University of Groningen

Crystal Structure and Biochemical Properties of a Novel Thermostable Esterase Containing an Immunoglobulin-Like Domain

Levisson, Mark; Sun, Lei; Hendriks, Sjon; Swinkels, Peter; Akveld, Twan; Bultema, Jelle B.; Barendregt, Arjan; van den Heuvel, Robert H. H.; Dijkstra, Bauke W.; van der Oost, John

[Link to publication in University of Groningen/UMCG research database](#)

JMB

Crystal Structure and Biochemical Properties of a Novel Thermostable Esterase Containing an Immunoglobulin-Like Domain

Mark Levisson[1]*†, **Lei Sun**[1,2]†, **Sjon Hendriks**[1], **Peter Swinkels**[1], **Twan Akveld**[1], **Jelle B. Bultema**[2], **Arjan Barendregt**[3], **Robert H. H. van den Heuvel**[3], **Bauke W. Dijkstra**[2], **John van der Oost**[1] and **Servé W. M. Kengen**[1]

[1]*Laboratory of Microbiology, Department of Agrotechnology and Food Sciences, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, The Netherlands*

[2]*Laboratory of Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands*

[3]*Biomolecular Mass Spectrometry and Proteomics Group, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Sorbonnelaan 16, 3584 CA Utrecht, The Netherlands*

Edited by M. Guss

Comparative analysis of the genome of the hyperthermophilic bacterium *Thermotoga maritima* revealed a hypothetical protein (EstA) with typical esterase features. The EstA protein was functionally produced in *Escherichia coli* and purified to homogeneity. It indeed displayed esterase activity with optima at or above 95 °C and at pH 8.5, with a preference for esters with short acyl chains (C2–C10). Its 2.6-Å-resolution crystal structure revealed a classical α/β hydrolase domain with a catalytic triad consisting of a serine, an aspartate, and a histidine. EstA is irreversibly inhibited by the organophosphate paraoxon. A 3.0-Å-resolution structure confirmed that this inhibitor binds covalently to the catalytic serine residue of EstA. Remarkably, the structure also revealed the presence of an N-terminal immunoglobulin (Ig)-like domain, which is unprecedented among esterases. EstA forms a hexamer both in the crystal and in solution. Electron microscopy showed that the hexamer in solution is identical with the hexamer in the crystal, which is formed by two trimers, with the N-terminal domains facing each other. Mutational studies confirmed that residues Phe89, Phe112, Phe116, Phe246, and Trp377 affect enzyme activity. A truncated mutant of EstA, in which the Ig-like domain was removed, showed only 5% of wild-type activity, had lower thermostability, and failed to form hexamers. These data suggest that the Ig-like domain plays an important role in the enzyme multimerization and activity of EstA.

© 2008 Elsevier Ltd. All rights reserved.

---

*Corresponding author. E-mail address: Mark.Levisson@wur.nl.

† M.L. and L.S. contributed equally to this work.

Present address: R.H.H. van den Heuvel, Biotech Analytical Development, Schering-Plough Corporation, Postbus 20, 5340 BH Oss, The Netherlands.

Abbreviations used: MAD, multiwavelength anomalous dispersion; PDB, Protein Data Bank; DEP, diethyl phosphate; NCS, noncrystallographic symmetry; DLS, dynamic light scattering; EM, electron microscopy; 7-ACA, 7-aminocephalosporanic acid; SeMet, selenomethionine; CAPS, 3-(cyclohexylamino) 1-propanesulphonic acid.

Introduction

Esterases constitute a large family of proteins with representatives in all domains of life. They catalyze the hydrolysis of ester bonds, resulting in the formation of an alcohol and a carboxylic acid. Most esterases belong to the α/β hydrolase family, which also contains lipases, and have a conserved catalytic triad that is usually composed of a serine, an aspartate, and a histidine.[1] Esterases differ, however, from lipases in that they show a preference for short-chain acyl esters (shorter than 10 carbon atoms), and that they are not active on micellar substrates.[1] The physiological function of these enzymes is often not clear. As biocatalysts, however, they are widely used in industrial processes because of their cofactor-independent activity, high regiospecificity and stereospecificity, and stability and activity in organic solvents.[2] In particular, esterases from thermophilic origin are potentially interesting for industrial applications, since most of them are stable in organic solvents and can withstand elevated temperatures.[3]

The hyperthermophilic anaerobic bacterium *Thermotoga maritima* grows optimally at temperatures between 55 °C and 90 °C. It can metabolize many simple and complex carbohydrates, including glucose, cellulose, and starch.[4] A bioinformatics analysis of the genome of *T. maritima*[5] revealed several open reading frames potentially encoding new thermostable esterases, including one (*est*A; TM0033) that was annotated as a hypothetical protein.

Multiple-sequence alignment suggested that EstA is composed of two domains:[6] a C-terminal domain with a predicted α/β hydrolase fold, and an N-terminal domain that has no homology to known proteins and therefore could not be assigned a function. A three-dimensional structure of EstA may provide better insight into the structure and function of this new N-terminal domain, as well as reveal the molecular basis of substrate recognition and catalysis by this enzyme. Therefore, EstA was expressed and crystallized, and its three-dimensional structure was determined. In addition, various biochemical properties and the quaternary structure in solution were determined.

Results

Identification and production

BLAST searches with sequences of known esterases from other hyperthermophilic microorganisms against the genome of *T. maritima* revealed several open reading frames potentially encoding new thermostable esterases, including one (*est*A; TM0033) that has been annotated as a hypothetical protein. The gene encodes a 395-amino-acid protein and has a calculated molecular mass of 44.5 kDa. A BLAST search of EstA at NCBI‡[7] revealed the

highest similarity to other hypothetical proteins and a few predicted peptidases and carboxylesterases. N-terminal sequence analysis using the SignalP 3.0 Server§ revealed that the first 16 amino acids form a signal peptide. The predicted mature gene was cloned into the expression vector pET-24d,[6] and the EstA enzyme was purified to homogeneity from heat-treated cell extracts of *Escherichia coli* BL21(DE3)/pSJS1244/pWUR350 by immobilized nickel affinity chromatography. Homogeneity of the protein was checked by SDS-PAGE, and a molecular subunit mass of 43 kDa (mature enzyme) was confirmed. Activity staining of washed SDS-PAGE gels with α-naphthyl acetate was used to confirm the identity of the EstA band (data not shown).

Overall structure

The native EstA crystal structure was solved in space group *H*32 with two molecules per asymmetric unit. The structure was determined by multiwavelength anomalous dispersion (MAD) phasing with selenium as anomalous scatterer and refined at 2.6 Å resolution. The molecule has an elongated shape with approximate dimensions of 40 Å × 35 Å × 75 Å. The EstA structural model includes residues 20–395, with the exception of residues D248-N251, which form a flexible loop. The structure can be subdivided into two domains: a catalytic domain (residues 158–395) and an immunoglobulin (Ig)-like domain (residues 20–157) (Fig. 1a).

The catalytic domain has the canonical architecture of an α/β hydrolase fold protein consisting of a central β-sheet of eight strands surrounded by helices (Fig. 1b), with strand β2 anti-parallel with the other strands. The β1 and β8 strands are approximately perpendicular to each other because of the counterclockwise twist of the β-sheet. Helices αA, αFa, and αFb, and the short 3_{10}-helices η4, η5 and η7 are on one side of the central β-sheet, and helices αB, αC, and αE, and the 3_{10}-helix η6 are on the other side. There is no lid structure present in the catalytic domain, as seen for some lipases.

The Ig-like domain is composed of two β-sheets, consisting of nine strands (distribution: IABFE/DCGH), sandwiched face to face (Fig. 1b). Ig-like domains are common in eukaryotic proteins; however, in bacteria, they have been found only in bacterial surface proteins, some glycoside hydrolases, and a few secreted peptidases (pfam family CL0159). EstA is the first example of an esterase consisting of both an α/β hydrolase domain and an Ig-like domain.

