Biochemical and Structural Analysis of Thermostable Esterases

Mark Levisson

Thesis supervisors:

Prof. dr. J. van der Oost
Personal chair at the Laboratory of Microbiology
Wageningen University

Prof. dr. W.M. de Vos
Professor of Microbiology
Wageningen University

Thesis co-supervisor:

Dr. S.W.M. Kengen
Assistant Professor at the Laboratory of Microbiology
Wageningen University

Other members:

Prof. dr. ir. H. Gruppen
Wageningen University

Dr. M.C.R Franssen
Wageningen University

Dr. T. Sonke
DSM, Geleen

Prof. dr. B.W. Dijkstra
University of Groningen

Dr. A.M.W.H Thunnissen
University of Groningen

Biochemical and Structural Analysis of Thermostable Esterases

Mark Levisson

TABLE OF CONTENTS

Aim and outline

AIM AND OUTLINE

This thesis describes the results of biochemical and structural analyses of thermostable esterases. The aim of this research project was to isolate and produce new esterases from hyperthermophilic microorganisms, to gain insight in their catalytic properties and to investigate their potential as biocatalysts in food-related conversions.

Many processes in industry are operated under conditions, which are unfavorable for many biocatalysts. In this regard, enzymes from hyperthermophiles are promising candidates since they generally have a high intrinsic stability. The initial project aimed also at investigating the behavior of thermostable enzymes under high pressure or microwave irradiation. These methods may open new ways for the tuning of enzyme reactions, and thus offers the possibility of expanding the area of biocatalysis in the food industry. However, after disappointing pilot experiments, technical problems and better insight into high pressure theory, this approach was abandoned. Therefore, the main focus in this thesis is on the isolation, and biochemical and structural characterization of distinct esterases from the bacterium *Thermotoga maritima* (chapter 2-6) and a lipase from the archaeon *Archaeoglobus fulgidus* (chapter 7). The information obtained in this study provides fundamental knowledge, which may be useful for an industrial application.

Chapter 1 - Introduction

The first chapter gives a general overview of all currently characterized carboxylic ester hydrolases from hyperthermophilic bacteria and archaea. The biochemical properties, structures and classification of these enzymes are discussed. In addition, approaches for the discovery of new carboxylic ester hydrolases are described.

Chapter 2 - Characterization and structural modeling of a new type of thermostable esterase from *Thermotoga maritima*

The second chapter describes the identification, heterologous production, purification and biochemical characterization of an esterase (EstD) from *T. maritima*. A structural model was constructed based on threading and provided insight into the active site and substrate binding. Residues involved in catalysis were verified by site-directed mutagenesis and inhibition studies. Phylogenetic analysis of EstD suggested a new family of esterases.

Chapter 3 - Crystallization and preliminary crystallographic analysis of an esterase with a novel domain from the hyperthermophile *Thermotoga maritima*

The third chapter describes the cloning, purification, crystallization and preliminary X-ray analysis of an esterase (EstA) from *T. maritima*. Native and selenomethionine-substituted EstA was crystallized by the hanging-drop vapour-diffusion method. Multiple wavelength anomalous data sets were collected to 2.6 Å resolution and an initial analysis is described.

Chapter 4 - Crystal structure and biochemical properties of a novel thermostable esterase containing an immunoglobulin-like domain

After obtaining its preliminary crystallographic data, as described in Chapter 3, EstA was investigated further, in order to provide insight into the function of the immunoglobulin-like domain and reveal the molecular basis of substrate recognition and catalysis. The structure of EstA, native and in complex with the competitive inhibitor paraoxon, is described and the quaternary structure was investigated using dynamic light scattering, mass-spectrometry and electron microscopy. At present, this is the only esterase that has been described to have an immunoglobulin-like domain.

Chapter 5 - The crystal structures of a thermostable acetyl esterase / cephalosporin C deacetylase from *Thermotoga maritima* in complex with PMSF and paraoxon reveal an altered conformation of the catalytic serine

Chapter five reports on the crystal structure of an acetyl esterase, presumably involved in xylan degradation. Insight into substrate binding was obtained from co-crystal structures with the inhibitors PMSF and paraoxon. Various biochemical properties and the positional specificity of the esterase was investigated.

Chapter 6 - Purification and partial characterization of a thermostable esterase from *Thermotoga maritima*

The sixth chapter describes the cloning, purification, crystallization and partial biochemical characterization of an esterase (EstB) from *T. maritima*. EstB was crystallized by hanging-drop vapour-diffusion and a dataset was collected to ~2.8 Å resolution. Its structure solution is ongoing.

Chapter 7 - Characterization of a lipase from the hyperthermophilic archaeon *Archaeoglobus fulgidus*

No true lipases, hydrolyzing long chain fatty acid esters, have thus far been identified in hyperthermophiles. The seventh chapter describes the identification, cloning, purification and partial characterization of a lipase (LipA) from *A. fulgidus*. Lipase activity on triacylglycerol esters was determined using qualitative plate assays. LipA was crystallized by hanging-drop vapour-diffusion and a dataset was collected to ~2.6 Å resolution. Its structure solution is ongoing.

Chapter 8 - Summary and general discussion

This final chapter is a brief summary of the findings described in this thesis. Discussed are the physiological role of esterases, the use of microwave irradiation and high pressure for biocatalysis, and the general aspects of the enzymes described in this thesis, with some concluding remarks.

1

Chapter

Introduction

INTRODUCTION

The synthesis of specific products by enzymes is a fundamental aspect of modern biotechnology. This biocatalytic approach has several advantages over traditional chemical engineering, such as higher product purity, fewer waste products, lower energy consumption and more selective reactions due to the high regio- and stereo-selectivity of enzymes [1]. One of the industrially most exploited and important groups of biocatalysts are the carboxylic ester hydrolases (EC 3.1.1.X) [2;3].

Carboxylic ester hydrolases are ubiquitous enzymes, which have been identified in all domains of life (Bacteria, Archaea and Eukaryotes), and in some viruses. In the presence of water, they catalyze the hydrolysis of an ester-bond resulting in the formation of an alcohol and a carboxylic acid. However, in an organic solvent they can catalyze the reverse reaction or a trans-esterification reaction (Figure 1.1) [4]. Most carboxylic ester hydrolases belong to the α/β-hydrolase family and share structural and functional characteristics, including a catalytic triad, an α/β-hydrolase fold, and a co-factor independent activity. The catalytic triad is conserved and is usually composed of a nucleophilic serine in a GXSXG pentapeptide motif (where X is any residue), and an acidic residue (aspartate or glutamate) that is hydrogen-bonded to a histidine residue [5;6;7;8].
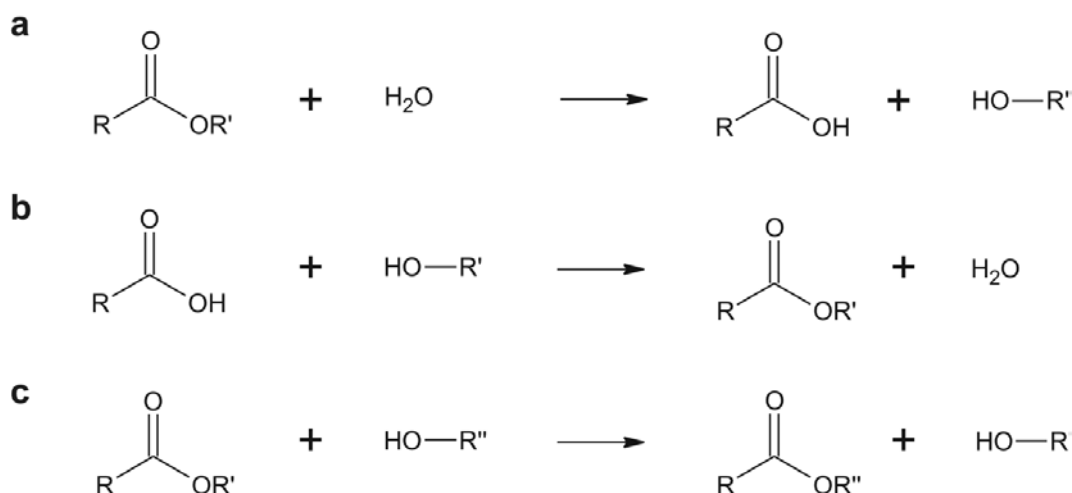
Figure 1.1: Reactions catalyzed by carboxylic ester hydrolases: a) hydrolysis, b) esterification and c) transesterification.

There are two well-known groups within the family of carboxylic ester hydrolases: lipases and esterases. Esterases differ 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 substrates that form micelles [9]. Other groups include, for instance, arylesterases and phospholipases. The physiological role of carboxylic ester hydrolases is often not known, but nevertheless, many have found applications in industry; amongst other in medical biotechnology, detergent

production, organic synthesis, biodiesel (methyl-ester) production, flavor and aroma synthesis, and other food related processes [10; 11].

The use of enzymes in industrial processes also has its restrictions. Many processes are operated at elevated temperatures or in the presence of organic solvents. These conditions are detrimental to most enzymes and therefore there is a growing demand for enzymes with an improved stability. In this regard, especially, enzymes from hyperthermophiles are promising candidates because these enzymes generally display a high intrinsic thermal and chemical stability [12]. In recent years, many new hyperthermophiles have been isolated and the genomes of a rapidly increasing number have been completely sequenced. Hyperthermophiles have proven to be a good source of new enzymes [13; 14; 15], including many putative esterases and lipases.

At this moment, most esterases and lipases used in industry are from mesophiles, basically, because they were the first to be identified and characterized. Esterases and lipases have only been isolated from a small number of hyperthermophiles (Table 1.1). An excellent review on thermostable carboxylesterases from hyperthermophiles appeared in 2004 [16]. However, since then, many new hyperthermophilic carboxylic ester hydrolases have been described. Therefore, in this review we aim to present an overview of the currently characterized carboxylic ester hydrolases from hyperthermophiles. We will focus on the identification of new carboxylic ester hydrolases, the biochemical properties and 3D-structures of characterized enzymes, and their classification. For details on the application of these enzymes, we refer to other reviews that cover this aspect extensively [2; 11; 13].

HYPERTHERMOPHILES

Hyperthermophiles are generally defined as micro-organisms that grow optimally at temperatures above 80°C [17]. They have been isolated from both terrestrial and marine environments, such as sulfur-rich solfataras (pH ranging from slightly alkaline to extremely acidic), hot-springs, oil field waters and hydrothermal vents at the ocean floor. Consequently, they show a broad physiological diversity, ranging from aerobic respirers to methanogens and saccharolytic heterotrophs [17; 18]. Hyperthermophiles can be found in both prokaryotic domains, *viz.* the Bacteria and the Archaea. In phylogenetic trees based on 16SrRNA they occupy the shortest and deepest lineages, suggesting they might be closely related to the common ancestor of all extant life [19]. For this reason and because they are a potential source of new biocatalysts, the genomes of several hyperthermophiles have been completely sequenced (Table 1.1).

All biomolecules of hyperthermophiles must be stabilized against thermal denaturation. The simplest approach for DNA stabilization would be to increase the GC-content of the DNA. However, it has been established that the GC-content of hyperthermophiles does not correlate

Table 1.1: List of completely sequenced hyperthermophiles (T-optimum > 80°C) and extreme thermophiles (no growth < 50°C)

Organism	Genome size (bp)	Number of ORFs	GC content (%)	Optimal growth (°C)	Esterases isolated
Bacteria					
Aquifex aeolicus VF5	1551335	1529	43.5	90	
Caldicellulosiruptor saccharolyticus DSM 8903	2970275	2679	35.3	70[a]	
Thermoanaerobacter tengcongensis MB4T / JCM 11007	2689445	2588	37.6	75[a]	+
Thermotoga lettingae TMO	2135342	2040	38.7	65[a]	
Thermotoga maritima MSB8	1860725	1858	46.2	80	+
Thermotoga petrophila RKU-1	1823511	1785	46.1	80	
Thermotoga sp. RQ2	1877693	1819	46.2	76-82	
Thermus thermophilus HB8	1849742	1973	69.4	75[a]	+
Thermus thermophilus HB27	1894877	1982	66.6	68[a]	+
Archaea					
Aeropyrum pernix K1	1669695	1700	56.3	90	+
Archaeoglobus fulgidus VC-16	2178400	2420	48.6	82	+
Desulfurococcus kamchatkensis 1221n	1365223	1475	-	85	
Hyperthermus butylicus DSM 5456	1667163	1602	53.7	95-106	
Methanocaldococcus jannaschii DSM 2661	1739933	1729	31.4	85	
Methanopyrus kandleri AV19	1694969	1687	61.2	98	
Methanothermobacter thermoautotrophicus Delta H	1751377	1873	49.5	65-70[a]	
Nanoarchaeum equitans Kin4-M	490885	536	31.6	90	
Pyrobaculum aerophilum IM2	2222430	2605	51.4	100	
Pyrobaculum arsenaticum PZ6	2121076	2298	58.3	95	
Pyrobaculum calidifontis JCM 11548	2009313	2149	57.2	90-95	+
Pyrobaculum islandicum DSM 4184	1826402	1978	49.6	100	
Pyrococcus abyssi GE5	1765118	1896	44.7	103	+
Pyrococcus furiosus DSM 3638	1908256	2125	40.8	100	+
Pyrococcus horikoshii OT3	1738505	1955	41.9	98	
Staphylothermus marinus F1	1570485	1570	35.7	92	
Sulfolobus acidocaldarius DSM 639	2225959	2292	36.7	70-75[a]	+
Sulfolobus tokodaii 7, JCM 10545	2694756	2825	32.8	80	+
Sulfolobus solfataricus P2	2992245	2977	35.8	80	+
Thermococcus onnurineus NA1	1847607	1976	51.3	80	
Thermococcus kodakaraensis KOD1	2088737	2306	52.0	85	
Thermofilum pendens Hrk 5	1781889	1824	57.7	88	
Thermoproteus neutrophilus V24Sta	1769823	1966	59.9	85	

Information concerning completely sequenced genomes and on-going sequence projects can be obtained at the GOLD Genomes Online Database (http://www.genomesonline.org) [23].
[a] Extreme thermophiles that are related to hyperthermophiles.

with the optimal growth temperatures (Table 1.1). Instead, other mechanisms are used to stabilize DNA, such as an increased intracellular electrolyte concentration, cationic DNA binding proteins, and DNA supercoiling [15]. Thus far, all completely sequenced hyperthermophiles have a reverse gyrase catalyzing a positive supercoiling of their DNA. A reverse gyrase is however not a prerequisite for hyperthermophilic life, but it can be seen as a marker for growth at high temperatures [20].

Proteins from hyperthermophiles have also been optimized for functioning at elevated temperatures. There is no single mechanism responsible for the stability of these hyperthermophilic proteins, rather, it can be attributed to multiple features. Features that

contribute to the stability of hyperthermophilic enzymes include (i) changes in amino acid composition, such as a decrease in the thermolabile residues asparagine and cysteine (ii) increased hydrophobic interactions, (iii) an increased number of ion pairs and salt bridge networks, (iv) reduction in the size of surface loops and of solvent-exposed surface, (v) as well as increased intersubunit interactions and oligomeric state [15; 18; 21]. Besides these structural adaptations proteins can also be stabilized by intracellular solutes, metabolites and sugars [22].

BIOMINING FOR NEW ENZYMES

Traditionally, new biocatalysts were discovered by a cumbersome screening of a wide variety of organisms for the desired activity. A modern variant is the metagenomics approach, which involves the extraction of genomic DNA from environmental samples, its cloning into suitable expression vectors and subsequent screening of the constructed libraries [24]. This approach has been successfully applied to isolate new biocatalysts, including carboxylic ester hydrolases from hyperthermophiles [25; 26]. This approach can potentially result in unique enzymes (no sequence similarity), but obviously depends on functional expression. At present, with many complete genome sequences available, bioinformatics has become an important tool in the discovery of new biocatalysts. This is a high-throughput approach for the identification, and *in silico* functional analysis, of more or less related sequences encoding potential biocatalysts. Sequence similarity, based on sequence alignments and motif searches, is most commonly used for assigning a function to new proteins [27].

Many sequences in the available databases have already been annotated as putative esterase or lipase. However, even more carboxylic ester hydrolases can be identified when BLAST and Motif searches, in combination with pair-wise comparison with sequences of known carboxylic ester hydrolases, are used. The advantage of this approach compared to traditional activity screening is the direct identification of new and diverse carboxylic ester hydrolases, which would otherwise not have been detected due to a low level of expression.

Such a bioinformatics approach has been successfully applied to identify new carboxylic ester hydrolase sequences in the completely sequenced genomes of several selected hyperthermophiles. In order to have as many possible candidates, also sequences that were assigned a different function, but did have the characteristics of carboxylic ester hydrolases, were included, such as acylpeptide hydrolases. The results can be found in Table 1.2. A typical strategy includes: BLAST-P searches [28] using sequences of known carboxylic ester hydrolases as template. In parallel, searching InterPro [29] for potential candidates. The resulting sequences can then be further analyzed (for conserved motifs and domains) using the NCBI Conserved Domain Search [30; 31].

Table 1.2: Identified sequences of potential carboxylic ester hydrolases in selected genomes

Microorganism	Locus tag	Genbank	Annotation (NCBI)	Residues	GxSxG
Aeropyrum pernix	APE1244	BAA80234	Hypothetical protein	583	GVSMG
	APE1832	BAA80835	Acylpeptide hydrolase/Esterase	659	GGSYG
	APE2361	BAA81374	Hydrolase, putative	279	GFSLG
	APE2441	BAA81456	Acylpeptide hydrolase/Esterase	595	GGSYG
Hyperthermus butylicus	Hbut_1071	ABM80914	Hypothetical protein	226	GLSVG
Pyrobaculum aerophilum	PAE2936	AAL64548	Hypothetical protein	194	GPSAS
	PAE3573	AAL65014	Hypothetical protein	196	GHSMG
Pyrobaculum calidifontis	Pcal_1307	ABO08731	Alpha/beta hydrolase	313	GDSAG
	Pcal_1997	ABO09412	Hypothetical protein	198	GHSMG
Pyrococcus abyssi	PAB1050	CAB50498	Lysophospholipase, putative	259	GHSLG
	PAB2176	CAB49187	Hypothetical esterase	286	GFSMG
Sulfolobus solfataricus	SSO0102	AAK40458	Esterase, tropinesterase	231	GHSIG
	SSO2262	AAK42427	Hypothetical protein	197	GASMG
	SSO2518	AAK42649	Esterase, putative	353	GESFG
	SSO2521	AAK42652	Lipase	311	GDSAG
	SSO2979	AAK43083	Hypothetical protein	320	GHSSG
	SSO3052	AAK43152	Hypothetical protein	210	GISGN
Thermoanaerobacter tengcongensis	TTE0035	AAM23348	Hypothetical protein	237	GDSIS
	TTE0419	AAM23703	Lysophospholipase	314	GHSFG
	TTE0552	AAM23828	Predicted hydrolase	279	GVSMG
	TTE0556	AAM23832	Predicted hydrolase	298	GWSMG
	TTE1809	AAM25001	Alpha/beta hydrolase	258	GLSMG
	TTE2321	AAM25462	Alpha/beta hydrolase	414	CHSMG
	TTE2547	AAM25672	Alpha/beta hydrolase	285	AHSFG
Thermococcus kodakaraensis	TK0522	BAD84711	Carbohydrate esterase	449	GSSLG
Thermotoga maritima	TM1022	AAD36099	Esterase	253	GLSMG
	TM1160	AAD36236	Esterase	306	GLSAG
	TM1350	AAD36421	Lipase, putative	259	GHSLG

PROPERTIES OF CHARACTERIZED ESTERASES

The first carboxylic ester hydrolase isolated and characterized from a hyperthermophile was a carboxylesterase from *Sulfolobus acidocaldarius* [32; 33]. Since then, many new esterases have been characterized. At this moment, most carboxylic ester hydrolases described from hyperthermophiles are esterases and only recently the first lipase from a hyperthermophile was identified (Levisson et al., manuscript in preparation). Esterases have been characterized from *Thermoanaerobacter tengcongensis*, *Thermotoga maritima*, *Thermus thermophilus*, *Aeropyrum pernix*, *Archaeoglobus fulgidus*, *Picrophilus torridus*, *Pyrobaculum calidifontis*, *Pyrococcus abyssi*, *Pyrococcus furiosus*, *Sulfolobus acidocaldarius*, *Sulfolobus shibatae*, *Sulfolobus solfataricus*, *Sulfolobus tokodaii* and from metagenomic libraries (Table 1.3).

Substrate preference

Enzymes are classified and named according to the type of reaction they catalyse (Enzyme Commission). The carboxylic ester hydrolases all catalyse the hydrolysis of carboxylic acid esters, but they can be further clustered into different groups based on their substrate of preference. Two well-known members of this family are the esterases and true lipases. The majority of the characterized hyperthermophilic carboxylic ester hydrolases are esterases. Lipases have been described for many mesophiles, mainly microbial and fungal, and are exploited for biotechnological applications [2; 34]. However, until recently no true lipase, hydrolyzing long chain fatty acid esters, had been identified in hyperthermophiles. The first lipase was characterized from the archaeon *A. fulgidus* (Levisson *et al.*, manuscript in preparation) (Table 1.3). This lipase shows maximal activity at a temperature of 95°C and has a half-life of 10 hours at 80°C. It displays highest activity with *p*-nitrophenyl-decanoate (*p*NP-C10) and is capable of hydrolyzing triacylglycerol esters of butyrate (C4), octanoate (C8), palmitate (C16) and oleate (C18). Two lipases from the thermophile *Thermosyntropha lipolytica*, LipA and LipB, have been characterized and are very stable at high temperatures [35]. Both enzymes show maximal activity at 96°C and have the highest activity with the triacylglycerol ester trioleate and *p*NP-C12. LipA and LipB retained 50% of their activity after 6 and 2h incubation at 100°C, respectively, indicating that these two lipases are the most thermostable ones so far reported. Unfortunately, attempts to clone the two lipases were unsuccessful. A few mesophilic lipases may operate at temperatures above 80°C, but they usually have short half-lifes. An exception is a mesophilic lipase that was isolated from *a Pseudomonas sp.*, which showed a half-life of over 13h at 90°C [36]. In comparison, the well-known lipase B from *Candida antartica* (CALB, Novozym 435) has a half-life of only 2 hours at 45°C [37].

Esterases have a preference for short to medium acyl-chain esters (Table 1.3). Several enzymes from hyperthermophiles have been tested for activity toward esters with various alcoholic moieties other than the standard *p*NP-esters or 4-methylumbelliferyl (4MU) esters (Figure 1.2). The esterase from *P. calidifontis* displays activity towards different acetate esters and showed highest activity on iso-butyl acetate [38]. Furthermore, it was able to hydrolyze *sec*- and *tert*-butyl acetate. At present, only few enzymes can catalyze the hydrolysis or the synthesis of tertiary esters. This is because known esterases and lipases cannot hydrolyse esters containing a bulky substituent near the ester carbonyl group.

Other esterases have been characterized for their ability to resolve mixtures of chiral esters. The kinetic resolution of the esterase Est3 from *S. solfataricus* P2 was investigated using (*R,S*)-ketoprofen methyl ester (Figure 1.2) [39]. The enzyme hydrolyzed the (*R*)-ester of racemic ketoprofen methylester and showed an enantiomeric excess of 80% with a conversion rate of 20% in 32h. In another study, the esterase SSo-Est1 from *S. solfataricus* P1 [40] was identified as homologue to the mesophilic *Bacillus subtilis* ThaiI-8 esterase (CNP) [41] and *Candida rugosa* lipase (CRL) [42], which are used for the chiral separation of racemic mixtures of 2-arylpropionic

Table 1.3: Biochemical properties of characterized carboxylic ester hydrolases

Microorganism	Enzyme	Locus tag	Preferred substrate [a]	K_M (µM)	K_{cat} (s⁻¹)	K_{cat}/K_M (s⁻¹ × µM)	T_{opt} (°C)	Optimal pH	Stability	Molecular mass (kDa)	PDB	References[b]
Bacteria												
Thermotoga maritima	Esterase	TM0033	pNP-C2 pNP-C8	105 27	115 37	1.095 1.370	95+	8.5	½ life of 1.5h at 100°C	267 (α6)	3DOH 3DOI	68; 70
Thermotoga maritima	Esterase	TM0053	pNP-C10	3.1	11	3.5	60	7	Retained 50% activity after 30 min at 80°C	40		71
Thermotoga maritima	Esterase	TM0053	pNP-C8	NR	NR	NR	90	9	NR	38.5		c
Thermotoga maritima	Acetyl esterase	TM0077	pNP-C2	185	57.5	0.311	100+	7.5	½ life of 2h at 90°C	222 (α6)	1VLQ	d
Thermotoga maritima	Esterase	TM0336	pNP-C5	66	10.2	0.155	95+	7	½ life of 1h at 100°C	44.5 (α1)		52
Thermoanaerobacter tengcongensis	Esterase	TTE0555	pNP-C3/C4	NR	NR	NR	70	9	½ life of 1.5h at 70°C	43		72
Thermosyntropha lipolytica	Lipase (LipA)	NR	pNP-C12	NR	NR	NR	96	9.4	Retained 50% activity after 6h at 100°C	50		35
Thermosyntropha lipolytica	Lipase (LipB)	NR	pNP-C12	NR	NR	NR	96	9.6	Retained 50% activity after 2h at 100°C	57		35
Thermus thermophilus HB8	Putative esterase	TT1662	NR	NR	NR	NR	NR	NR	NR	26	1UFO	73
Thermus thermophilus species	Esterases	NR	NR	NR	NR	NR	80	NR	NR	34 / 62		73; 74; 75; 76; 77; 78; 79
Archaea												
Aeropyrum pernix	Esterase / Acylpeptide hydrolase	APE1547	pNP-C8	43.3	6.6	0.152	90	8	½ life of 160h+ at 90°C	63 (α1)	1VE6 1VE7	65; 66; 67; 80; 81; 82; 83; 84
Aeropyrum pernix	Phospholipase Esterase	APE2325	pNP-C3	103	39	NR	90	NR	½ life of 1h at 100°C	18		85
Archaeoglobus fulgidus	Esterase	AF1041	NR	NR	NR	NR	NR	NR	NR			86
Archaeoglobus fulgidus	Esterase	AF1716	pNP-C6	11	1014	92	80	7.1	½ life of 1h at 85°C	35.5	1JJI	58; 59; 86; 87; 88; 89; 90; 91; 92
Archaeoglobus fulgidus	Esterase	AF1763	pNP-C18	876	47.5	0.054	70	10-11	Retained 40% activity after 30min at 40°C	53		93
Archaeoglobus fulgidus	Lipase	AF1763	pNP-C10	NR	NR	NR	95	11	½ life of 10h at 80°C	51		c
Archaeoglobus fulgidus	Esterase	NR	pNP-C4	NR	NR	NR	70	7-8	Retained 10% activity after 3h at 90°C	27.5		94
Picrophilus torridus	Esterase (EstA)	PTO0988	pNP-C2	NR	NR	NR	70	6.5	½ life of 21h at 90°C	66 (α3)		95
Picrophilus torridus	Esterase (EstB)	PTO1141	pNP-C2	NR	NR	NR	55	7	½ life of 10h at 90°C	81(α3)		95
Pyrobaculum calidifontis	Esterase	NR	pNP-C6	44.4	2620	59	90	7	½ life of 56min at 110°C	98 (α3)		38

Organism	Enzyme	Name	Substrate					Topt (°C)	pH	Stability	MW kDa (subunit)	PDB	Ref
Pyrococcus abyssi	Esterase	NR	pNP-C5	NR	NR	NR	NR	65-74	NR	½ life of 22h at 99°C / ½ life of 13min at 120°C	NR		54
Pyrococcus furiosus	Esterase	NR	4MU-C2	NR	NR	NR	NR	100	7.5	½ life of 34h at 100°C / ½ life of 2h at 120°C	NR		53
Pyrococcus furiosus	Lysophospholipase Esterase	PF0480	NR	NR	NR	NR	NR	70	7	NR	64 (α2)		96
Pyrococcus furiosus	Esterase	PF2001	4MU-C7	NR	NR	NR		60	7	Retained 100% activity after 2h at 75°C	48		97; 98
Sulfolobus acidocaldarius	Esterase	NR	pNP-C5	151.7	NR	NR	NR	NR	7.5-8.5	Retained 50% activity after 1h at 100°C	128 (α4)		32; 33; 97; 98
Sulfolobus acidocaldarius	Esterase	NR	NR	NR	NR	NR	NR	NR	NR	½ life of 45 min at 80°C	NR		99
Sulfolobus acidocaldarius	Phosphotriesterase	SACI2140	Methyl-paraoxon	1400	7.75	5.57×10^{-3}	NR	75	9	Retained 65% activity after 2h at 85°C	69 (α2)		47
Sulfolobus shibatae	Esterase	NR	NR	NR	NR	NR	NR	90	6	½ life of 20 min at 120°C	NR		55
Sulfolobus shibatae	Esterase	NR	pNP-C4	10	NR	NR	NR	NR	7-8	Retained 70% activity after 30 min at 90°C	90 (α3) 64 (α2)		100
Sulfolobus solfataricus MT4	Esterase	NR	pNP-C5	NR	NR	NR	NR	90+	6.5-7	½ life of 7h at 90°C	114 (α4)		101
Sulfolobus solfataricus MT4	Phosphotriesterase	NR	Methyl-paraoxon	205	1.3	6.34×10^{-3}	NR	95+	7-9	½ life of 1.5h at 100°C	35	2VC5 2VC7	45; 47; 102; 103
Sulfolobus solfataricus P1	Esterase	NR	4MU-C2* pNP-C6	45 *	1000 *	2.2 *	NR	95+	7.7	NR	33 (α1)		40; 43; 44; 104
Sulfolobus solfataricus P1	Esterase	NR	pNP-C8	71	14700	207.1	NR	85	8	Retained 41% activity after 120h at 80°C	34 (α1)		50
Sulfolobus solfataricus P1	Aryl esterase	NR	Paraoxon	5	597	119.4	NR	94	7	Retained 52% activity after 50h at 90°C	35 (α1)		46
Sulfolobus solfataricus P2	Esterase	SSO2493	pNP-C5	2100	46.3	21.1×10^{-3}	NR	80	7.4	½ life of 40 min at 80°C	96 (α3)		39
Sulfolobus solfataricus P2	Esterase	SSO2517 (SsoNΔ)	pNP-C6	50	2.5	0.05	NR	70	7.1	NR	NR		105; 106
Sulfolobus solfataricus P2	Esterase	SSO2517 (SsoNΔlong)	pNP-C6	30	34.5	1.15	NR	85	6.5	NR	34		106
Sulfolobus solfataricus P2	Esterase	NR	NR	NR	NR	NR	NR	75	8	½ life of 15 min at 100°C	100		107
Sulfolobus tokodaii strain 7	Esterase	ST0071	pNP-C4	0.53	127	0.239	NR	70	8	½ life of 40 min at 85°C	34		108
Metagenomic													
Metagenomic library	Esterase	NR	pNP-C6	700	1600	2.29	NR	95+	6	½ life of 20 min at 90°C	34	2C7B	25; 63; 64; 109
Metagenomic library	Phospholipase	NR	pNP-C4	140	574	4.101	NR	70	9	Retained 50% activity after 2h at 80°C	32		26
Metagenomic library	Esterase	NR	pNP-C5	120	110	0.921	NR	70	9	Retained 50% activity after 30 min at 80°C	29		26

[a] Abbreviations concerning substrates: pNP (p-nitrophenyl ester), 4MU (4-methylumbelliferyl ester); [b] References contain all current literature concerning the enzymes in this table; [c] Levisson *et al.*, manuscript in preparation, [d] Levisson *et al.*, unpublished results.

methyl esters. The enzyme was characterized biochemically for its ability to resolve mixtures of (*R,S*)-naproxen methyl ester under a variety of reaction environments [43; 44]. Sso-Est1 showed a specific reaction toward the (*S*)-naproxen ester in co-solvent reaction conditions with an enantiomeric excess of ≥90%.

In addition to esterases and lipases, other ester hydrolase types have been identified in hyperthermophiles, including two phosphotriesterases and an arylesterase that were found in *S. acidocaldarius, S. solfataricus* MT4 and P1, respectively [45; 46; 47]. The phosphotriesterases showed maximal activity on the organophosphate methyl-paraoxon (dimethyl *p*-nitrophenyl phosphate) and the arylesterase showed maximal activity on paraoxon (diethyl *p*-nitrophenyl phosphate) (Figure 1.2) (Table 1.3). Besides this phospho-esterase activity, also esterase activity (on *p*NP-esters) was observed for both enzymes. Stable organophosphate-degrading enzymes are of great interest for the detoxification of chemical warfare agents and agricultural pesticides.
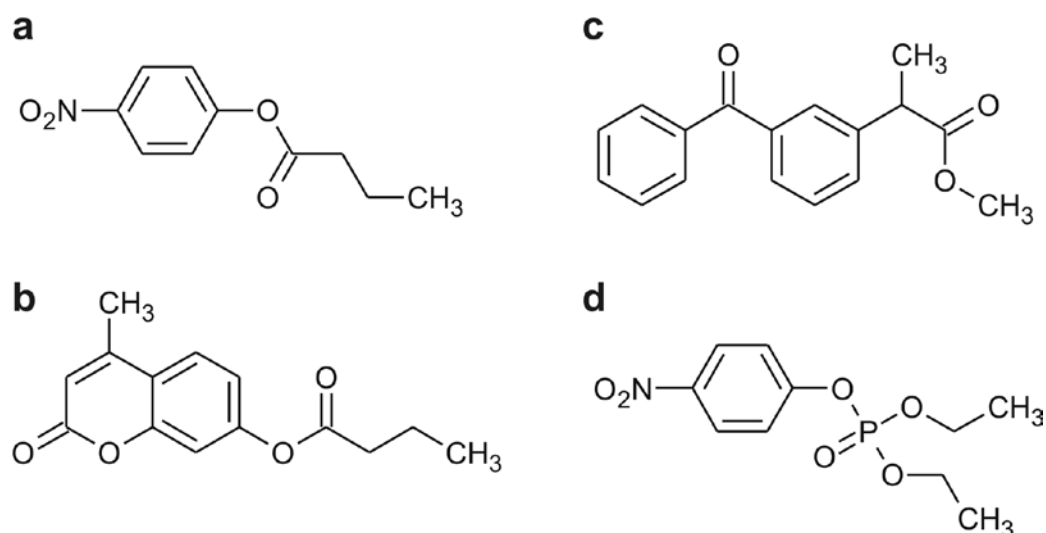

Figure 1.2: Substrates commonly used to test for esterase activity: a) *p*-nitrophenyl butyrate, b) 4-methylumbelliferyl butyrate, c) (R/S)-ketoprofen methyl ester and d) *p*-nitrophenyl diethyl phosphate.

Stability against chemicals

Stability and activity in the presence of organic solvents and detergents is an important property of an enzyme if it is to be used as a biocatalyst in industry. Several hyperthermophilic carboxylic ester hydrolases have been tested. The esterase from *P. calidifontis* [38] displays high stability in water-miscible organic solvents, and exhibited activity in 50% solutions of DMSO, methanol, acetonitrile, ethanol and 2-propanol. In addition, the enzyme retained almost full activity after 1 hour incubation in the presence of the above mentioned organic solvents at a concentration of 80%. In comparison, the lipases from the mesophiles *Pseudomonas sp.* B11-1 [48] and *Fusarium heterosporum* [49] were completely inactivated after incubation with acetonitrile.

In addition to stability against solvents the *Pyrobaculum* enzyme also has high thermal stability, with a half-life of approx. 1h at 110°C (Table 1.3). The esterase from *S. solfataricus* P1 [50] also displayed good stability against organic solvents, comparable to the enzyme from *P. calidifontis*. In addition, addition of 5% non-ionic detergents, such as Tween 20, stabilized the *Sulfolobus* enzyme. Moreover, the enzyme retained 45% and 98% activity in the presence of 5% SDS and 8M urea, respectively. The lipase from the mesophile *Penicillium expansum* shows a much lower stability against detergents or organic solvents [51]. The esterase EstD from *T. maritima* [52] does not display resistance to detergents and only retained 0% and 43% activity in the presence of 1% (v/v) SDS and 1% (v/v) Tween 20, respectively. However, EstD does show good resistance against organic solvents since it remained actity in the presence of 10% (v/v) solvents, which is comparable to the esterase from *P. calidifontis*. The esterase Est3 from *S. solfataricus* P2 displayed good resistance against mild detergents [39], it retained 51% and 99% activity in the presence of respectively 10% (w/v) Tween 60 and 10% (w/v) Tween 80, but displayed lower stability against organic solvents than the other three esterases described above.

Thermal stability

The most thermostable carboxylic ester hydrolase described to date is an esterase from *P. furiosus* [53] (Table 1.3). It is extremely stable with half-life's of 34 and 2 hours at 100 and 120°C, respectively. The enzyme has optimal activity at a temperature of 100°C, which is in good agreement with the optimal growth temperature of *Pyrococcus* (100°C). Highest activity was obtained with the substrate MU-C2, however, also little activity toward *p*NP-C18 was detected indicating it has a very broad substrate tolerance. Another very stable esterase was detected in crude extracts of *P. abyssi* [54]. This enzyme has a half-life of 22h and 13 minutes at 99 and 120°C, respectively. Maximal esterase activity was observed at least 65-74°C, however, temperatures above 74°C were not investigated due to instability of the substrate. The enzyme is active on a broad range of substrates, capable of hydrolyzing triacylglycerol esters and aromatic esters, but is restricted to short acyl chain esters of C2-C8 with an optimum for C5 fatty acid esters. Unfortunately, no sequence information has been reported for both *Pyrococcus* esterases. Most of the characterized carboxylic ester hydrolases from hyperthermophiles are optimally active at temperatures between 70 and 100°C (Table 1.3), which is often close to or above the host organism's optimal growth temperature. Furthermore, it is interesting to note that some carboxylic ester hydrolases, such as the esterase from *S. shibatae* [55] and the acetyl esterase from *T. maritima* (Levisson *et al.*, manuscript in preparation), after heterologous expression in *Escherichia coli*, show a transient activation during stability incubations, indicating they probably need a high temperature in order to fold properly. Compared to their mesophilic counterparts they perform similar functions, however due to intrinsic differences

hyperthermophilic enzymes are stable and can operate at higher temperatures. It is difficult to indicate exactly which factors contribute to this higher thermal stability since (as discussed before) many different factors are involved.

Structures

Most carboxylic ester hydrolases conform to a common structural organization: the α/β-hydrolase fold, which is also present in many other hydrolytic enzymes like proteases, dehalogenases, peroxidases and epoxide hydrolases [56]. The canonical α/β-hydrolase fold consists of an eight-stranded mostly parallel β-sheet, with the second strand anti-parallel. The parallel strands β3 to β8 are connected by helices, which pack on either side of the central beta-sheet. The sheet is highly twisted and bent so that it forms a half-barrel. The active site contains the catalytic triad consisting of the residues serine, aspartate and histidine (Figure 1.3a) [6; 7; 8]. The substrate-binding site is located inside a pocket on top of the central *β*-sheet that is typical of this fold. The size and shape of the substrate-binding cleft have been related to substrate specificity [57].

The three-dimensional structures of several hyperthermophilic esterases have been solved (Table 1.3) (Figure 1.3). The first reported structure of an hyperthermostable esterase was for the esterase AFEST of *A. fulgidus* (PDB: 1JJI) [58]. AFEST is an esterase that belongs to the hormone-sensitive lipase (HSL) group of esterases and lipases. The structure was refined to 2.2 Å resolution and showed that AFEST has the typical α/β-hydrolase fold. The active site is shielded by a cap region composed of five α-helices. Access to the active site of many lipases and some esterases is shielded by a mobile lid, whose position (closed or open) determines whether the enzyme is in an inactive or active conformation. AFEST is an esterase that prefers *p*NP-C6 as a substrate and shows maximal activity at 80°C. It is stable at high temperatures with a half-life of 1 hour at 85°C [59]. A comparison of the AFEST structure with its mesophilic and thermophilic homologues, Brefeldin A from *Bacillus substilus* (BFAE) (PDB: 1JKM) [60] and EST2 from *Alicyclobacillus acidocaldarius* (PDB: 1EVQ) [61], showed which structural features contribute to its thermal stability. The comparison revealed an increase in the number of intramolecular ion pairs, and a reduction in loop extensions and ratio of hydrophobic to charged surface residues [58; 62].

The structure of the esterase EstE1 was solved to 2.1 Å resolution (PDB: 2C7B) [63]. This enzyme, which was isolated from a metagenomic library, also belongs to the HSL group and is closely related with AFEST. EstE1 has the canonical architecture of the α/β-hydrolase fold and also contains a cap domain like other members of the HSL group [58]. It exhibits highest esterase activity on short acyl chain esters of length C6 and has a half-life of 20 minutes at 90°C [25]. The thermal stability of EstE1 seems to be achieved mainly by its dimerization through hydrophobic interactions and ion-pair networks that both contribute to the stabilization of
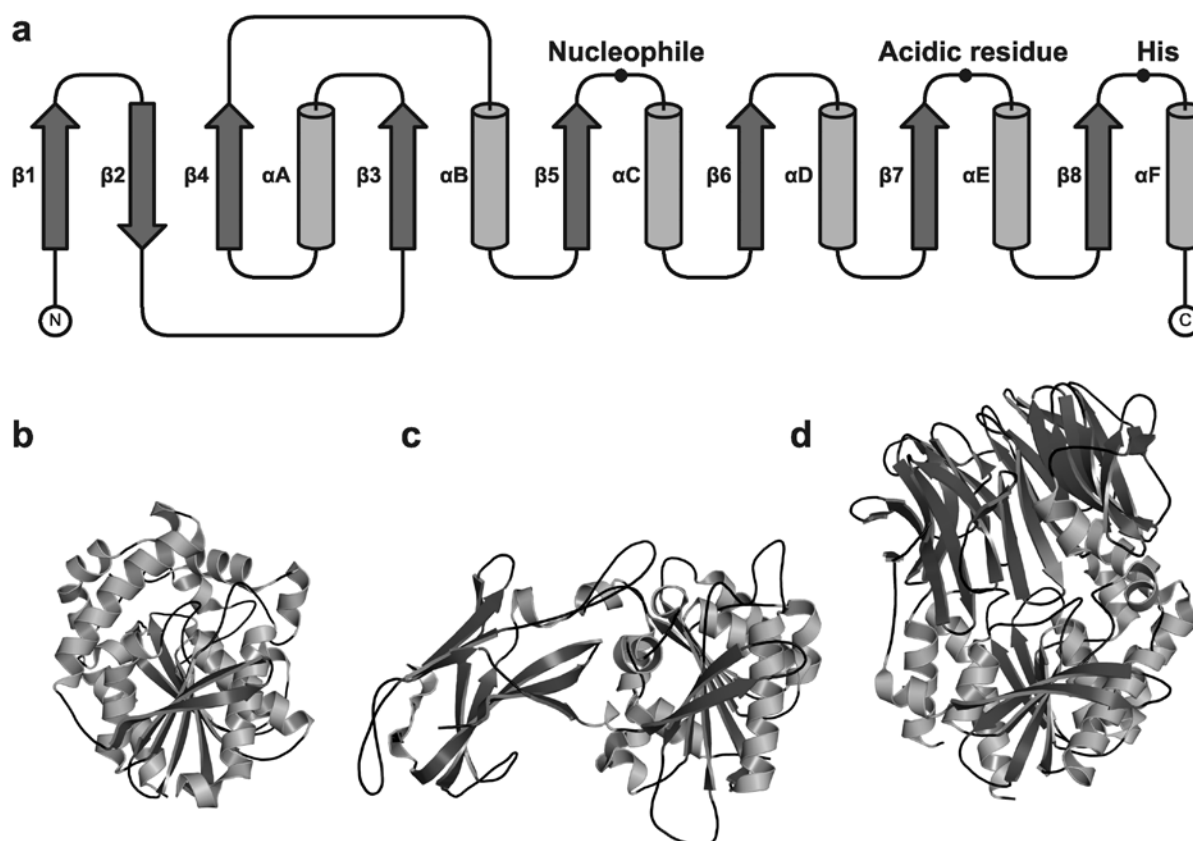
Figure 1.3 (in color on p.144): Canonical fold of α/β-hydrolases. In a) Topology diagram, with the strands indicated by red arrows and the helices by cyan cylinders. The positions of the catalytic residues are indicated. In b-d) the structures of three hyperthermophilic esterases: b) the carboxylesterase AFEST from *A. fulgidus* (pdb 1JJI), c) the esterase EstA from *T. maritima* (pdb 3DOH), and d) the acylpeptide hydrolase apAPH from *A. pernix* (pdb 1VE6).

EstE1 [64]. This strategy for thermostabilization is different from AFEST and shows that there are a variety of structural possibilities to acquire stability.

The crystal structure of an acylpeptide hydrolase (apAPH) from the archaeon *A. pernix* was solved to 2.1 Å resolution (PDB: 1EV6) [65]. Acylpeptide hydrolases are enzymes that catalyze the removal of an N-acetylated amino acid from blocked peptides. The enzyme shows an optimal temperature at 90°C for enzyme activity and is very stable at this temperature with a half-life of over 160 hours. It is active on a wide range of substrates, including *p*-nitroanilide (*p*NA) amino acids, peptides, and also *p*NP-esters with varying acyl chain lengths with an optimum for *p*NP-C6 [66]. The structure of the acylpeptide hydrolase / esterase apAPH belongs to the prolyl oligopeptidase family [65]. The structure is comprised of two domains, the N-terminal domain is a regular seven-bladed β-propeller and the C-terminal domain has the canonical α/β-hydrolase fold that contains the catalytic triad consisting of a serine, aspartate and histidine. It was shown that a single mutation (R526E), completely abolished the peptidase activity on Ac-Leu-*p*-nitroanilide of this enzyme while esterase activity on *p*NP-C8 was only halved [67]. Any mutation at the 526 site resulted in decreased peptidase activity due to loss of the ability of R526 to bind the peptidase substrate, while most of the mutants had increased esterase

activity due to a more hydrophobic environment of the active site. This result shows that enzymes can evolve such that they discriminate between substrates only by a single mutation.

The most recently elucidated structure belongs to an esterase, EstA, from *T. maritima* (PDB: 3DOH) [68]. The enzyme displayed optimal activity with short acyl chain esters at temperatures equal or higher than 95°C. Its structure was solved to 2.6 Å resolution and revealed a classical α/β-hydrolase domain, which contained the typical catalytic triad. Surprisingly, the structure also revealed the presence of an N-terminal immunoglobulin (Ig)-like domain. The combination of these two domains is unprecedented among both mesophilic and hyperthermophilic esterases. The function of this Ig-like domain was investigated and it was shown that it plays an important role in multimer formation, and in the stability and activity of EstA.

A high-resolution structure of an enzyme leads to a better understanding of its reaction mechanism, how it interacts with other proteins, what contributes to its stability, and may provide a basis for enzyme optimization and drug design. Because it is nowadays relatively easy to setup crystallization trials using commercially available screens and also because the current high-throughput crystallization projects are responsible for a large increase in the number of solved structures [69], it is expected that more structures of hyperthermophilic esterases will become available in future.

CLASSIFICATION

Enzymes can be classified on basis of their substrate preference, sequence homology and structural similarity. Classification of enzymes based on sequence alignments provides an indication of the evolutionary relationship between enzymes. Still, structural similarity is preserved much longer than sequence similarity during evolution. On the other hand, sequence homology and structural similarity are not always correlated with the substrate preference of an enzyme. All-in-all, classification of enzymes is not straightforward.

Several classifications of esterases and lipases into distinct families have been completed. In one such study, 53 bacterial esterases and lipases were classified into eight families based on their sequence similarity and some of their fundamental biological properties [110]. Many new esterases and lipases have since then been identified, including several, such as EstD from *T. maritima* [52], which could not be grouped into one of these eight families. Therefore, new families for these enzymes have been proposed. However, this early classification has provided a good basis for more refined classification of the esterases and lipases. Most of the recent studies are based on sequence and structural similarity and are accessible at online databases. Some relevant databases will be briefly discussed: The Lipase Engineering Database (LED), the Microbial Esterase and Lipase Database (MELDB), the Carbohydrate Active Enzyme (CAZy) database and the ESTHER database.

The LED (http://www.led.uni-stuttgart.de) combines information on sequence, structure and function of esterases, lipases and related proteins sharing the same α/β-hydrolase fold [111; 112]. The database contains more than 800 prokaryotic and eukaryotic sequences, which have been grouped into families based on multi-sequence alignments. The functionally relevant residues of each family have been annotated. The database was developed as a tool for protein engineering. The LED will be updated coming year (personal communication with Prof. dr. Juergen Pleiss). The classification will not change, but the number of proteins and families will increase substantially. MELDB (http://www.gem.re.kr/meldb) is a database that contains more than 800 microbial esterases and lipases [113]. The sequences in MELDB have been clustered into groups according to their sequence similarities and are divided in true esterase and lipase clusters. The database was developed in order to identify, conserved but yet unknown, functional domains/motifs and relate these patterns to the biochemical properties of the enzymes. According to the authors, new enzymes of other completely sequenced microbial strains will be added on a regular basis. CAZy (www.cazy.org) is a database that contains enzymes involved in the degradation, modification, or creation of glycosidic bonds [114]. One class of activities in this database is the carbohydrate esterases (CE). These enzymes remove ester-based modifications from carbohydrates. Carbohydrate esterases have been clustered into 15 families. These families have been created based on experimentally characterized proteins and sequence similarity. The database is continuously updated based on the available literature and structural information. The ESTHER database (http://bioweb.ensam.inra.fr/esther) contains more than 3500 sequences of enzymes belonging to the α/β-hydrolase fold [115]. These enzymes have been clustered into families based on sequence alignments. This database is updated regularly and furthermore contains information about the biochemical, pharmacological and structural properties of the enzymes.

2

Chapter

Characterization and structural modeling of a new type of thermostable esterase from *Thermotoga maritima*

ABSTRACT

A bioinformatic screening of the genome of the hyperthermophilic bacterium *Thermotoga maritima* for ester-hydrolyzing enzymes revealed a protein with typical esterase motifs, though annotated as a hypothetical protein. To confirm its putative esterase function the gene (*estD*) was cloned, functionally expressed in *E. coli* and purified to homogeneity. Recombinant EstD was found to exhibit significant esterase activity with a preference for short acyl chain esters (C_4-C_8). The monomeric enzyme has a molecular mass of 44.5 kDa and optimal activity around 95˚C and at pH 7. Its thermostability is relatively high with a half-life of 1 h at 100˚C, but less stable compared to some other hyperthermophilic esterases. A structural model was constructed with the carboxylesterase Est30 from *Geobacillus stearothermophilus* as template. The model covered most of the C-terminal part of EstD. The structure showed an α/β-hydrolase fold and indicated the presence of a typical catalytic triad consisting of a serine, aspartate and histidine, which was verified by site-directed mutagenesis and inhibition studies. Phylogenetic analysis showed that EstD is only distantly related to other esterases. A comparison of the active site pentapeptide motifs revealed that EstD should be grouped into a new family of esterases (Family 10). EstD is the first characterized member of this family.

INTRODUCTION

Enzymes play an important role in modern biotechnology because of their specificity, selectivity, efficiency and sustainability. One of the industrially most frequently used groups of biocatalysts are the esterases and lipases, which are exploited in various processes, such as the stereospecific hydrolysis of drugs and ester synthesis for food ingredients (flavors) [3; 4; 5; 7]. Esterases and lipases catalyse the hydrolysis of an ester bond resulting in the formation of an alcohol and a carboxylic acid. Both types of enzymes belong to the family of serine hydrolases and share structural and functional characteristics, including a catalytic triad, an α/b hydrolase fold and a cofactor independent activity. The catalytic triad usually consists of a nucleophilic serine in a GXSXG pentapeptide motif and an acidic residue (aspartic acid or glutamic acid) that is hydrogen bonded to a histidine residue [4; 7].

In the presence of water, esterases and lipases may be used for specific ester hydrolysis, but in anhydrous solvents the reverse reaction or a transesterification reaction becomes possible. The use of organic co-solvents, however, puts high constraints on the enzymes' stability, resulting in a growing demand for esterases with improved stability for industrial application. Enzymes from extremophiles and thermophiles in particular are promising in this respect because these enzymes have a high intrinsic thermal and chemical stability [18]. The hyperthermophilic archaea *Archaeoglobus fulgidus*, *Pyrococcus furiosus* and *Pyrobaculum calidifontis* have been shown to contain such thermostable esterases [38; 53; 59]. From the hyperthermophilic bacteria only few esterases have been described to date, viz. two acetyl xylan esterases from a *Thermoanaerobacterium* species [116], an esterase from *Thermoanaerobacter tengcongensis* [72] and recently a carboxylesterase from *Thermotoga maritima* [71].

Traditionally, active biocatalysts have been discovered by screening for the desired activity, but because of the availability of an ever increasing number of complete genome sequences, bioinformatics has become an important tool in the discovery and identification of novel industrial biocatalysts [16; 27]. In order to extract a maximal amount of information from the available genome sequences, conserved genes have been classified according to their homologous relationships, which resulted in the delineation of clusters of orthologous groups (COGs) [117; 118]. The purpose of the COG system is to facilitate the annotation of newly sequenced genomes and to functionally characterize individual proteins.

Here a bioinformatic analysis of the genome of the hyperthermophilic bacterium *Thermotoga maritima* was performed to find new thermostable esterases. Several open reading frames that potentially encode esterases or lipases were identified, including one (*estD*, TM0336) that has been annotated as a conserved hypothetical protein, although it does possess characteristics of an ester hydrolyzing enzyme. Interestingly, EstD belongs to a COG (1073) that comprises proteins only predicted to have an α/β hydrolase fold, whose function has not yet been experimentally determined. To confirm the anticipated function of EstD and to support COG1073 with experimental evidence, *estD* was cloned and expressed in *E. coli*.

The recombinant enzyme was characterized, including structural modeling and experimental analysis of the catalytic triad.

