53$, $\psi = -118$ for the catalytic serine of GCL. Ser^{87} is hydrogen bonded to NE2 of His^{285} (Fig. 3a). This histidine had been identified as a member of the catalytic triad on the basis of conservation amongst various *Pseudomonas* lipases, and confirmed to be essential for activity by site directed mutagenesis [3]. ND1 of this residue is hydrogen bonded to Asp^{263}, implicating this residue as the acidic member of the catalytic triad. Asp^{263} is totally conserved in members of the lipase family represented by PGL, but has previously been rejected as a candidate for the catalytic acidic residue on the basis of site directed mutagenesis [3]. Mutant $Asp^{263} \rightarrow$ Glu had approximately half wild-type V_{max} (1,800 LU · mg^{-1} compared to 4,000 LU · mg^{-1} for wild type), while mutant $Asp^{263} \rightarrow$ Ala was expressed at a lower level, but retained an activity of approximately 900 LU · mg^{-1}. Interestingly, the carboxyl of Asp^{263} is only 2.7 Å away from the carboxyl group of another acidic residue, Glu^{288} (Fig. 3a). Thus the acidic group of the catalytic triad in PGL is both dispensable, and in an unusual chemical environment. These two observations may be correlated: in the $Asp^{263} \rightarrow$ Ala mutant, it is possible that Glu^{288} either adopts the role of Asp^{263}, or alternatively serves to orient a water to perform this role. In other lipases and some related hydrolases (with the exception of HPL), the catalytic aspartate lies at the C-terminus of the β strand equivalent to β6, whereas in PGL, there are several secondary structural elements lying between the end of β6 and Asp^{263}.

Site directed mutagenesis has identified another aspartate, Asp^{241}, which is essential to production of active protein [3]. Mutant $Asp^{241} \rightarrow$ Ala had no detectable activity. In the crystal structure, this aspartate is one of the ligands of a calcium ion. This calcium site was observed as a very high peak in the early difference maps of all four subunits, and refined to a mean temperature factor of 20.2 $Å^2$ compared to an average value of 21.2 $Å^2$ for all protein atoms. The presence of a calcium site in the structure explains a number of biochemical features of PGL, notably the unexpectedly high pI (7.2 versus 6.0 predicted from the sequence [25]), and the irreversible inactivation of the protein on transfer to pH below 5.0. Loss of the calcium site, through either pH change or mutation of the Asp^{241} ligand, would be expected to decrease the protein's stability, and also to disrupt the local structure close to the active site. As


a
47
β2
44
33
α1
39
14
β1
11
61
α2
76
86
β3
81
S87
88
α3
99
129
α4
118
110
β4
104
137
α5
149
157
α6
165
169
α7
178
220
β7
214
209
β6
202
223
β8
228
240
α8
237
243
α9
255
271
α10
268
D263
H285
C269
276
β9
275
C190
196
β5
199
303
α11
317


