The Structure and Function of Esterases from Lactic Acid Bacteria

A thesis presented in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in the

Institute of Molecular BioSciences
Massey University
New Zealand

Matthew David Bennett
2007

Abstract

Compounds derived from the breakdown of glyceride esters of milk fat, such as free fatty acids and short chain esters, are recognised as playing an important role in the flavour of a range of fermented foods. Esterases, capable of hydrolysing ester bonds, and in some cases, synthesising them *via* an acyltransferase mechanism, typically enter the fermentation from the starter and adjunct lactic acid bacteria that are used to inoculate milk to initiate the fermentation process. With such an important role in the development of both desirable and undesirable flavours, understanding how these enzymes operate is essential for product control. In this study, the crystal structures of three lactic acid bacterial esterases were solved: EstA from *Lactococcus lactis*, and AA7 from *Lactobacillus rhamnosus* which are both capable of hydrolysis of short chain triglycerides as well as synthesising esters *via* a transferase mechanism, and AZ4, an esterase from *L. rhamnosus* which appears to be limited to hydrolysis reactions. Whilst all three were found to be members of the αβ hydrolase family, unique features were found for each enzyme, reflecting the large differences in their primary sequences, substrate specificities and activities.

EstA and AA7 were both found to have a shallow substrate binding cleft, bisected by the catalytic machinery. The divided binding cleft suggests that during a transferase reaction the transferred group binds in one pocket, with the donor and acceptor groups (dependant on the stage of catalysis) binding in the other.

In contrast, AZ4 was found to have a single deep substrate binding cavity, extending into the enzyme interior, with the catalytic residues located near its entrance. The absence of a second binding site for an acceptor is consistent with AZ4 having only one function – that of a hydrolase.

The structures presented in this study are the first three dimensional structures of esterases from lactic acid bacteria to be reported. Their analyses, both in native form, and complexed with a varity of ligands mimicking various stages of the reaction cycle have highlighted how this basic fold can be adapted to efficiently catalyse different reactions. More importantly, in the case of AZ4, these structures have suggested that there is a novel mechanism used by the esterases to promote the enzyme reaction to proceed to completion, by preventing a futile catalytic reaction.

Enzymes are things invented by biologists that explain things which otherwise require harder thinking.

Jerome Lettvin

Acknowledgements

Many people have supported me during my time working on this project. In no particular order, I would like to thank the following people for their support and assistance:

All the members of the X-Lab past and present, especially Celia, Jana, Judith, Fiona, Kun, Stuart, Reuben, Santanu, Rose and Kirsty, and the guys next door, Andrew, Alice, Carlene, Greg, Muthu and Anandan, for their support and cheer during the past four years. Damon who came on board this project in the closing stages and did some really good work on AA7 despite being somewhat thrown in the deep end.

Trevor Loo, biochemistry guru, for his help with machines, recipes and general good advice. The "X-ray and NMR boys" Simon, Joe and Martin, and Ursula for the constant natter in the computer lab, and many, many coffees/lunches/beers. Ann, Cynthia and Pat, the best secretaries in the world.

Rosauro Roldan and Emily Chen from Fonterra Palmerston North for assistance with instruments and methods.

Professor Ted Baker, and other members of the University of Auckland Maurice Wilkins Centre for Structural Biology, and Clyde Smith from Stanford Synchrotron Radiation Laboratory for arranging for the collection of MAD datasets.

Michael Hardman and Emily Parker for useful discussions about the ins and outs of enzyme catalysis. Mark Patchett for dicussions about molecular biology and catalysis. Bryan Anderson for his work on the all the enzymes, and for his assistance with software, space groups and unit cells and for showing me the ropes when it came to model building. Geoff Jameson for discussions about crystallography and enzyme mechanism in general, and for giving me the opportunity to learn about X-ray machine maintenance (and for then trusting me to do it!). Carol Taylor, David Harding, and especially David Lun, who put in a lot of hard work to kindly provide us with the inhibitors necessary to take the AA7 work to the next level.

Marie-Laure Delabre for her work on identifying, cloning and characterising the *Lb. rhamnosus* enzymes that have been integral to this project, and for her continued interest in this project.

My co-supervisor Ross Holland, who gave me the opportunity to start esterase research as a summer student in 2000, and showed faith that we would finally get some structures by going to great lengths to find the continued funding that has allowed our collaboration to finally yield some fruit.

My supervisor Gill Norris, whom I cannot thank enough for her support and assistance throughout this project, for being approachable at all times about anything, and for providing "tough love" when it was needed during the times of low motivation.

My parents for their love and support over the years, especially my mother for not balking (at least not visibly) at the idea of her son spending 10 years at university!

Finally Susan, for her love and support, and for understanding the late nights, and grumpiness when things were not working for me.

Matthew Bennett 24 October 2007

Table of Contents

List of Figures

Chapter 3

Chapter 4

Chapter 5

List of Tables

Abbreviations

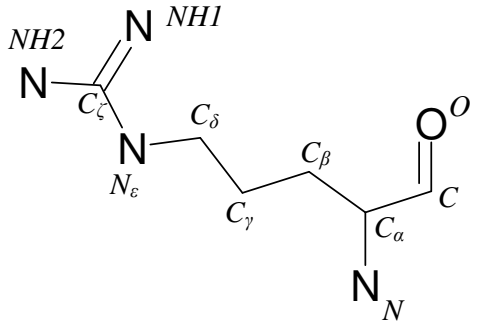
Amp	Ampicillin
BLAST	Basic Local Alignment Search Tool
ESRF	European Synchrotron Radiation Facility
FFA	Free Fatty Acids
GC	Gas Chromatography
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIC	Hydrophobic Interaction Chromatography
IEX	Ion Exchange Chromatography
IMAC	Immobilised Metal Affinity Chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
kDa	Kilo Daltons
LAB	Lactic Acid Bacteria
LB	Luria-Berteli Broth / Luria Broth / Lysogeny Broth
MAD	Multiple Anomalous Dispersion
MCS	Multiple Cloning Site / polylinker
MR	Molecular Replacement
MWt	Molecular Weight
NaAc	Sodium acetate
NCBI	National Centre for Biotechnology Information
nuc	Nucleophile
OD	Optical Density
ORF	Open Reading Frame
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
PEG	Polyethylene glycol
SAD	Single Anomalous Dispersion
SEC	Size Exclusion Chromatography
SDS-PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
Sp.	Species
SSRL	Stanford Synchrotron Radiation Laboratory
Subsp.	Subspecies
TRIS	Tris (hydroxymethyl) Aminomethane
WT	Wild Type

Amino Acid Abbreviations

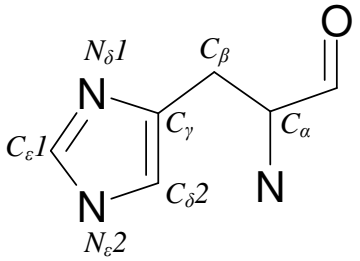
Amino acid	One letter code	Three letter code
Alanine	A	Ala
Arginine	R	Arg
Asparagine	N	Asn
Aspartate	D	Asp
Cysteine	C	Cys
Glutamine	Q	Gln
Glutamate	E	Glu
Glycine	G	Gly
Histidine	H	His
Isoleucine	I	Ile
Leucine	L	Leu
Lysine	K	Lys
Methionine	M	Met
Phenylalanine	F	Phe
Proline	P	Pro
Serine	S	Ser
Threonine	T	Thr
Tryptophan	W	Trp
Tyrosine	Y	Tyr
Valine	V	Val

Atom Numbering

This thesis uses the atom numbering system used by the Protein Data Bank (PDB). Greek letters are used to denote each atom outwards along the side chain from the C_α, with numbers used to differentiate between atoms where branching occurs. Examples are given below. It should be noted that in .pdb files, Roman letters are substituted for their Greek equivalents.



Arginine



Histidine

Nucleic Acid Abbreviations

One Letter Code	Base(s) Represented
A	Adenosine
C	Cytosine
G	Guanine
T	Thymidine
U	Uridine / Uracil
R	G or A
Y	T or C
K	G or T
M	A or C
S	G or C
W	A or T
B	G or T or C
D	G or A or T
H	A or C or T
V	G or C or A
N	Any

Chapter 1

Introduction and Literature Review

Chapter 1

Introduction and Literature Review

1.1 Introduction

Food preservation by fermentation is an art that has been practised for centuries. Legend has it that this method of storage began when the process of fermenting milk into cheese was discovered accidentally, when milk was kept in bags made from animal stomachs. These early cheese makers observed that the milk became a sludge, but was preserved from decay by a process that we now know as fermentation.
Since this time, the manufacture of fermented food products has come from being a little-understood art form to a highly controlled industrial process. The factors that affect cheese production have been studied closely, and a reasonable understanding of the biochemical changes that occur when a raw feedstock is fermented and aged into a final product has evolved.
The study of the enzymes involved in fermentation and maturation is an integral part of understanding this process. Although purification and characterisation of the enzymes representing some of the many different activities involved in the fermentation and aging of dairy products has been carried out (Section 1.2.3), this study, to our knowledge, is the first attempt to obtain structural information about the enzymes involved in the production of ester-based flavour compounds in a food product.
This review briefly covers what is known about the biochemistry of cheese and dairy product manufacture, but mainly focuses on esterases and other αβ hydrolases, and their mechanisms of action.

1.2 Cheese Manufacture

"Age is of no importance unless you are a cheese."

– *Billie Burke*

Cheese manufacture begins when "starter" bacteria are added to milk (either pasteurised in modern manufacturing plants or unpasteurised in more traditional methods), along with rennet, which contains a protease that begins milk protein digestion. Sometimes, controlled amounts of "adjunct" bacteria are also added, to produce altered flavours in the product. Starter bacteria and some adjunct species are collectively referred to as "dairy lactic acid bacteria" for their role in producing dairy products. The bacteria produce lactic acid as a by-product of the glycolytic cycle, and metabolise other milk compounds into new compounds. Additionally, most dairy products contain other species of lactic acid bacteria that are not starters or adjuncts, but inadvertently end up in the fermentation; they are referred to as non-starter lactic acid bacteria (NSLAB). The identity of all the species of NSLAB in any given fermentation is often unknown, and the NSLAB can either contribute to the product in desirable ways or cause defects in the product that may render it unfit for sale.

1.2.1 Fermentation

The lactic acid produced by the lactic acid bacteria causes the pH of the milk to decrease, leading to the precipitation of milk proteins and the expulsion of water from the resulting solid curd. Ultimately, the bacterial cells lyse in the cheese environment (low pH, low water activity and typically high salt), releasing intracellular enzymes directly into the immature curds, resulting in an increase in enzymatic activities such as lipolysis (Collins 2003) and peptidolysis (Law, Sharpe *et al.* 1974) in the product as cheese maturation progresses. After the initial fermentation stage, cheese is stored for a period of time to mature. During this period, the enzymes released from the

lysed starter bacteria slowly process milk compounds into the compounds that are responsible for the final flavour in the product. Although few intact starter bacteria remain in a fully matured cheese product, the contribution made by released enzymes to the final product is very important.

1.2.2 Dairy Lactic Acid Bacteria

Dairy lactic acid bacteria belong to a group of Gram-positive species that cover a wide range of genera including *Lactococcus*, *Streptococcus* and *Lactobacillus*. They all produce lactic acid as a by-product of their metabolism, and are generally harmless to humans (although it should be noted that most *Streptococcus* species are capable of causing human infections).

However, the species differ in their morphology and, more importantly, in their metabolism, with different species containing ranges of metabolic enzymes that are specifically tuned to their own metabolic needs, and that are therefore specific for a range of different substrates. Because of these variable activities and specificities, a spectrum of flavour compounds is produced that varies depending on the bacterial species present. It thus follows that different bacterial species can be used to manipulate the flavours of the final product. Although, in the past, this has largely been an empirical process, more recently, many of the enzymes thought to be involved in flavour development have been identified and studied in isolation to gain an understanding of their properties.

1.2.3 Lactic Acid Bacteria Catabolism and Flavour

Many cheese flavour compounds are produced by the activity of enzymes from lactic acid bacteria on compounds in the milk. Three major catabolic processes are important in cheese flavour: glycolysis, lipolysis and proteolysis.

The glycolytic process is an integral part of lactic acid bacterial metabolism, and results in the formation of lactic acid. Lactic acid is responsible for lowering the pH of the milk, coagulating it into a solid curd and inhibiting the growth of any bacteria present in the fermentation, whether they be starters or adventitious species. As glycolysis can occur only in intact cells, it occurs mainly in the initial fermentation stage of cheese manufacture. Lipolyis is the degradation of milk glycerides into fatty

acid compounds, many of which have a strong flavour. Proteolysis is the breakdown of milk proteins into peptides and amino acids, both of which can also be flavour components.

Although the production of flavours in products such as cheese involves the interaction of hundreds of milk and microbially derived compounds with tens, if not hundreds, of enzymes, the study of specific enzymes that make a significant addition to flavour will result in a better understanding of their contribution to flavour production. It has been noted that even enzymes that are weakly active in producing flavour compounds might have a significant impact on flavours over long maturation periods (Reiter, Fryer *et al.* 1966).

1.2.4 Control of Cheese Flavour

Although it is unlikely that it will ever be possible to investigate and understand the interactions of every single enzyme and compound present, many of the enzymes that make a significant contribution to desirable flavour properties have been studied. The outcome of such studies has resulted in better control of flavour development in cheese and other dairy products. One potential advantage is the acceleration of ripening, which, in hard cheeses, such as Cheddar, can take many months.

Some investigations into ways of accelerating the ripening process have been undertaken, including the simple addition of exogenous enzymes to alter peptide and amino acid levels (Law and Wigmore 1983). However, cheese maturation is a complex process and, with so many compounds and enzymes responsible for contributing to the final product, such a simple approach is unlikely to successfully produce a product that is indistinguishable from a traditionally matured product.

1.3 Esters and Esterases

1.3.1 Ester Compounds

A group of flavour compounds that has been studied in recent times is the esters. Lipids, which contain ester bonds, also contribute to flavour, and are differentiated from simple esters by their limited solubility in aqueous solution. Esters themselves are potent flavour compounds, which, when hydrolysed, liberate free fatty acids, which are also strong flavouring agents.

In dairy lactic acid bacteria, ester bond hydrolysis appears to be species dependent. For example, some subspecies of *Lactococcus lactis* have no lipase activity (Katz, Medina *et al.* 2002), but have enzymes that can hydrolyse smaller water soluble glycerides (Holland and Coolbear 1996).

These specificities are of interest to the dairy industry as they can be used to control the quantity and composition of either desirable or undesirable ester-derived flavour compounds in the final product. It is therefore logical that the relative amount and concentration of each bacterial species present in a fermentation can influence the balance of flavour compounds.

1.3.2 Glycerides

Glycerides, such as those found in milkfat, comprise up to three fatty acyl chains linked through an ester bond to a glyceride backbone. Glycerides are normally prefixed with tri, di or mono, depending on how many acyl substituents they contain, and also by the number of carbons present on the acyl chain. For simple glycerides with identical groups in all positions, common names such as tributyrin (four-carbon chains in all three positions) are used. To designate more complicated glycerides, particularly where different sized chains are present at each position, chemical names are used. Glycerides also have a stereocentre at the secondary position of the glyceride backbone, which can be used to distinguish the otherwise chemically identical primary positions of the backbone, forming the basis of the IUPAC naming system for these compounds (discussed in Section 1.4).

An additional complication is the usage of traditional common names for different fatty acids, often named after the source from which they were originally isolated. These common names are tabulated for reference in Appendix B.

1.3.3 Esterases and Lipases

In chemical terms, lipases and esterases catalyse exactly the same chemical reaction, the hydrolysis of an ester bond, but, whereas esterases catalyse the hydrolysis of a substrate in solution, lipases are most active on a substrate that is present in excess of its critical micelle concentration. Therefore, the general distinction between these enzymes is made on the basis of this difference in activity (Tsujita, Shirai *et al.* 1990). Most lipases and esterases can also be distinguished by their ability to exhibit interfacial activation, a phenomenon observed mainly in lipases. Interfacial activation allows the lipase to act upon a well-dispersed insoluble substrate (an emulsion). The enzyme has a hydrophobic lid that normally covers the binding site of the enzyme, but it opens in solutions that contain micelles, allowing the substrate to access the binding site. This means that lipase activity actually increases when the critical micelle concentration is exceeded, unlike esterase activity, which will reach a maximum before the substrate concentration reaches saturation. These differences in activity are easily observed in kinetic studies, as esterases exhibit the Michaelis-Menten kinetics of a fully soluble substrate, whereas lipases show a concentration-dependent sigmoidal curve (Sarda and Desnuelle 1958; Entressangles and Desnuelle 1968). Interfacial activation can often be used to distinguish between esterases and lipases, but it has been noted that there are exceptions to this generalisation (Tsujita, Shirai *et al.* 1990). An example is rat hepatic triglyceride lipase, which does not show any increase in activity when provided with a methyl butyrate substrate at a concentration exceeding solubility (Tsujita, Nakagawa *et al.* 1984).

1.3.4 Tributyrin Esterases

Tributyrin esterases are the most intensively studied esterases in the dairy industry as they are implicated in developing cheese flavours through the hydrolysis of butyl esters to produce potent flavour compounds such as butyric acid. This four-carbon carboxylic acid has a characteristic odour that is described variously as “cheesy” or

"rancid", depending on its concentration in a product, and its existence in cheese has been known for almost a century (Suzuki, Hastings *et al.* 1910). This changeability in the perception of dairy product odour and flavour depending on concentration makes it important to control the levels of fatty acids in dairy products. Although the synthetic glyceride tributyrin, used as a substrate for esterases because of its availability, is not found in milk, 3.9% of the acyl groups found in milkfat (by weight) are butyl groups (Parodi 1982) and these can potentially be liberated from the glyceride by enzymes.

In Cheddar cheese, butyric acid can represent 27% of the free fatty acids present after maturation (Forss and Patton 1966), indicating the importance of this class of esterase. It is also important that not too much butyric acid is present in a product, as this will lead to a rancid odour and flavour, similar to that found in aged butter; therefore, control of the production of this compound in a product, rather than its maximisation, is the goal.