A structural similarity search carried out using the DALI program[9] shows that EstA has structural similarity to some esterases, lipases, and peptidases, with Z-scores between 20.8 and 15 and RMSDs below 3.0. These scores are mainly based on structural similarity to the catalytic domain. The catalytic
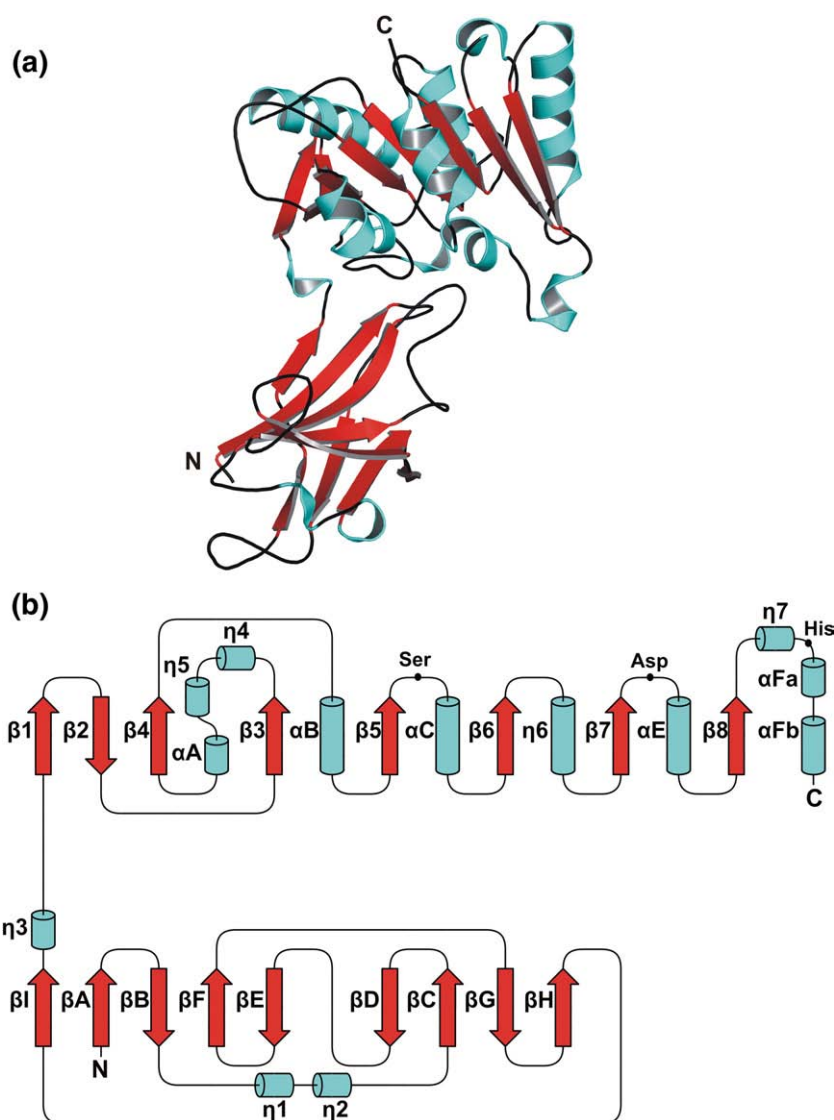
Fig. 1. (a) Overall fold of the esterase EstA. The N-terminal and C-terminal ends are indicated. The figure was generated using PyMOL.[8] (b) Topology diagram for EstA, with the helices displayed as cyan cylinders and the strands displayed as red arrows. The positions of the catalytic residues are indicated.

domain is structurally most similar to the catalytic domain of prolyl tripeptidyl aminopeptidase from *Porphyromonas gingivalis* [Protein Data Bank (PDB) 2D5L],[10] the catalytic domain of human dipeptidyl peptidase (PDB 1N1M),[11] and the feruloyl esterase domain of xylanase Z from *Clostridium thermocellum* (PDB 1JJF).[12] The structures are well superimposed, with RMSDs of 2.6 Å (for 204 aligned C^{α} atoms), 2.8 Å (200 aligned C^{α} atoms), and 2.6 Å (202 aligned C^{α} atoms), respectively.

A DALI search for structures similar to the Ig-like domain gave Z-scores of 6.4 for the Ig-like domain of Rab geranylgeranyltransferase (PDB 1DCE),[13] 5.4 for the Ig-like D1 domain of interleukin-4 receptor (PDB 1IAR),[14] and 5.2 for the β-galactosidase Ig-like domain 4 (PDB 1BGL).[15] None of the structures found with the DALI search belongs to an esterase. The DALI results confirm that the combination of an esterase domain and an Ig-like domain is unprecedented.

Inhibitor assays and structure of EstA complexed with paraoxon

Inhibition assays demonstrated that both phenyl-methylsulfonyl fluoride (PMSF) and diethyl 4-nitrophenyl phosphate (paraoxon) inhibit EstA activity, with 7% and 0% residual activity, respectively. For the determination of kinetic inhibition data, we followed the method described by Forsberg and Puu, which states that inhibition proceeds by the formation of a reversible Michaelis complex, followed by an irreversible step.[16] Inhibition can therefore be characterized by two parameters: a dissociation constant and a bonding rate constant. The inhibition kinetics for paraoxon were investigated in the presence of *p*-nitrophenyl octanoate and resulted in dissociation and rate constants of 33 μM and 0.14 s^{-1}, respectively. Compared to EST2 of *Alicyclobacillus acidocaldarius*, the dissociation constant is slightly higher, but the bonding rate constant

is comparable.[17] Inhibition kinetics for PMSF were not measurable in the presence of substrate. This is possibly a result of a high dissociation rate constant because inhibition was observed when EstA was preincubated with PMSF in the absence of substrate. Other chemical agents such as diethyl pyrocarbonate, dithiothreitol, divalent metal ions, and ethylenediaminetetraacetic acid did not influence EstA activity. EstA was cocrystallized with PMSF and paraoxon, but only crystals for the latter were obtained. Electron density maps for the paraoxon cocrystallized crystals displayed clear density for a diethyl phosphate (DEP) moiety covalently bound to the side chain of Ser286. The density revealed that the *p*-nitrophenol-leaving group of paraoxon had been cleaved off during cocrystallization, thereby leaving a tetrahedral product resembling the first transition state formed during ester hydrolysis. The native and paraoxon-bound structures superimpose with an RMSD of 0.4 Å. There are no significant differences between the two structures.

Quaternary structure

There are two molecules (protomers A and B) in the asymmetric unit that are related by noncrystallographic symmetry (NCS). Protomers A and B are essentially similar, with an RMSD of 0.3 over all C^α atoms. The refined model reveals that the two molecules form an interface of 280 Å^2 in each monomer, a value suggesting a low association constant. The interface involves four β-strands (βC, βD, βG, and βH) from the N-terminal Ig-like domain of both molecules. An intermolecular hydrogen bond (AspA56-LysB61) stabilizes the dimer. With the 3-fold crystallographic symmetry of space group *H*32, three dimers form a tightly packed hexamer, burying a total surface area of 3585 Å^2. The interdimer interfaces involve residues from αA, αE, β7, β8, and the loop between β2 and β3, as well as the N-termini of βF, βE, and the loop between them. Salt bridges and hydrogen bonds are formed between neighboring protomers such as ArgA213-GluA′362, ArgA222-GluA′360, GluA99-LysA′363, ArgA187-GluA′351, GlnA384-TyrA′358, and AspA182-LysA′347. On average, a surface area of 1454 Å^2 per monomer is buried upon hexamerization. Multiple-sequence alignment shows that the residues involved in the dimer and trimer interfaces are barely conserved, suggesting a novel mode of hexamer formation.

Native PAGE and size exclusion chromatography showed that multiple quaternary structures of EstA are present in solution. Therefore, native mass spectrometry—in which EstA was measured under nondenaturing conditions in ammonium acetate (pH 6.8)[18]—and dynamic light scattering (DLS) were performed and revealed that the purified EstA protein was present predominantly (>50%) as a hexamer in solution. However, to a minor extent, monomeric, dimeric, trimeric, and higher multimeric forms of EstA were also detected using mass spectrometry.

Electron microscopy (EM) and single-particle analysis were used to analyze the oligomeric structure of EstA in solution. Negatively stained specimens of purified EstA, with a final concentration of 30 μg ml^{-1}, were found suitable for single-particle image analysis. The analysis confirmed that EstA forms hexamers in solution (Fig. 2). Comparison of the EM projection maps with projection maps generated from the crystal structure shows great similarities. The generated top-view projection map shows three groups of densities that are recognizable in the EM projection map despite the limited details. Two EM side views are presented in Fig. 2: a side view (Fig. 2b) and a tilted side view (Fig. 2c). The projection in Fig. 2b is referred as side view because 2-fold symmetry can be imposed without altering the main characteristics in the projection map (not shown here). However, if this 2-fold symmetry is imposed on Fig. 2c, the overall shape changes. This indicates that the hexamer shown in Fig. 2c is slightly tilted due to different orientations on the carbon support film and is therefore referred to as tilted side view. It was possible to establish the rotational difference between these two views. Rotation of the proposed hexamer, followed by examination of the generated projection maps, revealed that the tilted side view is about 10° off the side view. The tilted side view provides further evidence of the proposed EstA hexamer. Altogether, EM and single-particle analysis showed that the hexameric structure of EstA in solution is composed of two trimers facing each other with the Ig-like domains, which is identical with the hexamer in the crystal structure.