b
His285
Asp263
Ser87

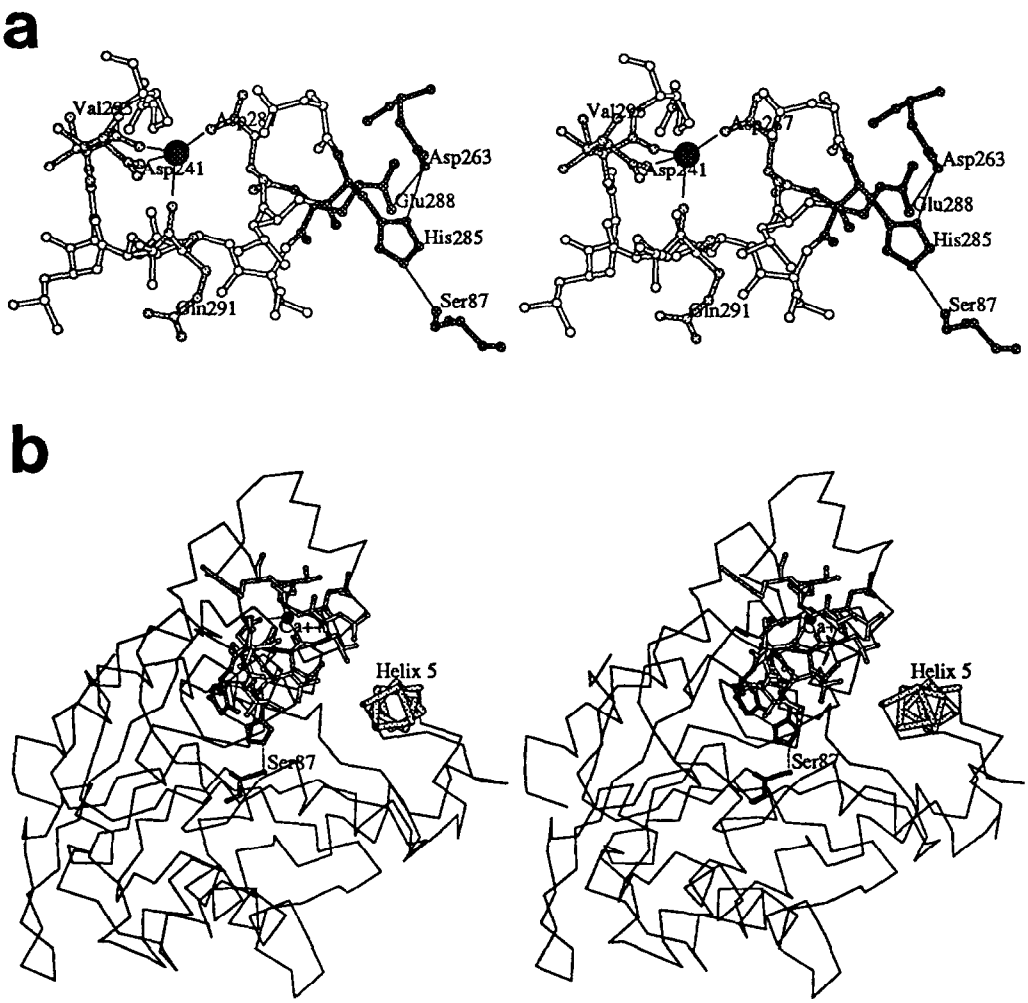


Fig. 3. Details of the calcium site, catalytic triad, and flap regions. (a) The catalytic residues Ser^{87}, Asp^{263}, and His^{285} are shown in a heavily stippled ball and stick representation, with solid lines representing hydrogen bonds between their side-chain atoms. Also shown is Glu^{288} which forms a 2.7 Å hydrogen bond with the side chain of Asp^{263}. The amino acids connecting His^{285} with residues of the calcium site are shown in an open ball and stick representation, with residues liganded to the dark calcium ion shown in a light stipple. (b) The calcium site and catalytic triad depiction above is related to the overall fold (thin trace), and particularly to the proposed helical lid, residues 137–149 (thick trace).

shown in Fig. 3a, three of the calcium ligands come from a part of the polypeptide which is only a few residues downstream of the catalytic His, His^{285} (ligands OD1-Asp^{287}, O-Gln^{291}, O-Val^{295}).

By analogy with lipase structures which have been solved in more than one conformation [9,10], it is expected that PGL has a lid able to move to reveal the active site, which in this structure is buried within the protein (Fig. 3b). Inspection of the structure shows that α5 (residues 137–149) is suitably located to serve as this lid. The two loops flanking this helix (residues 133–135, and 151–155) have the highest average main-chain temperature factors of the whole structure (>60 $Å^2$). In addition, the latter of these loop regions (residues 151–155), is the most protease-susceptible part of the PGL molecule [26]. This supports the assignment of α5 as the helical lid of PGL. If this is correct, then the helix of

←

Fig. 2. Topology and structure of PGL. (a) α helices and β strands of PGL as identified by the program DSSP [28] are depicted. In light grey are two short stretches of 3_{10} helix, residues 101–103, and 287–289. Residues of the catalytic triad are also depicted, with the order Ser-Asp-His in the primary structure. The two cysteines of the primary structure are involved in a disulphide bond which is indicated. (b) Three dimensional disposition of secondary structural elements in PGL. A cartoon representation of PGL showing the separation into the main hydrolase domain below, and two accessory domains above. The catalytic residues are labelled, and are situated at the top of the lower domain, well isolated from solvent.

PGL is considerably longer than the equivalent part of RML [4,9].

In summary, the structure of PGL has shown a new variation on the hydrolase theme, with an unusually short β sheet. This underlines the observation amongst hydrolase structures of a conserved core with a widely diverged periphery. It is in the peripheral regions, therefore, that the structures relating to system-specific properties of different lipases are to be expected. In the case of PGL, one such system-specific property is the interaction during in vivo folding and secretion with the product of the accessory gene LipB [27]. The crystal structure of PGL suggests an identity for the lid which controls substrate access to the active site, and the identities of the catalytic triad residues. The catalytic acidic group is not essential for lipase activity, and is in an unusual chemical environment. Further characterisation of this situation requires higher resolution data, and structural analysis of the Asp^{263} to alanine mutant.