1.3.5 Esterases as Synthetic Enzymes

A number of esterase-like enzymes have been shown to have transferase activities, in that they have been shown to be capable of catalysing the removal of a functional group from one molecule and ligating it to another molecule, in an aqueous solution. The cell-wall-synthesising enzymes of *Mycobacterium tuberculosis*, which have been well studied (Jackson, Raynaud *et al.* 1999), are examples of such enzymes, but a large number of serine-carboxypeptidase-like (SCPL) acyltransferases are also found in the secondary metabolic pathways of plants (Milkowski and Strack 2004).

Some of the tributyrin esterases of dairy lactic acid species have been shown to be able to synthesise esters from an acylglyceride substrate, in the presence of an alcohol acceptor (Liu, Baker *et al.* 2004). Such esters are of interest because they can produce either desirable or undesirable flavours in dairy products. These compounds generally have an odour that is described as "fruity", and ethyl butanoate, which can be produced from ethanol and tributyrin, has a characteristic pineapple-like odour. Ethyl butanoate is an extremely potent odorant, and can be detected by untrained human subjects at a concentration of only 0.015 ppm (Siek, Albin *et al.* 1969). In some types of cheese, such as Cheddar, the presence of ethyl butyrate is regarded as an undesirable flavour defect (Bills, Morgan *et al.* 1965). Different lactic acid bacteria

starter species can produce different concentrations of esters, resulting in distinct flavours in products. For example, *Lactococcus lactis* subsp. *lactis* is known to produce higher amounts of ethyl butanoate in cheese than *Lactococcus lactis* subsp. *cremoris* (Perry 1961), although both these subspecies carry very similar genes for tributyrin esterase (Fernandez, Beerthuyzen *et al.* 2000).

The phenomenon of ester synthesis by esterases has been investigated both *in vitro* (Liu, Baker *et al.* 2004) and *in situ* in a simulated cheese mix (Holland, Liu *et al.* 2002) for the tributyrin esterase EstA, which was investigated structurally in this current project.

1.3.6 Esterase Identification

A number of tributyrin esterases have been identified in lactic acid bacteria with purifications of enzymes reported from species such as *Lactococcus lactis* (Holland and Coolbear 1996), *Lactobacillus plantarum* (Gobbetti, Fox *et al.* 1997), *Lactobacillus fermentum* (Gobbetti, Smacchi *et al.* 1997) and *Streptococcus thermophilus* (Liu, Holland *et al.* 2001).

With the advent of large scale, high throughput genome sequencing, the possibility of identifying bacterial enzymes using a bioinformatic approach has allowed the recognition of potential candidate esterases on the basis of their sequence similarity to other known esterases (Ro, Hong *et al.* 2004), using sequence searching tools such as BLAST (Altschul, Madden *et al.* 1997).

1.3.6.1 Fungal Esterases

Esterase activities have been identified in some fungi, such as the cheese surface flora *Debaryomyces hansenii* (Besancon, Ratomahenina *et al.* 1995). Fungi are also known to produce lipases, some of which are utilised in the manufacture of highly ripened cheeses such as Roquefort. There are also fungal lipases that actually demonstrate esterase-like Michaelis-Menten kinetics when assayed in solution (Nini, Sarda *et al.* 2001). It should also be noted that, although the kinetic distinction between esterases and lipases is typically made on a substrate dissolved in an aqueous solution, dairy products are typically low in water, which could affect factors such as lipase

activation and might mean that an enzyme is permanently activated in the high lipid environment of a dairy product.

1.3.6.2 Microesterases

It is interesting to note that esterases have been isolated from the fungal species *Emericella nidulans* and *Talaromyces emersonii*, with the unusually small molecular masses of 4.1 and 1.6 kDa respectively (Fan and Mattey 1999). A 5-kDa, 56-residue esterase with a preference for butyl glycerides has also been isolated from *Candida lipolytica* (Mattey, Simoes *et al.* 1998). Mattey, Simoes *et al.* (1998) also found an esterase that is a mere 15 residues in length in *Bacillus stearothermophilus*. No sequence or structural information for these microenzymes is available, and, although there is some indication that the *Candida* enzyme uses a metal-ion-dependent mechanism, the ability of such a small peptide to catalyse a reaction is of great interest. The design of a synthetic microesterase that utilises a nucleophilic attack mechanism has helped to determine how this might happen (Nicoll and Allemann 2004). Although this synthetic peptide forms a defined helix-turn structure in solution, reminiscent of the nucleophile elbow of αβ hydrolases, all the reactive residues are found along one face of the helix, rather than as part of the turn, as in the globular esterase enzymes.

1.3.7 The Physiological Role of Lactic Acid Bacteria Esterases

The physiological role of most esterases in lactic acid bacteria has not yet been identified. Lactic acid bacteria in general have a low lipolytic activity (Meyers, Cuppert *et al.* 1996), and therefore must derive most of their energy from the catabolism of other compounds.

There is a possibility that some esterases may play a role in the structural maintenance of the cell wall. This hypothesis is based on two facts. Firstly, some *Lactococcus lactis* tributyrin esterase activity is found to be cell wall associated (Holland and Coolbear 1996), as is the secondary esterase PepX (Tan, Chapot-Chartier *et al.* 1992). Secondly, Antigen 85C, a transferase enzyme from *Mycobacterium tuberculosis*, which has a 30% protein sequence similarity to *Lactococccus* tributyrin esterase, has a well-documented role in cell wall synthesis (Jackson, Raynaud *et al.* 1999). It has also

been proposed that lactic acid bacteria esterases might have a role in detoxifying compounds, as they have some sequence similiarity to S-formylglutathione hydrolase, which is found in the formaldehyde detoxification pathway (Fernandez, Beerthuyzen *et al.* 2000). It is interesting that *Streptococcus mutans* contains a glucan-binding protein with lipase activity, which has been found to be able to hydrolyse free fatty acids from the lipoteichoic acid in the cell walls of another bacterial species (Shah and Russell 2004). Although this enzyme has a domain that is similar in structure to other esterase and lipase enzymes, it is not significantly related to *Lactococcus* EstA; however, it may nonetheless provide some insight into the possible roles of esterases in lactic acid bacteria species.

1.4 Substrate Stereochemistry

"Perhaps looking glass milk isn't good to drink."

– *Alice in "Through the Looking-Glass"*
(Lewis Carroll, 1871)

1.4.1 Glyceride Chirality

All glyceride-based molecules have a potential stereocentre, at position 2 of the glyceride backbone, that can have four chemically different groups attached. Whether the molecule is chiral is dependent on the substituents on each position. If the primary positions of a glyceride have identical chains present, the molecule is prochiral and does not exhibit stereochemistry, because the stereocentre does not have four different groups attached to it. However, a change at one of the primary positions will meet these conditions, producing a stereocentre and rendering the molecule chiral.

In acylglycerides, the acyl substituents are stereochemically distinguishable. A system for **s**tereospecifically numbering (*sn*) substituents *sn*-1, *sn*-2 and *sn*-3 was established (Hirshmann 1960), and adopted as the official IUPAC-IUB convention for defining triglycerides (1967).

This system is constructed by using the Fischer projection, which is a flat representation in which horizontal bonds project towards the viewer and vertical bonds project away from the viewer. In this representation, the glyceride backbone is drawn vertically, with the primary chains running to the right and the secondary chain running to the left. The chains can then be numbered by position from top to bottom, as shown in Figure 1.1.

Milkfat is an example of a chiral triacylglyceride, as it contains an acyl chain of different length on each of the three glyceride positions (Parodi 1982). Most other natural triglycerides also exhibit a non-random stereospecific distribution of acyl groups.

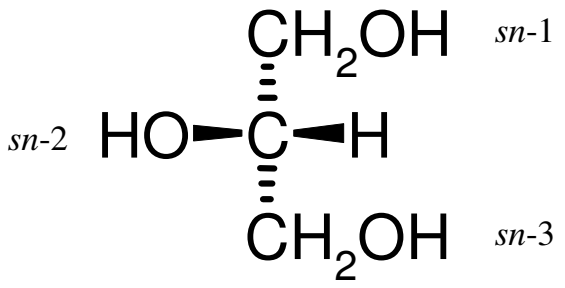


Figure 1.1 *Stereospecifically numbered glycerol, showing the stereochemical basis for this numbering using a Fischer projection.*

1.4.2 Substrate Selectivity

Most naturally occurring acylglycerides display a bias in the distribution of different sized acyl chains in the different positions of the glyceride, whether different sized acyl chains are found in the primary or secondary positions, or whether different sized acyl chains are distributed according to their stereospecific position on the glyceride backbone. In milkfat, there is a definite positional distribution of different sized acyl chains, an example being that C4 acyl chains are found exclusively at the *sn*-3 position in milk triglycerides (Parodi 1982). Because of this variety in glyceride distribution, three different kinds of substrate selectivity operate and are observed in enzymes that process glyceride-based substrates.

1.4.2.1 Typoselectivity

Typoselectivity is the selection of a particular type of acyl substituent, and occurs regardless of its position on the glyceride. At its most basic, this would be defined by a particular enzyme exhibiting a preference for a certain size of acyl chain. Broadly speaking, typoselectivity is what distinguishes esterases from lipases, although, in this case, it is largely based on the solubility differences between esters and lipids. Esterases are typoselective for short acyl substituents, whereas lipases are typoselective for longer acyl chains. Typoselectivity might arise from the length of acyl chain that can be accommodated in the substrate-binding cleft of the enzyme, but could also arise, such as in the case of lipases, from a preferential association with large insoluble acylglycerides rather than with smaller acylglycerides.

1.4.2.2 Regioselectivity

Regioselectivity is also described as positional selectivity, and arises when an enzyme has a preferential activity on either the primary (*sn*-1 or *sn*-3) or secondary (*sn*-2) acyl chains of a glyceride. Pancreatic lipase provides a good example of a regiospecific enzyme, as it has a preference for acyl chains at the primary positions rather than the secondary position (Mattson and Beck 1956). An examination of the structure of the enzyme, solved with methoxyundecylphosphinic acid, a C11 phosphonate-based inhibitor, bound in the active site, gave some insight into how this regioselectivity might occur. The crystallographic structure showed how a monoglyceride-like inhibitor binds in the symmetrical binding site, and showed that a triglyceride is likely to be able to bind equally well with either the *sn*-1 chain or the *sn*-3 chain in a position to be hydrolysed. However, it was obvious from the structure that it would be sterically more difficult for the *sn*-2 chain to be accommodated in the active site in a position amenable to hydrolysis (Egloff, Marguet *et al.* 1995).

1.4.2.3 Stereoselectivity

Stereoselectivity occurs when an enzyme preferentially hydrolyses either the *sn*-1 acyl chain or the *sn*-3 acyl chain. In a racemic triacylglyceride (that is, one with the same chemical groups on the *sn*-1 and *sn*-3 positions), both these groups are chemically identical, although they can be distinguished by their position relative to the

stereocentre by some enzymes. An example of this is rat lingual lipase, which cleaves the acyl chain in the *sn*-3 position of a triglyceride at twice the rate of that in the *sn*-1 position (Paltauf, Esfandi *et al.* 1974).

1.4.3 Determining Specificity

In practice, determining the selectivity of an acylglyceride hydrolase can be a complex process. Typoselectivity is probably the easiest characteristic to investigate, because of the availability of synthetic glyceride substrates. An enzyme can simply be assayed for activity against glycerides of different acyl chain length, as has been done for the EstA and AA7 enzymes examined in this project (Holland, Liu *et al.* 2005). When examining esterase activities, some care must be taken with the concentration of substrates used, as larger less soluble substrates will reach saturation at a lower concentration than shorter substrates. As a result, the enzyme will be presented with a lower concentration of substrate, resulting in a lower activity. This makes comparisons of results for different sized glycerides a risky prospect, although a comparison of different enzymes using identical substrates remains valid.

Determining regio- and stereo-selectivities is more difficult. Although, at first glance, the use of glycerides with different sized acyl substituents at each of the different backbone positions seems to be straightforward, it is possible that a particular chain might be preferentially hydrolysed by the enzyme based on typoselectivity, rather than a stereopreference for the position, giving a false impression of stereospecificity (Brockerhoff and Jensen 1974). Interestingly, it has also been shown that, for some microbial lipases, even the substituent at the secondary position of a glyceride can alter the stereopreference of the enzyme (Stadler, Kovac *et al.* 1995), indicating that caution needs to be applied when interpreting results.

Some researchers have used substrate analogue inhibitors, with the inhibitory group on a specific acyl group of the glyceride (Mannesse, de Haas *et al.* 1997). The inhibitor that shows the highest rate of inhibition then gives some information about the regio- and stereo-selectivity of the enzyme. However, caution must be taken, as these inhibitors are not true substrates, and therefore may not necessarily behave as such.

Chiral high performance liquid chromatography (HPLC) has been used to successfully separate 3,5-dinitrophenylurethane-derivatised diglyceride stereoisomers (Itabashi and Takagi 1987). This method may be able to be extended to the diglycerides of short chain esters, which would then allow the rate of removal of fatty acid chains from an intact glyceride such as tributyrin to be determined.
Some researchers have used radiolabelled triglycerides as substrates, and this could be extended to determining stereopreference using the method described by Karnovsky and Wolff (1960), as detailed by Brockerhoff and Jensen (1974). This work used glycerides labelled in the *sn*-1 position with ^{14}C as a substrate for pancreatic lipase and the distribution of the radiolabel was measured at the end of a reaction procedure. Only half the radiolabel was found in the final product. Had the enzyme been stereospecific, either all or none of the label would have been present.

As lipases that exhibit stereospecificity have the potential to synthesise useful chiral molecules, such as pharmaceuticals, their stereochemistry has been investigated quite extensively (reviewed by Muralidhar *et al.* (2001)). Many of the methods used in these various studies can potentially be extended to the study of esterases.

1.5 The αβ Hydrolase Family

1.5.1 αβ Hydrolases

All the enzymes examined in this project are members of the the αβ hydrolase family (Ollis, Cheah *et al.* 1992). This is a large group of enzymes with members in all domains of life, and with a diverse range of catalytic activities and specificities, including the originally identified acetylcholinesterase, carboxypeptidase II, dienelactone hydrolase, haloalkane dehalogenase and lipase (Ollis, Cheah *et al.* 1992). Several reviews covering the wide variety of catalytic activities and structural variations seen in the family have been published (Heikinheimo, Goldman *et al.* 1999; Nardini and Dijkstra 1999; Holmquist 2000), and the ESTHER database, which contains sequence and kinetic information about some family members, as well as a listing of the 586 αβ hydrolase structures available to date, has also been established (Hotelier, Renault *et al.* 2004). In more recent years, the huge amount of biochemical and structural information about this family has made review of the entire family a daunting task.

Despite a wide diversity in primary sequence, members of the family have a very similar structural core, which contains a variety of insertions, usually related to the particular substrate specificities and/or catalytic activities of the enzyme (Ollis, Cheah *et al.* 1992; Heikinheimo, Goldman *et al.* 1999). These insertions give rise to some notable features of individual enzymes. For example, lipases contain structural features in addition to the basic fold that give rise to interfacial activation, whereby association of the enzyme with an insoluble substrate micelle causes a conformational change that allows the substrate to access the catalytic machinery (Brzozowski, Derewenda *et al.* 1991). Another example of this diversity is acetylcholinesterase, which is one of a very small number of known enzymes that is considered to be catalytically perfect; that is, its rate of substrate turnover is controlled by the rate at which substrates can diffuse into the enzyme active site.

1.5.2 Overall Fold

The αβ hydrolases are globular proteins, composed of a central β sheet surrounded by an arrangement of α helices. In the prototype enzyme, shown in Figure 1.2, the β sheet consists of eight strands, all parallel, except for strand 2, which runs antiparallel to strand 1 and strand 4.

This polarity of the β strands is conserved in the core sheet of all αβ hydrolases, with deviations from this pattern occurring only in additional inserted strands. Since the initial work detailing the minimum requirements of the fold, the discovery that cutinase has only five strands in its sheet and four helices has set the minimal fold size (Martinez, De Geus *et al.* 1992; Heikinheimo, Goldman *et al.* 1999). Examples of enzymes that push the limits of the maximum size of the fold include acetylcholinesterases and some fungal lipases, which can contain up to 12 strands in the central twisted β sheet (Schrag, Li *et al.* 1991; Sussman, Harel *et al.* 1991).

