doi:10.1016/j.jmb.2009.05.017 J. Mol. Biol. (2009) xx, xxx–xxx

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding

Cammy K.-M. Chen[1,2]†, Guan-Chiun Lee[3]†, Tzu-Ping Ko[2], Rey-Ting Guo[2], Li-Min Huang[4], Hsiao-Jung Liu[3], Yi-Fang Ho[4], Jei-Fu Shaw[4,5]* and Andrew H.-J. Wang[1,2,6]*

[1]*Institute of Biochemical Sciences, National Taiwan University, Taipei 106, Taiwan*

[2]*Institute of Biological Chemistry, Academia Sinica, Taipei 115, Taiwan*

[3]*Department of Life Science, National Taiwan Normal University, Taipei 106, Taiwan*

[4]*Institute of Plant and Microbial Biology, Academia Sinica, Taipei 115, Taiwan*

[5]*Department of Food Science and Biotechnology, National Chung Hsing University, Taichung 402, Taiwan*

[6]*Core Facility for Protein Crystallography, Academia Sinica, Taipei 115, Taiwan*

Received 2 February 2009; received in revised form 27 April 2009; accepted 12 May 2009

Edited by G. Schulz

Several crystal structures of AFL, a novel lipase from the archaeon *Archaeoglobus fulgidus*, complexed with various ligands, have been determined at about 1.8 Å resolution. This enzyme has optimal activity in the temperature range of 70–90 °C and pH 10–11. AFL consists of an N-terminal α/β-hydrolase fold domain, a small lid domain, and a C-terminal β-barrel domain. The N-terminal catalytic domain consists of a 6-stranded β-sheet flanked by seven α-helices, four on one side and three on the other side. The C-terminal lipid binding domain consists of a β-sheet of 14 strands and a substrate covering motif on top of the highly hydrophobic substrate binding site. The catalytic triad residues (Ser136, Asp163, and His210) and the residues forming the oxyanion hole (Leu31 and Met137) are in positions similar to those of other lipases. Long-chain lipid is located across the two domains in the AFL–substrate complex. Structural comparison of the catalytic domain of AFL with a homologous lipase from *Bacillus subtilis* reveals an opposite substrate binding orientation in the two enzymes. AFL has a higher preference toward long-chain substrates whose binding site is provided by a hydrophobic tunnel in the C-terminal domain. The unusually large interacting surface area between the two domains may contribute to thermostability of the enzyme. Two amino acids, Asp61 and Lys101, are identified as hinge residues regulating movement of the lid domain. The hydrogen-bonding pattern associated with these two residues is pH dependent, which may account for the optimal enzyme activity at high pH. Further engineering of this novel lipase with high temperature and alkaline stability will find its use in industrial applications.

Keywords: lid; catalytic triad; interface; hinge; hydrophobic tunnel

*Corresponding authors. E-mail addresses: presid@dragon.nchu.edu.tw; ahjwang@gate.sinica.edu.tw.

† C.K.-M.C. and G.-C.L. contributed equally to this work.

Abbreviations used: AFL, *Archaeoglobusfulgidus* lipase; MIR, multiple isomorphous replacement; PEG, polyethylene glycol; pNPP, *p*-nitrophenyl palmitate; BSL, *Bacillus subtilis* lipase; mAFL, mature AFL; tAFL, C-domain-truncated AFL; pAFL, premature AFL; PL, pancreatic lipase; PDB, Protein Data Bank; TTR, transthyretin.

Introduction

Since the beginning of the 1990s, the crystal structures of several lipases have been determined.[1,2] The structures have attracted much interest because these enzymes can be used as biocatalysts not only for the hydrolysis of fats and oils but also for the synthesis of a large number of unnatural products.[3] The enzymes are frequently used for industrial-scale production of fine chemicals and pharmaceuticals.[3]

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017

Lipases are versatile in functions. In addition to their natural function of hydrolyzing carboxylic ester bonds, they also catalyze esterification, interesterification, and transesterification reactions in nonaqueous media. Such versatility makes them useful for a wide range of applications in the food, detergent, leather, pharmaceutical, textile, cosmetic, and paper industries.[4]

Lipases (EC 3.1.1.3) are triacylglycerol ester hydrolases that catalyze the hydrolysis of long-chain acylglycerols. Lipases and carboxylesterases differ in their substrate preferences: carboxylesterases prefer hydrophilic substrates whereas lipases favor long-chain hydrophobic fatty acid esters. Lipases generally have a broad substrate spectrum, though sometimes they possess a remarkable regioselectivity and enantioselectivity.[5] Lipases are widely distributed throughout the plant and animal kingdoms, as well as in molds and bacteria. Although lipases vary considerably in size, they all presumably have a similar α/β-hydrolase fold embedded with a catalytic triad consisting of serine, histidine, and aspartate (or glutamate). The serine residue is located in a strictly conserved β-Ser-α motif. The active site of lipase is often covered by an amphipathic helical lid domain, which moves away upon contact of the lipase with its substrate, thereby exposing hydrophobic binding pocket at the protein's surface.[2] In the presence of water–lipid interface, the lid opens and the enzyme activity is increased, a phenomenon called interfacial activation.[6]

Bacterial lipases have recently been classified into eight different families, with family I being the largest and consisting of six subfamilies.[7] These lipases usually show pronounced differences in regioselectivity and enantioselectivity despite a high degree of amino acid sequence homology.[2,7–10] *Archaeoglobus fulgidus* archaeabacteria are hyperthermophilic marine sulfate reducers found in hydrothermal environments. They were isolated from hot oil field waters from an oil production platform. They are anaerobes that utilize lactate, pyruvate, and valerate plus H_2 as carbon and energy sources with sulfate as electron acceptor.[11] Studies on lipases from hyperthermophilic bacteria such as *A. fulgidus* lipase (AFL) should provide additional insights into the identification of the structural elements important to protein's thermostability, while providing useful information for future applications in biocatalysis and biotechnology.[12,13]

The AFL is particularly interesting, since it not only is hyperthermophilic and alkalophilic but also contains a unique C-terminal domain. This enzyme showed optimal activity at 70–90 °C and between pH 10 and 11, which is among the most alkaline pH range detected for hydrolases.[14,15] Previously, the AFL gene was overexpressed in *Escherichia coli*, purified, and characterized.[14] On the basis of a secondary-structure-driven multisequence alignment, the N-terminal half of AFL was predicted to adopt the hydrolase topological fold and possess catalytic activity since a truncated version covering the 244 N-terminal residues retained measurable activity.[14] The function of the C-terminal domain in AFL was not clear. The large amounts of biochemical data and many crystal structures of lipases enabled good understanding of their catalytic mechanism, reaction selectivity, and substrate specificity, but none of the known bacterial lipase structures contains an extra C-terminal domain for long-chain substrate binding.[16–19] In the present work, we report the crystallographic analysis on a newly purified recombinant lipase cloned from the hyperthermophilic archaeon *A. fulgidus*, which was overexpressed in *E. coli*. The crystal structure of AFL at ~1.8 Å resolution showed that the AFL is the first lipase structure reported to contain a unique C-terminal domain that is essential for binding long-chain substrate.

Results and Discussion

X-ray structure determination

The structure of the enzyme was solved by multiple isomorphous replacement (MIR) method, and the refinement was first carried out using the 1.83-Å-resolution data of the S136A mutant (Supplementary Table 1). The structures of native AFL were determined in three different crystals, obtained by using isopropanol, polyethylene glycol (PEG) 4000, and PEG 8000 as the major precipitants and refined to 2.30, 1.77, and 3.10 Å resolution, respectively, all yielding low R and R_{free} values and stereochemical deviations (Table 1). There are two AFL monomers (A and B) in an asymmetric unit of the structures crystallized using isopropanol and PEG 4000, both in the space group $P2_12_12_1$ (Forms I and II). However, one monomer per asymmetric unit is observed in the structure crystallized using PEG 8000 in the space group $P4_3$ (Form III). The mutant S136A is also crystallized from isopropanol in the Form I unit cell.

The C-terminal His tag was not observed in any of the seven crystallographically independent monomer structures. When these seven independent AFL monomers are superimposed using the program sPDBV,[20] the root-mean-square deviations are 0.43–0.64 Å. In general, the respective N-terminal domain and the C-terminal domain superimpose well, except the lid region on top of the N-terminal domain (Supplementary Fig. 1a). Although the conformational change could be a result of crystal packing, it may also represent flexibility of the lid domain. In the AFL structures determined under three different crystallization conditions, we observed the binding of one Ca^{2+}, three Mg^{2+}, and one Cl^-. When these three structures are superimposed, we see the different ions all bind to the position around the top of the C-terminal domain, suggesting that these ions may help stabilize the structure of the C-terminal domain (Supplementary Fig. 1b).