Active site

The active site of EstA was identified by the location of the nucleophile serine Ser286 within the conserved pentapeptide sequence Gly-X-Ser-X-Gly (GLSMG). A classical catalytic triad—consisting of Ser286 as the nucleophile, His374 as the proton acceptor/donor, and Asp334 as the residue stabilizing the histidine—was found. Ser286 is located at the nucleophile elbow, a sharp turn between β5 and helix αC. In both native and paraoxon-bound structures, Ser286 has the energetically unfavorable main-chain conformation that is also observed in other α/β hydrolases and might provide an energy reservoir for catalysis.[19] The location of several glycine residues (Gly284, Gly288, and Gly289) in very close proximity to the catalytic Ser286 prevents steric hindrance in the sharp turn of the nucleophile elbow. Asp334 and His374 are located in loops between strand β7 and helix αE, and between strand β8 and helix αF, respectively. The oxyanion hole is formed by the backbone NH groups of Ala198 and Met287.

In the native structure, the Ser286 side chain has a hydrogen bond with the side chain of His374. The extra density near the side chain of Ser286 was interpreted as a sulfate ion, forming hydrogen bonds with the side chain of the active-site residue
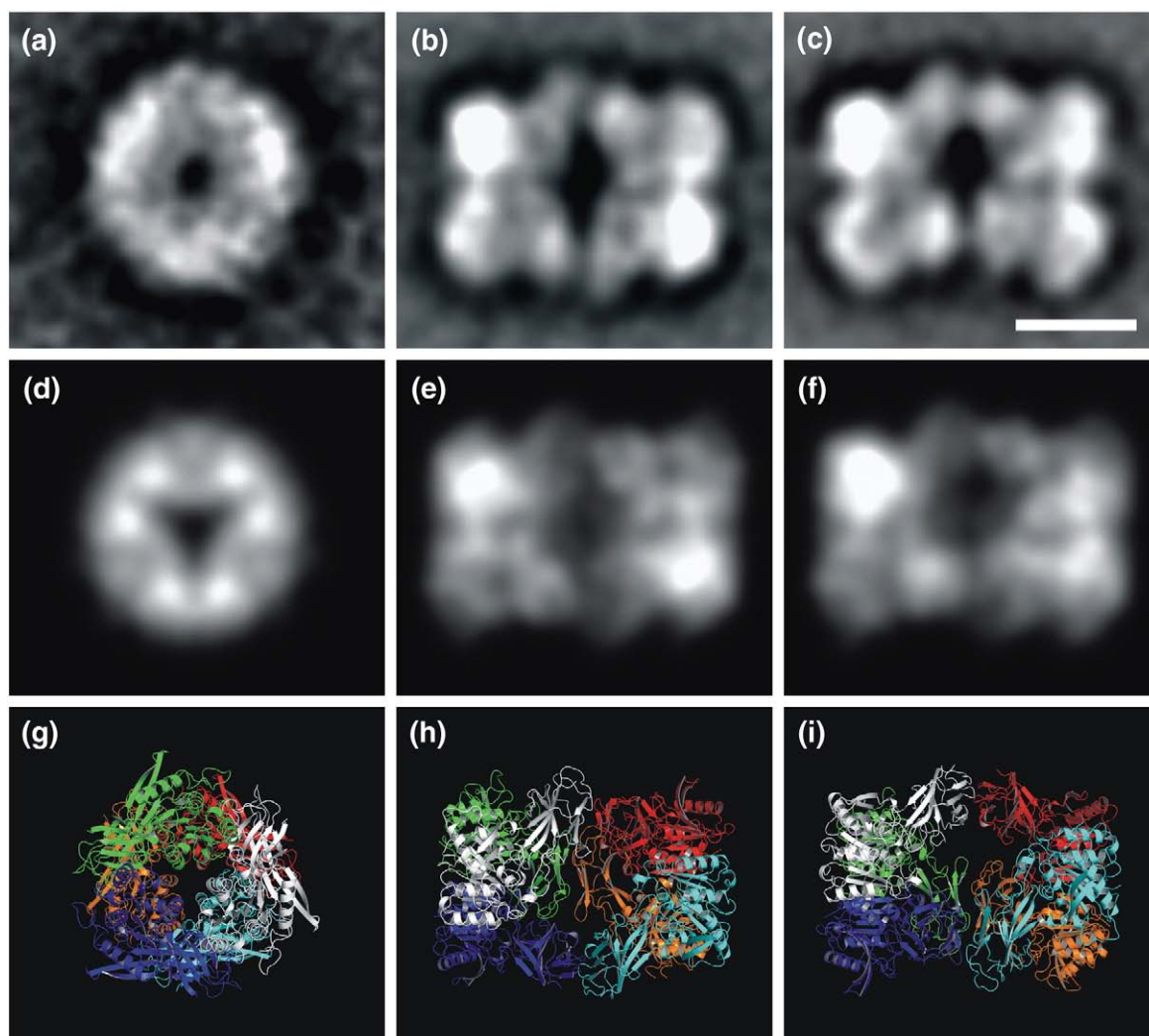
Fig. 2. Comparison of the EM projection maps of the EstA hexamer. Top view (a), side view (b), and 10° off side view (c). Two-dimensional projection maps obtained by statistical analysis and classification (a–c); the comparable two-dimensional projection maps with 15 Å resolution (d–f) generated from the proposed EstA hexameric structure (g–i).[8] The scale bar represents 50 Å.

Ser286 and the main-chain nitrogen atoms of Ala198 and Met287, and mimicking the oxyanion transition state (Fig. 3). Sulfate is present in the crystallization buffer and is commonly found as an adduct in other structures.

In the paraoxon-bound structure, paraoxon is stabilized by the covalent bond with Ser286, hydrogen-bonding interactions with the oxyanion hole, and a hydrogen bond to the side chain of His374 (Fig. 4). One of the two ethyl arms of bound paraoxon points toward the surface of EstA, while the other follows the groove of the acyl-binding pocket.

The catalytic triad and oxyanion hole are located at the end of a surface depression, characteristic of many α/β hydrolases (Fig. 5a). This groove extends by approximately 15 Å from the catalytic serine into the enzyme where the gorge is closed by Glu33 (S1 in Fig. 5b). The volume of the groove is ∼790 Å^3. The active-site serine residue is covered by a valine residue (Val336) and a phenylalanine residue (Phe116; of the Ig-like domain), resulting in a tunnel. The covered gorge extends to the other site of the catalytic serine by approximately 5 Å. The active-site gorge is slightly curved and is formed by the hydrophobic side chains of Ala198, Gly199, Leu285, Trp377, and Phe112. The tunnel is formed by the hydrophobic side chains of Val336, Val337, Phe116, Leu245, Phe246, and Phe89 (of the N-terminal domain of subunit B), and by the nonpolar side chain of Tyr290. A second side gorge with a volume of ∼210 Å^3 also provides access to the active site (S2 in Fig. 5b). The ∼7-Å-wide opening is lined by residues Leu245, Asp248, Arg249, Pro252, Phe253, Tyr290, Val336, Val337, Pro338, and Asn341.

Substrate specificity and kinetics

Kinetic parameters for the hydrolysis of *p*-nitrophenyl esters with varying acyl-chain lengths are given in Table 1. Catalytic activity (K_{cat}) decreases
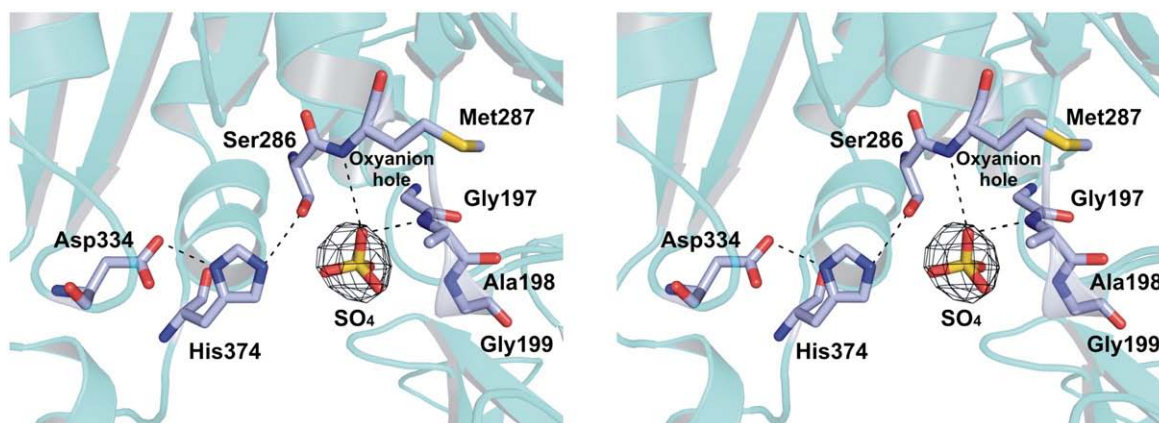
Fig. 3. Stereo view of the EstA catalytic site with the bound SO_4 adduct. The catalytic triad residues are shown as sticks, with the hydrogen bonds shown as dashed black lines. The observed density for the bound SO_4 is contoured. These images were generated using PyMOL.[8]

markedly with an increase in chain length, with only low levels of activity on the long-chain esters. The optimum substrate for an enzyme can be concluded from the catalytic efficiency (K_{cat}/K_m) of its conversion. According to this criterion, the esterase shows two optima: for *p*-nitrophenyl acetate and for *p*-nitrophenyl octanoate. However, one should realize that the natural substrate is not known, and that the *p*-nitrophenyl derivatives may give atypical kinetics.