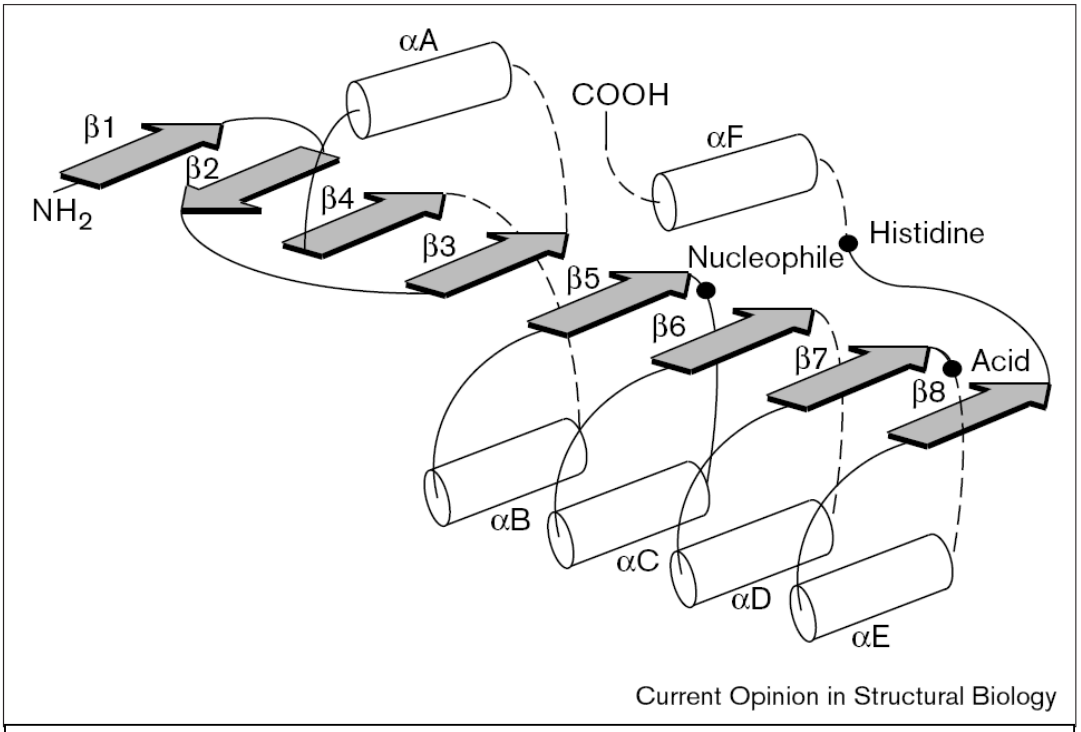


Figure 1.2 *The prototype αβ hydrolase fold, showing the relationship between secondary structural elements and the location of the residues that make up the catalytic triad. Dashed lines indicate areas where insertions are common. Diagram taken from Nardini and Dijkstra (1999).*

In the prototype, there are six helices, each occurring after strands 3 to 8, that surround the central sheet. It should be noted that it is common for additional helices to be inserted into the prototype fold, insertions being particularly common between

strands 6 and 7, and strands 7 and 8. Strands, when inserted, occur at either the N-terminus or the C-terminus of the prototype core, where they generally associate with the other end of the β sheet, extending it. Dienelactone hydrolase, a typical αβ hydrolase, is shown in Figure 1.3.

As well as the wide range of activities exhibited by different αβ hydrolases, these enzymes have a high sequence variability, which is so great that there is debate as to whether all αβ hydrolases are descended from a common ancestor (Ollis, Cheah *et al.* 1992), or whether they are an example of a convergent fold (Holmquist 2000). Regardless of whether the αβ fold is an example of fold convergence, or whether it simply shows that a particular fold is conserved because of its ability to be modified to catalyse a number of reactions, it is clear that the catalytic mechanism used by αβ hydrolases is an example of convergent evolution, as it has similarities in mechanism to that of serine proteases, despite having a completely unrelated fold (Brady, Brzozowski *et al.* 1990). This convergence is detailed in Section 1.8.

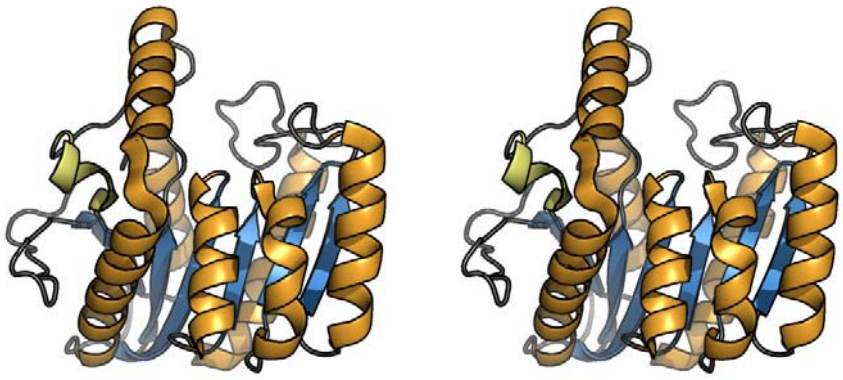


Figure 1.3 *Stereoview of the overall fold of dienelactone hydrolase, a typical αβ hydrolase enzyme (Pathak, Ngai et al. 1988).*
This enzyme has only one insertion compared, with the prototype fold, a short helix, which has been highlighted in yellow.

1.5.3 Substrate Binding

With a wide range of catalytic activities with different substrates, substrate-binding mechanisms in αβ hydrolases are diverse and generally reflect the characteristics of the substrate to be bound. Lipases, which tend to have substrate-binding sites high in non-polar and aromatic sidechains, to help bind hydrophobic substrates, are a good example of this. For example, the lipase from *Candida rugosa* contains a deep hydrophobic tunnel, suited for binding a long acyl chain (Grochulski, Li *et al.* 1993). Acetylcholinesterase has a deep "active-site gorge", which is made up of a large percentage of aromatic sidechains. This was surprising as it was originally thought that acetylcholinesterase would have a negatively charged binding site to attract the ammonium ion of the substrate to it. The authors of the first structure of this enzyme proposed that the ammonium ion is able to interact with the π electrons in the aromatic groups, thus encouraging the substrate to slide into the gorge (Sussman, Harel *et al.* 1991).

One common feature of substrate binding for all the family members is that all the substrate-binding regions of the enzymes are formed from loops inserted between the elements of the basic core of the protein. Changes in the sequences of these loops can thus accommodate the differences in activities and substrate specificities seen in different family members, while leaving the basic structure of the protein unaffected (Ollis, Cheah *et al.* 1992; Heikinheimo, Goldman *et al.* 1999; Nardini and Dijkstra 1999).

1.5.4 Catalytic Triad

αβ hydrolases contain a catalytic system that is made up of the sidechains of three residues (Brady, Brzozowski *et al.* 1990). The residues are always an acid (usually aspartate), a base (histidine) and a nucleophile (usually serine), found in the primary sequence in the well-defined order nucleophile–acid–histidine. Although well separated in the primary sequence, the residues come into close proximity when the protein folds (Figure 1.4). The residues are also found in distinct locations in the secondary structure, the nucleophile occurring between strand 5 and helix C, the acid occurring after strand 7 and the histidine occurring in a loop after strand 8 (towards the C-terminus of the enzyme). It is interesting to note that all the residues are found in loop regions, which potentially allows them to be more mobile than if they were

incorporated into regions of strand or helix. This may be important for controlling the directionality of catalysis, as discussed in Section 1.6.2.3.

1.5.4.1 Catalytic Triad Variations

Although the majority of αβ hydrolases contain a catalytic triad that contains serine, histidine and aspartate residues, variations occur at the nucleophile and acid positions. The serine nucleophile can be replaced with a cysteine residue, as occurs in dienelactone hydrolase (Pathak, Ngai *et al.* 1988; Pathak and Ollis 1990), or an aspartic acid residue, as occurs in haloalkane dehalogenase (Franken, Rozeboom *et al.* 1991).

A feature of the αβ hydrolase family, which is unique amongst the enzyme families that employ a catalytic triad, is the use of a glutamate sidechain as the catalytic acid (Schrag, Li *et al.* 1991; Heikinheimo, Goldman *et al.* 1999). LipA lipase from *Pseudomonas glumae* contains an additional glutamate sidechain near to its catalytic triad aspartate and histidine, which appears to be able either to take over the role of a point mutant of the catalytic aspartate or to assist in orienting a water to perform the acid role (Frenken, Egmond *et al.* 1992; Noble, Cleasby *et al.* 1993), giving both these acidic residues some degree of redundancy in terms of their roles in catalysis.

1.5.5 The Nucleophile Elbow

The nucleophile is found in a distinct sequence motif, initially reported as Sm–X–Nu–Sm–Sm (where Sm represents a residue with a small sidechain (Ollis, Cheah *et al.* 1992), although this is generally now described as G–X–Nu–X–G. This sequence motif allows the formation of a distinct structural arrangement, known as the nucleophile elbow. In three-dimensional terms, this structure consists of a short loop of three residues, found, as mentioned earlier, at the top of β strand 5, which makes a sharp γ turn downwards into α helix C. The sharp turn is possible only because of the two glycine residues in the motif. If these glycines were replaced with more bulky residues, their sidechains would sterically clash, preventing the formation of the tight turn. The sharp nucleophile elbow turn results in very high energy mainchain torsion angles, which results in the serine always being found in an unfavourable region of the Ramachandran plot. This has several advantages in terms of allowing the serine to

perform its catalytic role. Firstly, and most obviously, this mainchain conformation causes the serine to project outwards in an ε conformation, which is important for the catalytic mechanism, in that it allows the serine to access the scissile bond of the substrate (which would not be possible if it were flush or buried). A more subtle advantage of this strained conformation is that the tight turn allows the peptide bond immediately after the serine Cα to adopt a helical conformation, so that the amide groups point directly at the substrate-binding site. The top of the α helix that follows the nucleophile elbow is thus positioned so that the dipole generated by the helix can assist in the stabilisation of the oxyanion formed during catalysis (discussed in Section 1.5.6). The strand helix motif that allows this high energy conformation to exist is commonly found in several different protein folds, demonstrating its importance in catalytic mechanisms (Hakansson 2002).

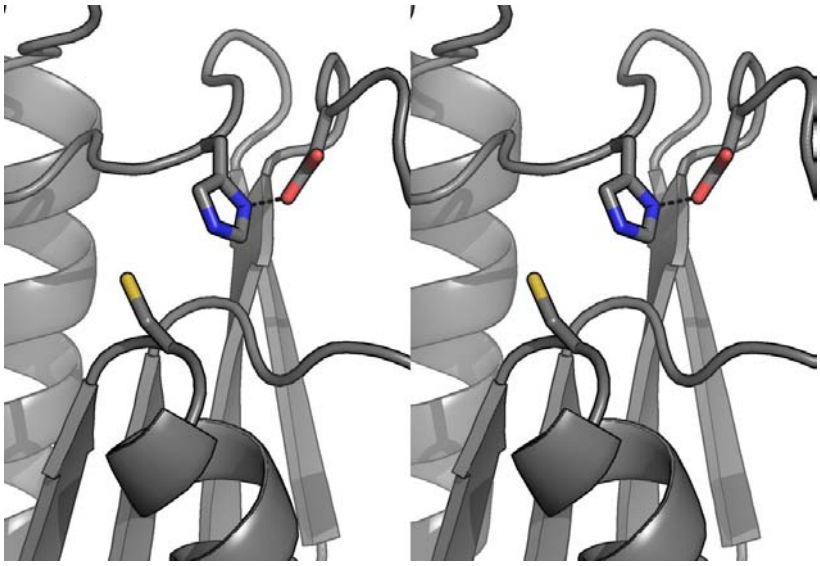


Figure 1.4 *Stereoview of catalytic triad of dienelactone hydrolase (Pathak, Ngai et al. 1988). The N-terminal half of the protein has been omitted for clarity. This enzyme contains a cysteine nucleophile, which can be seen at the top of the nucleophile elbow. The hydrogen bond between the acid and the histidine is shown. However, it should be noted that the histidine and the cysteine are too far apart in this structure for a hydrogen bond to occur, inferring that there must be some movement of the residues before the catalytic cycle can occur.*

1.5.6 The Oxyanion Hole

In addition to the hydrogen-bonding system, which induces the serine to become a better nucleophile, another structural feature, known as the oxyanion hole, lowers the energy required for the nucleophilic attack by stabilising the charge on the reaction intermediate. In αβ hydrolases, the oxyanion hole is generally formed by mainchain amides. Either one or two amides are contributed from an adjacent loop region that occurs at the top of strand 4.

In some lipases, such as that from *Geotrichum candidum*, the oxyanion hole is formed during the movement from the closed state to the open activated state (Schrag and Cygler 1993), a feature that appears to be limited to some, but not all, lipases (Grochulski, Li *et al.* 1994). In esterases, and other non-interfacially activated enzymes, the oxyanion hole usually appears to be preformed (Martinez, Nicolas *et al.* 1994). As the oxyanion intermediate is usually formed directly above the terminus of the helix that follows the nucleophile, it is also stabilised by the mainchain amide immediately after the nucleophile, as well as the N-terminal dipole of the helix. As mainchain amides are involved, testing their contribution to catalysis using mutagenesis is not straightforward. The only possibility is to replace the residues involved with proline residues, but, whereas this might work for the amide that is found on the loop near the catalytic site, a proline incorporated into the nucleophile elbow is likely to cause major structural disruption.

1.5.6.1 Oxyanion Hole Variation

An interesting variation on the oxyanion hole exists in some αβ hydrolases, whereby, rather than the mainchain amide normally located on a nearby loop, the tetrahedral state is stabilised by the hydroxyl group of a tyrosine sidechain, which is found in a similar location to the more common mainchain amide. Such an arrangement was first seen in mammalian prolyloligopeptidase (Fulop, Bocskei *et al.* 1998), but has since been observed in the bacterial enzymes cocaine esterase (Larsen, Turner *et al.* 2002) and X-prolyl dipeptidyl peptidase (Rigolet, Mechin *et al.* 2002). Theoretically, the hydroxyl group of a tyrosine sidechain is more acidic than that of a mainchain amide, and therefore should be better at providing a proton for oxyanion stabilisation. Site-directed mutagenesis studies in which this tyrosine has been mutated to a phenylalanine have served to verify the importance of the tyrosine hydroxyl in

catalysis (Szeltner, Renner *et al.* 2000; Turner, Larsen *et al.* 2002), and to provide an insight into the importance of the oxyanion hole. Whereas simple mutagenesis experiments to remove a particular functional group important to the catalytic cycle are thought to provide the proof of their role, this is not always the case. For example, in the kinetic study of the Y473F mutant of prolyloligopeptidase, the decrease in kinetic activity was smaller than expected, based on similar experiments carried out on subtilisin. It was hypothesised that this was due to the oxyanion becoming hydrated by a nearby water molecule during catalysis, which was able to bind only in such a position when the tyrosine hydroxyl was removed by the mutation to a phenylalanine (Szeltner, Renner *et al.* 2000; Szeltner, Rea *et al.* 2002).

1.5.7 The Tetrad Debate

As well as the three residues in the nucleophile–acid–base catalytic triad, an additional heavily conserved serine residue is found in a turn at the top of strand 6 in many αβ hydrolases, placing it immediately adjacent to the active site, and implicating it in catalysis (Thomas, Church *et al.* 1999). In cases where this serine is not conserved, it is often replaced by a free cysteine, such as in the structure of AA7 described in Chapter 4 of this thesis. Further evidence of the importance of this serine is implied from the observation that a serine exists in a similar position relative to the catalytic triad in the serine protease family (Barth, Wahab *et al.* 1993). This residue was not thought to be important to catalysis in the original description of αβ hydrolases because all of the limited number of structures available to the authors at the time of writing had glycine or alanine residues in this position (Ollis, Cheah *et al.* 1992).

Several hypotheses for the role of this residue were proposed, including the original theory of allowing a downward turn after strand 6 (Ollis, Cheah *et al.* 1992), orienting the water molecule that is involved in the second step of hydrolysis (Thomas, Church *et al.* 1999) and correctly orienting the acid of the catalytic triad, the last having the most experimental support. Using the well-characterised rat esterase as a model system, this extra serine was mutated in turn to glycine, alanine and cysteine. Whereas the serine to cysteine mutant was unstable and therefore not able to be examined, compared with the wild type enzyme, the glycine and alanine mutants showed 5- and 15-fold reductions in specific activity respectively. The rates of acylation and

deacylation of the alanine mutant enzyme were also tested. Whereas the rate of acylation was affected, the rate of deacylation (i.e. hydrolysis) was not, with the authors suggesting that the serine probably directs the optimal orientation of the nearby catalytic glutamic acid, and is unlikely to be involved directly in the hydrolytic step (Stok, Goloshchapov *et al.* 2004).

The potential role of this serine as a structural element should not be overlooked; in all αβ hydrolases, a sharp turn downwards (relative to strand 6) occurs at this point in the peptide chain, which is necessary to sterically clear the catalytic histidine. As the sidechain of the residue at this point projects into the closely packed interior of the protein, its small size is important because a large sidechain would disrupt the packing. Serine sidechains, as well as being small, can participate in hydrogen bonding, making them ideal as an anchor, in this case preventing movement of the end of strand 6, and enabling the bending of the protein chain at this point. Although there are a few structures with glycine or alanine at this position, and therefore no hydrogen bond acceptors, indicating that this tight turn can occur without the need for this anchoring in place, the vast majority have a serine or a cysteine, suggesting a structural advantage in this arrangement.

It is apparent from the evidence obtained that this “fourth residue in the catalytic triad” is likely to have a structural, rather than catalytic role, meaning that the αβ hydrolases retain their status as catalytic triad enzymes.

1.6 αβ Hydrolase Catalysis

1.6.1 Catalytic Mechanism

Like other catalytic-triad-based enzymes, αβ hydrolase esterases use a substituted-enzyme mechanism. In this mechanism, a group is transferred from a donor substrate on to a reactive residue on the enzyme, and then on to an acceptor, as shown in Figure 1.5. In hydrolysis, this final acceptor is water, but, for transferase activity, an acyl acceptor is utilised (discussed in Section 1.7). This differs from a two-substrate ternary complex enzyme, where both donor and acceptor are bound simultaneously and group transfer occurs directly from one to another. It is of importance to note that some of the features of transferase reactions occurring by this mechanism actually require an acceptor other than water to be bound to the enzyme prior to the initiation of enzyme substitution; however, as the donor and acceptor usually occupy overlapping sites during the different stages of catalysis, the acceptor must be shunted into position by a mechanical movement of some kind, between the substitution of the enzyme and transfer of the moving group on to the acceptor.