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017

t1.2 **Table 1.** Data collection and refinement statistics for AFL crystals

t1.3 Names	Native (isopropanol) [Form I]	S136A + *p*NPP [Form I]	Native (PEG 4000) [Form II]	Native (PEG 8000) [Form III]
t1.4 *Data collection*				
t1.5 Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P4_3$
t1.6 Resolution (Å)[a]	30–2.30 (2.38–2.30)	30–1.83 (1.90–1.83)	30–1.77 (1.83–1.77)	50–3.1 (3.21–3.10)
t1.7 Unit cell dimensions (Å)				
t1.8 *a*	89.44	91.20	52.29	99.55
t1.9 *b*	105.38	107.87	106.67	99.55
t1.10 *c*	117.00	118.90	175.89	59.02
t1.11 No. of reflections				
t1.12 Observed	276,752 (24,748)	605,017 (54,461)	346,404 (27,062)	77,200 (7272)
t1.13 Unique	50,018 (4866)	104,198 (10,255)	94,247 (8682)	10,557 (1031)
t1.14 Completeness (%)	98.6 (97.6)	99.8 (99.9)	97.9 (91.8)	99.8 (100.0)
t1.15 R_{merge} (%)	6.0 (27.2)	5.1 (42.9)	6.5 (34.4)	6.2 (40.0)
t1.16 $I/\sigma(I)$	37.0 (8.1)	31.8 (5.4)	23.9 (2.4)	35.8 (5.0)
t1.17				
t1.18 *Refinement*				
t1.19 No. of reflections	47,885 (4537)	98,823 (9164)	90,064 (7267)	10,287 (842)
t1.20 R_{work} (95% data)	0.182 (0.221)	0.187 (0.247)	0.178 (0.251)	0.184 (0.348)
t1.21 R_{free} (5% data)	0.238 (0.270)	0.223 (0.282)	0.214 (0.278)	0.252 (0.457)
t1.22 Geometry deviations				
t1.23 Bond lengths (Å)	0.017	0.020	0.019	0.014
t1.24 Bond angles (°)	1.8	1.8	1.9	1.9
t1.25 No. of protein atoms	7126	7202	7246	3632
t1.26 Mean *B* values (Å^2)	27.7	26.2	27.5	67.3
t1.27 No. of ligand atoms	38	32	29	13
t1.28 Mean *B* values (Å^2)	36.1	35.2	39.4	55.9
t1.29 No. of ions	2	2	5	1
t1.30 Mean *B* values (Å^2)	20.5	19.0	37.6	46.6
t1.31 No. of water molecules	782	1226	1255	214
t1.32 Mean *B* values (Å^2)	42.1	40.9	48.2	61.8
t1.33 Ramachandran plot (%)				
t1.34 Most favored	89.2	89.0	90.3	80.2
t1.35 Additionally allowed	10.3	10.5	9.2	19.3
t1.36 Other	0.5	0.5	0.5	0.5

t1.37 [a] Values in parentheses are from the highest-resolution shell.

Overall structure of AFL

The overall structure shows a clear bipartite architecture composed of an N-terminal domain and a C-terminal domain. The N-terminal domain (residues 1–237), which has an α/β-hydrolase fold, carries the active site, while the C-terminal domain (residues 238–474) forms a β-barrel structure and provides the substrate hydrocarbon tail binding site (Fig. 1a). The N-terminal domain shows a single compact domain that consists of six parallel β-strands forming a β-sheet, sandwiched by seven α-helices, with four helices on one side and three on the other side. The fold of the N-terminal domain of AFL resembles that of the core of the α/β-hydrolase enzymes, which is also seen in other lipases of known 3D structures. Lipase was traditionally believed to have a lid structure. By solving the crystal structure, we discovered the presence of a lid in the catalytic N-terminal domain to regulate the binding of the substrate. The AFL lid is composed of three helices, α3, α4, and α5 (residues 62–101), and two hinge residues, Asp61 and Lys101 (Fig. 1a), and it adopts a closed conformation with clear electron densities, suggesting a well-ordered structure. The active site is located at the bottom of a hydrophobic crevice covered by the lid. Normally, the lid and the hinges are associated with the interfacial activation of lipases, where the α-helical lid opens up by rotating the two connecting hinge regions.

The C-terminal domain (residues 238–474) is of the β-sandwich type and is formed by two layers of seven β-strands. The front antiparallel β-sheet consists of strands G to M, whereas the back β-sheet is made up of strands N to T. At the C-terminus, a substrate covering motif on top of the β-sandwich, composed of α12, α13, α14, and α15, forms a part of the hydrophobic substrate binding tunnel. Since the substrate covering motif does not cover the active site of AFL and no hinge can be identified, it cannot be recognized as another lid. The topology of AFL is shown in Supplementary Fig. 1c. The native structure of AFL turns out to be a fatty-acid-bound structure and this might be caused by the binding of the protein to some fatty acid fragments during its purification process. Since the different AFL structures we obtained seem to bind different lengths of fatty acid fragment, we suspect that the fatty acid might be further extended at the C-terminal end, reflecting multiple modes of binding. In the mutant structure of AFL (S136A) soaked with *p*-nitrophenyl palmitate (pNPP), we still see the binding of the fatty acid fragment. This might be because the highly hydrophobic tunnel in AFL binds to the fatty acid fragments tightly and prevents it

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017

from binding to other substrates such as pNPP. Alternatively, the pNPP head group is slowly hydrolyzed by the S136A mutant, which still possessed very low activity.

Comparison with *Bacillus subtilis* lipase

Comparisons of known crystal structures of lipases show that AFL has a similar catalytic

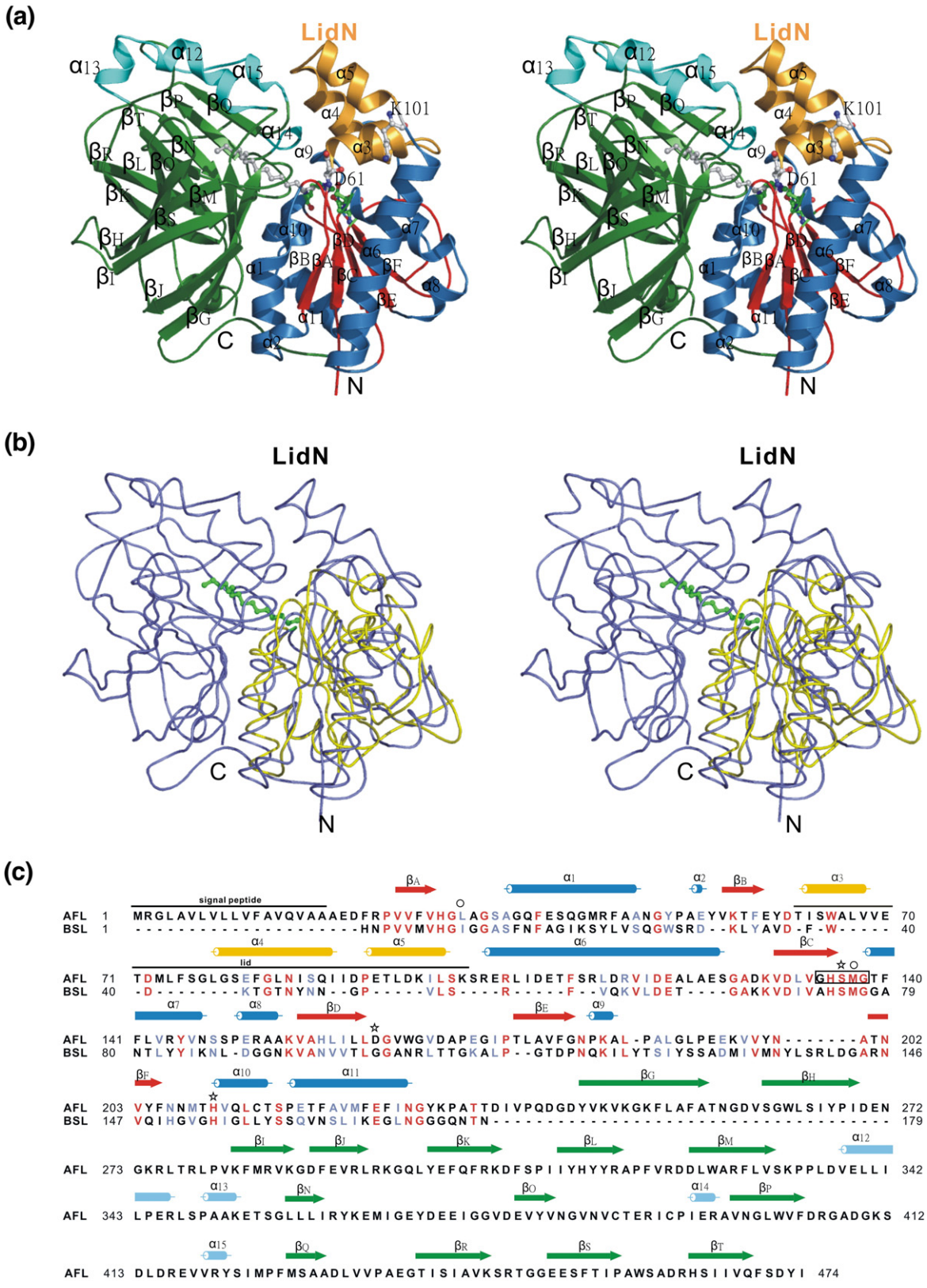

Fig. 1 (*legend on next page*)