Acknowledgements: We would like to thank Sir David Phillips for initiating structural studies on lipases in Oxford, and Dr. Elspeth Garman for help at the initial stages of data collection. This work was funded by E.C. Bridge Grant BIOT-CT090-0194. L.N.J. is a member of the Oxford Centre for Molecular Sciences which is funded by SERC and MRC. The coordinates will be deposited in the Brookhaven Protein Data bank.

REFERENCES

[1] Nishio, T., Takahashi, K., Tsuzuki, T., Yoshimoto, T., Kodera, Y., Matsushima, A., Saito, Y. and Inada, Y. (1992) J Biotechnology 8, 39–44.
[2] Thom, D., Swarthoff, T and Maat, J (1986) European Patent Applications 0 205 208 and 0 206 390.
[3] Frenken, L.G.J., Egmond, M.R., Batenburg, A.M., Bos, J.W., Visser, C. and Verrips, C.T. (1993) Applied and Environmental Microbiology, in press.
[4] Brady, L., Brzozowski, A.M., Derewenda, Z.S., Dodson, E., Dosdson, G., Tolley, S., Turkenburg, J.P., Christiansen, L., Huge-Jensen, B., Norskov, L., Thim, L. and Menge, U. (1990) Nature 343, 767–770.
[5] Schrag, J.D., Li, Y., Wu, S and Cygler, M. (1991) Nature 351, 761–764
[6] Winkler, F.K., D'Arcy, A. and Hunziker, W. (1990) Nature 343, 771–774.
[7] Ollis, D.L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S.M., Harel, M., Remington, S.J., Silman, I., Schrag, J., Sussman, J.L., Verschueren, K.H.G. and Goldman, A. (1992) Protein Eng. 5, 197–211.
[8] Cygler, M., Schrag, J D., Sussman, J.L., Harel, M., Silman, I., Gentry, M.K. and Doctor, B P. (1993) Prot. Sci. 2, 366–382.
[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] van Tilbeurgh, H., Egloff, M.-P., Martinez, C., Rugani, N., Verger, R. and Cambillau, C. (1993) Nature 362, 814–820
[11] Jorgensen, S., Skov, K.W. and Diderichsen, B. (1991) J. Bacteriol. 173, 559–567.
[12] Kordel, M., Hofmann, B., Schomburg, D. and Schmid, R.D. (1991) J. Bacteriol. 173, 4836–4841.
[13] Gilbert, E.J., Cornish, A. and Jones, C.W. (1991) J. Gen. Microbiol. 137, 2223–2229.
[14] Deveer, A.M.T J., Dijkman, R., Leuveling-Tjeenk, M., Vandenberg, B., Ransac, S., Batenburg, M., Egmond, M.R., Verheij, H.M. and de Haas, G.H. (1991) Biochemistry 30, 10034–10042.
[15] Cleasby, A., Garman, E., Egmond, M.R. and Batenburg, M. (1992) J. Mol. Biol 224, 281–282
[16] Howard, A. (1989) A Guide to Macromolecular X-ray Data Collection for the Nicolet Area Detector System, The Xengen System V1.3 (Program documentation).
[17] CCP4 (1979) The SERC (UK) Collaborative Computing Project no. 4, a Suite of Programs for Protein Crystallography, distributed from Daresbury Laboratory, Warrington, WA4 4AD, UK.
[18] Zhang, K.Y.J and Main, P. (1990) Acta Crystallogr. A46, 377–381
[19] Jones, T.A. (1992) in: Molecular Replacement, Proceedings of the CCP4 Study Weekend 31 January–1 February, 1992 (Dodson, E.J., Gover, S. and Wolf, W. Eds.) SERC, Daresbury Laboratory, Daresbury, Warrington, WA4 4AD
[20] Read, R.J. (1986) Acta Crystallogr. A42, 140–149.
[21] Jones, T.A., Zou, J.-Y., Cowan, S.W and Kjeldgaard, M. (1991) Acta Crystallogr. A47, 110–119.
[22] Brunger, A.T. (1992) Nature 355, 472–475.
[23] Morris, A.L., MacArthur, M.W., Hutchinson, E.G. and Thornton, J.M. (1992) Proteins: Structure, Function, and Genetics 12, 345–364.
[24] Brunger, A.T., Kuriyan, J. and Karplus, M. (1987) Science 235, 458–460
[25] Deveer, A.M Th.J. (1992) Mechanism of Activation of Lipolytic Enzymes, Ph D thesis, Rijksuniversiteit te Utrecht.
[26] Frenken, L G.J., Egmond, M.R., Batenburg, A.M. and Verrips, C.T. (1993) Protein Engineering, in press.
[27] Frenken, L.G.J (1993) Ph.D. thesis, Rijksuniversiteit te Utrecht.
[28] Kabsch, W. and Sander, C (1983) Biopolymers 22, 2577–2637