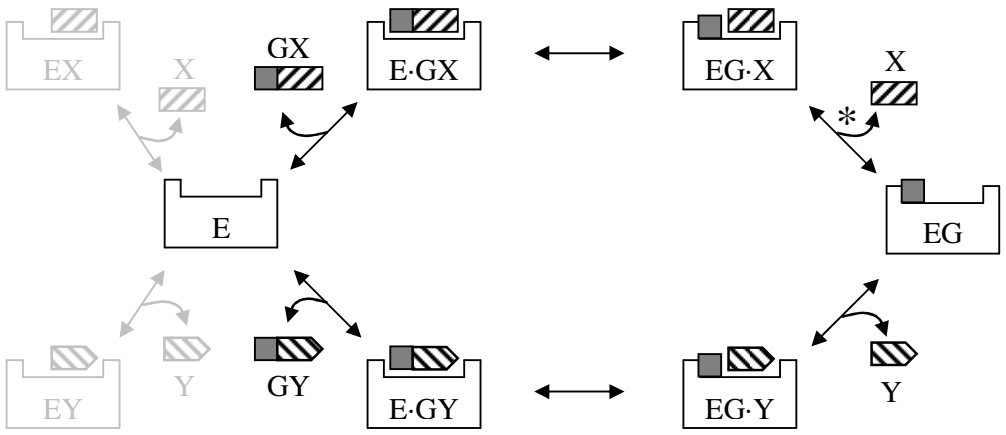


Figure 1.5 *Substituted-enzyme mechanism. E represents the enzyme. In the case of esterases, and assuming that the enzyme cycle operates in the anticlockwise direction, G is the acyl chain that is transferred, Y is the glyceride backbone and other attached acyl chains, and X is the alcohol acceptor. Grey areas represent unproductive binding of products/substrates. In acyltransferases, the binding step indicated with an asterisk can occur by physical movement of a previously bound acceptor. Adapted from Figure 7.3 (Cornish-Bowden 2004).*

1.6.2 Catalytic Chemistry

The chemical mechanisms of catalytic triad enzymes have been elucidated mainly from studies carried out on members of the serine protease family, but much of the mechanism can be extrapolated to the αβ hydrolases because of their very similar catalytic triad layout. The following section describing the enzyme mechanism is therefore mainly based around what is known from serine hydrolase research. The chemistry of the mechanism, described in the following sections, is shown in Figure 1.6.

1.6.2.1 Catalytic Triad Hydrogen-bonding Network and Proton Transfer

There has been much discussion about the mechanism used to effect hydrolysis in these enzymes. Originally, the double proton transfer or charge relay system was proposed, whereby a proton is abstracted from the histidine by the acid, which allows the deprotonated histidine to remove a proton from the catalytic serine, thus activating it (Blow, Birktoft *et al.* 1969). Although this mechanism keeps both the histidine and the aspartic acid neutral during catalysis, in order for the histidine to lose its proton to the acid, its pK_a has to be lower than that of the acid. Although the pK_a of a chemical group can be altered by its environment, histidine, with a pK_a of 6.08, would have to have its pK_a reduced to below that of aspartic acid (3.8) for this mechanism to be able to occur.

More recently, the single proton transfer mechanism has been proposed. In this system, the catalytic histidine is stabilised by the charge on the acid residue. This does not involve a proton transfer between the histidine and the acid, but rather suggests that the function of the acid is to stabilise the charge on the histidine that is generated when it abstracts a proton from the serine during nucleophilic activation. The interaction also plays a structural role in maintaining the alignment of the catalytic triad residues in an optimal position for catalysis. Experimental support for this mechanism comes from the fact that the pK_a of the histidine in the serine protease chymotrypsin is about 7.5, which means that it is likely to be protonated in the acidic environment in which chymotrypsin operates (Robillard and Shulman 1974).

A compromise between the double and single proton relay systems is the low energy barrier hydrogen bond hypothesis, which involves a short strong hydrogen bond being formed between the histidine and the acid. There is some evidence for this, as, in

some structures, the sidechains of the acid and the histidine are less than 2.6 Å apart. This is also supported by nuclear magnetic resonance (NMR) spectroscopy, which shows that such a bond is present (Frey, Whitt *et al.* 1994). Although the histidine does technically retain its proton in a low energy barrier hydrogen bond system, the proton nonetheless is virtually shared between the acid and histidine residues.
For an excellent discussion on the history and evidence for all three mechanisms, as well as serine proteases in general, the reader is directed to the review of Hedstrom (2002).

1.6.2.2 Nucleophilic Attack

Once activated, the nucleophile can then attack the scissile bond of the substrate, to form a tetrahedral enzyme–substrate oxyanion transition state, which is stabilised by hydrogen bonding between the oxyanion and the hydrogen bonds donated by the groups that form the oxyanion hole. The unstable tetrahedral intermediate then rapidly collapses into an acyl-enzyme intermediate. The bond that is formed between the nucleophile and the substrate during the nucleophilic attack remains, whereas the substrate scissile bond is broken during the collapse back to a trigonal centre, as the first reaction product is a better leaving group. This first product is then able to diffuse away from the enzyme.

1.6.2.3 Reaction Directionality

A mechanism must operate during the nucleophilic attack to prevent the tetrahedral state from simply collapsing back *via* the nucleophile and histidine base, which would result in breakage of the enzyme–substrate bond and lead to a futile cycle. It has long been hypothesised that this directionality must be orchestrated by hydrogen bonding, with the subsequent proton transfer being able to occur only in the direction that is favourable for the forward progress of catalysis. This results in a cycle of proton transfer, as opposed to an unproductive back-and-forth mechanism (Wang 1970).
The mechanism for how this directionality is enforced has been intensively studied in serine proteases, where it appears that the residues of the catalytic triad move apart slightly during the initial catalytic step, separating them just enough to prevent hydrogen bonding, thus preventing an overall reversal of the reaction.

Exactly how this occurs remained unclear for many years, despite the acid–base–nucleophile mechanism being well accepted. Even though a hydrogen bonding network between the triad members was proposed, early crystallographic studies showed that the nucleophile and the histidine base were too far apart and in orientations that were not conducive for hydrogen bonds to form (Matthews, Alden *et al.* 1977; James, Sielecki *et al.* 1980; Cohen, Silverton *et al.* 1981). Eventually, however, as more structures were solved to higher resolution, structures of serine proteases that showed that there was a hydrogen bond between the base and the nucleophile were obtained (Bode, Chen *et al.* 1983; Tsukada and Blow 1985). This was supported by an NMR experiment that also indicated the presence of the nucleophile–histidine bond (Bachovchin 1986). All these studies served to demonstrate that some mobility of the catalytic residues is possible in serine-protease-type enzymes. Recent reports of crystallographic structures at atomic resolution have shown that there indeed appears to be a subtle parting of the nucleophile and the histidine during catalysis that assists the reaction to achieve a productive cycle (Radisky, Lee *et al.* 2006).

Some movement is also seen in the catalytic histidine in αβ hydrolase structures, with the histidine in the structure of tabun-inhibited mouse acetylcholinesterase undergoing a displacement (Ekstrom, Akfur *et al.* 2006), whereas, in dienelactone hydrolase, the histidine becomes disordered when the enzyme is bound to phenylmethylsulphonyl fluoride (PMSF), suggesting that it is more mobile (Robinson, Edwards *et al.* 2000). Analysis of the structures of a number of αβ hydrolases has shown that it is not uncommon for the distance between the nucleophile hydroxyl and the histidine Nε1 to be too great for a hydrogen bond to form, implying that some kind of movement must occur for the enzyme catalytic system to be activated, and to encourage the reaction to proceed in a forward direction. Such a movement has been identified in *Bacillus subtilis* lipase, in which it appears that the nucleophile sidechain rotates into hydrogen-bonding range of the catalytic histidine, allowing enzyme activation (Kawasaki, Kondo *et al.* 2002).

The mechanism for disrupting this hydrogen-bonding network between the catalytic residues in αβ hydrolases was identified in this project, where the catalytic serine sidechain of the esterase AZ4 was found to rotate back out of hydrogen-bonding range of the catalytic histidine during the nucleophilic attack (Section 5.5). This new

rotamer is distal enough from the histidine to break the hydrogen-bonding system, and could be a mechanism to prevent a futile collapse of the oxyanion.

1.6.2.4 Acyl-enzyme Intermediate

Once the scissile bond has broken, the first enzyme product (in the case of esterases, this is an alcohol) is released from the enzyme as it is no longer physically bonded to the remainder of the substrate. The remainder of the substrate is bound to the enzyme in an acyl-enzyme intermediate. The formation of the acyl-enzyme intermediate helps to lower the energy required for the reaction, as, if it was necessary to precisely orient two substrates (in this case, the acylglyceride substrate and the water acceptor), there would be a higher entropy barrier because of the necessity of positioning both molecules. In its acylated state, the enzyme is required to precisely position only the acceptor for deacylation to proceed, as the donor is relatively rigidly positioned by being covalently bound to the enzyme (Spector 1973; Silverman 2000).

1.6.2.5 Deacylation of Enzyme

A water molecule enters the catalytic site of the enzyme, and gets close enough to the catalytic histidine to donate a proton to it in a process that is similar to the process that occurs in the activation of the serine residue in the first step of catalysis. The activated water then carries out a second nucleophilic attack, this time on the acyl-enzyme intermediate. A second tetrahedral intermediate is formed, again stabilised by the oxyanion hole and the dipole from the N-terminus of helix C. This time, however, the tetrahedral intermediate collapses, breaking the bond between the substrate and the catalytic serine. The serine is then reprotonated by the proton that was donated to the histidine to complete the cycle. The carboxylic acid product can then diffuse away from the enzyme, allowing the catalytic cycle to begin again.

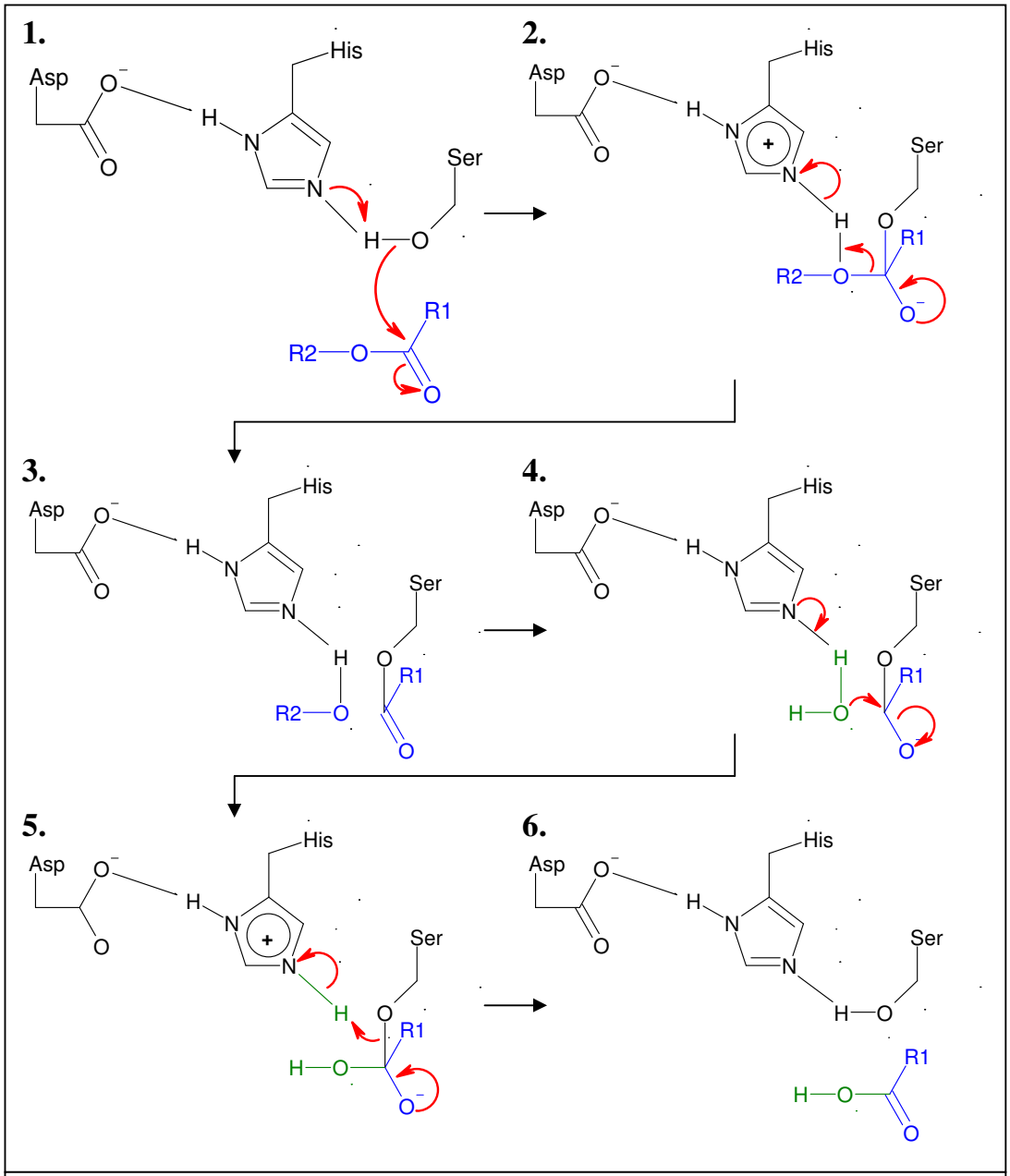


Figure 1.6 *Mechanism of ester hydrolysis.*

Substrate binds and carbonyl carbon is attacked by the enzyme nucleophile (1) to form a tetrahedral oxyanionic intermediate (2). The intermediate decomposes to form an acylated enzyme intermediate (3), and the first product (alcohol) is released. Water binds and performs nucleophilic attack on the acyl-enzyme (4), which forms a second tetrahedral anion (5). This decomposes to break the acyl-enzyme bond and liberate the second (carboxylate) product (6). Diagram adapted from Figure 15.23 (Voet and Voet 2004).

1.7 Transferases

1.7.1 Transferase Reaction

From a technical standpoint, ester hydrolysis is a two-substrate transferase reaction, in which an acyl chain is transferred from an acylglyceride on to a water acceptor. However, in aqueous solution, water is present in such a large concentration that, in terms of kinetic behaviour, it can be regarded as being present in excess, meaning that hydrolysis can be regarded as a single-substrate reaction (Cornish-Bowden 2004). However, for transfer reactions that occur in an aqueous setting, the acceptor substrate will not be present at a concentration even remotely close to that of water, meaning that the mechanism becomes more complex and that a transferase enzyme must have features that help group transfer on to a molecule other than water.

1.7.2 Transferase Catalytic Mechanism

In transferases, a chemical process known as alcoholysis occurs. The reaction proceeds identically to that of a hydrolase, through the attack of the serine nucleophile, to the acyl-enzyme intermediate stage. At this stage, rather than an activated water molecule attacking the acyl-enzyme intermediate to yield a final hydrolysed product, an alternative acceptor is activated by the histidine base in a similar way to the water of the hydrolytic reaction, to complete deacylation of the enzyme. This means that the acceptor must be bound in the catalytic site within hydrogen-bonding distance of the histidine, and that water must somehow be prevented from completing the hydrolysis. For alcohol acceptors, alcoholysis is virtually chemically identical to a hydrolysis reaction, except that an alcohol group is activated to become a nucleophile. Two general mechanisms have been proposed to ensure that a group transfer to an acceptor other than water occurs (Milkowski and Strack 2004).

1.7.2.1 Superior Nucleophilicity

It has been postulated that, if an acceptor is more easily activated than a water molecule to carry out a nucleophilic attack, the rate of product formed would be greater than the amount of substrate merely hydrolysed. In this case, the acceptor and solvent water molecules would compete to deacylate the acyl-enzyme intermediate. This mechanism might occur where alcohols are the acceptor, as they can act as a better nucleophile than water (De Simone, Mandrich *et al.* 2004).

1.7.2.2 Shielding of Catalytic Site from Water

The second method proposed involves the catalytic site being shielded from water molecules, which can occur by several mechanisms.

1.7.2.3 Water Shielding by Hydrophobic Environment

One way to achieve the exclusion of water from the catalytic site is shown by the mycolyl transferase Antigen 85B/C from *Mycobacterium tuberculosis*. This member of the αβ hydrolase family is involved in the synthesis of cell-wall compounds by the transfer of mycolic acid from trehalose-6-monomycolate on to a second mycolated trehalose molecule. Crystal structures of the protein complexed with various ligands identified a likely binding site for trehalose-6-monomycolate, an adjacent binding site for the mycolate intermediate and a third site, which was postulated to be for the acceptor trehalose-6-monomycolate. All of these sites were in or around the cleft that contains the catalytic machinery, providing a mechanism for the enzyme to anchor itself to the hydrophobic compounds of the cell wall, while at the same time excluding water from the catalytic site. Similarly, water can also be excluded from the active site when lipases are used to catalyse synthesis reactions in organic solvents, where water is simply not present in the reaction, in order to force the use of an alternative acceptor (Klibanov 1989; Klibanov 2001).

1.7.2.4 Water Shielding by Acceptor Binding

Water shielding may play a role in the alcoholytic mechanism of lactic acid bacteria esterases, and other transferases that operate in an aqueous environment. From the

results obtained in this project, it would appear that the binding of an acceptor to the enzyme may block water from entering the active site, thus preventing hydrolysis and ensuring that the alcoholytic reaction, rather than hydrolysis, proceeds. There is some evidence for this occurring in AA7 esterase, as discussed in Chapter 4.

Some protease inhibitors, such as bovine pancreatic trypsin inhibitor (BPTI), have also been shown to bind to their target hydrolases so tightly that they prevent water from entering the active site and completing the reaction (Voet and Voet 2004). This shows that similar strategies are used to exclude water in transferase reactions and to prevent completion of hydrolysis by inhibitors.

1.7.2.5 Water Exclusion by Internalisation of Active Site

Recently, the structure of an acyltransferase from *Mycobacterium smegmatis* was solved. This structure shows a unique mechanism of water exclusion, in that the enzyme forms an octamer in solution, with the catalytic site located internally. This prevents water from accessing the binding site, and is postulated to allow alcoholysis to occur in an aqueous solvent, at a great molar excess of water (Mathews, Soltis *et al.* 2007).

1.8 Convergent Evolution of Catalytic Triad Enzymes

1.8.1 Catalytic Triad Enzymes

As mentioned earlier, there are four well-characterised enzyme folds that share a similar catalytic triad. These folds are the αβ hydrolases, the serine proteases, subtilisin and papain. All these four families have distinctly different overall folds, but contain catalytic sites that are remarkably similar in their spatial arrangement and mechanism of action (Garavito, Rossmann *et al.* 1977; Brady, Brzozowski *et al.* 1990; Dodson and Wlodawer 1998).