The effect of temperature on EstA activity was studied using *p*-nitrophenyl valerate and octanoate as substrates. The esterase activity increased from 40 °C to 95 °C (Fig. 6a). An Arrhenius analysis for *p*-nitrophenyl valerate and octanoate resulted in a linear plot in the temperature range 40–80 °C (Fig. 6a), with a calculated activation energy for the formation of the enzyme–substrate complex of 9.5 kJ mol^{-1} and 13.5 kJ mol^{-1} respectively, comparable to the reported value.[20] EstA has a high resistance to thermal inactivation, with a half-life of approximately 1.5 h at 100 °C (data not shown). The optimal pH for the esterase was measured in the pH range

4.8–10.2 using the substrates *p*-nitrophenyl valerate and octanoate. The latter substrate was used due to the instability of *p*-nitrophenyl valerate at higher pH values. EstA displayed maximal activity at approximately pH 8.5 (Fig. 6b).

Hydrolysis of triacylglycerol esters was measured using a plate assay. EstA was able to hydrolyze tributyrin, but not longer acyl-chain esters. To visualize this, the structure of EstA in complex with the tributyrin transition state was manually modeled based on the structure of *Pseudomonas cepacia* lipase (PDB 4LIP) with the covalently bound tributyrin transition state analogue R_C-(R_P,S_P)-1,2-dibutylcarbamoylglycero-3-*O*-*p*-nitrophenyl butyl-phosphonate (data not shown).[21] According to the model, tributyrin could bind in the active site, resembling the way paraoxon binds. However, the size of the binding pocket is limited, thus making it impossible to bind longer acyl-chain triacylglycerols as in *P. cepacia* lipase. This observation is in agreement with the results of the plate assay.

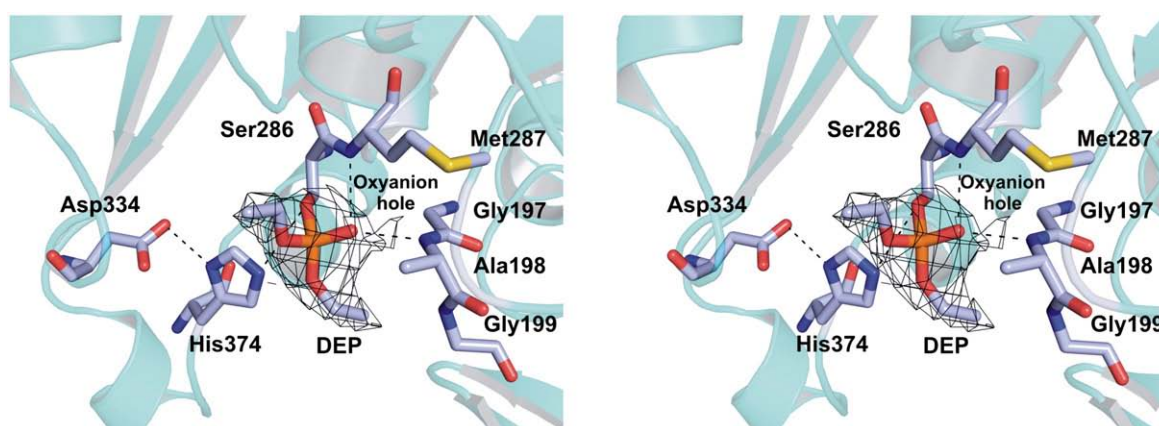EstA was investigated for its ability to remove acetyl groups from the substrates cephalosporin C,

Fig. 4. Stereo view of the EstA catalytic site with the DEP intermediate of the inhibitor paraoxon covalently bound to the catalytic serine. The catalytic triad residues are shown as sticks, with the hydrogen bonds shown as dashed black lines. The observed density for the bound DEP is contoured. These images were generated using PyMOL.[8]
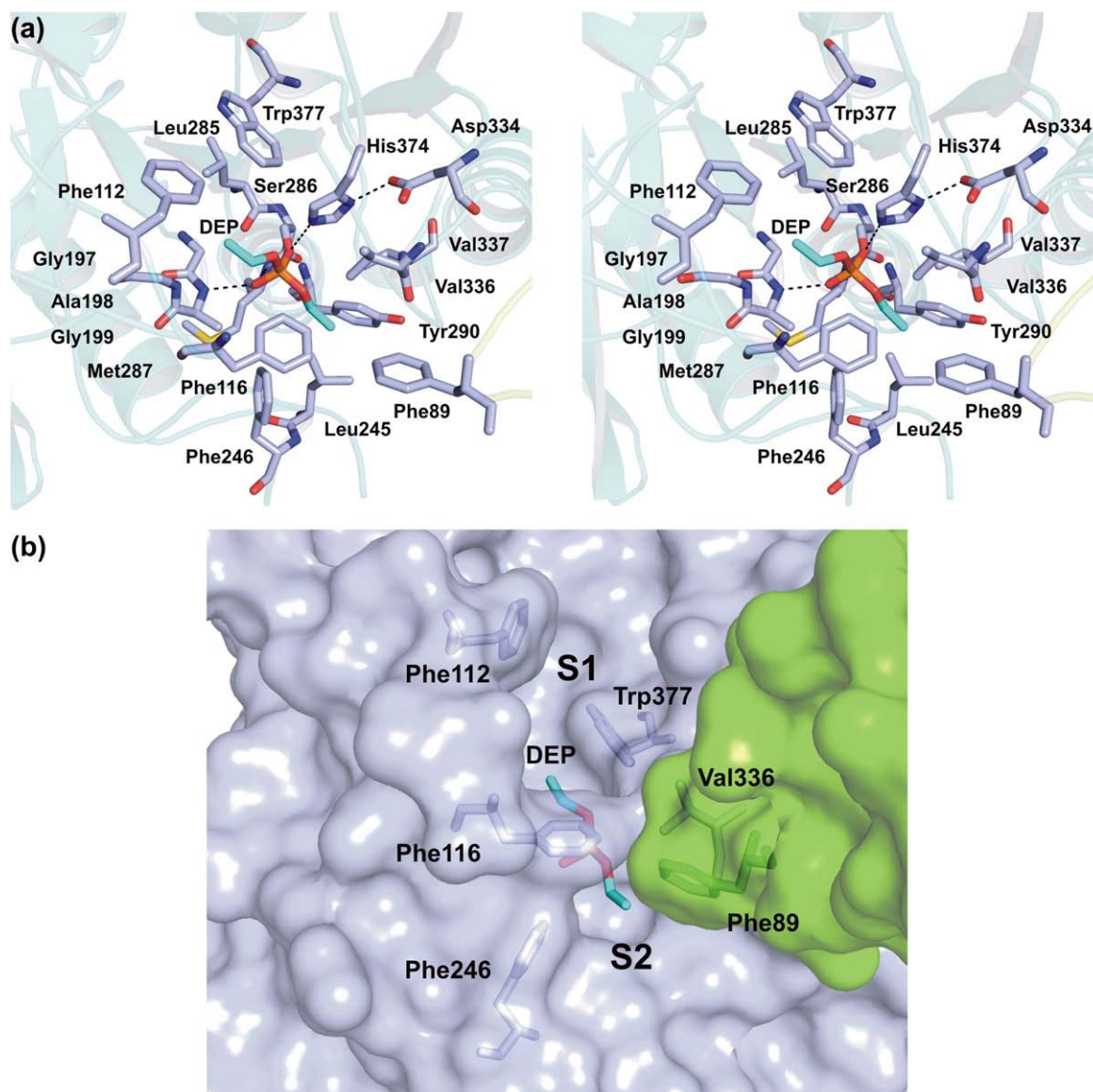
Fig. 5. Substrate-binding gorge. (a) Stereo view of the active site with the bound DEP intermediate covalently bound to the catalytic serine. Key residues of the catalytic gorge are shown as sticks, with the hydrogen bonds shown in dashed lines. (b) Surface representation of the EstA catalytic gorge, with some of the key residues and the intermediate DEP shown in stick mode. These images were generated using PyMOL.[8]

7-aminocephalosporanic acid (7-ACA), glucose pentaacetate, and *N*-acetyl-D-glucosamine. EstA displayed an activity of 20 U mg^{-1} on glucose pentaacetate. This activity is relatively low, suggesting

Table 1. Kinetic parameters for the hydrolysis of various *p*-nitrophenyl esters

p-Nitrophenyl esters	K_m (μM)	K_{cat} (s^{-1})	K_{cat}/K_m (s^{-1} mM^{-1})
Acetate (C2)	105 ± 10	115 ± 4	1095 ± 111
Butyrate (C4)	414 ± 37	99 ± 3	239 ± 23
Valerate (C5)	183 ± 17	85 ± 2	464 ± 43
Hexanoate (C6)	89 ± 14	79 ± 4	888 ± 147
Octanoate (C8)	27 ± 6	37 ± 3	1370 ± 324
Decanoate (C10)	8 ± 1	10 ± 0.4	1250 ± 164
Dodecanoate (C12)	6 ± 0.6	1.6 ± 0.03	267 ± 27

that EstA is not an oligosaccharide deacetylase. EstA was also able to hydrolyze acetyl groups from both cephalosporin C and 7-ACA with an activity of 80 U mg^{-1} for both substrates. Cephalosporin C and 7-ACA are not stable at high temperatures and are, therefore, not considered natural substrates. EstA was not able to remove the acetyl group from *N*-acetyl-D-glucosamine, indicating that it is specific for ester bonds and unable to cleave amide bonds.

