Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystal structure analysis of a fatty acid double-bond hydratase from *Lactobacillus acidophilus*

Anton Volkov,[a]‡ Sohail Khoshnevis,[b]‡§ Piotr Neumann,[b]‡ Cornelia Herrfurth,[a] Daniel Wohlwend,[b] Ralf Ficner[b]* and Ivo Feussner[a]

[a]Department for Plant Biochemistry, Albrecht-von-Haller-Institute for Plant Sciences, Georg-August-University, Untere Karspüle 2, 37073 Göttingen, Germany, and [b]Department of Molecular Structural Biology, Institute of Microbiology and Genetics, Georg-August-University, Justus-von-Liebig-Weg 11, 37077 Göttingen, Germany

‡ These authors contributed equally to this manuscript.
§ Current address: Cancer Biology Department, The Scripps Research Institute, Jupiter, FL 33458, USA.

Correspondence e-mail: rficner@gwdg.de

Received 6 December 2012
Accepted 10 January 2013

PDB References: LAH, SeMet derivative,4ia5; LA–LAH, 4ia6

Bacteria have evolved mechanisms for the hydrogenation of unsaturated fatty acids. Hydroxy fatty acid formation may be the first step in such a process; however, knowledge of the structural and mechanistic aspects of this reaction is scarce. Recently, myosin cross-reactive antigen was shown to be a bacterial FAD-containing hydratase which acts on the 9*Z* and 12*Z* double bonds of C_{16} and C_{18} non-esterified fatty acids, with the formation of 10-hydroxy and 10,13-dihydroxy fatty acids. These fatty acid hydratases form a large protein family which is conserved across Gram-positive and Gram-negative bacteria with no sequence similarity to any known protein apart from the FAD-binding motif. In order to shed light on the substrate recognition and the mechanism of the hydratase reaction, the crystal structure of the hydratase from *Lactobacillus acidophilus* (LAH) was determined by single-wavelength anomalous dispersion. Crystal structures of apo LAH and of LAH with bound linoleic acid were refined at resolutions of 2.3 and 1.8 Å, respectively. LAH is a homodimer; each protomer consists of four intricately connected domains. Three of them form the FAD-binding and substrate-binding sites and reveal structural similarity to three domains of several flavin-dependent enzymes, including amine oxidoreductases. The additional fourth domain of LAH is located at the C-terminus and consists of three α-helices. It covers the entrance to the hydrophobic substrate channel leading from the protein surface to the active site. In the presence of linoleic acid, the fourth domain of one protomer undergoes conformational changes and opens the entrance to the substrate-binding channel of the other protomer of the LAH homodimer. The linoleic acid molecule is bound at the entrance to the substrate channel, suggesting movement of the lid domain triggered by substrate recognition.

1. Introduction

Unsaturated fatty acids may be toxic to many bacteria owing to their deteriorating action on their cellular membranes and/or their blocking of fatty acid biosynthesis *via* the inhibition of enoyl-ACP reductase (Greenway & Dyke, 1979; Raychowdhury *et al.*, 1985; Zheng *et al.*, 2005). Hence, a mechanism for the hydrogenation of unsaturated fatty acids has evolved in bacteria. This process has been best described for the di-unsaturated fatty acid linoleic acid (LA; $18{:}2^{\Delta 9Z,12Z}$, where $x{:}y^{\Delta z}$ denotes a fatty acid with x C atoms and y double bonds in positions z counting from the carboxyl end), which is hydrogenated to the saturated stearic acid (18:0; Maia *et al.*, 2007). Different isomers of conjugated linoleic acid (CLA) and hydroxy fatty acids are considered to be intermediates in the hydrogenation process (Nam & Garnsworthy, 2007; Vossenberg & Joblin, 2003; Mosley *et al.*, 2002; Kellens *et al.*,

1986). However, knowledge of the structural and the mechanistic relationship between CLA and hydroxy fatty acid formation is still scarce (Kishino *et al.*, 2011). To date, only the crystal structure of the 10,12-CLA-producing fatty acid double-bond isomerase from *Propionibacterium acnes* (PAI) has provided structural and mechanistic insights into the isomerization reaction (Liavonchanka *et al.*, 2006, 2009).

In the last four years, the myosin cross-reactive antigen (MCRA) has been identified as an FAD-containing hydratase which acts on the 9*Z* and 12*Z* double bonds of C_{16} and C_{18} non-esterified fatty acids with the formation of 10-hydroxy and 10,13-dihydroxy fatty acids in several bacteria (Bevers *et al.*, 2009; Rosberg-Cody *et al.*, 2011; Volkov *et al.*, 2010; Joo, Seo *et al.*, 2012; Kim *et al.*, 2012; Yang *et al.*, 2013). MCRAs form a large family which is broadly conserved across different groups of Gram-positive and Gram-negative bacteria, including pathogenic and commensal species, and has no sequence similarity to any known proteins except for a short stretch of FAD-binding motif. Therefore, knowledge of the three-dimensional structure of an MCRA may lead to an increased understanding of the catalytic mechanism and facilitate further classification of the enzymes into a known protein family.

In the present study, we determined the crystal structure of an MCRA from *Lactobacillus acidophilus* NCFM, here denoted LAH (*L. acidophilus* hydratase), which was recently described as a hydratase (Yang *et al.*, 2013). Its structure revealed similar features to members of the family of flavin-containing amine oxidoreductases such as L-amino-acid oxidase from *Calloselasma rhodostoma* (Pawelek *et al.*, 2000) and PAI. The structural analysis provides the first structural implications of FAD and substrate binding across the superfamily of MCRAs.

2. Materials and methods

Chemicals were obtained from Sigma–Aldrich (Munich, Germany) and Carl Roth & Co. Agarose was from Biozym Scientific GmbH (Hessisch Oldendorf, Germany). All fatty acids were purchased from Sigma–Aldrich (Munich, Germany) or Cayman (Ann Arbor, USA). Acetonitrile was from Fisher Scientific UK Ltd (Loughborough, England). Restriction enzymes were provided by MBI Fermentas (St Leon-Roth, Germany).

2.1. Cloning and expression of recombinant LAH

For expression in *Escherichia coli*, the *lah* (*mcra*) ORF (GenBank accession No. AAV42528.1) was amplified with the following primers (forward primer 5′-GGGTTTCCATATGT-ATTATTCCAATGGTAATTACGAA-3′ and reverse primer 5′-CGCCTCGAGTTAGACTAAATTTGCTTCTTTAAGTA-GTTCTT-3′ containing *Nde*I and *Xho*I recognition sites, respectively) and cloned into pET24a and pET28a expression vectors (Novagen, USA), respectively, for the expression of nontagged and N-terminally His_6-tagged versions, respectively. Adding an N-terminal His_6 tag to the protein resulted in an additional 20 amino acids (MGSSHHHHHHSSGLVPR-GSH) preceding the native N-terminal methionine. The PCR reactions were performed using the Phusion Hot Start High Fidelity DNA Polymerase System (Finnzymes, Espoo, Finland) according to the manufacturer's protocol. The respective clones were analyzed by restriction analysis and sequencing.

The native LAH variants were expressed as described in Volkov *et al.* (2010). The expression of selenomethionine-substituted LAH was performed according to published protocols (Reuter & Ficner, 1999). After IPTG induction, the cells were transferred to 289 K and harvested by centrifugation (10 min at 9100*g*) after 16 h of induction.

2.2. Protein purification

Native and selenomethionine-substituted as well as N-terminally His-tagged LAH were purified as described for SPH in Volkov *et al.* (2010). DTT was added to all purification buffers at a concentration of 2 m*M* for the selenomethionine (SeMet) substituted LAH, except for the final buffer in which the protein was stored and used for crystallization (20 m*M* HEPES–NaOH pH 7.0).