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017

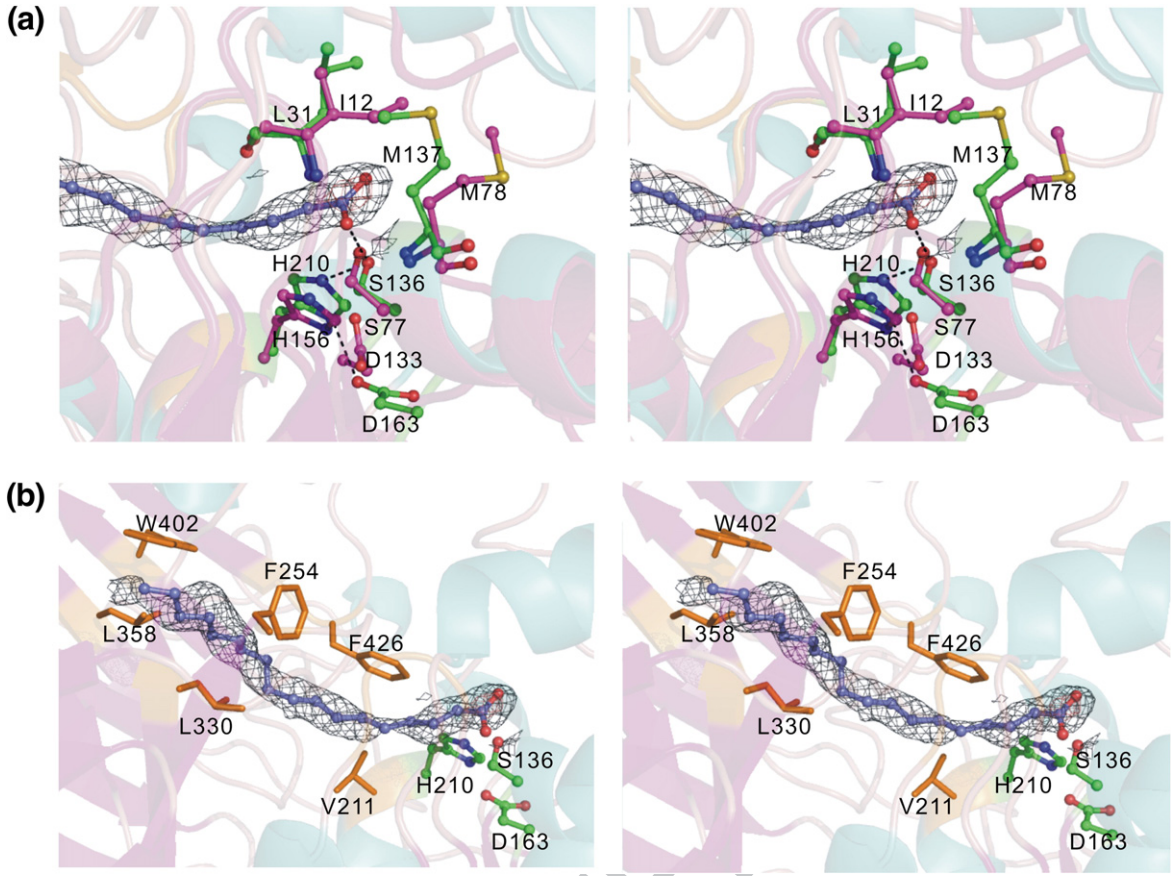

Fig. 2. Details of the interaction of AFL with its substrate in the active site. (a) A stereo view of the superimposition of the catalytic region of AFL and BSL. The figure shows the comparison of positions of the catalytic triad and oxyanion hole between the structures AFL (olive green) and BSL (purple). The AFL substrate is shown in blue as ball-and-stick structures. The $2F_o-F_c$ electron density maps are contoured at 1 σ for the AFL substrate. Hydrogen bonds between the triad of AFL and the substrate are depicted as black broken lines. (b) A stereo view of the substrate binding tunnel in AFL. The substrate is shown in blue as ball-and-stick structures, and the $2F_o-F_c$ electron density maps are contoured at 1 σ for the substrate. The side chains of the residues interacting with the substrate are shown as thicker lines (orange). The catalytic triad residues Ser136, Asp163, and His210 are shown in green as ball-and-stick structures.

domain. Further structural similarity search using DALI found that the structure of the *B. subtilis* lipase (BSL) exhibits the highest similarity to the N-terminal domain of AFL with a Z score of 19.7. BSL also shows the highest similarity (42%) and identity (23%) to AFL. They can be superimposed with an rmsd of 1.25 Å using 126 corresponding C^{α} atoms. However, the BSL structure lacks not only the C-terminal domain shown in the AFL structure but also a lid. The comparison of BSL and AFL shows that AFL lacks an α-helix between βD and βE and that it contains two additional α-helices α2 and α8 (Fig. 1b). Figure 1c shows the structure-based sequence alignment of BSL and AFL, revealing that the AFL structure contained an additional C-terminal domain and a lid domain.

Active site

Based on the sequence alignment with BSL, the amino acids Ser136, His210, and Asp163 or Asp169 of AFL were assumed to be part of the catalytic

Fig. 1. The overall structure of AFL and BSL and the comparison between them. (a) A stereo view of the overall structure of AFL. The N-terminal domain contains six parallel β-strands, forming a β-sheet (red), and are sandwiched by seven α-helices (blue) and a lid domain (orange). The C-terminal domain contains 14 antiparallel β-sheets (green) and a substrate covering domain (cyan). The substrate is shown in gray as a ball-and-stick model. The two hinge residues Asp61 and Lys101 are shown in gray as a ball-and-stick model. The catalytic triad is shown in green. (b) A stereo view of the superimposition of the AFL structure with BSL. The BSL structure (yellow) (PDB accession number 1r4z) shows a high degree of structural homology to the N-terminal α/β-hydrolase fold domain of AFL (blue). (c) Structure-based sequence alignment of AFL and BSL (PDB accession number 1r4z). The secondary structure of AFL is shown above the alignment [colored as in (a)]. β-Strands are shown as arrows and α-helices are shown as tubes. The numbering of all residues of the two proteins is shown on both sides of each line. Identical and similar amino acid residues are shown in red and sky blue, respectively. The lid and the signal peptide region are indicated. The stars indicate the catalytic triad residues, the circles indicate the oxyanion hole residues, and the box denotes the conserved GXSXG motif. The truncation site is shown with a pink arrow.

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017

triad.[14] Here, we revealed that the catalytic triad residues Ser136 (as a nucleophile), Asp163, and His210 are actually located at their canonical positions in the α/β-hydrolase fold.[21] Their positions and orientations are similar to the catalytic triad residues in BSL except that the position of Asp163 has a small deviation from Asp133 of BSL (Fig. 2a). This is true also for the oxyanion hole, which is formed by the backbone nitrogen atoms of the residue immediately following the nucleophile and a residue located between strand βA and helix α1. In AFL, the oxyanion hole is formed by the peptide NH moieties of Leu31 and Met137, which are in a conserved position as the oxyanion hole (formed by Ile12 and Met78) of BSL (Fig. 2a). This shows that AFL has a preformed oxyanion hole similar to that of BSL. The nucleophilic Ser136, which was covered by the lid in the closed form of AFL, is situated at the very sharp "nucleophile elbow", between strand βC and helix α7. The catalytic serine is in an ε characteristic conformation and is located in a tight turn with the G-H-S-M-G sequence belonging to the usual consensus sequence of the α/β-hydrolase fold family.

Acyl-binding site

The active-site nucleophile Ser136 lies at the entrance of a hydrophobic tunnel. The tunnel consists of a deep canyon 20 Å long and 7 Å wide (Fig. 2b). The cavity is formed in AFL between the N-terminal catalytic domain and the C-domain. Most residues of the tunnel wall are hydrophobic, the majority being leucines, phenylalanines, and valines (Fig. 2b). The structure of AFL with the fatty acid fragments shows the long-chain substrate fits in a hydrophobic tunnel lined by Val211, Phe254, Leu330, Leu358, Trp402, and Phe426. There is space for about 18 hydrocarbon units to be accommodated in this tunnel.

The two hinge residues Asp61 and Lys101 at two ends of the lid are located at the base of the lid domain. The opening of the lid may be induced by the hinge exerting a bending motion of the helical lid, thereby creating an increased solvent-exposed hydrophobic surface. By comparing the hinge points of the four structures of AFL we obtained, a pH-dependent conformational change at Lys101 was observed (Fig. 3a). Among the four structures, the conformation of Asp61 remains the same, but Lys101 of the structures determined at different pH values shows a different conformation. Detailed investigation on the interaction of the two hinge points shows that Lys101 interacts with Asp61 via hydrogen bonding with Ser64 under a more acidic crystallization condition (crystallized using isopropanol as the main precipitant, pH 4.6). However, Lys101 undergoes a 90° rotation to form a hydrogen bond with Glu109 in a more basic crystallization condition (crystallized using PEG 4000 and PEG 8000 as the main precipitant, pH 8.5 and 10.5, respectively) (Fig. 3b). It is reasonable to expect that under a more basic condition, the conformational change of Lys101 would render AFL to be more readily to open and bind to its substrate. The interaction of Lys101 with Glu109 instead of Ser64 probably enables the lid region in AFL to be more flexible and have the ability to exert a larger extent of movement. This might be the reason for AFL's alkalophilic property.