RESULTS

Identification and in silico analysis

T. maritima is a bacterium growing optimally at a temperature of 80°C. Its genome has been sequenced [119] and revealed 1,877 predicted coding regions, of which approximately 40% are still of unknown function. While performing BLAST searches with sequences of known esterases from other hyperthermophilic microorganisms against the *T. maritima* genome an amino acid sequence (accession code: TM0336) has been identified that had a pentapeptide consensus sequence, Gly-Xaa-Ser-Xaa-Gly, typical for serine hydrolases. The ORF was annotated as a conserved hypothetical protein [119]. The gene encodes a protein of 412 amino acids and has a calculated molecular mass of 46.5 kDa. BLAST-P analysis revealed the highest similarity to other hypothetical proteins and putative hydrolases. The most significant hits of a BLAST search analysis include a hypothetical protein of *Solibacter usitatus* (36% identity), a hypothetical protein of *Bacteroides fragilis* (33% identity) and putative hydrolases of several *Bacillus* species (up to 34% identity).

Analysis using Prosite InterProScan revealed a possible esterase domain (IPR000379) and lipase active site (IPR008262). A KEGG SSDB Motif Search showed that EstD is composed of two possible domains: an N-terminal domain (AA 17-121) which has homology to a MecA_N domain and a C-terminal domain (AA 150-400) which showed predicted domains for esterase or general hydrolase. The MecA gene is involved in bacterial resistance to antibiotics, however the N-terminal domain of MecA seems unlikely to have enzymatic activity and its role remains unknown [120]. The conserved domains present in the encoded protein were analyzed using the NCBI Conserved Domain Search. EstD belongs to the COG1073, comprising hydrolases of the alpha/beta superfamily. Furthermore, the C-terminal part of this protein is also related to COG1506 (Dipeptidyl aminopeptidases / acylaminoacyl-peptidases), COG1647 (Esterase/lipase) and COG2267 (Lyso-phospholipases), which are all subfamilies of the serine hydrolase family [121]. The characteristics of serine hydrolases include a tertiary structure called the α/β-hydrolase fold and a catalytic triad consisting of a serine, aspartate and histidine residue. A comparison of TM0336 with the amino acid sequences of the most significant hits in the BLAST search, as well as with the carboxylesterase Est2 of *Alicyclobacillus acidocaldarius* and the carboxylesterase Est30 of *Geobacillus stearothermophilus*, identified the three amino acids that potentially constitute the catalytic triad (Ser243, Asp347 and His378) (Figure 2.1).
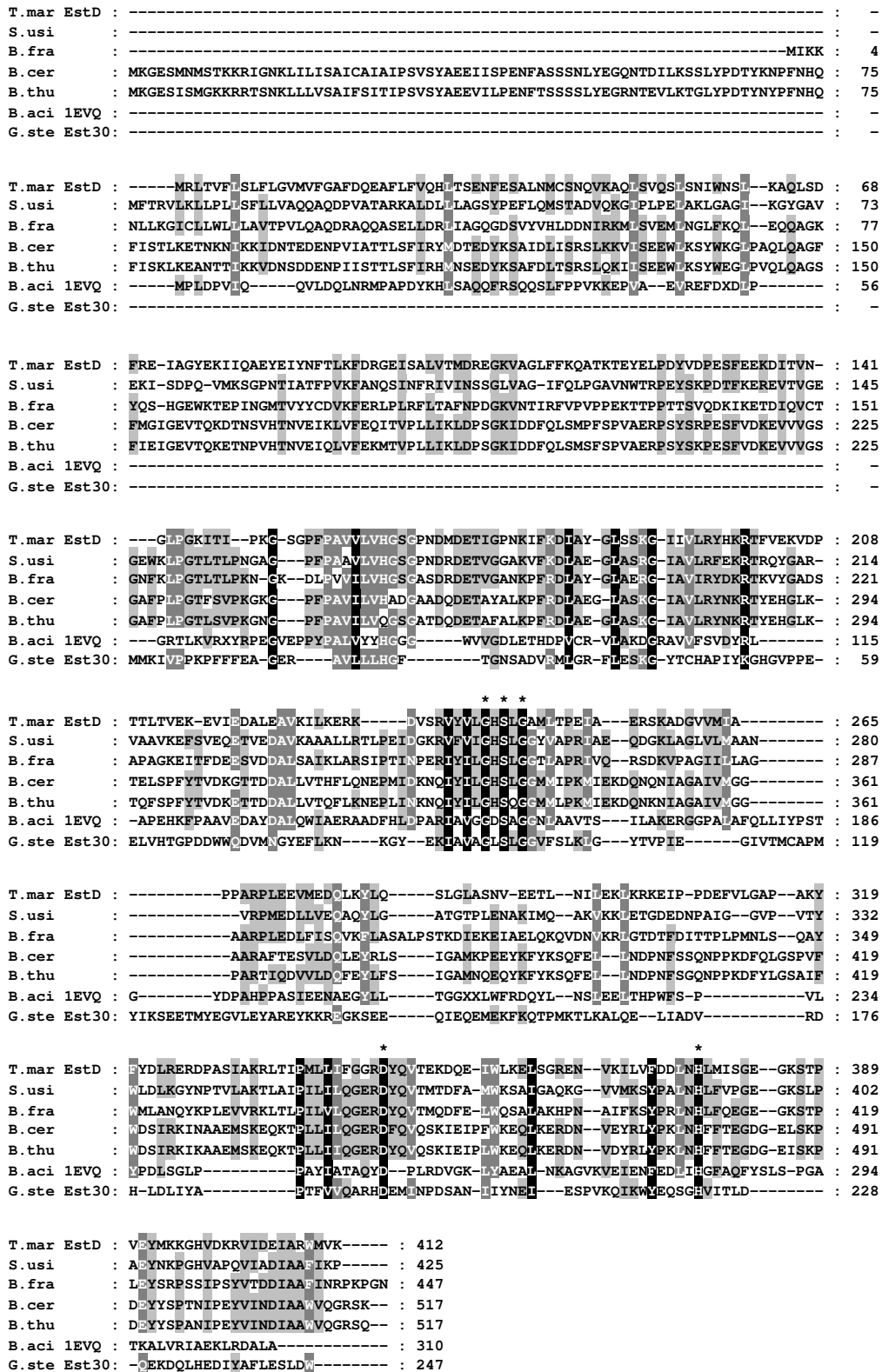
```
T.mar EstD : ----------------------------------------------------------------------- :   -
S.usi      : ----------------------------------------------------------------------- :   -
B.fra      : -------------------------------------------------------------------MIKK :   4
B.cer      : MKGESMNMSTKKRIGNKLILISAICAIAIPSVSYAEEIISPENFASSSNLYEGQNTDILKSSLYPDTYKNPFNHQ :  75
B.thu      : MKGESISMGKKRRTSNKLLLVSAIFSITIPSVSYAEEVILPENFTSSSSLYEGRNTEVLKTGLYPDTYNYPFNHQ :  75
B.aci 1EVQ : ----------------------------------------------------------------------- :   -
G.ste Est30: ----------------------------------------------------------------------- :   -


T.mar EstD : -----MRLTVFLSLFLGVMVFGAFDQEAFLFVQHLTSENFESALNMCSNQVKAQLSVQSLSNIWNST--KAQLSD :  68
S.usi      : MFTRVLKLLPLLSFLLVAQQAQDPVATARKALDLLAGSYPEFLQMSTADVQKGLPLPELAKLGAGL--KGYGAV :  73
B.fra      : NLLKGICLLWLLLAVTPVLQAQDRAQQASELLDRLIAGQGDSVYVHLDDNIRKMLSVEMLNGLFKQL--EQQAGK :  77
B.cer      : FISTLKETNKNLKKIDNTEDENPVIATTLSFIRYLDTEDYKSAIDLISRSLKKVLSEEWLKSYWKGLPAQLQAGF : 150
B.thu      : FISKLKEANTTLKKVDNSDDENPIISTTLSFIRHLNSEDYKSAFDLTSRSLQKILSEEWLKSYWEGLPVQLQAGS : 150
B.aci 1EVQ : -----MPLDPVLQ-----QVLDQLNRMPAPDYKHLSAQQFRSQQSLFPPVKKEPLA--ELREFDXDLP------- :  56
G.ste Est30: ----------------------------------------------------------------------- :   -


T.mar EstD : FRE-IAGYEKIIQAEYEIYNFTLKFDRGEISALVTMDREGKVAGLFFKQATKTEYELPDYVDPESFEEKDITVN- : 141
S.usi      : EKI-SDPQ-VMKSGPNTIATFPVKFANQSINFRIVINSSGLVAG-IFQLPGAVNWTRPEYSKPDTFKEREVTVGE : 145
B.fra      : YQS-HGEWKTEPINGMTVYYCDVKFERLPLRFLTAFNVPDGKVNTIRFVPVPPEKTTPPTTSVQDKIKETDIQVCT : 151
B.cer      : FMGIGEVTQKDTNSVHTNVEIKLVFEQITVPLLIKLDPSGKIDDFQLSMPFSPVAERPSYSRPESFVDKEVVVGS : 225
B.thu      : FIEIGEVTQKETNPVHTNVEIQLVFEKMTVPLLIKLDPSGKIDDFQLSMSFSPVAERPSYSKPESFVDKEVVVGS : 225
B.aci 1EVQ : ----------------------------------------------------------------------- :   -
G.ste Est30: ----------------------------------------------------------------------- :   -


T.mar EstD : ---GLPGKITI--PKG-SGPFPAVVLVHGSCPNDMDETIGPNKIFKDIAY-GLSSKG-IIVLRYHKRTFVEKVDP : 208
S.usi      : GEWKLPGTLTLPNGAG---PFPAAVLVHGSCPNDRDETVGGAKVFKDLAE-GLASRG-IAVLRFEKRTRQYGAR- : 214
B.fra      : GNFKLPGTLTLPKN-GK--DLLVVILVHGSCASDRDETVGANKPFRDIAY-GLAERG-IAVIRYDKRTKVYGADS : 221
B.cer      : GAFPLPGTFSVPKGKG---PFPAVILVHADKAADQDETAYALKPFRDLAEG-IASKG-IAVLRYNKRTYEHGLK- : 294
B.thu      : GAFPLPGTLSVPKGNG---PFPAVILVQGSCATDQDETAFALKPFRDIAE-GLASKG-IAVLRYNKRTYEHGLK- : 294
B.aci 1EVQ : ---GRTLKVRXYRPEGVEPPYPALVYYHGGC-----WVVGDLETHDPVCR-VLAKDGRAVLFSVDYRL------- : 115
G.ste Est30: MMKIVPPKPFFFEA-GER----AVLLLHGF---------TGNSADVRMLGR-FLESKG-YTCHAPIYKGHGVPPE- :  59

                                                             * * *
T.mar EstD : TTLTVEK-EVIEDALEAVKILKERK-----DVSRVYVLGHSLGAMLTPELA---ERSKADGVVMLA--------- : 265
S.usi      : VAAVKEFSVEQETVEDAVKAAALLRTLPEIDGKRVFVLGHSLGCYVAPRLAE--QDGKLAGLVLMAAN------- : 280
B.fra      : APAGKEITFDEESVDDALSAIKLARSIPTINPERLYILGHSLGGTLAPRLVQ--RSDKVPAGIILLAG------- : 287
B.cer      : TELSPFYTVDKGTTDDALLVTHFLQNEPMIDKNQLYILGHSLGGMLIPKMIEKDQNQNIAGAIVMGG-------- : 361
B.thu      : TQFSPFYTVDKETTDDALLVTQFLKNEPLIKNQLYILGHSLGGMLPKMIEKDQNKNIAGAIVLGG-------- : 361
B.aci 1EVQ : -APEHKFPAAVEDAYDALQWIAERAADFHLLPARLAVGCDSAGCNLAAVTS---ILAKERGGPALAFQLLIYPST : 186
G.ste Est30: ELVHTGPDDWWQDVMNGYEFLKN----KGY--EKLAVACLSLGGVFSLKLG---YTVPIE------GIVTMCAPM : 119


T.mar EstD : ----------PPARPLEEVMEDLLKLLQ-----SLGLASNV-EETL-NILEKLKRKEIP-PDEFVLGAP--AKY : 319
S.usi      : -----------VRPMEDLLVELAQLLG-----ATGTPLENAKIMQ--AKLKKLETGDEDNPAIG--GVP--VTY : 332
B.fra      : -----------AARPLEDLFISLVKELASALPSTKDIEKEIAELQKQVDNLKRLGTDTFDITTPLPMNLS--QAY : 349
B.cer      : -----------AARAFTESVLDQLELRLS----IGAMKPEEYKFYKSQFEL--NDPNFSSQNPPKDFQLGSPVF : 419
B.thu      : -----------PARTIQDVVLDQFELLFS----IGAMNQEQYKFYKSQFEL--NDPNFSGQNPPKDFYLGSAIF : 419
B.aci 1EVQ : G---------YDPAHPPASIEENAEGLLL-----TGGXXLWFRDQYL--NSLEELTHPWFS-P-----------VL : 234
G.ste Est30: YIKSEETMYEGVLEYAREYKKREGKSEE-----QIEQEMEKFKQTPMKTLKALQE--LIADV-----------RD : 176

                    *                                                      *
T.mar EstD : EYDLRERDPASIAKRLTIPMLLTFGGRDYQVTEKDQE-IWLKELSGREN--VKILVEDDLNHLMISGE--GKSTP : 389
S.usi      : WLDLKGYNPTVLAKTLAIPILILQGERDYQVTMTDFA--MWKSALGAQKG--VVMKSYPALNHLFVPGE--GKSLP : 402
B.fra      : WMLANQYKPLEVVRKLTLPILVLQGERDYQVTMQDFE--LWQSALAKHPN--AIFKSYPRLNHLFQEGE--GKSTP : 419
B.cer      : WDSIRKINAAEMSKEQKTPLLILQGERDFQVQSKIEIPFWKEQLKERDN--VEYRLYPKLNHFFTEGDG-ELSKP : 491
B.thu      : WDSIRKIKAAEMSKEQKTPLLILQGERDYQVQSKIEIPLWKEQLKERDN--VDYRLYPKLNHFFTEGDG-EISKP : 491
B.aci 1EVQ : MPDLSGLP----------PAYLATAQYD--PLRDVGK-LYAEALNKAGVKVEIENEDLIHGFAQFYSLS-PGA : 294
G.ste Est30: H-LDLIYA----------PTFVVQARHDEMLNPDSAN-LIYNEL---ESPVKQIKWYEQSGHVITLD-------- : 228


T.mar EstD : VEYMKKGHVDKRVIDEIARWMVK----- : 412
S.usi      : ADYNKPGHVAPQVIADIAAFIKP----- : 425
B.fra      : LEYSRPSSIPSYVTDDIAAFINRPKPGN : 447
B.cer      : DEYYSPTNIPEYVINDIAAWVQGRSK-- : 517
B.thu      : DEYYSPANIPEYVINDIAAWVQGRSQ-- : 517
B.aci 1EVQ : TKALVRIAEKLRDALA------------ : 310
G.ste Est30: -QEKDQLHEDIYAFLESLDW-------- : 247
```

Figure 2.1: Amino acid sequence alignment of EstD. Multiple sequence alignment of EstD (T.mar EstD) with the first BLAST hits: *Solibacter usitatus* (S.usi), *Bacteroides fragilis* (B.fra), *Bacillus cereus* (B.cer) and *Bacillus thuringiensis* (B.thu) and two characterized esterases: *Bacillus acidocaldarius* (B.aci 1EVQ) and *Geobacillus stearothermophilus* (G.ste Est30). The conserved GXSXG motif, and the aspartate and the histidine that constitute the catalytic triad are marked with an asterisk.

Cloning and purification of recombinant EstD

N-terminal sequence analysis using SignalP revealed that the first 18 amino acids form a signal peptide. The predicted mature gene was cloned into the expression vector pET-26b. The enzyme EstD was purified to homogeneity from heat-treated cell extracts of *E. coli* BL21(DE3)/pSJS1244/pWUR353 by immobilized metal affinity chromatography. The recombinant protein was purified 115-fold with a yield of 66%. Homogeneity of the protein was checked by SDS-PAGE and confirmed a molecular subunit mass of 44.5 kDa (mature enzyme) (Figure 2.2a). Activity staining of the SDS-PAGE gels using α-naphtyl acetate confirmed the identity of the EstD band (Figure 2.2b). Native-PAGE showed a single band that was confirmed to possess esterase activity by means of an activity stain. Size exclusion chromatography showed that the enzyme existed mainly as a monomer and to some extent as dimer, with estimated masses of 48 kDa and 93 kDa, respectively.
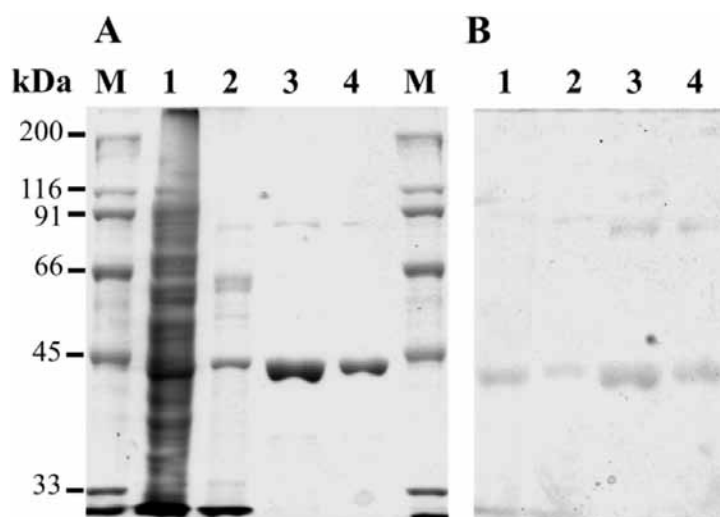
Figure 2.2: SDS-PAGE of EstD fractions. Samples were separated by SDS-PAGE in duplicate. One gel was stained with Coomassie brilliant blue (A) and the other was stained for activity using α-naphtyl acetate after renaturation (B). M: molecular weight standards, lane 1: cell free extract, lane 2: heat-stable cell free extract, lane 3: EstD after immobilized metal affinity chromatography and lane 4: purified EstD. A second band at approximately 90 kDa is corresponding to the EstD dimer. The dimer is believed to be catalytically active as well.

Substrate specificity and kinetics

The substrate specificity of purified EstD was analyzed using *p*-nitrophenyl esters. The highest specific activity with EstD was found towards short chain *p*-nitrophenyl esters of butyrate (C4) and valerate (C5). Little activity was found towards long chain *p*-nitrophenyl esters of decanoate (C10) to myristate (C14). In general, activity of the enzyme on shorter (≤C10) and longer fatty acids (≥C10) is referred to as esterase activity and lipase activity, respectively [7]. The kinetic properties of EstD were determined for *p*-nitrophenyl esters of acetate (C2), butyrate (C4), valerate (C5), octanoate (C8) and decanoate (C10) (Table 2.1). The catalytic efficiency

represented by the value of k_{cat}/K_m indicated that *p*-nitrophenyl valerate and *p*-nitrophenyl octanoate were the best substrates for EstD among the *p*-nitrophenyl esters tested. Hence, on the basis of its substrate profile EstD should be classified as an esterase.

Neither proteolytic activity using casein as substrate, nor peptidase activity when assayed with L-leucine *p*-nitroanilide and L-proline *p*-nitroanilide was detected.

Table 2.1: Kinetic parameters for hydrolysis of various *p*-nitrophenyl esters. Kinetic assays were performed in 50 mM citrate-phosphate buffer pH 7 at 70°C.

p-Nitrophenyl esters	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})
Acetate (C2)	0.148 ± 0.025	1.0 ± 0.05	6.8 ± 1.2
Butyrate (C4)	0.227 ± 0.017	14.9 ± 0.40	65.6 ± 5.2
Valerate (C5)	0.066 ± 0.006	10.2 ± 0.20	154.5 ± 14.4
Octanoate (C8)	0.011 ± 0.003	1.6 ± 0.15	145.5 ± 12.1
Decanoate (C10)	0.072 ± 0.012	1.3 ± 0.06	18.1 ± 0.5

Effect of temperature and pH on enzyme activity and thermal stability

The effect of temperature on EstD activity was studied using *p*-nitrophenyl valerate as a substrate. The esterase activity increased from 45°C upwards until 95°C (Figure 2.3). An Arrhenius analysis resulted in a linear plot in the temperature range of 45 to 85°C (inset in Figure 2.3), with a calculated activation energy for the formation of the enzyme/substrate complex of 15 kJ/mol. EstD has a high resistance to thermal inactivation, with a half-life value of approximately 1 h at 100°C. To determine the optimal pH for the esterase, the activity of EstD was measured in a pH range of 5 to 9. EstD displayed >70% of its maximal activity in the pH range of 5 to 9, with an optimal pH at approximately 7.0 (not shown).
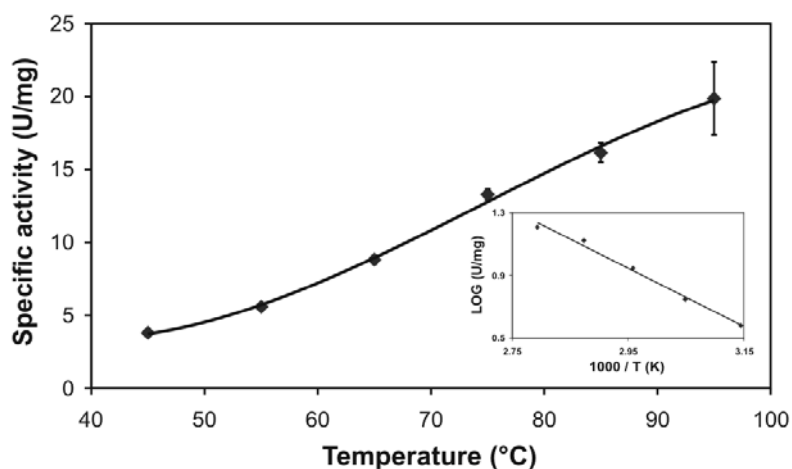
Figure 2.3: Effect of temperature on esterase activity. The effect of temperature on esterase activity was studied using *p*NP-valerate as a substrate at temperatures ranging from 45°C to 95°C. The inset shows the temperature dependence as an Arrhenius plot.

Table 2.2: Effect of inhibitors on EstD activity.

Inhibitors	Relative activity (%)
None	100
PMSF	4
DPC	53
N-ethylmaleimide	83
HgCl$_2$	3
EDTA	97
DTT	99
β-Mercaptoethanol	97

Effect of metals, detergents, solvents and inhibitors

The effect of metal ions on EstD activity was tested using various metal ions: Ca^{2+}, Ni^{2+}, Co^{2+}, Cu^{2+}, Fe^{2+}, Zn^{2+}, Mn^{2+} and Mg^{2+} at concentrations of 1 mM. No significant stimulation or reduction of activity of EstD was observed. The effect of inhibitors on EstD activity is shown in Table 2.2. PMSF, a serine protease inhibitor, strongly inhibited enzyme activity. DPC, a histidine modifier, also inhibited the reaction, albeit less pronounced than PMSF. This indicates that serine as well as histidine residues are important for EstD activity. Activity was also strongly inhibited by mercury chloride and to some extent by N-ethylmaleimide. In contrast, DTT did not affect enzyme activity and neither did EDTA, which agrees with the metal tests.

The effect of detergents and solvents on EstD activity was tested in the standard assay with final concentrations of either 1% or 10% (v/v) (Table 2.3). In the detergents test, activity was decreased by more than 50% when 1% Tween 20 was present and was completely inhibited by 1% SDS. Addition of the organic solvents methanol, ethanol and iso-propanol resulted in a decrease in activity ranging from more than 70% to less than 20% residual activity, respectively. On the other hand, addition of glycerol in the assay did not seem to have an effect on activity.

Table 2.3: Effect of detergents and solvents on EstD activity.

Detergents and solvents	Concentration (v/v %)	Relative activity (%)
Control	---	100
Methanol	10	73
Ethanol	10	37
2-Propanol	10	18
Glycerol	1	87
	10	96
DMSO	10	84
Tween 20	1	43
SDS	1	0

Structural modeling

In the absence of a three-dimensional structure of EstD, it was decided to build a 3D-model of EstD. Since there are no close structural homologs of EstD, modeling was based on threading. A model of EstD was made using the 3D-structural threading program PHYRE [122]. A threading algorithm seeks a template protein in a database that structurally fits well to a query sequence. Unlike homology modeling, a certain sequence similarity between the query sequence and a template protein is not necessary. Several structural fits were found. The thermophilic carboxylesterase Est30 of *Geobacillus stearothermophilus* (PDB code 1TQH) [123] was used to build the model of EstD. Est30 consists of 247 amino acid residues and the crystal structure showed a large domain with a modified alpha/beta hydrolase core including a seven-, rather than an eight-stranded beta sheet, and a smaller domain comprising three alpha helices. Like EstD, Est30 has a preference for short acyl chain substrates, with an optimum for C4-C8. The main difference between Est30 and EstD is their amino acid sequence length. The final model for EstD covered the C-terminal domain of EstD (amino acid residues 150 to 412). The schematic structural model consists of six α-helixes and has one central β-sheet made up of six β-strands (Figure 2.4a). The first and second β-strand of the α/β-hydrolase fold have not been modeled.

The quality of the model towards stereochemistry and geometry was analyzed by PROCHECK analysis [124]. The Ramachandran plot (not shown) indicated that most (92%) of the residues are in the core and allowed regions. Bond lengths, bond angles and torsion angles were evaluated with the What If program [125] and were considered good (a RMS z-score for a normally restrained data set is expected to be around 1.0). Bond lengths were found to deviate slightly less than normal from the mean standard bond length (a RMS z-score of 0.7). Bond angles and torsion angles were found to deviate normally (RMS z-scores around 1.0).

A first secondary structural alignment indicated the residues Ser243, Asp347 and His378 as the probable catalytic triad. In the obtained model, Ser243, Asp347 and His378 were indeed located in close proximity, most likely representing the actual active site. Ser243 is located within a nucleophile elbow connecting strand β5 and helix α3, while Asp347 and His378 are located on loops between β7-α7 and β8-α8, respectively (Figure 2.4a).

In the crystal structure of Est30, a covalently bound ligand is present. This ligand, propylacetate, was modeled into the active site of the EstD model. The ligand is covalently bound to the side-chain of Ser243, His378 acts as proton carrier and Asp347 is the charge relay network. The ligand is stabilized by hydrogen bond interactions with the amides of Leu244 and Gly164, which likely form the oxyanion hole (Figure 2.4b). The putative substrate binding pocket extends in a cleft on both sides of Ser243. The alcohol side of the substrate is in a groove pointed towards the entrance of the pocket and extends approximately 10 Å from Ser243. The acyl-side of the ligand fits in a less exposed pocket of approximately 6 Å wide and 9 Å long, consistent with the observed activity on substrates with acyl chain length C2-C12. The
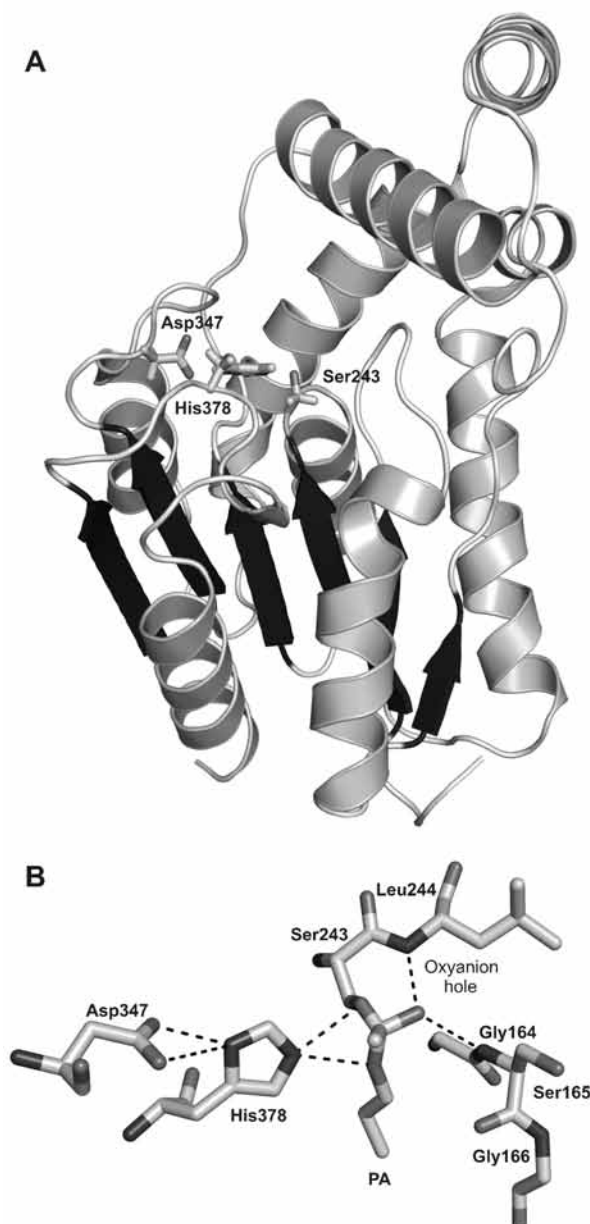
Figure 2.4: 3D model of EstD. (A) The overall structure of the C-terminal domain of EstD. The central β-sheet and surrounding α-helixes are shown in black and grey, respectively. Residues of the catalytic triad are indicated. (B) The active site region of the EstD model with bound ligand. Interatomic interactions are shown in dashed lines. The ligand, propylacetate (PA), is covalently bound to Ser243. The NH groups of Leu244 and Gly164 most likely form the oxyanion hole.

hydrophobic side-chains in this pocket are Met247, Ala265, Pro267, Ala268, Pro270, Leu271, Leu279, Phe320 and Val350. One polar residue Gln349 is located at the edges of the pocket and might have a role in substrate recognition. Gln349 and adjacent residues are well conserved in the closest homologues (supplementary material), suggesting an important structural role. The EstD substrate binding pocket is very similar to that of Est30 and structurally related esterases. This comprises an open accessible binding cleft and a relatively large cap domain, consisting of one small and two large helices on the N-terminal side of the central β-sheet. This structural similarity between EstD and Est30 corresponds with their very similar substrate preference.

To confirm the predictions of the catalytic triad, these residues were substituted by site-directed mutagenesis. The mutants Ser243Ala, Asp347Asn and His378Asn were expressed and purified using heat-treatment. The enzymes remained stable during heat-treatment. However, no activity was observed with the mutants, confirming the importance of these three residues for the activity of EstD.

DISCUSSION

In this contribution the cloning, expression, and characterization of a new type of esterase from the hyperthermophilic bacterium *T. maritima* is described. The encoding gene (*estD*) was originally annotated as a hypothetical protein, but a more detailed sequence analysis revealed the presence of an α/β hydrolase fold and a nucleophilic serine in a pentapeptide motif, suggesting a possible role in ester hydrolysis. After functional expression in *E. coli* the esterase activity could indeed be confirmed. When EstD was assayed with *p*-nitrophenyl esters it showed a preference for substrates with shorter chain lengths, indicating it should be classified as an esterase and not as a lipase. Highest activity was seen on esters of butyrate and valerate, which is comparable to esterases from other hyperthermophiles, *viz*. *T. tengcongensis* [72], *Sulfolobus solfotaricus* [39], *Sulfolobus shibatae* [100] and *Sulfolobus tokodaii* [108]. The determined k_{cat} values of EstD, however, were found to be 100-1000-fold lower when compared to the hyperthermophilic esterases. The K_m, on the other hand, was relatively low. The low k_{cat} may indicate that the artificial *p*-nitrophenyl substrates differ substantially from the enzyme's natural substrate. However, the physiological function of EstD is not known, as is the case for most described esterases.

As to be expected for a hyperthermophilic enzyme, EstD showed a temperature optimum around 95˚C, which is comparable to that of the *P. furiosus* esterase [53] and the *P. calidifontis* esterase [38]. The Arrhenius plot for EstD was linear at temperatures ranging from 45-85°C, indicating that the conformation of EstD does not change throughout this temperature range. The enzyme was very stable at high temperatures, with a half-life of approximately one hour at 100°C. EstD is less stable than the esterase from *P. furiosus* (half-life value of 34 h at 100˚C [53]), but substantially more stable than the esterase from *T. tengcongensis* (half-life value of 15 min at 80˚C [72]) or the esterase from *T. maritima* (half-life value of 30 min at 80°C [71]), which makes EstD the most stable bacterial esterase to date. EstD exhibited activity in the presence of 10% organic solvents which is comparable to the activity of the *P. calidifontis* carboxylesterase [38]. The high thermal stability and activity in the presence of organic solvents makes EstD an attractive catalyst for future applications in industry.

To gain more knowledge on the presence of essential catalytic or structural amino acids, EstD activity was tested upon incubation with various chemicals. The inhibition by PMSF and DPC indicated that serine and histidine residues might be involved at the catalytic site of the

enzyme, in agreement with the anticipated catalytic triad. Different metals and EDTA did not inhibit activity indicating that there is no requirement for divalent cations. The inhibition by HgCl$_2$ and N-ethylmaleimide suggests that the only free thiol group which is present (Cys42), is important for the correct functioning of the enzyme. The presence of a single thiol makes oxidation to a disulfide not possible, which is confirmed by the observation that neither DTT nor β-mercaptoethanol enhanced the activity of the enzyme. The single cysteine is not included in the EstD model; however it may be close to active site residues and as such can influence activity when modified with chemicals. Altogether, the inhibition pattern is similar to that described for the esterases from *P. calidifontis* [38], *S. solfataricus* [39] and *T. maritima* [71].

Based on the alignment and the site-directed mutagenesis experiments EstD was shown to contain the typical catalytic triad, consisting of a serine in a GXSXG pentapeptide, an acidic aspartate, and a histidine residue. The structural modeling was expected to be difficult due to the lack of 3-D structures of homologous esterases. Despite the very low sequence identity (16% identity over the C-terminal part); EstD could be modeled using Est30 from *G. stearothermophilis* as a template. However, modeling was only possible with the C-terminal domain of EstD, which also contains all the active site residues. The N-terminal domain of EstD has similarity to the MecA N-terminus but could not be modeled. The function of the N-terminus remains unclear. It might be involved in selection of the substrates, either by binding of the substrate or by narrowing the entrance to the active site.

The low sequence homology of EstD to characterized proteins was the reason that it was initially annotated as a hypothetical protein. Nevertheless, the results described here show that EstD has esterase activity and also exhibits the typical structural features of this type of enzyme. Bacterial esterases and lipases have been classified into eight families based on a comparison of their amino acid sequences and some fundamental biological properties [110]. Enzymes in Family 1 are called true lipases and are further classified into six subfamilies.

```
Consensus esterase/lipase          GxSxG

T. maritima           (EstD)    VLGHSLGA        NP_228147
B. fragilis                     ILGHSLGG        BAD47626
B. fragilis                     ILGHSLGG        CAH06543
S. usitatus                     VIGHSLGG        ZP_00519815
B. cereus                       ILGHSLGA        NP_978269
B. thuringensis                 ILGHSLGA        ZP_00739952
C. phytofermentans              VLGHSLGG        ZP_01352053

Consensus                       LGHSLG

Consensus family 1              GHSQG
Consensus family 2              GDSL
Consensus family 3              GxSMGGGG
Consensus family 4              GDSAGG
Consensus family 5              GxSMGG
Consensus family 6              GFSQG

P. lemoignei          (PhaZ7)   AHSMGV          AAK07742
G. stearothermophilus (Est30)   GLSLGG          AAN81911
```

Figure 2.5: Alignment of the esterase/lipase pentapeptide motif of EstD with related enzymes and consensus sequences. Consensus sequences of the different lipase and esterase families 110 and the two enzymes discussed in the text, PhaZ7 126 and Est30 123, are indicated.

Enzymes belonging to Family 2 - 8 are esterases. However, a homology search with the EstD sequence against public databases revealed the highest similarity to hypothetical proteins and putative hydrolases that are not grouped in any of the eight families. Moreover, EstD showed no sequence identity to any of the members of the previously classified families of microbial lipases and esterases. A phylogenetic analysis showed that EstD is indeed grouped into a new separate family (data not shown), which also includes enzymes from several *Bacillus* species, *B. fragilis* and *S. usitatus*. This divergence from the current families can be viewed best by aligning the pentapeptide consensus sequences (Figure 2.5). EstD and related sequences show a high pentapeptide homology (GHSLG), which is different from the consensus of the esterase families. These data suggest that EstD is a member of a new family of esterases, designated as Family 10. EstD is the third esterase that cannot be grouped into one of the eight families. Because of absence of significant amino acid homology, Handrick *et al.* [126], suggested that PhaZ7 of *Paucimonas lemoignei* should be classified into a new family of esterases (Family 9: extracellular PHA depolymerases) and also Liu *et al.* [123], suggested that Est30 of *G. stearothermophilus* represents a new family of carboxylesterases (Figure 2.5). EstD is the first characterized member of the proposed new family, and as such also the first characterized enzyme of COG1073, which will contribute to a better understanding of the function of the other enzymes in this COG.

MATERIALS AND METHODS

Chemicals
All chemicals were purchased from Sigma-Aldrich (St Louis, MO, USA) or Acros Organics (Geel, Belgium). The restriction enzymes were obtained from Invitrogen (Carlsbad, CA, USA). *Pfu* Turbo and T4 DNA ligase were purchased from Invitrogen and Stratagene (La Jolla, CA, USA), respectively.

Strains and plasmids
The vector pGEM-t-easy (Promega, Madison, WI, USA) was used for the cloning of PCR products. For heterologous expression, the vector pET-26b (Kanamycin-resistant; Novagen, San Diego, CA, USA) and the tRNA helper plasmid pSJS1244 (Spectinomycin-resistant / [127; 128]) were used. *Escherichia coli* strain XL1-Blue (Stratagene) was used as a host for cloning. *E. coli* strain BL21(DE3) (Novagen) was used as an expression host. Both strains were grown under standard conditions [129] following the instructions of the manufacturers.

Data mining
The genome of *T. maritima* MSB8 [119] was screened for possible esterases and lipases. Sequences coding for esterases and lipases were identified by performing BLAST searches

with sequences from characterized esterases / lipases (http://www.ncbi.nlm.nih.gov/blast/) [28] and Motif (http://www.expasy.org/prosite/) searches. The conserved domains were analyzed with CD-SEARCH (http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) [30] and KEGG SSDB Motif Search (http://www.genome.jp/kegg/ssdb/) [130]. The N-terminal sequence analysis of the translational product of TM0336 was done using the SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP/) [131]. Phylogenetic analysis was performed by aligning EstD, close homologues and sequences of the esterase and lipase families using the Tcoffee server (http://www.igs.cnrs-mrs.fr/Tcoffee/tcoffee_cgi/index.cgi) [132]. The alignment was further corrected by hand. A bootstrapped phylogenetic tree was constructed and displayed using the neighbor-joining method with TreeView version 1.6.5 [133]. A three dimensional structure of EstD was modeled using the PHYRE Protein Fold Recognition Server (http://www.sbg.bio.ic.ac.uk/phyre/) [122]. The model was evaluated for stereochemical quality using the programs PROCHECK (http://www.biochem.ucl.ac.uk/~roman/procheck/procheck.html) [124] and What If (http://swift.cmbi.kun.nl/WIWWWI/) [125]. PyMol was used to analyze and visualize the structure [134].

Cloning and expression

The gene TM0336 (Genbank accession number NP_228147) was PCR-amplified, without the sequence encoding its signal peptide (the first 18 amino acids) and its stop codon using chromosomal DNA of *T. maritima* as a template and the following two primers: 5'-GCGGCGCCATATGGATCAGGAAGCGTTTCTC-3' (sense, underlined *Nde*I restriction site) and 5'-GCGCGCTCGAGTTTTACCATCCACCTGGC-3' (anti-sense, underlined *Xho*I restriction site). The PCR product generated was modified using the A-tailing procedure [135] and ligated into the pGEM-t-easy vector. *E. coli* XL1-blue was transformed with this construct (pWUR349). The recombinant plasmid was digested by *Nde*I and *Xho*I and the product was purified and inserted into pET-26b digested with the same restriction enzymes. The construct was designed with a hexahistidine-tag engineered at the C-terminus of the enzyme to facilitate purification. Subsequently, *E. coli* BL21(DE3), harboring the tRNA helper plasmid pSJS1244, was transformed with the resulting plasmid (pWUR353). The sequence of the expression clone was confirmed by sequence analysis of both DNA strands.

Mutagenesis

Mutants of EstD were created in order to confirm the identity of the active site residues. Mutants Ser243Ala, Asp347Asn and His378Asn were generated using Quickchange (Stratagene) site directed mutagenesis with the following primers 5'- GTGCTGG GACACGCCCTCGGTGCGATGC-3' and 5'- GCATCGCACCGAGGGCGTGTCCCAGCAC-3', 5'- GATCTTCGGCGGCAGAAACTACCAGGTGACTG-3' and 5'- CAGT CACCTGGTAGTTACTGCCGCCGAAG-3', 5'- CGACGATCTCAATAACTTGATGATTT CAGG-3' and 5'- CTCCTGAAATCATCAAGTTATTGAGATCGTCG-3', respectively (the underlining indicates the modified codon). Mutations were confirmed by sequence analysis of both DNA strands.

Production and purification

E. coli BL21(DE3)/pSJS1244 was transformed with pWUR353. A single colony was used to inoculate 4 mL Luria-Bertani medium containing kanamycin and spectinomycin (both 50 µg · mL^{-1}) and incubated overnight at 37°C while shaking. Next, the preculture was used to inoculate (1:1000) two times 1500 mL Luria-Bertani medium containing kanamycin and spectinomycin (both 50 µg · mL^{-1}) in 2-L conical flasks and incubated in a rotary shaker at 37°C for 8 h. The culture was then induced by adding isopropyl-ß-D-thiogalactopyranoside (IPTG) to a final concentration of 0.1 mM. The culture was further incubated at 37°C for another 16 h. Cells were harvested by centrifugation at 10000 *g* and 4°C for 15 min. The cell pellet was resuspended in 25 mL lysis buffer (50 mM Tris-HCl buffer (pH 7.8), 300 mM NaCl, 10 mM imidazole), and passed twice through a French press at 110 MPa. The crude cell extract was DNase treated for 30 min at room temperature to become less viscous. The extract was centrifuged at 43000 *g* and 4°C for 25 min. 20 mL lysis buffer was added to the resulting supernatant (cell free extract) and heated for 25 min at 70°C and subsequently centrifuged at 43000 *g* and 4°C for 25 min. The supernatant (heat-stable cell free extract) was filtered (0.45 µm) and applied at a flow rate of 2 mL · min^{-1} to a Ni-chelating column (20 mL) equilibrated in 50 mM Tris-HCl buffer (pH 7.8) containing 300 mM NaCl. The column was washed with 20 mM imidazole in the same buffer and subsequently proteins were eluted with a linear gradient of 20-500 mM imidazole and fractions (2 mL) were collected. The most active fractions were pooled and applied at a flow rate of 10 mL · min^{-1} to a Hi-prep desalting column (53 mL), equilibrated in 50 mM Tris-HCl buffer (pH 7.8) containing 150 mM NaCl in order to remove imidazole. Fractions of 5 mL were collected.

Size exclusion chromatography

The molecular mass of the purified enzyme was determined by size exclusion chromatography on a Superdex 200 high-resolution 10/30 column (24 mL) (Amersham Biosciences, Piscataway, NJ, USA) equilibrated in 50 mM Tris-HCl (pH 7.8) containing 100 mM NaCl. Two hundred microliters of enzyme solution in 50 mM Tris-HCl and 150 mM NaCl (pH 7.8) buffer was loaded at a flow rate of 0.7 mL · min^{-1} onto the column and fractions (0.5 mL) were collected. Proteins used for calibration were blue dextran 2000 (>2,000 kDa), ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa).

SDS-PAGE, Native PAGE and activity staining

SDS-PAGE was performed with gels containing 10% acrylamide using a MiniProtean III system (BioRad, Hercules, CA, USA). Samples containing loading buffer (0.1 M sodium phosphate buffer, 4% SDS, 10% 2-mercaptoethanol, 20% glycerol, pH 6.8), were prepared by heating for 10 min at 100°C. Gels were stained with Coomassie brilliant blue. The molecular mass was estimated using the BioRad broad range protein marker. Native PAGE was performed with gels

containing 6% acrylamide. Native PAGE and SDS-PAGE gels were stained for esterase activity by a modified version of the staining technique of Sobek [32]. A renaturation procedure was carried out after SDS-PAGE by incubating the gel two times for 15 min in 50 mM Tris-HCl (pH 7.8) / iso-propanol (4:1, v/v%), subsequently rinsed three times for 15 min in 50 mM Tris-HCl (pH 7.8) and then rinsed again with water. The gel was stained at 37°C in the dark by incubating it in a 100 mL solution of 50 mM Tris-HCl (pH 7.8) buffer containing 50 mg of Fast Blue BB plus and 1 mL of acetone solution containing 10 mg of α-naphtyl acetate. When esterase active bands began to color deep brown, the reactions were stopped by rinsing the gel with tap water, followed by fixation in 3% (v/v) acetic acid.

Enzyme assays

Esterase activity was determined by measuring the amount of *p*-nitrophenol released during enzymatic hydrolysis of different *p*-nitrophenyl esters. The release of *p*-nitrophenol was continuously monitored at 405 nm using a Hitachi UV2001 spectrophotometer with a temperature controlled cuvette holder. Unless otherwise indicated, in a standard assay, esterase activity was measured with 0.2 mM *p*-nitrophenyl valerate (*p*NP-C5) as a substrate in 50 mM citrate-phosphate buffer (pH 7) containing 1% isopropanol at 70°C. Stock solutions of *p*-nitrophenyl esters were prepared by dissolving substrates in iso-propanol. After pre-incubation, the reaction was started by adding enzyme to the reaction mix. One unit of esterase activity was defined as the amount of protein releasing 1 μmol of *p*-nitrophenol from pNP-C5 per minute. Measurements were corrected for background hydrolysis in the absence of enzyme. Measurements were carried out at least three times and the molar extinction coefficient of *p*-nitrophenol was determined for every condition prior to each measurement. Activity was determined from the initial rate of the hydrolysis reaction. The protein concentration was measured at 280 nm using a NanoDrop ND-1000 Spectrophotometer (NanoDrop, Wilmington, DE, USA).

Peptidase activity was assayed with 0.2 mM L-leucine *p*-nitroanilide and L-proline *p*-nitroanilide as substrates in a standard assay as described above.

The proteolytic activity of EstD was assayed using 1% (w/v) casein in 50 mM Tris-HCl (pH 8). Casein hydrolysis assays were performed for up to 1 h at 70°C. The reaction was terminated with 10% (v/v) trichloroacetic acid and incubated on ice for 30 min. The absorbance of the centrifuged supernatant was measured at 280 nm. A blank without esterase was incubated under the same conditions.

Acyl chain length preference

Substrate specificity of the enzyme towards the acyl chain length of different *p*-nitrophenyl esters was investigated by using *p*-nitrophenyl acetate, *p*-nitrophenyl butyrate, *p*-nitrophenyl valerate, *p*-nitrophenyl octanoate, *p*-nitrophenyl decanoate, *p*-nitrophenyl dodecanoate, and *p*-nitrophenyl myristate in the standard assay.

pH and temperature optimum

The effect of pH on esterase activity was studied by measuring activities on *p*-nitrophenyl valerate for a pH range of 4.0 to 9.5. The buffers used were 50 mM citrate-phosphate (pH 4.0 to 8.0) and 50 mM CAPS buffer (pH 9.5). The effect of temperature on esterase activity was studied in the range of 45°C to 95°C using 1 mM *p*-nitrophenyl valerate in the standard assay. 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 [136]).

Thermostability

Enzyme thermostability was determined by incubating the enzyme in a 50 mM Tris-HCl, 150 mM NaCl (pH 7.8) buffer at 100°C for various time intervals. Residual activity was assayed under the standard condition.

Inhibition studies

The effect of metal ions on esterases activity was determined using different metal salts (CaCl$_2$, NiCl$_2$, CuCl$_2$, MnCl$_2$, MgCl$_2$, FeSO$_4$ and ZnSO$_4$) at final concentrations of 1 mM using the standard activity assay. The activity of EstD without addition of metal ions was defined as 100%. The effect of inhibitors on esterase activity was determined using ethylenediaminetetraacetic acid (EDTA), dithiothreitol (DTT), b-mercaptoethanol and mercuric chloride. The effect of modifying agents for serine and histidine was determined using phenylmethylsulfonyl fluoride (PMSF) and diethyl pyrocarbonate (DPC), respectively. The enzyme was preincubated in 50 mM citrate phosphate buffer (pH 7) in the presence of the inhibitor (1 mM) at 37°C for 60 min. Subsequently, samples were cooled on ice and the residual activities were measured using the standard method. Stability against organic solvents and detergents was measured in the presence of 1% solvents and detergents within the standard activity assay, viz. glycerol, sodium dodecyl sulphate (SDS), Tween 20 and 10% solvents and detergents, viz. methanol, ethanol, 2-propanol, glycerol and dimethylsulfoxide (DMSO).

Kinetic measurements

The EstD kinetic parameters K_m and V_{max} were calculated from multiple measurements (substrate concentrations used were 0.001, 0.005, 0.01, 0.02, 0.05, 0.1, 0.15, 0.2, 0.4, 0.6, 0.8, 1.0 mM) by a computer-aided direct fit to the Michaelis-Menten curve (Tablecurve 2D, version 5.0).

ACKNOWLEDGEMENTS

Chapter 3

Crystallization and preliminary crystallographic analysis of an esterase with a novel domain from the hyperthermophile *Thermotoga maritima*

ABSTRACT

A predicted esterase (EstA) with an unusual new domain from the hyperthermophilic bacterium *Thermotoga maritima* has been cloned and over-expressed in *Escherichia coli.* The purified protein has been crystallized by the hanging-drop vapour diffusion technique, in the presence of lithium sulphate and polyethylene glycol 8,000. Selenomethionine-substituted EstA crystals were obtained under the same condition and three different wavelength data sets were collected to 2.6 Å resolution. The crystal belongs to the space group *H32*, with unit-cell parameters $a = b = 130.2$ Å, $c = 306.2$ Å. There are two molecules in the asymmetric unit with a V_M of 2.9 Å^3 Da^{-1} and 58% solvent content.

INTRODUCTION

Esterases and lipases catalyse the hydrolysis of ester bonds resulting in the formation of an alcohol and a carboxylic acid. Both enzyme types share structural characteristics, including an α/β-hydrolase fold and a conserved catalytic triad that is usually composed of a serine, an aspartate and a histidine residue [7]. Esterases differ from lipases in that they show a preference toward short-chain acyl esters (shorter than 10 carbon atoms) [7]. Ester-hydrolyzing enzymes are attractive biocatalysts, because they have a cofactor independent activity, a high regio- and stereo-specificity and are generally rather stable and active in organic solvents. There is a special interest from industry in esterases from thermophilic origin since these enzymes are both stable in organic solvents and function at elevated temperatures [16].

The hyperthermophilic anaerobic bacterium *Thermotoga maritima* was isolated from geothermally heated marine sediments at Vulcano, Italy. The organism grows optimally between 55 and 90°C and can metabolize many simple and complex carbohydrates, including glucose, cellulose and starch [119]. During a bioinformatic analysis of the genome of *T. maritima* several open reading frames potentially encoding new thermostable esterases were found, including one (*est*A, TM0033) annotated as a hypothetical protein. EstA was produced in *Escherichia coli* and found to exhibit esterase activity with a preference for esters of short chain fatty acids. It consists of 395 amino acid residues and has a predicted molecular mass of 44.7 kDa. A BLAST search of EstA at NCBI (http://www.ncbi.nlm.nih.gov/BLAST/) [28] reveals the highest similarity to other hypothetical proteins and a few predicted peptidases and carboxylesterases. An alignment with the obtained homologous sequences revealed that EstA is composed of two domains. The C-terminal domain of EstA has a predicted α/β hydrolase fold and includes the characteristic conserved Ser-Asp-His catalytic triad. However, the amino acid sequence of the N-terminal domain has no homology to known proteins in the database and has therefore an unknown function and structure. It is tempting to speculate that the N-terminal domain might be involved in substrate binding and recognition. Therefore, determination of the structure of EstA may provide insight into the function of this new N-terminal domain and reveal the molecular basis of substrate recognition and catalysis by this enzyme.

This paper describes the cloning, purification, crystallization and preliminary X-ray analysis of *T. maritima* EstA as a first step in the structure determination of this new esterase.

RESULTS AND DISCUSSION

Cloning, expression and purification

The gene encoding EstA was amplified by PCR, without the predicted signal peptide (the first 16 amino acids), and without its stop codon (fused to a

(His)$_6$-tag), using chromosomal DNA of *T. maritima* as a template and the two primers BG1962 (5'-GCGCCATGGAGGATGTTACTGTGAAGAGTG-3') and BG1963 (5'-GCGCTCGAGTCTACTTTGTTCAAACAGCCAC-3') introducing *Nco*I and *Xho*I restriction sites. The generated PCR product was digested by *Nco*I and *Xho*I and the product was purified and ligated into pET-24d digested with the same restriction enzymes, resulting in the plasmid pWUR350. The construct was designed with a hexa-histidine-tag engineered at the C-terminus of the enzyme to facilitate purification. *E. coli* BL21(DE3)/pSJS1244 was transformed with pWUR350.

A single colony was used to inoculate 4 ml Luria-Bertani medium containing kanamycin and spectinomycin (both 50 μg ml^{-1}) and incubated overnight at 37°C while shaking. The preculture was used to inoculate (1:1000) 2L of the same medium and growth was continued for 8 hours (an OD$_{600}$ above 2.0 was reached). Subsequently, the culture was induced by adding IPTG (isopropyl-ß-D-thiogalactopyranoside) to a final concentration of 0.5 mM. The culture was further incubated at 37°C for 16 hours.