Mutational studies

Five residues, all near the active site or part of the active-site gorge, were changed to alanines by site-directed mutagenesis in order to analyze their importance for EstA activity. The residues selected for mutagenesis were three phenylalanines (Phe89,
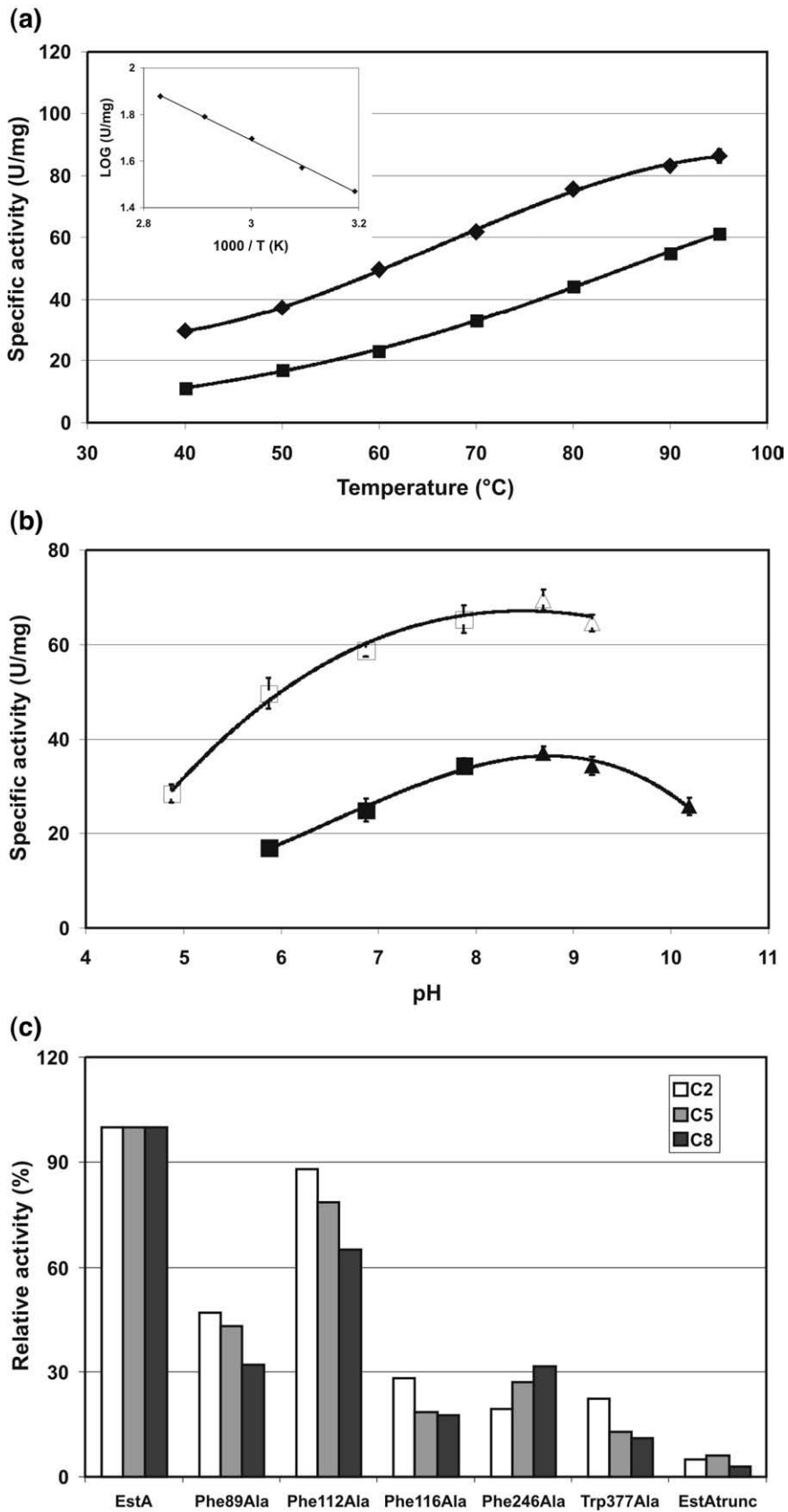
(a)

(b)

(c)

Fig. 6 (*legend on next page*)

Phe112, and Phe116) located on loops of the Ig-like domain and two conserved residues (Phe246 and Trp377) of the catalytic domain. Phe89 is located on a loop (Tyr77-Tyr85) coming from the Ig-like domain of a subunit B (EstA multimer) and is a part of the side gorge. Phe116 is located at the top of a very long loop (Leu100-Ile126), leading all the way back to the active site and covering the active site. Phe112 is on the same long loop and is part of the groove. The conserved Phe246 is part of the tunnel, and Trp377 is part of the active-site gorge.

The mutants Phe89Ala, Phe112Ala, Phe116Ala, Phe246Ala, and Trp377Ala were expressed and purified. The mutants were analyzed by mass spectrometry, which revealed that all mutants were present as hexamer in solution. Esterase activity was determined using substrates with different acyl-chain lengths (*p*NP-C2, *p*NP-C5, and *p*NP-C8) (Fig. 6c). A significant reduction in activity was observed for all five mutants, confirming their importance for the activity of EstA. Their activity relative to EstA with the substrate *p*NP-C5 was, respectively, approximately 30% for Phe89Ala, 80% for Phe112Ala, 20% for Phe116Ala, 25% for Phe246Ala, and 15% for Trp377Ala. As can be seen in Fig. 6c, the effect was more pronounced for longer acyl-chain substrates.

In addition, a truncated mutant of EstA coding only for the catalytic domain (EstAtrunc: Asp158-Arg395) was constructed. EstAtrunc was expressed, purified, and found to have a relative activity of only 5% compared to EstA (Fig. 6c). Furthermore, it had a much lower temperature optimum (60 °C), had lower thermostability (a half-life at 90 °C of 15 min), and was present as a monomeric structure (analyzed by mass spectrometry).

Discussion

A hypothetical protein with esterase features from the hyperthermophilic bacterium *T. maritima* was produced heterologously and proved to exhibit ester-hydrolyzing activity. The highest activities were found on *p*-nitrophenol derivatives with short acyl chains (C2 and C4). In accordance, the enzyme also showed activity with tributyrin, but not with triacylesters with longer chains. However, because of a high catalytic efficiency for the acetate, as well as for the octanoate-*p*-nitrophenol derivative, we cannot exclude that the physiological substrate may contain acyl chains up to C8. Nevertheless, these data indicate that EstA should be classified as an esterase (<C10), not as a lipase.

The three-dimensional structure of the protein was determined at 2.6 Å resolution. Analysis of the EstA structure confirms that it is a member of the α/β hydrolase family, with a conserved Ser-Asp-His catalytic triad comprising Ser286, Asp334, and His374. The active site can be accessed via a gorge flanked with predominantly hydrophobic residues. The structure was found to be composed of two clearly distinguishable domains: a C-terminal domain containing the active site and an unusual N-terminal domain resembling Igs. Such a combination of an esterase domain or an α/β hydrolase domain with an Ig-like domain is new and has not, as such, yet been described.

Analysis of the quaternary structure by gel filtration, mass spectrometry, and DLS revealed that EstA predominantly exists as a hexamer in solution. The crystal structure also shows a hexameric arrangement composed of two trimers. EM demonstrated that the hexamer in solution is identical with the hexamer in the crystal and is constructed as a dimer of trimers, with the N-terminal Ig-like domains facing each other. Esterases often have a trimeric structure, as was, for example, described for the thermostable esterase from *Bacillus circulans*[22] and the thermostable esterase from *Sulfolobus shibatae*.[23] A hexameric structure, however, is rather unusual. Being at the interface of the two trimers, the Ig-like domain apparently has a function in multimerization.