The substrate binding tunnel is covered by the lid domain and a substrate covering motif that lies directly on top of the hydrophobic tunnel (Fig. 4a). It is interesting to note that the four α-helices on the top of the C-terminal domain lie directly on the top of the substrate binding site and block the contact

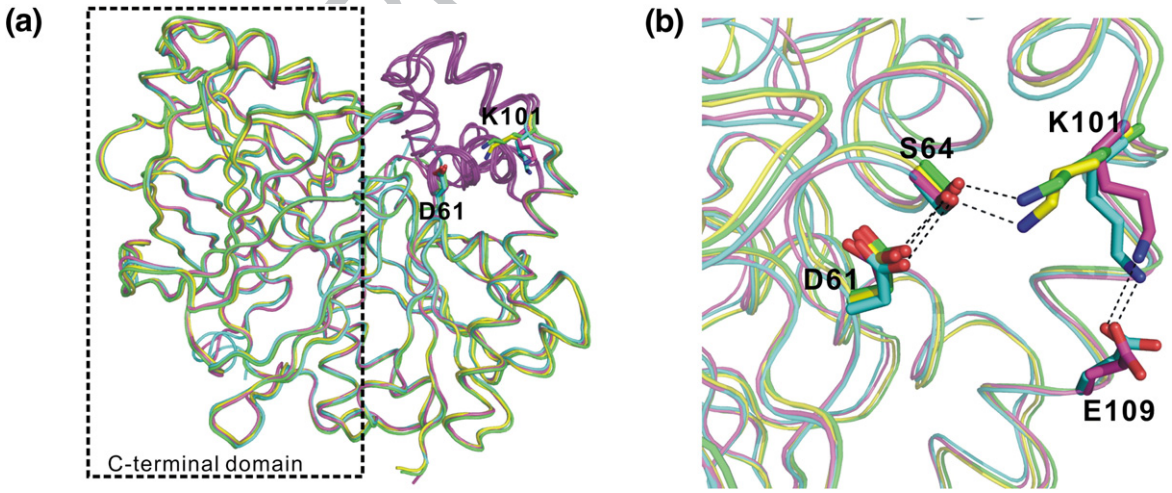

Fig. 3. The pH-dependent conformational change of the hinge residues. (a) Superimposition of the subunits of native AFL (isopropanol) (yellow), mutant AFL S136A (isopropanol) (green), native AFL (PEG 4000) (cyan), and native AFL (PEG 8000) (magenta). The lid domain is shown as a purple loop. The hinge residues Asp61 and Lys101 of each of the four AFL structures are shown as sticks. (b) Details of the hinge residues. The hinge residues and the interacting residues Ser64, Asp61, and Glu109 of each of the four AFL structures are shown as sticks. Hydrogen bonds between the hinge residues and the interacting residues are depicted as black broken lines.

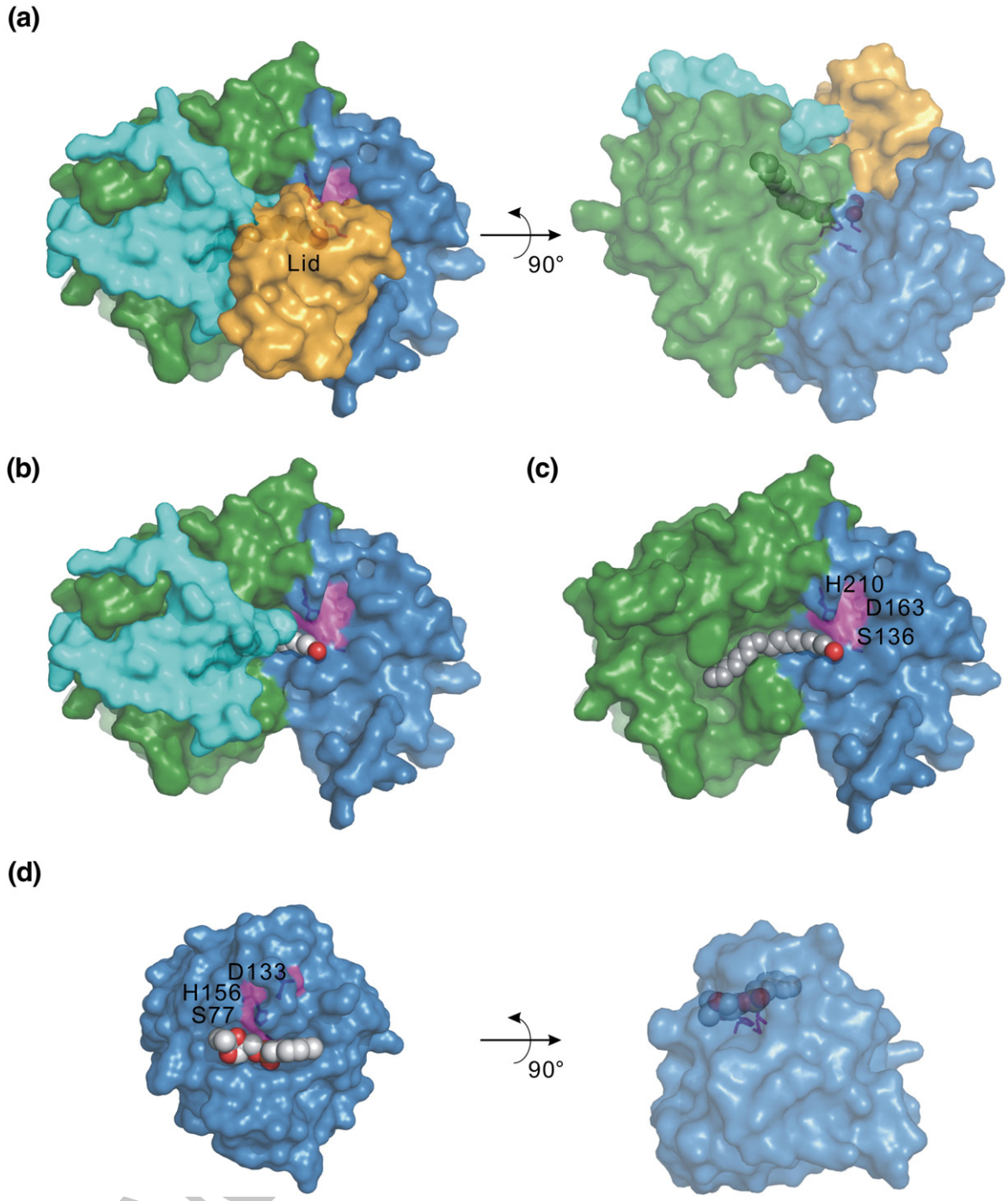

Fig. 4. The substrate binding site of AFL and BSL. (a) The top view and the side view of AFL. Left: the top view of the AFL structure is shown as a surface with the N-terminal domain in blue, the C-terminal domain in green, the lid domain in orange, and the substrate covering motif in cyan. Right: the side view of the AFL structure is obtained by rotating the top view by 90°, with a horizontal axis in the plane of paper. The substrate is shown as a gray sphere. (b) The top view of AFL without the lid domain. The top view of AFL structure is colored as in (a). The catalytic triad is shown as a purple line. (c) The top view of AFL without the lid domain and the substrate covering motif. The top view of the AFL structure is shown in a surface colored as in (a). The catalytic triad is shown as a purple stick. (d) The top view and the side view of BSL. Left: the BSL structure is shown as the surface in blue with its substrate shown as gray spheres and the triad shown as a purple stick. Right: the side view of BSL is obtained by rotating the top view by 90°, with a horizontal axis in the plane of paper.

between the substrate and the outer environment. Access of the substrate to the binding pocket is only possible through an opening of the lid on top of the N-terminal domain. The four helices in the lid might undergo conformational change due to the contact with substrate to allow the entrance of substrate (Fig. 4b). In the AFL lipase, we observed that the carbonyl group of the ligand is facing the catalytic triad and the acyl part extends to the C-terminal domain (Fig. 4c). We also observed some space at the

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017

end of the tunnel toward the C-terminal domain occupied by some water molecule. This shows that the tunnel might be solvent accessible. Whether some conformational change of the substrate covering motif is necessary for the substrate to enter the binding pocket awaits further investigation.