1.8.2 Comparison of αβ Hydrolases and Serine Proteases

The serine protease family has been well studied, and was the first subset of the so-called "catalytic triad" enzymes to be characterised and to have the mechanisms unravelled. On the basis of these findings, predictions about the catalytic system of lipases were made almost 20 years before the first structure was reported (Brockerhoff 1973; Brady, Brzozowski *et al.* 1990). Structural analyses of many members of both enzyme families have allowed detailed comparisons of both their catalytic and binding sites to be made, which have led to a detailed understanding of individual mechanisms.

Both families utilise an acid–base–nucleophile catalytic triad, operating with similar residues and the same chemistry. An interesting difference between the two enzyme families is that the catalytic triad of αβ hydrolases is effectively a mirrored version of that seen in serine proteases (Brady, Brzozowski *et al.* 1990), as shown in Figure 1.7. This mirroring of residues includes the "fourth catalytic residue" serine, which, in both serine proteases and αβ hydrolases, has a similar (although reversed) spatial relationship to the catalytic acid sidechain, and appears to be required for maximal catalytic activity in both types of enzymes (McGrath, Vasquez *et al.* 1992; Stok, Goloshchapov *et al.* 2004).

Although the catalytic groups of the triad are superimposable (taking into account the mirrored nature of the triad), the secondary structural elements that make up the catalytic triad are very different in each of the two enzyme families.

Perhaps the most striking difference between the enzymes is the lack of a nucleophile elbow in the serine proteases. Rather than being held in a projecting conformation by a tight turn, the nucleophilic serine is located on a loop, in a favoured Ramachandran conformation. The serine amide contributes to the oxyanion hole, as in the αβ hydrolases, but the second oxyanion hole amide comes from the mainchain amide of the residue after the nucleophile instead of an amide from an adjacent loop. The other residues of the catalytic triad are also found in different secondary structural units compared with those in αβ hydrolases; the histidine is found on a short helix, the acid is found in a loop region and the "fourth catalytic residue" serine is found near the end of a β strand.

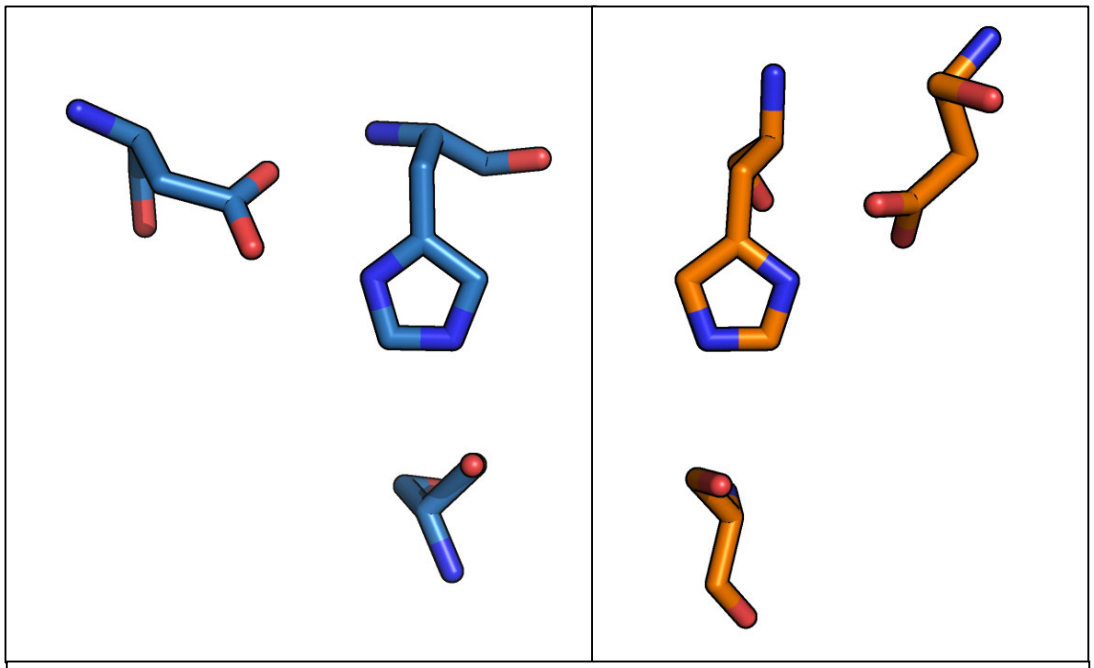


Figure 1.7 *Catalytic triads of chymotrypsin (left) and AZ4 αβ hydrolase (right). Note the mirror image relationship between the two catalytic triads. The acid is the topmost residue, with the histidine in the centre and serine at the bottom.*

1.9 Microbial Esterase and Lipase Classification

1.9.1 The Eight Class System

The most recent published attempt to develop a classification system for bacterial lipolytic enzymes, including the esterases, categorises the bacterial carboxylesterases and lipases into eight classes (I–VIII), with class I then divided into six subfamilies (Arpigny and Jaeger 1999). This classification was based on both the sequences of the enzymes and known biological properties, with enzymes grouped as follows.

1.9.1.1 Family I and Subfamilies

Family I contains what are termed "true lipases", which exhibit interfacial activation, with Subclass I.1 containing enzymes of mass 30–32 kDa and Subclass I.2 containing slightly larger lipases of about 33 kDa. Expression of these two subclasses also requires the usage of lipase-specific foldases (see Section 1.9.2 for a more detailed review). Subclass I.3 contains larger enzymes (50–60 kDa), which, unlike the Subclass I.1 and I.2 enzymes, contain no cysteine residues. Classes I.4, I.5 and I.6 are all made up of enzymes from Gram-positive species. Small lipases (≈ 20 kDa), represented by enzymes from *Bacillus subtilis* and *Bacillus pumilus*, are grouped into Class I.4, and staphylococcal enzymes of ≈ 70 kDa are grouped into Class I.5. Finally, other lipases from *Propionibacterium* and *Streptomyces* are grouped into Class I.6.

1.9.1.2 Family II

Family II enzymes are alternatively referred to as GDS(L) lipases, as the sequence motif about the serine nucleophile is not the conventional GXSXG motif. An interesting member of this family is the esterase from *Streptomyces scabies*; this esterase has a catalytic diad, rather than a triad, with the aspartate acid being replaced by a mainchain carbonyl group (Wei, Schottel *et al.* 1995). Interestingly, the group members appear not to exhibit the αβ hydrolase fold as defined by Ollis et al. (1992), although this is an assumption based only on the group's sequence similarity to the *S. scabies* enzyme.

1.9.1.3 Family III

Family III lipases are similar to human platelet-activating factor acetylhydrolase (PAF-AH), showing a 20% sequence similarity, and are represented by the enzyme cloned by Cruz et al. (1994).

1.9.1.4 Family IV

This class contains enzymes that are similar to mammalian hormone sensitive lipase (HSL). There is also some similarity to hepatic and pancreatic lipases, and acetylcholinesterases (Hemila, Koivula *et al.* 1994).

1.9.1.5 Family V

Family V members bear more similarity to non-lipolytic enzymes, such as epoxide hydrolases, dehalogenases and haloperoxidase. However, these enzymes are still αβ hydrolases despite these differences in activity.

1.9.1.6 Family VI

This family contains small carboxylesterases, with a mass between 23 and 26 kDa. They have a minimal version of the αβ hydrolase fold, with some similarity to eukaryote lysophospholipases.

1.9.1.7 Family VII

Family VII enzymes are large (55 kDa) enzymes with sequence similarity to the eukaryotic acetylcholinesterases.

1.9.1.8 Family VIII

This family contains enzymes that are similar to the β-lactamases, and do not appear to have the conserved GXSXG catalytic serine motif characteristic of the αβ hydrolases. They also lack conservation of the order of catalytic residues (usually Ser–Asp–His), which suggests that this family are not αβ hydrolases.

1.9.2 Lipase-specific Foldases

In some Gram-negative bacterial species, such as *Pseudomonas cepacia*, extracellular lipases from Classes I.1 and I.2 are found in an operon with a lipase-specific foldase (Lif) that is essential for lipase activity (Jorgensen, Skov *et al.* 1991). Lifs, which have a transmembrane region, have been shown to act as a chaperone to promote the correct folding and secretion of lipases, as they allow refolding of denatured target proteins (Hobson, Buckley *et al.* 1993), and appear to anchor to the cell wall so that their active domain is exposed to the periplasm. More recently, the structure of *Burkholderia glumae* lipase in complex with its chaperone has been published, showing the chaperone, a mainly α-helical structure, to be wrapped around the outside of the hydrolase (Pauwels, Lustig *et al.* 2006). Similar chaperones have not been identified in Gram-positive species, and correspondingly the enzymes covered in this project do not appear to require any chaperones for successful expression in *Escherichia coli*. However, it should be noted that, as the dairy lactic acid bacteria esterases are intracellular, perhaps these chaperones are reserved solely for assisting in passing Gram-negative secreted lipases through the cell membrane.

1.10 Inhibition Studies

1.10.1 Inhibitors of Esterases and Lipases

Enzymology has often relied on the use of different inhibitors to bind to the active site, and/or the catalytic machinery of enzymes to gain an insight into the chemistry of their reactions and their substrate-binding requirements.

Non-specific inhibitors can be used to determine the chemistry and the types of residues involved in catalysis. For example, diisopropylphosphofluoridate (DFP) was used to show that a single reactive serine sidechain was present in chymotrypsin (Oosterbaan, Kunst *et al.* 1958) and acetylcholinesterase (Augustinsson and Nachmansohn 1949). It is interesting to note that many chemical warfare nerve agents are designed to inactivate acetylcholinesterase by forming adducts with the catalytic serine (for an example, see the recent structure of tabun-bound mouse acetylcholinesterase (Ekstrom, Akfur *et al.* 2006)).

The catalytic mechanism in lipases was also partially elucidated by experiments in which the inhibitor diethyl *p*-nitrophenyl phosphate was complexed to several enzymes. This work indicated that a single serine sidechain was always acylated by the inhibitor, showing that it is also the reactive residue in this class of enzymes (Maylie, Charles *et al.* 1969; Maylie, Charles *et al.* 1972).

Such small chemical agents tend to be non-specific for the enzyme they inactivate, and may be really useful only in the determination of the active site residues, as they may not trigger any changes that take place during binding of a larger, authentic substrate. For this reason, inhibitors that use similar chemical mechanisms for inactivation, but that also more closely resemble a substrate of the targeted enzyme, have been developed.

1.10.2 Substrate Analogue Inhibitors

Substrate analogues cover a wide range of compounds. The simplest substrate-based inhibitors include some inhibitory peptides, such as Ile–Pro–Ile, which appears to be a very poor substrate for dipeptidyl peptidase IV. Because the peptide is hydrolysed very slowly, it acts as an inhibitor (Rahfeld, Schierhorn *et al.* 1991).

Natural peptide inhibitors also exist for most proteases that contain disulphide linkages that join the two products of the cleavage. This prevents the two peptide fragments from separating, so that the hydrolytic products remain bound to the enzyme, preventing the association of further substrates (Birk 1983).
More complex inhibitors that are specific for particular enzymes have been synthesised; they mimic certain binding determinants of a physiological substrate, but contain either no scissile bonds or groups that, when partially processed by the enzyme, produce a stable complex with the enzyme. These are often known as transition state analogues.

1.10.3 Transition State Analogues

> ***"I think that enzymes are molecules that are complementary in structure to the activated complexes of the reactions that they catalyse, that is, to the molecular configuration that is intermediate between the reacting substances and the products of reaction for these catalysed processes."***
>
> – *Linus Pauling (1948)*

Pauling's hypothesis, that enzymes function by stabilising a high energy intermediate enzyme–substrate complex formed during catalysis, was shown to be essentially correct with the publication of the structure of lysozyme, the first high resolution enzyme structure, which contained clues as to how the transition state was stabilised by the enzyme (Blake, Johnson *et al.* 1967). Enzymes catalyse a reaction by binding to the substrate much more tightly in its transition state than in its native conformation. This has a stabilising effect, and led to the idea that enzyme inhibitors could be produced by making molecules that are similar in structure to those adopted by a substrate during the transition state of catalysis (Wolfenden 1969).
Such inhibitors are produced by designing a compound that superficially resembles a substrate for the enzyme in question, but contains a group that, although structurally resembling the transition state of the enzyme, cannot be processed. Although these

adducts are not permanent, their lifetime is usually many magnitudes longer than a normal transition state, making enzyme inhibition effectively permanent over the course of a typical experiment.

1.10.4 Inhibitor Design

In αβ hydrolases, the trigonal reaction centre bond of the substrate changes to a tetrahedral geometry in the two transition states of the reaction, a conformation that is stabilised by the interactions made between the enzyme–substrate complex and the mainchain amides of the oxyanion hole. Thus, if the reaction centre in a substrate is replaced by a group having tetrahedral geometry, a compound with a high affinity for the catalytic site can be created.

This has been achieved by substituting the reactive carbon of a substrate with boron, which is able to form a stable negatively charged tetrahedral state (Lienhard, Secemski *et al.* 1971), but has more commonly been done with phosphorus. As phosphorus is pentavalent, it can form a stable tetrahedral compound, while being able to undergo a transition that results in acylation of the enzyme in an imitation of the mechanism that occurs during the acylation of the enzyme with a normal substrate. When the phosphate reaction centre is linked to a good leaving group, this bond is preferentially broken during the collapse of the transition oxyanion, leaving the phosphate covalently bonded to the nucleophile. The resultant complex is resistant to deacylation, forming a long-lived inhibited state, which can be regarded as irreversible. Figure 1.8 demonstrates how the tetrahedral anionic transition state seen in esterase and lipase catalysis is mimicked by the adduct formed by enzymatic reaction with a phosphonate-based inhibitor, and Figure 1.9 demonstrates the mechanism by which the inhibited state is established.

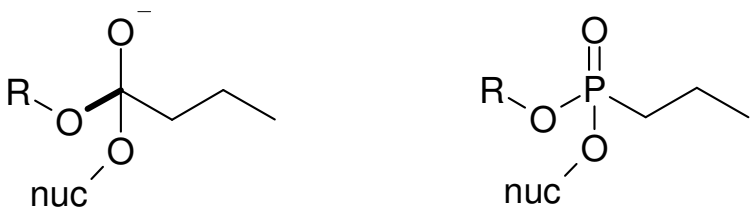


Figure 1.8 *Transition States*

Left: Tetrahedral oxyanion transition state, with scissile bond shown in bold.

Right: Tetrahedral phosphonate adduct

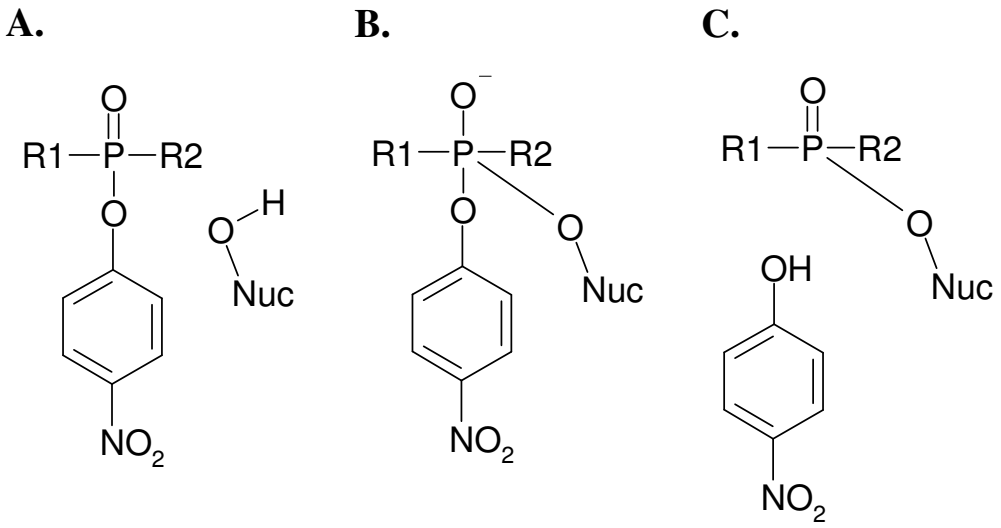


Figure 1.9 *Establishment of a tetrahedral phospho-enzyme complex.*

A. Nucleophilic attack.

B. Oxyanion intermediate.

C. Tetrahedral phospho-enzyme complex and liberated p*-nitrophenol product.*

1.10.5 Lipases and Inhibitors

Initial studies on αβ hydrolases using transition state analogues were carried out, as described earlier in this review, using DFP (acetylcholinesterase) (Augustinsson and Nachmansohn 1949) and diethyl *p*-nitrophenyl phosphate (pancreatic lipase) (Maylie, Charles *et al.* 1969), but exactly how these compounds inhibit enzymes was not known until the structure of the lipase from *Rhizomucor miehei* in complex with *n*-hexylphosphonate ethyl ester was reported (Brzozowski, Derewenda *et al.* 1991). The inhibitor was a short chain ester analogue, and, although useful for giving insights into mechanisms, only superficially resembled a true glyceride-based ester substrate. Mannesse, Boots et al. (1995) set out to produce inhibitors that used the same phosphate-based inhibitory centre, but that resembled a triglyceride in all other respects (Mannesse, Boots *et al.* 1995). As triacylglycerides contain three ester bonds, in addition to the replacement of the scissile bond with a phosphate group, the two remaining ester bonds can be replaced with non-hydrolysable bonds, so that their attached acyl chains are preserved in reactions with the enzyme. To achieve this, Mannesse, Boots et al. (1995) used carbamoyl linkages; other groups have used ether linkages (Stadler, Zandonella *et al.* 1996; Taylor, O'Connor *et al.* 2001). To completely mimic a triacylglyceride, inhibitors that retain the ester linkages at the non-inhibitory positions have also been made (Marguet, Cavalier *et al.* 1999). Figure 1.10 compares the structures of the products of some of these syntheses with tributyrin.