Cells were harvested by centrifugation at 10,000 x *g* for 10 min. The cell pellet was resuspended in 30 ml buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 10 mM imidazole). The cells were disrupted by passing twice through a French press at 110 MPa. The crude cell extract was treated with DNAse I at room temperature for 30 minutes and subsequently centrifuged at 43,000 x *g* for 30 minutes in order to remove cell debris. The supernatant was heated at 70°C for 30 min and then centrifuged to remove the precipitated proteins. The supernatant was filtered and applied to a Ni-chelating column packed with 20 mL Ni-NTA His·Bind Resin (Novagen) equilibrated in 50 mM Tris-HCl buffer (pH 7.8) containing 300 mM NaCl. The column was washed with 20 mM imidazole in the same buffer and subsequently proteins were eluted with a linear gradient of 20-500 mM imidazole in 50 mM Tris-HCl, pH 7.5, 300 mM NaCl. Fractions containing esterase activity were pooled and applied to a Hi-prep desalting column (Amersham Biosciences) equilibrated with 50 mM disodium phosphate buffer, pH 7.5. Homogeneity of the protein was checked by SDS-PAGE electrophoresis and activity staining of the SDS-PAGE gels using α-naphtyl acetate as described previously [52]. The protein concentration was determined at 280 nm using a NanoDrop ND-1000 Spectrophotometer (NanoDrop). For preparation of the selenomethionine-substituted EstA (SeMet-EstA), the overproducing strain was grown as described previously [137] and the enzyme was purified as described above. EstA and SeMet-EstA were assayed for esterase activity using *p*-nitrophenyl-valerate as a substrate (not shown).

Crystallization

The purified protein was dialyzed against 10 mM potassium phosphate buffer (pH 7.5) and concentrated to 15 mg ml^{-1}. Crystallization was performed using the hanging-drop vapour-diffusion method at room temperature. Initial crystallization conditions were screened using

Crystal Screen and Crystal Screen II (Hampton Research). Drops consisting of equal volumes (1 µl) protein and reservoir solution were equilibrated over 500 µl reservoirs. Crystals of EstA were obtained using a reservoir solution consisting of 1.0 M lithium sulphate monohydrate, 2% (*w/v*) polyethylene glycol 8,000 (Crystal Screen condition 49). SeMet-EstA crystals were obtained under the same condition (Figure 3.1).
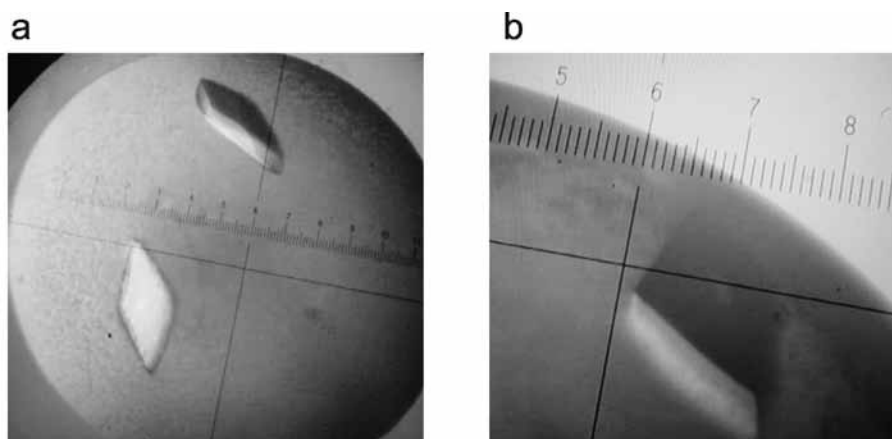
Figure 3.1: (a) Crystals of native EstA and (b) Crystals of SeMet-EstA.

Table 3.1: Data-collection and processing statistics for EstA

	Native	SeMet peak	SeMet inflection	SeMet remote
Wavelength (Å)	1.0000	0.9791	0.9793	0.9757
Resolution limit (Å)	50.0 − 2.6 (2.74-2.6) [a]	50.0 −2.6 (2.74-2.6)	50.0 −2.6 (2.74-2.6)	50.0 − 2.6 (2.74-2.6)
Space group	H32	H32	H32	H32
Unit-cell parameters (Å)				
$a = b$	130.2	131.0	131.0	131.0
c	306.2	306.8	306.8	306.8
Observed reflections	173066	353388	358408	358161
Unique reflections	31079	31295	31587	31457
Completeness (%)	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)
<I/s(I)>	14.2(3.6)	22.3 (4.6)	22.1 (4.4)	24.8 (5.0)
Redundancy	5.6(5.7)	11.4(11.7)	11.3(11.4)	11.4(11.3)

[a] Values in parentheses correspond to the highest resolution shell;

[b] $R_{merge} = \Sigma_h \Sigma_l |I_{hl} - <I_h>| / \Sigma_h \Sigma_l <I_h>$, where I_l is the lth observation of reflection h and $<I_h>$ is the weighted average intensity for all observations l of reflection h.

Data collection and preliminary X-ray analysis

For cryoprotection, crystals were soaked for a few seconds in reservoir solution containing 20% (*w/v*) glycerol. The crystals were mounted in a cryoloop and subsequently flash-frozen in liquid nitrogen. X-ray data were collected at 100K on beamline ID29 at the ESRF, Grenoble. A native data set was collected to 2.6 Å resolution. The crystal belongs to space group *H32*, with unit-cell parameters $a = b = 130.2$ Å, and $c = 306.2$ Å. There are two molecules in the asymmetric unit with a V_M of 2.9 Å³ Da⁻¹ and 58% solvent content [138]. Crystals of SeMet-EstA showed a well defined Se *K* absorption edge by fluorescence scanning. A single SeMet-EstA crystal was used for the MAD data collection at the peak (0.9791 Å), inflection (0.9793 Å) and remote (0.9557 Å) wavelengths to 2.6 Å resolution. All data were indexed and integrated using *MOSFLM* [139] and scaled with anisotropic scaling corrections using *SCALA* [140] within the *CCP*4 suite [141]. Processing statistics for the native and SeMet data sets are shown in Table 3.1.

Structure determination is currently ongoing. However, a first model shows that there are indeed two domains present in EstA: a domain resembling the α/β-hydrolase fold and a domain resembling the immunoglobulin fold. It is speculated that the α/β-hydrolase domain is the catalytically active domain and that the immunoglobulin domain might be involved in substrate binding or cell adhesion. Structure refinement combined with biochemical analyses is expected to provide better insights into the function of this esterase.

4
Chapter

Crystal structure and biochemical properties of a novel thermostable esterase containing an immunoglobulin-like domain

ABSTRACT

Comparative analysis of the genome of the hyperthermophilic bacterium *Thermotoga maritima* revealed a hypothetical protein with typical esterase features (EstA). The EstA protein was functionally produced in *E. coli* and purified to homogeneity. It indeed displayed esterase activity with optima at or above 95˚C and at pH 8.5 and 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-like (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 to 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 the wild-type activity, had a lower thermostability, and failed to form hexamers. These data suggest that the Ig-like domain plays an important role in enzyme multimerization and activity of EstA.

INTRODUCTION

Esterases constitute a large family of proteins having 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 residue [7]. Esterases differ, however, from lipases in that they show a preference toward short-chain acyl esters (shorter than 10 carbon atoms), and that they are not active on micellar substrates [7]. 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 regio- and stereo-specificity, and their stability and activity in organic solvents [3]. 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 [16].

The hyperthermophilic anaerobic bacterium *Thermotoga maritima* grows optimally at temperatures between 55 and 90 °C. It can metabolize many simple and complex carbohydrates, including glucose, cellulose and starch [142]. A bioinformatics analysis of the genome of *Thermotoga maritima* [119] revealed several open reading frames potentially encoding new thermostable esterases, including one (*est*A, TM0033) which was annotated as a hypothetical protein.

Multiple-sequence alignment suggested that EstA is composed of two domains [70], a C-terminal domain with a predicted α/β hydrolase fold and a N-terminal domain, which 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, 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 protein of 395 amino acids and has a calculated molecular mass of 44.5 kDa. A BLAST search of EstA at NCBI (http://www.ncbi.nlm.

nih.gov/BLAST [28]; revealed highest similarity to other hypothetical proteins and a few predicted peptidases and carboxylesterases. N-terminal sequence analysis using the SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP) revealed that the first 16 amino acids form a signal peptide. The predicted mature gene was cloned into the expression vector pET-24d [70], and the EstA enzyme was purified to homogeneity from heat-treated cell extracts of *E. coli* BL21(DE3) ⁄ pSJS1244 ⁄pWUR350 by immobilized nickel affinity chromatography. Homogeneity of the protein was checked by SDS-PAGE and confirmed a molecular subunit mass of 43 kDa (mature enzyme). Activity staining of the washed SDS-PAGE gels with α-naphthyl acetate was used to confirm the identity of the EstA band (not shown).

Overall structure

The native EstA crystal structure was solved in space group *H32* with 2 molecules per asymmetric unit. The structure was determined by multi-wavelength 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 x 35 Å x 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-like (Ig-like) domain (residues 20-157) (Figure 4.1a).

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 (Figure 4.1b), with strand β2 antiparallel to the other strands. The β1 and β8 strands are approximately perpendicular to each other because of the counter-clockwise twist of the β-sheet. Helices αA, αFa, αFb, and the short 3_{10}-helices η4 and η5 are on one side of the central β-sheet, and helices αB, αC, αE, and the two 3_{10}-helices η5 and 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 (Figure 4.1b). Ig-like domains are common in eukaryotic proteins, but 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 [143] 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 domain is structurally most similar to the catalytic domain of prolyl tripeptidyl aminopeptidase from *Porphyromonas gingivalis* (pdb 2D5L [144]), the catalytic domain of human dipeptidyl peptidase (pdb 1N1M [145]) and the feruloyl esterase domain of xylanase Z from *Clostridium thermocellum* (pdb 1JJF [146]). The structures are well superimposed,
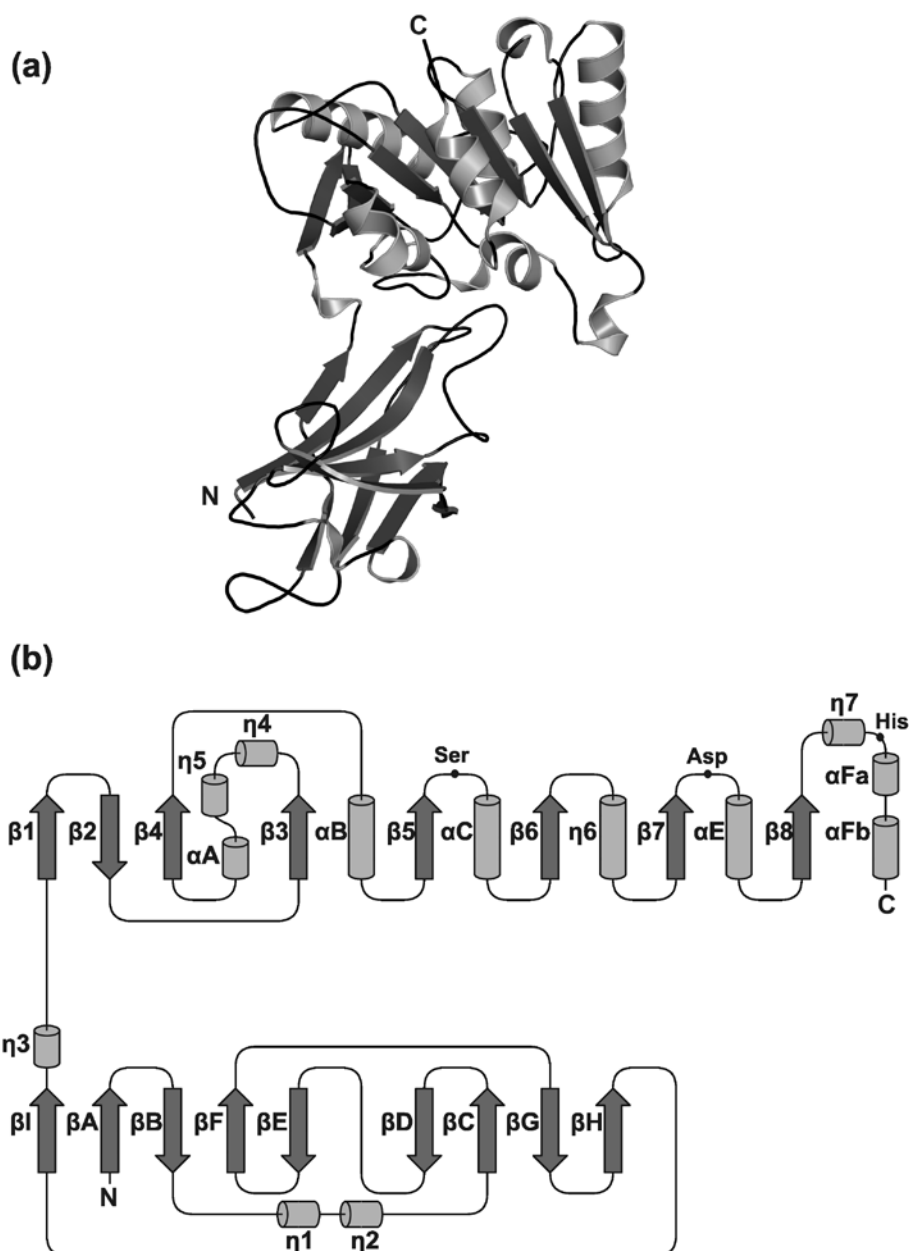
Figure 4.1 (in color on p.145): (a) Overall fold of the esterase EstA. The N- and C-terminal ends are indicated. The figure was generated using Pymol [134]. (b) Topology diagram for EstA with the helices displayed as cyan cylinders and the strands as red arrows. The positions of the catalytic residues are indicated.

with RMSDs of 2.6 Å (for 204 aligned Cα-atoms), 2.8 Å (200 aligned Cα atoms), and 2.6 Å (202 aligned Cα 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 [147]), 5.4 for the Ig-like D1 domain of Interleukin-4 receptor (pdb 1IAR [148]) and 5.2 for the β-galactosidase Ig-like domain 4 (pdb 1BGL [149]). 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 phenylmethylsulfonyl fluoride (PMSF) and diethyl 4-nitrophenyl phosphate (paraoxon) inhibit EstA activity, with respectively 7% and 0% residual activity. 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 [150]. 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 [151]. Inhibition kinetics for PMSF were not measurable in the presence of substrate. Possibly, this is a result of a high dissociation 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 EDTA did not influence EstA activity. EstA was co-crystallized with PMSF and paraoxon, but only crystals for the latter were obtained. Electron density maps for the paraoxon co-crystallized crystals displayed clear density for a diethyl phosphate 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 co-crystallization, thereby leaving a tetrahedral product resembling the first transition state formed during ester hydrolysis. The native and the 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 (protomer A and protomer B) in the asymmetric unit which are related by non-crystallographic symmetry. Protomers A and B are essentially similar, with an RMSD between them of 0.3 over all Cα atoms. The refined model reveals that the two molecules form an interface of 280 Å^2 in each monomer, a value that suggests a low association constant. The interface involves four beta-strands (βC, βD, βG, and βH) from the N-terminal Ig-like domain of both molecules. An intermolecular hydrogen bond (AspA56 to 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 3,585 Å^2. The inter-dimer 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 to GluA'362, ArgA222 to GluA'360, GluA99 to LysA'363, ArgA187 to GluA'351, GlnA384 to TyrA'358, and AspA182 to LysA'347. On average, 1,454 Å^2 surface area
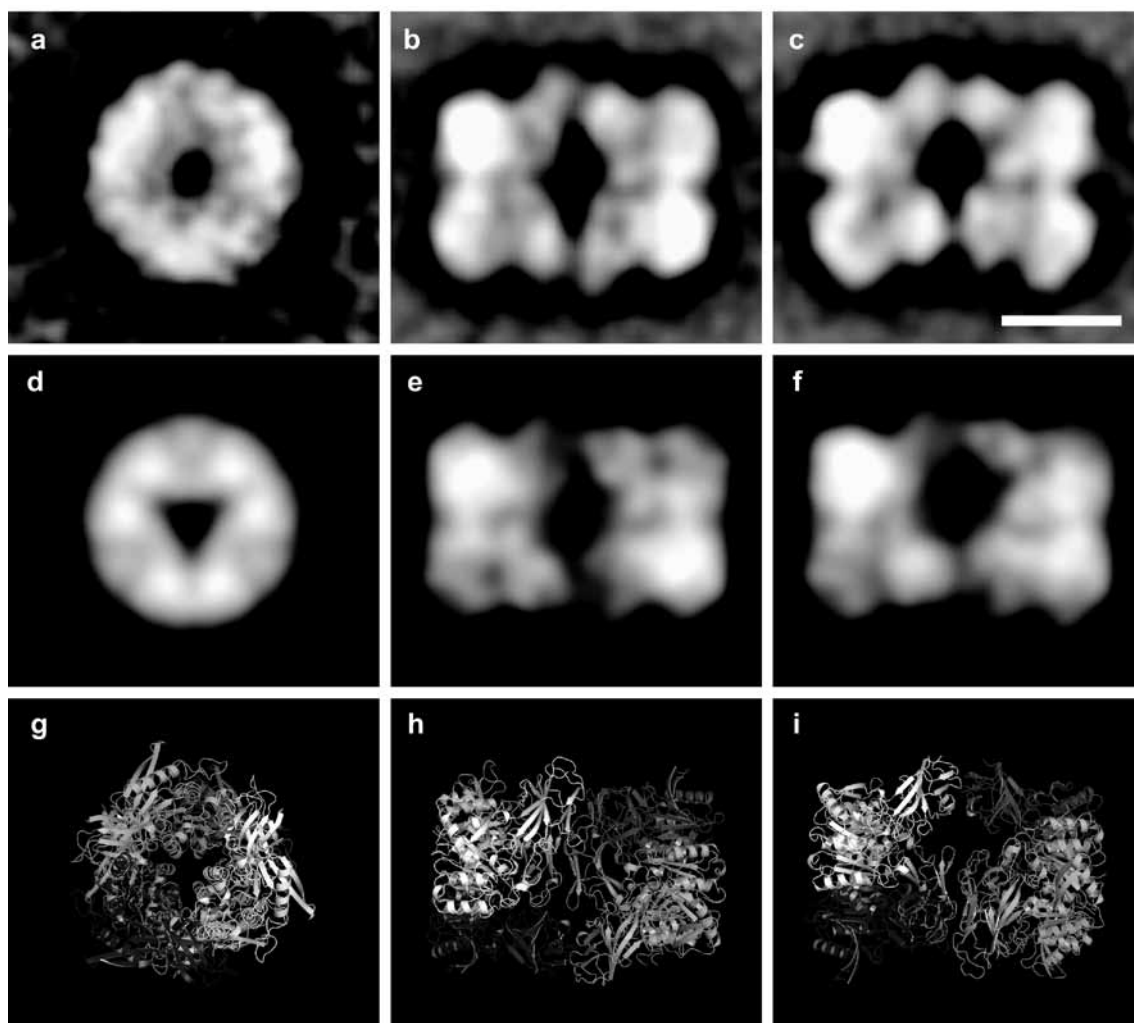
Figure 4.2 (in color on p.146): Comparison of the EM projection maps of the EstA hexamer. Top view (a), side view (b) and 10 degrees off side view (c). Two-dimensional projection maps obtained by statistical analysis and classification (a - c), the comparable two-dimensional projection maps with 15 Å resolution (d - f) generated from the proposed EstA hexameric structure (g - i) [134]. The scale bar equals 50 Å.

per monomer is buried upon hexamerization. A 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 non-denaturing conditions in ammonium acetate at pH 6.8 [152], and dynamic light scattering were performed and revealed that the purified EstA protein was present predominantly (>50%) as a hexamer in solution. However, to a minor extent also mono-, di-, tri, and higher multimeric forms of EstA were detected using mass-spectrometry.

Electron microscopy 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 were found suitable for single particle image analysis. The analysis

confirmed that EstA forms hexamers in solution (Figure 4.2). 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 which are recognizable in the EM projection map despite the limited details. Two EM side views are presented in Figure 4.2, a side view (b) and a tilted side view (c). The projection in Figure 4.2b is referred as side view because two-fold symmetry can be imposed without altering the main characteristics in the projection map (not shown here). However, if this two-fold symmetry is imposed on Figure 4.2c, the overall shape changes. This indicates that the hexamer shown in Figure 4.2c is slightly tilted due to different orientations on the carbon support film and therefore referred as tilted side view. It was possible to establish the rotational difference between these two views. Rotating the proposed hexamer followed by examination of the generated projection maps revealed that the tilted side view is about 10 degrees off from the side view. The tilted side view provides further evidence of the proposed EstA hexamer. Altogether, electron microscopy 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 to 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 was found, consisting of Ser286 as the nucleophile, His374 as the proton acceptor/donor and Asp334 as the residue that stabilizes the histidine. Ser286 is located at the nucleophile elbow, a sharp turn between β5 and helix αC. In both native and the paraoxon-bound structures, Ser286 has the energetically unfavorable main chain conformation which is also observed in other
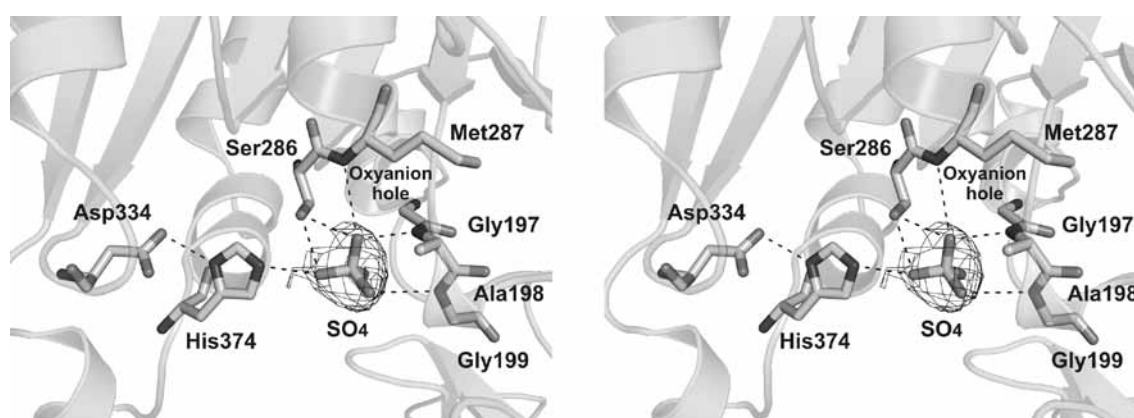
Figure 4.3 (in color on p.147): Stereoview of the EstA catalytic site with the bound SO$_4$ adduct shown. The catalytic triad residues are shown as sticks with the hydrogen bonds shown as dashed black lines. Observed density for the bound SO$_4$ is contoured. These images were generated using Pymol [134].
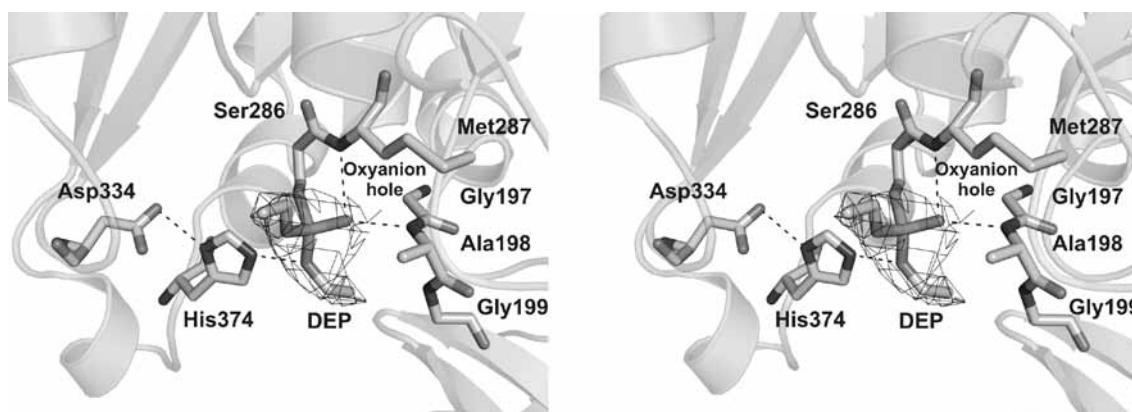
Figure 4.4 (in color on p.147): Stereoview of the EstA catalytic site with the diethyl phosphate (DEP) intermediate of the inhibitor paraoxon covalently bound to the catalytic serine. The catalytic triad residues are shown as sticks with the hydrogen bonds shown as dashed black lines. Observed density for the bound diethyl phosphate is contoured. These images were generated using Pymol [134].

α/β hydrolases and might provide an energy reservoir for catalysis [153]. 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 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. Extra density near the side-chain of Ser286 was interpreted as a sulfate-ion, forming hydrogen bonds to the side chain of the active site residue Ser286 and the main chain nitrogen atoms of Ala198 and Met287, mimicking the oxyanion transition state (Figure 4.3). 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 by a hydrogen bond to the side-chain of His374 (Figure 4.4). 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 for many α/β hydrolases (Figure 4.5a). This groove extends approximately 15 Å from the catalytic serine into the enzyme where the gorge is closed by Glu33 (S1 in Figure 4.5b). 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 for 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 non-polar side-chain of Tyr290. A second side gorge with a volume of ~210 Å^3 also provides access to the active site (S2 in Figure 4.5b). The ~7 Å wide opening
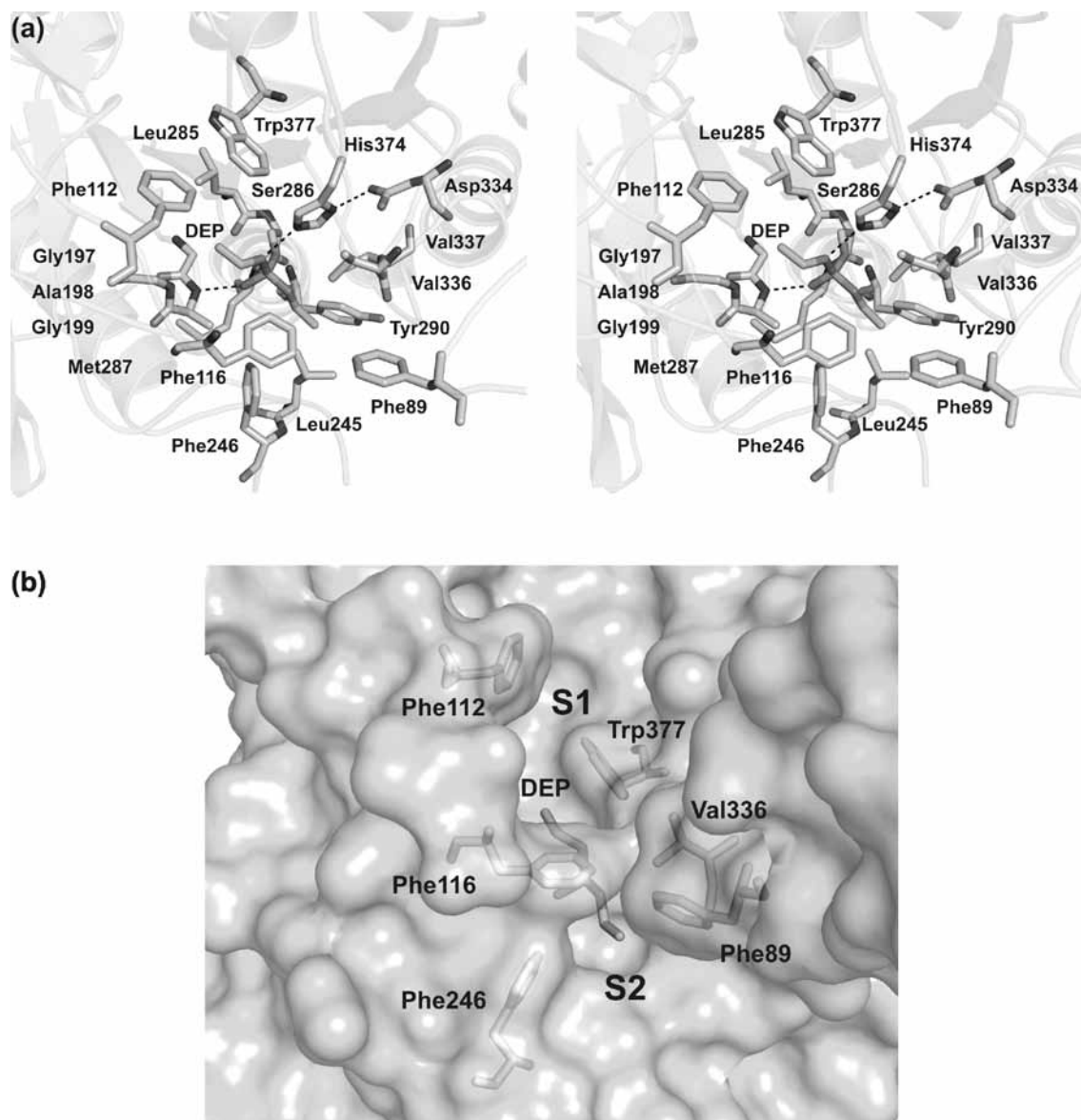
Figure 4.5 (in color on p.148): Substrate binding gorge. (a) Stereoview of the active site with the bound diethyl phosphate intermediate covalently bound to the catalytic serine. Key residues of the catalytic gorge are shown as sticks with the hydrogen bonds shown in dashed lines. (b) Surface representation 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 [134].

is lined by the 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 length are given in Table 4.1. Catalytic activity (K_{cat}) decreases markedly with an increase in chain length with only low levels of activity on the long chain esters. The optimum substrate for an

Table 4.1: Kinetic parameters for 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

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 upwards until 95 °C (Figure 4.6a). An Arrhenius analysis for *p*-nitrophenyl valerate and octanoate resulted in a linear plot in the temperature range of 40-80 °C (Figure 4.6a), with a calculated activation energy for the formation of the enzyme-substrate complex of 9.5 and 13.5 kJ/mol respectively, which is comparable to a reported value [52]. EstA has a high resistance to thermal inactivation, with a half-life value of approximately 1.5 h at 100°C (not shown). The optimal pH for the esterase was measured in the pH range of 4.8 to 10.2 using the substrates *p*-nitrophenyl valerate and octanoate. The latter substrate was used due to instability of *p*-nitrophenyl valerate at higher pH values. EstA displayed maximal activity at approximately pH 8.5 (Figure 4.6b).

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 on the basis of the structure of *Pseudomonas cepacia* lipase (pdb 1LIP) with the covalently bound tributyrin transition state analogue R_C-(R_P,S_P)-1,2-dibutylcarbamoylglycero-3-*O*-*p*-nitrophenyl butylphosphonate (not shown) [154]. 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 *Pseudomonas 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, 7-aminocephalosporanic acid (7-ACA), glucose pentaacetate and N-acetyl-D-glucosamine. EstA displayed an activity of 20 U/mg on glucose pentaacetate. This activity is relatively low, suggesting that EstA is not an oligosaccharide deacetylase. EstA was also able
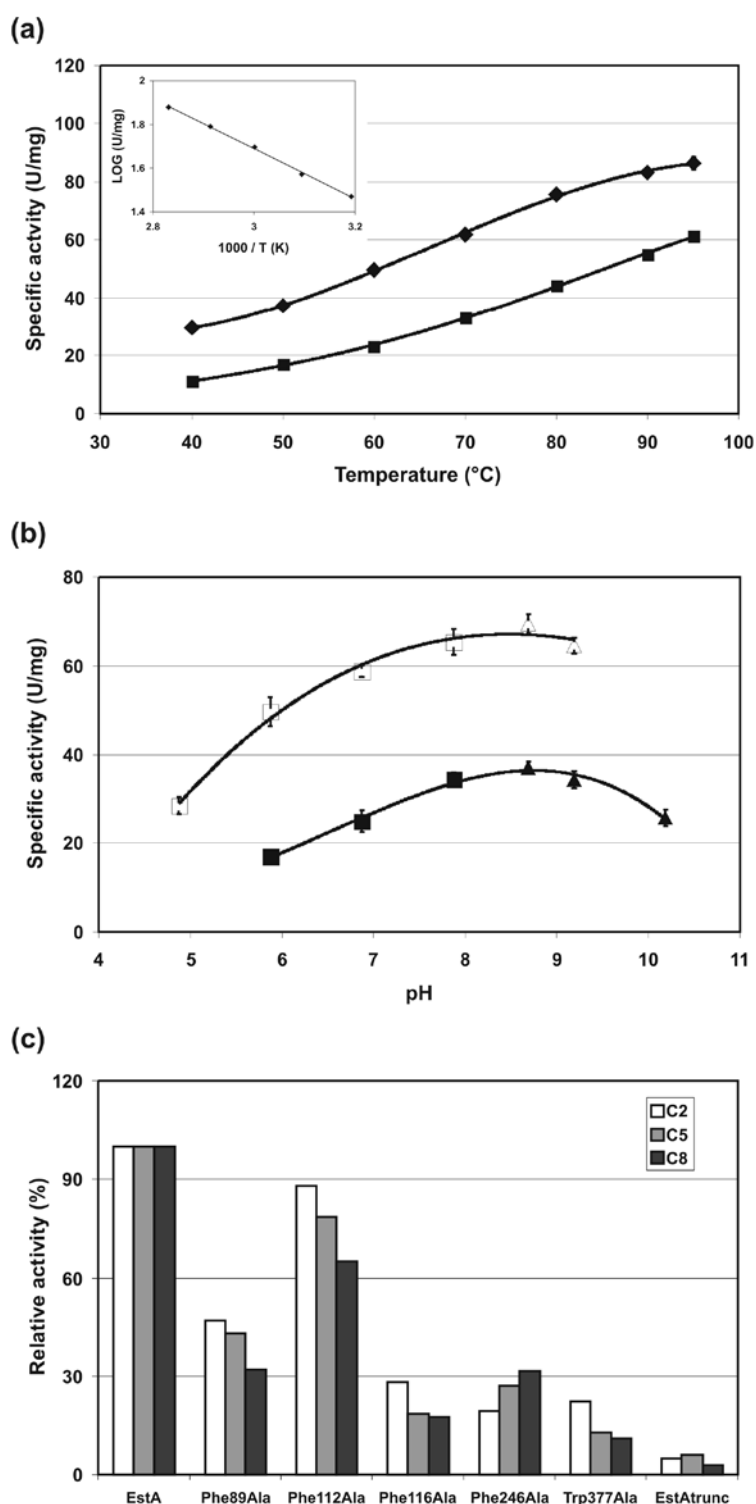
Figure 4.6: Effect of temperature, pH and mutations on esterase activity. (a) The effect of temperature on esterase activity was studied using *p*-nitrophenyl-valerate (■) and -octanoate (♦) as a substrate at temperatures ranging from 40°C to 95°C. The inset shows the temperature dependence for *p*-nitrophenyl-valerate as an Arrhenius plot. (b) The effect of pH on esterase activity was studied using *p*-nitrophenyl-valerate (□ citrate phosphate buffer and Δ CAPS buffer) and -octanoate (■ citrate phosphate buffer and ▲ CAPS buffer) as substrates at pH values ranging from pH 4.8 to pH 10.2. (c) The effect of mutations on esterase activity was studied using *p*-nitrophenyl-acetate (C2), -valerate (C5) and -octanoate (C8) at pH 8 and 70°C.

to hydrolyze acetyl groups from both cephalosporin C and 7-ACA with an activity of 80 U/mg for both substrates. Cephalosporin C and 7-ACA are not stable at high temperature and are therefore not considered as natural substrates. EstA was not able to remove the acetyl group from N-acetyl-D-glucosamine, indicating 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, 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 towards the active site and covers 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 and revealed that all mutants were present as a hexamer in solution. Esterase activity was determined using substrates with different acyl chain lengths, respectively pNP-C2, pNP-C5 and pNP-C8 (Figure 4.6c). Significant reduction in activity was observed for all five mutants, confirming their importance for the activity of EstA. Their relative activity to EstA with the substrate pNP-C5 was, respectively, approx. 30% for Phe89Ala, 80% for Phe112Ala, 20% for Phe116Ala, 25% for Phe246Ala and 15% for Trp377Ala. As can be seen in Figure 4.6c, the effect was more pronounced for longer acyl chain substrates.

In addition, a truncated mutant of EstA was constructed coding only for the catalytic domain (EstAtrunc: Asp158-Arg395). EstAtrunc was expressed and purified and was found to have a relative activity of only 5% compared to EstA (Figure 4.6c). Furthermore, it had a much lower temperature optimum (60°C), lower thermostability (a half-life at 90°C of 15 minutes) 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. 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 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 a lipase.

The three-dimensional structure of the protein was determined at 2.6 Å resolution. Analysis of the structure of EstA confirms that it is a member of the α/β hydrolase family, with a conserved Ser-Asp-His catalytic triad, comprised of 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 immunoglobulins. Such a combination of an esterase- or α/β hydrolase domain with an Ig-like domain is new and has, as such, not yet been described.

Analysis of the quaternary structure by gel filtration, mass-spectrometry and dynamic light scattering revealed that EstA predominantly exists as hexamer in solution. The crystal structure also shows a hexameric arrangement, composed of two trimers. Electron microscopy demonstrated that the hexamer in solution is identical to 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 *e.g.* described for the thermostable esterase from *Bacillus circulans* [155] and the thermostable esterase from *Sulfolobus shibatae* [100]. 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 [156; 157]. 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 was constructed only composed of the C-terminal catalytic domain. 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 that was 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 the thermostability [158].

The active site is accessible via a gorge, typical for 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 substrate or an exit for one of the reaction product following 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 better 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 of the active site were mutated to alanines. The specific activity of each mutant was decreased indicating they all were important for activity. Least affected was mutant Phe112Ala, however, this residue is located furthest 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 stabilization of the active site can, however, not 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 the activity of EstA, although not essential, since Phe89Ala still has 30% activity.

In conclusion, the structural and biochemical characterization of EstA showed that it is an unusual esterase, which is composed of a conserved C-terminal catalytic domain and an unprecedented N-terminal immunoglobulin-like domain. The Ig-like domain presumably plays a role in multimerization of EstA into an unusual hexameric structure. Additionally, it may also participate in 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 *Thermotoga 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 [70]. The purified native and SeMet-EstA proteins were dialyzed against 10 mM potassium phosphate buffer at pH 7.5 and concentrated to 15 mg/ml. 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, 2% w/v polyethylene glycol 8,000. 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 one and two weeks, respectively for the native and SeMet-EstA. A crystal of EstA in complex with its inhibitor paraoxon was obtained by incubating the enzyme (15 mg/ml) supplemented with 0.2 mM paraoxon for 1h at room temperature and then set up crystallization as described above. Crystals were obtained within two weeks.

Data collection

For cryoprotection, crystals were soaked for a few seconds in reservoir solution containing 20% (*w/v*) glycerol. The crystals were mounted in a cryoloop and subsequently flash-frozen in liquid nitrogen. X-ray data were collected at 100K on beamline ID29 at the ESRF, Grenoble. A native data set was collected to 2.6 Å resolution. The crystal belongs to space group *H32*, with unit-cell parameters $a = b = 130.2$ Å, and $c = 306.2$ Å. There are two molecules in the asymmetric unit with a V_M of 2.9 Å^3 Da^{-1} and 58% solvent content [138]. Crystals of SeMet-EstA showed a well defined Se *K* absorption edge by fluorescence scanning. A single SeMet-EstA crystal was used for the MAD data collection at the peak (0.9791 Å), inflection (0.9793 Å) and remote (0.9557 Å) wavelengths to 2.6 Å resolution. Data were indexed and integrated with MOSFLM [139] and scaled using SCALA [141].

Structure determination and refinement

The EstA structure was solved by MAD phasing with the MAD data from the SeMet-EstA crystal and the native data set, using HKL2MAP [159]. Eight selenium sites in the asymmetric unit of the crystal were found and were used to calculate phases to 2.6 Å resolution (Table 4.2). However, initial density maps were of generally poor quality and not suitable for tracing the structure. Phases were improved using RESOLVE [160; 161], allowing the identification of non-crystallographic symmetry (NCS). NCS averaging and solvent flattening were performed using the program DM [162] of the CCP4 suite, giving an electron density map of better quality. Autobuilt models from RESOLVE and ARP/wARP [163] were combined to give a starting model, comprising 170 residues from the total of 380 with only 20 residues assigned into sequence properly. The structure was then manually rebuilt in Coot [164] and refined using CNS [165] and REFMAC [166]. Strict NCS restraints were applied during the earlier stages of the refinement, and were released at later stages. In the final stages of refinement, solvent molecules were added using ARP-wARP [163] and were manually inspected in Coot [164]. 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 the model. The complex structure was rebuilt in Coot [164] and refined using REFMAC [166]. 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 $R_{work} = 22.0\%$ and $R_{free} = 26.2\%$.

In the final models (native and paraoxon-complexed crystal structures), residues 16-20 of both molecules and residues 249-250 of chain B are missing because of poor density. The geometry of both models was monitored using PROCHECK [124], with the native and paraoxon complex models having 85.2 and 82.2% of their residues in the most favored regions of the Ramachandran plot, respectively. Cartoon representations were generated using PYMOL [134].

Table 4.2: 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 = *b*	130.2	131.0	131.0	131.0	130.5
c	306.2	306.8	306.8	306.8	304.5
Observed reflections	173066	353388	358408	358161	160768
Unique reflections	31079	31295	31587	31457	20405
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)
<I/s(I)>	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					
No. 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)
No. of reflections	28373				19311
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
Rms deviation					
Bond lengths (Å)	0.018				0.016
Bond angles (°)	1.88				1.77

[a.] Values in parentheses correspond to the highest resolution shell;

[b.] $R_{merge} = \sum_h \sum_l |I_{hl} - <I_h>|/ \sum_h \sum_l <I_h>$, where I_l is the *l*th observation of reflection h and $<I_h>$ is the weighted average intensity for all observations *l* of reflection h.

Native mass spectrometry

Native mass spectrometry measurements were performed in positive ion mode using an Electrospray Ionisation Time-of-Flight (ESI-ToF) instrument (LC-T; Micromass, Manchester, U.K.) equipped with a Z-spray nano-electrospray 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 by 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 7.0-7.3 mbar, and nano-electrospray 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 [167; 168]. The applied voltages on the needle and sample cone were 1,300 V and 150 V, respectively. All spectra were mass calibrated by using an aqueous solution of cesium iodide (20 mg/ml). Buffer exchange of EstA samples to 100 mM ammonium acetate pH 6.8, was performed by using ultra filtration units with a cut-off of 5,000 Da (Millipore, Bedford). EstA was diluted to 5 μM and measured at room temperature.

Dynamic light scattering (DLS)

DLS was performed at room temperature using the EstA sample ready for crystallization. The EstA particles were monodisperse, with a molecular weight of approximately 285 kDa, which suggests that the 43 kDa subunits assemble into a hexamer in solution.

Electron microscopy 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 CCD camera on a Philips CM12 electron microscope (Fei, Eindhoven, The Netherlands) operated at 120 kV using ''GRACE'' software for semi-automated specimen selection and data acquisition [169]. The final magnification was 100,000x with a pixel size (after binning the images) of 3.0 Å at the specimen level. About 8,700 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 particles projections (96 x 96 pixel frame) were subjected to multi-reference alignment and reference-free alignment procedure, multivariate statistical analysis followed by and hierarchical ascendant classification [170]. 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. 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 [134] for possible hexameric structures which resembled the projection maps obtained with electron microscopy 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 [134]. 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 [171].

Enzyme assays

Esterase activity was measured as described previously [52]. 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 of 5-11. The buffers used were 50 mM citrate-phosphate (pH 5-8) and 50 mM CAPS buffer (pH 9.5-11). The effect of temperature on esterase activity was studied in the range of 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 for citrate-phosphate buffer and -0.018 pH/°C for CAPS buffer [136]. Enzyme thermostability was determined by incubating the enzyme in a 50 mM Tris-HCl, 150 mM NaCl (pH 7.8) buffer at 100 °C for various time intervals. Residual activity was assayed under the standard condition.

Plate assays were used for the detection of activity on triacylglycerol-esters. Agar plates containing tributyrin, trioctanoate, triolein, palm oil and olive (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 (HPLC) by measuring the amount of acetic acid released from the substrates cephalosporin C, 7-aminocephalosporanic acid and glucose-pentaacetate. The reaction mixture contained 0.9 ml 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 stop solution (100 mM H_2SO_4 and 30 mM crotonate) and placing the sample on ice. The conditions for HPLC were as follows: column, KC811-Shodex; detection, RI and UV detector; solvent, 3 mN H_2SO_4; flow rate, 1.5 ml/min; temperature, 30°C; internal standard, crotonate. One unit of enzyme activity was defined as the amount of enzyme that releases one μmol of acetic acid per minute.

The effect of inhibitors and metals was studied by pre-incubating EstA with 1 mM inhibitor in 50 mM citrate phosphate buffer (pH 8) at 37°C for 1 hour. Subsequently, samples were placed on ice and residual activity was measured using the standard assay. Activity of the enzyme without inhibitor was defined as 100%. Inhibition kinetics of paraoxon were determined as described for the acetylcholinesterase from electric eel [150].

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-CGGAGGTCTCTCTTTCAG-CATTGCCAACAACCGTGGAAAGTACG-3 and BG2487 5-CGTACTTTCCACGGTTGTTGGCAATGCT-GAAAGAGAGACCTCCG-3, BG2488 5-CACTCAAACACCATTGTTGCCGGGGCCTAATTTTCTCAA-

CACCCG-3 and BG2645 5-CGGGTGTTGAGAAAATTAGGCCCGGCAACAATGGTGTTTGAGTG-3, BG2490 5-CCATTGTTTTCGGGCCTAATGCGCTCAACACCCGCATGAAACTGG-3 and BG2646 5-CCAGTTTCATGCGGGTGTTGAGCGCATTAGGCCCGAAAACAATGG-3, BG2492 5-CAGCAGCTG-GTCCACACTCGCCACTGACAGGGAAAATCC-3 and BG2647 5-GGATTTTCCCTGTCAGTGGCGAGT-GTGGACCAGCTGCTG-3, BG2494 5-GGTGGGATCCACACGGATCGGCGATACCTACCTACGAG-3 and BG2648 5-CTCGTAGGTAGGTATCGCCGATCCGTGTGGATCCCACC-3, respectively (the underlining indicates the modified codon). A truncated EstA mutant was generated using the following primers BG2496 5-CCGCCATGGATTTCCTCGCATTCACTTTCAAAGACC-3 and BG1963 5-GCGCTC-GAGTCTACTTTGTTCAAACAGCCAC-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 Protein Data Bank with accession codes 3DOH for native EstA and 3DOI for EstA in complex with paraoxon.

ACKNOWLEDGEMENT

5

Chapter

The crystal structures of a thermostable acetyl esterase / cephalosporin C deacetylase from *Thermotoga maritima* in complex with PMSF and paraoxon reveal an altered conformation of the catalytic serine

Levisson, M., Han, G.W., Deller, M.C., Biely, P., Hendriks, S., Xu, Q., Ten Eyck, L., Flensburg, C., Roversi, P., Miller, M., McMullan, D., von Delft, F., Kreusch, A., Elsliger, M., Deacon, A., Lesley, S., van der Oost, J., Wilson, I. & Kengen, S.W.M.
Manuscript in preparation.

ABSTRACT

Deacetyl cephalosporins are valuable building blocks for the production of semisynthetic β-lactam antibiotics. These compounds can be produced from cephalosporin C or 7-aminocephalosporanic acid by the action of specific esterases, belonging to the carbohydrate esterase family seven (CE7). Members of this family are active towards both cephalosporin C and acetylated xylo-oligosaccharides. The acetyl esterase AceA from *Thermotoga maritima*, belonging to CE7, was functionally produced in *Escherichia coli* and purified to homogeneity. AceA was found to be active on a variety of acetylated compounds, including cephalosporin C. On artificial *p*-nitrophenyl-substrates esterase activity was confined to short chain acyl esters (C2-C3), and showed optima at or above 100°C and at pH 7.5. The positional specificity of AceA was investigated using 4-nitrophenyl-β-D-xylopyranoside monoacetates as substrate in a β-xylosidase-coupled assay. AceA hydrolyzed acetate from positions 2, 3 and 4 with the same efficiency. Its selenomethionine-substituted and native structures were solved at 2.1 Å and 2.5 Å resolution, respectively, and revealed a classic α/β-hydrolase fold. AceA forms a "doughnut"-like hexamer in solution, with two small tunnels leading to an inner cavity, to which the six catalytic centers are exposed. AceA is irreversibly inhibited by the inhibitors PMSF and paraoxon. Structures of AceA in complex with PMSF and paraoxon were determined to 2.4 Å and 2.1 Å, respectively, and confirmed that both inhibitors bind covalently to the catalytic serine. Remarkably, upon binding of inhibitor the catalytic serine adopts an altered conformation. It is speculated that this transition is necessary for good hydrogen bonding to the oxyanion hole.

INTRODUCTION

Thermotoga maritima is a hyperthermophilic bacterium growing optimally at 80°C and able to metabolize a variety of simple and complex carbohydrates, including glucose, sucrose, starch, cellulose, and xylan [142]. Its carbohydrate utilization potential was revealed by its genome sequence [119]. The xylan degrading system of *T. maritima* has been studied using microarrays [119; 172; 173], and several genes encoding transporters, xylanases, and a β-xylosidase have been identified. Among the enzymes with a differential expression pattern in the microarray was a predicted acetyl xylan esterase (locus tag TM0077, *ace*A) [172; 174]. In hardwood xylans many of the xylose residues are acetylated at the 2 and / or 3 positions. Therefore, for the complete degradation of xylan, esterases / deacetylases are required in addition to xylanases and xylosidases [175].

Presently, esterases and deacetylases active on carbohydrate substrates have been classified into 16 families by Henrissat and coworkers (Carbohydrate-Active enZymes Server (CAZy)) [114]. According to this classification, the predicted acetyl xylan esterase from *T. maritima* belongs to family 7 of the carbohydrate esterases (CE7). Enzymes in the CE7 family are rather unusual in that they, in addition to the acetyl xylan esterase activity, display a high specific activity towards the antibiotic cephalosporin C [176]. Cephalosporin C belongs to the class of β-lactam antibiotics that also includes penicillin. These antibiotics affect cell growth by inhibiting the penicillin-binding-protein that cross-links peptide glycans required for cell wall formation [177]. The production of deacetylated cephalosporins is of great interest because these compounds are valuable building blocks for the production of semi-synthetic β-lactam antibiotics [178; 179].

In order to disclose the catalytic capacity of the predicted acetyl xylan esterase from *T. maritima* and to gain a better insight into the structure and function of the family 7 carbohydrate esterases, TM0077 was functionally expressed and its three-dimensional structure, native and in complex with the inhibitors PMSF and paraoxon, was determined. In addition, various biochemical properties and the positional specificity of the esterase was investigated.

RESULTS

In silico analysis

AceA was selected as one of the model proteins by the Joint Center for Structural Genomics (JCSG), as part of the NIH-funded Protein Structure Initiative (PSI), to establish an automated pipeline for large-scale protein production and structure determination [180]. The protein consists of 325 amino acids and has a calculated molecular mass of 37 kDa. N-terminal sequence analysis, using the SignalP 3.0 server, revealed that AceA has no signal sequence and

is therefore believed to be an intracellular enzyme. Analysis of the gene organization shows that the AceA gene (TM0077) co-localizes with genes encoding a xylanase (TM0070) [181], 5 ABC transporter components (TM0071-TM0075), and a β-xylosidase (TM0076) [182]. BLAST-P analysis showed that AceA has highest similarity with putative acetyl esterases, acetyl xylan esterases and cephalosporin C deacetylases. Among the BLAST results there was also a hit with an acetyl xylan esterase-related protein from *T. maritima* (locus tag: TM0435). AceA was compared with other members of the CE7-family using a structure-based multi-sequence alignment. The three catalytic residues Ser188, Asp274, and His303, were identified because of their complete conservation. The nucleophilic serine is located within a conserved pentapeptide consensus sequence, Gly-Xaa-Ser-Gln-Gly, typical for this family. Previously, a signature sequence motif, RGQ...GxSQG...HE (at the appropriate spacing), was suggested for the CE7 family based on an alignment of 12 amino acid sequences [183]. We have confirmed this motif in the multi-sequence alignment using 50 sequences belonging to the CE7 family.

Overall structure

The crystal structure of seleno-methionine incorporated AceA (AceA-SM) was determined to 2.1 Å by multi-wavelength anomalous dispersion (MAD) (Table 5.2). The AceA-SM structure was solved in space group $P2_1$ and has twelve AceA molecules per asymmetric unit. Native AceA and the PMSF (AceA-PMS) and paraoxon (AceA-DEP) co-crystallized structures were determined to 2.5 Å, 2.4 Å and 2.1 Å, respectively (Table 5.2). They were solved by molecular replacement, using AceA-SM as a searching molecule, and belong to space group $P2_12_12_1$ containing six AceA molecules in the asymmetric unit. In each monomer of the hexamer a calcium ion is bound by Lys22, Glu26, and Asp25 through a bridging water molecule. A superposition of the AceA-SM and native AceA structures gives a root mean square deviation (RMSD) of 0.185 Å, which implies that both overall structures are nearly identical.