The function of the Ig-like domain in other bacterial enzymes has been proposed to be substrate binding, directing a substrate to the catalytic groove or cell adhesion.[24,25] The latter option seems unlikely for EstA, with the Ig-like domains facing each other. To elucidate the function of the Ig-like domain of EstA, a truncated mutant composed only of the C-terminal catalytic domain was constructed. The resulting EstAtrunc was still active, but had lost 95% of its activity and was no longer able to form hexamers. This again points to a role of the Ig-like domain in multimerization. The inability to form multimers may also be the reason for the reduced stability observed at higher temperatures. Apolar residues at the interface become exposed to the solvent and may contribute to the observed loss of stability. Multimerization is a phenomenon often described for enzymes from (hyper)thermophiles and is regarded as one of different mechanisms to increase thermostability.[26]

The active site is accessible via a gorge, typical of many α/β hydrolases. Unusual, however, is that the active site is also accessible via a second side gorge. This side gorge could possibly provide an access for the substrate or an exit for one of the reaction pro-

Fig. 6. Effect of temperature, pH, and mutations on esterase activity. (a) The effect of temperature on esterase activity was studied using *p*-nitrophenyl-valerate (■) and *p*-nitrophenyl-octanoate (♦) as substrates at temperatures ranging from 40 °C to 95 °C. Inset: The temperature dependence for *p*-nitrophenyl-valerate as an Arrhenius plot. (b) The effect of pH on esterase activity was studied using *p*-nitrophenyl-valerate (□, citrate phosphate buffer; Δ, CAPS buffer) and *p*-nitrophenyl-octanoate (■, citrate phosphate buffer; ▲, CAPS buffer) as substrates at pH values in the range 4.8–10.2. (c) The effect of mutations on esterase activity was studied using *p*-nitrophenyl-acetate (C2), *p*-nitrophenyl-valerate (C5), and *p*-nitrophenyl-octanoate (C8) at pH 8 and 70 °C.

ducts following a nucleophilic attack and formation of the intermediate. On the other hand, compared to other esterases and lipases, the EstA active-site pocket is unique in its closure by Val336 and Phe116. Considering that Phe116 is located in a flexible loop, as indicated by a high *B*-factor for the fragment Phe112-Leu117, it is possible that, upon substrate binding, Phe116 could change its conformation and open the tunnel, making it more accessible.

Besides Phe116, the active site is surrounded by a set of aromatic residues. To disclose their function, five residues (Phe89, Phe112, Phe116, Phe246, and Trp377) in proximity to the active site were mutated to alanines. The specific activity of each mutant was decreased, indicating that they all are important for activity. Least affected was mutant Phe112Ala; however, this residue is located farthest from the active site. The most pronounced inhibiting effects were seen with the more hydrophobic longer-chain esters. This observation suggests that the hydrophobic residues facilitate the entrance of the substrate along the gorge, which would hold most for the longer acyl chains. On the other hand, a more general role of the residues in the stabilization of the active site, however, cannot be excluded. One of the mutated residues, Phe89, is located on a loop coming from the Ig-like domain of a subunit B. This interaction of the Ig-like domain of one subunit with the active site of another subunit also supports the view that multimerization is important for—although not essential to—the activity of EstA, since Phe89Ala still has 30% activity.

In conclusion, the structural and biochemical characterization of EstA showed that it is an unusual esterase composed of a conserved C-terminal catalytic domain and an unprecedented N-terminal Ig-like domain. The Ig-like domain presumably plays a role in the multimerization of EstA into an unusual hexameric structure. Additionally, it may also participate in the catalysis of EstA by guiding the substrate to the active site. Further mutagenesis and biochemical studies are needed for a better understanding of the role of the N-terminal domain.

Materials and Methods

Protein production and crystallization

The *T. maritima estA* gene (locus tag TM0033) was cloned into the expression vector pET24d without the predicted signal peptide (the first 16 amino acids). EstA was expressed and purified as described.[6] The purified native and selenomethionine (SeMet) EstA proteins were dialyzed against 10 mM potassium phosphate buffer (pH 7.5) and concentrated to 15 mg ml^{-1}. EstA was crystallized by hanging-drop vapor-diffusion at room temperature. Crystals of EstA were obtained using a reservoir solution consisting of 1.0 M lithium sulphate monohydrate and 2% wt/vol polyethylene glycol 8000. Drops consisting of equal volumes (1 μl) of protein and reservoir solution were equilibrated over 500-μl reservoirs. Crystals suitable for X-ray diffraction were obtained within 1 week and 2 weeks, respectively, for the native and SeMet-EstA

proteins. A crystal of EstA in complex with its inhibitor paraoxon was obtained by incubating the enzyme (15 mg ml^{-1}) supplemented with 0.2 mM paraoxon for 1 h at room temperature and then by setting up crystallization as described above. Crystals were obtained within 2 weeks.

Data collection

For cryoprotection, crystals were soaked for a few seconds in a reservoir solution containing 20% (wt/vol) glycerol. The crystals were mounted in a cryoloop and subsequently flash-frozen in liquid nitrogen. X-ray data were collected at 100 K on beamline ID29 at the European Synchrotron Radiation Facility (Grenoble, France). A native data set was collected to 2.6 Å resolution. The crystal belongs to space group *H*32, with unit cell parameters *a* = 130.2 Å, *b* = 130.2 Å, and *c* = 306.2 Å. There are two molecules in the asymmetric unit that have a V_M of 2.9 Å^3 Da^{-1} and a solvent content of 58%.[27] Crystals of SeMet-EstA showed a well-defined Se *K* absorption edge by fluorescence scanning. A single SeMet-EstA crystal was used for MAD data collection at the peak (0.9791 Å), inflection (0.9793 Å), and remote (0.9557 Å) wavelengths up to 2.6 Å resolution. Data were indexed and integrated with MOSFLM[28] and scaled using SCALA.[29]

Structure determination and refinement

The EstA structure was solved by MAD phasing with MAD data from the SeMet-EstA crystal and the native data set, using HKL2MAP.[30] Eight selenium sites in the asymmetric unit of the crystal were found and used to calculate phases to 2.6 Å resolution (Table 2). However, initial density maps were of generally poor quality and not suitable for tracing the structure. Phases were improved using RESOLVE,[31,32] allowing the identification of NCS. NCS averaging and solvent flattening were performed using the program DM[33] of the CCP4 suite, giving an electron density map of better quality. Autobuilt models from RESOLVE and ARP/wARP[34] were combined to give a starting model comprising 170 residues from a total of 380, with only 20 residues assigned into sequence properly. The structure was then manually rebuilt in Coot[35] and refined using CNS[36] and REFMAC.[37] Strict NCS restraints were applied during the earlier stages of the refinement and released at later stages. In the final stages of refinement, solvent molecules were added using ARP-wARP[34] and manually inspected in Coot.[35] A sulfate ion from the reservoir solution was clearly visible at high contour level in the omit map. The final refinement R_{work} is 19.7%, and R_{free} is 26.7%.

The structure of the EstA–paraoxon complex was determined by molecular replacement, which was performed with MOLREP using the native structure as model. The complex structure was rebuilt in Coot[35] and refined using REFMAC.[37] The $F_o - F_c$ electron density map and omit density map displayed clear density for paraoxon and were used to assign the head of the paraoxon molecule. No water molecules were picked for this low-resolution data set. The model with paraoxon was refined to a final R_{work} of 22.0% and an R_{free} of 26.2%.