2.3. Cofactor identification

Cofactor identification was performed by UV–Vis spectroscopy and HPLC-MS/MS analysis as described in Volkov *et al.* (2010) for the SPH enzyme. The purified LAH had a yellow colour, which was lost after gel filtration on a Superdex S200 26/60 pg column (GE Healthcare, Munich, Germany) or after extensive washing of anion-exchange or affinity-bound forms of LAH. Nevertheless, the N-terminally His_6-tagged protein retained a strong yellow colour and was essentially pure, with an identical hydratase activity to nontagged LAH.

2.4. GC-MS analysis of lipids

To determine the substrate and product specificity, the LAH enzymes were incubated with the respective fatty acids and the products were extracted and analyzed by GC-MS or GC. The reaction conditions and the setup used were identical to those described for the SPH enzyme (Volkov *et al.*, 2010).

2.5. Crystallization and crystal structure determination

Both native and SeMet-derivative LAH were crystallized using the sitting-drop vapour-diffusion technique (24-well 'Big' Sitting Drop Crystallization Plate, XtalQuest) by mixing protein and reservoir solutions in a 1:1 ratio (2 µl droplet size, 500 µl reservoir volume). Initially, clusters of needles were obtained from N-terminally His-tagged native LAH at a protein concentration of 10 mg ml^{-1} exclusively in conditions D2 and D4 of JBScreen Classic 7 (Jena Bioscience, Jena, Germany). Since optimization of condition D2 (35% MPD, 0.1 *M* imidazole–HCl, pH 8.0) failed to yield well diffracting crystals, untagged LAH protein at a concentration of 12 mg ml^{-1} was used for further crystallization trials with condition D4 (47% MPD, 0.1 *M* HEPES–NaOH, pH 7.5) in combination with microseeding. Well diffracting crystals of tag-free native LAH were obtained using 37% MPD, 0.1 *M*

HEPES–KOH, pH 7.9 with 10 m*M* phenol as an additive and 3.6 m*M* linoleic acid. SeMet-derivative crystals were grown in 37–38% MPD, 0.1 *M* HEPES–KOH, pH 7.85 with 20 m*M* NaBr as an additive. Prior to data collection, crystals were washed in a cryoprotectant solution consisting of the reservoir solution diluted with glycerol at 10%(*v*/*v*) and flash-cooled by immersion in liquid nitrogen. No systematic improvement in the quality of the diffraction pattern was observed using glycerol as an additional cryoprotectant; however, the data sets used for structure determination were obtained from glycerol-treated crystals. Both SAD and native data sets were collected at 100 K on beamline 14.1 at BESSY, Berlin, Germany. The oscillation photographs were indexed, integrated and merged using the *XDS* package (Kabsch, 2010*a*,*b*) to final resolutions of 2.3 and 1.8 Å for SeMet-derivative and native LAH, respectively. Prior to use in structural refinement, a randomly selected 5% of the reflections were set aside for the calculation of $R_{\rm free}$ as a quality monitor (Brünger, 1993).

The structure was solved by SeMet SAD phasing using the peak data set within the *Auto-Rickshaw* software pipeline (Panjikar *et al.*, 2005). Within this pipeline, *SHELXD* (Sheldrick, 2008) was used for the heavy-atom (HA) search and clearly found 36 selenium sites. HA refinement and phase calculations were carried out using *MLPHARE* from *CCP*4 (Winn *et al.*, 2011), followed by density modification and phase extension using *DM* (Cowtan, 1994) and *RESOLVE* (Terwilliger, 2000). The initial model was built by *ARP*/*wARP* (Morris *et al.*, 2004; Perrakis *et al.*, 1999). The model

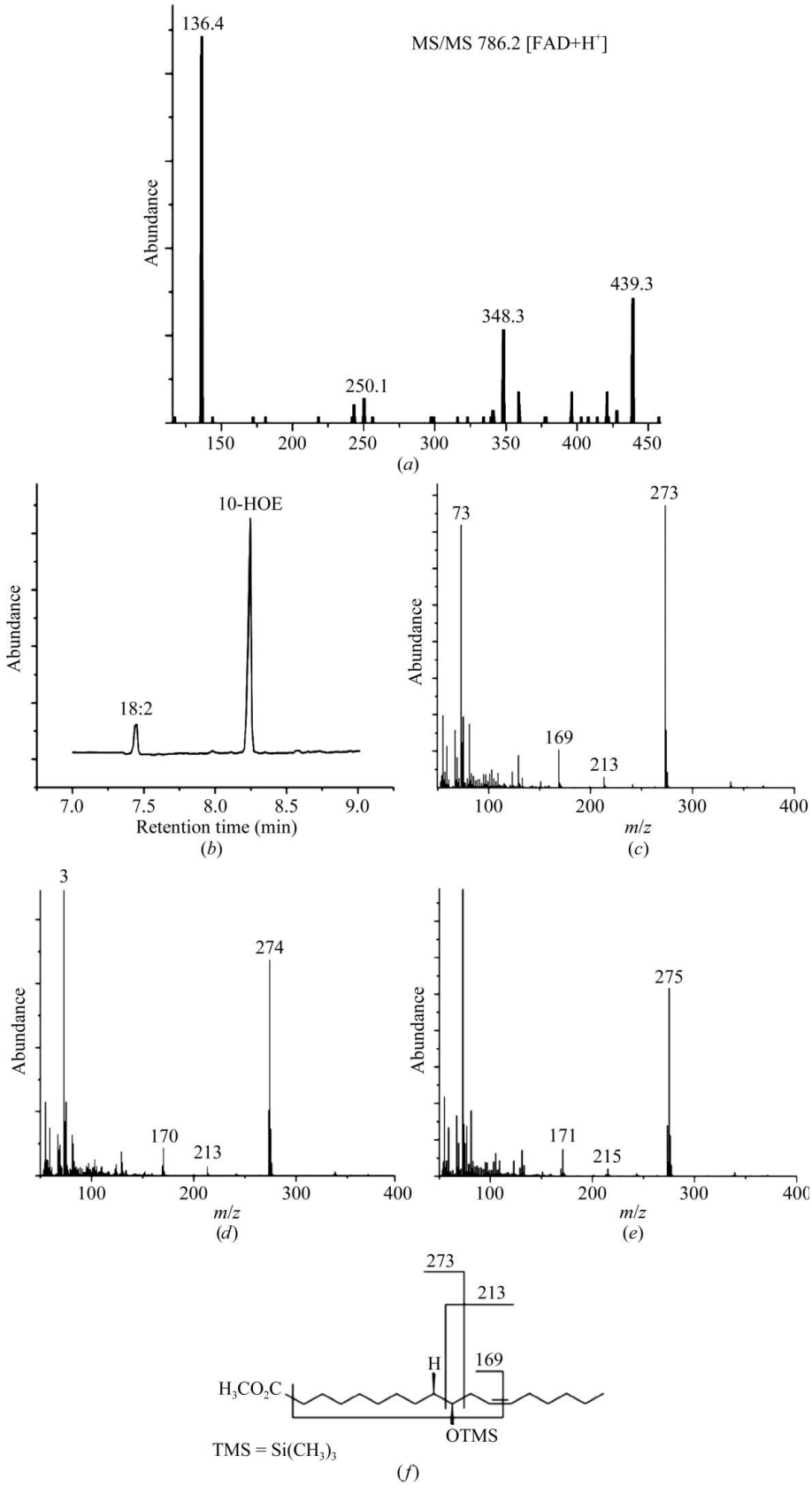


Figure 1
Characterization of LAH. (*a*) MS/MS spectrum of the molecular ion M^+ = 786.1 corresponding to $[FAD+H]^+$. The supernatant from heat-precipitated N-terminally His_6-tagged LAH was desalted and concentrated for further analysis by HPLC-MS/MS. (*b*)–(*e*) Determination of the origin of the hydroxyl group in 10-HOE. LA was used as a substrate and the reaction products were analyzed by GC-MS. The total ion chromatogram of the product analysis is shown in (*b*). Mass spectra of 10-HOE after reaction in H_2O (*c*), in ^2H_2O (*d*) and in $H_2{}^{18}O$ (*e*) show that one water molecule is added to the Δ9 double bond. The mass of the C_{10} fragment (*m*/*z* = 273) changes owing to isotopic labelling with either ^2H at position C9 (+1 Da) or ^{18}O at position C10 (+2 Da). The structure of 10-HOE and its fragmentation pattern is shown in (*f*).

was manually completed in *Coot* (Emsley *et al.*, 2010) and was refined with *PHENIX* (Adams *et al.*, 2010), including TLS, weight optimization and bulk-solvent mask optimization. The native LAH structure was solved using the model of SeMet-substituted LAH subjected to rigid-body refinement (R and R_{free} factors of 0.2589 and 0.2885, respectively) followed by simulated annealing in *PHENIX*. The resulting model was completed by iterative cycles of model building in *Coot* and refinement in *PHENIX* using the same refinement strategy as used for SeMet-substituted LAH.