1.10.6 Organophosphate Chirality

As well as the stereocentre inherent in the glyceride backbone, non-symmetrical organophosphates have an additional stereocentre about the phosphorus atom, because it can have four different groups attached. This means that R and S enantiomers are produced during the synthesis of these inhibitors. This added stereochemistry can affect the efficacy of the inhibitor, depending on the direction in which it binds relative to the catalytic site, which in turn can be affected by the substrate-mimicking groups attached to the inhibitory centre. Because of this directionality, in some cases, the enantiomer will not react with the enzyme, as the carbonyl group attached to the reaction centre will not be oriented correctly for the formation of an oxyanion in the oxyanion hole during the nucleophilic attack. R and S enantiomers can be separated

by chiral HPLC, although this is not always successful for triglyceride analogues (D Lun, personal communication, 2007).

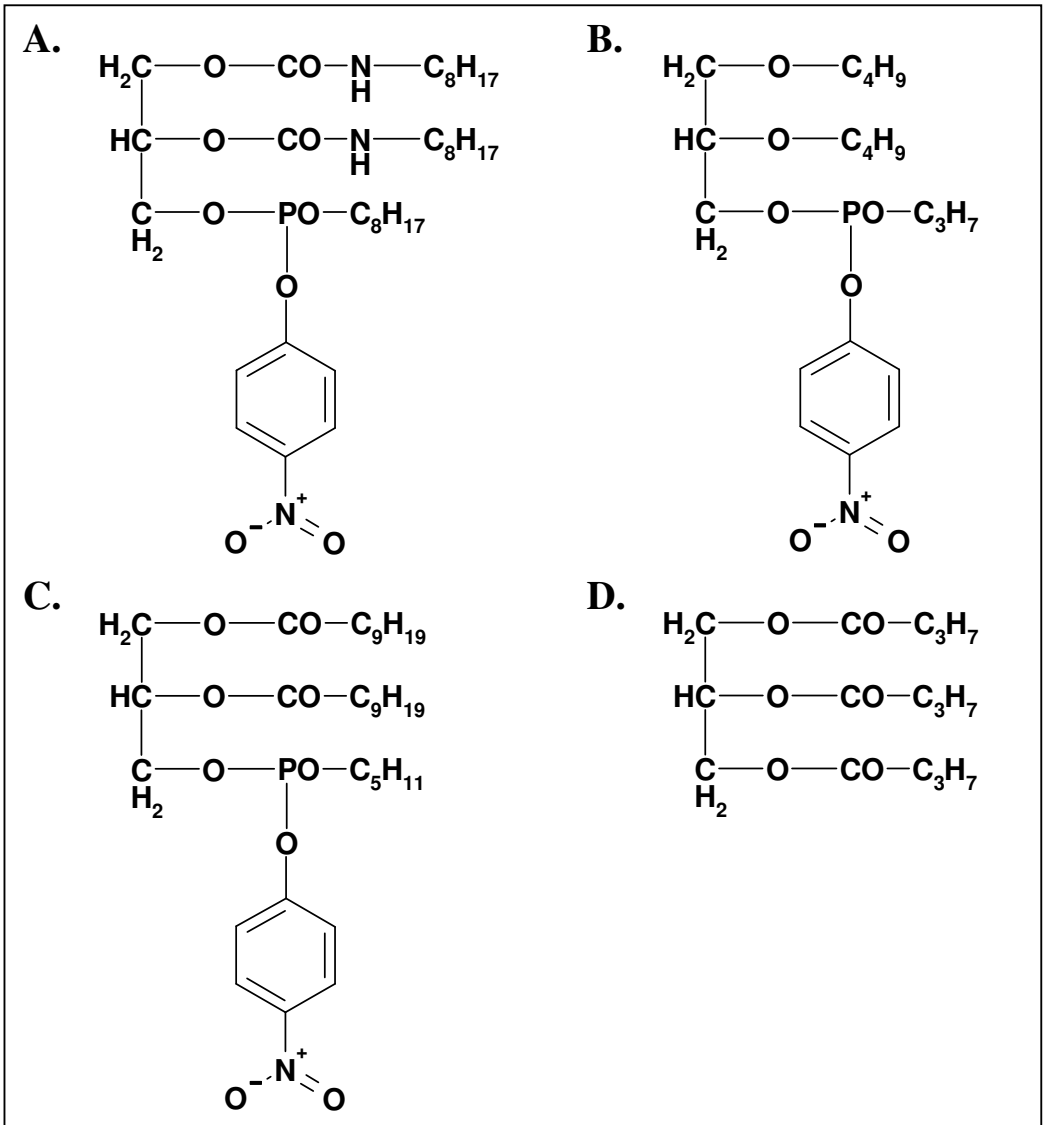


Figure 1.10 *Phosphonate triglyceride analogues.*

- *A. 1,2-Dioctylcarbamoylglycero-3-O-*p*-nitrophenyl octylphosphonate (Mannesse, Boots et al. 1995).*
- *B. 1,2-di-O-butyl-*sn*-glycerol-3-O-*p*-nitrophenyl-*n*-propylphosphonate (Taylor, O'Connor et al. 2001).*
- *C. 1,2-diacyl-*sn*-glycerophosphonate (Marguet, Cavalier et al. 1999).*
- *D. Tributyrin.*

Note that no attempt has been made to represent stereochemistry in these diagrams.

1.11 Research Aims

Despite being first isolated from dairy lactic acid bacteria more than a decade ago (Holland and Coolbear 1996), and diffracting crystals of EstA being obtained shortly thereafter (Zheng 1998), there has been little progress to date on producing structural information about this, or other more recently characterised enzymes. The primary aim of this project is therefore to produce structural information about esterases that have been identified as being of interest to the dairy industry. If this information can be related this to what is known about the activities of the enzymes it will allow a better overall understanding of their function, and contribution to ester based flavour compounds in dairy fermentations. As a secondary aim, it was hoped that structural solutions would lead to the possibility of rational substrate binding site design to allow the fine tuning of desirable activities to be engineered into the enzymes.

Three enzymes were chosen for structural study: EstA, from *L. lactis*, AA7 from *Lb. rhamnosus*, and AZ4 from *Lb. rhamnosus*.

EstA was selected, as it has been characterised to a reasonable extent (Holland and Coolbear 1996; Shaw 1999). EstA has both a hydrolase and transferase activity, and might therefore give information about how these activities are carried out by the enzyme. EstA is a member of Class 3 of the bacterial lipase classification system (Arpigny and Jaeger 1999), meaning it has sequence similarity to extracellular lipases.

AA7 was also selected as it has also been characterised as having both hydrolase and transferase activities (Holland, Liu *et al.* 2005; Delabre, Ng *et al.* 2006), and a structural comparison to EstA might give further indications as to how these different activities are carried out by the same enzyme. AA7 is a member of Class IV of the lipase class system, and is therefore a Hormone Sensitive Lipase analogue (HSL).

AZ4 was the final enzyme selected, as it is appears to exclusively perform a hydrolase activity (M-L Delabre, personal communication). Comparison of the structure of this enzyme to the other two might give insights into the difference between these two reactions, and the structural features that lead to each. AZ4 also appears to have relatively narrow typoselectivity, with little activity seen on substrates with substituent chains larger than 4 carbons in length. AZ4, with similarity to the recently solved structure of Est30, falls outside the lipase classification system of Jaeger and Arpigny (1999).

1.12 Summary

The coupling of traditional dairy product manufacture with modern technology and analytical techniques has led to a better understanding of the processes involved in fermentations, and in the maturation of dairy products. In particular, the study of tributyrin esterases, involved in the generation of free fatty acids and esters in products such as cheese, has led to a better understanding of how these enzymes function at a biochemical level.

Tributyrin esterases act on glycerides, either to hydrolyse them to produce a free fatty acid or to catalyse the transfer of a fatty acyl group on to an alcohol acceptor to produce a short chain ester. Despite a number of studies of these enzymes, no physiological role has yet been established.

Glycerides have a chiral centre, located at the secondary carbon of the glyceride backbone. Enzymes acting on glycerides exhibit a range of specificities for the glyceride chain on which they act, classified by typoselectivity for a preferred length of acyl chain, regioselectivity for primary or secondary glyceride positions and stereoselectivity for either of the primary glyceride positions.

The tributyrin esterases are members of the αβ hydrolase family, a protein fold that is widespread both in its phylogenetic distribution and in the range of activities that its members can catalyse. The αβ hydrolases have a conserved twisted β-sheet core, with six conserved helices found on either side of the sheet. Variations on the fold are very common, with insertions typically occurring in certain “hotspots” in the fold, while preserving the core fold. The αβ hydrolases have a nucleophile–base–acid.catalytic triad that is in a similar spatial arrangement, although mirrored, to the catalytic triad that is found in the serine proteases such as chymotrypsin. Whereas the base and acid residues of the triad are found in loop regions, the nucleophile is found in a tight γ hairpin turn, called the “nucleophile elbow”. This positions the nucleophile in a high energy mainchain conformation, but allows the sidechain to project into the substrate-binding site and to access the reaction centre. A hydrogen-bonding network is formed between the sidechains of the triad residues, which results in polarisation and

activation of the nucleophilic sidechain to perform a nucleophilic attack on the substrate. During the attack, the substrate assumes a tetrahedral transition state that is stabilised by several mainchain amides that form a pocket known as the oxyanion hole, before collapsing back to a trigonal state, breaking the scissile bond of the substrate and producing an acylated enzyme intermediate. This acylated intermediate is then subjected to a second nucleophilic attack, carried out by either a water molecule (hydrolysis) or an alcohol (alcoholysis). Alcoholysis can occur in aqueous solution as a result of one of several possible mechanisms that result in the exclusion of water from, but access of the acceptor to, the catalytic site.

Inhibitors that are substrate analogues have been used to characterise microbially sourced lipases and esterases, in biochemical and structural studies. These inhibitors often have a phosphonate instead of a carbon at the reaction centre. The phosphonate is able to form a stable tetrahedral intermediate, which structurally resembles the tetrahedral oxyanion transition state of catalysis. Various groups that resemble a substrate are attached to the reaction centre to assist in binding of the inhibitor, and to provide a structural model of how a substrate might bind to the enzyme.

Chapter 2

The Structure of EstA from *Lactococcus lactis*

Chapter 2

The Structure of EstA from *Lactoccus lactis*

2.1 Introduction

EstA is an esterase that was originally isolated from the dairy lactic acid bacterium *Lactococcus lactis* (Holland and Coolbear 1996), and was the first dairy bacterial esterase to be studied in any depth, although tributyrin esterase activity had been shown to exist in lactic acid bacteria much earlier (Oterholm, Ordal et al. 1968).

2.1.1 Genetic Study of EstA

After its identification and biochemical characterisation, the gene for EstA was cloned and sequenced (Fernandez, Beerthuyzen et al. 2000). This work also established a plasmid-based expression system for EstA in a homologous host cell, and produced a cell line that was a null mutant for EstA expression.

Like the other esterases examined in this study, the role of EstA in the metabolism of the organism is as yet unknown. EstA null mutant cells appear to be viable and to grow normally, indicating that the enzyme is not required for growth in rich media situations (Fernandez, Beerthuyzen et al. 2000). It was suggested that, because of sequence similarities between EstA and S-formylglutathione hydrolase, EstA may be involved in detoxification of metabolites.

2.2 Crystallographic Methods

Crystallographic methods were carried out as described in Appendix A. Procedures specific to this chapter are detailed below. It should be noted that the work described in Sections 2.2.1 and 2.2.2 were carried out by other investigators prior to this work.

2.2.1 *Lactococcus* EstA Native Data Collection (1998)

Recombinant EstA was prepared from a homologous expression system and was crystallised prior to this work (Zheng 1998), and a native dataset to 2.2 Å was collected from four of these original crystals using capillary-mounted crystals (Table 2.1), with the data collected at room temperature (G. Norris, personal communication, 2007). Despite prolonged attempts, diffracting crystals of the protein could not be obtained from subsequent purifications (G. Norris, personal communication 2004). This lack of reproducibility meant that techniques for obtaining phase information, such as multiple isomorphous replacement (MIR) or multiple anomalous dispersion (MAD), could not be used. A search for a suitable model for a molecular replacement solution was made, with Antigen 85B from *Mycobacterium tuberculosis* (Anderson, Harth et al. 2001) being identified as a potential template, with models made using SWISS-MODEL (Schwede, Kopp et al. 2003) and by making a polyalanine model based on the coordinates of Antigen 85B. This was not successful at obtaining a solution.

An alternative approach was to identify an EstA homologue, to clone, isolate and purify the protein, to crystallise it and to solve the structure to generate a model to solve the structure of EstA by molecular replacement.

2.2.2 Prior Solution of *Streptococcus pyogenes* EstA Structure (2005)

Prior to this project, a homologue of EstA was identified in the *Streptococcus pyogenes* genome (Ferretti, McShan et al. 2001) using BLAST (Altschul, Madden et al. 1997). This enzyme has a 56% identity and 71% similarity by protein sequence to *L. lactis* EstA. As well as being sequentially, and likely to be structurally similar,

genomic DNA of a reduced pathogenecity strain of *S. pyogenes* was available. It was therefore anticipated that this would by the easiest route to obtaining a suitable model for a molecular replacement solution.

The ORF encoding *S. pyogenes* EstA was cloned into a plasmid-based expression system using polymerase chain reaction (PCR)-based methods (Trevor Loo, personal communication, 2005). The recombinant protein was produced in *Escherichia coli* and purified, and diffracting crystals were obtained. The native crystallisation conditions were then used (with some optimisation) to produce selenomethionine-labelled crystals, suitable for structural solution using MAD (T. Loo, personal communication, 2005).

MAD datasets were collected at the European Synchrotron Radiation Facility (ESRF) (Ted Baker, University of Auckland). Initial phase estimates were calculated from these datasets using SOLVE/RESOLVE (Terwilliger and Berendzen 1999), and a model was produced after several rounds of rebuilding using COOT (Emsley 2004) followed by refinement using REFMAC5 (Murshudov, Vagin et al. 1997). Restrained refinement was used in all cases, with the exception of the refinement used to generate maps from MOLREP solutions, which used rigid body refinement (B. Anderson, personal communication, 2006). The completed model was provided (B. Anderson) to allow the solution of the *L. lactis* EstA from the previously native dataset obtained by previous investigators (J. Zheng).

2.2.3 Co-crystallisation of Enzyme–Inhibitor Complex

Attempts were made to crystallise EstA in complex with a phosphonate-based tributyrin analogue ((2R)-1,2-di-O-butyl-*sn*-glycerol-3-O-*p*-nitrophenyl-*n*-propylphosphonate), kindly donated by D. Lun, D.R. Harding and C.M. Taylor. Crystallisation conditions were screened using sparse matrix screens (Hampton Research Crystal Screen and Crystal Screen II and Molecular Dimensions Structure Screen I and Screen II), and several lead conditions were identified and were refined to yield crystals. When these were tested, they were found to diffract only to low resolution (4–5 Å).

2.2.4 Solution of *Lactococcus lactis* EstA Structure

Using the structure of *S. pyogenes* EstA as a model, a molecular replacement solution was calculated for *L. lactis* EstA using MOLREP (Vagin and Teplyakov 1997), with a score of 0.470 over the nearest incorrect solution which had a score of 0.427. The Matthews coefficient of the unit cell was calculated and found to be 2.61 for a 4 molecule asymmetric unit, which gave an estimated solvent content of 53%.

The solution had four molecules per asymmetric unit, as expected from the Matthews coefficient, which were labelled A–D. The solution was examined to check that there were no intermolecular clashes between molecules in the unit cell. Rigid body refinement was carried out using REFMAC (Murshudov, Vagin *et al.* 1997), which, although did not appear to refine well, with an R-factor of 47%, yielded an interpretable density map.

Initially, a single molecule was built by changing individual amino acids to reflect the *Lactococcus* EstA sequence. This molecule was used as a model to rerun MOLREP, to produce a solution with the correct amino acid sequence present in all four molecules in the asymmetric unit. In order to preserve conformational differences between the individual molecules, molecular averaging was not carried out. Instead, each molecule in the asymmetric unit was rebuilt using COOT, and labelled A, B, C and D. After several rounds of rebuilding and refinement using REFMAC (Murshudov, Vagin *et al.* 1997) the R-factor had converged, and the model was completed.

2.2.5 Unbuilt Regions

Because of disorder at the termini, Met 1 was left unbuilt on all molecules, and Ser 158 was not built in Chain B. All or some of the sidechains of residues A255 Glu, B231 His, C99 Lys, C157 Lys, C186 Arg, C193 Lys and D157 Lys were also omitted. All these sidechains were located on the surface of the molecule.

Because of gaps in the electron density, there were regions in molecule D that could not be built in the final model. Molecule D also appeared to be in a different conformation from the other three molecules in the asymmetric unit. The residues left unbuilt were 78–89, 169–178 and 186–189, all of which had no mainchain density in evidence.

2.2.6 Model Quality

The overall quality of the model was checked using PROCHECK (Laskowski, MacArthur et al. 1993) and BAVERAGE (unpublished computer program, E. Dodson, 1991). The quality of the model was deemed to be satisfactory, as the structure fell within acceptable geometric parameters, as determined by PROCHECK. The R-factor and R-free were 19.66% and 21.10% respectively, indicating that the structure had not been over-refined.

Overall, Chain B had a slightly lower B-factor than Chains A and C, and Chain D had a significantly higher B-factor. This difference between Chain D and the other molecules was probably due to conformational differences seen between the structures.

Three residues were outliers in terms of acceptable Ramachandran mainchain conformations: Gln 188 and Asn 190 from Chain C, and Asp 76 from Chain D. Chain C had difficult-to-interpret mainchain density in the region 188–190, whereas Asp 76 (Chain D) had very weak density associated with it, meaning that building and refining this residue in a lower energy conformation was not possible. It should be noted that all three residues would require only slight adjustment of their mainchain torsions to bring them back into an allowed conformation.