The AceA structure resembles the canonical α/β-hydrolase fold, which has been described as a central twisted eight-stranded β-sheet surrounded by α-helices on both sides, with β2 antiparallel to the other strands. Compared with the canonical α/β-hydrolase fold, AceA has an insertion of three helices after β6 and an extension of the N-terminus (Figure 5.1). Insertions after β6 or β7 are common for α/β-hydrolases and are thought to help shape the substrate-binding site [8]. The N-terminus is extended with two helices and an antiparallel β-strand, which lines up with the other eight β-strands, extending the central sheet. Therefore, the second β-strand is equivalent to β1 in the canonical fold. The nine-stranded β-sheet is highly twisted, and β1 and β9 are approximately 130°-rotated relative to each other. Helices αA, αA1, αC, αD, αE, αE1, αE2, αE3, αF, and the 3_{10}-helix η1 are on one side of the central β-sheet, and helices αB and αG are on the other side. Helices αB and αC start with a 3_{10}-helix.
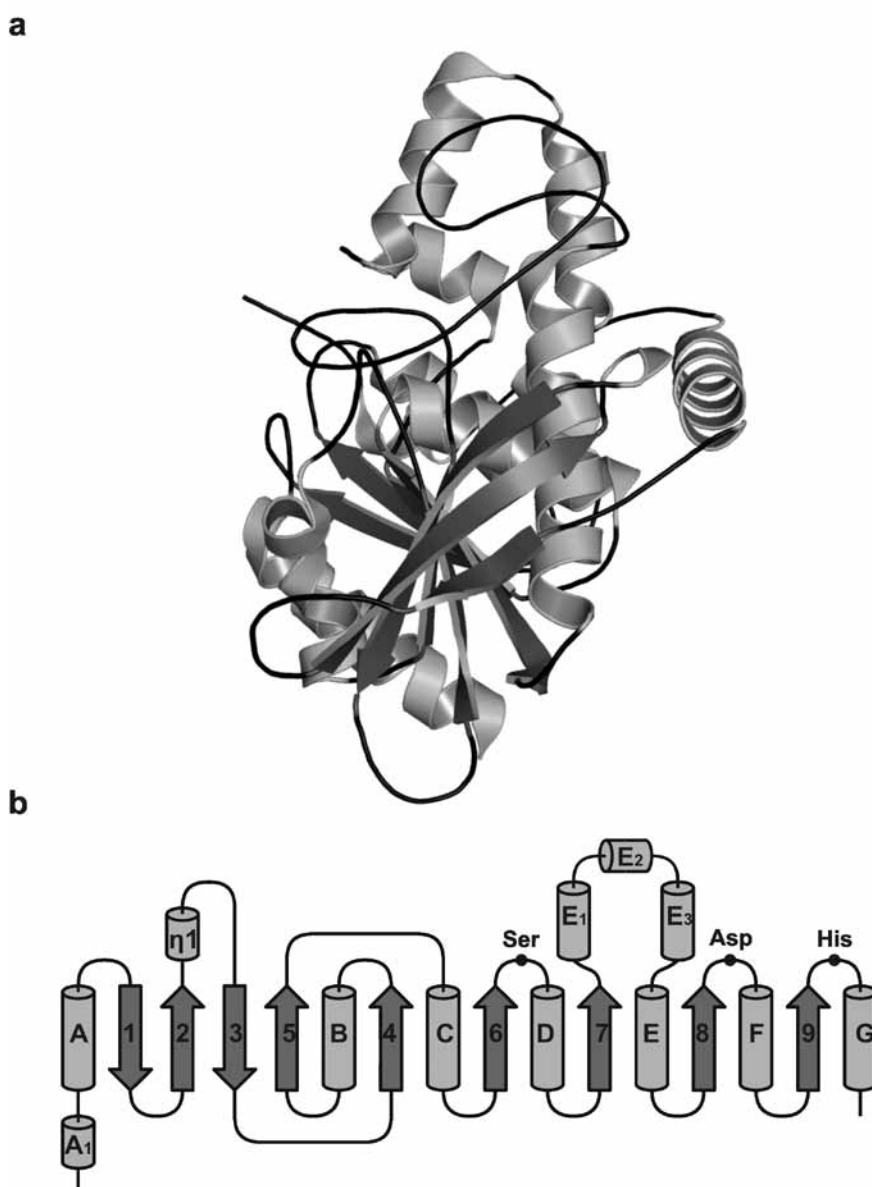
Figure 5.1 (in color on p.149): (a) Overall fold of the acetyl esterase AceA. The Figure was generated using Pymol (DeLano) [134]. (b) Topology diagram for AceA, with the helices displayed as cyan cylinders and the strands displayed as red arrows. The positions of the catalytic residues are indicated. The figure was generated using TopDraw [184].

Quaternary structure

There is one hexamer in the asymmetric unit of native AceA. Crystallographic packing analysis using PQS [185] indicated that the physiological multimeric state of AceA is a hexamer. Size exclusion chromatography coupled with static light scattering did confirm this.

The assembly of the AceA hexamer was studied using PISA (EBI) [186]. This analysis indicated that the interfaces between protomers AB = CD = EF and AF = BC = DE play an essential role in complex formation (CSC scores of 1). The interface between protomers A and
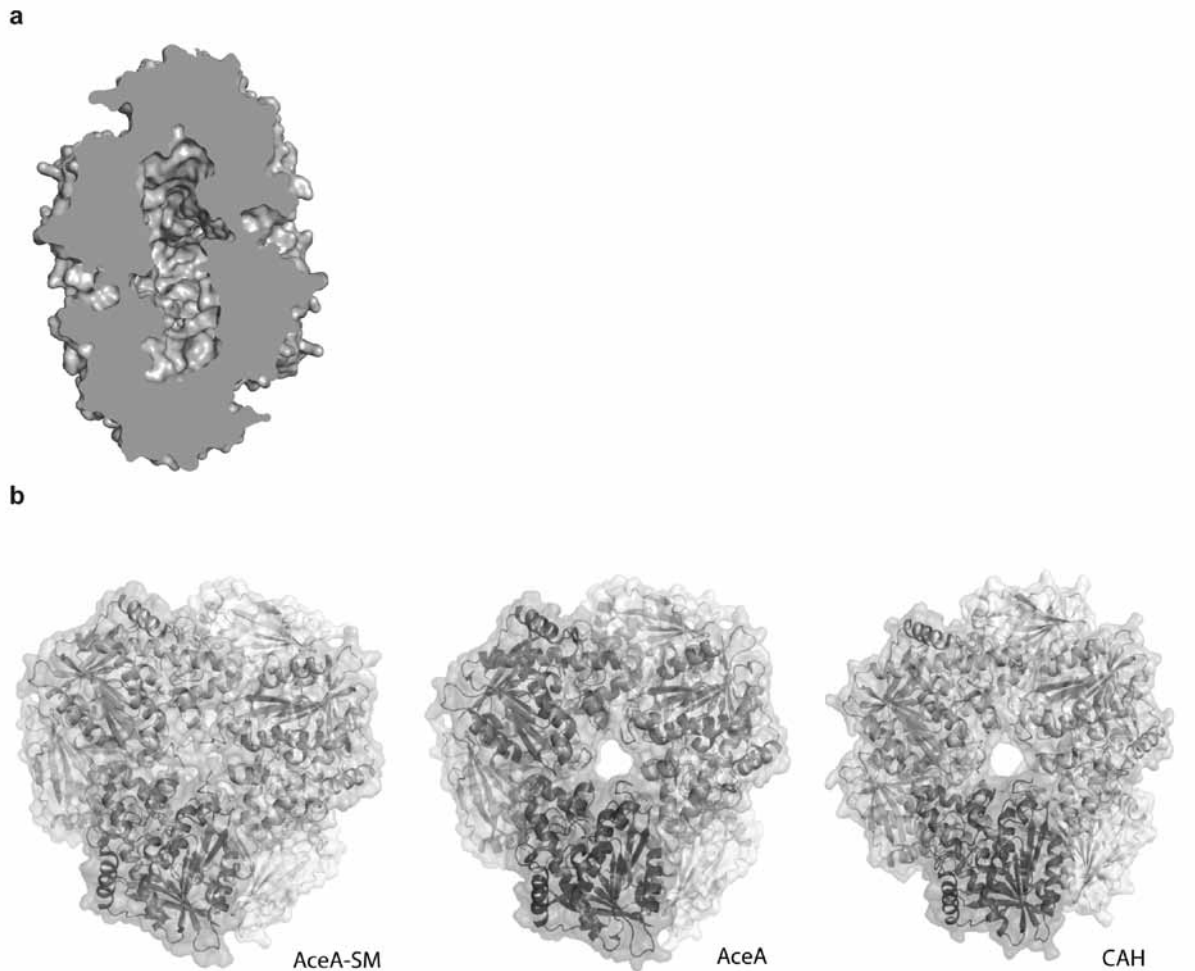
Figure 5.2 (in color on p.150): (a) A "sliced" image of the AceA hexamer showing the two entrances on each side of the 'doughnut' and the internal cavity, and (b) Surface representation of the biological unit of AceA-SM and AceA from *T. maritima* and the cephalosporin C deacetylase (CAH) from *B. subtilus* [183], with each monomer in a different color.

B is stabilized by on average 7 hydrogen bonds and has a buried surface area of 1024 Å^2. The interface between protomers A and F is stabilized by on average 17 hydrogen bonds and has a buried surface area of 1079 Å^2. Other interfaces are formed between protomers AC = AE = BD = BF = CE = DF, burying on average 514 Å^2. The multiple sequence alignment of AceA showed that the residues involved in the interfaces are not conserved. The hexamer has a total buried surface area of 18,860 Å^2, which is approximately 30% of the total surface area. On average, 3,143 Å^2 per monomer is buried upon complex formation.

Molecular surface representations of the AceA hexamer shows a 'doughnut'-like shape. The six active sites are located at the interior of the hexamer, where they line an oval-shaped cavity. Access to this cavity is via two entrances, one on each site of the hexamer (Figure 5.2a). The two entrances are approximately 13 Å wide and result in a short tunnel spanning approximately 10 Å to the inner cavity. This is similar in the hexamer of AceA and CAH [183]. Interestingly, the AceA-SM structure does not contain the two holes in the hexamer, due to very flexible N-terminal residues blocking the entrances (Figure 5.2b).

Comparison with related structures

A structural similarity search was performed using the program DALI [187]. Chain A of the AceA structure was used as a search model and similarity was found with cephalosporin C deacetylases, acylamino-releasing enzymes, dipeptidyl peptidases and some esterases and lipases. AceA is structurally most similar with the cephalosporin C deacetylase (CAH) from *B. subtilus* (PDB: 1ODS) [183] and with an acylpeptide hydrolase/esterase apAPH from *Aeropyrum pernix* K1 (PDB: 1VE6) [65]. The sequence identity between AceA and CAH after the structural alignment is 41% and the two structures are aligned with a Z-score of 46 and an RMSD of 1.5 Å over 312 C_α atoms. The sequence identity with apAPH is 17% with a Z-score of 23.3 and an RMSD of 2.3 Å over 230 C_α atoms. Superposition of AceA with CAH and apAPH is shown in Figure 5.3.
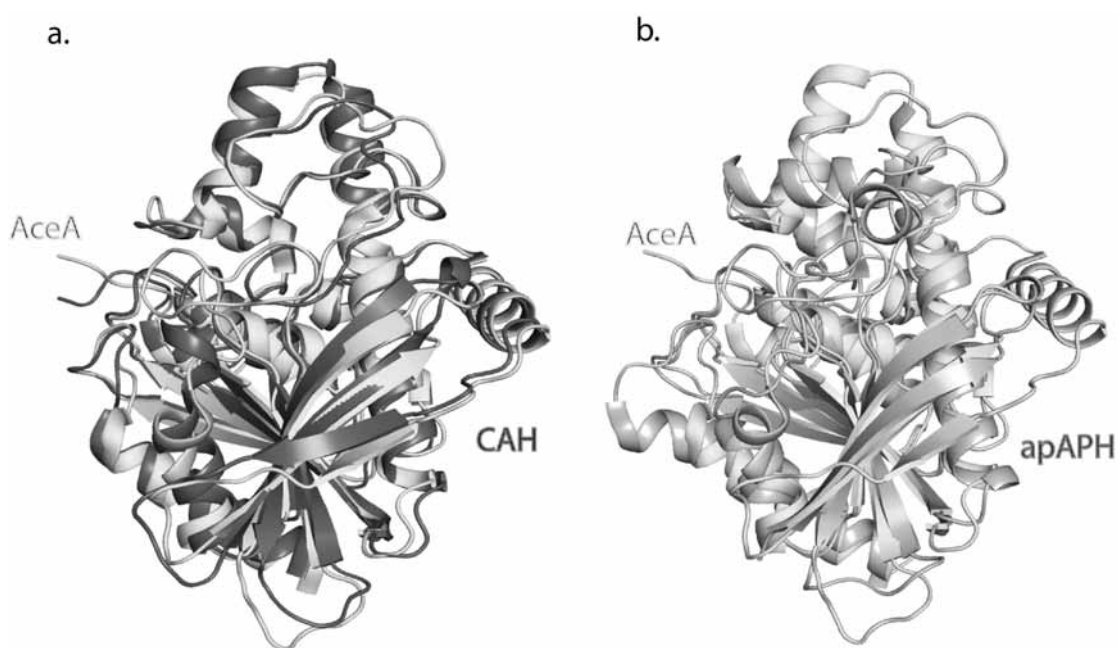
Figure 5.3 (in color on p.151): Superposition of AceA (*yellow*) with (a) the cephalosporin C deacetylase (CAH) from *B. subtilus* (PDB: 1ODS; *blue*) [183] and (b) the α/β-hydrolase domain of the acylpeptide hydrolase/esterase apAPH from *A. pernix* K1 (PDB: 1VE6; *light blue*) [65].

Inhibitor assays and structures of AceA complexed with PMSF and paraoxon

Inhibition assays showed that both diethyl *p*-nitrophenyl phosphate (paraoxon) and phenylmethylsulfonyl fluoride (PMSF) inhibit AceA activity with 1 and 45% residual activity, respectively. The inhibition kinetics for PMSF and paraoxon were investigated in the presence of *p*-nitrophenyl acetate, as described previously [68]. This resulted in dissociation and rate constants of 690 µM and 0.017 s^{-1} for PMSF, and 960 µM and 0.21 s^{-1} for paraoxon, respectively. These results are in good agreement with the inhibition assay. No significant stimulation

or reduction of activity of AceA was observed in the presence of divalent metal ions or ethylenediaminetetraacetic acid (EDTA).

In order to obtain more information about possible conformational changes occurring during catalysis, AceA was co-crystallized with the inhibitors PMSF and paraoxon. The electron density map of AceA in complex with PMSF showed clear density for the PMSF modification. The F-atom was cleaved from the PMSF molecule during the binding reaction and the phenylmethyl sulfonyl (PMS) moiety is covalently bound to the O^γ atom of Ser188. The native and PMS-bound structure superimpose with an RMSD of 0.114 Å. Electron density maps for the paraoxon co-crystallized crystals displayed clear density for a diethyl-phosphate (DEP) moiety covalently bound to the O^γ atom of Ser188. The *p*-nitrophenol group of paraoxon has been cleaved of during co-crystallization, thereby leaving a tetrahedral product reminiscent of the first transition state formed during carboxyl ester hydrolysis. The native and paraoxon-bound structure superimpose with an RMSD of 0.143 Å. Attempts to obtain co-crystallizations of AceA with the substrate cephalosporin C, even at a low temperature of 4 °C, were unsuccessful.

Analysis of the active site

The active site of AceA has a classic catalytic triad, consisting of Ser188 as the nucleophile, His303 as the proton acceptor/donor, and Asp274 as the acidic residue stabilizing the histidine. The catalytic serine Ser188 is located within a conserved pentapeptide sequence, Gly-X-Ser-X-Gly (GGSQG), characteristic for esterases and lipases. The positions of Ser188, Asp274, and His303 are consistent with the canonical fold of the α/β-hydrolase family. Ser188 is located at the nucleophile elbow, a sharp turn between β6 and helix αC. The presence of five glycine residues (Gly186, Gly187, Gly190, Gly191, Gly192) in close proximity to Ser188 prevents steric hindrance in the sharp turn of the nucleophile elbow. Asp274 and His303 are located at loops between β8 and helix αF, and between β9 and helix αG, respectively. The oxyanion hole is formed by the backbone amide groups of Tyr92 and Gln189.

In the native structure, the Ser188 side chain makes a hydrogen bond with the imidazole ring of His303. Extra density was observed near the side chain of Ser188 and was interpreted as a chloride ion on the basis of its electron density and geometry. The chloride ion is bound at the entrance of the oxyanion hole, forming hydrogen bonds with the backbone amides of the residues Tyr92 and Gln189 (Figure 5.4a). In the PMSF inhibited structure, the sulfonyl group of the inhibitor makes hydrogen bonds with the backbone amides of the residues Tyr92 and Gln189. The phenyl ring of the inhibitor is located in the small active site groove surrounded by the hydrophobic residues Tyr92, Trp124, Pro228, Ile276, and His303 (Figure 5.4b). In the paraoxon inhibited structure, the DEP moiety is stabilized by hydrogen-bonding interactions with the oxyanion hole. One of the two ethyl arms of bound paraoxon points toward the large pocket of the protein, while the other follows the groove of the small pocket. The two ethyl

arms are stabilized by packing against the side chains of Tyr92, Trp124, Pro228, Ile276, and His303 (Figure 5.4c).

The catalytic triad and oxyanion hole are located at one end of a surface depression. This pocket has a nearly oval shape and extends from the catalytic serine approximately 15 Å into one end of the enzyme and spans approximately 12 Å. A small pocket of approximately 5 Å long extends to the other side of the catalytic serine. The volume of both pockets together is 1082 Å^3 (CASTp analysis; [188]). The substrate-binding pocket is bordered by residues from the helices αB and αG, and its base is formed by residues from β-strands 5, 6, and 7. The large pocket consists of the following residues: Gln88, Tyr89, Ile90, Gly91, Asp101, Trp102, Phe104, Trp105, Cys112, Trp124, Leu125, Ala185, Gly186, Gly187, Ser188, Leu208, Asp210, Val211, Ser269, His303, Glu304, Gly305, Gly306, Gly307, Ser308, Gln310, Ala311, and Gln314. The small groove consists of the following residues: Tyr92, Ser188, Gln189, Phe213, and Pro228. The overall pocket is hydrophobic, although it does have some polar residues (Gln88, Asp210, and Gln314) which may interact with the substrate.
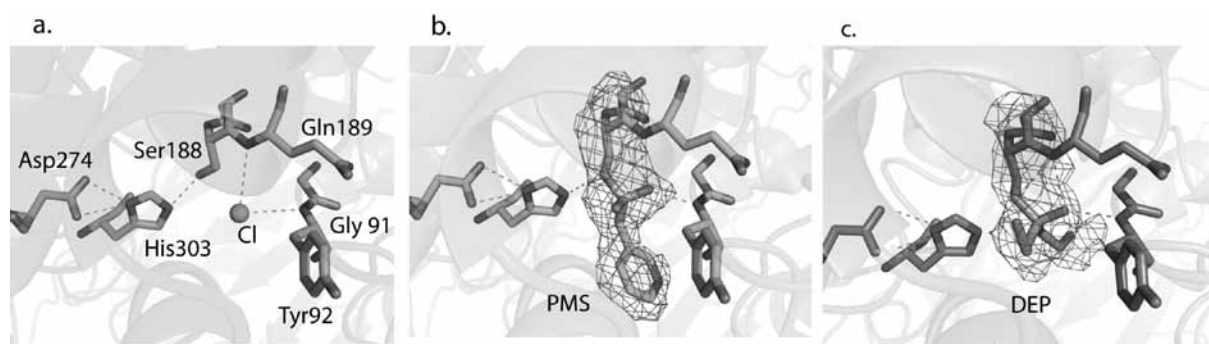
Figure 5.4 (in color on p.151): A view of the AceA catalytic site. (a) Native AceA with the bound chloride ion, (b) AceA in complex with PMSF and (c) AceA in complex with paraoxon. The catalytic residues are shown as sticks, with the hydrogen bonds shown as dashed lines. Green (AceA), cyan (AceA-PMS) or blue (AceA-DEP) were used for the carbon atoms, red was used for oxygen atoms. The omit maps for Ser188 with the phenylmethyl sulfonyl (PMS) and diethylphosphate (DEP) moieties are contoured at 1σ level.

Two conformations of the catalytic serine

Although no large conformational changes were observed upon binding of PMSF or paraoxon, compared to native AceA, a different conformation of the catalytic serine was observed (Figure 5.5). In the native AceA structure, the catalytic serine shows a conformation in which the Ser188 O$^\gamma$ is in the plane of the imidazole ring of His303, most commonly observed in the resting state of esterases and lipases. The Ser188 (O$^\gamma$) side chain makes a hydrogen bond with the imidazole ring (Nε2) of His303, which are 2.50 Å apart from each other (Conformation A). In the PMSF and paraoxon bound structures, the conformation of the catalytic serine changes in which the Ser188 O$^\gamma$ rotated over 110°, increasing the distance to the imidazole ring to 3.12 Å and 2.93 Å for PMSF and paraoxon bound structures, respectively (Conformation B). In the

AceA-SM structure, the catalytic serine was observed to be also in conformation B, with a distance to the imidazole ring of 3.04 Å. Extra electron density was identified in the AceA-SM structure, suggesting that there could exist a partially occupied acyl intermediate on Ser188. However, the density was not clear enough and, as a result, water molecules were modeled instead. The transition of conformation A to B seems to take place upon binding of substrate. No rearrangements of other residues of the active site were observed in the A to B transition.
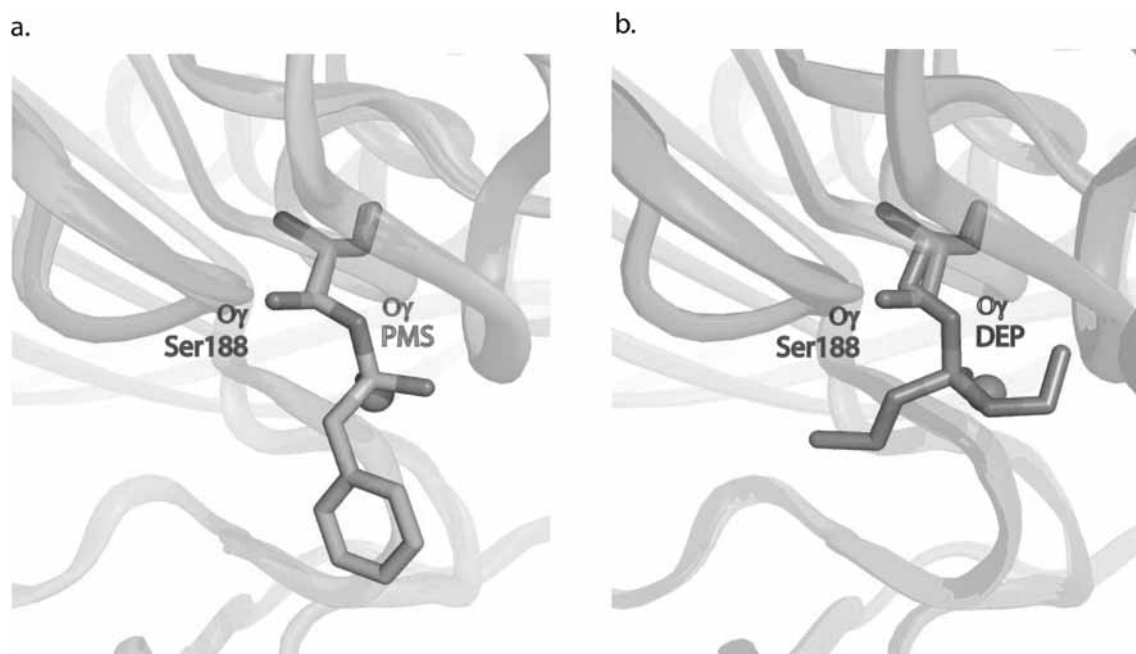
Figure 5.5 (in color on p.152): Movement of Ser188 O^γ from conformation A to B. The O^γ atom of the Ser188 is rotated about 110° to the opposite direction between native AceA and the complexed structures. Conformation A in AceA (green), and conformation B in AceA-PMS (cyan) and AceA-DEP (blue) are shown. Similar colors for each structure were used as in Figure 5.4.

Enzyme activity

Activity of AceA was investigated using *p*-nitrophenol esters with varying acyl-chain length, ranging from C2 to C18. However, AceA is only active on the short-chain *p*-nitrophenol esters of acetate and propionate and did not hydrolyze esters with acyl-chains longer than four carbons. The kinetic parameters for the hydrolysis of *p*-nitrophenyl C2 and C3 were determined (Table 5.1). The optimum substrate for an enzyme can be deduced from the catalytic efficiency (k_{cat}/K_m) of its conversion. According to this criterion both substrates are preferred equally.

The effect of temperature on activity was studied using *p*-nitrophenyl acetate as substrate. The esterase activity increased from 40°C upwards until 100°C (Figure 5.6a). An Arrhenius analysis resulted in a linear plot in the temperature ranges of 40-60°C and 60-100°C (6a; inset), with calculated activation energies for the formation of the enzyme substrate-enzyme complex of 33.7 and 21.9 kJ/mol, respectively. The nick in the Arrhenius

a.

b.

c.
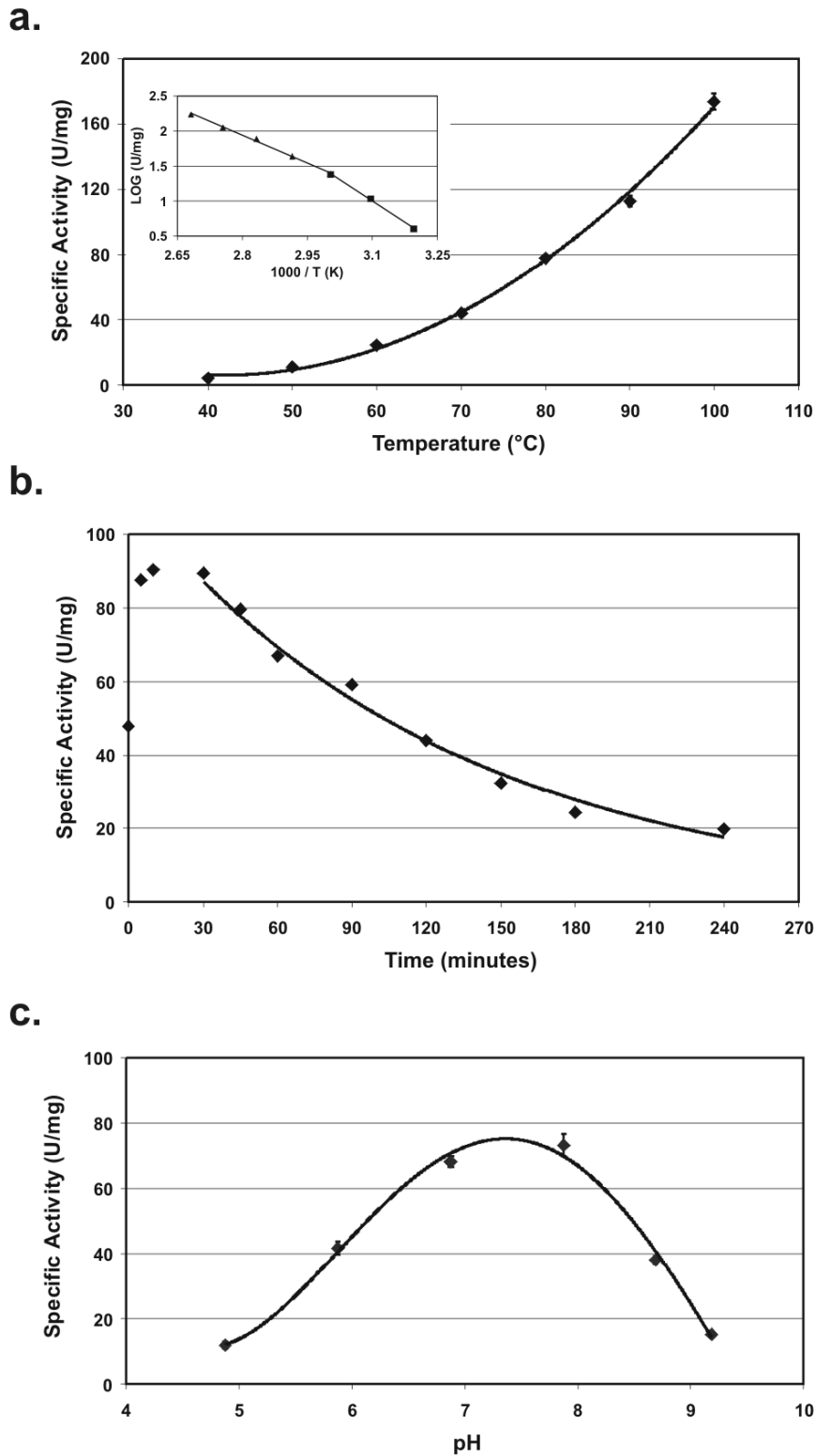
Figure 5.6: Effect of temperature and pH on esterase activity. (a) The effect of temperature on esterase activity was studied using *p*NP-C2 as a substrate at temperatures ranging from 40-100°C. The inset shows the temperature dependence as an Arrhenius plot. (b) Thermal stability of AceA at 90°C. (c) The effect of pH on esterase activity was studied using *p*NP-C2 as a substrate at pH values in the range of 4.8-9.2.

plot is probably caused by a conformational change. AceA has good resistance to thermal inactivation. An approximately 50% transient increase in activity is seen during the first 10 to 20 minutes when the enzyme is incubated at 90°C. After 30 minutes, inactivation of function occurs by first order kinetics with a half-life of approximately 120 minutes (Figure 5.6b). A transient activation has also been seen for other thermophilic esterases, such as the esterase from *Sulfolobus shibatae* [55], and it is believed that a high temperature is needed in order to obtain an optimal conformation. AceA was not stable at 100°C, resulting in a half-life value of less than 5 minutes. The effect of pH on activity was measured in the pH range of 4.8 to 9.2 using the substrate *p*-nitrophenyl acetate. AceA displayed maximum activity at approximately pH 7.5 (Figure 5.6c), which is comparable to other CE7 esterases such as the acetyl xylan esterases from *Thermoanaerobacterium* sp. strain JW/SL-YS485 [116].

Table 5.1: Kinetic parameters for hydrolysis of various esters

Ester	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (s^{-1} mM^{-1})
*p*NP-Acetate	0.185	57.5	310.8
*p*NP-Propionate	0.137	41.2	300.7
2-Ac *p*NP-Xylp	2.2	35.5	16.1
3-Ac *p*NP-Xylp	3.9	64.6	16.6
4-Ac *p*NP-Xylp	3.9	87.7	22.5

Positional specificity

The positional specificity of AceA was tested on three monoacetates of 4-nitrophenyl β-D-xylopyranoside. For that reason, the β-xylosidase XloA [182] (TM0076) from *T. maritima* was cloned. The thermostable enzyme was heterologously expressed, purified to homogeneity and activity of XloA was confirmed by measuring the release of *p*-nitrophenol from the substrate 4-nitrophenyl β-D-xylopyranoside (data not shown). The β-xylosidase was not active on the three monoacetates of 4-nitrophenyl β-D-xylopyranoside. In the XloA-coupled assay, AceA hydrolyzed acetate from positions 2, 3 and 4 of 4-nitrophenyl β-D-xylopyranoside at similar rates and also displayed an equal affinity for each substrate. The results are summarized in Table 5.1.

In addition, AceA was investigated for its ability to remove acetyl groups from the substrates 7-aminocephalosporanic acid (7-ACA), cephalosporin C, glucose pentaacetate, N-acetyl-D-glucosamine, xylan and acetylated xylan. AceA has no activity on acetylated and non-acetylated xylan polymers, indicating that it indeed is an acetyl esterase and not an acetyl xylan esterase. As expected from an acetyl esterase, AceA displayed high activity on glucose pentaacetate with a turnover number of 2680 s^{-1}. Like other members of CE7, AceA was also

able to hydrolyze the acetyl groups from both cephalosporin C and 7-ACA with a turnover number of 376 s^{-1} and 1140 s^{-1}, respectively. However, it is unlikely that both compounds are natural substrates, because the stability of these compounds at the growth temperature of *T. maritima* is very low. AceA was not able to hydrolyze the acetyl group from N-acetyl-D-glucosamine, indicating that it is specific for ester bonds and unable to hydrolyze amide bonds.

DISCUSSION

Deacetyl cephalosporins are valuable building blocks for the production of semisynthetic β-lactam antibiotics. These compounds are derived from cephalosporin C or 7-aminocephalosporanic acid via enzymatic or chemical processes [178]. Esterases and deacetylases belonging to the carbohydrate esterase family 7 are unusual in that they are active towards both acetylated xylo-oligosaccharides and the antibiotic cephalosporin C. CE7 family members have also been identified in genomes of thermophilic bacteria. Such thermostable esterases may be valuable in the preparation of derivatives of β-lactam antibiotics. One of these, a putative CE7 acetylxylan esterase (AceA) from the hyperthermophilic bacterium *T. maritima*, was functionally produced in *Escherichia coli* and purified to homogeneity.

The three-dimensional structure of AceA has been determined at 2.6 Å resolution and revealed a classic α/β-hydrolase fold with an insertion of three helices after β6 and an extension of the N-terminus with a β-strand and two helices. The crystal structure has a hexameric arrangement composed of three dimers. The hexamer has a "doughnut"-like shape, with two entrances on each side of the hexamer leading to an internal cavity where the six catalytic centers reside. The N-terminal extension and the inserted helices are important structural features, because they contribute to shaping the internal cavity and the two entrances. Interestingly, the two entrances are blocked in the AceA-SM structure due to a very flexible N-terminal peptide chain. The electron densities are also poor in this region. Analysis of the quaternary structure by gel filtration and static light scattering showed that AceA also exists as a hexamer in solution.

In the structure of AceA a calcium ion is bound at the beginning of helix αA. The calcium ion makes hydrogen bonds with Lys22, Glu26 and Asp25 via a water molecule. In each monomer of the hexamer a calcium ion is bound. However, no significant stimulation or reduction of activity of AceA was observed in the presence of calcium ions or EDTA. Therefore, it seems that these calcium ions are not important for activity. No calcium is present in the *B. subtilus* CAH structure [183], however, the residues Lys22, Glu26 and Ser25 are conserved and may act as an calcium binding site in this structure as well. Calcium is present in the crystallization buffer and could therefore be a crystallization adduct. On the other hand, calcium may help stabilize the structure.

Structures of AceA in complex with the inhibitors PMSF and paraoxon were obtained at 2.5 Å and 2.1 Å resolution, respectively. Upon binding of PMSF or paraoxon, the reaction mechanism of AceA is blocked at the acylation step via the formation of a tetrahedral intermediate. The negatively charged oxygen of the tetrahedral intermediate, derived from the substrate oxyanion, is stabilized by hydrogen bonds with the two backbone amide groups of Tyr92 and Gln189, resulting in the formation of the oxyanion hole. Remarkably, upon binding of both inhibitors, the catalytic serine appears to adopt an altered conformation. In the PMSF-bound and paraoxon-bound structures the Ser188 O^γ is rotated over 110°, compared to Ser188 O^γ in the native structure. It is speculated that this conformational change of the Ser188 O^γ in AceA is required for correct hydrogen bonding to the backbone amides of Tyr92 and Gln189, hence forming the oxyanion hole and thereby stabilizing the transition state intermediate. The oxyanion hole is an essential feature for catalysis in α/β-hydrolases [6; 8; 56]. Alternate conformations of the catalytic serine have also been observed in other esterase structures, such as the *Fusarium solani* cutinase [189], the *Penicillium purpurogenum* acetyl xylan esterase [190], the *Bacillus subtilis* lipase [191], and the *Aspergillus niger* feruloyl esterase [192]. However, in contrast to the above enzymes, the conformational change of the AceA catalytic serine seems to be solely induced by binding of an inhibitor. We expect that upon binding of a substrate, Ser188 will also adopt an altered conformation.

On artificial *p*-nitrophenyl-esters AceA activity was confined to short acyl chain esters (C2 and C3). AceA is active on a variety of acetylated compounds, with highest specific activity on glucose penta-acetate. No activity was detected on xylan or acetylated xylan. This implies that AceA should be classified as an acetyl esterase, and not as an acetyl xylan esterase [193]. In good agreement with this observation, the lack of any apparent signal sequence suggests that AceA is not a secreted protein. This predicted intracellular location of AceA is incompatible with a role in the deacetylation of xylan, an extracellular substrate. In addition, it is tempting to speculate that the tunnel entrances in the AceA hexamer may function as a filter, disallowing the access of large substrates to the active sites.

Three different monoacetates of 4-nitrophenyl-β-D-xylopyranoside were tested as substrates for AceA in a β-xylosidase-coupled assay. This assay enabled to examine the positional specificity of the enzyme. AceA hydrolyzed acetate at positions 2, 3 and 4 with the same efficiency. This is the first esterase from the CE7-family tested for its positional specificity. The esterase *Ct*Axe from *Clostridium thermocellum* belonging to CE4 shows a clear preference for hydrolyzing acetate at the C2 position of 4-nitrophenyl β-D-xylopyranoside [194]. The *Penicillium purpurogenum* AXE II esterase belonging to CE5 also has a preference for acetate at C2 of 4-nitrophenyl β-D-xylopyranoside [195]. The lack of preference for a specific position of the acetate group correlates with the broad substrate specificity for the CE7 esterases.

Based on the results of this study, the likely main biological function of AceA is removing the remaining acetyl groups on the short acetylated end products of xylan degradation that are transported into the cytoplasm, and thus allowing access of a β-xylosidase to the

xylose oligomer. Yet, it cannot be excluded that AceA will also act on other small acetylated compounds.

AceA has a relatively high specific activity on the substrates 7-ACA and cephalosporin C. This is approximately ten fold higher than the acetyl xylan esterase from *B. pumilus* [196] or the acetyl esterase from *Thermoanaerobacterium* sp. strain JW/SL YS485 [197]. AceA has a higher hydrolytic activity on 7-ACA compared to cephalosporin C, as has been described for other CE7 esterases [183; 196; 197]). The optimum temperature (≥100°C) and thermal stability ($t_{1/2}$ = 2h at 90°C) of AceA are considerably higher than those reported for other characterized CE7 esterases, such as the *Thermoanaerobacterium* enzyme that has a temperature optimum of 80°C and a half-life of 1h at 75°C [116]. Recently, the substrate specificity of the acetyl xylan esterase from *P. purpurogenum* was changed from a clear preference for acetate, to a range of fatty acid esters of up to at least 14 carbons [195]. In future it might be possible to engineer the structure of AceA as well, and enable the (de)acetylation of cephalosporins at the C10 position with various acyl chains. Altogether, the high stability and activity and the relatively broad specificity of AceA on 7-ACA and cephalosporin C suggests a possible application in the production of new semi-synthetic antibiotics.

MATERIALS AND METHODS

Gene cloning

The gene encoding TM0077 (GenBank: AAD35171.1, GI:4980565; SwissProt: Q9WXT2) was amplified by polymerase chain reaction (PCR) from genomic DNA using *PfuTurbo* DNA polymerase (Stratagene) and primers corresponding to the predicted 5' and 3' ends. The PCR product was cloned into plasmid pMH1, which encodes an expression and purification tag (MGSDKIHHHHHH) at the amino terminus of the protein. The cloning junctions were confirmed by DNA sequencing.

AceA-SM protein production and purification

Protein production was performed in a selenomethionine-containing medium using the *Escherichia coli* methionine auxotrophic strain DL41. At the end of fermentation, lysozyme was added to the culture to a final concentration of 1 mg/mL, and the cells were harvested. After one freeze/thaw cycle, the cells were sonicated in Lysis Buffer [50 mM Tris pH 7.9, 50 mM NaCl, 1 mM $MgCl_2$, 0.25 mM Tris(2-carboxyethyl)phosphine hydrochloride (TCEP), and the lysate was centrifuged at 3,400 x g for one hour. The soluble fraction was applied to nickel-chelating resin (GE Healthcare) pre-equilibrated with Equilibration Buffer [50 mM potassium phosphate pH 7.8, 300 mM NaCl, 10% (v/v) glycerol, 0.25 mM TCEP] containing 20 mM imidazole, the resin was washed with Equilibration Buffer containing 40 mM imidazole, and the protein was eluted with Elution Buffer [20 mM Tris pH 7.9, 300 mM imidazole, 10% (v/v) glycerol, 0.25 mM

TCEP]. The eluate was buffer exchanged with Buffer Q [20 mM Tris pH 7.9, 5% (v/v) glycerol, 0.25 mM TCEP] containing 50 mM NaCl and applied to a RESOURCE Q column (GE Healthcare) pre-equilibrated with the same buffer. The protein was eluted using a linear gradient of 50 to 500 mM NaCl in Buffer Q and purified further with a HiLoad 16/60 Superdex 200 column (GE Healthcare), using Crystallization Buffer [20 mM Tris pH 7.9, 150 mM NaCl, 0.25 mM TCEP] as the mobile phase. For crystallization assays the peak Superdex 200 fractions were concentrated to ~15 mg/mL by centrifugal ultrafiltration (Millipore). Molecular weight and oligomeric state of AceA were determined using a 1 cm × 30 cm Superdex 200 column (GE Healthcare) in combination with static light scattering (Wyatt Technology). The mobile phase consisted of 20 mM Tris pH 8.0, 150 mM NaCl, and 0.02% (w/v) sodium azide.

Native AceA production and purification

For protein production, cells were grown in LB medium for 8 hours (an OD600 well above 2.0 was reached). Subsequently, the culture was induced by adding 0.15% L-arabinose and incubated another 16 hours at 37°C. Cells were harvested by centrifugation at 10,000 x g for 20 min. The cell pellet was resuspended in 30 ml of Lysis Buffer 2 [50 mM Tris-HCl pH 8.0, 50 mM NaCl, 10 mM imidazole, 0.25 mM TCEP]. The cells were disrupted by two passages through a French press at 110 MPa. The crude cell extract was treated with DNAse I at room temperature for 30 min and subsequently centrifuged at 43,000 x g for 30 min in order to remove cell debris. The supernatant was heated at 70°C for 25 min and then centrifuged to remove the precipitated proteins. The supernatant was filtered and loaded onto a nickel-chelating column packed with 20 ml of Ni-NTA His-Bind Resin (Novagen) and equilibrated in 50 mM Tris-HCl pH 8.0, 300 mM NaCl, 2% (v/v) glycerol, and 0.25 mM TCEP. The column was washed with 20 mM imidazole in the same buffer, and proteins were subsequently eluted with a linear gradient of 20-500 mM imidazole in the same buffer. Fractions containing esterase activity were pooled and loaded onto a HiPrep Desalting column (GE Healthcare) equilibrated with 20 mM Tris-HCl pH 8.0, 150 mM NaCl, and 0.25 mM TCEP. The homogeneity of the protein was checked by SDS-PAGE, and activity staining of the SDS-PAGE gel was performed using α-naphtyl acetate, as described previously [52]. The protein concentration was determined at 280 nm using a NanoDrop ND-1000 Spectrophotometer.

Crystallization

Crystals of selenomethionine-substituted AceA were obtained by hanging drop vapor diffusion against a 250 µl reservoir solution consisting of 20% PEG-300, 10% glycerol, 0.1 M phosphate-citrate pH 4.2, and 0.2 M $(NH_4)_2SO_4$ pH 4.5, at 20°C. Drops consisted of 0.5 µl protein and 0.5 µl reservoir.

Native AceA was crystallized using nano-drop vapour diffusion techniques against a crystallization reagent consisting of 0.2 M calcium acetate hydrate, 20% (w/v) PEG 3350, pH 7.3 at 20°C. Protein was concentrated to 22.8 mg/ml and 7-aminodesacetoxycephalosporanic

acid added to a molar ratio of 1:3 (protein:ligand). Drops consisted of 100 nl protein and 100 nl of crystallization reagent and a 60 µl reservoir of crystallization reagent. Crystals of AceA in complex with its inhibitors PMSF and paraoxon were obtained at 4°C in the same condition as the native crystals. AceA protein and PMSF / paraoxon were added to a molar ratio of 1:3 (protein:inhibitor).

Data collection

For cryoprotection, glycerol was added to the AceA-SM crystal to a final concentration of 15% (v/v). The crystal was mounted in a cryoloop and subsequently flash-frozen in liquid nitrogen. X-ray data were collected at 100 K on beamline BL9-2 at the SSRL (Stanford) using a Quantum 210 CCD detector (ADSC). A AceA-SM data set was collected to 2.1 Å resolution and the data indexed in space group $P2_1$ (Table 5.2), with unit cell parameters a = 152.6 Å, b = 130.95 Å, and c = 157.82 Å. There are twelve molecules in the asymmetric unit. Data were indexed and integrated with DENZO and then scaled with the program SCALEPACK [198].

Native AceA crystals were transferred to crystallization reagent complimented with 10% (v/v) ethylene glycol and flash-cooled to 100K and data collected at beamline 5.0.3 of the ALS. All data were processed with the HKL2000 package [198]. The native data set was collected to 2.5 Å resolution and the data indexed in space group $P2_12_12_1$ with unit cell parameters a = 103.5 Å, b = 103.8 Å, and c = 221.0 Å. (Table 5.2). There are six molecules in the asymmetric unit.

Structure solution and refinement

The AceA-SM structure was solved by MAD phasing using two wavelength MAD datasets. For initial phasing, SHELXD was used to find candidate SeMet substract sites. Attempts to complete phasing were unsuccessful due to the translational non-crystallographic symmetry (NCS) (not recognized initially). Self-consistent sets (partial sets) were found using the CCP4 program PROFESSS and a few more SeMet sites were found by the SHELXD program, and added into these partial sets. The AutoSHARP run did not complete because of an over-simplistic view of NCS. The solution was achieved, but a reasonable electron density map was not obtained. Trials of ARP/wARP did not succeed initially. However, the SHARP run was completed. Updates of the ARP/wARP program, concerning docking of side chains, were used in the chain tracing. The structure was refined with BUSTER using tight NCS to an R/R$_{free}$ of 24.2/24.7. The continuous model building was performed using O [199] and the structure was further refined using CNS [165] and REFMAC5 [200]. Refinement statistics are summarized in Table 5.2. The final model contains residues 2-323 (A, B, C, D, E, F, G, H, I, J, K and L chains) in the asymmetric unit. A main chain torsion angle analysis using MolProbity [201] showed that 97.4% and 100% of all residues are in favored and allowed regions of the Ramachandran plot, respectively. GlnB120, which is on the border of the Ramachandran outlier, is supported by electron density as well as the NCS model.

Table 5.2: Summary of crystal parameters, data collection, and refinement statistics

	AceA-SM (λ₁ MADSe)	AceA-SM (λ₂ MADSe)	AceA native	AceA-PMS	AceA-DEP
Space group	$P2_1$		$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters	a=152.64Å b=130.95Å c=157.82Å α=90.00° β=118.90° γ=90.00°		a=103.46Å b=103.79Å c=221.02Å α=90.00° β=90.00° γ=90.00°	a=103.57Å b=104.50Å c=221.61Å α=90.00° β=90.00° γ=90.00°	a=103.80Å b=104.43Å c=221.64Å α=90.00° β=90.00° γ=90.00°
Data collection	λ₁ MADSe	λ₂ MADSe			
Wavelength (Å)	0.9791	0.9183	0.9765	0.9765	0.9765
Resolution range (Å)	67.20-2.10	67.20-2.10	48.795-2.500	48.970-2.400	48.970-2.120
Number of observations	558,708	559,681	1222016	765546	989949
Number of unique reflections	236,781	237,982 (13,660)	83045	94681	123070
Completeness (%)	92.6 (66.7)[a]	92.4 (65.7)[a]	100.0 (100.0)	100.0 (100.0)	89.8 (53.5)
Mean I/σ (I)	7.2 (2.5)[a]	9.0 (2.3)	14.4 (2.9)	11.5 (3.4)	15.3 (2.2)
R_{sym} on I (%)	0.139 (0.598)	0.135 (0.649)	20.700 (109.700)	18.000 (67.400)	9.500 (51.900)
Highest resolution shell (Å)	2.21-2.10	2.21-2.10	2.56-2.50	2.46-2.40	2.18-2.12
Model and refinement statistics					
Resolution range (Å)	29.56-2.10		48.795-2.500	48.970-2.400	48.970-2.120
Data set used in refinement	λ₁ MADSe		first column of data	first column of data	first column of data
Cutoff criteria	|F| > 0		|F| > 0	|F| > 0	|F| > 0
R_{cryst}	0.186		0.165	0.159	0.167
R_{free}	0.223		0.215	0.215	0.209
No. of reflections (total)	278,371		83045	94680	122994
No. of reflections (test)	14,726[b]		4200	4742	6188
Completeness (% total)	92.8		100.0	100.0	89.8
Stereochemical parameters					
Restraints (RMS observed)					
Bond angle	1.48		1.516	1.564	1.431
Bond length	0.018		0.017	0.017	0.014
Average isotropic B-value	27.89		24.64	19.21	19.58
ESU based on R_{free}	0.170		0.248	0.225	0.177
Water molecules / ligands	2,464 / 1		479 / 17	877 / 11	919 / 11
PDB ID	1VLQ				

[a] Highest resolution shell. ESU = Estimated overall coordinate error [141; 205]

$R_{sym} = \sum_i |I_i - <I_i>| / \sum_i |I_i|$ where I_i is the scaled intensity of the ith measurement and $<I_i>$ is the mean intensity for that reflection.

$R_{cryst} = \sum |F_{obs}| - |F_{calc}|| / \sum |F_{obs}|$ where F_{calc} and F_{obs} are the calculated and observed structure factor amplitudes, respectively.

R_{free} = as for R_{cryst}, but for 5.0 % of the total reflections chosen at random and omitted from refinement.

[b] Typically, the number of unique reflections used in refinement is slightly less than the total number that were integrated and scaled. Reflections are excluded due to systematic absences, negative intensities, and rounding errors in the resolution limits and cell parameters.

The native AceA structure was solved by molecular replacement using PHASER [202; 203] with the AceA-SM coordinates (pdb: 1VLQ) as a search model. Six molecules were successfully located and the structure was further refined with Refmac [200] using tight NCS to an R/R$_{free}$ of 0.181/0.215. Iterative cycles of refinement and building were performed with Refmac5, Phenix [166; 204] and Coot [164]. All other crystallographic manipulations were carried out with the CCP4 package [141]. Refinement statistics are summarized in Table 5.2. The final model contains residues 3-324 (A, B, C, D and F chains) and 3-323 (E chain) in the asymmetric unit. A main chain torsion angle analysis using MolProbity [201] showed that 97.1% and 100% of all residues are in favored and allowed regions of the Ramachandran plot, respectively. Glu134 and Gly135 of B chain, and Asn302 of B, C and D chains are located on the border line of the Ramachandran outlier. The total secondary structure content in terms of α-helical, 3$_{10}$-helical, and β-strand elements is 33.2, 2.2, and 18.9%, respectively. The model backbone as well as side-chain geometries fully conform to standard protein stereo-chemical parameters.

Enzyme assays

Esterase activity using *p*-nitrophenyl esters was measured as described previously [52]. In short, in a standard assay, activity was measured with 0.2 mM *p*-nitrophenyl acetate as substrate in 50 mM citrate-phosphate (pH 6) 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. Kinetic parameters were determined by fitting the data, obtained from multiple measurements, with a computer-aided direct fit to the Michaelis–Menten curve (Tablecurve 2d, version 5.0).

The effect of pH on esterase activity was studied in the pH range of 5 to 10. The buffers used were 50 mM citrate-phosphate (pH 5-8) and 50 mM CAPS (3-(cyclohexylamino) 1-propanesulphonic acid) buffer (pH 9.5-10). The pH of the buffers was set at room temperature, and temperature corrections were made using their temperature coefficients: -0.0028 pH/°C for citrate-phosphate buffer and -0.018 pH/°C for CAPS buffer. The effect of temperature on esterase activity was studied in the range of 40-100°C using 0.2 mM *p*-nitrophenyl acetate as substrate. Enzyme thermostability was determined by incubating the enzyme in a 50 mM Tris-HCl, 150 mM NaCl (pH 7.8) buffer at 90°C and 100°C for various time intervals. Residual activity was assayed in a standard assay.

The effect of inhibitors was studied by pre-incubating AceA with 1 mM inhibitor in 50 mM citrate-phosphate buffer (pH 8) at 37°C for 1 hour. Subsequently, samples were placed on ice and residual activity was measured using the standard assay. Activity of the enzyme without inhibitor was defined as 100%. Inhibition kinetics of PMSF and paraoxon were determined as described for the acetylcholinesterase from electric eel [150]. Deacetylase activity was determined using high-performance liquid chromatography (HPLC) by measuring the amount of acetic acid released from the substrates cephalosporin C, 7-aminocephalosporanic

acid, and glucose-pentaacetate. 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 mN H_2SO_4 and 30 mM crotonate) and placing the sample on ice. The conditions for HPLC were as follows: column, KC811 Shodex; detection, RI and UV detectors; solvent, 3 mN H_2SO_4; flow rate, 1.5 ml/min; temperature, 30°C; internal standard, crotonate. One unit of enzyme activity was defined as the amount of enzyme that releases one µmol of acetic acid per minute.

Xylan was acetylated by the method described by Johnson [206]. Activity on xylan was measured quantitatively using DMSO extracted xylan (1% polysaccharide solution in 0.1 M sodium phosphate buffer pH 6) at 60°C [207].

Enzyme-coupled assay

The positional specificity of AceA was investigated using an enzyme-coupled assay on monoacetylated 4-nitrophenyl β-D-xylopyranosides as described by Biely *et al.*, 2004 [208]. The β-xylosidase XloA (locus tag: TM0076) from *T. maritima* was cloned into the vector pET24d in frame with a C-terminal 6xHis tag. The enzyme was expressed and purified as described above for native AceA. Activity of XloA was confirmed by measuring the release of *p*-nitrophenol at 405 nm from the substrate 4-nitrophenyl β-D-xylopyranoside (NPh-Xyl).

The enzyme-coupled assay was performed at 60°C in a total volume of 125 µl, which contained 0.1 M sodium phosphate (pH 6 or 7), 2-*O*-, 3-*O*-, or 4-*O*-acetyl NPh-Xyl, the β-xylosidase XloA, and AceA. Stable 50 times-concentrated stock solutions of the substrates were prepared in DMSO. The reaction was started by the addition of 2.5 µl of a stock solution to a preheated reaction mixture consisting of phosphate buffer, auxiliary β-xylosidase XloA (~0.6 nkat), and AceA. The reaction was terminated by the addition of 800 µl of a 2% solution of Na_2CO_3. Liberated *p*-nitrophenol was determined at 405 nm against substrate and enzyme blanks. A short incubation time for activity determination was used to suppress acetyl migration on the xylopyranosyl-ring, which is significant at pH 6 or 7 [209]. The kinetic constants were determined at pH 7 and 60°C with a reaction time of 2 minutes.

Structure validation and deposition

Analysis of the stereochemical quality of the model was accomplished using AutoDepInputTool [210], MolProbity [201], SFcheck 4.0 [211], and WHATIF 5.0 [125]. Protein quaternary structure was analyzed using the PQS server [185]. Figures were prepared with PyMOL (DeLano Scientific) [134]. The structural data of AceA-SM has been deposited in the RCSB Protein Data Bank with accession code 1VLQ. The structural data of native AceA and AceA complexed with PMSF and paraoxon will be deposited after a final Quality Control check (QC-check) at the JCSG.