The atomic coordinates and structure factors have been deposited in the Protein Data Bank with accession codes 4ia5 (apo LAH) and 4ia6 (LA–LAH).

Table 1
Substrate specificity of LAH.

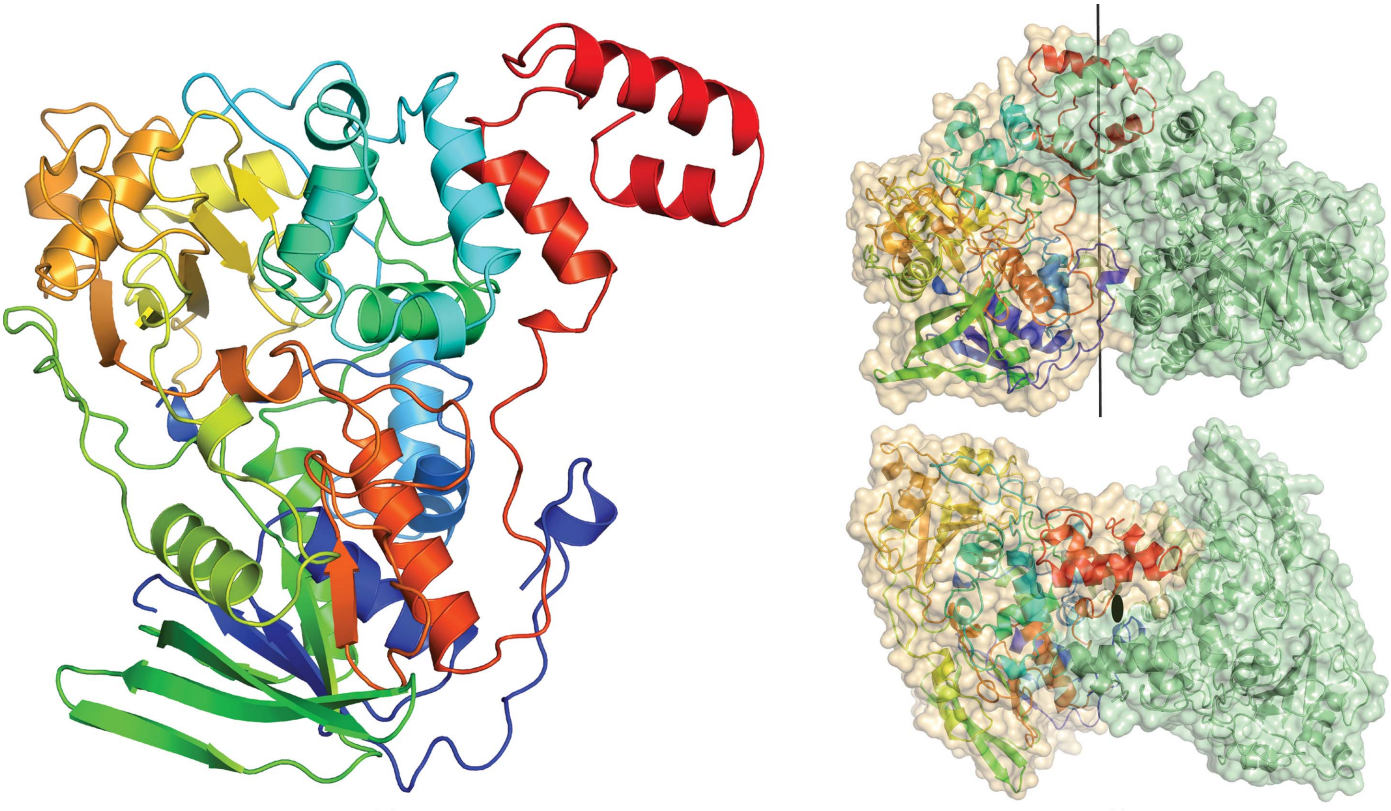
The products were identified by GC-MS analysis after chloroform:methanol (2:1) extraction of the samples; the reactions were performed as described in §2. Two repetitions of each experiment were conducted.

Substrate	Product(s)	Retention time (min)	Fragmentation pattern
14:1$^{\Delta 9Z}$	No products detected		
16:1$^{\Delta 9Z}$	10-Hydroxyhexadecanoate	6.88	273, 187, 169
18:1$^{\Delta 9Z}$	10-Hydroxyoctadecanoate	8.1	273, 215, 169
18:1$^{\Delta 9E}$	No products detected		
18:2$^{\Delta 9Z,\Delta 12Z}$	(12*Z*)-10-Hydroxy-12-octadecenoate	8.33	273, 213, 173, 169
18:3$^{\Delta 9Z,\Delta 12Z,\Delta 15Z}$	(12*Z*,15*Z*)-10-Hydroxy-12,15-octadecadienoate	8.77	273, 241, 169
18:3$^{\Delta 6Z,\Delta 9Z,\Delta 12Z}$	No products detected		
20:3$^{\Delta 11Z,\Delta 14Z,\Delta 17Z}$	No products detected		
20:4$^{\Delta 5Z,\Delta 8Z\Delta 11Z,\Delta 14Z}$	No products detected		
18:2$^{\Delta 9Z,\Delta 12Z}$ methyl ester	No products detected		
Linoleyl-CoA	No products detected		
Dilinoleylphosphatidylcholine	No products detected		
Trilinoleylglycerol	No products detected		

3. Results

3.1. Protein purification, characterization and crystallization

LAH is a protein composed of 591 residues (calculated molecular weight of 67.7 kDa) which could be purified as a yellow-coloured homodimer. Absorption spectroscopy and HPLC-MS/MS analysis revealed the presence of a noncovalently bound oxidized FAD (Fig. 1*a*). The reaction of LAH with LA led to the formation of 10-hydroxyoleic acid (10-HOE), which according to GC-MS was the only reaction product (Figs. 1*b* and 1*c*). The characteristic ion of m/z = 273 (M^+) represents a C_{10}



(*a*) (*b*)

Figure 2
Structure of LAH monomer and dimer. (*a*) Ribbon diagram of LAH protomer (symmetric protomer) in rainbow colouring (from blue at the N-terminus to red at the C-terminus). (*b*) LAH dimer; the intradimer twofold axis of symmetry is depicted as a black stick. (*c*) LAH dimer viewed along the dimerization axis.

fragment in which the 10-hydroxy group is modified to the trimethylsilyl ether (Fig. 1*f*). The detailed substrate profile of LAH with respect to double-bond position, configuration and the polarity of the fatty acid head group is summarized in Table 1. It confirms that LAH, like other hydratases, is only active towards free fatty acids and requires a (9*Z*) double bond, whereas (9*E*), (11*E*) and (11*Z*) double bonds were not hydrated. Notably, double bonds before C9 abolished LAH activity. The chain length of the substrate was limited to 16 and 18 C atoms, respectively. The origin of the hydroxyl group was analyzed in detail for 10-HOE. 10-HOE showed a mass shift of +1 (Figs. 1*d* and 1*e*; m/z = 273 *versus* 274; M^+) for the C_{10} fragment in the presence of ^2H_2O and a +2 Da shift for the C_{10} fragment (Fig. 1*f*; m/z = 275; M^+) after incubation with $H_2{}^{18}O$. The observed increases in mass confirm that the hydroxy group originates from a water molecule.