Data Collection Statistics					
Resolution Range	90–2.5 Å				
Space Group	$C222_1$ (# 20)				
Cell Parameters					
a, b, c (Å)	76.45	179.82	176.42		
α, β, γ (°)	90	90	90		
Unique Reflections (Total Reflections)	40870 (42703)				
Completeness (Outer Shell 2.50-2.59Å)	96.1 (86.4)				
Redundancy	2.95 (1.6)				
Mean I / σ	7.7 (3)				
$\mathbf{R_{merge}}$ *	0.11 (0.16)				
Model and Refinement Statistics					
No. Residues Total (Water Molecules)	1228 (234)				
B-factor Analysis by Molecule	**A**	**B**	**C**	**D**	**Water**
Mainchain Å^2	24.26	22.31	24.26	34.15	–
Sidechain Å^2	24.68	22.87	24.64	33.96	–
Overall Å^2	24.47	22.60	24.45	34.05	21.47
Average all atoms Å^2	26.058				
Refinement					
Reflections used	37806				
Reflections R_{free}	2003 (random selection)				
Resolution limits	90 – 2.5 Å				
RMSD bond lengths (Å)	0.005 Å				
RMSD bond angles (°)	0.001 °				
R-factor	0.197				
R-free	0.241				
Ramachandran Analysis					
Most favoured	774 (89.2%)				
Additionally allowed	83 (9.6%)				
Generously allowed	8 (0.9%)				
Disallowed	3 (0.3%)				

Table 2.1 *Data, model and refinement statistics.*

*$\left(\Sigma|I-\{I\}|/ \Sigma\{I\} \right)$ *Where I = intensity of a given reflection and {I} = mean of the intensity for multiple measurements of that reflection*

2.3 The Crystal Structure of EstA

Lactococcus lactis EstA has a typical αβ hydrolase fold (Ollis, Cheah et al. 1992), consisting of an eight-stranded β sheet, surrounded by ten α helices. Like the other αβ hydrolase structures covered in this project, it contains insertions additional to the prototype fold.

2.3.1 Excursions from Basic αβ Hydrolase Fold

Helix A, which is involved in the quaternary association of EstA monomers, is split into two, helix A_1 (residues 51–57) and helix A_2 (residues 59–64), and three additional helices are inserted after strand 6, which is the most common location for insertions in the basic αβ hydrolase fold (Ollis, Cheah et al. 1992). These are designated $D`_1$ (residues 153–158), $D`_2$ (residues 160–167) and $D`_3$ (residues 170–174), in order of primary sequence. The overall fold of EstA is shown in stereo in Figure 2.1.

2.3.2 Catalytic Triad and Active Site

Like other members of the αβ hydrolase family, EstA has a catalytic triad consisting of a serine nucleophile (Ser 121), a histidine base (His 231) and an aspartic acid (Asp 202), which are located distantly in the primary sequence but are in close proximity in the final folded enzyme, as shown in Figure 2.2.

Ser 121 is found in a tight turn at the top of strand 5, in a disallowed Ramachandran conformation that is typical for the nucleophilic residues in members of this enzyme family. The Oγ of the Ser 121 sidechain is within hydrogen-bonding distance of the εN of His 231 (2.75 Å), which in turn is within hydrogen-bonding distance of Asp 202 (2.74 Å), forming the charge stabilisation network used to activate Ser 121. Immediately adjacent to Ser 121 is the oxyanion hole, formed by mainchain amides of Met 122 and Met 46. The carbonyl oxygen of Gly 45 is hydrogen bonded to the Nδ1 of His 44 (2.81 Å), which holds the peptide bond between Gly 45 and Met 46 in an orientation that allows the Met 46 amide to point towards the oxyanion hole, reminiscent of, but not identical to, the arrangement seen in members of the hormone-

specific lipase family (De Simone, Galdiero et al. 2000). The sidechain of Met 46 appears to occupy two alternative rotamers. One rotamer brings the Sδ of Met 46 to within 3.00 Å of the Oγ of Ser 121, whereas, in the second conformation, the Met 46 sidechain has rotated away from Ser 121, increasing this distance to 6.25 Å. It is tempting to hypothesise that the conformational change of the Met 46 sidechain might be involved in substrate binding, by acting as a "gate" that becomes displaced when the substrate binds, because, when it is placed closest to Ser 121, substrate binding would almost certainly be prevented. No additional electron density that could be associated with bound substrate, acyl intermediate or product was observed in or around the catalytic site.

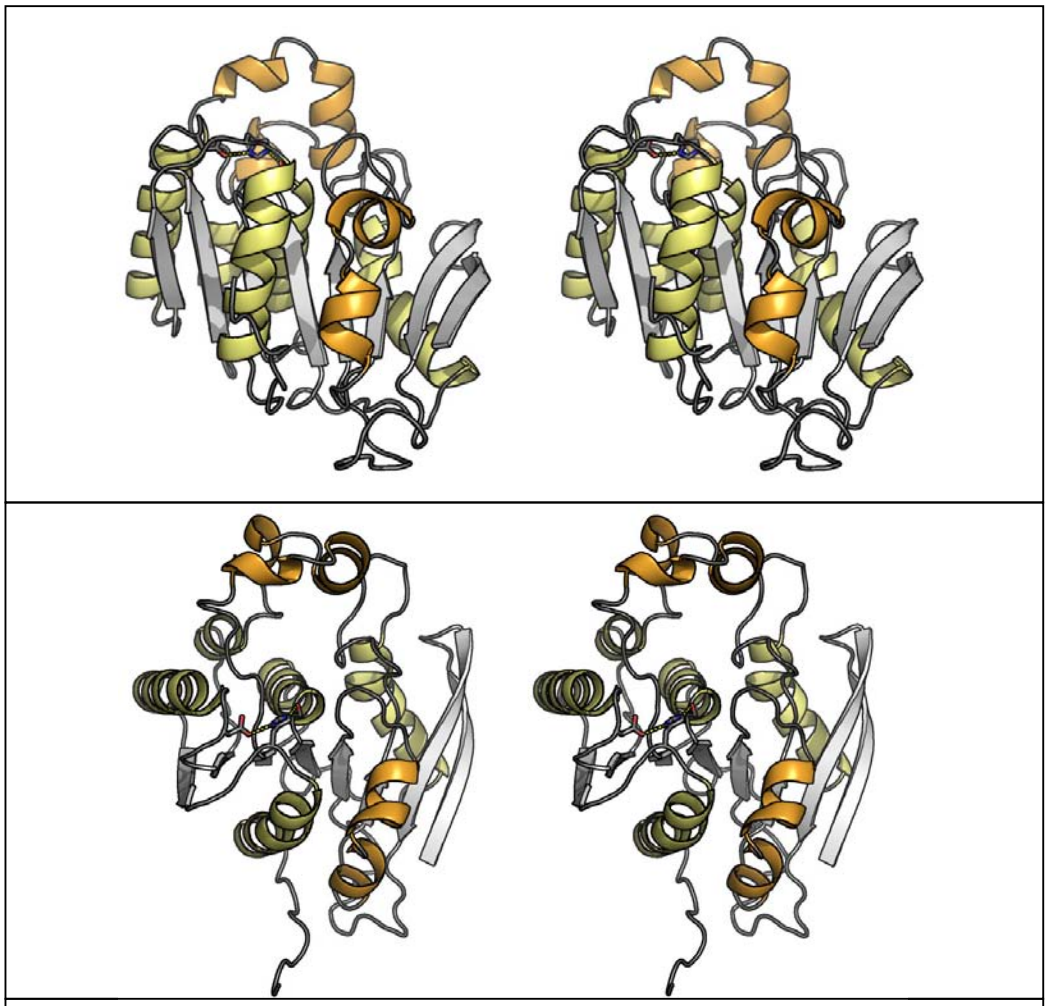


Figure 2.1 *Cartoon representation of a single EstA monomer. The β sheet runs right to left. The A_1, A_2, $D`_1$, $D`_2$ and $D`_3$ helices that are additional secondary structural elements to the prototype fold are shown in bright orange.*

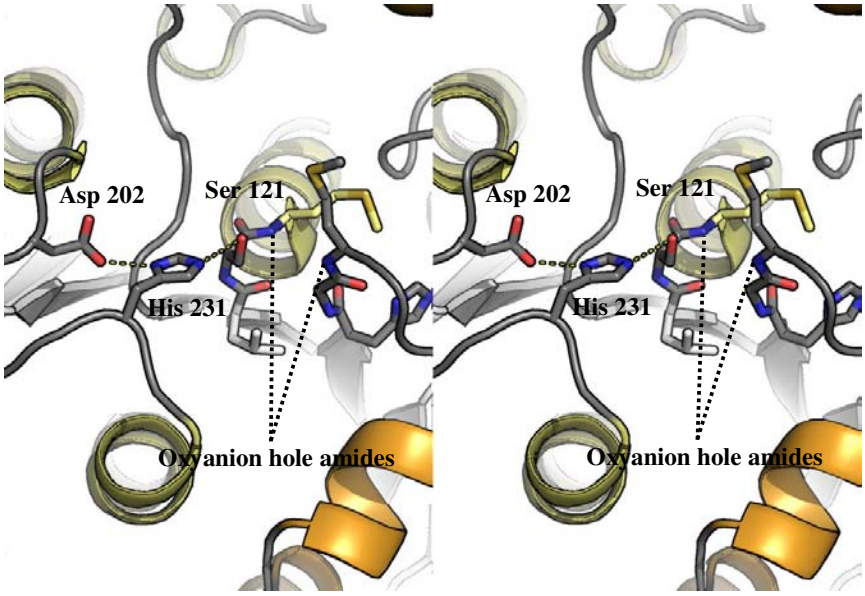


Figure 2.2 *Catalytic site of EstA showing catalytic triad and other residues involved in catalysis. Hydrogen bonds are shown by dotted lines (Asp–His = 2.74 Å; His–Ser = 2.75 Å).*

2.3.3 Quaternary Structure

EstA forms a tetramer that is composed of a dimer of dimers in crystal form. This is supported by earlier work that found, according to size exclusion chromatography, that EstA runs at a molecular weight consistent with it being a trimer or tetramer in solution (Shaw 1999). In addition, both the S. pyogenes EstA and the L. lactis EstA form very similar quarternary structures when symmetry operations are performed, despite having different unit cells and asymmetric units.

The biological tetramer differs in spatial arrangement from that of the four monomers seen in the asymmetric unit (Figure 2.3). Whereas the asymmetric unit consists of four monomers arranged in a row, the dimer of dimers is generated by crystallographic symmetry operations to give a globular arrangement of monomers. The asymmetric unit consists of monomers A, B, C and D, but interaction with another asymmetric unit (notated as A`, B`, C` and D`) occurs to generate a physiological tetramer. The symmetry operations that generate the physiological dimer of dimers produce two tetramers, one containing monomers of types A and B and one containing monomers of types C and D. In the biological dimer, molecules pair with a monomer of the

opposite type from another asymmetric unit (i.e. A pairs with B` and C pairs with D`). Because of this interaction between adjacent asymmetric units, tetramers that contain only two molecule types are formed, with A–B` pairing to A`–B, and C–D` pairing to C`–D, giving AB`BA` and CD`DC` dimer of dimer pairs, as shown in Figure 2.3. It should be noted that monomers A and B from a given asymmetric unit interact with monomers C and D from a different asymmetric unit to form a repeating interaction throughout the crystal.

The surface area of the interface between each monomer in a dimer is 2211 $Å^2$, and the surface area of the interface between the dimer units that form the tetramer is 7328 $Å^2$.

Because only two molecules of the four in the asymmetric unit are represented in each tetramer, and molecule D appears to be in a unique "closed" conformation compared with the other three monomers, one out of every two tetramers is found in an entirely "open" state (containing monomers A and B) and one is found in a half "open" and half "closed" state (containing monomers C and D), as shown in Figure 2.4.

How this situation arises in the crystal lattice is unclear. It is likely that both of the forms seen in the crystal represent a different physiological conformation of the enzyme and that these two conformations are in equilibrium in solution. The closed form may be preferentially incorporated into the position occupied by monomer D during crystallisation because of crystal contacts. There are distinct B-factor differences between the four monomers, especially molecule D, which has a higher B-factor than the three "open" molecules, and molecule B, with a slightly lower B-factor, although there are no obvious specific differences between the environment around molecules B and D and that around molecules A and C.

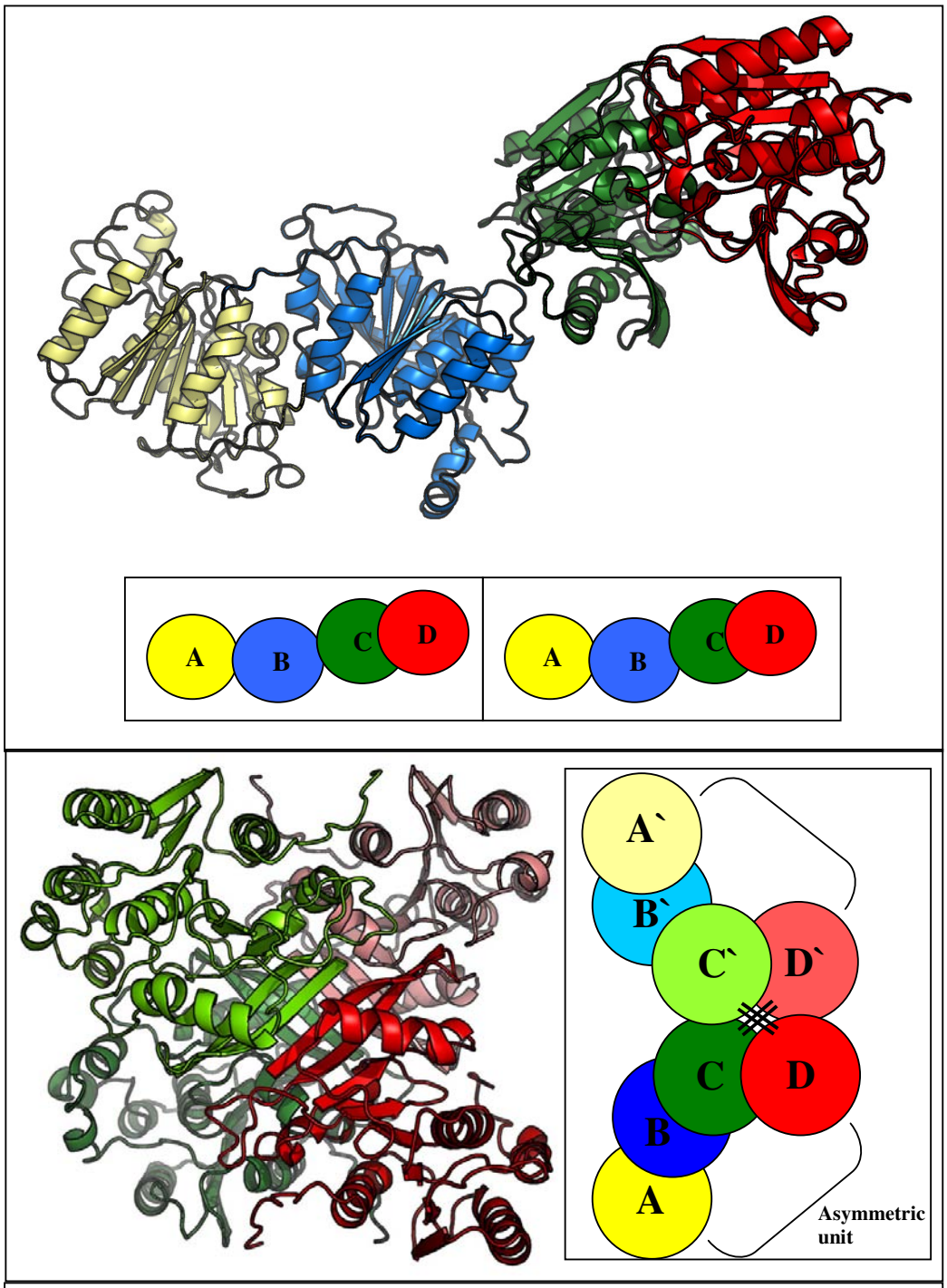


Figure 2.3 *EstA asymmetric and biological unit.*

Upper panel: Cartoon of EstA crystallographic asymmetric unit. Inset: Stereo schematic indicating monomers making up the unit. Lower panel: Cartoon of EstA biological unit, with only CDC`D` shown. Inset: Schematic indicating monomers making up the unit, and the layout of the other molecules in each asymmetric unit. Colouring is as for the top panel, with molecules from the second asymmetric unit shown in lighter tones (and/or indicated by primes). The dimerisation interface is represented by double lines.

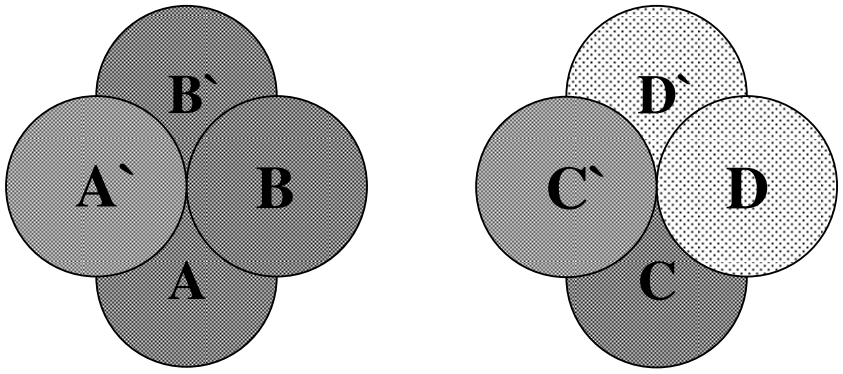


Figure 2.4 *Representative tetramers of EstA, showing fully open and half open forms, and their monomeric composition.* = *open;* = *closed.*

2.3.3.1 Dimer Interface

The dimer of monomers is formed by a molecule dimerising with a molecule from another asymmetric unit. This dimerisation occurs mainly via the central β sheet core of the fold, with β strand 1 from each monomer forming an antiparallel sheet with its partner from the other monomer (Figure 2.3, lower panel). This association means that, upon dimerisation, a continuous twisted β sheet is formed through the core of the dimer. This β sheet interaction extends from the mainchain amide of Val 3 to the mainchain carbonyl of Glu 7 of each monomer. In addition, interactions between the top of the top B helices of each monomer and the loop that spans between strands 2 and 3 occur. These involve reciprocated van der Waals' interactions between the sidechains of Phe 105 and Pro 106, and Phe 104 and Tyr 23 of each of the monomers in the dimer. Also hydrogen bonds between the mainchain amide of Tyr 9 and the Oε2 of Glu 25 (2.79 Å) and between the mainchain carbonyl of Asn 103 and the Nδ2 of Asn 107 (3.09 Å) are present.