ACKNOWLEDGMENTS

6

Chapter

Purification and partial characterization of a thermostable esterase from *Thermotoga maritima*

Levisson, M., Gielen, P., Deller, M.C., Kengen, S.W.M. & van der Oost, J.
Manuscript in preparation.

ABSTRACT

A bioinformatics analysis of the genome of *T. maritima* revealed the presence of several genes potentially encoding esterases and lipases, including one (*est*B, TM0053) that was annotated as a putative esterase. The esterase encoding gene was cloned and functionally expressed in *E. coli*. EstB was found to exhibit esterase activity with a preference for medium acyl chain esters (C8-C10). The enzyme is a hexamer in solution. EstB showed optimal activity around 95°C and at pH 9. Crystals were obtained and a dataset was collected to ~2.8 Å. The structure solution is ongoing.

INTRODUCTION

Esterases and lipases belong to a diverse group of hydrolases with representatives in all domains of life. They catalyze the hydrolysis (or synthesis) of ester bonds, resulting in the formation of an alcohol and a carboxylic acid. Most esterases and lipases belong to the α/β-hydrolase family and share structural and functional characteristics. In general, they have a co-factor independent activity, an α/β-hydrolase structural fold, and a conserved catalytic triad that is usually composed of a serine acting as the nucleophile, a histidine as the proton acceptor/ donor, and an aspartate or glutamate as the acidic residue stabilizing the histidine [5; 6; 7; 8].

Esterases differ from lipases in that they show a preference for short chain acyl esters (≤C10), whereas lipases show a preference for long chain acyl esters (≥C10) and are active on water-insoluble substrates [9]. It is noteworthy that lipases are often also capable of hydrolyzing the esterase substrates [7]. In industry, esterases and lipases are applied in various processes, such as the stereospecific hydrolysis of drugs and ester synthesis for food ingredients [3; 5; 10; 11]. In particular, enzymes from thermophilic origin are potentially interesting for industrial applications, since most of them can withstand elevated temperatures and show an increased stability in organic solvents compared to enzymes of mesophilic origin [12].

Thermotoga maritima is a hyperthermophilic bacterium growing optimally at 80°C and is able to metabolize many simple and complex carbohydrates, including glucose, starch, and xylan [142]. A bioinformatics analysis of the genome of *T. maritima* revealed the presence of several genes potentially encoding esterases and lipases, including one (*est*B, TM0053) that was annotated as a putative esterase. In this paper, the identification, cloning, expression, and partial characterization of the esterase EstB from *T. maritima* is reported. This enzyme has been described in another study and was assigned the function of carboxylesterase [71]. The results of the current study, however, revealed several different characteristics.

RESULTS

Identification and *in silico* analysis

A bioinformatic analysis of the genome of *T. maritima*, led to the identification of several open reading frames potentially encoding new thermostable esterases and lipases, including one (*est*B, TM0053) that was annotated as a putative esterase. The gene encodes a protein of 364 amino acids and has a calculated molecular mass of 40.8 kDa. N-terminal sequence analysis using the SignalP 3.0 Server revealed that the first 22 amino acids form a signal peptide.

BLAST-P analysis, using the EstB amino acid sequence as a template, showed that EstB has highest similarity with other putative esterases and hypothetical proteins belonging to the *Thermotogales* family. A comparison of EstB with the amino acid sequences of these

Figure 6.1: Amino acid sequence alignment EstB. Multiple sequence alignment of EstB with the most significant BLAST hits: *Thermotoga sp.* RQ2 (TRQ2_0893), *Thermotoga petrophila* RKU-1 (Tpet_0871), *Marinitoga piezophila* KA3 (MPKA3_1323), *Thermotoga neapolitana* DSM 4359 (CTN_0641), *Thermosipho melanesiensis* BI429 (Tmel_1912), *Fervidobacterium nodosum* Rt17-B1 (Fnod_1333), *Thermosipho africanus* TCF52B (THA_206), and the lipase from *Pseudomonas glumae* (1TAH). The probable catalytic triad, consisting of Ser162, Asp309 and His330, is marked with an asterisk.

most significant hits in the BLAST-search, as well as with the sequence of the triacylglycerol lipase from *Pseudomonas glumae* (pdb: 1TAH) [212], identified three amino acids that potentially constitute the characteristic catalytic triad (Ser162, Asp309, and His330) (Figure 6.1). The potential catalytic serine lies within a conserved pentapeptide motif A/G-X-S-X-G (AHSMG), which is called the nucleophile elbow and is typical for esterases and lipases. The levels of similarity with other hits in the BLAST search were much lower and were restricted to the region around the catalytic serine.

Cloning and purification

The gene encoding EstB was amplified by PCR, without the predicted signal peptide (the first 22 amino acids) and without its stopcodon (fused to a His_{6x} tag), using chromosomal DNA of *T. maritima* as a template. The gene was cloned into the expression vector pET-24d (designated pWUR351), and subsequently expressed in *E. coli* BL21(DE3)/pSJS1244. The enzyme EstB was purified to homogeneity from heat-treated cell extracts by immobilized metal affinity chromatography. The homogeneity of the protein was checked by SDS-PAGE and activity staining of the SDS-PAGE gels using α-naphtyl-acetate (Figure 6.2). On the SDS-PAGE and activity stained gels a main band corresponding to the mature EstB monomer (38.4 kDa) was visible, however, also other bands belonging to multimers of EstB were visible. Native-PAGE showed a single band on gel indicating that EstB is present in one quaternary conformation. Size exclusion chromatography showed that EstB eluted as a single peak corresponding to an estimated mass of 253 kDa. This suggests that EstB is most probably present as a hexamer.
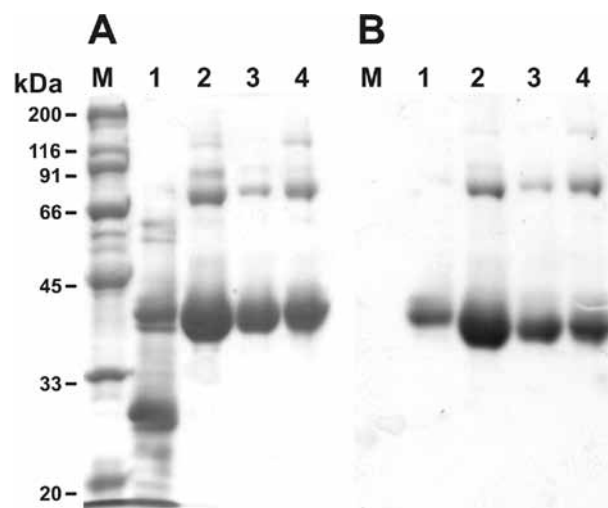

Figure 6.2: SDS-PAGE of EstB fractions. Samples were separated by SDS-PAGE in duplicate. One gel was stained with Coomassie Brilliant Blue (A) and the other was stained for activity using α-naphtyl acetate after renaturation (B). Lane M, relative molecular mass standards; lane 1, heat-stable cell free extract; lane 2, EstB after immobilized metal affinity chromatography; lane 3, two-times diluted EstB after immobilized metal affinity chromatography; lane 4, purified EstB. Multiple bands are visible corresponding to EstB multimers. Both the monomer (38.4 kDa) and multimers (76.8 and 115.2 kDa) of EstB are believed to be catalytically active.
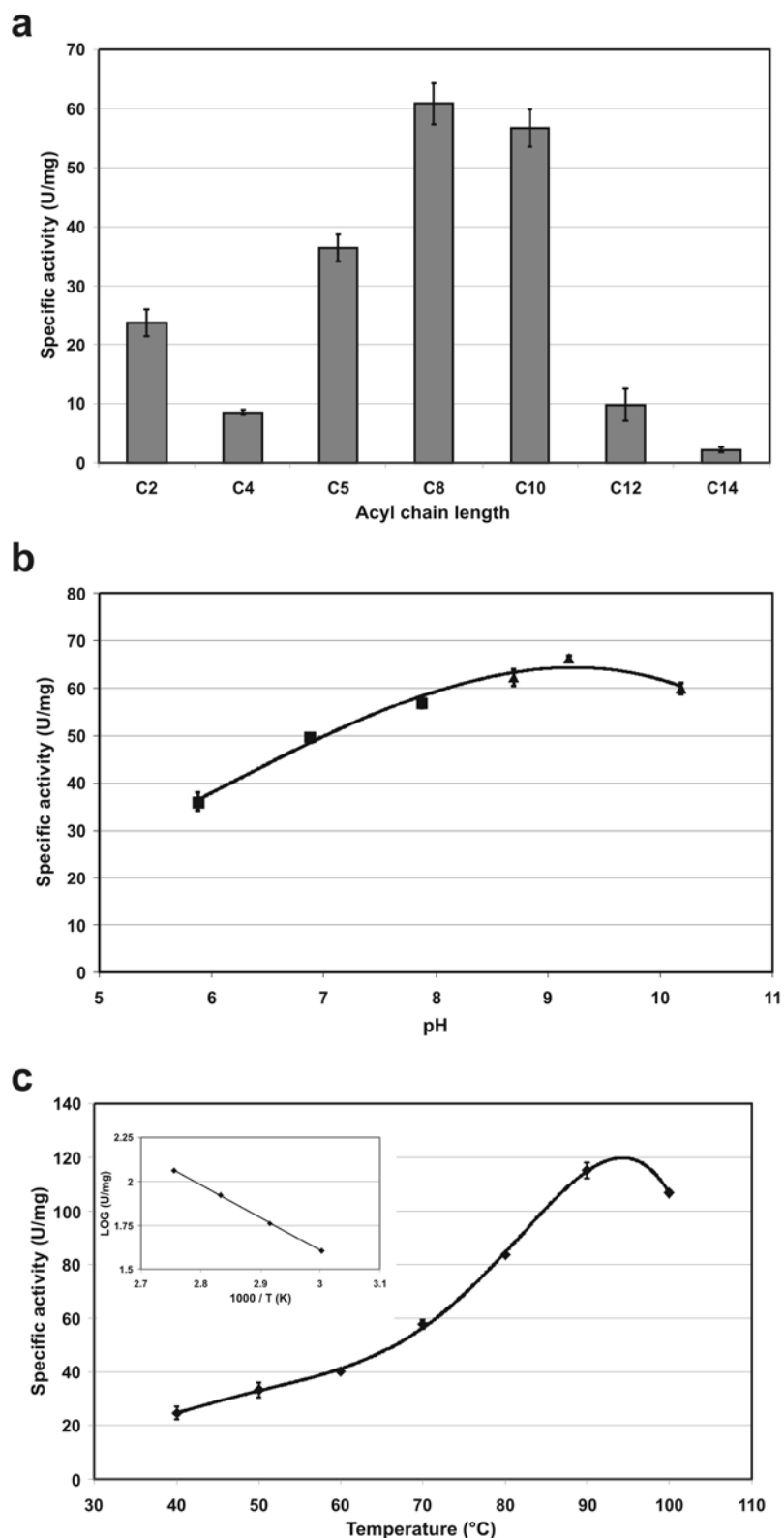
Figure 6.3: Substrate preference and effect of temperature and pH on activity. (a) Substrate preference, (b) The effect of pH on esterase activity was studied using *p*-nitrophenyl octanoate (citrate-phosphate buffer (■) and CAPS buffer (▲)) as substrate at pH values in the range of 5.8-10.2. (c) The effect of temperature on esterase activity was studied using *p*-nitrophenyl octanoate as substrate at temperatures ranging from 40-100°C. Inset: the temperature dependence for *p*-nitrophenyl octanoate as an Arrhenius plot.

Biochemical properties

The substrate specificity of purified EstB was investigated using *p*-nitrophenol (*p*NP) esters with acyl chains of different lengths, ranging from C2 to C18 (Figure 6.3a). The highest hydrolytic activity was measured with the substrates *p*NP-C8 and *p*NP-C10. Very low activity (≤ 1 U/ mg) was measured with the two long acyl chain esters *p*NP-C16 and C18. Based on these results, EstB is not a lipase but an esterase.

The effect of pH on esterase activity was determined at 70°C using the substrate *p*NP-C8. The activity was measured over the pH range of pH 4.8 to 10.3. EstB showed highest activity at approximately pH 9 (Figure 6.3b). No hydrolytic activity could be measured at a pH value of 4.8.

The effect of temperature on esterase activity was determined with a citrate-phosphate buffer (pH 8) and substrate *p*NP-C8. The activity was measured in the range of 40 to 100°C. Activity of EstB increased with temperature and the highest hydrolytic activity was measured at approximately 95 °C (Figure 6.3c). An Arrhenius analysis resulted in a linear plot in the temperature range of 60-90°C (Figure 6.3c, inset), with a calculated activation energy for the formation of the enzyme/substrate complex of 15 kJ/mol. The thermostability of EstB was determined by measuring the residual activity after incubation at 100°C at different time points. EstB has no resistance to thermal inactivation at 100°C (half-life ≤ 1 min).

Preliminary crystallization results

Purified EstB was used for crystallization trials, at the Joint Center for Structural Genomics (JCSG) automated pipeline for structure determination [180], testing 400 conditions. Crystals of native EstB were obtained using a reservoir solution consisting of 3.4 M 1,6 Hexanediol; 0.2 M MgCl2; 0.1 M Tris pH 8.5 at 20°C. A dataset was collected to ~2.8 Å resolution. The crystal belongs to spacegroup P1 and molecular replacement is ongoing. However, at this moment the models look too dissimilar for solution and therefore seleno-methionine incorporated EstB may be required to allow multi-wavelength anomalous dispersion (MAD).

DISCUSSION

In this paper the cloning, expression and partial characterization of an esterase from the hyperthermophilic bacterium *T. maritima* is described. During a bioinformatic analysis of the genome of *T. maritima* the gene *TM0053* was identified as a putative esterase. Sequence analysis allowed the identification of a probable catalytic triad, Ser162, Asp309 and His330. The recombinant protein was heterologously expressed in *E. coli* to fairly high levels. Purification

of EstB to homogeneity required only a few steps (heat treatment, Ni-chelating column and a desalting column).

When the recombinant enzyme was assayed using *p*-nitrophenyl esters, it exhibited highest activity towards *p*-nitrophenyl octanoate (C8) and decanoate (C10), which is comparable to esterases from other hyperthermophiles, viz. *Aeropyrum pernix* [66] and *Sulfolobus solfataricus* [50]. Low activities were observed for esters with acyl chain lengths of C14 or higher, suggesting that the enzyme is an esterase with specificity for substrates of medium chain length.

EstB showed a high temperature optimum around 95°C, which is comparable to several other esterases of *T. maritima* [50; 52; 68] and other hyperthermophiles such as the esterase from *Pyrobaculum calidifontis* [38], the esterase from *Sulfolobus shibatae* [55], the esterase from *Sulfolobus solfataricus* [40] and an esterase from a metagenomic library [25]. Unlike, the esterases EstA and EstD of *T. maritima* [52; 68], EstB was not stable at 100°C. However, it is expected that EstB will display good stability around 80°C, which is the optimal growth temperature of *T. maritima* [142].

The esterase showed maximal activity around pH 9, with approximately 55% and 90% of the maximal activity at pH 5.8 and 10.2, respectively. This broad pH spectrum suggests that the enzyme could be utilized in various applications. A high pH optimum has also been observed for other esterases from hyperthermophiles, such as the esterase from *Thermoanaerobacter tengcongensis* [72] and an esterase from a metagenomic library [26].

In a previous study by Kakugawa *et al.* (2006) [71], EstB was characterized as well, but some of the reported features differed considerably from our current findings. The enzyme showed a lower temperature optimum (60°C instead of 95°C) and also the pH optimum differed (pH 7.5 instead of pH 9). There are small differences between both studies in respectively the cloning, purification and assay methods, which may result in the differences observed. In the earlier study, the signal peptide of EstB was not completely removed during cloning. The first 14 N-terminal amino acid residues were removed during cloning, instead of the first 22 amino acid residues in this study.

Determination of the structure of EstB may provide insight into the function of this enzyme and reveal the molecular basis of substrate recognition and catalysis by this enzyme. Therefore, crystallization trials for EstB have been initiated. Crystal were obtained in this screen and resulted in a dataset to ~2.8 Å. Structure determination is ongoing.

As to the physiological role of this esterase, one can only speculate. For most characterized esterases a specific function has not been attributed. Most probably, the enzyme is involved in the extracellular breakdown of esterified fatty acids, which would agree with the presence of a signal sequence and the broad pH range.

MATERIALS AND METHODS

Data mining

The genome of *T. maritima* MSB8 [119] was screened for potential esterases and lipases. Sequences coding for esterases and lipases were identified by performing BLAST-P searches with sequences from characterized esterases and lipases at the National Center for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/blast/) [28]. Multiple sequence alignments were constructed using the Tcoffee Expresso(3DCoffee) server (http://www.tcoffee.org/) [132; 213] and manually adjusted using the program GeneDoc. The N-terminal sequence analysis of the translational product of TM0053 was performed using the SignalP 3.0 Server (http://www.cbs. dtu.dk/services/SignalP/) [214].

Cloning

The gene encoding EstB (locus tag: TM0053; GenBank: AAD35147) was PCR-amplified, without the sequence encoding its predicted signal peptide (the first 22 amino acids) and without its stop codon, using chromosomal DNA of *T. maritima* as a template and the two primers BG1964 5'-GCGCCATGGAATCTACACCCACACTCTACGATGTGGTGG-3' and BG1965 5'-GCGGCGGCCGCTTGCGATCCCCCCTTC-3' introducing respectively a *Nco*I and *Not*I restriction site. The generated PCR product was digested by *Nco*I and *Not*I and the product was purified and ligated into pET-24d (Novagen) digested with the same restriction enzymes, resulting in the plasmid pWUR351. The construct was designed with a 6xhis-tag at the C-terminus of the enzyme to facilitate purification. The sequence of the expression clone was confirmed by sequence analysis of both DNA strands.

Expression and purification

E. coli BL21(DE3)/pSJS1244 was transformed with pWUR351. A single colony was used to inoculate 4 ml Luria-Bertani medium containing kanamycin and spectinomycin (both 50 µg / ml) and incubated overnight at 37°C with shaking. The preculture was used to inoculate (1:1000) 1L of the same medium and growth was continued for 8 h (an OD_{600} above 2.0 was reached). Subsequently, the culture was induced by adding IPTG (isopropyl β-D-1-thiogalactopyranoside) to a final concentration of 0.5 mM. The culture was incubated for a further 16 h at 37°C. Cells were harvested by centrifugation at 10,000xg for 15 min. The cell pellet was resuspended in 30 ml lysis buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 10 mM imidazole). The cells were disrupted by passing twice through a French press at 110 MPa. The crude cell extract was treated with DNAseI at room temperature for 30 min and subsequently centrifuged at 43,000xg for 20 min in order to remove cell debris. The supernatant was heated at 70°C for 15 min and then centrifuged at 43,000xg for 20 min in order to remove the precipitated proteins. The supernatant was filtered and applied to a nickel-chelating column packed with 20 ml Ni-NTA His-Bind Resin (Novagen) and equilibrated in 50 mM Tris-HCl buffer pH 7.5 containing 300 mM

NaCl. The column was washed with 20 mM imidazole in the same buffer and proteins were subsequently eluted with a linear gradient of 20-500 mM imidazole in 50 mM Tris-HCl pH 7.5, 300 mM NaCl. Fractions containing esterase activity were pooled and applied onto a Hi-Prep desalting column (GE Healthcare) equilibrated with 50 mM Tris-HCl buffer pH 7.5. The homogeneity of the protein was checked by SDS-PAGE and activity staining of the SDS-PAGE gels using α-naphtyl acetate, as described previously [52].

Size exclusion chromatography

The molecular mass of the purified esterase was determined by size exclusion chromatography on a Superdex 200 high-resolution 10/30 column (24 mL) (GE Healthcare) equilibrated in 50 mM Tris/HCl (pH 7.5) containing 100 mM NaCl. Two hundred microliters of enzyme solution (1.4 mg/ml) in 50 mM Tris-HCl buffer pH 7.5 was loaded at a flow rate of 0.6 mL·min^{-1} onto the column and fractions (0.5 mL) were collected. Proteins used for calibration were blue dextran 2000 (> 2000 kDa), ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa).

Enzyme assays

Esterase activity was determined by measuring the amount of *p*-nitrophenol released during enzymatic hydrolysis of different *p*-nitrophenyl esters. The release of *p*-nitrophenol was continuously monitored at 405 nm using a Hitachi UV2001 spectrophotometer with a temperature controlled cuvette holder. Unless otherwise indicated, in a standard assay, esterase activity was measured with 0.2 mM *p*-nitrophenyl octanoate (*p*NP-C8) as a substrate in 50 mM citrate-phosphate buffer (pH 8.0) containing 1% isopropanol at 70 °C. Stock solutions of *p*-nitrophenyl esters were prepared by dissolving substrates in isopropanol. After preincubation, the reaction was started by adding enzyme to the reaction mix. One unit of esterase activity was defined as the amount of protein releasing 1 μmol·min^{-1} of *p*-nitrophenol from *p*NP-C8. Measurements were corrected for background hydrolysis in the absence of enzyme. Measurements were carried out at least three times and the molar extinction coefficient of *p*-nitrophenol was determined for every condition prior to each measurement. Activity was determined from the initial rate of the hydrolysis reaction. The protein concentration was measured at 280 nm using a NanoDrop ND-1000 Spectrophotometer.

Acyl chain length preference

Substrate specificity of the enzyme towards the acyl chain length of different *p*-nitrophenyl esters was investigated by using *p*-nitrophenyl acetate (C2), *p*-nitrophenyl butyrate (C4), *p*-nitrophenyl valerate (C5), *p*-nitrophenyl octanoate (C8), *p*-nitrophenyl decanoate (C10), *p*-nitrophenyl dodecanoate (C12), *p*-nitrophenyl myristate (C14), *p*-nitrophenyl palmitate (C16) and *p*-nitrophenyl stearate (C18) in the standard assay.

Effect of pH and temperature

The effect of pH on esterase activity was studied by measuring activities on *p*-nitrophenyl octanoate for a pH range of 5-11. The buffers used were 50 mm citrate-phosphate (pH 5.0-8.0) and 50 mm CAPS buffer (pH 9.5-11.0). 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 the phosphate buffer, and -0.018 pH·°C^{-1} for CAPS buffer). The effect of temperature on esterase activity was studied in the range 40-95 °C using *p*-nitrophenyl octanoate in the standard assay. Enzyme thermostability was determined by incubating the enzyme in a 50 mm Tris-HCl (pH 7.5) buffer at 100 °C for various time intervals. Residual activity was assayed under the standard condition.

Crystallization and data collection

Purified EstB in 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 0.25 mM tris(2-carboxyethyl) phosphine (TCEP) was used for crystallization trials at the Joint Center for Structural Genomics (JCSG) automated pipeline for structure determination. 400 JCSG crystallization conditions were screened against samples containing ligands (PMSF, capric acid, paraoxon and apo) at 20°C and 4°C. Protein was concentrated to ~18 mg/ml in crystallization buffer (20 mM Tris-HCl, pH 7.8 and 150 mM NaCl) and ligand added to a final concentration of ~4.7 mM (molar ratio 10:1, ligand:protein). Crystals of EstB were obtained by hanging drop vapor diffusion against a 60 µl reservoir solution consisting of 3.4 M 1,6 Hexanediol; 0.2 M MgCl2; 0.1 M Tris-HCl pH 8.5 at 20°C. Drops consisted of 150 nl protein and 150 nl reservoir. A native dataset was collected to ~2.8 Å resolution. Crystals were flash-cooled to 100K and data collected at beamline 5.0.3 of the ALS. All data were processed with the HKL2000 package [198] and all other crystallographic manipulations were carried out with the CCP4 package [141]. The crystal belongs to spacegroup P1 with unit cell dimensions of a = 97.346 Å, b = 97.680 Å, c =98.752 Å, a=91.625°, b=102.794° and g=119.563°. These dimensions suggest 9 molecules are present in the asymmetric unit with a solvent content of 45.8% and a Matthews coefficient of 2.3 A^3/Da. Molecular replacement is ongoing, however, thus far the models look too dissimilar for solution and therefore seleno-methionine incorporated EstB may be needed.

ACKNOWLEDGEMENT

Characterization of a lipase from the hyperthermophilic archaeon *Archaeoglobus fulgidus*

Levisson, M., Hendriks, S., Siemerink, M.A.J., Deller, M.C., Kengen, S.W.M. & van der Oost, J. *Manuscript in preparation*.

ABSTRACT

The genome sequence of *A. fulgidus* suggested the presence of a gene potentially encoding a lipase (*lipA*, AF1753). The lipase encoding gene was cloned and functionally expressed in *E. coli*. LipA was found to exhibit lipase activity with a preference for medium acyl chain esters (C10). Qualitative plate assays showed that LipA has highest activity on the triacylglycerol esters of octanoate and palmitate. LipA showed optimal activity around 95°C and at pH 11. A structural model was constructed using the lipase A from *Bacillus subtilus* as template. The model indicated a non-canonical catalytic triad, consisting of Ser136, Asp163 and His210, with the aspartate located after strand β6. Crystals of LipA were obtained and a dataset was collected to ~2.6 Å. The structure solution is ongoing.

INTRODUCTION

Lipases (EC 3.1.1.3), like esterases, belong to the family of carboxylic ester hydrolases [7]. In the presence of water they catalyze the hydrolysis of an ester into its corresponding alcohol and carboxylic acid. However, in an organic solvent they can catalyze the ester-synthesis or the transesterification reaction [4]. Esterases are defined as enzymes that catalyze the hydrolysis of short chain fatty acid esters (<10 carbon atoms) and lipases are defined as enzymes that catalyze the hydrolysis of longer chain fatty acid esters (>10 carbon atoms) and are active on micellar substrates [9]. Both esterases and lipases share structural and functional characteristics, such as the α/β-hydrolase fold, a co-factor independent activity, and a conserved catalytic triad usually consisting of a nucleophilic serine in a GxSxG pentapeptide motif, and an acidic residue (aspartate or glutamate) that is hydrogen bonded to a histidine [6; 8; 56].

Lipases are frequently applied in industry, e.g. in organic synthesis, biodiesel production and medical biotechnology [2]. These processes are often operated at elevated temperatures or in the presence of organic solvents, which are conditions that are detrimental to most enzymes. In this respect, enzymes from thermophilic origin are promising because they can withstand elevated temperatures and display a high intrinsic stability in organic solvents [12]. Hence, it would be of interest to search for lipases in hyperthermophilic bacteria and archaea. Most microorganisms living at elevated temperatures belong to the domain of the archaea, however, to our knowledge no lipase has been cloned and characterized from archaea thus far. Recently, it was shown that several species of halophilic archaea exhibit lipase activity, but none of the enzymes responsible for this activity was purified and characterized [215; 216].

Archaeoglobus fulgidus is a hyperthermophilic sulphate-reducing archaeon isolated from hydrothermal areas near Vulcano (Italy) [217]. The organism grows optimally at 83°C and can grow on various carbon sources such as fatty acids, amino acids, organic acids, and CO [218; 219]. Its genome has been sequenced [219], and the annotation suggested the presence of a putative lipase (*lipA*, AF1763). To confirm the anticipated function of LipA, the corresponding gene was cloned and heterologously expressed in *Escherichia coli*, and its biochemical properties were investigated. This enzyme has been described in another study and was assigned the function of carboxylesterase [93]. The results in the present study, however, revealed different characteristics, amongst others showing that LipA is a true lipase.

RESULTS

In silico analysis

The entire genome of *A. fulgidus* was analyzed for the presence of new esterases and lipases. The genome sequence revealed the presence of an ORF (AF1763) that was annotated as a

putative lipase. It has a conserved pentapeptide sequence (GxSxG), which is typical for esterases and lipases, and is predicted to have an α/β-hydrolase fold. The gene encodes a 474-amino-acid protein with a calculated molecular mass of 52.8 kDa.

A homology search at JGI IMG showed that LipA has highest similarity with two putative lipases from respectively *Frankia alni* ACN14a (32% identity) and *Frankia sp.* CcI3 (30% identity). Furthermore, similarity was found with a conserved hypothetical protein from *Aurantimonas sp.* SI85-9A1 (26% identity) and with lipases from *Bacillus pumilus* (24% identity) and *B. subtilus* (23% identity). A BLAST search of LipA at NCBI against the non-redundant database confirmed these findings. Similarity to other sequences in the database is limited to the N-terminal part (amino acids 1-229) of LipA. This N-terminal part is predicted to have the α/β-hydrolase fold and has highest similarity to various putative esterases and lipases. The C-terminal part (AA 230-474) of LipA only shares similarity with the two sequences of the *Frankia* species. It is has no homology with other sequences and is therefore believed to be a new domain.

Active site analysis

A multiple sequence alignment of LipA with the two putative lipases from *Frankia* species and several *Bacillus* lipases showed that the active site of LipA is potentially composed of Ser136, which is located within a pentapeptide motif (Gly-Xaa-Ser-Xaa-Gly), and His210 (Figure 7.1). Both the serine and histidine residue are not conserved in the sequences from *Frankia*, suggesting that these enzymes are probably not esterases or lipases. The alignment was not conclusive about the acidic residue involved in LipA catalysis. Potential candidates are Asp163, Asp169, Glu172, Glu193 and Glu194.

In the absence of a crystal structure, a three-dimensional model of LipA was built using the 3D-structural threading program PHYRE [220]. The program was able to construct a model only for the N-terminal part of LipA (Asp21-Thr234) (Figure 7.2a). The structure of lipase A from *Bacillus subtilus* [191] (pdb: 1ISP) was used as a template to build the model. The model was constructed with an E-value of 3.2e-13 and an estimated precision of 100%. No good model could be obtained for the C-terminal part of LipA, suggesting it could be a new domain. The model showed that Ser136 and His210 are both located in close proximity of each other, and are at positions according to the canonical α/β-hydrolase fold. Ser136 is centrally located within a nucleophilic elbow after the carboxyl end of the fifth strand of the central β-sheet and His210 is located on a loop between the eighth β-strand and helix α8. There is, however, no acidic residue located at the "usual" position after the seventh β-strand. An acidic residue (Asp163) is located after strand six and is in close proximity of His210 and Ser136 (Figure 7.2b).
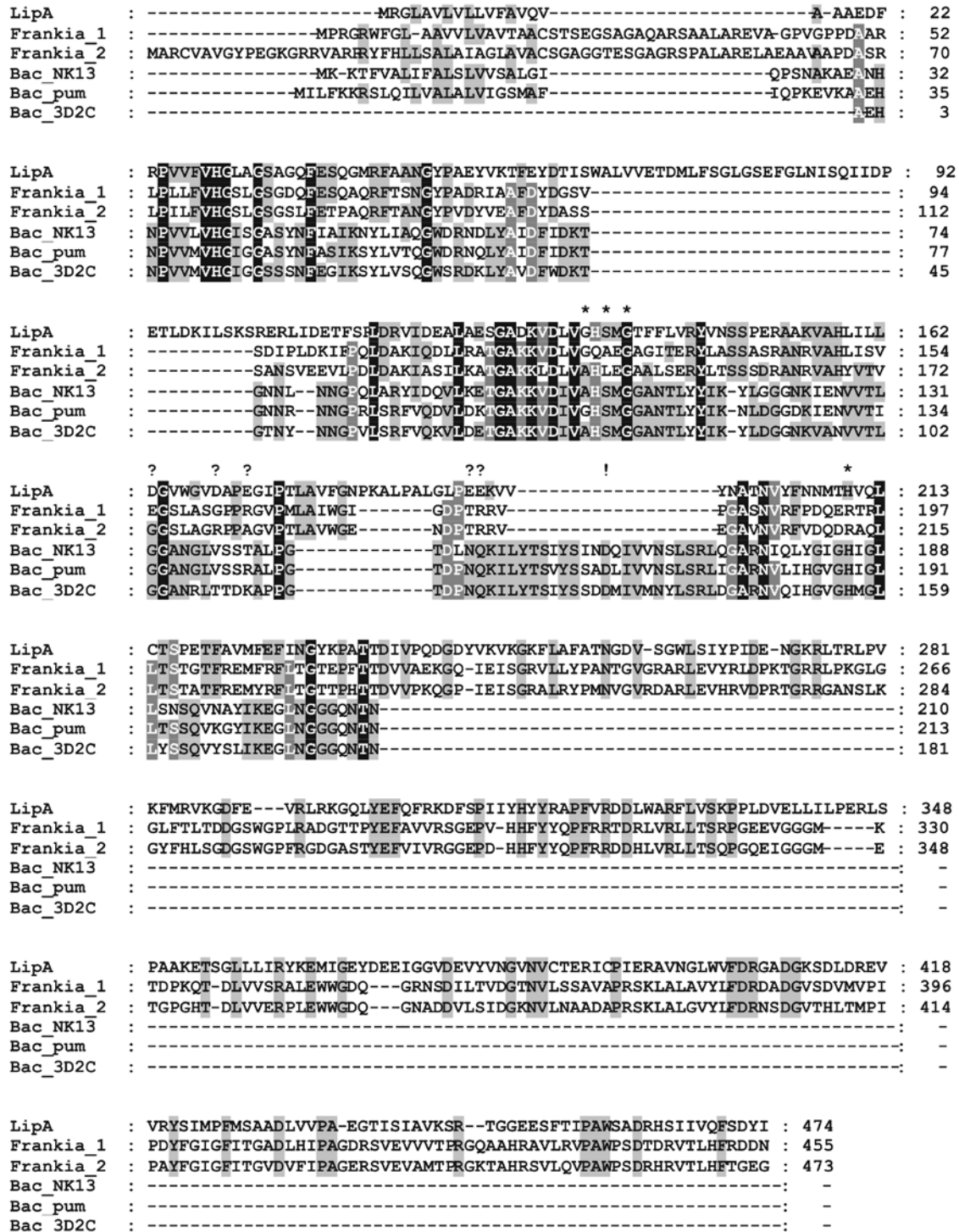
Figure 7.1: Amino acid sequence alignment LipA. Multiple sequence alignment of LipA with the first BLAST hits: *Frankia alni* ACN14a (Frankia 1), *Frankia sp.* CcI3 (Frankia 2), *Bacillus sp.* NK13 (Bac_NK13), *Bacillus pumilus* (Bac_pum) and the lipase from *Bacillus subtilis* (Bac_3D2R). The GXSXG motif containing the catalytic serine and the histidine that is part of the catalytic triad are marked with an asterisk. The catalytic acidic residue from the *Bacillus* enzymes is indicated with an exclamation mark, and the potential candidates in LipA (Asp163, Asp169, Glu172, Glu193 and Glu194) are indicated with a question mark.

Cloning and production

The N-terminal sequence of LipA was analyzed using SignalP version 3.0 and it was predicted that the first 20 amino acids form a signal peptide. The *LipA* gene was cloned into the expression vector pET-24d without its signal peptide and in frame with a C-terminal 6xHis-tag (pWUR364). The enzyme was purified to homogeneity from heat-treated cell extracts of *E. coli* BL21(DE3)/pSJS1244/pWUR364 by immobilized metal affinity chromatography. SDS-PAGE of purified LipA was performed to confirm homogeneity of the protein and to determine the molecular mass of the recombinant protein. Results obtained from SDS-PAGE confirmed a molecular subunit mass of 51 kDa (mature enzyme). Furthermore, it showed that the purification with metal affinity chromatography was efficient (Figure 7.3a). Activity staining of the SDS-PAGE gels using α-naphtyl acetate confirmed the identity of the LipA band (Figure 7.3b). Native-PAGE showed a single band and size exclusion chromatography suggested that the enzyme exists as a monomer with an estimated mass of 50 kDa (not shown).
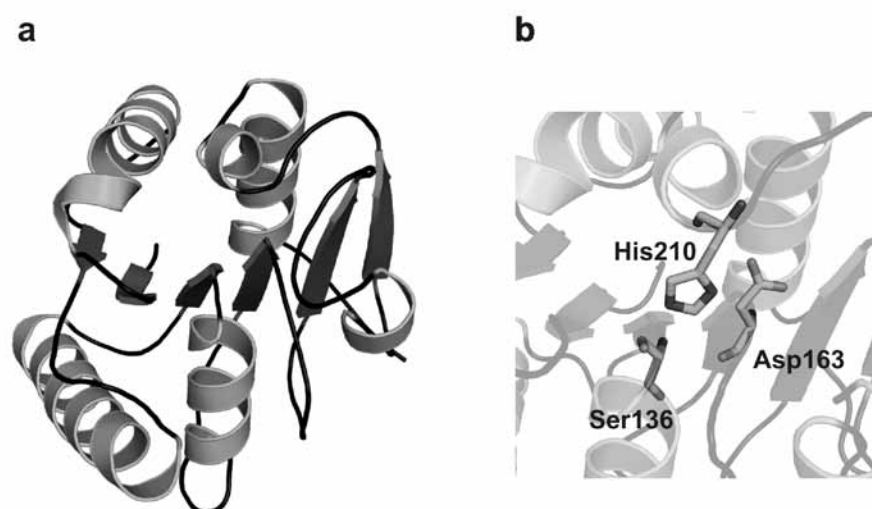

Figure 7.2 (in color on p.152): 3D model of LipA. In a) The overall structure of the N-terminal domain of LipA, and b) Residues of the potential catalytic triad.

Enzyme activity

Substrate specificity of purified LipA was determined using *p*-nitrophenyl esters. LipA showed hydrolytic activity towards most of the *p*-nitrophenyl esters tested (C3-C18) (Figure 7.4a). The highest specific activity with LipA was found for the *p*-nitrophenyl ester of decanoate (C10). No activity was found for the short chain *p*-nitrophenyl ester of acetate (C2) and lowest activity was found towards the *p*-nitrophenyl esters of propianate (C3), butyrate (C4), and the long chain *p*-nitrophenyl ester of stearate (C18).

A qualitative plate assay was done in order to determine activity of LipA on triacylglycerol esters. LipA showed activity towards tributyrate, trioctanoate, palm oil, olive oil and triolein. Substrate hydrolysis causes the formation of a halo around the spot where enzyme was added to the plate. The largest and clearest halos were observed for the substrates trioctanoate and palm oil.
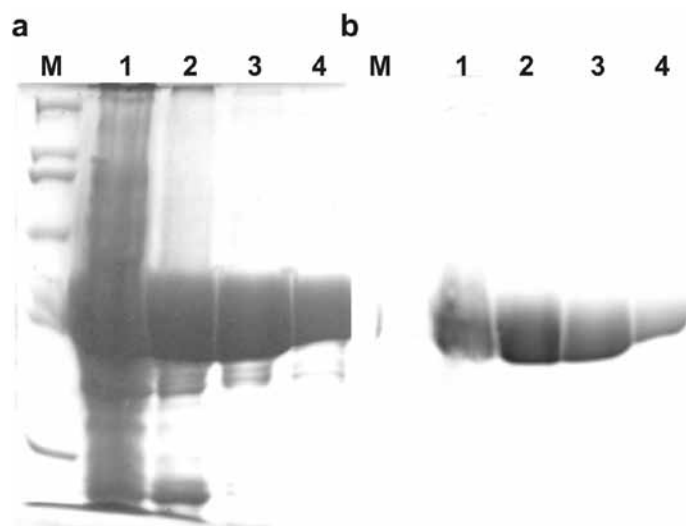

Figure 7.3: SDS-PAGE of LipA fractions. Samples were separated by SDS-PAGE in duplicate. One gel was stained with Coomassie brilliant blue (a) and the other was stained for activity using α-naphtyl acetate after renaturation (b). M: molecular weight standards, lane 1: cell free extract, lane 2: heat-stable cell free extract, lane 3: LipA after immobilized metal affinity chromatography and lane 4: purified LipA.

Influence of temperature and pH

The effect of temperature on the activity of LipA was investigated over a temperature range from 40 to 100°C using *p*-nitrophenyl octanoate as a substrate (Figure 7.4b). The hydrolytic activity increased from 40°C upwards with an optimum at 90°C. An Arrhenius analysis resulted in a linear plot in the temperature range of 60-90°C (not shown), with a calculated activation energy for the formation of the enzyme/substrate complex of 16.5 kJ·mol^{-1}. LipA was found to have a high resistance to thermal inactivation, with half-life values of approximately 20 minutes at 90°C and 10 hours at 80°C.

The effect of pH on the activity of LipA was investigated over a pH range from 4.8 to 12 using *p*-nitrophenyl myristate (C14) as a substrate (Figure 7.4c). The substrate *p*NP-C14 was used because of its stability at higher pH values. Maximal activity of LipA was measured at approximately pH 11. LipA was active over the complete range of pH values examined.
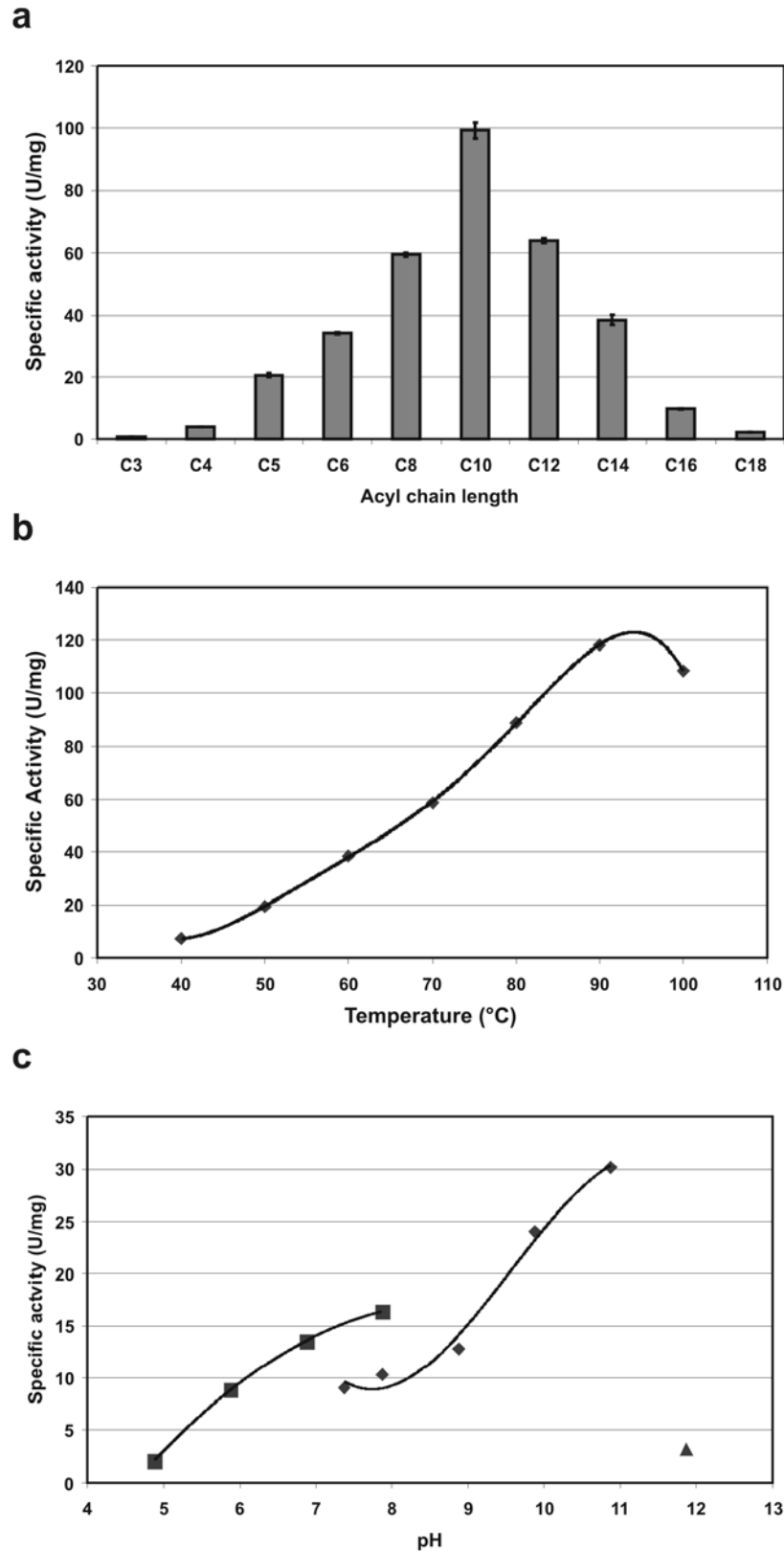
Figure 7.4: Biochemical properties of LipA. (a) Substrate preference, (b) The effect of temperature on esterase activity was studied using *p*-nitrophenyl octanoate as substrate at temperatures ranging from 40-100°C, and (c) The effect of pH on esterase activity was studied using *p*-nitrophenyl myristate as substrate at pH values in the range of 4.8-11.8.

Effect of inhibitors and divalent metal ions

The effect of various inhibitors on LipA activity was investigated. Phenylmethylsulfonyl fluoride (PMSF), a serine protease inhibitor, strongly inhibited enzyme activity, with no residual activity remaining in the presence of 1 mM inhibitor. Paraoxon inhibited LipA to a lower extent with less than 40% residual activity remaining in the presence of 1 mM inhibitor and no activity remaining in the presence of 10 mM. Activity was also strongly inhibited by mercury chloride (30% residual activity). Other chemical agents such as diethyl pyrocarbonate (DPC), dithiothreitol (DTT) and EDTA did not affect LipA activity. The effect of various metal ions was investigated by measuring LipA activity in the presence of the metal ions Ca^{2+}, Ni^{2+}, Co^{2+}, Zn^{2+}, Mn^{2+} and Mg^{2+} at concentrations of 1 mM. No significant stimulation or reduction of activity of LipA was observed.

Preliminary crystallization results

Purified LipA was used for crystallization trials, at the Joint Center for Structural Genomics (JCSG) automated pipeline for structure determination, testing 400 conditions. Crystals of native LipA were obtained using a reservoir solution consisting of 1.5M sodium chloride and 10% (v/v) ethanol at 20°C. A dataset was collected to ~2.6 Å resolution. The crystal belongs to spacegroup C2 and molecular replacement is ongoing.

DISCUSSION

This paper describes the cloning and characterization of a thermostable lipase from *A. fulgidus*. In a previous study by Rusnak *et al.* (2005) [93], LipA was partially characterized as well, but the reported features differed considerably from our current findings. The enzyme showed a lower temperature optimum (70°C instead of 90°C) and a much lower thermostability ($t_{1/2}$ of 25 min at 40°C compared to the here reported $t_{1/2}$ of 10 hours at 80°C). Especially the latter result is dubious in view of the high growth temperature of *A. fulgidus*. The previous data suggested that LipA should be classified as a carboxylesterase, whereas our data clearly classified LipA as lipase. These differences may be explained by the fact that in the earlier study the signal peptide of LipA was not removed during cloning. LipA does possess a signal peptide and is therefore believed to be an extracellular enzyme. Usually, when such an enzyme is transported across the cell membrane, the signal peptide is removed in the process resulting in a mature protein. Not removing the signal peptide potentially influences the biochemical properties of LipA.

The multiple-sequence alignment suggested that LipA is composed of two domains: an N-terminal catalytic domain and a C-terminal domain that has no homology to known proteins and therefore cannot be assigned a function. In order to investigate the function of this C-terminal domain, a truncated mutant of LipA, in which the C-terminal domain was removed, was made. LipAtrunc showed lower activity than wild-type LipA on *p*NP-C10, suggesting that the C-terminal domain plays an important role in activity of LipA. Further biochemical analyses are ongoing.

Most esterases and lipases belong to the α/β-hydrolase family and have a conserved catalytic triad that is composed of a nucleophile, an acidic residue and a conserved histidine [6; 8; 56]. The catalytic triad typically comprises Ser, Asp and His, although variations have been observed. In LipA, the Ser136 and His210 are conserved, yet the acidic residue is not conserved. Several potential candidates were identified: Asp163, Asp169, Glu172, Glu193 and Glu194. In the absence of a crystal structure, a three-dimensional model of LipA was built using the 3D-structural threading program PHYRE [220]. In the obtained model, Ser136 and His210 are both located in close proximity of each other, and are at positions according to the canonical α/β-hydrolase fold. The acidic residue (Asp163) is located at a non-canonical position, after strand six, but in close proximity of His210 and Ser136. In the study by Rusnak *et al.* (2005) [93], the amino acids Ser136, Asp163 and His210 were mutated, resulting in complete loss of activity of LipA. Therefore, the model's prediction of Asp163 as potential active site residue is in good agreement with the mutagenesis data. This topological shift of the acidic residue has been observed in other enzymes belonging to the α/β-hydrolase fold [8]. A limited number of lipases have been described that have the acidic catalytic site residue located after the sixth strand, such as the human pancreatic lipases (HPL) and the lipoprotein lipases [221]. A three-dimensional structure of LipA may provide better insight into this topological shift of the acidic residue, as well as give information on the function of the C-terminal part of LipA. Therefore, LipA was crystallized by hanging-drop vapour-diffusion and a dataset was collected to ~2.6 Å resolution. Its structure solution is ongoing.

The expression level of the recombinant enzyme was high and the two-step purification (heat treatment and metal-chelating column) was successful. When LipA was assayed with *p*-nitrophenyl esters, it showed a preference for substrates with long-chain lengths. Highest activity was observed with *p*NP-decanoate. Furthermore, LipA was able to hydrolyze both short and long chain triacylglycerol esters. These data indicate that LipA should be classified as a lipase.

As expected from an enzyme derived from a hyperthermophile, LipA has a high temperature optimum at 90°C and is stable for 10h at 80°C. The optimal activity at high pH values (10-11) and high thermal stability and activity make LipA a good candidate for future applications in industry. This is the first lipase that has been characterized from a hyperthermophilic archaeon.

NOTE ADDED IN PROOF

Recently, the structure of LipA was solved by another group [222] (Figure 7.5a). We could solve our data set by molecular replacement using the corresponding pdb coordinates (2ZYI). Our unit cell and space group are different, however the two structures are nearly identical with an RMSD of 0.28 Å (Figure 7.5b). Residues 71-86 (lid structure) are disordered in our structure, in agreement with the fact that no ligand is bound. The LipA structure can indeed be subdivided into two domains: an N-terminal domain belonging to the α/β-hydrolase fold that contains the active site residues, and a C-terminal β-barrel domain involved in binding of lipids. The structure also shows that, as predicted, the catalytic triad consists of Ser136, Asp163, and His210.
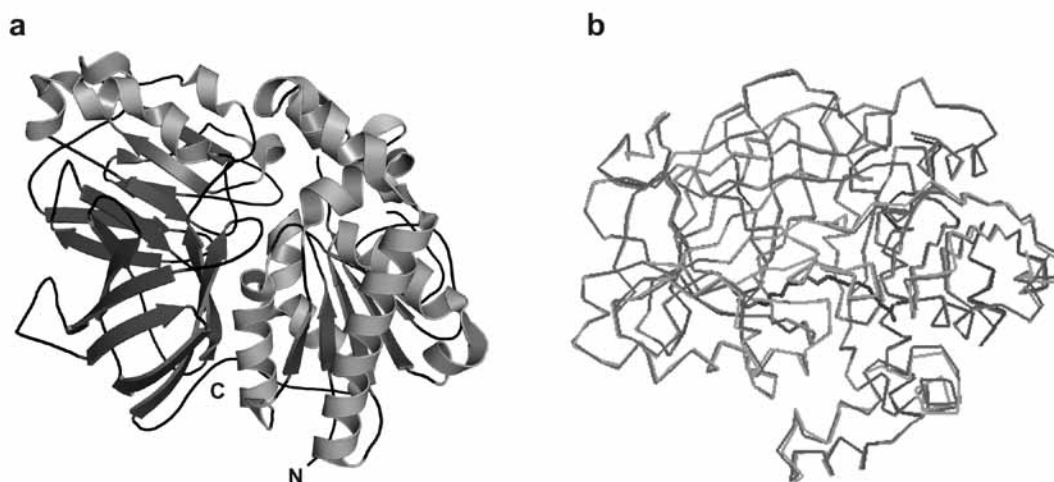

Figure 7.5 (in color on p.153): (a) Overall fold of the lipase (pdb: 2ZYI). The N-terminal and C-terminal ends are indicated. The helices are displayed in cyan and the strands in red. (b) Superposition of our LipA (blue) and 2ZYI (green). A stearic acid that was present in the 2ZYI structure is shown in red.

MATERIALS AND METHODS

Data mining

The genome of *A. fulgidus* was screened for sequences coding possible esterases and lipases. Esterases and lipases were identified by performing BLAST searches at the National Center for Biotechnology Information (NCBI) with sequences from characterized esterases and lipases (http://www.ncbi.nlm.nih.gov/blast/) [28; 223]. Homologues sequences were searched at the DOE Joint Genome Institute (JGI) Integrated Microbial Genomes (IMG) (http://img.jgi.doe. gov). The N-terminal sequence analysis of the translational product of AF1763 was performed using the SignalP 3.0 Server (http://www.cbs.dtu.dk/services/SignalP/) [214]. Multiple sequence alignments were constructed using the Tcoffee server (http://www.igs.cnrs-mrs.fr/Tcoffee/tcoffee_cgi/index.cgi) [132] and manually assessed using GeneDoc.

Cloning

The gene encoding LipA (locus tag AF1763) was PCR-amplified, without the sequence encoding its predicted signal peptide (the first 20 amino acids) and without its stop-codon, using chromosomal DNA of *A. fulgidus* as a template and the two primers BG2064 5'-CGGCCATGGAAGACTTTAGACCGGTAGTGTTTGTGC-3' and BG2065 5'-GGCGCTCGAGAATGTAATCCGAAAACTGCACGATAATCG-3' introducing respectively a *Nco*I and *Xho*I restriction site. The generated PCR product was digested by *Nco*I and *Xho*I and the product was purified and ligated into the expression vector pET-24d digested with the same restriction enzymes, resulting in the plasmid pWUR364. The construct was designed with a hexahistidine tag engineered at the C-terminus of the enzyme to facilitate purification. The sequence of the expression clone was confirmed by sequence analysis of both DNA strands.