Interestingly, purification of a nontagged version of LAH resulted in loss of the yellow colour, while N-terminally His_6-tagged LAH remained yellow. Both tagged and nontagged versions of LAH were used in crystallization attempts, but single crystals were only obtained for nontagged LAH by vapour diffusion using 40%(*v*/*v*) MPD as precipitant.

3.2. Overall structure

The crystals of LAH belonged to space group $P2_12_12$ and contained two molecules in the asymmetric unit, corresponding to a solvent content of 54% (Fig. 2). Since no structure of a homologous protein was known, the LAH crystal structure was solved *de novo* by single-wavelength anomalous dispersion using an SeMet derivative crystal. The crystal structures of ligand-free selenomethionine-derivatized LAH (apo LAH) and of native LAH cocrystallized with linoleic acid (LA–LAH) were determined at resolutions of 2.3 and 1.8 Å, respectively. Both structures comprised the entire length of the protein, with the exception of a flexible loop between residues 61 and 72, and were refined with good stereochemistry to reasonable *R* factors (see Table 2 for details). Analysis of the contacts between the two LAH monomers (32 hydrogen bonds and 20 salt bridges) using the *PISA* software (Krissinel & Henrick, 2007) revealed that 9.7% (2289 Å^2) of the surface of each monomer becomes buried upon dimer formation. The solvation free-energy gain upon formation of the interface, Δ^1G, as predicted by *PISA*, was -20.4 kcal mol^{-1}, suggesting that two monomers stably bind to each other to form a dimer. This was supported by gel-filtration chromatography, which demonstrated the presence of both monomeric and dimeric LAH in solution (data not shown).

A search for proteins with overall structural similarity to LAH using the *DALI* server (Holm & Rosenström, 2010) yielded several FAD-binding proteins with high statistical significance, including L-amino-acid oxidase (LAAO) from *C. rhodostoma* (Pawelek *et al.*, 2000), protoporphyrinogen oxidase from *Bacillus subtilis* (POBS; Qin *et al.*, 2010), 6-hydroxy-L-nicotine oxidase from *Arthrobacter nicotinovorans* (Kachalova *et al.*, 2010) and PAI (Liavonchanka *et al.*, 2006), all with *Z*-scores above 17 and a sequence identity to LAH of 10–13% (Fig. 3).

Table 2
Data-collection and refinement statistics.

Values in parentheses are for the outer shell.

	SeMet-derivative apo LAH	Native LA–LAH
Data collection		
Space group	$P2_12_12$	$P2_12_12$
Unit-cell parameters		
a (Å)	168.75	165.10
b (Å)	78.97	77.74
c (Å)	108.77	104.81
Wavelength (Å)	0.97977	0.98410
Resolution (Å)	44.79–2.22 (2.42–2.22)	48.73–1.80 (1.90–1.80)
Observed reflections	374784 (15159)	530308 (62004)
Unique reflections	122663 (9536)	122391 (15622)
Multiplicity	3.1 (1.6)	4.3 (4.0)
Completeness (%)	88.4 (55.6)	97.6 (84.1)
$\langle I/\sigma(I)\rangle$	17.8 (3.9)	13.9 (2.4)
R_{merge}† (%)	4.5 (18.7)	9.6 (65.2)
$CC_{1/2}$‡	99.8 (94.0)	99.7 (72.0)
Phasing		
Anomalous correlation‡ (%)	82 (37)	
Mean anomalous difference (SigAno)‡	2.57 (1.09)	
No. of sites/% of substituted Met	36/100	
FOM after *DM* and phase extension	0.861	
Refinement		
Resolution (Å)	44.79–2.22	48.73–1.80
R_{work} (%)	21.38 (31.78)	15.37 (25.58)
R_{free}§ (%)	25.85 (36.56)	19.15 (26.90)
No. of atoms		
Total	9939	10974
Protein	9375	9331
Solvent	474	1431
Ligands and ions	90	212
B factors (Å^2)		
Average	64.7	20.5
Protein	65.4	18.5
Solvent	51.4	31.8
Ligands and ions	68.8	34.2
R.m.s.d. from ideal		
Bond lengths (Å)	0.007	0.012
Bond angles (°)	1.194	1.376
Ramachandran plot (%)		
Favoured	98.17	98.16
Outliers	0.09	0.00
Allowed	1.74	1.84
PDB code	4ia5	4ia6

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ Calculated with *XSCALE* (Karplus & Diederichs, 2012). § R_{free} was calculated for a randomly chosen 5% of reflections that were not included in the refinement.

Based on these similarities, four intricately connected domains could be assigned for LAH. The first three domains (residues 1–538) build up the major part of the protein and, despite low sequence similarity, reveal remarkable structural similarity to the distinct domains of the FAD-dependent enzymes noted above (Fig. 3). Domain 1 (residues 1–119, 218–291, 312–335 and 471–538) is a mixed α/β domain composed of a parallel five-stranded β-sheet packed between two α-helices on one side and a three-stranded antiparallel β-sheet on the other side, resembling a variant of the Rossmann fold found in many dinucleotide-binding proteins (Fig. 3, blue). Domain 2 (residues 122–134, 292–311 and 336–470) consists of an antiparallel β-sheet flanked by three α-helices and forms the