In the final models (native and paraoxon-complexed crystal structures), residues 16–20 of both molecules and residues 249–250 of chain B were missing because of poor density. The geometry of both models was monitored using PROCHECK,[38] with the native and paraoxon complex models having 85.2% and 82.2% of their residues

Table 2. Data collection, phasing, and refinement statistics

	Native	SeMet peak	SeMet inflection	SeMet remote	EstA paraoxon complex
Data collection					
Wavelength (Å)	1.0000	0.9791	0.9793	0.9757	1.0000
Resolution range (Å)	50–2.6 (2.74–2.6)[a]	50–2.6 (2.74–2.6)	50–2.6 (2.74–2.6)	50–2.6 (2.74–2.6)	50–3.0 (3.16–3.0)
Space group	H32	H32	H32	H32	H32
Unit cell parameters (Å)					
a	130.2	131.0	131.0	131.0	130.5
b	130.2	131.0	131.0	131.0	130.5
c	306.2	306.8	306.8	306.8	304.5
Observed reflections	173,066	353,388	358,408	358,161	160,768
Unique reflections	31,079	31,295	31,587	31,457	20,405
Completeness (%)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
R_{merge}[b]	0.087 (0.430)	0.086 (0.498)	0.082 (0.412)	0.075 (0.437)	0.152 (0.643)
$\langle I/\sigma(I)\rangle$	14.2 (3.6)	22.3 (4.6)	22.1 (4.4)	24.8 (5.0)	13.4 (3.0)
Redundancy	5.6 (5.7)	11.4 (11.7)	11.3 (11.4)	11.4 (11.3)	7.9 (8.1)
MAD phasing					
Number of Se sites		8			
Figure of merit		0.5			
Refinement					
Resolution range (Å)	20–2.6 (2.66–2.6)				35–3.0 (3.08–3.0)
Number of reflections	28,373				19,311
R_{work} (%)	19.7 (31.2)				22.0 (30.7)
R_{free} (%)	26.7 (35.5)				26.2 (35.6)
Average *B*-factors					
Protein	36.8				45.8
Water	33.2				—
Ligand	49.5				39.6
RMSD					
Bond lengths (Å)	0.018				0.016
Bond angles (°)	1.88				1.77

[a] Values in parentheses correspond to the highest-resolution shell.
[b] $R_{merge} = \sum_h \sum_l |I_{hl} - \langle I_h \rangle| / \sum_h \sum_l \langle I_h \rangle$, where I_l is the *l*th observation of reflection *h*, and $\langle I_{\boldsymbol{h}} \rangle$ is the weighted average intensity for all observations *l* of reflection *h*.

in the most favored regions of the Ramachandran plot, respectively. Cartoon representations were generated using PyMOL.[8]

Native mass spectrometry

Native mass spectrometry measurements were performed in positive ion mode using an Electrospray Ionization Time-of-Flight instrument (LC-T; Micromass, Manchester, UK) equipped with a Z-spray nanoelectrospray ionization source. Needles were made from borosilicate glass capillaries (Kwik-Fil; World Precision Instruments, Sarasota, FL) on a P-97 puller (Sutter Instruments, Novato, CA), coated with a thin gold layer using an Edwards Scancoat (Edwards Laboratories, Milpitas, CA) six Pirani 501 sputter coater. To produce intact ions *in vacuo* from EstA in solution, the ions were cooled by increasing the pressure in the first vacuum stages of the mass spectrometer. In addition, efficient desolvation was needed to sharpen the ion signals in order to withdraw the oligomeric states of EstA from the mass spectrum. Therefore, source pressure conditions were raised to values ranging from 7.0 mbar to 7.3 mbar, and nanoelectrospray voltages were optimized for transmission of large protein complexes. The pressure in the interface region was adjusted by reducing the pumping capacity of the rotary pump by closing the speed valve.[39,40] The applied voltages on the needle and sample cone were 1300 V and 150 V, respectively. All spectra were mass-calibrated using an aqueous solution of cesium iodide (20 mg ml^{-1}). Buffer exchange of EstA samples with 100 mM ammonium acetate (pH 6.8) was performed using ultrafiltration units with a cutoff of 5000 Da (Millipore, Bedford, MA). EstA was diluted to 5 μM and measured at room temperature.

Dynamic light scattering

DLS was performed at room temperature using the EstA sample that was ready for crystallization. The EstA particles were monodisperse with a molecular mass of approximately 267 kDa, suggesting that the 43-kDa subunits assemble into a hexamer in solution.

EM and single-particle analysis

Samples of purified EstA were negatively stained with 2% uranyl acetate on glow-discharged carbon-coated copper grids. Images were recorded with a Gatan 4 K slow-scan charge-coupled device camera on a Philips CM12 electron microscope (Fei, Eindhoven, The Netherlands) operated at 120 kV, using GRACE software for semiautomated specimen selection and data acquisition.[41] The final magnification was 100,000×, with a pixel size (after binning the images) of 3.0 Å at the specimen level. About 8700 single particles were selected and extracted from 600 electron micrographs. Single-particle analysis was performed with the GRoningen Image Processing (GRIP) software package on a PC cluster. The single-particle projections (96×96 pixel frame) were subjected to multireference alignment and reference-free alignment procedure, multivariate statistical analysis, and hierarchical ascendant classification.[42] From the whole data set,

almost 45% of single particles were assigned to top views of the EstA hexamer, whereas the remaining ∼55% resembled side views. The final two-dimensional projection maps of the presented views of the EstA hexamer were calculated from the best-resolved classes, which represented about 20% of the whole data.

The packing of EstA within the crystal was examined with PyMOL software[8] for possible hexameric structures that resembled the projection maps obtained with EM and single-particle analysis. Possible structures were manually fitted into the EM projections. Different views of the EstA hexamer were finally displayed using PyMOL software.[8] Truncated versions at 15 Å resolution and two-dimensional projection maps of the generated model for the EstA hexamer were generated using routines from the EMAN package.[43]

Enzyme assays

Esterase activity was measured as described previously.[20] In short, in a standard assay, activity was measured with 0.2 mM *p*-nitrophenyl octanoate as substrate in 50 mM citrate phosphate (pH 8) at 70 °C. The amount of *p*-nitrophenol liberated was measured continuously at 405 nm on a Hitachi U-2001 spectrophotometer with a temperature-controlled cuvette holder. Extinction coefficients of *p*-nitrophenol were determined prior to each measurement.

The effect of pH on esterase activity was studied in the pH range 5–11. The buffers used were 50 mM citrate phosphate (pH 5–8) and 50 mM CAPS (3-(cyclohexylamino) 1-propanesulphonic acid) buffer (pH 9.5–11). The effect of temperature on esterase activity was studied in the range 40–95 °C. The pH of the buffers was set at 25 °C, and temperature corrections were made using their temperature coefficients: -0.0028 pH $°C^{-1}$ for citrate-phosphate buffer, and -0.018 pH $°C^{-1}$ for CAPS buffer.[44] Enzyme thermostability was determined by incubating the enzyme in 50 mM Tris–HCl and 150 mM NaCl (pH 7.8) buffer at 100 °C for various time intervals. Residual activity was assayed under standard conditions.

Plate assays were used for the detection of activity on triacylglycerol esters. Agar plates containing tributyrin, trioctanoate, triolein, palm oil, and olive oil (1%, vol/vol) were prepared, and 5-mm-wide holes were perforated. The resulting holes were loaded with purified enzyme and incubated at 50 °C. Activity was detected by zones of clearance around the holes.

Deacetylase activity was determined using high-performance liquid chromatography by measuring the amount of acetic acid released from the substrates cephalosporin C, 7-ACA, glucose pentaacetate, and N-acetyl-D-glucosamine. The reaction mixture contained 0.9 ml of substrate solution (dissolved in 50 mM Tris–HCl, pH 7.5) and 0.1 ml of enzyme solution, and was incubated at 37 °C for various time intervals. The reaction was stopped by adding 0.2 ml of stop solution (100 mM H_2SO_4 and 30 mM crotonate) and placing the sample on ice. The conditions for high-performance liquid chromatography were as follows: column, KC811-Shodex; detection, RI and UV detectors; solvent, 3 mM H_2SO_4; flow rate, 1.5 ml min^{-1}; temperature, 30 °C; internal standard, crotonate. One unit of enzyme activity was defined as the amount of enzyme that releases 1 μmol of acetic acid per minute.

The effect of inhibitors was studied by preincubating EstA with 1 mM inhibitor in 50 mM citrate phosphate buffer (pH 8) at 37 °C for 1 h. Subsequently, samples were placed on ice, and residual activity was measured using the standard assay. The activity of the enzyme without

inhibitor was defined as 100%. Inhibition kinetics of paraoxon were determined as described for the acetylcholinesterase from electric eel.[16]

Kinetic parameters were determined by fitting the data obtained from multiple measurements by a computer-aided direct fit to the Michaelis–Menten curve (Tablecurve 2d, version 5.0).

Mutagenesis

Mutants of EstA were created to identify the function of the Ig-like domain. Mutants Phe89Ala, Phe112Ala, Phe116Ala, Phe246Ala, and Trp377Ala were generated using Quikchange (Stratagene) site-directed mutagenesis with the following primers: BG2486 5-CGGAGGTCTCT-CTTTCAGCATTGCCAACAACCGTGGAAAGTACG-3 and BG2487 5-CGTACTTTCCACGGTTGTTGGCAATGC-TGAAAGAGAGACCTCCG-3; BG2488 5-CACTCAAAC-ACCATTGTTGCCGGGCCTAATTTTCTCAACACCCG-3 and BG2645 5-CGGGTGTTGAGAAAATTAGGCCCGGC-AACAATGGTGTTTGAGTG-3; BG2490 5-CCATTGTTTTC-GGGCCTAATGCGCTCAACACCCGCATGAAACTGG-3 and BG2646 5-CCAGTTTCATGCGGGTGTTGAGCGCAT-TAGGCCCGAAAACAATGG-3; BG2492 5-CAGCAGCT-GGTCCACACTCGCCACTGACAGGGAAAATCC-3 and BG2647 5-GGATTTTCCCTGTCAGTGGCGAGTGTGGA-CCAGCTGCTG-3; and BG2494 5-GGTGGGATCCACAC-GGATCGGCGATACCTACCTACGAG-3 and BG2648 5-CTCGTAGGTAGGTATCGCCGATCCGTGTGGATCC-CACC-3, respectively (with modified codons underlined). A truncated EstA mutant was generated using the primers BG2496 5-CCGCCATGGATTTCCTCGCATTCACTTTCA-AAGACC-3 and BG1963 5-GCGCTCGAGTCTACTTTG-TTCAAACAGCCAC-3. The sequence of the mutants was confirmed by sequence analysis of both DNA strands.