To understand the binding of a lipid substrate in AFL, we compared the structure of AFL with the bound fatty acid fragment with the BSL structure complexed with the covalently bound inhibitor Rc-IPG phosphonate inhibitor.[22] The acyl part binding pocket in AFL is much larger when compared with that of BSL, and consequently, the enzyme–substrate interactions are more extensive than those in BSL. In addition, the ligand binding direction in the active site of AFL and BSL is opposite to each other. In other

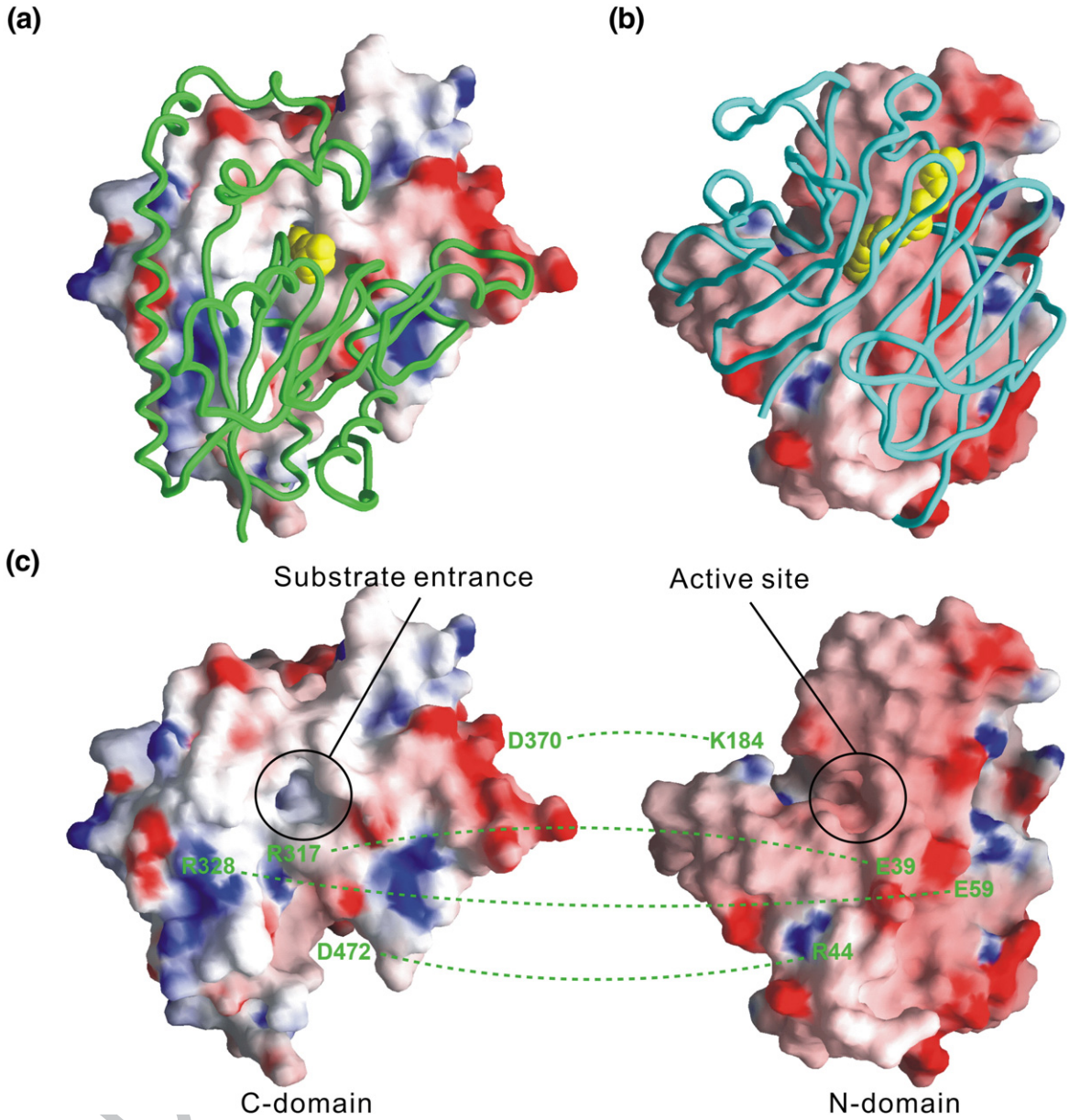

Fig. 5. The domain–domain interface. The molecular surface of the C-terminal and the N-terminal domain color-coded according to electrostatic potential displayed with GRASP.[37] In (a) and (b), one domain is shown as a surface representation, colored according to the electrostatic potentials, and another domain is shown as a worm tracing. Positive charge is in blue and negative charge is in red. (a) The electrostatic potential distribution of the C-terminal domain and the substrate binding site. This figure is viewed by rotating the molecule in Fig. 1 by −90°, with a vertical axis in the plane of the paper. The C-terminal domain of AFL is shown as a surface with the N-terminal domain shown as worm tracing in green. The substrate is shown as yellow spheres. (b) The electrostatic potential distribution of the N-terminal domain and the substrate binding position. This structure is obtained by rotating (a) by 180°, with a vertical axis in the plane of paper. The N-terminal domain of AFL is shown as a surface with the C-terminal domain shown as worm tracing in cyan. (c) The interaction between the two domains. An associated interaction can be formed by rotating each molecule by 90° inwards around the vertical axis. The two surfaces are complementary in charge. The interacting residues between the two domains are indicated.

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017

words, the acyl part of BSL ligand extends to the catalytic N-terminal domain, which is in a completely opposite direction to that of AFL (Fig. 4d).

The domain–domain interface

The residues involved in the interactions can be identified by calculation of the solvent-accessible surface areas before and after domain–domain interaction. The exposed and buried surfaces were calculated with Areaimol in CCP4 using a solvent probe radius of 1.4 Å.[23] The solvent-accessible surface areas calculated for each of the domains are 2049 Å^2 or 17% for the C-terminal domain and 2031 Å^2 or 18% for the N-terminal domain of AFL (Supplementary Table 2). Figure 5 shows the molecular surface at the interface of the two domains with color coded by electrostatic potential. The interacting surfaces are highly hydrophobic and spotted with charged areas (Fig. 5a and b). Interestingly, the charges are complementary from the two domains. There are four ion pairs (Glu39-Arg317, Arg44-Asp472, Glu59-Arg328, and Lys184-Asp370) involved in the surface interactions (Fig. 5c). The detailed interactions of the four ion pairs are shown in Supplementary Fig. 2. The electrostatic interactions must play one of the most important roles in the domain–domain interactions with which the tight bipartite structure is formed, resulting in a highly thermostable structure. In addition to the complementary charge pair between the two domains, extensive hydrophobic interaction also plays an important role in stabilizing the structure of AFL. There are 12 and 21 amino acids from the N- and C-terminal domains, respectively, participating in the hydrophobic interaction (Supplementary Table 2). Thus, our analysis suggests that the higher thermostability of mature AFL (mAFL) as compared to C-domain-truncated AFL (tAFL) may be attributed to a larger buried surface area and extra electrostatic networks embedded in the nonpolar interface between the C-terminal and the N-terminal domains in the mAFL enzyme.

AFL is a true lipase that contains a unique C-domain structure for substrate binding and catalysis

As shown in Fig. 6a, interfacial activation effect of AFL was observed with tricaprylin as a substrate. A drastic increase in lipase activity occurred when the solubility limit of tricaprylin was exceeded. This means that the apparent rate of hydrolysis correlates with the degree of micellar formation. Consistent with observations from other lipases, the lid conformation may change from the closed to the open form in the presence of lipid interface. Together with the evidence of the lid structure in AFL, we classify AFL as a lipase rather than as a carboxylesterase.

To obtain insight into the function of the C-terminal domain and the substrate preference of AFL, we performed activity measurements using

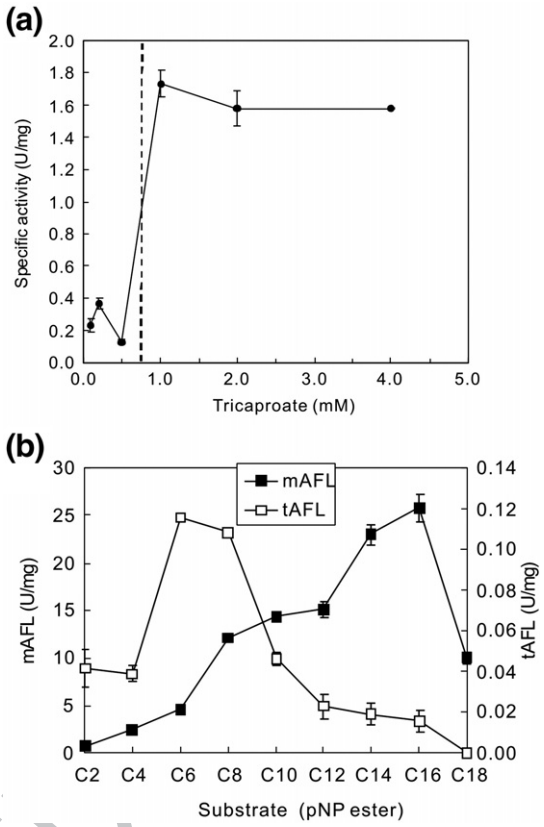

Fig. 6. (a) Interfacial activation effect of AFL. The enzyme was assayed in the presence of different tricaprylin concentrations under the conditions of lipase activity assay. The broken line indicates the limit of solubility of tricaprylin. (b) Substrate specificity of recombinant AFL (mAFL, ■) and tAFL (□) in the hydrolysis of *p*-nitrophenyl (pNP) esters containing fatty acids of various chain lengths. One unit of esterase activity is the amount of enzyme that hydrolyzes 1.0 μmol of pNP ester per minute at 60 °C and pH 8.0. Values are means±SD from three independent experiments.

pNP ester as the substrate to compare the substrate spectrum between the mAFL (no signal peptide) and tAFL. Activity measurements are provided in Fig. 6b. The mAFL preferentially hydrolyzed esters of long-chain fatty acids (C10–C16) and especially those of C16. In contrast, tAFL preferentially hydrolyzed esters of short-chain fatty acids (C4–C10) and especially those of C6, but with a much lower activity. Without the C-terminal domain, AFL lost its function to hydrolyze long-chain ester substrates. Although mAFL shows its highest activity at a substrate length of C16, it still retains a lower activity to substrate with longer chain length such as C18. These results suggest that the C-terminal domain of AFL is essential for its binding to the long-chain substrate.