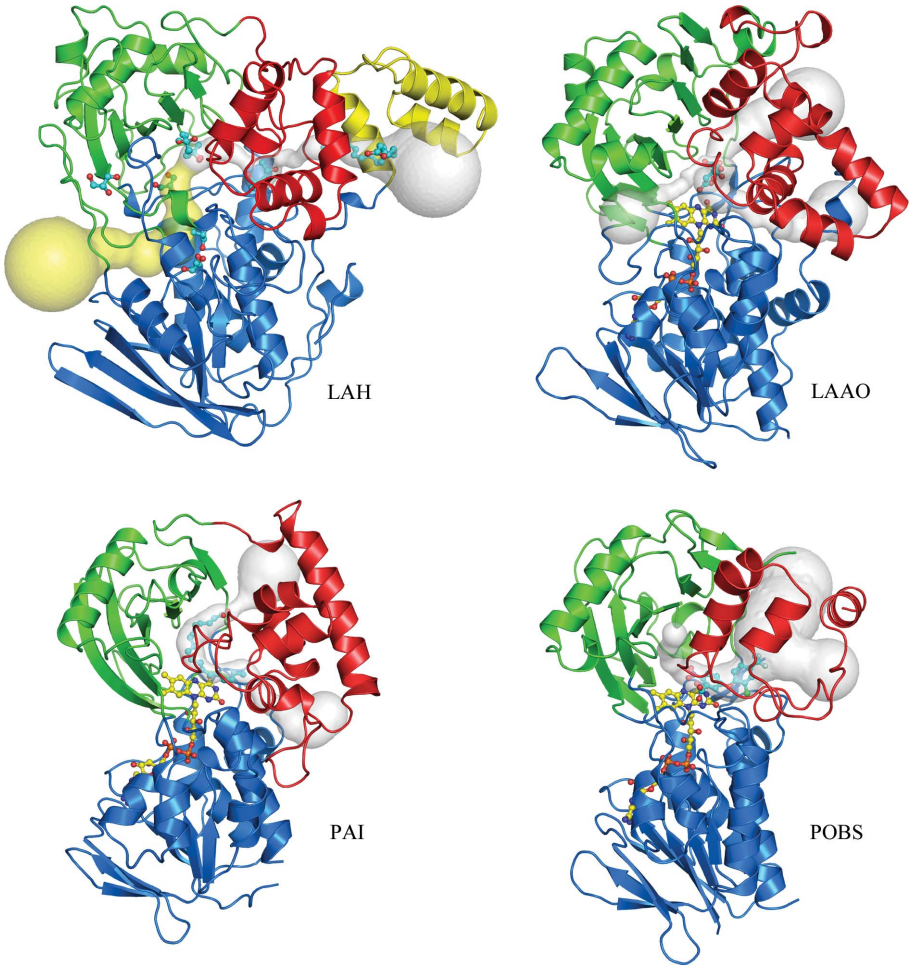


Figure 3
Comparison of LAH with three structurally related FAD-binding proteins. Domains 1, 2 and 3 are coloured marine, light green and red, respectively. The additional domain in LAH (domain 4) is depicted in yellow. Bound FAD and substrate/solvent molecules are shown in ball-and-stick representation and are coloured yellow and cyan, respectively. Channels leading to the known (PAI, LAAO and POBS) or putative (LAH) active site were calculated using the *MOLE* software (Petrek *et al.*, 2007) and are depicted as transparent white surfaces. The putative FAD-binding site of LAH is depicted as a yellow surface.

FAD- and substrate-binding sites in concert with domain 1 (Fig. 3, green). Domain 3 (residues 135–217) is mostly α-helical and closely resembles the substrate-binding domain of monoamine oxidases (Fig. 3, red). Domain 4 (residues 539–591) is made up of three α-helices (h1, residues 540–551; h2, residues 564–576; h3, residues 580–588) connected by two loops, corresponds to the C-terminal region of the protein and is juxtaposed with the substrate-binding domain 3 (Fig. 3, yellow). It has no structural equivalent domain in FAD-containing oxidases or isomerases (see above), but shows structural similarity to the N-terminal lid domain of the active site of the long-chain acylglycerol lipase from the alkalo-hyperthermophile *Archaeoglobus fulgidus* (Chen *et al.*, 2009).

3.3. Structural differences between LAH monomers

LAH forms a symmetric dimer in which its protomers, in the absence of substrate, do not show significant conformational differences. This is reflected by a low root-mean-square deviation (r.m.s.d.) of 0.63 Å as calculated for all C^{α} atoms. Domain 4 of apo LAH follows the local intradimer twofold axis of symmetry (Figs. 2*b* and 2*c*), even though it exhibits higher flexibility and forms different numbers of crystal contacts (0 *versus* 16 for the two protomers of the apo LAH homodimer). However, in the presence of the LA molecule the helices h2 and h3 of domain 4 that are not involved in any crystal contacts in apo LAH are repositioned in one LA–LAH protomer (Fig. 4*a*). In this asymmetric protomer, residues 556–574, which form a loop and part of helix h2 in apo LAH, are not defined in the electron-density map and are most likely to be disordered, while the remaining and traceable C-terminal α-helix h3 is extended by five residues and therefore comprises residues 575–588 (Fig. 4*a*). In addition, the partially refolded helix h3 shows a significant (∼9 Å) displacement in comparison with the position of the equivalent fragment observed in the symmetric protomer (Fig. 4*a*).

3.4. Substrate-binding mode

Comparison of the apo LAH and LA–LAH structures reveals the displacement of the last two helices of domain 4 in the asymmetric protomer, of which only the most C-terminal α-helix is traceable. The rearrangement of domain 4 in the LA–LAH asymmetric protomer exposes some of the surface area of the substrate-binding domain (domain 3) of the other protomer (symmetric protomer) of the homodimer, implying functional relevance of this oligomeric state (Figs. 4*b*, 5*b* and 5*c*). The observed displacement generates a cavity forming the entrance to a channel (Fig. 5*b*) which runs from the surface of the protein down to a cleft formed at the interface of domains 1–3. This cleft is known to be the FAD- and substrate-binding site in structurally related proteins (Fig. 3). In the difference Fourier electron-density map ($mF_o - DF_c$ at the 3σ level) an elongated electron density which can accommodate an aliphatic chain comprising seven C atoms has been observed in the symmetric protomer of LA–LAH with exposed surface of domain 3 (Fig. 5*a*). This electron density starts from the channel entrance and continues down towards the interior of the protomer and has been interpreted as partially disordered molecule of linoleic acid, a seven-C-atom hydrophobic chain of which is bound inside the channel while the rest of the molecule is disordered inside the cavity. The accumulation of

positively charged residues (three arginines and eight lysines) on the surface of domain 4 (Fig. 5*b*) in concert with its conformational flexibility may facilitate the recruitment of a fatty acid molecule to LAH by the formation of a salt bridge to its carboxyl group (Fig. 5*c*). Analysis of the substrate-binding sites of PAI, POBS and LAAO revealed that their substrates bind in a cleft positioned close to the isoalloxazine moiety of FAD which, similar to as is observed for LAH, is accessible through the channel formed *inter alia* by domain 3 (Fig. 3). Therefore, it is tempting to speculate that upon recognition of the carboxyl group of a fatty acid molecule the C-terminal helices constituting domain 4 act as a lid which can be displaced to further open up a channel through which the substrate can reach its binding site at the interface between domains 1–3. In addition, the hydrophobic side chains that line the interior of this channel argue in favour of the fatty acid-binding properties of this channel. In LAH, this cleft is filled with two MPD molecules (Fig. 3) which, together with the fact that LAH was crystallized at 40% MPD, could explain why all attempts to obtain a complex of LAH with substrate bound deeper in its active site failed.

3.5. FAD-binding pocket of LAH

The crystallized LAH lacks FAD, as it easily dissociated from the protein during purification (see above). Despite the low sequence identity between the Rossmann folds of structurally related proteins identified by a *DALI* search, a conserved FAD-binding sequence motif $GXGXXGX_{18–21}EX_5GGX_{14–16}G$ of the GR_1 family of FAD-binding proteins is identifiable as in the hydratase from *Streptococcus pyogenes* (Volkov *et al.*, 2010). The first two glycine residues are located in the P1 loop, which connects the first β-strand and α-helix of the Rossmann fold and guides the sugar moiety of the FAD into its binding pocket (Fig. 3, yellow cavity). The glutamate side chain resides in the loop right after the second β-strand and further positions the sugar moiety by hydrogen bonding to its hydroxyl groups. The last glycine residue is positioned in a loop which creates the part of the pocket in which the isoalloxazine ring resides. The loop covering the binding cleft for the sugar and phosphate moieties of the FAD in PAI, POBS and LAAO is missing in both LAH crystal structures as it is not defined in the electron-density maps. Two conserved glycine residues (the fourth and fifth glycines in the motif) are located in this loop in LAH. Hence, it is tempting to speculate that in the presence of FAD this loop adopts the same turn as observed in the structures of PAI, POBS and LAAO, and that it becomes disordered upon loss of FAD. The fact that the length of the traceable part of this loop is different in both protomers would argue against the possibility of this loop being removed by proteases. The positioning of the seven