Expression and purification

E. coli BL21(DE3)/pSJS1244 was transformed with pWUR364. A single colony was used to inoculate 4 ml Luria-Bertani medium containing kanamycin and spectinomycin (both 50 μg / ml) and incubated overnight at 37°C with shaking. The preculture was used to inoculate (1:1000) 1L of the same medium and growth was continued for 8 h (an OD_{600} above 2.0 was reached). Subsequently, the culture was induced by adding IPTG (isopropyl β-D-1-thiogalactopyranoside) to a final concentration of 0.5 mM. The culture was incubated for a further 16 h at 37°C. Cells were harvested by centrifugation at 10000g for 15 min. The cell pellet was resuspended in 25 ml lysis buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 10 mM imidazole). The cells were disrupted by passage twice through a French press at 110 MPa. The crude cell extract was treated with DNAseI at room temperature for 20 min and subsequently centrifuged at 43000g for 30 min in order to remove cell debris. The supernatant was heated at 70°C for 15 min and then centrifuged to remove the precipitated proteins. The supernatant was filtered and applied onto a nickel-chelating column packed with 20 ml Ni-NTA His-Bind Resin (Novagen) and

equilibrated in 50 mM Tris-HCl buffer pH 7.8 containing 300 mM NaCl. The column was washed with 20 mM imidazole in the same buffer and proteins were subsequently eluted with a linear gradient of 20-500 mM imidazole in 50 mM Tris-HCl pH 7.5, 300 mM NaCl. Fractions containing esterase activity were pooled and applied onto a Hi-Prep desalting column (Amersham Biosciences) equilibrated with 50 mM Tris-HCl buffer pH 7.5. The homogeneity of the protein was checked by SDS-PAGE and activity staining of the SDS-PAGE gels using α-naphtyl acetate, as described previously [52].

Size exclusion chromatography

The molecular mass of the purified esterase was determined by size exclusion chromatography on a Superdex 200 high-resolution 10/30 column (24 mL) (Amersham Biosciences) equilibrated in 50 mM Tris/HCl (pH 7.5) containing 100 mM NaCl. Two hundred microliters of enzyme solution in 50 mM Tris/HCl and 150 mm NaCl (pH 7.5) buffer was loaded at a flow rate of 0.6 mL·min^{-1} onto the column and fractions (0.5 mL) were collected. Proteins used for calibration were blue dextran 2000 (> 2000 kDa), ferritin (440 kDa), catalase (232 kDa), aldolase (158 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), chymotrypsinogen A (25 kDa), and ribonuclease A (13.7 kDa).

Enzyme assays

Esterase activity was determined by measuring the amount of *p*-nitrophenol released during enzymatic hydrolysis of different *p*-nitrophenyl esters. The release of *p*-nitrophenol was continuously monitored at 405 nm using a Hitachi UV2001 spectrophotometer with a temperature controlled cuvette holder. Unless otherwise indicated, in a standard assay, esterase activity was measured with 0.2 mm *p*-nitrophenyl octanoate (*p*NP-C8) as a substrate in 50 mm citrate-phosphate buffer (pH 8) containing 1% isopropanol at 70 °C. Stock solutions of *p*-nitrophenyl esters were prepared by dissolving substrates in isopropanol. After preincubation, the reaction was started by adding enzyme to the reaction mix. One unit of esterase activity was defined as the amount of protein releasing 1 μmol·min^{-1} of *p*-nitrophenol from *p*NP-C8. Measurements were corrected for background hydrolysis in the absence of enzyme. Measurements were carried out at least three times and the molar extinction coefficient of *p*-nitrophenol was determined for every condition prior to each measurement. Activity was determined from the initial rate of the hydrolysis reaction. The protein concentration was measured at 280 nm using a NanoDrop ND-1000 Spectrophotometer

Acyl chain length preference

Substrate specificity of the enzyme towards the acyl chain length of different *p*-nitrophenyl esters was investigated by using *p*-nitrophenyl acetate (C2), *p*-nitrophenyl propionate (C3), *p*-nitrophenyl butyrate (C4), *p*-nitrophenyl valerate (C5), *p*-nitrophenyl octanoate (C8), *p*-nitrophenyl decanoate (C10), *p*-nitrophenyl dodecanoate (C12), *p*-nitrophenyl myristate

(C14), *p*-nitrophenyl palmitate (C16) and *p*-nitrophenyl stearate (C18) in the standard assay.

Plate assays with triacylglycerol esters

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.

pH and temperature optimum

The effect of pH on esterase activity was studied by measuring activities on *p*-nitrophenyl myristate for a pH range of 5-12. The buffers used were 50 mm citrate-phosphate (pH 4.0–8.0), 50 mm CAPS buffer (pH 9.5-12) and 50 mM phosphate buffer (pH12). 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 phosphate buffer, and -0.018 pH·°C^{-1} for CAPS buffer). The effect of temperature on esterase activity was studied in the range 40-100 °C using *p*-nitrophenyl octanoate in the standard assay. Enzyme thermostability was determined by incubating the enzyme in a 50 mm Tris-HCl, 150 mm NaCl (pH 8) buffer at 80° and 90 °C for various time intervals. Residual activity was assayed under the standard condition.

Inhibition studies

The effect of divalent metal ions on activity was determined using different metal salts (CaCl$_2$, NiCl$_2$, CoCl$_2$, MnCl$_2$, MgCl$_2$, and ZnSO$_4$) at final concentrations of 1 mm using the standard activity assay. The activity of LipA without addition of metal ions was defined as 100%. The effect of inhibitors (EDTA, DTT, DPC, HgCl$_2$, PMSF and paraoxon) was studied by preincubating LipA with 1 mM inhibitor in 50 mM TRIS-HCl 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%.

Mutagenesis

A truncated LipA mutant was generated using the primers BG2064 5'-CGGCCATGGAAGACTTTAGACCGGTAGTGTTTGTGC-3' and BG2649 5-CCGCTCGAGCCGTTTATGAACTCAAACATCACC -3 introducing respectively a *Nco*I and *Xho*I restriction site. The generated PCR product was digested by *Nco*I and *Xho*I and the product was purified and ligated into the expression vector pET-24d digested with the same restriction enzymes. The construct was designed with a hexahistidine tag engineered at the C-terminus of the enzyme to facilitate purification. The sequence of the expression clone was confirmed by sequence analysis of both DNA strands. LipAtrunc was produced and purified as described for LipA.

Crystallization and data collection

Purified LipA in 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, and 0.25 mM TCEP was used for crystallization trials at the Joint Center for Structural Genomics (JCSG) automated pipeline for structure determination. 400 JCSG crystallization conditions were screened at 20°C and 4°C. Protein was concentrated to ~19 mg/ml in crystallization buffer (20 mM Tris-HCl, pH 7.8 and 150 mM NaCl) with NDBS195 added to a final concentration of 1M. Crystals of LipA were obtained by hanging drop vapor diffusion against a 60 μl reservoir solution consisting of 1.5 M sodium chloride and 10% (v/v) ethanol at 20°C. Drops consisted of 100 nl protein and 100 nl reservoir. A native dataset was collected to ~2.6 Å resolution. Crystals were flash-cooled to 100K and data collected at beamline 5.0.3 of the ALS. All data were processed with the HKL2000 package [198] and all other crystallographic manipulations were carried out with the CCP4 package [141]. The crystal belongs to spacegroup C2 with unit cell dimensions of A=134.8, B=90.5 Å, C=105.9 Å, a=90°, b=124.7° and g=90°. These dimensions suggest 2 molecules are present in the asymmetric unit with a solvent content of 52.5% and a Matthews coefficient of 2.6 A^3/Da. Molecular replacement is ongoing.

ACKNOWLEDGEMENT

8

Chapter

Summary and general discussion

CARBOXYL ESTER HYDROLASES

Biocatalysts play an important role in modern biotechnology because of their specificity, selectivity, efficiency and sustainability. One of the industrially most exploited and important groups of biocatalysts are the carboxylic ester hydrolases. There is a special interest from industry in carboxylic ester hydrolases of thermophilic origin since these enzymes generally display a high intrinsic thermal and chemical stability [12]. This thesis describes the results of biochemical and structural analyses of thermostable esterases, derived from a hyperthermophilic bacterium (*Thermotoga maritima*) and a hyperthermophilic archaeon (*Archaeoglobus fulgidus*). The information obtained in this study provides fundamental knowledge, which may act as a basis for modern methods of enzyme engineering, with the aim to improve the applicability of these enzymes.

Chapter 1 of this thesis gives an overview of the carboxyl ester hydrolases from hyperthermophilic bacteria and archaea that have been characterized up to now. Carboxyl ester hydrolases are important biocatalysts with applications in medical biotechnology, detergent production, organic synthesis, biodiesel (methyl-ester) production, flavor and aroma synthesis, and other food related processes [10; 11]. At this moment, most esterases and lipases used in industry are obtained from mesophiles ($T_{opt} < 60°C$), basically, because they were the first to be identified and characterized. Analyses of the genome sequences of hyperthermophilic bacteria and archaea ($T_{opt} \geq 80°C$) revealed that the hyperthermophilic world also displays a broad repertoire of esterases that deserves isolation and further characterization.

ESTERASES AND LIPASES

Although there are two well-known groups of carboxylic ester hydrolases, *viz*. the esterases and lipases, the distinction between both groups is not always clear. In fact, the only discriminative feature which is still valid, is that esterases show a preference for short-chain acyl esters (shorter than 10 carbon atoms) and that lipases are active on long-chain acyl esters and are active on micellar substrates [9]. Both types of enzymes share structural and functional characteristics, including a catalytic triad, an α/β-hydrolase fold and a cofactor independent activity [7]. The catalytic triad usually consists of a nucleophilic serine in a GXSXG pentapeptide motif and an acidic residue (aspartic acid or glutamic acid) that is hydrogen bonded to a histidine residue [6; 8; 56]. In **chapter 1** it is shown that the majority of the characterized hyperthermophilic carboxylic ester hydrolases are esterases. Only recently a true lipase, hydrolyzing long chain fatty acid esters, was identified in the hyperthermophilic archaeon *A. fulgidus*. Thus far, only one comprehensive screening approach has been described, in this study 160 thermophilic or hyperthermophilic anaerobic archaea and bacteria were isolated from deep-sea hydrothermal vents and screened for esterase activity, of which 47 were found to be esterase positive [54]. To

the best of our knowledge, no screens have been reported that were aimed to specifically find lipases in hyperthermophiles.

Thermotoga maritima esterases

Most of the characterized esterases in this thesis are obtained from *Thermotoga maritima*. *T. maritima* is a hyperthermophilic anaerobic bacterium isolated from geothermally heated marine sediments at Vulcano, Italy. The organism grows between 55°C and 90°C ($T_{opt} = 80°C$) and can metabolize a variety of simple and complex carbohydrates, including glucose, sucrose, cellulose, xylan and starch [142]. We performed a bioinformatics analysis of the genome of *T. maritima* [119] and several open reading frames potentially encoding new esterases were identified: a hypothetical protein (locus tag: TM0033), a putative esterase (TM0053), a predicted acetyl xylan esterase (TM0077), a hypothetical protein (TM0336), a putative esterase (TM1160), and a putative lipase (TM1350) (Figure 8.1).
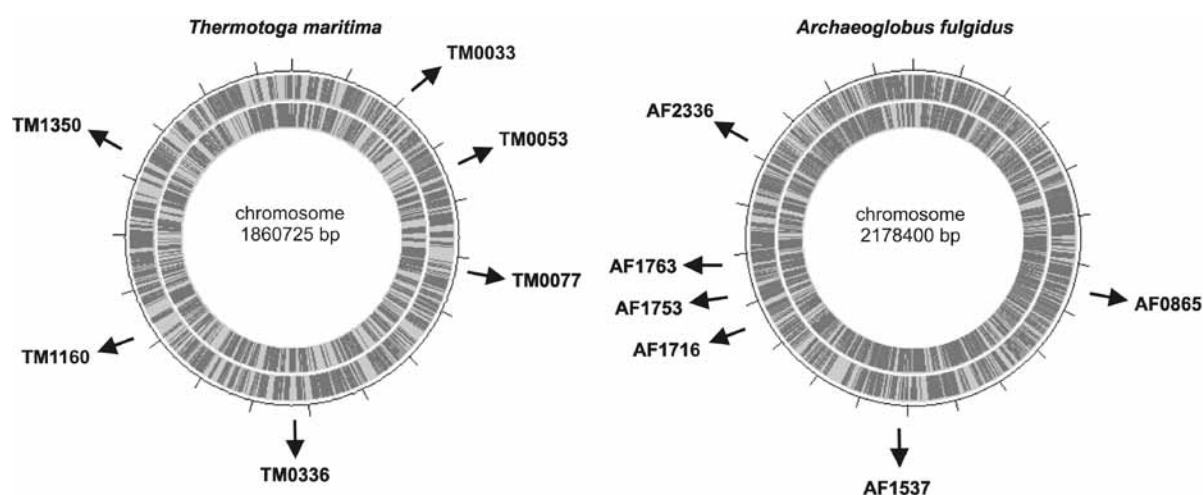

Figure 8.1: Circular representations of the *Thermotoga maritima* and *Archaeoglobus fulgidus* chromosomes. Potential esterase and lipase genes are indicated by their locus tags.

One of the first enzymes that was cloned and found active in an initial screen was EstD (TM0336). **Chapter 2** describes the identification, heterologous production, purification and biochemical characterization of this novel esterase. The enzyme was originally annotated as a hypothetical protein, but a more detailed sequence analysis revealed the presence of an α/β-hydrolase fold and a nucleophilic serine in a pentapeptide motif, suggesting a possible role in ester hydrolysis. We found that EstD exhibited significant esterase activity with a preference for short acyl chain esters (C4–C8). As to be expected for a hyperthermophilic enzyme, EstD showed a temperature optimum around 95°C, and was stable at high temperatures, with a

half-life of 1h at 100°C. This is comparable to other described hyperthermophilic esterases, such as the *Pyrococcus furiosus* esterase [53] and the *Pyrobaculum calidifontis* esterase [38]. In the absence of a crystal structure, a structural model was constructed using the carboxylesterase Est30 from *Geobacillus stearothermophilus* [123] as a template. The model covered most of the C-terminal part of EstD. The model showed an α/β-hydrolase fold and indicated the presence of a typical catalytic triad consisting of Ser243, Asp347 and His378.

Several classifications of esterases and lipases into distinct families have been completed. In one such study, 53 bacterial esterases and lipases were classified into eight families based on their sequence similarity and some of their fundamental biological properties [110]. Phylogenetic analysis showed that EstD is only distantly related to other esterases and could not be grouped into one of these eight families. Therefore, a new family for this enzyme was proposed. Most of the recent classification studies are based on sequence and structural similarity and are accessible at online databases. For a more detailed discussion of relevant databases we refer the reader to **Chapter 1**.

EstA is another esterase (TM0033) that was annotated as a hypothetical protein, but had typical esterase features. A multiple-sequence alignment suggested that EstA is composed of two domains: 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 no function could be assigned. To gain insight into the function of this new N-terminal domain and reveal the molecular basis of substrate recognition and catalysis by this enzyme, crystallization trials were initiated. Hence, in **Chapter 3** the cloning, purification, crystallization and preliminary X-ray analysis of EstA is described. Native and selenomethionine-substituted EstA were purified to apparent homogeneity and were crystallized by the hanging-drop vapour-diffusion method. Crystals grew optimally in 1.0 M lithium sulfate monohydrate and 2% (w/v) polyethylene glycol 8000. Multiple wavelength anomalous data sets were collected to 2.6 Å resolution.

After obtaining its preliminary crystallographic data, the structure and various biochemical properties of EstA were determined, as described in **Chapter 4**. EstA displayed esterase activity with a preference for esters with short acyl chains (C2-C10). The structure of EstA was indeed found to be composed of two distinguishable domains: a conserved C-terminal domain containing the catalytic site and an N-terminal domain resembling the immunoglobulin fold. Analysis of the quaternary structure by mass-spectrometry and dynamic light scattering revealed that EstA predominantly exists as a hexamer in solution. Electron microscopy confirmed that the hexamer in solution is identical with the hexamer in the crystal, which is formed by two trimers, with the immunoglobulin-like domains facing each other. Different functions of immunoglobulin-like domains in other bacterial enzymes have been proposed: substrate binding, directing a substrate to the catalytic groove or cell adhesion [156; 157]. In EstA, the immunoglobulin-like domain appears to play an important role in multimer formation and activity (see disccusion below).

In **Chapter 5** we investigated the structure and various biochemical properties of an

acetyl esterase / cephalosporin C deacetylase AceA (TM0077), belonging to the carbohydrate esterase family seven (CE7) [114]. Enzymes belonging to CE7 are unusual in that they are active towards both acetylated xylo-oligosaccharides and the antibiotic cephalosporin C [176]. AceA was found to be active on a variety of acetylated compounds, including cephalosporin C. The activity on artificial p-nitrophenyl-substrates was confined to short chain acyl esters (C2-C3). The positional specificity of AceA was investigated using 4-nitrophenyl-β-D-xylopyranoside monoacetates as substrate in a β-xylosidase-coupled assay [208]. AceA hydrolyzed acetate from positions 2, 3 and 4 with the same efficiency. A selenomethionine-substituted and a native structure were solved to 2.1 Å and 2.5 Å, respectively, revealing a classical α/β-hydrolase fold. AceA forms a hexamer both in the crystal and in solution. The hexamer has a "doughnut"-like shape, with two entrances on each side of the hexamer leading to an inner cavity, to which the six catalytic centers are exposed. AceA is irreversibly inhibited by the inhibitors PMSF and paraoxon. Structures of AceA in complex with PMSF and paraoxon were solved at 2.4 Å and 2.1 Å, respectively, confirming that both inhibitors bind covalently to the catalytic serine. Surprisingly, upon binding of these substrate analogs, the catalytic serine adopts an altered conformation. Alternate conformations of the catalytic serine have also been observed in other esterase structures, such as the *Fusarium solani* cutinase [189], the *Penicillium purpurogenum* acetyl xylan esterase [190], the *Bacillus subtilis* lipase [191], and the *Aspergillus niger* feruloyl esterase [192]. However, in contrast to the above enzymes, the conformational change of the AceA catalytic serine seems to be solely induced by binding of substrate. We speculate that this transition is required for proper hydrogen bonding to the oxyanion hole, and thus stabilizing the tetrahedral intermediate during catalysis.

In **Chapter 6** we described the cloning, purification, crystallization and partial biochemical characterization of the esterase EstB (TM0053). EstB exhibited esterase activity with a preference for medium acyl chain esters (C8-C10), and with optimal activity around 95°C and at pH 9. Gel filtration indicated that EstB is also present as a hexamer in solution. EstB was crystallized by hanging-drop vapour-diffusion and a dataset was collected to ~2.8 Å resolution. Its structure solution is ongoing.

Archaeoglobus fulgidus lipase

The only true lipase we discovered was obtained from *Archaeoglobus fulgidus. A. fulgidus* is a hyperthermophilic sulphate reducing archaeon initially isolated from hydrothermal areas near Vulcano Island, Italy. The organism grows optimally at 83°C and can use various carbon sources such as fatty acids, amino acids, organic acids, and CO [218; 219]. *Archaeoglobus* species are also commonly found in reservoir waters of offshore oil fields, where they are believed to be responsible for 'reservoir souring' [224; 225]. The genome of *A. fulgidus* was found to contain several open reading frames potentially encoding new esterases: a carboxylesterase (est-1)

(locus tag: AF0865), a carboxylesterase (est-2) (AF1537), a carboxylesterase (AFEST) (AF1716) [59], a lysophospholipase (AF1753), a putative lipase (AF1763) [93], and a carboxylesterase (est-3) (AF2336) (Figure 8.1).

The anticipated function of the putative lipase LipA (AF1753) was investigated. In **Chapter 7** the identification, cloning, purification and partial characterization of this lipase is described. LipA was found to exhibit lipase activity with a preference for medium acyl chain esters (C10) and qualitative plate assays showed that LipA has highest activity on the triacylglycerol esters of octanoate and palmitate. This indicated that LipA should indeed be classified as a lipase. LipA showed optimal activity around 95°C and at pH 11. In several lipases access to the active site is regulated by the position of a "lid structure", which can influence the activity and selectivity of the enzyme [226]. Whether LipA also has a lid structure is not known, a crystal structure may provide insight into substrate recognition and catalysis of this enzyme. LipA has been crystallized by hanging-drop vapour-diffusion and a dataset was collected to ~2.6 Å resolution. Its structure solution is ongoing. This is the first lipase isolated from a hyperthermophilic archaeon.

GENERAL ASPECTS

Various carboxylic ester hydrolases have been investigated in this thesis and several others were treated in the review of **Chapter 1**, and one may wonder whether we can distill some general conclusions with respect to their hyperthermophilic nature.

One obvious feature of all investigated enzymes is that they show a high thermostability, with half-lives ranging from 10h at 80°C to 1.5h at 100°C. This characteristic is also the reason for the interest from industry for this group of enzymes. The high thermostability is a result of subtle changes in the amino acid composition, generally correlating with a high chemical stability [12]. In most cases, however, no specific residues can be identified that govern thermostability of the esterases discussed. The difference in stability observed between our reported characterization of LipA and an earlier description of LipA ($t_{1/2}$ of 25 min at 40°C) [93], can be attributed to the presence of the signal peptide. Cloning of the *lip*A gene without the signal sequence resulted in a significant increase in the temperature optimum and thermal stability (**Chapter 7**). A similar observation was made for EstB (**Chapter 6**) and an earlier description of EstB (T_{opt} = 60°C; pH_{opt} = 7.5) [71], in which the signal peptide was incompletely removed, resulting in a lower temperature optimum and a different pH-optimum. These data indicate that even a short N-terminal sequence can substantially affect the properties of an enzyme. We speculate that a not or incomplete removed signal peptide may prevent multimerization of the enzyme, and as such decrease the thermal stability. If the signal peptide is in close proximity of the active site it may also influence catalysis of the enzyme.

With respect to the active site, the hyperthermophilic esterases, described here, all

show the typical α/β hydrolase fold, and a conserved catalytic triad consisting of a nucleophilic serine in a pentapeptide motif, an aspartate, and a histidine. The only unusual features that were found concern the conformational change of the catalytic serine in AceA upon inhibitor binding, and the non-canonical position, after strand β6, of the catalytic aspartate in LipA. These variations may contribute to improved positioning of active site residues to accommodate certain substrates.

Concerning the quaternary structure it appears that most of the esterases occur as hexamer. Only EstD is present mostly as a monomer (and to a lower extent as dimer). Commonly, both mesophilic and thermophilic esterases are present as a monomer, or higher multimers such as dimer, trimer or tetramer. However, a hexameric structure is rather unusual. Multimerization is a phenomenon often described for enzymes from hyperthermophiles and is regarded as one of different mechanisms to increase thermostability [158]. Whether this is also the case for the esterases is hard to say. For the esterases EstA and AceA, we believe that multimerization most probably also plays an important role in the activity of these enzymes. For instance, the mutation of Phe89Ala, which is located on a loop leading to the active site of another subunit (EstA hexamer), and did result in decreased activity (**Chapter 4**).

One aspect that was hardly discussed in the previous chapters, concerns the physiological role of this group of enzymes. Carboxyl ester hydrolases are ubiquitous enzymes, which have been identified in all domains of life (Bacteria, Archaea and Eukaryotes), and in some viruses. However, for most of the characterized enzymes the physiological role has not been assigned. As discussed above, many of these enzymes have a signal peptide, suggesting that they are transported across the cytoplasmic membrane, and function to hydrolyze extracellular ester compounds, providing carboxylic acids and/or alcohols to be assimilated by the cell. However, there are also esterases without a signal peptide that are presumably active on esterified substrates that have been transported into the cell. The physiological role of esterases acting on carbohydrates is often better understood, e.g. the acetyl esterase of *T. maritima* is believed to be involved in the degradation of xylan, deacetylating xylo-oligosaccharides (as discussed in **Chapter 5**).

The presence of LipA in *A. fulgidus* suggests that the organisms is able to hydrolyze esterified long-chain fatty acids. *A. fulgidus* has been isolated from hot oil reservoirs, but little is known about the natural substrates of the hyperthermophiles in the hot oil reservoirs. Initial experiments with *A. fulgidus* demonstrated, that it is able to grow anaerobically on oil components, such as crude oil, olive oil and arachnic acid, by sulfate reduction [227]. The electron donor was not identified, but was most likely a fatty acid. Genome analyses revealed that *A. fulgidus* has many putative genes encoding enzymes involved in ß-oxidation as well as esterase and lipase genes [219], suggesting it can grow on fatty acids and oils. Growth of *A. fulgidus* has been tested on straight-chain fatty acids from C4 to C18, olive oil, palm oil and sunflower oil. *A. fulgidus* was able to grow on all these substrates using sulfate as terminal electron acceptor (personal communication dr. Diana Sousa).

In several publications, it was shown that the production of lipases and esterases is induced by adding fatty acids or oils to the growth medium, e.g. the studies by Dalmau *et al.* (2000), and Choi & Lee (2001) [228; 229]. To investigate the role of esterases and lipases in *A. fulgidus*, we cultured it on several substrates that could potentially induce the expression of esterases and lipases. The different substrates tested were glycerol, butyrate, octanoate, palmitate, tributyrate, trioctanoate, palm oil, olive oil, and lactate, with thiosulfate as terminal electron receptor. We observed growth on all the substrates tested, except glycerol. A glycerol transporter, a glycerol kinase and glycerol-3-phosphate dehydrogenase have been annotated in the genome sequence [219]. Apparently, *A. fulgidus* is unable to grow on glycerol as a sole carbon source. To investigate whether certain esterase and lipase genes were specifically induced, a microarray experiment was set up, in which *A. fulgidus* was grown batchwise and in a chemostat on lactate, octanoate, trioctanoate and palm oil. However, growth in the chemostat was not successful for the substrates trioctanoate and palm oil due to very slow growth rates and low cell densities compared to growth in serum flasks. Anaerobic hydrolysis of lipids to glycerol and fatty acids occurs rapidly, subsequent fatty acid degradation via ß-oxidation proceeds rather slowly [230]. Due to time limitations, the microarray experiments were terminated (Levisson M., Sousa D.Z. and Henstra A.M., unpublished results).

HIGH PRESSURE AND MICROWAVE IRRADIATION

The initial project aimed also at investigating the behavior of thermostable enzymes under high pressure or microwave irradiation. These methods may open new ways for the tuning of enzyme reactions, and thus offers the possibility of expanding the area of biocatalysis in the food industry.

High pressure can have a kinetic and an equilibrium effect on enzymatic reactions. The kinetic effect is determined by the activation volume of the reaction, meaning that a negative activation volume will result in an increase of the reaction rate. The equilibrium effect is dependent on the reaction volume in such a way that pressure will shift the equilibrium towards the state of lowest volume [231]. Thus far, high pressure has hardly been used for process optimization, mostly because of the instability of enzymes at high pressure. However, recently it was shown that high pressures (up to 1000 MPa) can be applied when an enzyme from a hyperthermophilic microorganism is used [232]. The use of enzymes to catalyze the synthesis of ester bonds is of great interest to industry. However, the synthesis of an ester-bond between an alcohol and a fatty acid is a condensation reaction in which water and an ester are formed. This will lead to an increase in the reaction volume and thus, an unfavorable equilibrium. Therefore, ester synthesis under high pressure is not feasible. On the other hand, the equilibrium for the hydrolysis of an ester will be positively influenced by high pressure. Microwave irradiation has been successfully used to accelerate reactions in organic chemistry

[233; 234]. There are two mechanisms associated with microwave heating: dipole polarization and ionic conduction [233; 235]. A polar molecule (dipole or ion) will continuously align itself with the fluctuating electric field. This causes molecular friction and the energy is lost as heat. During ionic conduction, as the molecules fluctuate under influence of the microwave field, they collide with their neighboring molecules or atoms. These collisions cause agitation or motion, creating heat [236]. Lately, microwave irradiation has also been used as a tool in biocatalysis. Several effects on enzymatic activity have been reported, such as improved stability [237], improved rates [238; 239], improved conversion of substrate (Gelo-Pujic 1997), improved activity at suboptimal temperatures [240], an improved affinity, and kinetic resolution [239; 241; 242]. In contrast, there are also several publications that question the beneficial effects of microwave irradiation and showed that the observed effects were purely thermal and not related to the microwave field [243; 244].

In the course of this project, the esterases EstA and EstB from *T. maritima* and the lipase LipA from *A. fulgidus* have been investigated for their ability to stereo-selectively acylate racemic alcohols. Among the various enzymes and reactions evaluated, the enzyme LipA catalyzed the synthesis of 2-octyl acetate using (*R,S*)-2-octanol and vinyl-acetate as substrates, and 2-pentyl acetate using (*R,S*)-2-pentanol and vinyl-acetate as substrates (Figure 8.2). The latter reaction was studied at 70°C comparing conventional heating in a waterbath and microwave irradiation in a CEM Discover. Care was taken that the experiments in both systems were performed under identical circumstances, only changing the mode of heating. In addition, the effect of two different solvents, toluene and heptane, was tested. The reaction performed in heptane had a two-fold higher rate compared to toluene. We also found that, in both toluene and heptane, the heating mode had no significant effect on the rate of the reaction. This was also observed in a study by Réjasse *et al.* (2004) [237]. The effect of solvent and the heating mode on kinetic resolution is ongoing.
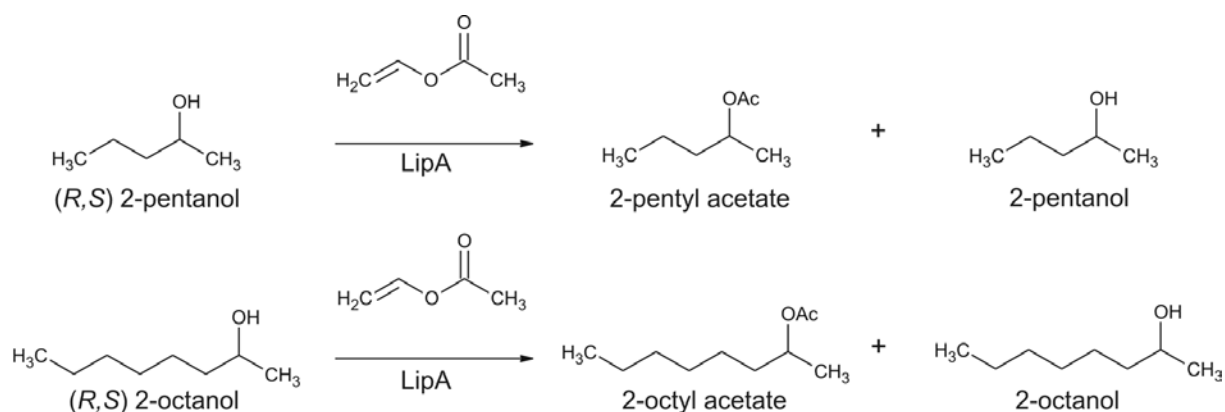

Figure 8.2: Transesterification of (*R,S*)-2-pentanol and (*R,S*)-2-octanol with vinyl acetate by LipA. The kinetic resolution of LipA is not known.

APPLICATIONS

The enzymes described in this thesis have been tested in a platform, amongst ≥400 other commercially available esterases and lipases, at an industrial company for multiple applications. Most reactions concerned the selective hydrolysis of a racemic ester, resulting in an optically pure product. Because of intellectual property reasons no details can be given on the exact composition of these reactions. Nevertheless, in various of the reactions tested, one or more of the hyperthermophilic enzymes performed better than the other enzymes in the platform. Especially, LipA, EstA and EstB often behaved best in screens of ≥400 different esterases. The enzymes showed good conversion, selectivity, and / or activity with the selected substrates. These results indicate that there is a lot of potential in the use of carboxyl ester hydrolases from hyperthermophiles for industrial applications.

NOVEL DEVELOPMENTS AND FUTURE PERSPECTIVES

In recent years, many new hyperthermophilic bacteria and archaea have been isolated. The genomes of several of these hyperthermophiles have been sequenced and in future this number will increase rapidly due to forthcoming sequencing projects (GOLD genomes online; [23]). This increase in sequence information will accelerate the identification of new carboxylic ester hydrolases with new properties. Hitherto, traditional screening has been used to identify new enzymes, however, bioinformatics and metagenome-screening will contribute more and more to this identification process. A major drawback of metagenome screening is that in order to function well, the genes of interest need to be functionally expressed in the heterologous screening host. Therefore, recently a new two-host fosmid system for functional screening of (meta)genomic libraries from extreme thermophiles was developed [245]. This system allows the construction of large-insert fosmid libraries in *E. coli* and the transfer of the recombinant libraries to the extreme thermophile *T. thermophilus*, in combination with high-throughput screening (robotics). This system has been proven to have a higher level of functional expressed genes and may be of value in the identification of new carboxylic ester hydrolases from hyperthermophiles. However, in addition to the identification of new carboxylic ester hydrolases also their characterization is indispensable.

The classification of esterases into families is an ongoing process and due to the rapid increase in new esterase gene sequences, many of the current databases are incomplete. A promising approach is the superfamily based approach, which combines theoretical and experimental data and can reveal much information about a protein family [246]. A new completely automatic program capable of constructing these superfamily systems is 3DM [247]. This program is able to create a new superfamily of the carboxylic ester hydrolases based on structural and sequence similarity. In addition, superfamily systems generated by 3DM have

also been proven to be powerful tools for understanding and predicting rational modification of proteins [248].

Many new sequences and protein structures of carboxylic ester hydrolases are becoming available. These structures can provide a basis for modern methods of enzyme engineering. Apart from exploiting natural evolution, optimization by computational design and laboratory evolution is a major challenge for obtaining tailor-made enzymes for industrial applications [249]. In the past, these methods have been successfully used to alter enzymes to meet specific demands, including increased stability, activity, and enantioselectivity [250]. In future, the identification of new esterases, together with advanced engineering strategies and high-throughput screening systems will lead to improved thermostable esterases able to perform a vast array of reactions.

References

1. **Rozzell, J. D.** (1999). Commercial scale biocatalysis: myths and realities. *Bioorg Med Chem* **7**, 2253-61.
2. **Hasan, F., Shah, A. A. & Hameed, A.** (2006). Industrial applications of microbial lipases. *Enzyme and Microbial Technology* **39**, 235-251.
3. **Jaeger, K. E. & Eggert, T.** (2002). Lipases for biotechnology. *Curr Opin Biotechnol* **13**, 390-7.
4. **Krishna, S. H. & Karanth, N. G.** (2002). Lipases and lipase-catalyzed esterification reactions in nonaqueous media. *Catalysis Reviews* **44**, 499-591.
5. **Bornscheuer, U. T.** (2002). Microbial carboxyl esterases: classification, properties and application in biocatalysis. *FEMS Microbiol Rev* **26**, 73-81.
6. **Heikinheimo, P., Goldman, A., Jeffries, C. & Ollis, D. L.** (1999). Of barn owls and bankers: a lush variety of alpha/beta hydrolases. *Structure* **7**, R141-6.
7. **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**, 315-51.
8. **Nardini, M. & Dijkstra, B. W.** (1999). Alpha/beta hydrolase fold enzymes: the family keeps growing. *Curr Opin Struct Biol* **9**, 732-7.
9. **Chahinian, H., Nini, L., Boitard, E., Dubes, J. P., Comeau, L. C. & Sarda, L.** (2002). Distinction between esterases and lipases: a kinetic study with vinyl esters and TAG. *Lipids* **37**, 653-62.
10. **Panda, T. & Gowrishankar, B. S.** (2005). Production and applications of esterases. *Appl Microbiol Biotechnol* **67**, 160-9.
11. **Salameh, M. A. & Wiegel, J.** (2007). Lipases from extremophiles and potential for industrial applications. *Adv Appl Microbiol* **61**, 253-83.
12. **Gomes, J. & Steiner, W.** (2004). The biocatalytic potential of extremophiles and extremozymes. *Food Technology and Biotechnology* **42**, 223-235.
13. **Atomi, H.** (2005). Recent progress towards the application of hyperthermophiles and their enzymes. *Curr Opin Chem Biol* **9**, 166-73.
14. **Egorova, K. & Antranikian, G.** (2005). Industrial relevance of thermophilic Archaea. *Curr Opin Microbiol* **8**, 649-55.
15. **Unsworth, L. D., van der Oost, J. & Koutsopoulos, S.** (2007). Hyperthermophilic enzymes - stability, activity and implementation strategies for high temperature applications. *Febs J* **274**, 4044-56.
16. **Atomi, H. & Imanaka, T.** (2004). Thermostable carboxylesterases from hyperthermophiles. *Tetrahedron: Asymmetry* **15**, 2729-2735.
17. **Stetter, K. O.** (1996). Hyperthermophilic procaryotes. *FEMS Microbiology Reviews* **18**, 149-158.
18. **Vieille, C. & Zeikus, G. J.** (2001). Hyperthermophilic enzymes: sources, uses, and molecular mechanisms for thermostability. *Microbiol Mol Biol Rev* **65**, 1-43.
19. **Stetter, K. O.** (2006). Hyperthermophiles in the history of life. *Philos Trans R Soc Lond B Biol Sci* **361**, 1837-42; discussion 1842-3.
20. **Atomi, H., Matsumi, R. & Imanaka, T.** (2004). Reverse gyrase is not a prerequisite for hyperthermophilic life. *J Bacteriol* **186**, 4829-33.
21. **Robinson-Rechavi, M., Alibes, A. & Godzik, A.** (2006). Contribution of electrostatic interactions, compactness and quaternary structure to protein thermostability: lessons from structural genomics of *Thermotoga maritima*. *J Mol Biol* **356**, 547-57.
22. **Santos, H. & da Costa, M. S.** (2002). Compatible solutes of organisms that live in hot saline environments. *Environ Microbiol* **4**, 501-9.
23. **Liolios, K., Mavromatis, K., Tavernarakis, N. & Kyrpides, N. C**. (2008). The Genomes OnLine Database (GOLD) in 2007: status of genomic and metagenomic projects and their associated metadata. *Nucleic Acids Res* **36**, D475-9.
24. **Lorenz, P. & Eck, J.** (2005). Metagenomics and industrial applications. *Nat Rev Microbiol* **3**, 510-6.
25. **Rhee, J. K., Ahn, D. G., Kim, Y. G. & Oh, J. W.** (2005). New thermophilic and thermostable esterase with sequence similarity to the hormone-sensitive lipase family, cloned from a metagenomic library. *Appl Environ Microbiol* **71**, 817-25.
26. **Tirawongsaroj, P., Sriprang, R., Harnpicharnchai, P., Thongaram, T., Champreda, V.,**

Tanapongpipat, S., Pootanakit, K. & Eurwilaichitr, L. (2008). Novel thermophilic and thermostable lipolytic enzymes from a Thailand hot spring metagenomic library. *J Biotechnol* **133**, 42-9.

27. **Kwoun Kim, H., Jung, Y. J., Choi, W. C., Ryu, H. S., Oh, T. K. & Lee, J. K.** (2004). Sequence-based approach to finding functional lipases from microbial genome databases. *FEMS Microbiol Lett* **235**, 349-55.

28. **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**, 3389-402.

29. **Hunter, S., Apweiler, R., Attwood, T. K., Bairoch, A., Bateman, A., Binns, D., Bork, P., Das, U., Daugherty, L., Duquenne, L., Finn, R. D., Gough, J., Haft, D., Hulo, N., Kahn, D., Kelly, E., Laugraud, A., Letunic, I., Lonsdale, D., Lopez, R., Madera, M., Maslen, J., McAnulla, C., McDowall, J., Mistry, J., Mitchell, A., Mulder, N., Natale, D., Orengo, C., Quinn, A. F., Selengut, J. D., Sigrist, C. J., Thimma, M., Thomas, P. D., Valentin, F., Wilson, D., Wu, C. H. & Yeats, C.** (2009). InterPro: the integrative protein signature database. *Nucleic Acids Res* **37**, D211-5.

30. **Marchler-Bauer, A., Anderson, J. B., Cherukuri, P. F., DeWweese-Scott, C., Geer, L. Y., Gwadz, M., He, S. Q., Hurwitz, D. I., Jackson, J. D., Ke, Z. X., Lanczycki, C. J., Liebert, C. A., Liu, C. L., Lu, F., Marchler, G. H., Mullokandov, M., Shoemaker, B. A., Simonyan, V., Song, J. S., Thiessen, P. A., Yamashita, R. A., Yin, J. J., Zhang, D. C. & Bryant, S. H.** (2005). CDD: a conserved domain database for protein classification. *Nucleic Acids Research* **33**, D192-D196.

31. **Marchler-Bauer, A., Anderson, J. B., Chitsaz, F., Derbyshire, M. K., DeWeese-Scott, C., Fong, J. H., Geer, L. Y., Geer, R. C., Gonzales, N. R., Gwadz, M., He, S., Hurwitz, D. I., Jackson, J. D., Ke, Z., Lanczycki, C. J., Liebert, C. A., Liu, C., Lu, F., Lu, S., Marchler, G. H., Mullokandov, M., Song, J. S., Tasneem, A., Thanki, N., Yamashita, R. A., Zhang, D., Zhang, N. & Bryant, S. H.** (2009). CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Res* **37**, D205-10.

32. **Sobek, H. & Gorisch, H.** (1988). Purification and characterization of a heat-stable esterase from the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius*. *Biochem J* **250**, 453-8.

33. **Sobek, H. & Gorisch, H.** (1989). Further kinetic and molecular characterization of an extremely heat-stable carboxylesterase from the thermoacidophilic archaebacterium *Sulfolobus acidocaldarius*. *Biochem J* **261**, 993-8.

34. **Gupta, R., Gupta, N. & Rathi, P.** (2004). Bacterial lipases: an overview of production, purification and biochemical properties. *Appl Microbiol Biotechnol* **64**, 763-81.

35. **Salameh, M. A. & Wiegel, J.** (2007). Purification and characterization of two highly thermophilic alkaline lipases from *Thermosyntropha lipolytica*. *Appl Environ Microbiol* **73**, 7725-31.

36. **Rathi, P., Bradoo, S., Saxena, R. K. & Gupta, R.** (2000). A hyper-thermostable, alkaline lipase from *Pseudomonas* sp. with the property of thermal activation. *Biotechnology Letters* **22**, 495-498.

37. **Suen, W. C., Zhang, N., Xiao, L., Madison, V. & Zaks, A.** (2004). Improved activity and thermostability of *Candida antarctica* lipase B by DNA family shuffling. *Protein Eng Des Sel* **17**, 133-40.

38. **Hotta, Y., Ezaki, S., Atomi, H. & Imanaka, T.** (2002). Extremely stable and versatile carboxylesterase from a hyperthermophilic archaeon. *Appl Environ Microbiol* **68**, 3925-31.

39. **Kim, S. & Lee, S. B.** (2004). Thermostable esterase from a thermoacidophilic archaeon: purification and characterization for enzymatic resolution of a chiral compound. *Biosci Biotechnol Biochem* **68**, 2289-98.

40. **Sehgal, A. C., Callen, W., Mathur, E. J., Short, J. M. & Kelly, R. M.** (2001). Carboxylesterase from *Sulfolobus solfataricus* P1. *Methods Enzymol* **330**, 461-71.

41. **Quax, W. J. & Broekhuizen, C. P.** (1994). Development of a new *Bacillus* carboxyl esterase for use in the resolution of chiral drugs. *Appl Microbiol Biotechnol* **41**, 425-31.

42. **Lee, E. G., Won, H. S. & Chung, B. H.** (2001). Enantioselective hydrolysis of racemic naproxen methyl ester by two-step acetone-treated *Candida rugosa* lipase. *Process Biochemistry* **37**, 293-8.

43. **Sehgal, A. C. & Kelly, R. M.** (2002). Enantiomeric resolution of 2-aryl propionic esters with hyperthermophilic and mesophilic esterases: contrasting thermodynamic mechanisms. *J Am Chem Soc* **124**, 8190-1.

44. **Sehgal, A. C. & Kelly, R. M.** (2003). Strategic selection of hyperthermophilic esterases for resolution of 2-arylpropionic esters. *Biotechnol Prog* **19**, 1410-6.

45. **Merone, L., Mandrich, L., Rossi, M. & Manco, G.** (2005). A thermostable phosphotriesterase from the archaeon *Sulfolobus solfataricus*: cloning, overexpression and properties. *Extremophiles* **9**, 297-305.

46. **Park, Y. J., Yoon, S. J. & Lee, H. B.** (2008). A novel thermostable arylesterase from the archaeon *Sulfolobus solfataricus* P1: purification, characterization, and expression. *J Bacteriol* **190**, 8086-95.

47. **Porzio, E., Merone, L., Mandrich, L., Rossi, M. & Manco, G.** (2007). A new phosphotriesterase from *Sulfolobus acidocaldarius* and its comparison with the homologue from *Sulfolobus solfataricus*. *Biochimie* **89**, 625-36.

48. **Choo, D. W., Kurihara, T., Suzuki, T., Soda, K. & Esaki, N.** (1998). A cold-adapted lipase of an Alaskan psychrotroph, *Pseudomonas* sp. strain B11-1: gene cloning and enzyme purification and characterization. *Appl Environ Microbiol* **64**, 486-91.

49. **Shimada, Y., Koga, C., Sugihara, A., Nagao, T., Takada, N., Tsunasawa, S. & Tominaga, Y.** (1993). Purification and characterization of a novel solvent-tolerant lipase from *Fusarium heterosporum*. *Journal of Fermentation and Bioengineering* **75**, 349-352.

50. **Park, Y. J., Choi, S. Y. & Lee, H. B.** (2006). A carboxylesterase from the thermoacidophilic archaeon *Sulfolobus solfataricus* P1; purification, characterization, and expression. *Biochim Biophys Acta* **1760**, 820-8.

51. **Stocklein, W., Sztajer, H., Menge, U. & Schmid, R. D.** (1993). Purification and properties of a lipase from *Penicillium expansum*. *Biochim Biophys Acta* **1168**, 181-9.

52. **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**, 2832-42.

53. **Ikeda, M. & Clark, D. S.** (1998). Molecular cloning of extremely thermostable esterase gene from hyperthermophilic archaeon *Pyrococcus furiosus* in *Escherichia coli*. *Biotechnol Bioeng* **57**, 624-9.

54. **Cornec, L.** (1998). Thermostable esterases screened on hyperthermophilic archaeal and bacterial strains isolated from deep-sea hydrothermal vents: Characterization of esterase activity of a hyperthermophilic archaeum, *Pyrococcus abyssi*. *Journal of Marine Biotechnology* **6**, 104-110.

55. **Huddleston, S., Yallop, C. A. & Charalambous, B. M.** (1995). The identification and partial characterisation of a novel inducible extracellular thermostable esterase from the archaeon *Sulfolobus shibatae*. *Biochem Biophys Res Commun* **216**, 495-500.

56. **Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M., Harel, M., Remington, S. J., Silman, I., Schrag, J. & et al.** (1992). The alpha/beta hydrolase fold. *Protein Eng* **5**, 197-211.

57. **Pleiss, J., Fischer, M. & Schmid, R. D.** (1998). Anatomy of lipase binding sites: the scissile fatty acid binding site. *Chem Phys Lipids* **93**, 67-80.

58. **De Simone, G., Menchise, V., Manco, G., Mandrich, L., Sorrentino, N., Lang, D., Rossi, M. & Pedone, C.** (2001). The crystal structure of a hyper-thermophilic carboxylesterase from the archaeon *Archaeoglobus fulgidus*. *J Mol Biol* **314**, 507-18.

59. **Manco, G., Giosue, E., D'Auria, S., Herman, P., Carrea, G. & Rossi, M.** (2000). Cloning, overexpression, and properties of a new thermophilic and thermostable esterase with sequence similarity to hormone-sensitive lipase subfamily from the archaeon *Archaeoglobus fulgidus*. *Arch Biochem Biophys* **373**, 182-92.

60. **Wei, Y., Contreras, J. A., Sheffield, P., Osterlund, T., Derewenda, U., Kneusel, R. E., Matern, U., Holm, C. & Derewenda, Z. S.** (1999). Crystal structure of brefeldin A esterase, a bacterial homolog of the mammalian hormone-sensitive lipase. *Nat Struct Biol* **6**, 340-5.

61. **De Simone, G., Galdiero, S., Manco, G., Lang, D., Rossi, M. & Pedone, C.** (2000). A snapshot of a transition state analogue of a novel thermophilic esterase belonging to the subfamily of mammalian hormone-sensitive lipase. *J Mol Biol* **303**, 761-71.

62. **Mandrich, L., Pezzullo, M., Del Vecchio, P., Barone, G., Rossi, M. & Manco, G.** (2004). Analysis of thermal adaptation in the HSL enzyme family. *Journal of Molecular Biology* **335**, 357-369.

63. **Byun, J. S., Rhee, J. K., Kim, N. D., Yoon, J., Kim, D. U., Koh, E., Oh, J. W. & Cho, H. S.** (2007). Crystal structure of hyperthermophilic esterase EstE1 and the relationship between its dimerization and

thermostability properties. *BMC Struct Biol* **7**, 47.

64. **Rhee, J. K., Kim, D. Y., Ahn, D. G., Yun, J. H., Jang, S. H., Shin, H. C., Cho, H. S., Pan, J. G. & Oh, J. W.** (2006). Analysis of the thermostability determinants of hyperthermophilic esterase EstE1 based on its predicted three-dimensional structure. *Appl Environ Microbiol* **72**, 3021-5.

65. **Bartlam, M., Wang, G., Yang, H., Gao, R., Zhao, X., Xie, G., Cao, S., Feng, Y. & Rao, Z.** (2004). Crystal structure of an acylpeptide hydrolase/esterase from *Aeropyrum pernix* K1. *Structure* **12**, 1481-8.

66. **Gao, R., Feng, Y., Ishikawa, K., Ishida, H., Ando, S., Kosugi, Y. & Cao, S.** (2003). Cloning, purification and properties of a hyperthermophilic esterase from archaeon *Aeropyrum pernix* K1. *Journal of Molecular Catalysis. B, Enzymatic* **24**, 1-8.

67. **Wang, Q., Yang, G., Liu, Y. & Feng, Y.** (2006). Discrimination of esterase and peptidase activities of acylaminoacyl peptidase from hyperthermophilic *Aeropyrum pernix* K1 by a single mutation. *J Biol Chem* **281**, 18618-25.

68. **Levisson, M., Sun, L., Hendriks, S., Swinkels, P., Akveld, T., Bultema, J. B., Barendregt, A., van den Heuvel, R. H., Dijkstra, B. W., van der Oost, J. & Kengen, S. W.** (2009). Crystal structure and biochemical properties of a novel thermostable esterase containing an immunoglobulin-like domain. *J Mol Biol* **385**, 949-62.

69. **Fox, B. G., Goulding, C., Malkowski, M. G., Stewart, L. & Deacon, A.** (2008). Structural genomics: from genes to structures with valuable materials and many questions in between. *Nat Methods* **5**, 129-32.

70. **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 Struct Biol Cryst Commun* **63**, 777-9.

71. **Kakugawa, S., Fushinobu, S., Wakagi, T. & Shoun, H.** (2007). Characterization of a thermostable carboxylesterase from the hyperthermophilic bacterium *Thermotoga maritima*. *Appl Microbiol Biotechnol* **74**, 585-91.

72. **Zhang, J., Liu, J., Zhou, J., Ren, Y., Dai, X. & Xiang, H.** (2003). Thermostable esterase from *Thermoanaerobacter tengcongensis*: high-level expression, purification and characterization. *Biotechnol Lett* **25**, 1463-7.

73. **Murayama, K., Shirouzu, M., Terada, T., Kuramitsu, S. & Yokoyama, S.** (2005). Crystal structure of TT1662 from *Thermus thermophilus* HB8: a member of the alpha/beta hydrolase fold enzymes. *Proteins* **58**, 982-4.

74. **Domínguez, A., Fucinos, P., Rúa, M. L., Pastrana, L., Longo, M. A. & Sanromán, M. A.** (2007). Stimulation of novel thermostable extracellular lipolytic enzyme in cultures of *Thermus* sp. *Enzyme and Microbial Technology* **40**, 187-194.

75. **Dominguez, A., Pastrana, L., Longo, M. A., Rúa, M. L. & Sanroman, M. A.** (2005). Lipolytic enzyme production by *Thermus thermophilus* HB27 in a stirred tank bioreactor. *Biochemical Engineering Journal* **26**, 95-99.

76. **Dominguez, A., Sanroman, A., Fucinos, P., Rua, M. L., Pastrana, L. & Longo, M. A.** (2004). Quantification of intra- and extra-cellular thermophilic lipase/esterase production by *Thermus* sp. *Biotechnol Lett* **26**, 705-8.

77. **Fucinos, P., Abadín, C. M., Sanromán, A., Longo, M. A., Pastrana, L. & Rúa, M. L.** (2005). Identification of extracellular lipases/esterases produced by *Thermus thermophilus* HB27: Partial purification and preliminary biochemical characterisation. *Journal of Biotechnology* **117**, 233-241.

78. **Fucinos, P., Dominguez, A., Sanroman, M. A., Longo, M. A., Rua, M. L. & Pastrana, L.** (2005). Production of thermostable lipolytic activity by *Thermus* species. *Biotechnol Prog* **21**, 1198-205.

79. **Fucinos, P., Rúa, M. L., Longo, M. A., Sanromán, M. & Pastrana, L.** (2008). Thermal spring water enhances lipolytic activity in *Thermus thermophilus* HB27. *Process Biochemistry* **43**, 1383-1390.

80. **Gao, R.-j., Xie, G.-q., Zhou, J., Feng, Y. & Cao, S.-g.** (2006). Denaturing effects of urea and guanidine hydrochloride on hyperthermophilic esterase from *Aeropyrum pernix* K1. *Chemical Research in Chinese Universities* **22**, 168-172.

81. **Yang, G., Bai, A., Gao, L., Zhang, Z., Zheng, B. & Feng, Y.** (2009). Glu88 in the non-catalytic domain of acylpeptide hydrolase plays dual roles: Charge neutralization for enzymatic activity and formation of salt bridge for thermodynamic stability. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* **1794**, 94-102.

82. **Zhang, G., Gao, R., Zheng, L., Zhang, A., Wang, Y., Wang, Q., Feng, Y. & Cao, S.** (2006). Study on the relationship between structure and enantioselectivity of a hyperthermophilic esterase from archaeon *Aeropyrum pernix* K1. *Journal of Molecular Catalysis B: Enzymatic* **38**, 148-153.

83. **Zhang, G.-r., Gao, R.-j., Zhang, A.-j., Rao, L. & Cao, S.-g.** (2007). High-throughput screening of the enantioselectivity of hyperthermophilic mutant esterases from archaeon *Aeropyrum pernix* K1 for resolution of (R,S)-2-octanol acetate. *Chemical Research in Chinese Universities* **23**, 319-324.