Data bank accession codes

The structural data have been deposited in the RCSB PDB with accession code 3DOH for native EstA and accession code 3DOI for EstA in complex with paraoxon.

References

1. 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.

2. Jaeger, K. E. & Eggert, T. (2002). Lipases for biotechnology. *Curr. Opin. Biotechnol.* **13**, 390–397.

3. Atomi, H. & Imanaka, T. (2004). Thermostable carboxylesterases from hyperthermophiles. *Tetrahedron: Asymmetry*, **15**, 2729–2735.

4. Huber, R. & Hannig, M. (2006). Thermotogales. *Prokaryotes*, **7**, 899–922.

5. Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H. *et al.* (1999). Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature*, **399**, 323–329.

6. Sun, L., Levisson, M., Hendriks, S., Akveld, T., Kengen, S. W., Dijkstra, B. W. & van der Oost, J. (2007). Crystallization and preliminary crystallographic analysis of an esterase with a novel domain from the hyperthermophile *Thermotoga maritima*. *Acta Crystallogr. Sect. F*, **63**, 777–779.

7. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402.

8. DeLano, W. L. (2002). The PyMOL Molecular Graphics System DeLano Scientific, San Carlos, CA.

9. Holm, L. & Sander, C. (1996). The FSSP database: fold classification based on structure–structure alignment of proteins. *Nucleic Acids Res.* **24**, 206–209.

10. Ito, K., Nakajima, Y., Xu, Y., Yamada, N., Onohara, Y., Ito, T. *et al.* (2006). Crystal structure and mechanism of tripeptidyl activity of prolyl tripeptidyl aminopeptidase from *Porphyromonas gingivalis*. *J. Mol. Biol.* **362**, 228–240.

11. Rasmussen, H. B., Branner, S., Wiberg, F. C. & Wagtmann, N. (2003). Crystal structure of human dipeptidyl peptidase IV/CD26 in complex with a substrate analog. *Nat. Struct. Biol.* **10**, 19–25.

12. Schubot, F. D., Kataeva, I. A., Blum, D. L., Shah, A. K., Ljungdahl, L. G., Rose, J. P. & Wang, B. C. (2001). Structural basis for the substrate specificity of the feruloyl esterase domain of the cellulosomal xylanase Z from *Clostridium thermocellum*. *Biochemistry*, **40**, 12524–12532.

13. Zhang, H., Seabra, M. C. & Deisenhofer, J. (2000). Crystal structure of Rab geranylgeranyltransferase at 2.0 A resolution. *Structure*, **8**, 241–251.

14. Hage, T., Sebald, W. & Reinemer, P. (1999). Crystal structure of the interleukin-4/receptor alpha chain complex reveals a mosaic binding interface. *Cell*, **97**, 271–281.

15. Jacobson, R. H., Zhang, X. J., DuBose, R. F. & Matthews, B. W. (1994). Three-dimensional structure of beta-galactosidase from *E. coli*. *Nature*, **369**, 761–766.

16. Forsberg, A. & Puu, G. (1984). Kinetics for the inhibition of acetylcholinesterase from the electric eel by some organophosphates and carbamates. *Eur. J. Biochem.* **140**, 153–156.

17. Febbraio, F., D'Andrea, S. E., Mandrich, L., Merone, L., Rossi, M., Nucci, R. & Manco, G. (2008). Irreversible inhibition of the thermophilic esterase EST2 from *Alicyclobacillus acidocaldarius*. *Extremophiles*, **12**, 719–728.

18. van den Heuvel, R. H. & Heck, A. J. (2004). Native protein mass spectrometry: from intact oligomers to functional machineries. *Curr. Opin. Chem. Biol.* **8**, 519–526.

19. Goettig, P., Groll, M., Kim, J. S., Huber, R. & Brandstetter, H. (2002). Structures of the tricorn-interacting aminopeptidase F1 with different ligands explain its catalytic mechanism. *EMBO J.* **21**, 5343–5352.

20. Levisson, M., van der Oost, J. & Kengen, S. W. (2007). Characterization and structural modeling of a new type of thermostable esterase from *Thermotoga maritima*. *FEBS J.* **274**, 2832–2842.

21. Lang, D. A., Mannesse, M. L., de Haas, G. H., Verheij, H. M. & Dijkstra, B. W. (1998). Structural basis of the chiral selectivity of *Pseudomonas cepacia* lipase. *Eur. J. Biochem.* **254**, 333–340.

22. Kademi, A., Ait-Abdelkader, N., Fakhreddine, L. & Baratti, J. (2000). Purification and characterization of a thermostable esterase from the moderate thermophile *Bacillus circulans*. *Appl. Microbiol. Biotechnol.* **54**, 173–179.

23. Ejima, K., Liu, J., Oshima, Y., Hirooka, K., Shimanuki, S., Yokota, Y. *et al.* (2004). Molecular cloning and characterization of a thermostable carboxylesterase from an archaeon, *Sulfolobus shibatae* DSM5389: nonlinear kinetic behavior of a hormone-sensitive lipase family enzyme. *J. Biosci. Bioeng.* **98**, 445–451.

24. Bork, P., Holm, L. & Sander, C. (1994). The immunoglobulin fold: structural classification, sequence patterns and common core. *J. Mol. Biol.* **242**, 309–320.

25. Halaby, D. M. & Mornon, J. P. (1998). The immunoglobulin superfamily: an insight on its tissular, species, and functional diversity. *J. Mol. Evol.* **46**, 389–400.

26. Korkhin, Y., Kalb, A. J., Peretz, M., Bogin, O., Burstein, Y. & Frolow, F. (1999). Oligomeric integrity—the structural key to thermal stability in bacterial alcohol dehydrogenases. *Protein Sci.* **8**, 1241–1249.

27. Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497.

28. Leslie, A. G. W. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4+ ESF-EAMCB Newsl. Protein Crystallogr.* **26**, 27–33.

29. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sect. D*, **50**, 760–763.

30. Pape, T. & Schneider, T. R. (2004). HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Crystallogr.* **37**, 843–844.

31. Terwilliger, T. C. (2000). Maximum-likelihood density modification. *Acta Crystallogr. Sect. D*, **56**, 965–972.

32. Terwilliger, T. C. (2003). Automated side-chain model building and sequence assignment by template matching. *Acta Crystallogr. Sect. D*, **59**, 45–49.

33. Cowtan, K. (1994). 'dm': An automated procedure for phase improvement by density modification. *Joint CCP4 ESF-EACBM Newsl. Protein Crystallogr.* **31**, 34–38.

34. Perrakis, A., Morris, R. & Lamzin, V. S. (1999). Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6**, 458–463.

35. Emsley, P. & Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D*, **60**, 2126–2132.

36. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W. *et al.* (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. Sect. D*, **54**, 905–921.

37. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. Sect. D*, **53**, 240–255.

38. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291.

39. Tahallah, N., Pinkse, M., Maier, C. S. & Heck, A. J. (2001). The effect of the source pressure on the abundance of ions of noncovalent protein assemblies in an

electrospray ionization orthogonal time-of-flight instrument. *Rapid Commun. Mass Spectrom.* **15**, 596–601.

40. van Berkel, W. J., van den Heuvel, R. H., Versluis, C. & Heck, A. J. (2000). Detection of intact megadalton protein assemblies of vanillyl-alcohol oxidase by mass spectrometry. *Protein Sci.* **9**, 435–439.

41. Oostergetel, G. T., Keegstra, W. & Brisson, A. (1998). Automation of specimen selection and data acquisition for protein electron crystallography. *Ultramicroscopy*, **74**, 47–59.

42. van Heel, M., Gowen, B., Matadeen, R., Orlova, E. V., Finn, R., Pape, T. *et al.* (2000). Single-particle electron cryo-microscopy: towards atomic resolution. *Q. Rev. Biophys.* **33**, 307–369.

43. Ludtke, S. J., Baldwin, P. R. & Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128**, 82–97.

44. Beynon, R. J. & Easterby, J. S. (2003). Buffer Solutions Taylor and Francis Group, London.