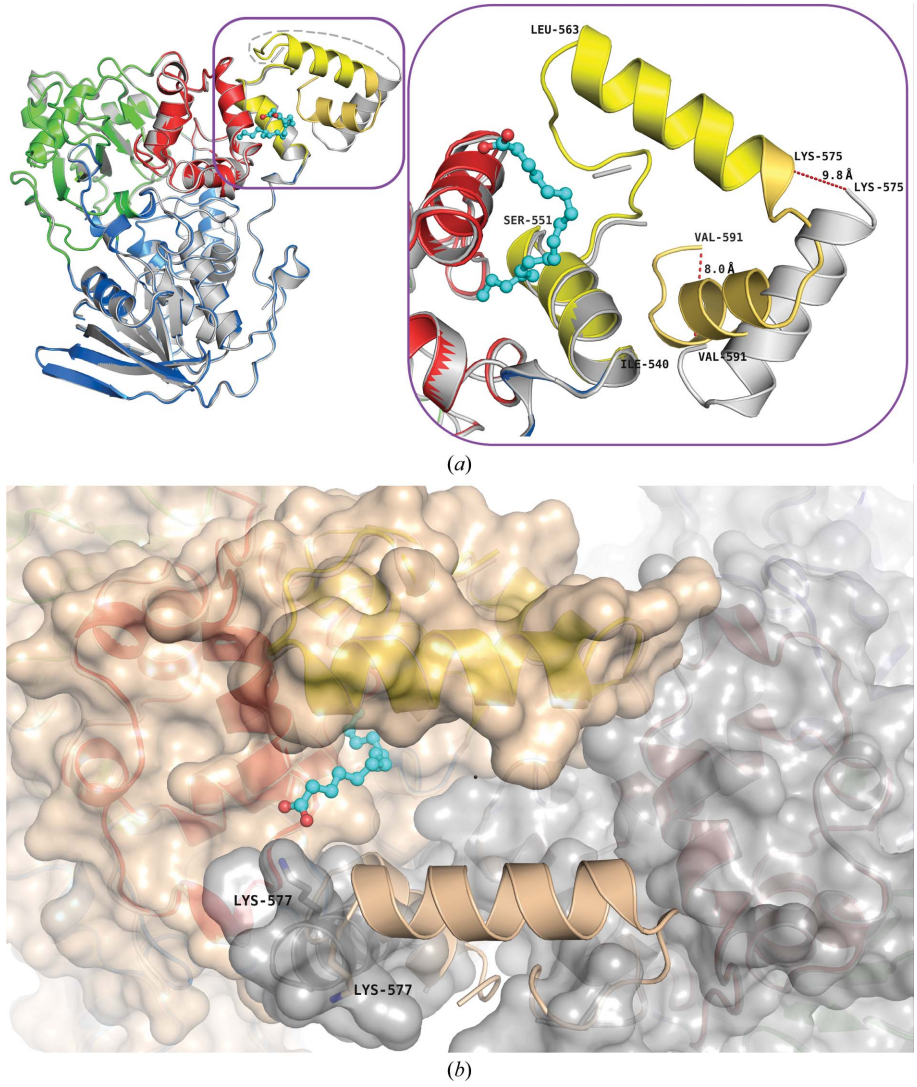


Figure 4
Different conformations of domain 4 observed in LAH monomers. (*a*) A superposition of two LAH protomers differing in the conformation of domain 4. Domains 1, 2, 3 and 4 of the apo LAH protomer (symmetric protomer) are coloured marine, light green, red and yellow, respectively. The LA–LAH protomer (asymmetric protomer) exhibiting a different conformation of domain 4 is coloured grey. The missing residues of domain 4 (556–574) are marked as a grey dashed line and the remaining traceable fragment of domain 4 (elongated C-terminal α-helix, residues 575–591) is shown as a cartoon representation. Red dashed lines indicate distances (in Å) between the positions of equivalent residues (575 and 591) in both conformations. (*b*) Superposition of apo LAH and LA–LAH dimers viewed along the dimerization axis. Domains 1, 2, 3 and 4 of apo LAH are coloured marine, light green, red and yellow/wheat, respectively. The protomomers of LA–LAH are coloured grey and dark grey for the symmetric and the asymmetric protomer, respectively. Residue Lys575 is labelled in both the symmetric protomer (apo LAH, loop region) and the asymmetric protomer (LA–LAH, α-helix). The linoleic acid molecule is shown in ball-and-stick representation.

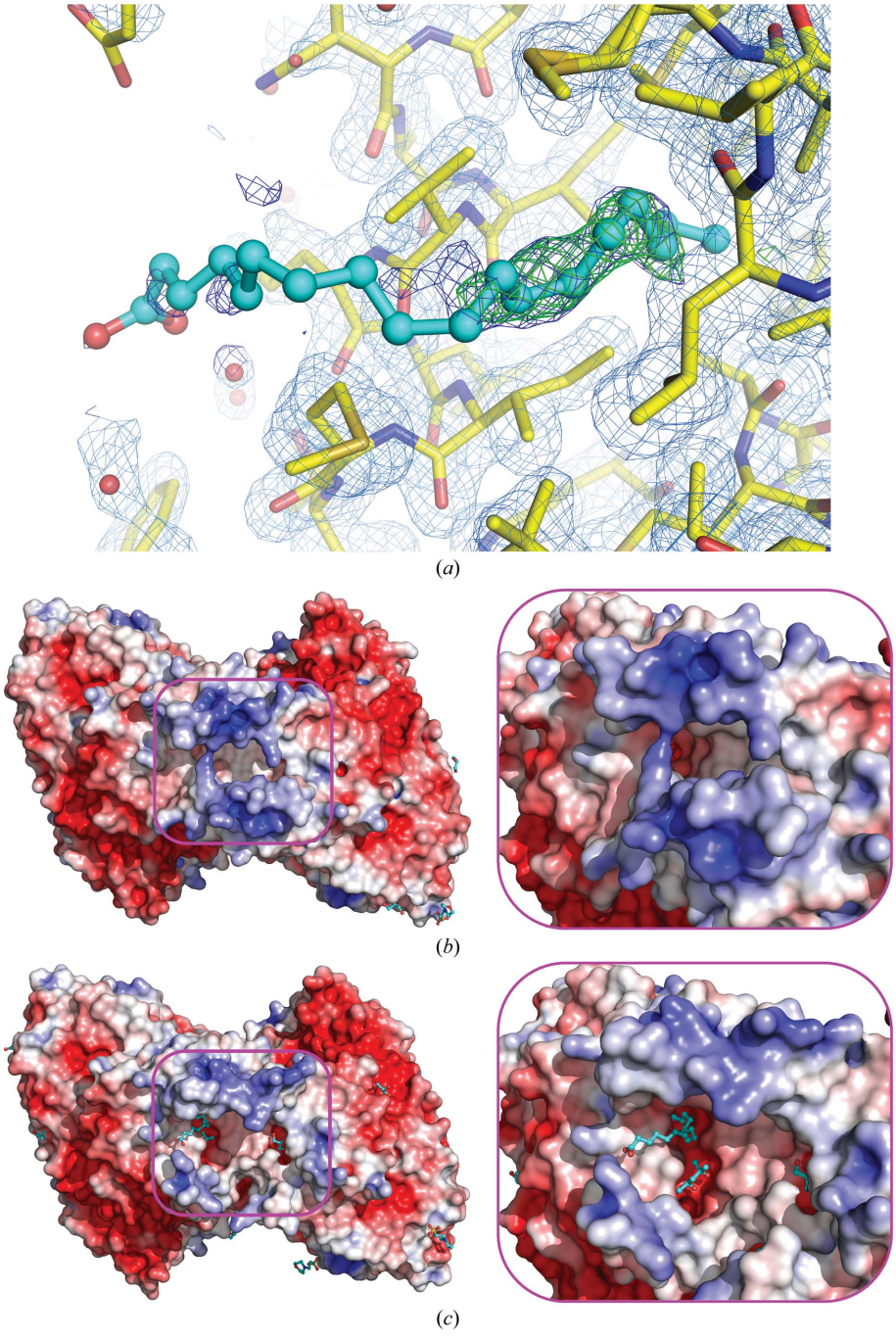


Figure 5
Linoleic acid bound at the entrance to the LAH substrate channel. (*a*) The electron density observed at the substrate-entrance channel. The simulated-annealing (SA) OMIT $mF_o - DF_c$ map at 3σ is coloured green, the SA OMIT $2mF_o - DF_c$ map at 1σ is coloured dark blue and the difference $2mF_o - DF_c$ map at 1σ is coloured light blue and mostly covers hydrophobic residues that form the entrance to the channel. The modeled LA molecule is shown in ball-and-stick representation and is coloured cyan. The difference electron density in the substrate channel has been interpreted as a partially disordered molecule of LA, the seven-C-atom hydrophobic chain of which is bound inside the channel while the rest of the molecule is disordered inside the cavity and shows no difference electron density. (*b*) The electrostatic surface potential of the apo LAH dimer coloured from red ($-5kT$/e) to blue ($5kT$/e) shows distinct basic patches on the surface of domain 4 (viewed along the dimerization axis). Organic molecules are shown as cyan sticks with O atoms in red. On the right, a detailed view of the substrate-channel entrance is shown that is closed by domain 4. (*c*) Electrostatic surface potential of the LA–LAH dimer viewed and coloured as in (*b*). On the right, a closer view of the opened entrance to the substrate channel is shown.