84. **Zhang, Z., Zheng, B., Wang, Y., Chen, Y., Manco, G. & Feng, Y.** (2008). The conserved N-terminal helix of acylpeptide hydrolase from archaeon *Aeropyrum pernix* K1 is important for its hyperthermophilic activity. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* **1784**, 1176-1183.

85. **Wang, B., Lu, D., Gao, R., Yang, Z., Cao, S. & Feng, Y.** (2004). A novel phospholipase A2/esterase from hyperthermophilic archaeon *Aeropyrum pernix* K1. *Protein Expr Purif* **35**, 199-205.

86. **Ro, H. S., Hong, H. P., Kho, B. H., Kim, S. & Chung, B. H.** (2004). Genome-wide cloning and characterization of microbial esterases. *FEMS Microbiol Lett* **233**, 97-105.

87. **D'Auria, S., Herman, P., Lakowicz, J. R., Bertoli, E., Tanfani, F., Rossi, M. & Manco, G.** (2000). The thermophilic esterase from *Archaeoglobus fulgidus*: structure and conformational dynamics at high temperature. *Proteins* **38**, 351-60.

88. **Del Vecchio, P., Elias, M., Merone, L., Graziano, G., Dupuy, J., Mandrich, L., Carullo, P., Fournier, B., Rochu, D., Rossi, M., Masson, P., Chabriere, E. & Manco, G.** (2009). Structural determinants of the high thermal stability of SsoPox from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Extremophiles*.

89. **Del Vecchio, P., Graziano, G., Granata, V., Barone, G., Mandrich, L., Rossi, M. & Manco, G.** (2002). Denaturing action of urea and guanidine hydrochloride towards two thermophilic esterases. *Biochem J* **367**, 857-63.

90. **Del Vecchio, P., Graziano, G., Granata, V., Barone, G., Mandrich, L., Rossi, M. & Manco, G.** (2003). Effect of trifluoroethanol on the conformational stability of a hyperthermophilic esterase: a CD study. *Biophys Chem* **104**, 407-15.

91. **Manco, G., Camardella, L., Febbraio, F., Adamo, G., Carratore, V. & Rossi, M.** (2000). Homology modeling and identification of serine 160 as nucleophile of the active site in a thermostable carboxylesterase from the archaeon *Archaeoglobus fulgidus*. *Protein Eng* **13**, 197-200.

92. **Manco, G., Carrea, G., Giosuè, E., Ottolina, G., Adamo, G. & Rossi, M.** (2002). Modification of the enantioselectivity of two homologous thermophilic carboxylesterases from *Alicyclobacillus acidocaldarius* and *Archaeoglobus fulgidus* by random mutagenesis and screening. *Extremophiles* **6**, 325-331.

93. **Rusnak, M., Nieveler, J., Schmid, R. D. & Petri, R.** (2005). The putative lipase, AF1763, from *Archaeoglobus fulgidusis* is a carboxylesterase with a very high pH optimum. *Biotechnol Lett* **27**, 743-8.

94. **Kim, S. B., Lee, W. & Ryu, Y. W.** (2008). Cloning and characterization of thermostable esterase from *Archaeoglobus fulgidus*. *J Microbiol* **46**, 100-7.

95. **Hess, M., Katzer, M. & Antranikian, G.** (2008). Extremely thermostable esterases from the thermoacidophilic euryarchaeon *Picrophilus torridus*. *Extremophiles* **12**, 351-64.

96. **Chandrayan, S. K., Dhaunta, N. & Guptasarma, P.** (2008). Expression, purification, refolding and characterization of a putative lysophospholipase from *Pyrococcus furiosus*: retention of structure and lipase/esterase activity in the presence of water-miscible organic solvents at high temperatures. *Protein Expr Purif* **59**, 327-33.

97. **Almeida, R. V., Alquéres, S. M. C., Larentis, A. L., Rössle, S. C., Cardoso, A. M., Almeida, W. I., Bisch, P. M., Alves, T. L. M. & Martins, O. B.** (2006). Cloning, expression, partial characterization and structural modeling of a novel esterase from *Pyrococcus furiosus*. *Enzyme and Microbial*

Technology **39**, 1128-1136.

98. **Almeida, R. V., Branco, R. V., Peixoto, B., Lima, C. d. S., Alqueres, S. M. C., Martins, O. B., Antunes, O. A. C. & Freire, D. M. G.** (2008). Immobilization of a recombinant thermostable esterase (Pf2001) from *Pyrococcus furiosus* on microporous polypropylene: Isotherms, hyperactivation and purification. *Biochemical Engineering Journal* **39**, 531-537.

99. **Arpigny, J. L., Jendrossek, D. & Jaeger, K. E.** (1998). A novel heat-stable lipolytic enzyme from *Sulfolobus acidocaldarius* DSM 639 displaying similarity to polyhydroxyalkanoate depolymerases. *FEMS Microbiol Lett* **167**, 69-73.

100. **Ejima, K., Liu, J., Oshima, Y., Hirooka, K., Shimanuki, S., Yokota, Y., Hemmi, H., Nakayama, T. & Nishino, T.** (2004). Molecular cloning and characterization of a thermostable carboxylesterase from an archaeon, *Sulfolobus shibatae* DSM5389: non-linear kinetic behavior of a hormone-sensitive lipase family enzyme. *J Biosci Bioeng* **98**, 445-51.

101. **Morana, A., Di Prizito, N., Aurilia, V., Rossi, M. & Cannio, R.** (2002). A carboxylesterase from the hyperthermophilic archaeon *Sulfolobus solfataricus*: cloning of the gene, characterization of the protein. *Gene* **283**, 107-15.

102. **Elias, M., Dupuy, J., Merone, L., Lecomte, C., Rossi, M., Masson, P., Manco, G. & Chabriere, E.** (2007). Crystallization and preliminary X-ray diffraction analysis of the hyperthermophilic *Sulfolobus solfataricus* phosphotriesterase. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **63**, 553-5.

103. **Elias, M., Dupuy, J., Merone, L., Mandrich, L., Porzio, E., Moniot, S., Rochu, D., Lecomte, C., Rossi, M., Masson, P., Manco, G. & Chabriere, E.** (2008). Structural basis for natural lactonase and promiscuous phosphotriesterase activities. *J Mol Biol* **379**, 1017-28.

104. **Sehgal, A. C., Tompson, R., Cavanagh, J. & Kelly, R. M.** (2002). Structural and catalytic response to temperature and cosolvents of carboxylesterase EST1 from the extremely thermoacidophilic archaeon *Sulfolobus solfataricus* P1. *Biotechnol Bioeng* **80**, 784-93.

105. **Mandrich, L., Merone, L., Pezzullo, M., Cipolla, L., Nicotra, F., Rossi, M. & Manco, G.** (2005). Role of the N-terminus in enzyme activity, stability and specificity in thermophilic esterases belonging to the HSL family. *Journal of Molecular Biology* **345**, 501-512.

106. **Mandrich, L., Pezzullo, M., Rossi, M. & Manco, G.** (2007). SSoNDelta and SSoNDeltalong: two thermostable esterases from the same ORF in the archaeon *Sulfolobus solfataricus*? *Archaea* **2**, 109-15.

107. **Chung, Y., Park, C. & Lee, S.** (2000). Partial purification and characterization of thermostable esterase from the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Biotechnology and Bioprocess Engineering* **5**, 53-56.

108. **Suzuki, Y., Miyamoto, K. & Ohta, H.** (2004). A novel thermostable esterase from the thermoacidophilic archaeon *Sulfolobus tokodaii* strain 7. *FEMS Microbiol Lett* **236**, 97-102.

109. **Byun, J. S., Rhee, J. K., Kim, D. U., Oh, J. W. & Cho, H. S.** (2006). Crystallization and preliminary X-ray crystallographic analysis of EstE1, a new and thermostable esterase cloned from a metagenomic library. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**, 145-7.

110. **Arpigny, J. L. & Jaeger, K. E.** (1999). Bacterial lipolytic enzymes: classification and properties. *Biochem J* **343**, 177-83.

111. **Fischer, M. & Pleiss, J.** (2003). The Lipase Engineering Database: a navigation and analysis tool for protein families. *Nucleic Acids Res* **31**, 319-21.

112. **Pleiss, J., Fischer, M., Peiker, M., Thiele, C. & Schmid, R. D.** (2000). Lipase engineering database: Understanding and exploiting sequence-structure-function relationships. *Journal of Molecular Catalysis B: Enzymatic* **10**, 491-508.

113. **Kang, H. Y., Kim, J. F., Kim, M. H., Park, S. H., Oh, T. K. & Hur, C. G.** (2006). MELDB: a database for microbial esterases and lipases. *FEBS Lett* **580**, 2736-40.

114. **Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. & Henrissat, B.** (2009). The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res* **37**, D233-8.

115. **Hotelier, T., Renault, L., Cousin, X., Negre, V., Marchot, P. & Chatonnet, A.** (2004). ESTHER, the

database of the alpha/beta-hydrolase fold superfamily of proteins. *Nucleic Acids Res* **32**, D145-7.

116. **Shao, W. L. & Wiegel, J.** (1995). Purification and characterization of 2 thermostable acetyl xylan esterases from *Thermoanaerobacterium* sp strain JW/SL-YS485. *Applied and Environmental Microbiology* **61**, 729-733.

117. **Tatusov, R. L., Fedorova, N. D., Jackson, J. D., Jacobs, A. R., Kiryutin, B., Koonin, E. V., Krylov, D. M., Mazumder, R., Mekhedov, S. L., Nikolskaya, A. N., Rao, B. S., Smirnov, S., Sverdlov, A. V., Vasudevan, S., Wolf, Y. I., Yin, J. J. & Natale, D. A.** (2003). The COG database: an updated version includes eukaryotes. *Bmc Bioinformatics* **4**, -.

118. **Tatusov, R. L., Koonin, E. V. & Lipman, D. J.** (1997). A genomic perspective on protein families. *Science* **278**, 631-637.

119. **Nelson, K. E., Clayton, R. A., Gill, S. R., Gwinn, M. L., Dodson, R. J., Haft, D. H., Hickey, E. K., Peterson, J. D., Nelson, W. C., Ketchum, K. A., McDonald, L., Utterback, T. R., Malek, J. A., Linher, K. D., Garrett, M. M., Stewart, A. M., Cotton, M. D., Pratt, M. S., Phillips, C. A., Richardson, D., Heidelberg, J., Sutton, G. G., Fleischmann, R. D., Eisen, J. A., White, O., Salzberg, S. L., Smith, H. O., Venter, J. C. & Fraser, C. M.** (1999). Evidence for lateral gene transfer between Archaea and bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**, 323-329.

120. **Lim, D. & Strynadka, N. C. J.** (2002). Structural basis for the beta-lactam resistance of PBP2a from methicillin-resistant *Staphylococcus aureus*. *Nature Structural Biology* **9**, 870-876.

121. **Holmquist, M.** (2000). Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms. *Curr Protein Pept Sci* **1**, 209-35.

122. **Bates, P. A., Kelley, L. A., MacCallum, R. M. & Sternberg, M. J.** (2001). Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM. *Proteins* **Suppl 5**, 39-46.

123. **Liu, P., Wang, Y. F., Ewis, H. E., Abdelal, A. T., Lu, C. D., Harrison, R. W. & Weber, I. T.** (2004). Covalent reaction intermediate revealed in crystal structure of the *Geobacillus stearothermophilus* carboxylesterase Est30. *Journal of Molecular Biology* **342**, 551-561.

124. **Laskowski, R. A., MacArthur, M. W. , Moss, D. S. & Thornton, J. M.** (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of applied crystallography* **26**, 283-291.

125. **Vriend, G.** (1990). WHAT IF: a molecular modeling and drug design program. *J Mol Graph* **8**, 52-6, 29.

126. **Handrick, R., Reinhardt, S., Focarete, M. L., Scandola, M., Adamus, G., Kowalczuk, M. & Jendrossek, D.** (2001). A new type of thermoalkalophilic hydrolase of *Paucimonas lemoignei* with high specificity for amorphous polyesters of short chain-length hydroxyalkanoic acids. *J Biol Chem* **276**, 36215-24.

127. **Kim, R., Sandler, S. J., Goldman, S., Yokota, H., Clark, A. J. & Kim, S. H.** (1998). Overexpression of archaeal proteins in *Escherichia coli*. *Biotechnology Letters* **20**, 207-210.

128. **Sorensen, H. P., Sperling-Petersen, H. U. & Mortensen, K. K.** (2003). Production of recombinant thermostable proteins expressed in *Escherichia coli*: completion of protein synthesis is the bottleneck. *J Chromatogr B Analyt Technol Biomed Life Sci* **786**, 207-14.

129. **Sambrook, J., Maniatis, T. & Fritsch, E. F.** (1989). Molecular cloning : a laboratory manual 2nd edit. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

130. **Kanehisa, M.** (2002). The KEGG database. *In Silico Simulation of Biological Processes* **247**, 91-103.

131. **Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S.** (2004). Improved prediction of signal peptides: SignalP 3.0. *Journal of Molecular Biology* **340**, 783-795.

132. **Notredame, C., Higgins, D. G. & Heringa, J.** (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology* **302**, 205-217.

133. **Page, R. D. M.** (1996). TreeView: An application to display phylogenetic trees on personal computers. *Computer Applications in the Biosciences* **12**, 357-358.

134. **DeLano, W. L.** (2002). The Pymol molecular graphics system. *DeLano Scientific, San Carlos, CA*.

135. **Kobs, G.** (1997). Cloning blunt-end DNA fragments into the pGEM-T vector systems. *Promega notes* **62**, 15.

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

137. **Akerboom, J., Turnbull, A. P., Hargreaves, D., Fisher, M., de Geus, D., Sedelnikova, S. E., Berrisford, J. M., Baker, P. J., Verhees, C. H., van der Oost, J. & Rice, D. W.** (2003). Purification, crystallization and preliminary crystallographic analysis of phosphoglucose isomerase from the hyperthermophilic archaeon *Pyrococcus furiosus*. *Acta Crystallographica Section D-Biological Crystallography* **59**, 1822-1823.

138. **Matthews, B. W.** (1968). Solvent content of protein crystals. *J Mol Biol* **33**, 491-7.

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

140. **Evans, P. R.** (1993) Data reduction.

141. **Collaborative Computational Project Number 4.** (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**, 760-3.

142. **Huber, R., Hannig, M.** (2006). *Thermotogales*. *Prokaryotes* **7**, 899-922.

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

144. **Ito, K., Nakajima, Y., Xu, Y., Yamada, N., Onohara, Y., Ito, T., Matsubara, F., Kabashima, T., Nakayama, K. & Yoshimoto, T.** (2006). Crystal structure and mechanism of tripeptidyl activity of prolyl tripeptidyl aminopeptidase from *Porphyromonas gingivalis*. *J Mol Biol* **362**, 228-40.

145. **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.

146. **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**, 12524-32.

147. **Zhang, H., Seabra, M. C. & Deisenhofer, J.** (2000). Crystal structure of Rab geranylgeranyltransferase at 2.0 A resolution. *Structure* **8**, 241-51.

148. **Hage, T., Sebald, W. & Reinemer, P.** (1999). Crystal structure of the interleukin-4/receptor alpha chain complex reveals a mosaic binding interface. *Cell* **97**, 271-81.

149. **Jacobson, R. H., Zhang, X. J., DuBose, R. F. & Matthews, B. W.** (1994). Three-dimensional structure of beta-galactosidase from *E. coli*. *Nature* **369**, 761-6.

150. **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.

151. **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**, 719-28.

152. **van den Heuvel, R. H. & Heck, A. J.** (2004). Native protein mass spectrometry: from intact oligomers to functional machineries. *Curr Opin Chem Biol* **8**, 519-26.

153. **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. *The EMBO journal* **21**, 5343-5352.

154. **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. *European journal of biochemistry* **254**, 333-340.

155. **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**, 173-9.

156. **Bork, P., Holm, L. & Sander, C.** (1994). The immunoglobulin fold: Structural classification, sequence patterns and common core. *J Mol Biol* **242**, 309-20.

157. **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.

158. **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 science* **8**, 1241-1249.

159. **Pape, T., Schneider, T.R.** (2004). HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *Journal of applied crystallography* **37**, 843-844.

160. **Terwilliger, T. C.** (2000). Maximum-likelihood density modification. *Acta Crystallogr D Biol Crystallogr* **56**, 965-72.

161. **Terwilliger, T. C.** (2003). Automated side-chain model building and sequence assignment by template matching. *Acta Crystallogr D Biol Crystallogr* **59**, 45-9.

162. **Cowtan, K.** (1994). 'dm': An automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31**, 34-38.

163. **Perrakis, A., Morris, R. & Lamzin, V. S.** (1999). Automated protein model building combined with iterative structure refinement. *Nat Struct Biol* **6**, 458-63.

164. **Emsley, P. & Cowtan, K.** (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-32.

165. **Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L.** (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-21.

166. **Murshudov, G. N., Vagin, A. A. & Dodson, E. J.** (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-55.

167. **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.

168. **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**, 435-9.

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

170. **van Heel, M., Gowen, B., Matadeen, R., Orlova, E. V., Finn, R., Pape, T., Cohen, D., Stark, H., Schmidt, R., Schatz, M. & Patwardhan, A.** (2000). Single-particle electron cryo-microscopy: towards atomic resolution. *Q Rev Biophys* **33**, 307-69.

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

172. **Chhabra, S. R., Shockley, K. R., Conners, S. B., Scott, K. L., Wolfinger, R. D. & Kelly, R. M.** (2003). Carbohydrate-induced differential gene expression patterns in the hyperthermophilic bacterium *Thermotoga maritima*. *Journal of Biological Chemistry* **278**, 7540-7552.

173. **Conners, S. B., Montero, C. I., Comfort, D. A., Shockley, K. R., Johnson, M. R., Chhabra, S. R. & Kelly, R. M.** (2005). An expression-driven approach to the prediction of carbohydrate transport and utilization regulons in the hyperthermophilic bacterium *Thermotoga maritima*. *Journal of Bacteriology* **187**, 7267-7282.

174. **VanFossen, A. L., Lewis, D. L., Nichols, J. D. & Kelly, R. M.** (2008). Polysaccharide degradation and synthesis by extremely nermophilic anaerobes. *Incredible Anaerobes: From Physiology to Genomics to Fuels* **1125**, 322-337.

175. **Biely, P., Mackenzie, C. R., Puls, J. & Schneider, H.** (1986). Cooperativity of esterases and xylanases in the enzymatic degradation of acetyl xylan. *Bio-Technology* **4**, 731-733.

176. **Topakas, E. & Paul, C.** (2007). Microbial xylanolytic carbohydrate esterases. In *Industrial Enzymes*, pp. 83-97.

177. **Weil, J., Miramonti, J. & Ladisch, M. R.** (1995). Cephalosporin-C - Mode of action and biosynthetic-pathway. *Enzyme and Microbial Technology* **17**, 85-87.

178. **Barends, T. R. M., Yoshida, H. & Dijkstra, B. W.** (2004). Three-dimensional structures of enzymes useful for beta-lactam antibiotic production. *Curr Opin Biotechnol* **15**, 356-363.

179. **Martínez-Martínez, I., Montoro-García, S., Lozada-Ramírez, J. D., Sánchez-Ferrer, Á. & García-Carmona, F.** (2007). A colorimetric assay for the determination of acetyl xylan esterase or cephalosporin C acetyl esterase activities using 7-amino cephalosporanic acid, cephalosporin C,

or acetylated xylan as substrate. *Analytical Biochemistry* **369**, 210-217.

180. **Lesley, S. A., Kuhn, P., Godzik, A., Deacon, A. M., Mathews, I., Kreusch, A., Spraggon, G., Klock, H. E., McMullan, D., Shin, T., Vincent, J., Robb, A., Brinen, L. S., Miller, M. D., McPhillips, T. M., Miller, M. A., Scheibe, D., Canaves, J. M., Guda, C., Jaroszewski, L., Selby, T. L., Elsliger, M. A., Wooley, J., Taylor, S. S., Hodgson, K. O., Wilson, I. A., Schultz, P. G. & Stevens, R. C.** (2002). Structural genomics of the *Thermotoga maritima* proteome implemented in a high-throughput structure determination pipeline. *PNAS* **99**, 11664-11669.

181. **Jiang, Z. Q., Kobayashi, A., Ahsan, M. M., Lite, L., Kitaoka, M. & Hayashi, K.** (2001). Characterization of a thermostable family 10 endo-xylanase (XynB) from *Thermotoga maritima* that cleaves p-nitrophenyl-beta-D-xyloside. *Journal of Bioscience and Bioengineering* **92**, 423-428.

182. **Xue, Y. M. & Shao, W. L.** (2004). Expression and characterization of a thermostable beta-xylosidase from the hyperthermophile, *Thermotoga maritima*. *Biotechnology Letters* **26**, 1511-1515.

183. **Vincent, F., Charnock, S. J., Verschueren, K. H. G., Turkenburg, J. P., Scott, D. J., Offen, W. A., Roberts, S., Pell, G., Gilbert, H. J., Davies, G. J. & Brannigan, J. A.** (2003). Multifunctional xylooligosaccharide/cephalosporin C deacetylase revealed by the hexameric structure of the *Bacillus subtilis* enzyme at 1.9 angstrom resolution. *Journal of Molecular Biology* **330**, 593-606.

184. **Bond, C. S.** (2003). TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics* **19**, 311-312.

185. **Henrick, K. & Thornton, J. M.** (1998). PQS: a protein quaternary structure file server. *Trends in Biochemical Sciences* **23**, 358-361.

186. **Krissinel, E. & Henrick, K.** (2007). Inference of macromolecular assemblies from crystalline state. *Journal of Molecular Biology* **372**, 774-797.

187. **Holm, L., Kaariainen, S., Rosenstrom, P. & Schenkel, A.** (2008). Searching protein structure databases with DaliLite v.3. *Bioinformatics* **24**, 2780-2781.

188. **Dundas, J., Ouyang, Z., Tseng, J., Binkowski, A., Turpaz, Y. & Liang, J.** (2006). CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Research* **34**, W116-W118.

189. **Longhi, S., Czjzek, M., Lamzin, V., Nicolas, A. & Cambillau, C.** (1997). Atomic resolution (1.0 angstrom) crystal structure of *Fusarium solani* cutinase: Stereochemical analysis. *Journal of Molecular Biology* **268**, 779-799.

190. **Ghosh, D., Sawicki, M., Lala, P., Erman, M., Pangborn, W., Eyzaguirre, J., Gutierrez, R., Jornvall, H. & Thiel, D. J.** (2001). Multiple conformations of catalytic serine and histidine in acetylxylan esterase at 0.90 angstrom. *Journal of Biological Chemistry* **276**, 11159-11166.

191. **Kawasaki, K., Kondo, H., Suzuki, M., Ohgiya, S. & Tsuda, S.** (2002). Alternate conformations observed in catalytic serine of *Bacillus subtilis* lipase determined at 1.3 angstrom resolution. *Acta Crystallographica Section D-Biological Crystallography* **58**, 1168-1174.

192. **McAuley, K. E., Svendsen, A., Patkar, S. A. & Wilson, K. S.** (2004). Structure of a feruloyl esterase from *Aspergillus niger*. *Acta Crystallographica Section D-Biological Crystallography* **60**, 878-887.

193. **Williamson, G., Kroon, P. A. & Faulds, C. B.** (1998). Hairy plant polysaccharides: a close shave with microbial esterases. *Microbiology-Sgm* **144**, 2011-2023.

194. **Biely, P., Mastihubova, M. & Puchart, V.** (2007). The vicinal hydroxyl group is prerequisite for metal activation of *Clostridium thermocellum* acetylxylan esterase. *Biochimica Et Biophysica Acta-General Subjects* **1770**, 565-570.

195. **Colombres, M., Garate, J. A., Lagos, C. F., Araya-Secchi, R., Norambuena, P., Quiroz, S., Larrondo, L., Perez-Acle, T. & Eyzaguirre, J.** (2008). An eleven amino acid residue deletion expands the substrate specificity of acetyl xylan esterase II (AXE II) from *Penicillium purpurogenum*. *Journal of Computer-Aided Molecular Design* **22**, 19-28.

196. **Degrassi, G., Kojic, M., Ljubijankic, G. & Venturi, V.** (2000). The acetyl xylan esterase of *Bacillus pumilus* belongs to a family of esterases with broad substrate specificity. *Microbiology-Uk* **146**, 1585-1591.

197. **Lorenz, W. W. & Wiegel, J.** (1997). Isolation, analysis, and expression of two genes from *Thermoanaerobacterium* sp. strain JW/SL YS485: a beta-xylosidase and a novel acetyl xylan

esterase with cephalosporin C deacetylase activity. *Journal of Bacteriology* **179**, 5436-5441.

198. **Otwinowski, Z. & Minor, W.** (1997). Processing of X-ray diffraction data collected in oscillation mode. *Macromolecular Crystallography, Pt A* **276**, 307-326.

199. **Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M.** (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47**, 110-9.

200. **Winn, M. D., Murshudov, G. N. & Papiz, M. Z.** (2003). Macromolecular TLS refinement in REFMAC at moderate resolutions. *Methods Enzymol* **374**, 300-21.

201. **Davis, I. W., Murray, L. W., Richardson, J. S. & Richardson, D. C.** (2004). MOLPROBITY: structure validation and all-atom contact analysis for nucleic acids and their complexes. *Nucleic Acids Res* **32**, W615-9.

202. **McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J.** (2005). Likelihood-enhanced fast translation functions. *Acta Crystallographica Section D-Biological Crystallography* **61**, 458-464.

203. **Storoni, L. C., McCoy, A. J. & Read, R. J.** (2004). Likelihood-enhanced fast rotation functions. *Acta Crystallographica Section D-Biological Crystallography* **60**, 432-438.

204. **Adams, P. D., Grosse-Kunstleve, R. W., Hung, L. W., Ioerger, T. R., McCoy, A. J., Moriarty, N. W., Read, R. J., Sacchettini, J. C., Sauter, N. K. & Terwilliger, T. C.** (2002). PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallographica Section D-Biological Crystallography* **58**, 1948-1954.

205. **Tickle, I. J., Laskowski, R. A. & Moss, D. S.** (1998). Error estimates of protein structure coordinates and deviations from standard geometry by full-matrix refinement of gamma B- and beta B2-crystallin. *Acta Crystallographica Section D-Biological Crystallography* **54**, 243-252.

206. **Johnson, K. G., Harrison, B. A., Schneider, H., Mackenzie, C. R. & Fontana, J. D.** (1988). Xylan-hydrolyzing enzymes from *Streptomyces* spp. *Enzyme and Microbial Technology* **10**, 403-409.

207. **Biely, P., Mackenzie, C. R. & Schneider, H.** (1988). Production of acetyl xylan esterase by *Trichoderma reesei* and *Schizophyllum commune*. *Canadian Journal of Microbiology* **34**, 767-772.

208. **Biely, P., Mastihubova, M., la Grange, D. C., van Zyl, W. H. & Prior, B. A.** (2004). Enzyme-coupled assay of acetylxylan esterases on monoacetylated 4-nitrophenyl beta-D-xylopyranosides. *Analytical Biochemistry* **332**, 109-115.

209. **Mastihubova, M. & Biely, P.** (2004). Lipase-catalysed preparation of acetates of 4-nitrophenyl beta-D-xylopyranoside and their use in kinetic studies of acetyl migration. *Carbohydrate Research* **339**, 1353-1360.

210. **Yang, H., Guranovic, V., Dutta, S., Feng, Z., Berman, H. M. & Westbrook, J. D.** (2004). Automated and accurate deposition of structures solved by X-ray diffraction to the Protein Data Bank. *Acta Crystallogr D Biol Crystallogr* **60**, 1833-9.

211. **Vaguine, A. A., Richelle, J. & Wodak, S. J.** (1999). SFCHECK: a unified set of procedures for evaluating the quality of macromolecular structure-factor data and their agreement with the atomic model. *Acta Crystallogr D Biol Crystallogr* **55**, 191-205.

212. **Noble, M. E. M., Cleasby, A., Johnson, L. N., Egmond, M. R. & Frenken, L. G. J.** (1993). The crystal-structure of triacylglycerol lipase from *Pseudomonas glumae* reveals a partially redundant catalytic aspartate. *Febs Letters* **331**, 123-128.

213. **O'Sullivan, O., Suhre, K., Abergel, C., Higgins, D. G. & Notredame, C.** (2004). 3DCoffee: Combining protein sequences and structures within multiple sequence alignments. *Journal of Molecular Biology* **340**, 385-395.

214. **Emanuelsson, O., Brunak, S., von Heijne, G. & Nielsen, H.** (2007). Locating proteins in the cell using TargetP, SignalP and related tools. *Nature Protocols* **2**, 953-971.

215. **Boutaiba, S., Bhatnagar, T., Hacene, H., Mitchell, D. A. & Baratti, J. C.** (2006). Preliminary characterisation of a lipolytic activity from an extremely halophilic archaeon, *Natronococcus* sp. *Journal of Molecular Catalysis B-Enzymatic* **41**, 21-26.

216. **Ozcan, B., Ozyilmaz, G., Cokmus, C. & Caliskan, M.** (2009). Characterization of extracellular esterase and lipase activities from five halophilic archaeal strains. *Journal of Industrial Microbiology & Biotechnology* **36**, 105-110.

217. **Stetter, K. O., Lauerer, G., Thomm, M. & Neuner, A.** (1987). Isolation of extremely thermophilic sulfate reducers - Evidence for a novel branch of archaebacteria. *Science* **236**, 822-824.

218. **Hartzell, P. & Reed, D.** (2006). The genus *Archaeoglobus*. In *The Prokaryotes,* pp. 82-100.

219. **Klenk, H. P., Clayton, R. A., Tomb, J. F., White, O., Nelson, K. E., Ketchum, K. A., Dodson, R. J., Gwinn, M., Hickey, E. K., Peterson, J. D., Richardson, D. L., Kerlavage, A. R., Graham, D. E., Kyrpides, N. C., Fleischmann, R. D., Quackenbush, J., Lee, N. H., Sutton, G. G., Gill, S., Kirkness, E. F., Dougherty, B. A., McKenney, K., Adams, M. D., Loftus, B., Peterson, S., Reich, C. I., McNeil, L. K., Badger, J. H., Glodek, A., Zhou, L. X., Overbeek, R., Gocayne, J. D., Weidman, J. F., McDonald, L., Utterback, T., Cotton, M. D., Spriggs, T., Artiach, P., Kaine, B. P., Sykes, S. M., Sadow, P. W., DAndrea, K. P., Bowman, C., Fujii, C., Garland, S. A., Mason, T. M., Olsen, G. J., Fraser, C. M., Smith, H. O., Woese, C. R. & Venter, J. C.** (1997). The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. *Nature* **390**, 364-370.

220. **Bennett-Lovsey, R. M., Herbert, A. D., Sternberg, M. J. E. & Kelley, L. A.** (2008). Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. *Proteins-Structure Function and Bioinformatics* **70**, 611-625.

221. **Schrag, J. D., Winkler, F. K. & Cygler, M.** (1992). Pancreatic lipases: evolutionary intermediates in a positional change of catalytic carboxylates? *J Biol Chem* **267**, 4300-3.

222. **Chen, C. K., Lee, G. C., Ko, T. P., Guo, R. T., Huang, L. M., Liu, H. J., Ho, Y. F., Shaw, J. F. & Wang, A. H.** (2009). Structure of the alkalohyperthermophilic Archaeoglobus fulgidus lipase contains a unique C-terminal domain essential for long-chain substrate binding. *J Mol Biol* **390**, 672-85.

223. **Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J.** (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403-10.

224. **Nilsen, R. K., Beeder, J., Thorstenson, T. & Torsvik, T.** (1996). Distribution of thermophilic marine sulfate reducers in North Sea oil field waters and oil reservoirs. *Applied and Environmental Microbiology* **62**, 1793-1798.

225. **Stetter, K. O., Huber, R., Blochl, E., Kurr, M., Eden, R. D., Fielder, M., Cash, H. & Vance, I.** (1993). Hyperthermophilic Archaea are thriving in deep North-Sea and Alaskan oil-reservoirs. *Nature* **365**, 743-745.

226. **Secundo, F., Carrea, G., Tarabiono, C., Gatti-Lafranconi, P., Brocca, S., Lotti, M., Jaeger, K. E., Puls, M. & Eggert, T.** (2006). The lid is a structural and functional determinant of lipase activity and selectivity. *Journal of Molecular Catalysis B-Enzymatic* **39**, 166-170.

227. **Stetter, K. O. & Huber, R.** (1999). The role of hyperthermophilic prokaryotes in oil fields. In *Microbial biosystems: new frontiers. Proceedings of the 8th International Symposium on Microbial Ecology* (Bell, C. R., Brylinsky, M. & Johnson-Green, P., eds.). Atlantic Canada Society for Microbial Ecology, Halifax, Canada.

228. **Choi, Y. J. & Lee, B. H.** (2001). Culture conditions for the production of esterase from *Lactobacillus casei* CL96. *Bioprocess and Biosystems Engineering* **24**, 59-63.

229. **Dalmau, E., Montesinos, J. L., Lotti, M. & Casas, C.** (2000). Effect of different carbon sources on lipase production by *Candida rugosa*. *Enzyme and Microbial Technology* **26**, 657-663.

230. **Sousa, D. Z., Pereira, M. A., Stams, A. J. M., Alves, M. M. & Smidt, H.** (2007). Microbial communities involved in anaerobic degradation of unsaturated or saturated long-chain fatty acids. *Applied and Environmental Microbiology* **73**, 1054-1064.

231. **Mozhaev, V. V., Heremans, K., Frank, J., Masson, P. & Balny, C.** (1996). High pressure effects on protein structure and function. *Proteins-Structure Function and Genetics* **24**, 81-91.

232. **Bruins, M. E., Janssen, A. E. M. & Boom, R. M.** (2006). Equilibrium shifts in enzyme reactions at high pressure. *Journal of Molecular Catalysis B-Enzymatic* **39**, 124-127.

233. **Kappe, C. O.** (2008). Microwave dielectric heating in synthetic organic chemistry. *Chemical Society Reviews* **37**, 1127-1139.

234. **Lidstrom, P., Tierney, J., Wathey, B. & Westman, J.** (2001). Microwave assisted organic synthesis - a review. *Tetrahedron* **57**, 9225-9283.

235. **Kappe, C. O.** (2004). Controlled microwave heating in modern organic synthesis. *Angewandte Chemie-International Edition* **43**, 6250-6284.

236. **Kappe, C. O. & Stadler, A.** (2005). *Microwaves in organic and medicinal chemistry*, Wiley-VCH, Weinheim.

237. **Rejasse, B., Lamare, S., Legoy, M. D. & Besson, T.** (2004). Stability improvement of immobilized *Candida antarctica* lipase B in an organic medium under microwave radiation. *Organic & Biomolecular Chemistry* **2**, 1086-1089.

238. **Parker, M. C., Besson, T., Lamare, S. & Legoy, M. D.** (1996). Microwave radiation can increase the rate of enzyme-catalysed reactions in organic media. *Tetrahedron Letters* **37**, 8383-8386.

239. **Yu, D. H., Wang, Z., Chen, P., Jin, L., Cheng, Y. M., Zhou, J. G. & Cao, S. G.** (2007). Microwave-assisted resolution of (R,S)-2-octanol by enzymatic transesterification. *Journal of Molecular Catalysis B-Enzymatic* **48**, 51-57.

240. **Young, D. D., Nichols, J., Kelly, R. M. & Deiters, A.** (2008). Microwave activation of enzymatic catalysis. *Journal of the American Chemical Society* **130**, 10048-+.

241. **Carrillo-Munoz, J. R., Bouvet, D., Guibe-Jampel, E., Loupy, A. & Petit, A.** (1996). Microwave-promoted lipase-catalyzed reactions. Resolution of (+/-)-1-phenylethanol. *Journal of Organic Chemistry* **61**, 7746-7749.

242. **Yu, D. H., Chen, P., Wang, L., Gu, Q., Li, Y., Wang, Z. & Cao, S. G.** (2007). A chemo-enzymatic process for sequential kinetic resolution of (R,S)-2-octanol under microwave irradiation. *Process Biochemistry* **42**, 1312-1318.

243. **Herrero, M. A., Kremsner, J. M. & Kappe, C. O.** (2008). Nonthermal microwave effects revisited: On the importance of internal temperature monitoring and agitation in microwave chemistry. *Journal of Organic Chemistry* **73**, 36-47.

244. **Kuhnert, N.** (2002). Microwave-assisted reactions in organic synthesis - Are there any nonthermal microwave effects? *Angewandte Chemie-International Edition* **41**, 1863-+.

245. **Angelov, A., Mientus, M., Liebl, S. & Liebl, W.** (2009). A two-host fosmid system for functional screening of (meta)genomic libraries from extreme thermophiles. *Syst Appl Microbiol*.

246. **Folkertsma, S., van Noort, P., Van Durme, J., Joosten, H. J., Bettler, E., Fleuren, W., Oliveira, L., Horn, F., de Vlieg, J. & Vriend, G.** (2004). A family-based approach reveals the function of residues in the nuclear receptor ligand-binding domain. *J Mol Biol* **341**, 321-35.

247. **Joosten, H. J., Han, Y., Niu, W., Vervoort, J., Dunaway-Mariano, D. & Schaap, P. J.** (2008). Identification of fungal oxaloacetate hydrolyase within the isocitrate lyase/PEP mutase enzyme superfamily using a sequence marker-based method. *Proteins* **70**, 157-66.

248. **Leferink, N. G., Fraaije, M. W., Joosten, H. J., Schaap, P. J., Mattevi, A. & van Berkel, W. J.** (2009). Identification of a gatekeeper residue that prevents dehydrogenases from acting as oxidases. *J Biol Chem* **284**, 4392-7.

249. **Boersma, Y. L., Droge, M. J. & Quax, W. J.** (2007). Selection strategies for improved biocatalysts. *Febs Journal* **274**, 2181-2195.

250. **Dalby, P. A.** (2007). Engineering enzymes for biocatalysis. *Recent Pat Biotechnol* **1**, 1-9.

Appendix I: Co-author affiliations

CO-AUTHOR AFFILIATIONS

Sjon Hendriks, Twan Akveld, Peter Swinkels, Paul Gielen, Lei Sun, Servé W.M. Kengen, John van der Oost, Laboratory of Microbiology, Department of Agrotechnology and Food Sciences, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, The Netherlands.

Lei Sun, Bauke W. Dijkstra, Laboratory of Biophysical Chemistry – Protein Crystallography, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands.

Jelle B. Bultema, Laboratory of Biophysical Chemistry – Electron Microscopy, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands.

Arjan Barendregt, Robert H.H. van den Heuvel [a], 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.

Peter Biely, Institute of Chemistry, Slovak Academy of Sciences, Dubravska cesta 9, SK-845 38 Bratislava, Slovakia.

Gye Won Han [c]**, Marc C. Deller** [c]**, Qingping Xu** [d]**, Daniel McMullan, Mitchell D. Miller** [c]**, Frank von Delft** [c]**, Marc-André Elsliger** [c]**, Ashley M. Deacon** [c]**, Scott A. Lesley** [c]**, Ian A. Wilson** [c], Joint Center for Structural Genomics (JCSG), 10550 N. Torrey Pines Road, CA 92037, La Jolla, USA

Lynn Ten Eyck, San Diego Supercomputer Center, University of California San Diego (UCSD), 9500 Gilman Drive, CA 92093-0505, La Jolla, USA

Claus Flensburg, Pietro Roversi [b], Global Phasing Ltd., Sheraton House, Castle Park, Cambridge, UK

Andreas Kreusch, Scott A. Lesley [c], Genomics Institute of the Novartis Research Foundation, 10675 John Jay Hopkins Drive, CA 92121, San Diego, USA

Present address:
[a] Biotech Analytical Development, Schering-Plough Corporation, 5340 BH, Oss, The Netherlands
[b] Laboratory of Molecular Biophysics, Department of Biochemistry, University of Oxford, South Parks Road, Oxford, UK

Additional affiliations:
[c] Department of Molecular Biology, The Scripps Research Institute, 10550 N. Torrey Pines Road, CA 92037, La Jolla, USA
[d] Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Stanford University, 2575 Sand Hill Road, CA 94025, Menlo Park, USA
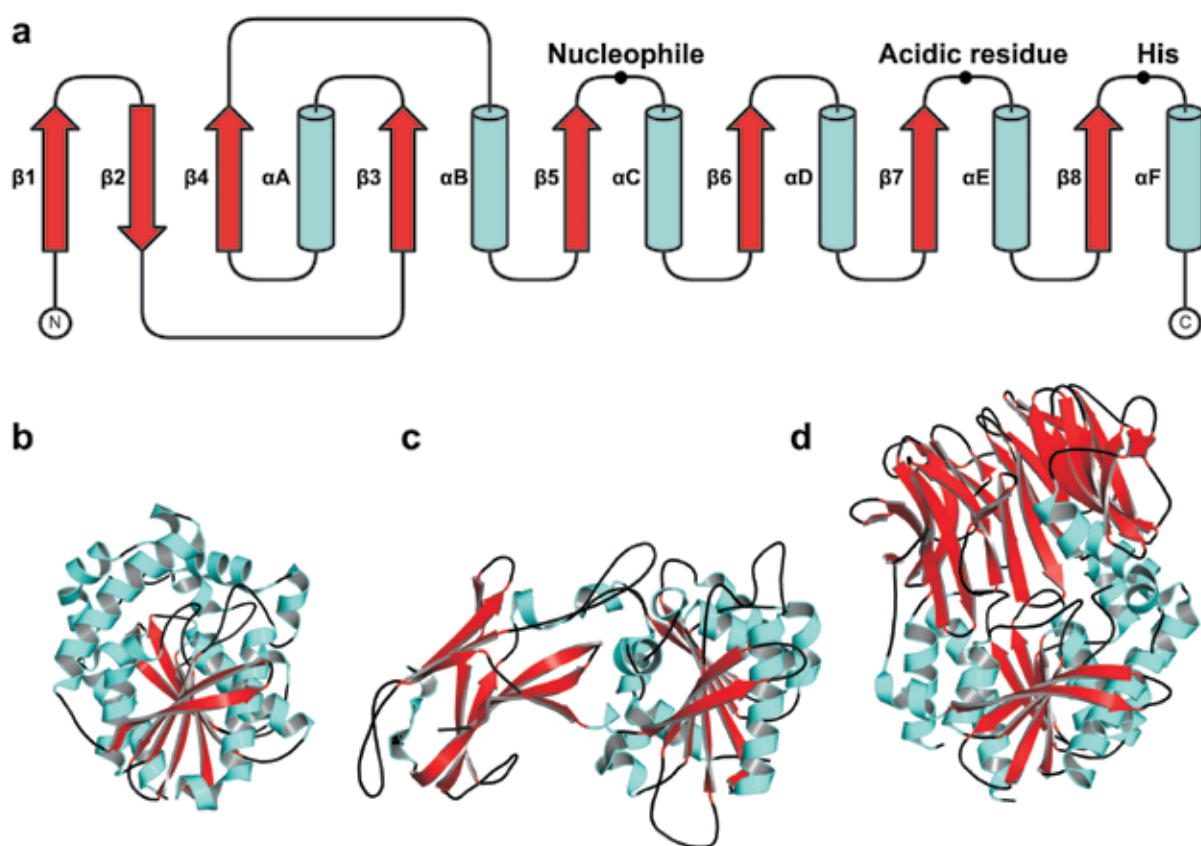
Appendix II: Color figures

Figure 1.3: Canonical fold of α/β-hydrolases. In a) Topology diagram, with the strands indicated by red arrows and the helices by cyan cylinders. The positions of the catalytic residues are indicated. In b-d) the structures of three hyperthermophilic esterases: b) the carboxylesterase AFEST from *A. fulgidus* (pdb 1JJI), c) the esterase EstA from *T. maritima* (pdb 3DOH), and d) the acylpeptide hydrolase apAPH from *A. pernix* (pdb 1VE6).
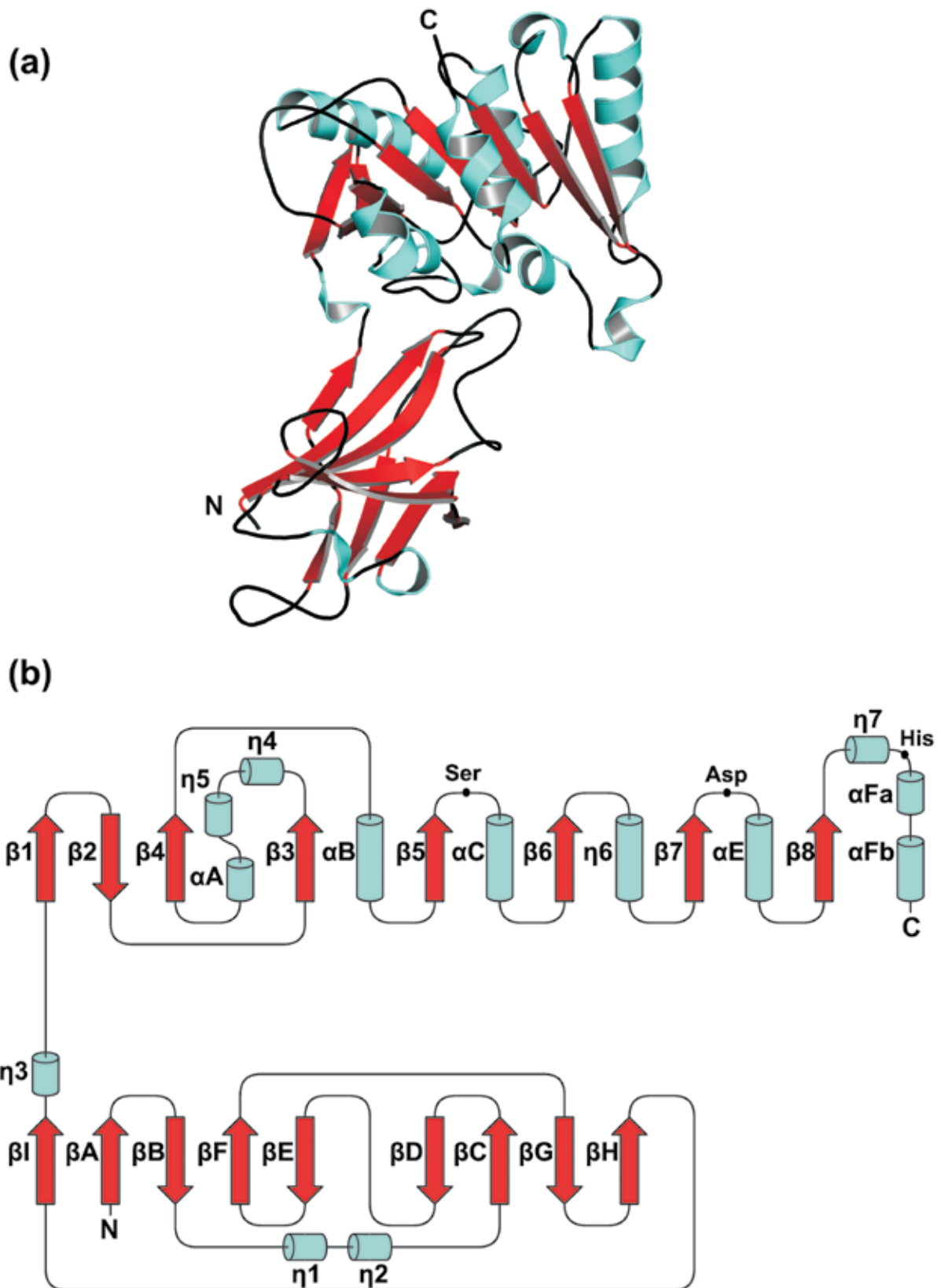
Figure 4.1: (a) Overall fold of the esterase EstA. The N- and C-terminal ends are indicated. The figure was generated using Pymol. (b) Topology diagram for EstA with the helices displayed as cyan cylinders and the strands as red arrows. The positions of the catalytic residues are indicated.
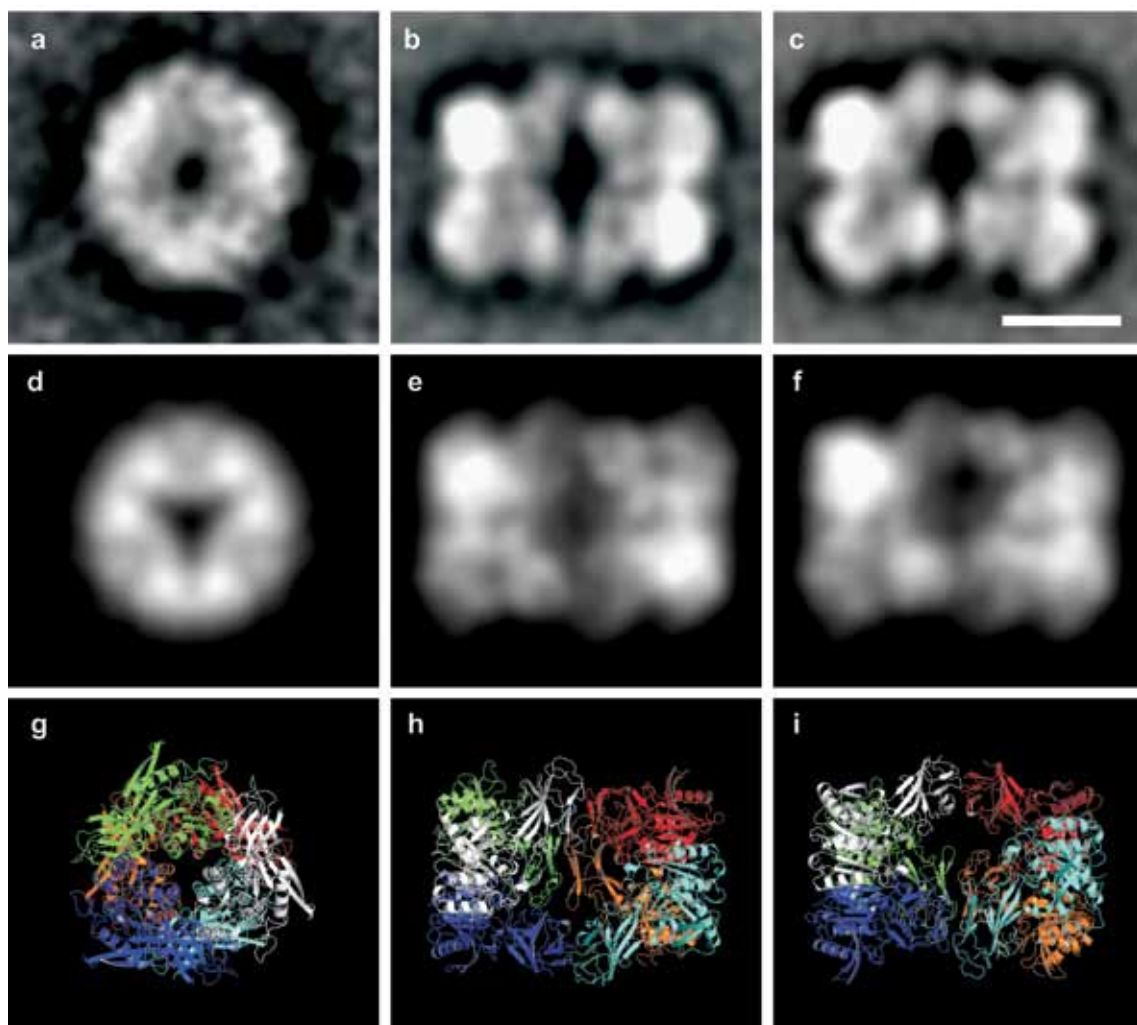
Figure 4.2: Comparison of the EM projection maps of the EstA hexamer. Top view (a), side view (b) and 10 degrees off side view (c). Two-dimensional projection maps obtained by statistical analysis and classification (a - c), the comparable two-dimensional projection maps with 15 Å resolution (d - f) generated from the proposed EstA hexameric structure (g - i). The scale bar equals 50 Å.
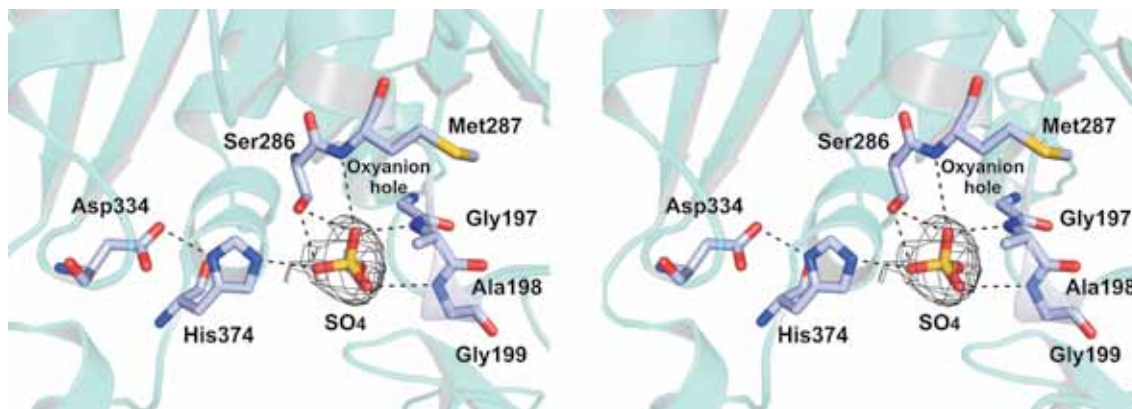
Figure 4.3: Stereoview of the EstA catalytic site with the bound SO_4 adduct shown. The catalytic triad residues are shown as sticks with the hydrogen bonds shown as dashed black lines. Observed density for the bound SO_4 is contoured. These images were generated using Pymol.
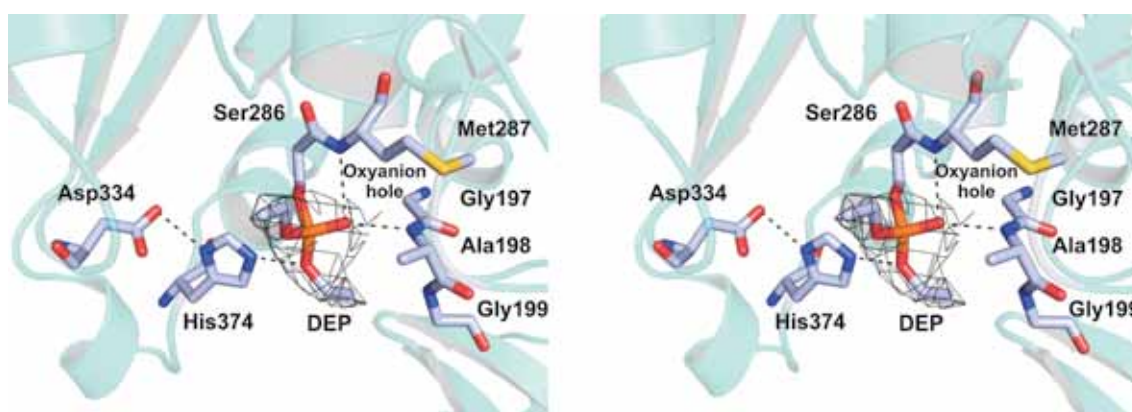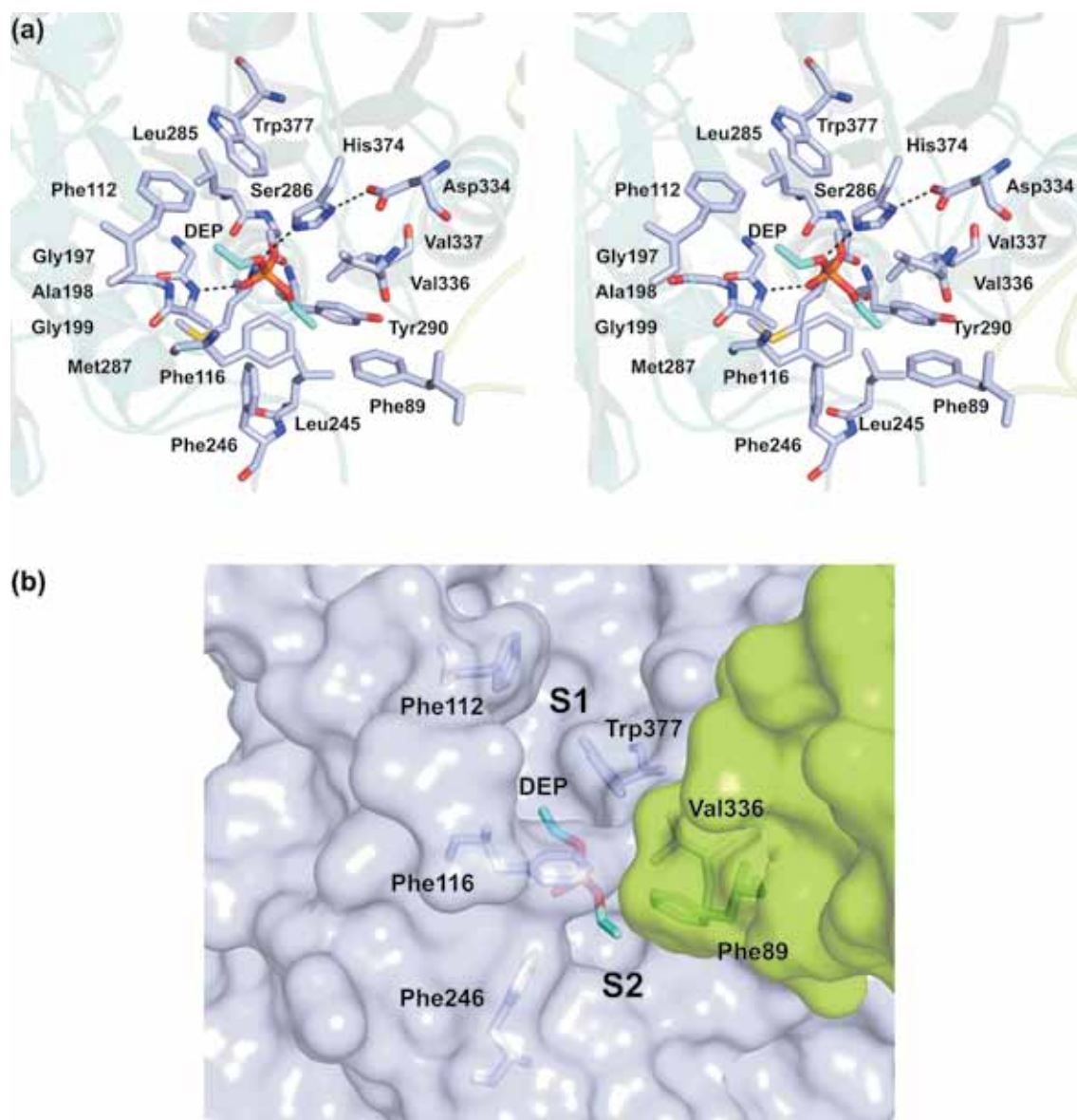

Figure 4.4: Stereoview of the EstA catalytic site with the diethyl phosphate (DEP) intermediate of the inhibitor paraoxon covalently bound to the catalytic serine. The catalytic triad residues are shown as sticks with the hydrogen bonds shown as dashed black lines. Observed density for the bound diethyl phosphate is contoured. These images were generated using Pymol.