To further explore the function of the unique C-terminal domain, we also investigated its role in thermostability and pH stability. We discovered that the deletion of the C-terminal domain reduces the optimal temperature from 90 to 80 °C (Fig. 7a). The lack of the C-terminal domain also causes the

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017

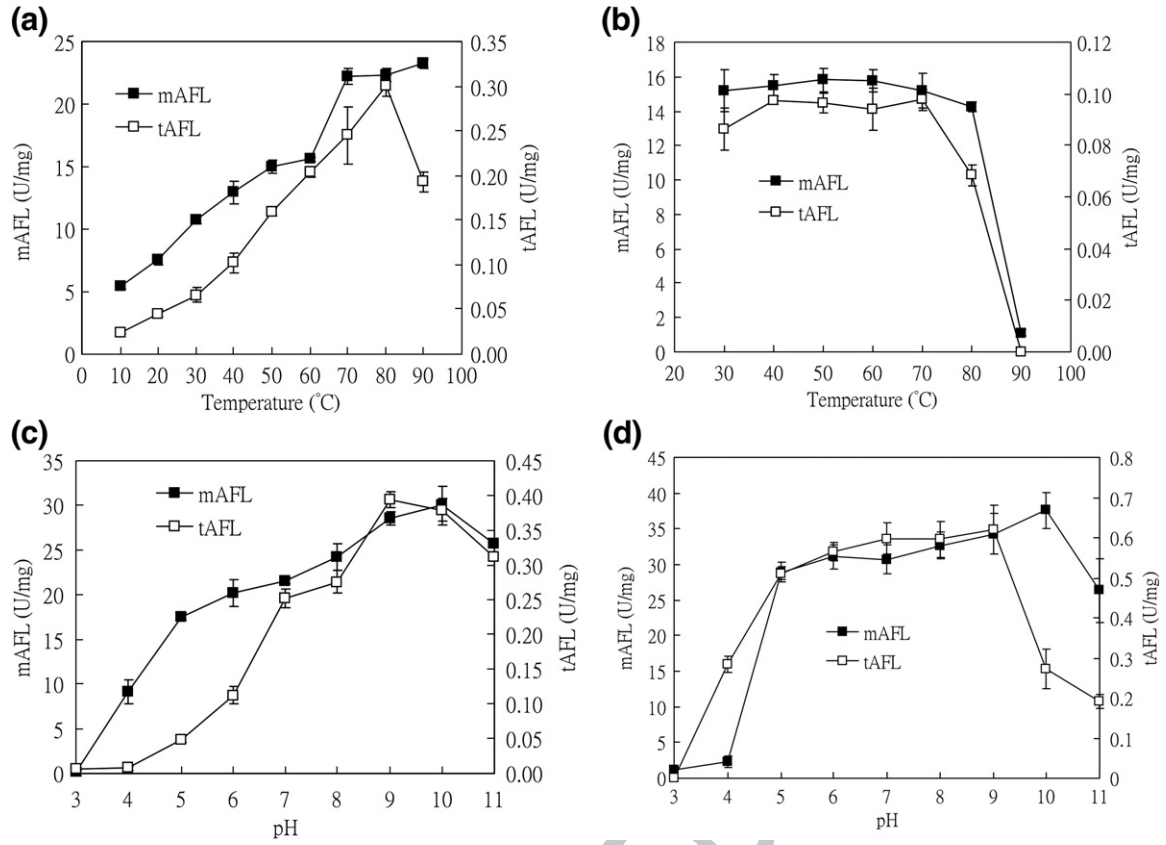

Fig. 7. Optimum temperature, thermostability, optimum pH, and pH-stability assay of recombinant AFL (mAFL, ■) and tAFL (□). (a) Determination of optimum temperature. The optimal temperature was investigated in the range of 10–90 °C at pH 8.5 using *p*-nitrophenyl caprate (0.2 mM) as a substrate. Optimal temperature: 90 °C, 23.27 U/mg for mAFL; 80 °C, 0.30 U/mg for tAFL. (b) Thermostability assay. To analyze the thermostability of AFL, we incubated the lipase enzyme for 30 min at various temperatures in the range of 30–90 °C. Residual activity was determined at 60 °C, pH 8.5, using pNPP (0.2 mM) as a substrate. (c) Determination of optimum pH. The optimal pH was investigated in the pH range of 3.0–11.0 at 60 °C using Good's buffer [50 mM each of *N*,*N*-bis(2-hydroxyethyl)glycine, CAPS, sodium acetate, and Bis-Tris propane] containing 0.5% Triton X-100. pH effects were assessed using pNPP (0.2 mM) as a substrate for mAFL and *p*-nitrophenyl caproate (0.5 mM) for tAFL. Optimal pH: pH 10.0, 30.02 U/mg for mAFL; pH 9.0, 0.39 U/mg for tAFL. (d) pH-stability assay. The pH stability of AFL was determined by incubating the enzyme preparation in buffers of different pH values in the range of 3.0–11.0 for 2 h at 60 °C. Residual activity was determined at pH 8.5 at 60 °C using pNPP (0.2 mM) as a substrate for mAFL and *p*-nitrophenyl caproate (0.5 mM) for tAFL.

protein to be less thermostable (the activity drastically reduced above 70 °C in tAFL instead of 80 °C in mAFL) (Fig. 7b). However, the deletion also changes the optimal pH and pH stability of mAFL. It reduces the optimal pH from 10.0 to 9.0 (Fig. 7c) and causes the protein to be less pH stable (the activity drastically reduced above pH 10.0 in tAFL instead of pH 11.0 in mAFL) (Fig. 7d). Compared to the C-terminal domain's role in thermostability and pH stability, it appears that this domain plays a more important role in catalytic efficiency. Under the optimal temperature and optimal pH, the activity of mAFL is around 77 times higher than that of tAFL. Moreover, the deletion also slightly changes the optimal temperature and pH of mAFL. Therefore, it is likely that the C-terminal domain not only plays an important role in substrate specificity and catalytic efficiency but also attributes partly to its thermophilicity and alkalophilicity.

In an earlier article,[14] the authors used an inactive premature form of AFL (pAFL), which contains a signal peptide and shows low activity towards long-chain triglycerides. We have measured the specific activities (in units per milligram) of mAFL and pAFL using the colorimetric pNPP assay. mAFL shows a 34-fold higher activity than pAFL at 60 °C. As shown in Fig. 7a, mAFL shows high activities in the temperature range from 70 to 90 °C, while pAFL has the highest activity at 70 °C in the article by Rusnak *et al.*[14] In Fig. 7b, mAFL has high residual activity after incubation at the temperatures up to 80 °C (at least 90%), while pAFL shows low residual activity even at 40 °C (about 40%). Obviously, these two forms of AFL display different optimal temperatures and thermostabilities.

The presence of a C-terminal non-catalytic domain of about 200 residues for substrate binding distinguishes AFL from all other lipases whose structures have been determined. The AFL is the first lipase structure discovered so far to contain a C-terminal domain, which enables the binding of long-chain substrate. Although there are several lipases with a

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017

C-terminal domain, their C-domain function and composition are totally different from that of AFL. For example, pancreatic lipases (PLs) contain a C-terminal domain belonging to the C2 family, which is a β-sandwich composed of around 100 amino acids. The β-sandwich C-terminal domain of PL plays an important part in the binding process between the PL and colipase, the specific PL cofactor.[24] The C-terminal domains of AFL and PL not only differ in their function but also show a large difference in their size and the interaction with the N-terminal domain. The C-terminal domain of AFL forms extensive interaction with the N-terminal domain to enable the binding of long-chain substrate and help protein stability. However, the C-domain of PL does not interact extensively with the N-domain; neither does it participate in substrate binding. A search of the C-terminal domain of AFL with DALI uncovered a number of similar structures in the Protein Data Bank (PDB) with *Z* scores ranging from 7.0 to 2.0. Among these structures, many have not been functionally characterized and, hence, did not provide any useful functional information. However, 5 out of the top 10 hits with known functions are transthyretins (TTRs).[25–29] TTR is a tetrameric protein composed of four identical subunits, each consisting of 8 antiparallel β-strands, in contrast to the C-terminal domain of AFL with 14 β-strands. The substrate binding modes of AFL and TTR also show a large difference. The function of TTR is to transport thyroid hormones and other hydrophobic substances in the blood. Although the C-terminal domain of AFL does not superimpose well with the TTR proteins, their similarity in folding pattern infers that the C-terminal domain of AFL might be able to bind or transport hydrophobic substances.