residues comprising the G*X*G*XX*G*X* motif is quite similar in the LAH, PAI, POBS and LAAO structures, suggesting that LAH would exploit a similar FAD-binding mode as observed in the other three enzymes. Analysis of the electrostatic potential distribution around the putative FAD-binding cleft of LAH reveals positively charged patches that favour FAD binding, further supporting this hypothesis. It is noteworthy that in the structure of LAH the positions of the isoalloxazine ring and the phosphate moiety of the FAD molecule are occupied by two glycerol molecules.

4. Discussion

The biological significance of bacterial double-bond hydratases has recently been demonstrated for the hydratase from *S. pyogenes* M49, which was shown to be involved in oleic acid detoxification and bacterial virulence (Volkov *et al.*, 2010). However, to date the underlying structural and mechanistic enzymatic details have only been characterized for the structurally related fatty acid double-bond isomerase PAI (Liavonchanka *et al.*, 2006, 2009). Our crystal structure analysis of LAH allowed us to assign it as a putative member of the flavin-containing amine oxidoreductase family with a well conserved Rossmann fold (Fig. 3). We have shown that FAD was bound to LAH (Fig. 1*a*) but was easily lost during purification, which is reminiscent of the hydratase from *S. pyogenes* M49 (Volkov *et al.*, 2010). Both enzymes have a high degree of overall sequence identity and similarity (68 and 82%, respectively).

A *DALI* search revealed structural similarity of LAH to several FAD-binding proteins including, amongst others, LAAO from *C. rhodostoma* and PAI from *P. acnes*, with *Z*-scores above 17 and r.m.s.d. values of 3.3 and 3.8 Å, respectively (Fig. 3). The highest degree of homology was found for the FAD-binding domains of these enzymes, which are represented by a variation of the Rossmann fold. Sequence analysis of the FAD-binding domains revealed a conserved dinucleotide-binding motif G*X*G*XX*G$(X)_{17/23}$E surrounded by

hydrophobic patches. Significant divergence in the sequence of the motif was observed within the lactobacilli subfamily of MCRAs; thus, sequences from all three clusters showed differences either in the third glycine position or in the length of the linker region between the third glycine and the glutamate or aspartate residues [lactobacilli 1, G*X*G*XX*G/A/S$(X)_{21}$E; lactobacilli 2, G*X*G*XX*N$(X)_{23}$D; flavin-containing amine oxidoreductase 3, G*X*G*XX*G$(X)_{17/18}$E; Volkov *et al.*, 2010]. Generally, families of FAD-containing enzymes do not show such a high degree of divergence inside the FAD-binding motif or surrounding sequences as that observed across the MCRA superfamily. It is known that variations in the linker region or the exchange of glutamate to aspartate can occur in a number of FAD-containing enzymes. However, substitutions at the third glycine position are rather rare and may lead to problems with FAD binding, but are found in the glutathione reductase family (Dym & Eisenberg, 2001). These observations are supported by the recent analysis of an oleate hydratase from *Macrococcus caseolyticus* (Joo, Jeong *et al.*, 2012). Here, the authors showed that FAD is essential for hydratase activity and that the first two glycines and the aspartate residue are essential for FAD binding.

Comparison of the LAH structure with those of structurally related FAD-binding enzymes revealed the presence of an additional distinct domain (domain 4) comprising 52 residues which is located at the C-terminus of LAH and is folded into three α-helices connected by two hinge loops (Fig. 3, yellow). In LAH, domain 4 is positioned close to domain 3, which is known to form the substrate-binding channel in concert with domain 2 in the structures of homologues. These channels lead from the protein surface to the active site located in the interior of the monomer close to the isoalloxazine ring of FAD (Fig. 3, white channels). Among the identified structures of homologues, PAI shares a similar biological function with LAH as it may participate in the detoxification of fatty acids of similar length (C_{18}). PAI has been crystallized in apo and product-bound forms and revealed a hydrophobic C-shaped channel of ~30 Å in length which is controlled by a unique substrate-gating mechanism involving side-chain switches. In the PAI structure, positively charged patches generated by several Lys and Arg residues could be localized near the channel entrance in domain 3, which might serve as an initial recognition site for the fatty acid carboxylate group. In contrast, the entrance to the substrate channel in LAH is covered by the LAH-specific domain 4, which also provides lysines and arginines to form the positively charged patches on the surface (Figs. 5*b* and 5*c*). The observed conformational flexibility and refolding ability of domain 4 (Fig. 4*a*) suggests its possible role as a plastic lid that blocks and opens the access to the entrance of the hydrophobic substrate-binding channel. Conformational changes are most likely to be induced by recognition of the negatively charged substrate carboxylate group, which can be sensed by numerous positively charged residues on the surface of domain 4. Such a mode of action closely resembles the role of a flexible amphipathic helical lid domain covering the active site of lipases which moves away upon contact of the lipase with its substrate, thereby exposing a hydrophobic binding pocket at the surface of the protein (Jaeger *et al.*, 1994). Interestingly, a lipase from the archeon *A. fulgidus* (AFL), in addition to harbouring the active-site α/β hydratase domain, carries a unique lipid-binding C-terminal β-barrel domain (Chen *et al.*, 2009). The association of these two domains in AFL covers most of the hydrophobic substrate-binding site except for the catalytically active residues. This generates a hydrophobic tunnel of ~20 Å in length that is able to accommodate about 18 hydrocarbon units with the catalytic triad lying at the entrance of the hydrophobic tunnel, thus generating higher specificity towards long-chain substrates. The overall domain organization of AFL (N-terminal catalytic domain, C-terminal lipid-binding β-barrel domain and N-terminal lid) resembles the association of domains 3, 2 and 4 in LAH, respectively (Fig. 6). Hence, the formation of hydrophobic substrate channels seems to be a general mechanism that has evolved in order to modulate the substrate specificity of enzymes involved in the biosynthesis of lipids and the products of their degradation (Böttcher & Bornscheuer, 2010).

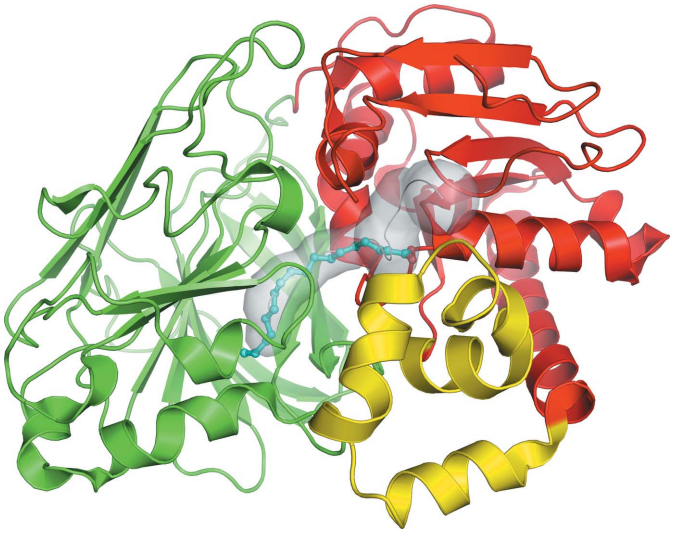


Figure 6
The overall domain organization of AFL. The N-terminal catalytic domain, C-terminal lipid-binding β-barrel domain and N-terminal lid are coloured red, green and yellow, respectively. The bound LA molecule is shown in ball-and-stick representation. The channel leading to the active site was calculated using the *MOLE* software and is depicted as a transparent white surface.