Figure 4.5: Substrate binding gorge. (a) Stereoview of the active site with the bound diethyl phosphate 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.
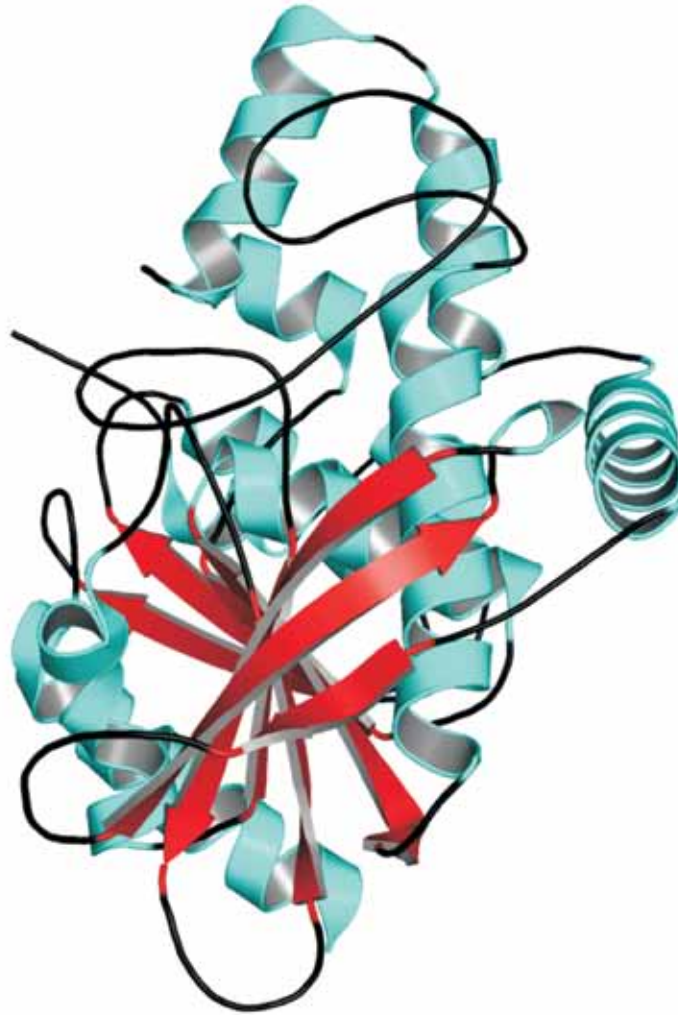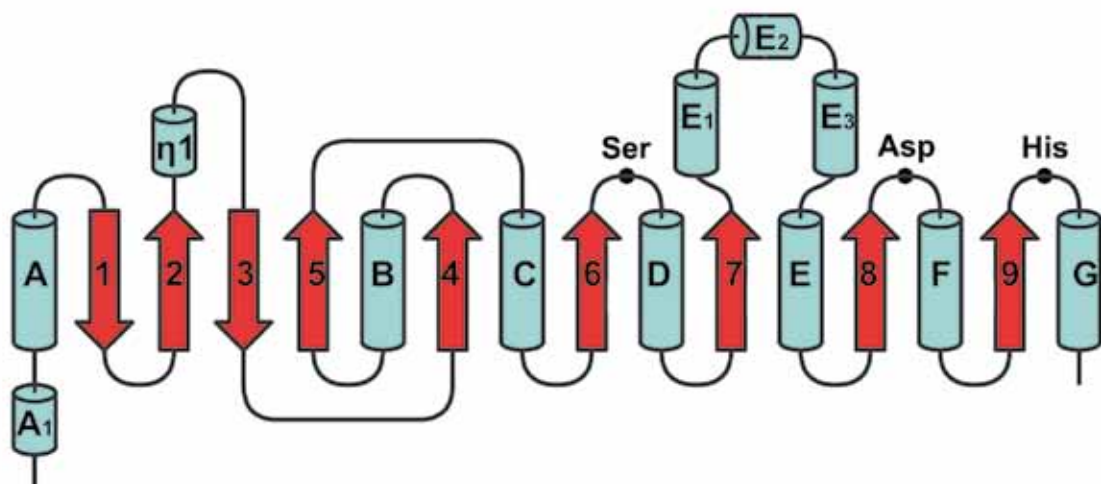
Figure 5.1: (a) Overall fold of the acetyl esterase AceA. The Figure was generated using Pymol (DeLano). (b) Topology diagram for AceA, with the helices displayed as cyan cylinders and the strands displayed as red arrows. The positions of the catalytic residues are indicated. The Figure was generated using TopDraw.
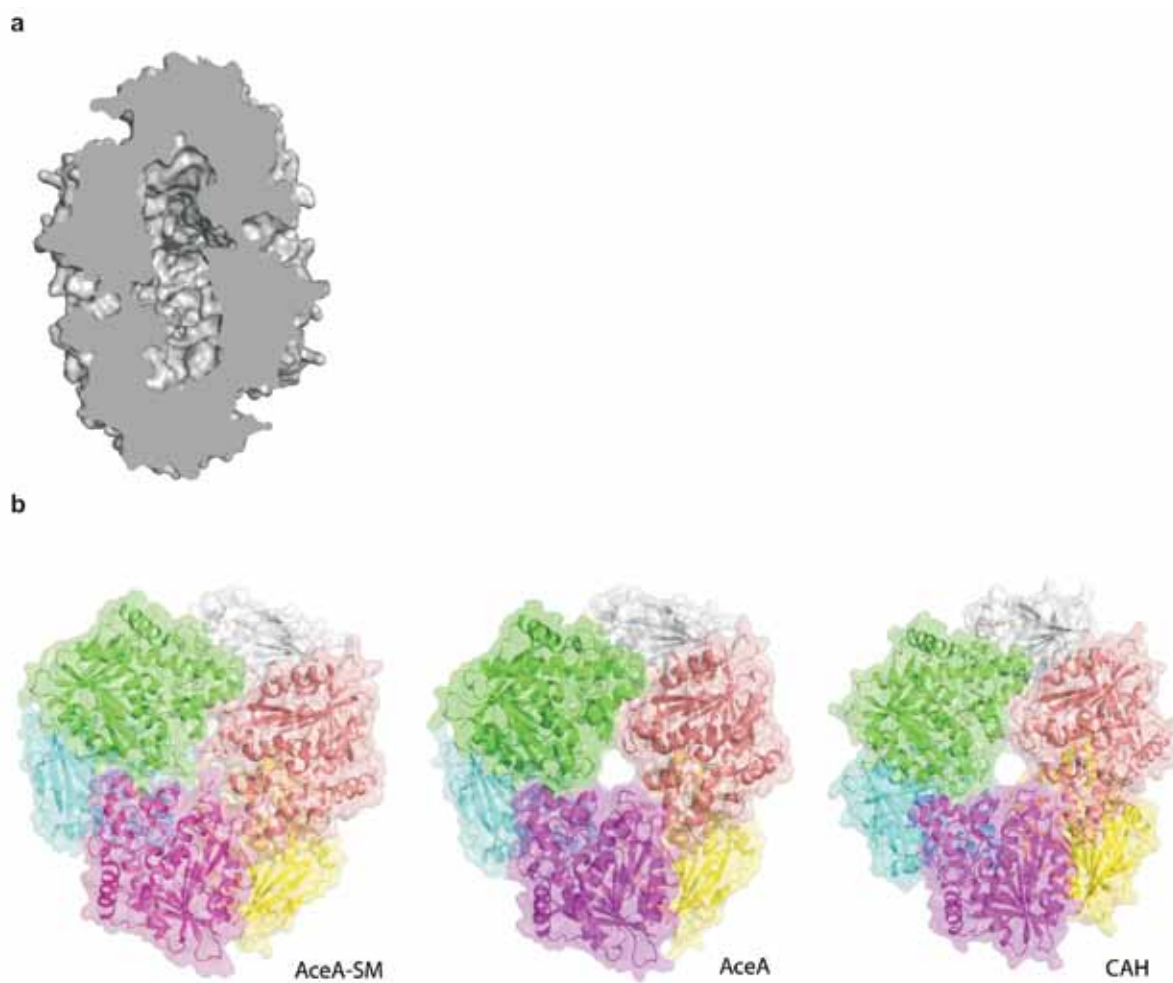
Figure 5.2: (a) A "sliced" image of the AceA hexamer showing the two entrances on each side of the 'doughnut' and the internal cavity, and (b) Surface representation of the biological unit of AceA-SM, AceA and cephalosporin C deacetylase (CAH) from *B. subtilus* with each monomer in a different color.
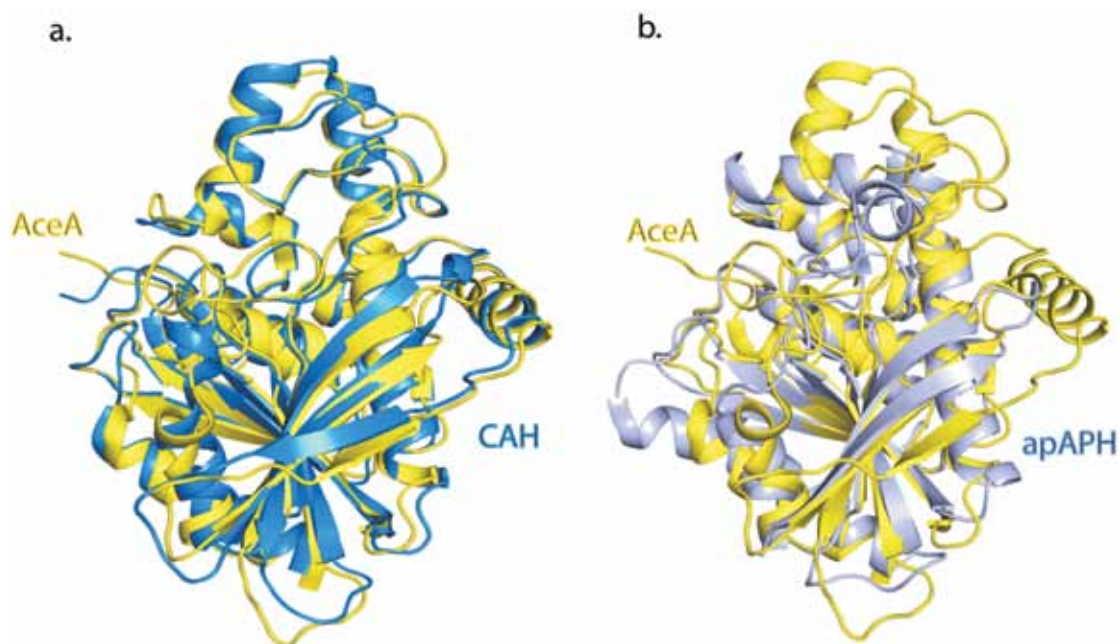
Figure 5.3: Superposition of AceA (*yellow*) with (a) the cephalosporin C deacetylase (CAH) from *B. subtilus* (PDB: 1ODS; *blue*) and (b) the α/β-hydrolase domain of the acylpeptide hydrolase/esterase apAPH from *A. pernix* K1 (PDB: 1VE6; *light blue*).
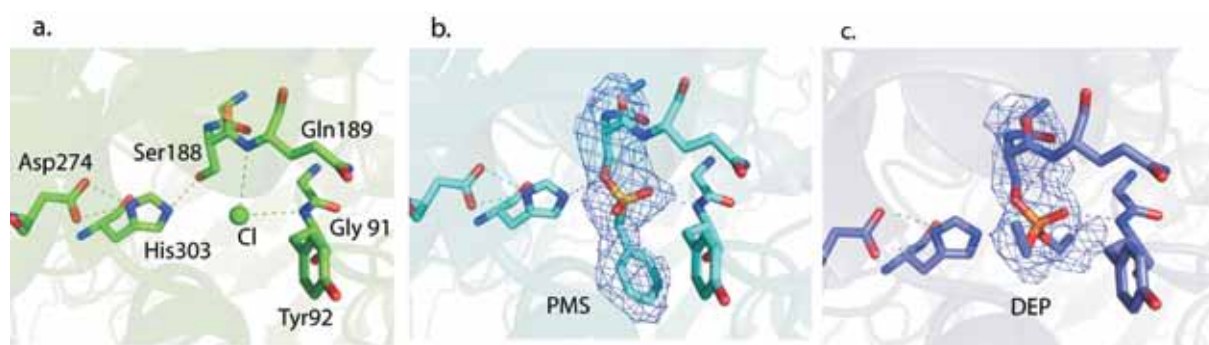

Figure 5.4: A view of the AceA catalytic site. (a) Native AceA with the bound chloride ion, (b) AceA in complex with PMSF and (c) AceA in complex with paraoxon. The catalytic residues are shown as sticks, with the hydrogen bonds shown as dashed lines. Green (AceA), cyan (AceA-PMS) or blue (AceA-DEP) were used for the carbon atoms, red was used for oxygen atoms. The omit maps for Ser188 with the phenylmethyl sulfonyl (PMS) and diethylphosphate (DEP) moieties are contoured at 1σ level.
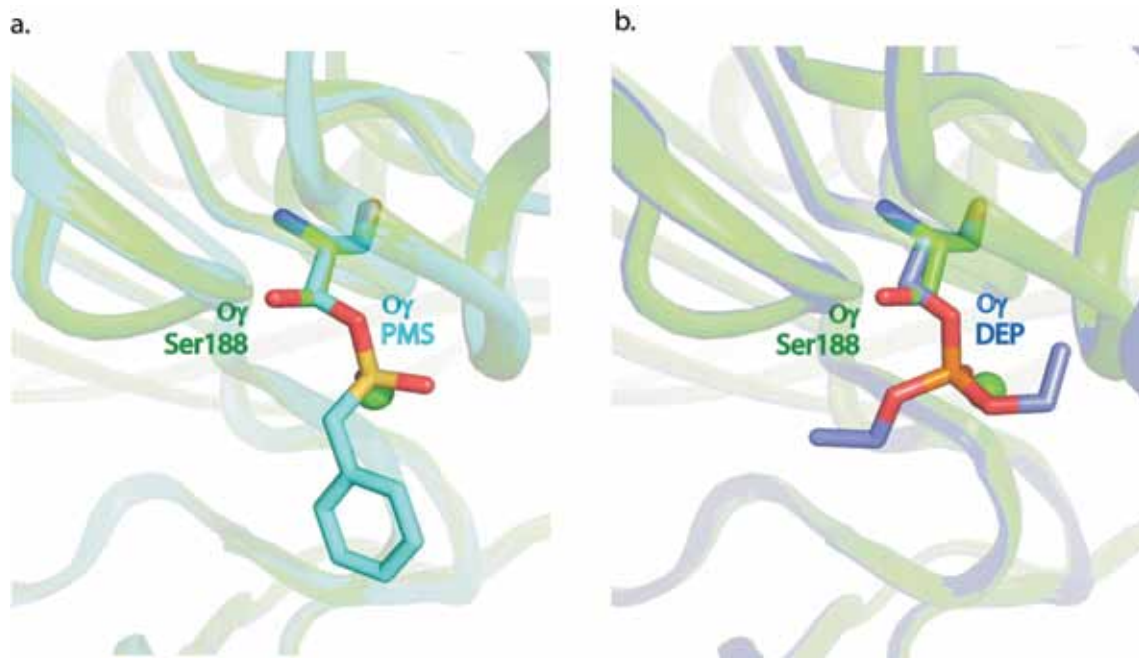
Figure 5.5: Movement of Ser188 O$^\gamma$ from conformation A to B. The O$^\gamma$ atom of the Ser188 is rotated about 110° to the opposite direction between native AceA and the complexed structures. Conformation A in AceA (green), and conformation B in AceA-PMS (cyan) and AceA-DEP (blue) are shown. Similar colors for each structure were used as in Figure 5.4.
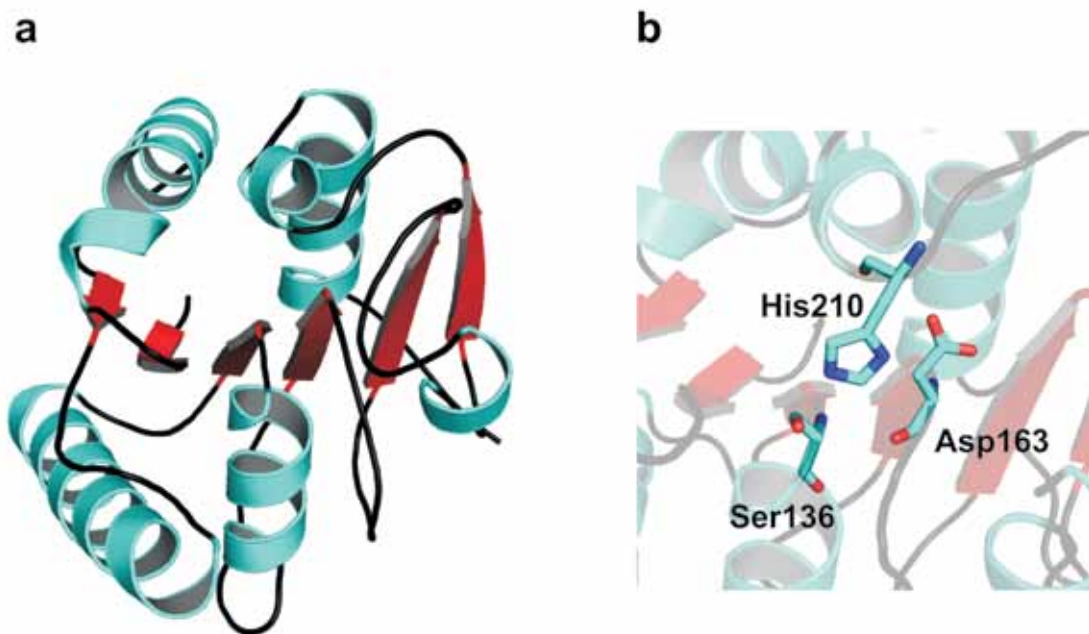

Figure 7.2: 3D model of LipA. In a) The overall structure of the N-terminal domain of LipA, and b) Residues of the potential catalytic triad.
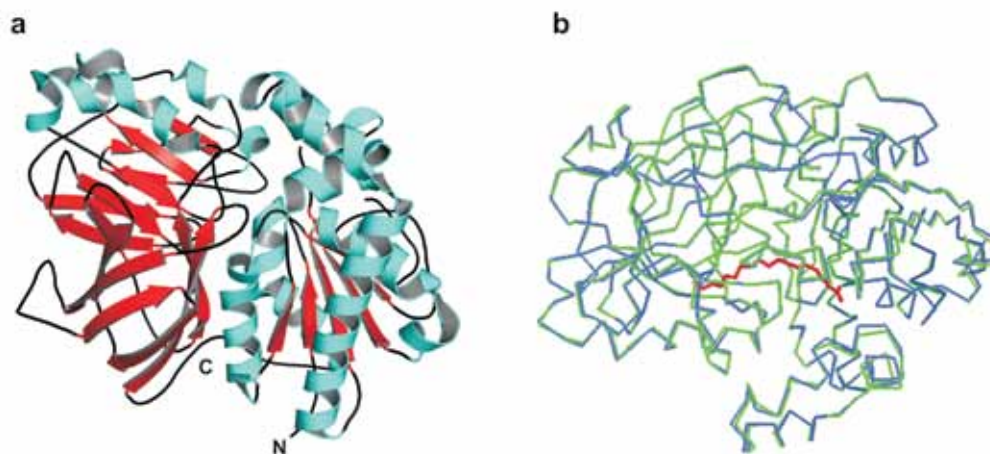
Figure 7.5: (a) Overall fold of the lipase (pdb: 2ZYI). The N-terminal and C-terminal ends are indicated. The helices are displayed in cyan and the strands in red. (b) Superposition of our LipA (blue) and 2ZYI (green). A stearic acid that was present in the 2ZYI structure is shown in red.

Nederlandse samenvatting

NEDERLANDSE SAMENVATTING

Een veelgebruikte techniek in de moderne biotechnologie is de synthese van producten met behulp van enzymen. Enzymen genieten vaak de voorkeur boven chemische processen vanwege een hogere productzuiverheid, minder restproducten, laag energieverbruik en selectievere omzettingen. Een belangrijke groep enzymen, die veel gebruikt wordt in de industrie, zijn de carboxylester-hydrolases. Deze enzymen katalyseren de hydrolyse van esterverbindingen in een waterige omgeving, maar in organische oplosmiddelen kunnen ze eveneens de synthese en de trans-esterificatie-reactie katalyseren.

De carboxylester-hydrolases worden veelal onderverdeeld in twee groepen, namelijk de esterases en de lipases. Het onderscheid tussen beide groepen is niet altijd duidelijk. Het enige nog geldige onderscheid is dat esterases een voorkeur hebben voor korte keten acyl-esters (korter dan 10 koolstofatomen) en dat lipases een voorkeur hebben voor lange keten acyl-esters (langer dan 10 koolstofatomen) en ook actief zijn op in water onoplosbare substraten (micellen). Beide enzymgroepen delen echter structurele en functionele kenmerken, waaronder een katalytische triade, een α/β-hydrolase-vouwing, en een cofactor onafhankelijke activiteit. De katalytische triade bestaat meestal uit een nucleofiele serine in een pentapeptide motief (GxSxG), en een zuur residu (asparaginezuur of glutaminezuur) dat via een waterstofbrug gebonden is aan een histidine-residu.

Het gebruik van enzymen in de biotechnologie heeft ook nadelen. Veel reacties vinden plaats onder hoge temperaturen en/of in organische oplosmiddelen. Aangezien deze omstandigheden schadelijk zijn voor de meeste enzymen is er juist een vraag naar meer stabiele enzymen. De enzymen van hyperthermofielen zijn daarom een goede keuze omdat die vaak van nature al een thermische en chemische stabiliteit vertonen. Hyperthermofielen zijn micro-organismen die optimaal leven bij temperaturen boven de 80°C.

Dit proefschrift beschrijft de resultaten van de biochemische en structurele analyse van verschillende thermostabiele esterases, afkomstig uit een hyperthermofiele bacterie (*Thermotoga maritima*) en een hyperthermofiele archaeon (*Archaeoglobus fulgidus*). De in deze studie verkregen informatie verschaft fundamentele kennis, die kan fungeren als basis voor moderne methoden van enzym-engineering, met als doel het verbeteren van de toepasbaarheid van deze enzymen.

Hoofdstuk één van dit proefschrift geeft een overzicht van alle tot nu toe gekarakteriseerde carboxylester-hydrolases uit hyperthermofiele bacteriën en archaea. Carboxylester-hydrolases zijn belangrijke biokatalysatoren met toepassingen in de medische biotechnologie, in wasmiddelen, in organische synthese, in biodieselproductie (methylester), in smaak- en aromasynthese en in andere voedingsgerelateerde processen. Op dit moment, zijn de meeste door de industrie gebruikte esterases afkomstig uit mesofielen (T_{opt} <60°C) omdat zij als eerste zijn geïdentificeerd en gekarakteriseerd. Analyse van de genoomsequenties van hyperthermofiele bacteriën en archaea laat zien dat ook de hyperthermofiele wereld een breed repertoire aan esterases heeft, die zowel isolatie als een verdere karakterisatie verdienen.

Thermotoga maritima esterases

Het merendeel van de gekarakteriseerde esterases beschreven in dit proefschrift zijn afkomstig uit *Thermotoga maritima*. *T. maritima* is een hyperthermofiele anaërobe bacterie die geïsoleerd is uit geothermaal verwarmde mariene sedimenten bij Vulcano, Italië. Het micro-organisme groeit tussen de 55°C en 90°C (T_{opt} = 80°C) en kan een aantal eenvoudige en complexe koolhydraten metaboliseren, zoals glucose, sucrose, cellulose, xylaan en zetmeel. Door middel van een bioinformatica-analyse van het genoom van *T. maritima* werden verschillende genen, die potentieel coderen voor nieuwe esterases, geïdentificeerd, namelijk: een hypothetisch eiwit (locus tag: TM0033), een vermeend esterase (TM0053), een voorspeld acetyl xylaan esterase (TM0077), een hypothetisch eiwit (TM0336), een vermeend esterase (TM1160), en een vermeend lipase (TM1350).

Hoofdstuk twee beschrijft de identificatie, heterologe productie, zuivering en biochemische karakterisatie van een esterase (EstD, TM0336). Het enzym was oorspronkelijk geannoteerd als een hypothetisch eiwit. Echter, een meer gedetailleerde sequentieanalyse onthulde de aanwezigheid van een α/β-hydrolase-vouwing en een nucleofiele serine in een pentapeptide motief, hetgeen duidt op een mogelijke rol in esterhydrolyse. EstD heeft inderdaad esterase-activiteit met een voorkeur voor acyl-esters met een korte keten (C4-C8). Zoals verwacht van een hyperthermofiel enzym heeft EstD een optimale activiteit bij 95°C en is stabiel bij hoge temperaturen, met een half-waarde tijd van 1 uur bij 100°C. Een model van de eiwitstructuur werd geconstrueerd en gaf inzicht in het actieve centrum en mogelijke substraatbinding van dit enzym. De residuen die betrokken zijn bij de katalyse zijn door mutatie- en inhibitiestudies bevestigd. Een fylogenetische analyse van EstD toonde aan dat dit enzym slechts ver verwant is aan andere esterases en daarom hebben we een nieuwe familie voor dit enzym voorgesteld.

Het **derde hoofdstuk** beschrijft het kloneren, zuiveren, kristalliseren en de voorlopige structuuranalyse van een esterase (EstA, TM0033). EstA is ook een esterase dat in eerste instantie werd aangeduid als een hypothetisch eiwit, maar tevens typische esterase-kenmerken heeft. Een meervoudige sequentievergelijking suggereerde dat EstA is samengesteld uit twee domeinen: een C-terminaal domein met een voorspelde α/β-hydrolase-vouwing en een N-terminaal domein dat geen homologie heeft met bekende eiwitten en dus geen functie kan worden toegewezen. Zodoende werden kristallisatieonderzoeken ingezet om zowel de functie van dit nieuwe N-terminale domein te ontrafelen, als om inzicht te krijgen in de moleculaire basis van substraatherkenning en katalyse door dit enzym. Natief en selenomethionine gesubstitueerd EstA werden gekristalliseerd door middel van de "hanging-drop vapour-diffusion" (hangende-druppel dampdiffusie) methode. Kristallen groeiden optimaal in 1,0 M lithiumsulfaat monohydraat en 2% (w/v) polyethyleenglycol 8000. Een dataset werd verzameld tot 2.6 Å resolutie en een eerste analyse wordt beschreven.

Na deze eerste structuuranalyse, werden de structuur en de verschillende biochemische eigenschappen van EstA in meer detail onderzocht. Dit staat beschreven in **hoofdstuk vier**.

EstA vertoonde esterase-activiteit met een voorkeur voor esters met korte acyl-ketens (C2-C10). De structuren van zowel natief EstA, als EstA in complex met de competitieve remmer paraoxon, zijn bepaald. De structuur van EstA blijkt, zoals was voorspeld, te zijn samengesteld uit twee verschillende domeinen: een geconserveerd C-terminaal domein met het katalytische centrum en een N-terminaal domein, gelijkend op de immunoglobuline-vouwing. Uit analyse van de quaternaire structuur met behulp van massaspectrometrie en dynamic light scattering (dynamische licht verstrooiing) bleek dat EstA in oplossing voornamelijk bestaat als hexameer. Elektronenmicroscopie bevestigde dat de hexameer in oplossing identiek is aan de hexameer in het kristal. Het immunoglobuline-achtige domein van EstA lijkt een belangrijke rol te spelen in het vormen van de EstA hexameer en in de activiteit van dit enzym.

In **hoofdstuk vijf** onderzochten we de kristalstructuur en verschillende biochemische eigenschappen van een acetylesterase / cephalosporine C deacetylase (AceA, TM0077), behorende tot de carbohydraat esterase familie zeven (CE7). Enzymen die behoren tot CE7 zijn ongebruikelijk omdat zij actief zijn op zowel geacetyleerde xylo-oligosachariden als op het antibioticum cephalosporine C. AceA blijkt actief te zijn op een verscheidenheid aan geacetyleerde verbindingen, waaronder ook cephalosporine C. De activiteit op onnatuurlijk *p*-nitrofenyl-ester substraten is beperkt tot korte keten acyl-esters (C2-C3). De kristalstructuur van AceA werd opgehelderd en onthulde een klassieke α/β-hydrolase-vouwing. Inzicht in substraatbinding werd verkregen door co-kristallisatie-structuren met de remmers PMSF en paraoxon. Na binding van deze remmers veranderde de katalytische serine van conformatie. Wij speculeren dat deze verandering nodig is, zodat de waterstofbruggen met de "oxyanion hole" correct gevormd kunnen worden en zodoende, stabilisatie van de tetrahedral intermediair tijdens katalyse bewerkstelligen

In **hoofdstuk zes** staat het kloneren, zuiveren, kristalliseren, en een gedeeltelijke biochemische karakterisatie van een esterase (EstB, TM0053) beschreven. EstB vertoont esterase-activiteit met een voorkeur voor de middellange keten acyl-esters (C8-C10). EstB heeft een optimale activiteit bij ongeveer 95°C en een pH van 9. Gelfiltratie toont aan dat EstB in oplossing waarschijnlijk ook aanwezig is als een hexameer. EstB werd gekristalliseerd door middel van de "hanging-drop vapour-diffusion" methode en een dataset is verzameld tot ~2,8 Å resolutie. Er wordt momenteel gewerkt aan het oplossen van de structuur.

Archaeoglobus fulgidus lipase

Het enige echte lipase dat we ontdekt hebben is afkomstig uit *Archaeoglobus fulgidus*. *A. fulgidus* is een hyperthermofiel sulfaat reducerend archaeon die aanvankelijk is geïsoleerd nabij het eiland Vulcano, Italië. Het organisme groeit optimaal bij 83°C en kan leven op diverse koolstofbronnen zoals vetzuren, aminozuren, organische zuren, en CO. *Archaeoglobus* soorten zijn ook aangetroffen in reservoirwater van offshore olievelden, waar ze verantwoordelijk worden gehouden voor "reservoirverzuring". Het genoom van *A. fulgidus* blijkt verscheidene

genen te bevatten die potentieel coderen voor nieuwe esterases: een carboxylesterase (est-1) (locus tag: AF0865), een carboxylesterase (est-2) (AF1537), een carboxylesterase (AFEST) (AF1716), een lysofosfolipase (AF1753), een voorspeld lipase (AF1763), en een carboxylesterase (est-3) (AF2336).

De functie van het voorspelde lipase (LipA, AF1753) werd onderzocht. In **hoofdstuk 7** wordt het kloneren, zuiveren en een gedeeltelijke karakterisatie van dit lipase beschreven. LipA blijkt lipase-activiteit te vertonen met een voorkeur voor middellange keten acyl-esters (C10). Met behulp van kwalitatieve plaattesten is gebleken dat LipA de hoogste activiteit heeft met triacylglycerolesters van octanoaat en palmitaat. Hieruit is duidelijk geworden dat LipA inderdaad een lipase is. LipA heeft optimale activiteit bij ongeveer 95°C en bij een pH van 11. In verscheidene lipases is de toegang tot het actieve centrum geregeld door de positie van een "dekselstructuur", die van invloed kan zijn op de activiteit en de selectiviteit van het enzym. Of LipA ook een dekselstructuur heeft is niet bekend, maar een kristalstructuur kan inzicht geven in substraatbinding en katalyse van dit enzym. LipA is gekristalliseerd met behulp van de "hanging-drop vapour-diffusion" methode en een dataset is verzameld tot ~2,6 Å resolutie. Er wordt momenteel gewerkt aan het oplossen van de structuur. Dit is het eerste lipase dat is geïsoleerd uit een hyperthermofiel archaeon.

Nieuwe ontwikkelingen en vooruitzichten

In de afgelopen jaren zijn veel nieuwe hyperthermofiele bacteriën en archaea geïsoleerd. De genoomsequentie van verschillende van deze hyperthermofielen is bepaald en in de toekomst zal dit aantal snel toenemen als gevolg van het grote aantal lopende sequencing projecten (GOLD genomes online). Deze toename van sequentie-informatie zal de identificatie van nieuwe carboxylester-hydrolases, met nieuwe eigenschappen, versnellen. Tot nu toe wordt voor de identificatie van nieuwe enzymen veeal een traditionele screening gebruikt, echter, in de toekomst zal bioinformatica en metagenoom-screening meer en meer bijdragen aan het identificatieproces. Een groot nadeel van metagenoom-screening is dat de eiwitten die van belang zijn functioneel moeten worden geproduceerd in de heterologe host. Daarom is er onlangs een nieuw systeem ontwikkeld voor het functioneel screenen van (meta)genomische banken van extreme thermofielen. Dit systeem maakt het mogelijk om fosmid-banken in *E. coli* te maken en dan over te brengen naar de extreem thermofiele *T. thermophilus*. Dit systeem heeft bewezen een groter aantal genen functioneel tot expressie te brengen en kan in de toekomst waardevol zijn bij de identificatie van nieuwe carboxylester-hydrolases uit hyperthermofielen. In aanvulling op de identificatie van nieuwe carboxylester-hydrolases is ook hun karakterisatie onontbeerlijk.

De indeling van esterases in families is een voortdurend proces en als gevolg van de snelle toename in nieuwe esterase-gensequenties zijn veel van de huidige databases onvolledig. Een nieuwe veelbelovende aanpak is gebaseerd op superfamilies, waarbij

theoretische en experimentele gegevens worden verenigd en daardoor veel informatie kan onthullen over een eiwitfamilie. Een nieuw volledig automatisch programma dat in staat is deze superfamilie-systemen te bouwen is 3DM. Dit programma heeft de mogelijkheid om een nieuwe superfamilie van de carboxylester-hydrolases te maken, gebaseerd op structurele en sequentiële overeenkomsten. Daarnaast hebben superfamilie-systemen, gegenereerd door 3DM, ook bewezen krachtige hulpmiddelen te zijn voor het rationeel modificeren van eiwitten.

Veel nieuwe sequenties en eiwitstructuren van carboxylester-hydrolases zullen beschikbaar komen. Deze structuren kunnen een basis vormen voor moderne methoden van enzym-engineering. Afgezien van het benutten van de natuurlijke evolutie, zijn optimalisatie door computational design en laboratoriumevolutie een grote uitdaging voor het verkrijgen van op maat gemaakte enzymen voor industriële toepassingen. In het verleden, zijn deze methoden met succes gebruikt om enzymen aan te passen voor specifieke eisen, zoals een verbeterde stabiliteit, activiteit en enantioselectiviteit. In de toekomst zal de identificatie van nieuwe esterases, samen met geavanceerde engineering-technieken en high-throughput screening-systemen leiden tot betere thermostabiele esterases die in staat zullen zijn om een breed scala aan reacties te katalyseren.

Dankwoord

DANKWOORD

Mijn proefschrift is dan eindelijk af!!! Graag wil ik daarom van de gelegenheid gebruik maken om alle mensen te bedanken die, ieder op hun eigen manier, hebben bijgedragen aan het tot stand komen van dit proefschrift.

Op de eerste plaats wil ik graag mijn dagelijkse begeleider en tevens co-promotor Servé Kengen bedanken. Servé, ik heb het erg getroffen met een begeleider zoals jij. Bedankt, dat ik (na te kloppen) altijd bij je binnen kon lopen met mijn vragen en problemen. Jouw raad, optimisme en vertrouwen waren voor mij onmisbaar. Ook wil ik graag mijn promotor John van der Oost van harte bedanken. John, ik ben blij dat ik alweer enige jaren in jouw groep mag werken, eerst als student, toen AIO en nu als PostDoc. Jouw enorme enthousiasme en optimisme inspireerden mij en hadden een positief effect op mijn werk. Zowel jij als Servé hebben mij de vrijheid gegeven om mijn onderzoek grotendeels zelf in te vullen en daarvoor ben ik jullie beiden erg dankbaar. Mijn tweede promotor, Willem de Vos, wil ik ook graag bedanken. Willem, je stond wat verder af van mijn onderzoek, maar ik heb dankbaar gebruik gemaakt van je commentaar en adviezen wanneer mijn onderzoek ter sprake kwam.

Natuurlijk, nu mijn kamergenoten, paranimfen, fotograaf en bovenal goede vrienden Marco, John en Marcel. Ik wil jullie bedanken omdat jullie willen optreden tijdens mijn promotie. Mede dankzij jullie ging ik met plezier naar mijn werk. We hebben samen veel lol gehad, zowel op onze kamer als tijdens de lunch en koffiepauzes. Samen lachen om de grappen van John en Marco, het brouwen van speciale biertjes, tafeltennis na de lunch, barbecues in Wageningen en Haaksbergen, ik heb het allemaal samen met jullie meegemaakt. Marcel, het reizen door Zuid-Afrika was super en zeker voor herhaling vatbaar. John, wanneer gaan we het volgende biertje brouwen? En Marco, jij hebt het momenteel druk met schrijven en ik wens je daarom veel succes met het afronden van je eigen proefschrift.

Verder wil ik graag alle mensen van de BacGen groep bedanken voor alle hulp en de gezellige tijd. Ans, en later Sjon, bedankt voor jullie ondersteuning op het lab. Stan, jij bent van alle markten thuis, bedankt voor je raad en adviezen. Harmen, als bioinformatica-goeroe, heb ik je regelmatig moeten raadplegen. Bart, Mark M. en Katrin, voor de "social talk" tijdens lunch en koffiepauzes. Suzanne en Bram, voor de gezelligheid op de werkkamer. Alle andere Bacgenners: Jaapie, Jasper W., Johan, Ronnie, Krisztina, Hao, Colin, Marke, Fabian, Lei, Karin, Matthijs, Pawel, Magnus, Edze, Pierpaolo, Ratnesh, Teunie, Rie, Nicolas, Vincent en Ingrid, allemaal bedankt!!!

Voor de dagelijkse gang van zaken op de vakgroep ben ik dank verschuldigd aan Wim, Nees, Gosse, Jannie, Renée, Francis en Anja. Verder wil ik ook iedereen in één van de andere

groepen binnen Microbiologie (MicPhys, MolEco en FunGen) bedanken voor de goede sfeer op het lab, tijdens de koffie en lunchpauzes, labuitjes, etc.

Tijdens mijn onderzoek heb ik samengewerkt met een groot aantal mensen van zowel binnen als buiten Wageningen.

Ik wil graag Geb Visser en Maurice Franssen van Organische Chemie bedanken voor hun hulp, advies en samenwerking. Ondanks meerdere pogingen voor biokatalyse onder magnetronstraling, heeft dit helaas niet geresulteerd in de beoogde verschillen in enzym kinetiek.

Marieke Bruins en Anja Janssen van Proceskunde wil ik graag bedanken voor het nuttige overleg betreffende het gebruik van hoge druk. Het gebruik van hoge druk voor de productie van esters bleek een lastige opgave.

Anne-Meint Henstra, Diana Sousa en Melike Balk van Microbiologie wil ik graag bedanken voor hun hulp met het kweken van *Thermotoga maritima* en *Archaeoglobus fulgidus*.

I want to thank Alessandra Pasquo and Valerio Consalvi from the University of Rome "La Sapienza" for their crystallization experiments on the EstD enzyme.

I would like to thank Lei Sun, Bauke Dijkstra and Jelle Bultema from Groningen University. Lei, it was really great working together with you on the crystallization of the esterase EstA. Bauke, bedankt voor al je hulp en advies. Jelle, bedankt dat je EstA onder de elektronenmicroscoop wilde bekijken. Dit gaf echt de doorslag van hoe de hexameer eruit ziet.

Arjan Barendregt en Robert van den Heuvel van Universiteit Utrecht wil ik graag bedanken voor de natieve massaspectrometrie metingen.

I want to thank Gye Won Han, Marc Deller, Scott Lesley and Ian Wilson from the Joint Center of Structural Genomics. Dear Gye, I have been very lucky that we met each other at the Thermotoga 2007 Meeting in San Diego. If we had not exchanged email-addresses, chapter 5 of this thesis, and also our collaboration with the JCSG, would not have existed. I really enjoyed our collaboration on AceA and I hope we can publish our results in a good journal. I also have to thank Scott Lesley for extending our collaboration to other targets than AceA. And Marc, I really appreciate your efforts to crystallize our target proteins. Thanx a lot!!!

I would like to thank Peter Biely from the Institute of Chemistry in Bratislava for his activity measurements with AceA. Although we never met, you have been a great help from the moment I contacted you.

Elise Meulenbroeks en Hans Kierkels van DSM Geleen wil ik graag bedanken voor het testen van onze esterases en lipases voor verschillende toepassingen.

Dit boekje zou een stuk minder gevuld zijn als ik gedurende mijn AIO-schap niet de hulp had gehad van een aantal enthousiaste studenten. Ik ben mijn studenten dan ook erg dankbaar voor hun inzet: Jojo, Sjon, Paul, Zou, Twan, en Peter van harte bedankt!! Sjon, jij hebt eerst als student en later als analist in onze groep hard gewerkt aan mijn project en ik ben dan ook blij en trots dat jouw naam meerdere van mijn hoofdstukken en publicaties siert.

Verder wil ik graag de mede organisatoren van de studiereis naar California bedanken. Marco, Bart, Miguel, Arno, Rob en Filipe, bedankt voor het helpen organiseren van een geweldige reis. Ook wil ik graag alle deelnemers en bezochte groepen in California bedanken, want mede dankzij jullie was deze reis fantastisch!!!

Tevens wil ik mijn vrienden in Oss en in Wageningen bedanken voor de welkome afleiding van het onderzoek. Bedankt voor alle etentjes, kaartavonden, spellenavonden, het stappen in Oss en Wageningen, etc. Hierbij wil ik speciaal Ramon bedanken. Gracias por todo, amigo! Christian, het is altijd gezellig in de International Club met jou. En Nicole, jij gaf het goede voorbeeld al een paar maanden geleden door doctor te worden. Bedankt voor je hulp en de gezellige gesprekken over ons werk.

Ik wil graag mijn familie bedanken voor hun interesse in mijn werk. In het bijzonder mijn neef Pieter. Jouw interesse en luisterend oor tijdens onze etentjes wordt door mij zeer gewaardeerd.

Tot slot, mijn "thuisfront". Pa, ma, Karen & Marc, ik neem aan dat jullie nog steeds niet goed begrijpen waar ik eigenlijk de afgelopen vijf jaar mee bezig ben geweest. Toch hebben jullie alles gevolgd en steeds geïnteresseerd geluisterd naar wat ik te vertellen had. Ik wil jullie graag danken omdat jullie er steeds voor me zijn en omdat ik altijd op jullie kan rekenen!

Mark
September 2009

CURRICULUM VITAE

Mark Levisson werd op 4 november 1980 geboren te Oss. In juli 1999 behaalde hij het VWO diploma aan het Mondriaan College te Oss. In september van datzelfde jaar begon hij met de opleiding Biologie en Medisch Laboratorium Onderzoek aan de Hogeschool van Arnhem en Nijmegen. Tijdens deze studie heeft hij 10 maanden stage gelopen bij de sectie Moleculaire Genetica van Industriële Micro-organismen aan de Wageningen Universiteit. Onder begeleiding van Marco J.L. de Groot en dr. George J.G. de Ruijter werd onderzoek verricht naar de pentose katabole routes van de schimmel *Aspergillus niger*. In juli 2002 behaalde hij zijn diploma's in de afstudeerrichtingen Biochemie en Biotechnologie.

Vervolgens besloot hij om verder te studeren aan de Wageningen Universiteit en begon in september 2002 aan de studie Biotechnologie. Tijdens deze studie specialiseerde hij zich in de richting Cellulaire en Moleculaire Biotechnologie. Gedurende zijn studie deed hij een eerste afstudeeronderzoek bij de vakgroep Virologie. Hier werd onder begeleiding van Etienne Bucher en prof. dr. Rob Goldbach onderzoek verricht naar de suppressie van RNA silencing door plantenvirussen. Vervolgens ging hij voor een tweede afstudeeronderzoek naar de vakgroep Microbiologie waar hij, onder begeleiding van Suzanne van Loo en dr. John van der Oost, onderzoek verrichtte naar de biochemische eigenschappen van thermostabiele aldolasen.

In september 2004 behaalde hij het Master diploma Biotechnologie, waarna hij in oktober begon aan zijn promotieonderzoek bij het Laboratorium voor Microbiologie aan de Wageningen Universiteit. Onderzoek werd verricht naar de biochemische en structurele eigenschappen van hyperthermofiele esterases. De resultaten van dit onderzoek staan beschreven in dit proefschrift.

Sinds januari 2009 werkt hij als postdoctoraal onderzoeker voor het Dutch Separation Technology Institute (DSTI) te Amersfoort, maar is gedetacheerd bij het Laboratorium voor Microbiologie aan de Wageningen Universiteit. Hier houdt hij zich bezig met het ontwikkelen van nieuwe affiniteits liganden voor suikers, peptiden en eiwitten.

PUBLICATION LIST

de Groot, M.J.L., van den Dool, C., Wosten, H.A.B., **Levisson, M.**, vanKuyk, P.A., Ruijter, G.J.G. & de Vries, R.P. (2007) Regulation of pentose catabolic pathway genes of *Aspergillus niger*. Food Technology and Biotechnology 45 (2), p. 134 - 138.

Levisson, M., van der Oost, J. & Kengen, S.W.M. (2007) Characterization and structural modeling of a new type of thermostable esterase from *Thermotoga maritima*. FEBS Journal 274 (11), p. 2832 - 2842.

Sun, L., **Levisson, M.**, Hendriks, S., Akveld, T., Kengen, S.W.M., Dijkstra, B.W. & van der Oost, J. (2007) Crystallization and preliminary crystallographic analysis of an esterese with a novel domein from the hyperthermophile *Thermotoga maritima*. Acta Crystallographica Section F. Structural Biology and Crystallization Communications 63 (9), p. 777 - 779.

Wolterink-van Loo, S., **Levisson, M.**, Cabrières, M.C., Franssen, M.C.R. & van der Oost, J. (2008) Characterization of a thermostable dihydrodipicolinate synthase from *Thermoanaerobacter tengcongensis*. Extremophiles 12 (3), p. 461 - 469.

Levisson, M., Sun, L., Hendriks, S., Swinkels, P., Akveld, T., Bultema, J.B., Barendregt, A., van den Heuvel, R.H.H., Dijkstra, B.W., van der Oost, J. & Kengen, S.W.M. (2009) Crystal structure and biochemical properties of a novel thermostable esterase containing an immunoglobulin-like domain. Journal of Molecular Biology 385 (3), p. 949 - 962.

Levisson, M., van der Oost, J. & Kengen, S.W.M. (2009) Carbohydrate ester hydrolases from hyperthermophiles. Extremophiles 13 (4), p. 567 - 581.

Levisson, M., Han, G.W., Deller, M.C., Biely, P., Hendriks, S., Xu, Q., Ten Eyck, L., Flensburg, C., Roversi, P., Miller, M., McMullan, D., von Delft, F., Kreusch, A., Elsliger, M., Deacon, A., Lesley, S., van der Oost, J., Wilson, I. & Kengen S.W.M. The crystal structures of a thermostable acetyl esterase / cephalosporin C deacetylase from *Thermotoga maritima* in complex with PMSF and paraoxon reveal an altered conformation of the catalytic serine. In preparation.

Levisson, M., Gielen, P., Deller, M.C., Kengen, S.W.M. & van der Oost, J. Purification and partial characterization of a thermostable esterase from *Thermotoga maritima*. In preparation.

Levisson, M., Hendriks, S., Siemerink, M.A.J., Deller, M.C., Kengen, S.W.M. & van der Oost, J. Characterization of a lipase from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. In preparation.

OVERVIEW OF COMPLETED TRAINING ACTIVITIES

Graduate School VLAG	Year

Discipline specific activities
Courses

Organic Chemistry, Wageningen University, Wageningen	2004
Advanced Course Food Enzymology, VLAG, Wageningen	2005
Advanced Course Bio-nanotechnology, VLAG, Wageningen	2005
Advanced Course on Biocatalysis, BODL, Delft	2005

Meetings

ALW-platform Molecular Genetics, Lunteren	2004
CW / NWO Protein Research, Lunteren (poster presentation)	2004
ALW-platform Molecular Genetics, Lunteren	2005
CW / NWO Protein Research (poster presentation)	2005
Biocat 2006, Hamburg, Germany (poster presentation)	2006
Biocatalysis in the Food and Drinks Industries, Wageningen (poster presentation)	2006
CW / NWO Protein Research, Lunteren (oral presentation)	2006
Thermotoga Workshop, San Diego, USA (invited speaker)	2007
SIM Annual Meeting, Denver, USA (invited speaker)	2007
Nederlands Biotechnologie Congres, Ede (oral presentation)	2008
Extremophiles 2008, Cape Town, South-Africa (oral presentation)	2008

General courses

VLAG PhD Introduction Week	2004

Optionals

Preparation PhD research proposal	
Group Meetings, Wageningen,	2004-8
PhD/Postdoc Meetings, Wageningen, 2004-2008	2004-8
Project Meetings, Wageningen, 2004-2008	2004-8
Journal Club	2008
Organizing PhD Study Trip to California, USA	2005-6
PhD Study Trip to California, USA	2006