In conclusion, the unique C-terminal domain in AFL enables it to bind and react with long-chain substrates that probably form in the growing environment of *A. fulgidus*. This unique domain also helps to increase AFL's catalytic efficiency to a large extent. Furthermore, the presence of the extensive inter-domain interactions contributes to the thermostability of AFL. As deduced from the biochemical data and crystal structures, we defined the role of the C-terminal β-barrel domain of the AFL as an anchoring domain for its substrate as well as providing the stability force. The observation of the pH-dependent conformational changes associated with the two hinge residues connecting the lid domain and the catalytic domain may explain the alkalophilic nature of this novel lipase.

Materials and Methods

Materials

PfuTurbo DNA polymerase was obtained from Invitrogen (Carlsbad, CA). The plasmid miniprep kit, DNA gel extraction kit, and Ni^{2+}-NTA resin were purchased from Qiagen (Valencia, CA). The protein expression kit [including the pET23a vector and competent JM109 and BL21 (DE3) cells] was obtained from Novagen (Madison, WI). The QuikChange Site-Directed Mutagenesis Kit was obtained from Stratagene, Inc. (La Jolla, CA). All commercial buffers and reagents were of the highest grade.

Protein expression and purification of the truncated, native, and mutant AFL

AFL gene was amplified by PCR from genomic DNA of *A. fulgidus* with primers 5′-AFL (5′-ATACAT**ATG**GCGGAAGACTTTAGACCG-3′) and 3′-AFL (5′-CAGTCGACA**TTA**AATGTAATCCGAAAA-3′) (restriction sites are underlined). The PCR product encoding AFL was cloned into expression vector pET-23a (Novagen), and DNA sequencing was performed to confirm the appropriate orientation. This construct, with a $(His)_6$ tag attached to the C-terminal, was transferred into *E. coli* BL21 (DE3) (Novagen) competent cells.

The hexa-His-tagged wild-type and mutant AFL proteins were overexpressed in BL21 (DE3). Cells were cultured in Difco Luria–Bertani broth containing 50 mg/L ampicillin to an optical density at 600 nm of 0.5–0.6 and then induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside to induce protein expression. Cells were grown for 15 h at 30 °C. The cells were then harvested by centrifugation at 4000*g* for 15 min and disrupted by a French press in a buffer containing 25 mM Tris–HCl (pH 7.5), 150 mM NaCl, and 20 mM imidazole. The homogenate was centrifuged at 27,000*g* for 30 min, and the supernatant was recovered as a crude extract, which was then heated at 60 °C for 30 min to denature unwanted proteins. The denatured proteins were removed by centrifugation at 27,000*g* for 30 min. The supernatant containing the thermo-tolerant AFL was recovered, and the cell-free extract was loaded onto a Ni^{2+}-NTA column, which had been equilibrated with lysis buffer. The column was washed with lysis buffer, and the $(His)_6$-tagged AFL was subsequently eluted by a linear gradient of imidazole concentration from 30 to 500 mM. The enzyme was dialyzed 3 times each with 5 L of buffer (20 mM Tris–HCl, pH 8.0) for 6 h. The resulting AFL solution was further purified by applying it to a diethylaminoethyl cellulose column (2.2 cm × 50 cm) equilibrated with 20 mM Tris–HCl (pH 8.0). The column was washed with 500 mL of the same buffer containing 50 mM NaCl, and AFL proteins were then eluted with 1000 mL of a linear gradient of 50 to 500 mM NaCl in the same buffer. The purified $(His)_6$-tagged AFL was dialyzed three times against 5 L of buffer (20 mM Tris–HCl, pH 8.0), concentrated by Centricon Plus-20 with 10 kDa cutoff membrane (Millipore, Massachusetts, USA), and stored at −80 °C. The purities (>95%) were estimated by SDS/PAGE.

Crystallization and data collection for AFL

AFL crystals were obtained in three different conditions using the hanging drop method (Hampton Research, Laguna Niguel, CA). Native (isopropanol) and S136A AFL crystals were obtained by mixing 2 μL of an AFL solution (5–10 mg/mL AFL in 20 mM Tris–HCl, pH 8.0) with 2 μL of mother liquor (0.2 M $CaCl_2$, 0.1 M sodium acetate, pH 4.6, and 10% isopropanol), and equilibrating with 500 μL of the mother liquor. Native (PEG 4000) AFL crystal was crystallized by mixing 2 μL of an AFL solution (5–10 mg/mL AFL in 20 mM Hepes, pH 7.5) with 2 μL of mother liquor (0.16 M $MgCl_2$, 0.08 M Tris–HCl, pH 8.5, 24% PEG 4000, and 20% glycerol). The native crystals

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017

(PEG 8000) were obtained by mixing 2 μL of an AFL solution (5–10 mg/mL AFL in 20 mM Hepes, pH 7.5) with 2 μL of mother liquor (20% PEG 8000, 100 mM CAPS, pH 10.5, and 200 mM NaCl). The Form I crystals grew to 0.5 mm × 0.2 mm × 0.2 mm in 4 days, at room temperature, and then were soaked briefly in solution containing crystallization buffer with 10% (v/v) perfluoropolyether as cryoprotectant.

X-ray diffraction data were collected at beam line BL13B1 of the National Synchrotron Radiation Research Center (NSRRC, Hsinchu, Taiwan), Taiwan Contract BL12B2 station at SPring-8 (Hyogo, Japan), and Advanced Light Source (Lawrence Berkeley National Laboratory). Diffraction data were processed and scaled by using the program HKL2000.[30] The data sets we obtained can be divided into three different crystal forms. The native AFL crystal (isopropanol), the S136A mutant crystal, the AFL-$KAu(CN)_2$, and the AFL-K_2PtCl_4 crystals belonged to the $P2_12_12_1$ space group with typical unit cell parameters of $a=89$–91 Å, $b=105$–107 Å, and $c=117$–118 Å (Form I). Each asymmetric unit contained two AFL molecules. The native (PEG 4000) crystal belonged to the $P2_12_12_1$ space group with unit cell parameters of $a=52$ Å, $b=106$ Å, and $c=175$ Å (Form II). There are also two AFL molecules in each asymmetric unit. The other native AFL crystal (PEG 8000) in the $P4_3$ space group had typical unit cell parameters of $a=b=99.55$ Å and $c=59.02$ Å (Form III). Each asymmetric unit contained one AFL molecule.

Structure determination and refinement

Two AFL MIR data sets were collected at beam line BL13B1 of the NSRRC. The phase angles of the AFL crystal were calculated using $KAu(CN)_2$ and K_2PtCl_4 heavy-atom derivative data by the MIR method using SOLVE[31] (Supplementary Table 1). The MIR map at 1.83 Å was subject to maximum likelihood density modification followed by autotracing using RESOLVE.[32] An initial model was built using O.[33] Density modification with non-crystallographic symmetry averaging, using the program CNS,[34] improved the phase angles and enabled construction of a more complete model by RESOLVE. Manual building of the remaining model and further refinement were carried out with the programs O, XtalView,[35] and CNS. Prior to use in structural refinements, 5% of the randomly selected reflections were set aside for calculating R_{free} as a quality monitor.[36] The other three crystal structures were determined by using molecular replacement. Refinement statistics of AFL are summarized in Table 1. The residues in the generously allowed and disallowed region are 0.5%. This is because residues Ser136/Ala136 and Ala255 form glycine-like conformation in all four AFL structures. The glycine-like conformation plays an important role in the function of the active site. The molecular figures were produced by using PyMOL (DeLano Scientific)‡.

Site-directed mutagenesis of AFL

AFL mutant S136A was prepared by using a QuikChange Site-Directed Mutagenesis Kit in conjunction with the AFL gene template in the pET-23a (+) vector. The basic procedure of mutagenesis utilizes a supercoiled double-stranded DNA vector with an insert of interest and two synthetic oligonucleotide primers containing the desired mutation. The mutation was confirmed by sequencing the entire AFL mutant gene of the plasmid obtained from the overnight culture. The correct construct was subsequently transformed to *E. coli* BL21 (DE3) for protein expression and purification.

Construction of the C-terminal-truncated mutant

To truncate the AFL open reading frame by engineering a terminator after amino acid T235, we PCR amplified a DNA fragment from the full-length AFL using a sense primer (5′-ATACATATGGCGGAAGACTTTAGACCG-3′) and an antisense primer (5′-CAGTCGACA**TTA**GGTTGTTGCTGGCTTGTA-3′) (restriction sites are underlined). The sequence of the antisense primer was complementary to nucleotides 688 to 705 of the AFL open reading frame and included one engineered stop codon. An NdeI-SalI-treated DNA fragment of PCR product was ligated into the pET-23a (+) vector.