5. Conclusion

The LAH structures share structural similarity with several flavin-dependent enzymes, including amine oxidoreductase from *C. rhodostoma* and PAI. The observed similarity allowed the assignment of four intricately connected domains in LAH, of which three have structural equivalents in flavin-dependent enzymes and form a hydrophobic substrate channel running from the protein surface to the active site located close to the isoalloxazine ring of FAD. The additional domain located at the C-terminus of LAH acts as a lid that covers the substrate-channel entrance. Interestingly, movement of the lid domain

of one protomer opens the entrance of the substrate-binding channel of the other protomer forming the LAH homodimer. Positively charged residues located on the surface of domain 4 might facilitate initial recognition of the fatty acid carboxylate group, upon which conformational changes that lead to opening of the substrate-channel entrance occur. The LA–LAH complex structure allowed localization of the substrate-channel entrance and unveiled the flexibility and refolding abilities of the lid domain. Numerous attempts to crystallize LAH with FAD, free fatty acids and a FAD–fatty acid complex were not successful and left the question of the localization of the active site open.

This work was supported by the PhD program 'Molecular Biology and Molecular Sciences and Biotechnology of Crop Plants'. AV and SK are recipients of Georg-Christoph-Lichtenberg stipends of the PhD program 'Molecular Biology' and the Göttingen Graduate School of Neurosciences and Molecular Biology (GGNB). We thank Helmholtz-Zentrum Berlin (HZB, Germany) for the allocation of synchrotron-radiation beamtime and for financial support.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Bevers, L. E., Pinkse, M. W., Verhaert, P. D. & Hagen, W. R. (2009). *J. Bacteriol.* **191**, 5010–5012.
Böttcher, D. & Bornscheuer, U. T. (2010). *Curr. Opin. Microbiol.* **13**, 274–282.
Brünger, A. T. (1993). *Acta Cryst.* D**49**, 24–36.
Chen, C. K.-M., Lee, G.-C., Ko, T.-P., Guo, R.-T., Huang, L.-M., Liu, H.-J., Ho, Y.-F., Shaw, J.-F. & Wang, A. H.-J. (2009). *J. Mol. Biol.* **390**, 672–685.
Cowtan, K. (1994). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **31**, 34–38.
Dym, O. & Eisenberg, D. (2001). *Protein Sci.* **10**, 1712–1728.
Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. (2010). *Acta Cryst.* D**66**, 486–501.
Greenway, D. L. & Dyke, K. G. (1979). *J. Gen. Microbiol.* **115**, 233–245.
Holm, L. & Rosenström, P. (2010). *Nucleic Acids Res.* **38**, W545–W549.
Jaeger, K. E., Ransac, S., Dijkstra, B. W., Colson, C., van Heuvel, M. & Misset, O. (1994). *FEMS Microbiol. Rev.* **15**, 29–63.
Joo, Y.-C., Jeong, K.-W., Yeom, S.-J., Kim, Y.-S., Kim, Y. & Oh, D.-K. (2012). *Biochimie*, **94**, 907–915.
Joo, Y.-C., Seo, E.-S., Kim, Y.-S., Kim, K.-R., Park, J.-B. & Oh, D.-K. (2012). *J. Biotechnol.* **158**, 17–23.
Kabsch, W. (2010*a*). *Acta Cryst.* D**66**, 125–132.
Kabsch, W. (2010*b*). *Acta Cryst.* D**66**, 133–144.
Kachalova, G. S., Bourenkov, G. P., Mengesdorf, T., Schenk, S., Maun, H. R., Burghammer, M., Riekel, C., Decker, K. & Bartunik, H. D. (2010). *J. Mol. Biol.* **396**, 785–799.
Karplus, P. A. & Diederichs, K. (2012). *Science*, **336**, 1030–1033.
Kellens, M. J., Goderis, H. L. & Tobback, P. P. (1986). *Biotechnol. Bioeng.* **28**, 1268–1276.
Kim, B.-N., Joo, Y.-C., Kim, Y.-S., Kim, K.-R. & Oh, D.-K. (2012). *Appl. Microbiol. Biotechnol.* **95**, 929–937.
Kishino, S., Park, S.-B., Takeuchi, M., Yokozeki, K., Shimizu, S. & Ogawa, J. (2011). *Biochem. Biophys. Res. Commun.* **416**, 188–193.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Liavonchanka, A., Hornung, E., Feussner, I. & Rudolph, M. G. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 2576–2581.
Liavonchanka, A., Rudolph, M. G., Tittmann, K., Hamberg, M. & Feussner, I. (2009). *J. Biol. Chem.* **284**, 8005–8012.
Maia, M. R., Chaudhary, L. C., Figueres, L. & Wallace, R. J. (2007). *Antonie Van Leeuwenhoek*, **91**, 303–314.
Morris, R. J., Zwart, P. H., Cohen, S., Fernandez, F. J., Kakaris, M., Kirillova, O., Vonrhein, C., Perrakis, A. & Lamzin, V. S. (2004). *J. Synchrotron Rad.* **11**, 56–59.
Mosley, E. E., Powell, G. L., Riley, M. B. & Jenkins, T. C. (2002). *J. Lipid Res.* **43**, 290–296.
Nam, I. S. & Garnsworthy, P. C. (2007). *J. Appl. Microbiol.* **103**, 551–556.
Panjikar, S., Parthasarathy, V., Lamzin, V. S., Weiss, M. S. & Tucker, P. A. (2005). *Acta Cryst.* D**61**, 449–457.
Pawelek, P. D., Cheah, J., Coulombe, R., Macheroux, P., Ghisla, S. & Vrielink, A. (2000). *EMBO J.* **19**, 4204–4215.
Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.
Petrek, M., Kosinová, P., Koca, J. & Otyepka, M. (2007). *Structure*, **15**, 1357–1363.
Qin, X., Sun, L., Wen, X., Yang, X., Tan, Y., Jin, H., Cao, Q., Zhou, W., Xi, Z. & Shen, Y. (2010). *J. Struct. Biol.* **170**, 76–82.
Raychowdhury, M. K., Goswami, R. & Chakrabarti, P. (1985). *J. Appl. Bacteriol.* **59**, 183–188.
Reuter, K. & Ficner, R. (1999). *Acta Cryst.* D**55**, 888–890.
Rosberg-Cody, E., Liavonchanka, A., Göbel, C., Ross, R. P., O'Sullivan, O., Fitzgerald, G. F., Feussner, I. & Stanton, C. (2011). *BMC Biochem.* **12**, 9.
Sheldrick, G. M. (2008). *Acta Cryst.* A**64**, 112–122.
Terwilliger, T. C. (2000). *Acta Cryst.* D**56**, 965–972.
Volkov, A., Liavonchanka, A., Kamneva, O., Fiedler, T., Goebel, C., Kreikemeyer, B. & Feussner, I. (2010). *J. Biol. Chem.* **285**, 10353–10361.
Vossenberg, J. L. van de & Joblin, K. N. (2003). *Lett. Appl. Microbiol.* **37**, 424–428.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yang, B., Chen, H., Song, Y., Chen, Y., Zhang, H. & Chen, W. (2013). *Biotechnol. Lett.* **35**, 75–81.
Zheng, C. J., Yoo, J.-S., Lee, T.-G., Cho, H.-Y., Kim, Y.-H. & Kim, W.-G. (2005). *FEBS Lett.* **579**, 5157–5162.