Enzyme assay

The initial rates of lipase activities were measured spectrophotometrically (Hitachi U-2010) using *p*-nitrophenyl esters as substrates. The hydrolysis of *p*-nitrophenyl esters was carried out at 60 °C in 500 μL of reaction buffer (20 mM Tris–HCl, 0.1% gum arabic, and 0.2% sodium deoxycholate, pH 8.5) containing the correspondent *p*-nitrophenyl ester (0.2 mM). The increase in absorbance was recorded for 10 min at 405 nm. One unit of activity was defined as the quantity of enzyme necessary to release 1 μmol of *p*-nitrophenol per minute under the above conditions.

Interfacial activation assay

The lipolytic activity was evaluated titrimetrically using tricaprylin ($C_{8:0}$) as a substrate. The release of free fatty acids was monitored continuously by titration using 10 mM NaOH with a pH-Stat (Radiometer TIM854, Copenhagen, Bagsvaerd, Denmark), adjusted to a constant end-point value of pH 8.0. The substrate emulsification reagent contains 60 mM NaCl, 0.4 mM Tris, 10.8% glycerol, and 0.4% gum arabic. Each assay was carried out in a 60 °C thermostatted reaction vessel containing 20 mL of various concentrations of substrate emulsion. One unit of lipase activity was defined as the amount of lipase necessary to produce 1 μmol of fatty acid per minute under the assay conditions.

PDB accession numbers

The coordinates of native AFL (isopropanol), native AFL (PEG 4000), native AFL (PEG 8000), and mutant AFL S136A (isopropanol) have been deposited in the PDB and are available under accession numbers 2ZYI, 2ZYR, 2ZYS, and 2ZYH, respectively.

Acknowledgements

We thank NSRRC (Taiwan), Advanced Light Source (USA), and SPring-8 (Japan) for beam time allocation and Hui-Ling Shr at the National Core

‡ http://www.pymol.org

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017

Facility of X-ray Crystallography, the Institute of Biological Chemistry, Academia Sinica, for crystal screening.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2009.05.017

References

1. Schrag, J. D. & Cygler, M. (1997). Lipases and alpha/beta hydrolase fold. *Methods Enzymol.* **284**, 85–107.
2. Jaeger, K. E., Ransac, S., Dijkstra, B. W., Colson, C., van Heuvel, M. & Misset, O. (1994). Bacterial lipases. *FEMS Microbiol. Rev.* **15**, 29–63.
3. Bornscheuer, U. T., Altenbuchner, J. & Meyer, H. H. (1999). Directed evolution of an esterase: screening of enzyme libraries based on pH-indicators and a growth assay. *Bioorg. Med. Chem.* **7**, 2169–2173.
4. Houde, A., Kademi, A. & Leblanc, D. (2004). Lipases and their industrial applications: an overview. *Appl. Biochem. Biotechnol.* **118**, 155–170.
5. Jaeger, K. E. & Reetz, M. T. (1998). Microbial lipases form versatile tools for biotechnology. *Trends Biotechnol.* **16**, 396–403.
6. Nardini, M. & Dijkstra, B. W. (1999). Alpha/beta hydrolase fold enzymes: the family keeps growing. *Curr. Opin. Struct. Biol.* **9**, 732–737.
7. Arpigny, J. L. & Jaeger, K. E. (1999). Bacterial lipolytic enzymes: classification and properties. *Biochem. J.* **343**, 177–183.
8. Gilbert, E. J. (1993). *Pseudomonas* lipases: biochemical properties and molecular cloning. *Enzyme Microb. Technol.* **15**, 634–645.
9. Svendsen, A., Borch, K., Barfoed, M., Nielsen, T. B., Gormsen, E. & Patkar, S. A. (1995). Biochemical properties of cloned lipases from the *Pseudomonas* family. *Biochim. Biophys. Acta*, **1259**, 9–17.
10. Gupta, R., Gupta, N. & Rathi, P. (2004). Bacterial lipases: an overview of production, purification and biochemical properties. *Appl. Microbiol. Biotechnol.* **64**, 763–781.
11. Beeder, J., Nilsen, R. K., Rosnes, J. T., Torsvik, T. & Lien, T. (1994). *Archaeoglobus fulgidus* isolated from hot North Sea oil field waters. *Appl. Environ. Microbiol.* **60**, 1227–1231.
12. Turner, P., Mamo, G. & Karlsson, E. N. (2007). Potential and utilization of thermophiles and thermostable enzymes in biorefining. *Microb. Cell Fact.* **6**, 9–23.
13. Yano, J. K. & Poulos, T. L. (2003). New understandings of thermostable and peizostable enzymes. *Curr. Opin. Biotechnol.* **14**, 360–365.
14. Rusnak, M., Nieveler, J., Schmid, R. D. & Petri, R. (2005). The putative lipase, AF1763, from *Archaeoglobus fulgidus* is a carboxylesterase with a very high pH optimum. *Biotechnol. Lett.* **27**, 743–748.
15. Dartois, V., Coppée, J. Y., Colson, C. & Baulard, A. (1994). Genetic analysis and overexpression of lipolytic activity in *Bacillus subtilis*. *Appl. Environ. Microbiol.* **60**, 1670–1673.
16. Jaeger, K. E., Dijkstra, B. W. & Reetz, M. T. (1999). Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases. *Annu. Rev. Microbiol.* **53**, 315–351.
17. Gotz, F., Verheij, H. M. & Rosenstein, R. (1998). Staphylococcal lipases: molecular characterisation, secretion, and processing. *Chem. Phys. Lipids*, **93**, 15–25.
18. Rosenstein, R. & Gotz, F. (2000). Staphylococcal lipases: biochemical and molecular characterization. *Biochimie*, **82**, 1005–1014.
19. Gupta, R., Gupta, N. & Rathi, P. (2004). Bacterial lipases: an overview of production, purification and biochemical properties. *Appl. Microbiol. Biotechnol.* **64**, 763–781.
20. Guex, N. & Peitsch, M. C. (1997). SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.
21. Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M. *et al.* (1992). The alpha/beta hydrolase fold. *Protein Eng.* **5**, 197–211.
22. Droge, M. J., Boersma, Y. L., van Pouderoyen, G., Vrenken, T. E., Ruggeberg, C. J., Reetz, M. T. *et al.* (2006). Directed evolution of *Bacillus subtilis* lipase A by use of enantiomeric phosphonate inhibitors: crystal structures and phage display selection. *ChemBioChem*, **7**, 149–157.
23. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **50**, 760–763.
24. Chahinian, H., Sias, B. & Carriere, F. (2000). The C-terminal domain of pancreatic lipase: functional and structural analogies with c2 domains. *Curr. Protein Pept. Sci.* **1**, 91–103.
25. Zanotti, G., Cendron, L., Ramazzina, I., Folli, C., Percudani, R. & Berni, R. (2006). Structure of zebra fish HIUase: insights into evolution of an enzyme to a hormone transporter. *J. Mol. Biol.* **363**, 1–9.
26. Muziol, T., Cody, V. & Wojtczak, A. (2001). Comparison of binding interactions of dibromoflavonoids with transthyretin. *Acta Biochim. Pol.* **48**, 885–892.
27. Jung, D. K., Lee, Y., Park, S. G., Park, B. C., Kim, G. H. & Rhee, S. (2006). Structural and functional analysis of PucM, a hydrolase in the ureide pathway and a member of the transthyretin-related protein family. *Proc. Natl Acad. Sci. USA*, **103**, 9790–9795.
28. Eneqvist, T., Lundberg, E., Karlsson, A., Huang, S., Santos, C. R., Power, D. M. & Sauer-Eriksson, A. E. (2004). High resolution crystal structures of piscine transthyretin reveal different binding modes for triiodothyronine and thyroxine. *J. Biol. Chem.* **279**, 26411–26416.
29. Hennebry, S. C., Law, R. H., Richardson, S. J., Buckle, A. M. & Whisstock, J. C. (2006). The crystal structure of the transthyretin-like protein from *Salmonella dublin*, a prokaryote 5-hydroxyisourate hydrolase. *J. Mol. Biol.* **359**, 1389–1399.
30. Otwinowski, Z. & Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326.
31. Terwilliger, T. C. & Berendzen, J. (1999). Automated MAD and MIR structure solution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **55**, 849–861.
32. Terwilliger, T. (2004). SOLVE and RESOLVE: automated structure solution, density modification and model building. *J. Synchrotron Radiat.* **11**, 49–52.
33. 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 Crystallogr., Sect. A: Found. Crystallogr.* **47**, 110–119.

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017

34. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **54**, 905–921.
35. McRee, D. E. (1999). XtalView/Xfit—a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**, 156–165.
36. Brunger, A. T. (1993). Assessment of phase accuracy by cross validation—the free *R*-value. Methods and applications. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **49**, 24–36.
37. Nicholls, A., Sharp, K. A. & Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281–2961.

Please cite this article as: Chen, C.K.-M. *et al.*, Structure of the Alkalohyperthermophilic *Archaeoglobus fulgidus* Lipase Contains a Unique C-Terminal Domain Essential for Long-Chain Substrate Binding, *J. Mol. Biol.* (2009), doi:10.1016/j.jmb.2009.05.017