2.3.3.2 Tetramer Interface

The tetramer is formed by one dimer associating with another dimer, through hydrophobic and charge interactions. For a given monomer, interactions with both monomers of another dimer occur. Hydrophobic interactions between C Leu 7 and C` Met 87 and C` Met 15, and between C Ile 55 and D Val 3 occur. A number of ion pairs are also evident: C Lys 18 Nζ – C` Glu 50 (2.39 Å) and C Arg 56 Nε – D Glu 60 Oε2 (2.82 Å). Because the tetramer is generated by crystallographic symmetry, all bonds and associations between monomers are reciprocated. Interestingly, all these interactions involve residues in helices A_1 and A_2, which could be the reason why helix A is split in EstA. Because it is split, a close association of the residues involved in tetramerisation is possible, whereas, if the helix was a single extended unit, as is more conventional for the fold, steric clashes would occur.

2.3.4 Conformational Changes between Open Form and Closed Form

Conformational change between the open and closed forms is seen in the $D`_1$ and $D`_2$ helices, but, because of disorder in the peptide chain, no density was seen in the closed form for residues 169–178, which contain helix $D`_3$. Interestingly, residues 148–179 contain the $D`_1$, $D`_2$ and $D`_3$ helices, which represent insertions into the basic αβ hydrolase fold.

The conformational change in this region appears to be driven mainly by a rotation and twisting movement of the $D`_2$ helix towards the active site of the enzyme. As an example of the magnitude of this motion, the α carbon of Asn 159 moves 10.23 Å, bringing the sidechain of the residue over the top of the P2 binding site.

Figure 2.5 shows the differences in Phi and Psi between residues 147- and 167, and the sum of these differences. From this plot, it can be seen that there are three regions that flex during the conformational change, allowing movement to occur. These are residues 151–152, 154–159 and 164–166. It is perhaps unsurprising that three of the six residues that have significant mainchain torsional differences are glycine, which has few limitations on its mainchain conformation. These differences in torsion angle are reflected in the displacement of the α carbon positions during the change (Figure 2.6).

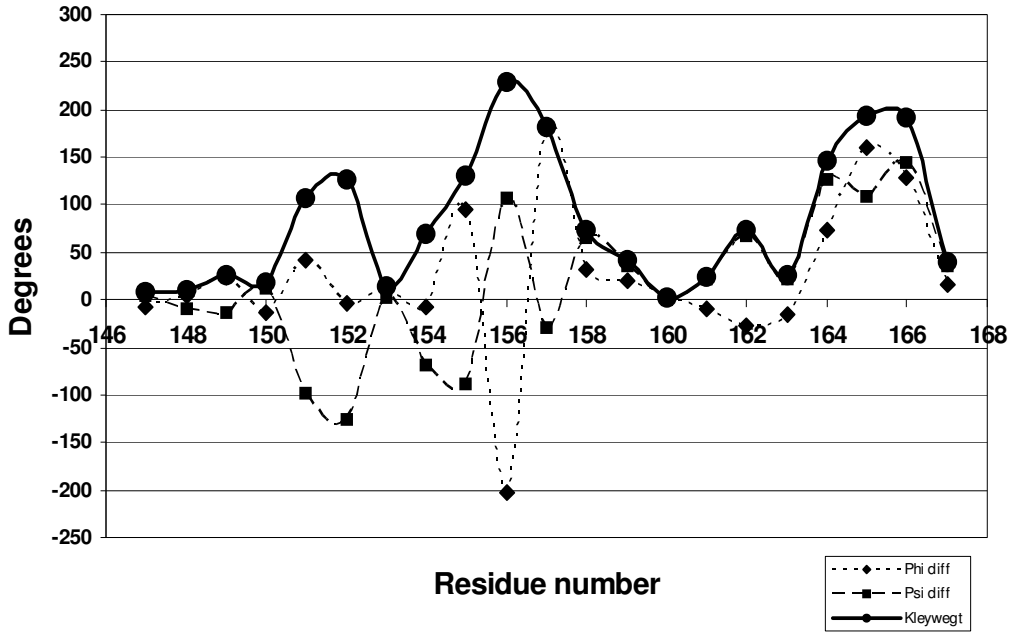


Figure 2.5 *Graphical representation of Kleywegt (Ramachandran difference) plot (Kleywegt 1996) . Phi and Psi differences were calculated using ($Phi_{closed} - Phi_{open}$) and ($Psi_{closed} - Psi_{open}$). The Kleywegt series was calculated as the hypotenuse of Phi and Psi differences using $\sqrt{((Phi_{closed} - Phi_{open})^2 + (Psi_{closed} - Psi_{open})^2)}$.*

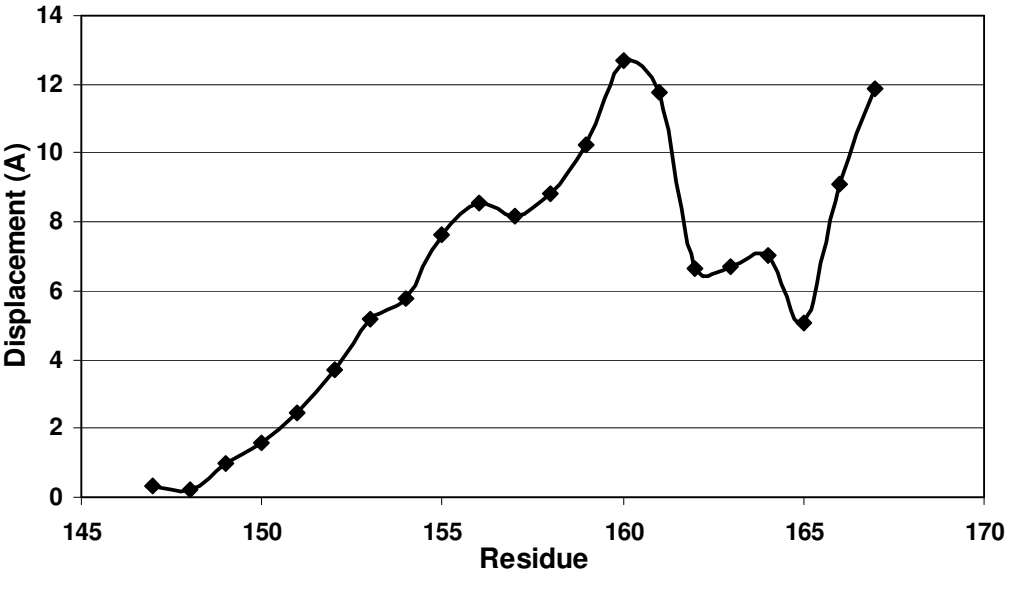


Figure 2.6 *Displacement of α carbons by residue during conformational change in the mobile loop.*

Although the residues in between these flexible regions appear to have relatively stable conformations, the interactions made by several of the sidechains change during the conformational change (Figure 2.7). On the D` insertion, residues Trp 163 and Phe 167 interact laterally. The resultant aromatic plane is capped by the sidechains Ile 166 and Tyr 162, which interact with it via van der Waals' interactions. The sidechain of Phe 173 is found underneath the plane, and also interacts with Phe 167 and with Phe 149, which interacts with Trp 163, all through van der Waals' interactions. The front of the plane interacts with Phe 79, Tyr 125 and Val 81, and, together, these form quite an extensive hydrophobic pocket.

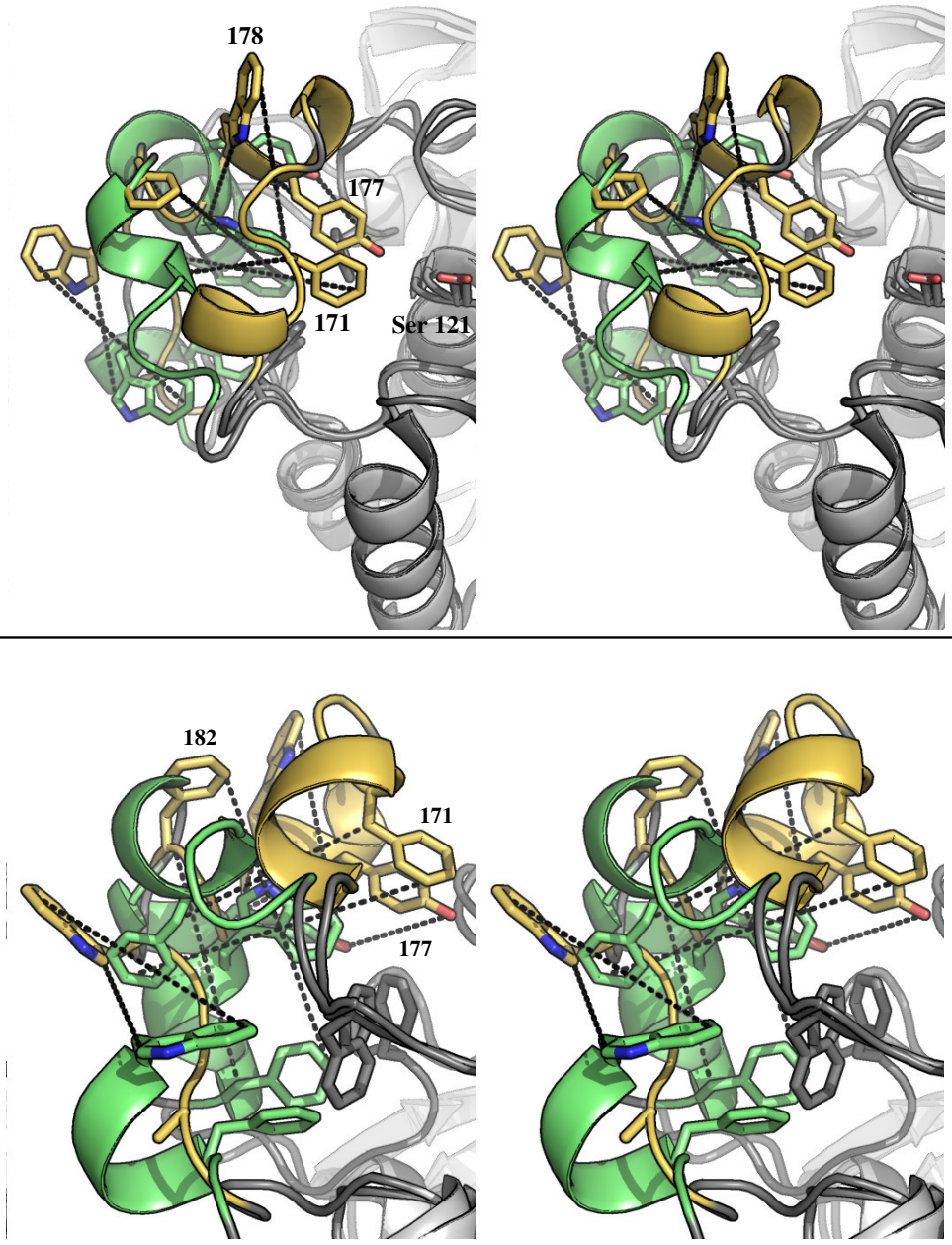


Figure 2.7 *Stereoviews of changes in the P2 pocket and cavity during conformational change from the open state (green) to the closed state (yellow). Key residue sidechains are shown as sticks, with the range of movement tracked by dotted lines. Upper panel: View into substrate-binding cleft. Lower panel: View upwards from the opposite direction, to convey how great the rearrangements of residues are in this region.*

The conformational change to the closed state causes both Phe 167 and Trp 163 to rotate away from the hydrophobic pocket, outwards towards the protein surface. During this change, the sidechain of Tyr 162, which is found stacked on top of the sidechain of Trp 163 in the open state, rotates into the same position and plane that was previously occupied by Trp 163. Phe 156 moves from a position on the outside of the D`$_1$ helix to form a π stack with Tyr 162. These two sidechains are bracketed on one side by the sidechain of Met 152, on the other side by Val 146 and Met 122 and on the top by Met 46. Phe 203 also rotates towards Phe 156, where the Cζs of each sidechain are within van der Waals' distance of each other (3.64 Å).

The open-to-closed conformational change also causes the D`$_1$ helix to partially unwind so that the helical region spans only residues 152–156 (compared with residues 152–158 in the open form). The D`$_2$ helix appears to stay intact, undergoing a rotation of ≈ 100^o as a whole. As the D`$_3$ helix (included in the unbuilt region from residues 169 to 178) is disordered in the closed state, it is not clear what kind of conformational change it undergoes. What can be seen is that residues Glu 179 and Gly 168, which are defined with good electron density, change their position to become separated by a distance of 7 Å. Such a change in position can occur only if the D`$_3$ helix unwinds to form a loop. Disorder caused by the enzyme changing from one conformation to the other may be responsible for the absence of strong electron density in this region of the closed structure.

A displacement between residues 74 and 77 in the loop formed by residues 74–91 is also seen, between strand 2 and helix A. Density for the peptide chain is not visible for residues 78–89, suggesting disorder. In the open form, this region is closely associated with the D`2 helix, making it likely that the changes observed are linked to the conformational shift of the D` helices. In the open state, the loop comprising residues 74–91 appears to be anchored by the interaction of Phe 79 and Val 81 with Trp 163 and Phe 167, as described above. Some stabilisation of the 83–91 loop also appears to be due to Tyr 85, the sidechain of which is found in a π-stacked interaction with the Tyr 85 from the other open monomer of the tetramer (i.e. in the C monomer, Tyr 85 from one dimer associates with C` Tyr 85 from the other dimer), as shown in Figure 2.8. It is assumed that, in the closed form of the enzyme, when this loop becomes disordered, this interaction is disrupted. Such a specific interaction that seems to be dependent on the enzyme being in an open state might represent a means of communication between the two dimers of each tetramer, indicating possible

cooperative substrate binding. In such a scenario, substrate binding to one monomer would cause it to adopt an open conformation. The loops that are disordered during the closed state would become ordered, moving Tyr 85 into position to interact with its counterpart on the other dimer, and, as a consequence, pulling the 74–90 loop into the conformation found in the open enzyme. This stabilisation would, in turn, induce the 147–179 loop into the alternative conformation seen in the open state, forming the substrate-binding pocket. Whether cooperative binding occurs in EstA is unclear as the appropriate kinetic work has not been carried out. It should also be noted, that, as the natural substrates for EstA are not known, it is impossible at this time to verify that the enzyme is an allosteric enzyme.

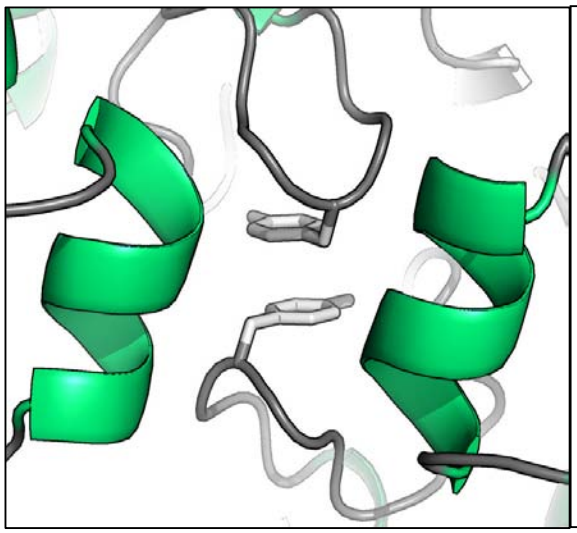


Figure 2.8 *Tyr 85 stacking interaction. This occurs between the two open state monomers that are members of different dimers in the tetramer, suggesting that it is a mechanism for interdimer communication in the tertiary structure. Both monomers are coloured the same, but the tyrosine from one monomer is at the top and the tyrosine from the other is at the bottom.*

2.3.5 Substrate-binding Cleft Changes

The conformational change in the D` insertion (residues 148–179) could be linked to substrate binding, as the change from a closed state to an open state appears to enlarge the active-site cleft, allowing the substrate access to the catalytic machinery.

In the open state, the substrate-binding cleft is a well-defined trench that is bisected by the active-site serine. Although unrelated, serine proteases such as trypsin and chymotrypsin have a similar substrate binding trench and catalytic machinery arrangement. It is common practice to divide the serine protease substrate binding trench into to halves, denoted P1 and P2, and this convention will be used where appropriate for the esterases reported here. Although the exact orientation of the substrate binding site relative to the catalytic machinery does vary slightly between the different esterases, in broad terms the P1 pocket shall be defined as running in the

direction pointed to by the Nε1 of the catalytic histidine, while the P2 pocket runs in the opposite direction.

In EstA therefore the P1 pocket extends over the top of β strand 4, towards the centre of the EstA tetramer, whereas the P2 pocket runs in the opposite direction, towards the outside of the monomer (Figure 2.9). The residues involved in forming the binding pockets are summarised in Table 2.2.

P1 Binding Pocket Residues		**P2 Binding Pocket Residues**	
D Val 3	C Ser 47	C Met 46	C Asn 155
D Glu 60	C Ser 52	C Phe 79	C Met 152
D Lys 28	C Trp 53	C Tyr 162	C Phe 203
C` Glu 7	C Trp 233	C Trp 163	
C` Asn 16	C Trp 236	C Phe 149	
C Ile 55	C Leu 120	C Val 146	
C Arg 56		C Leu 204	

Table 2.2 *Residues involved in forming the substrate-binding cleft in monomer C of EstA. The residues are indicated by chain number, type and primary sequence.*

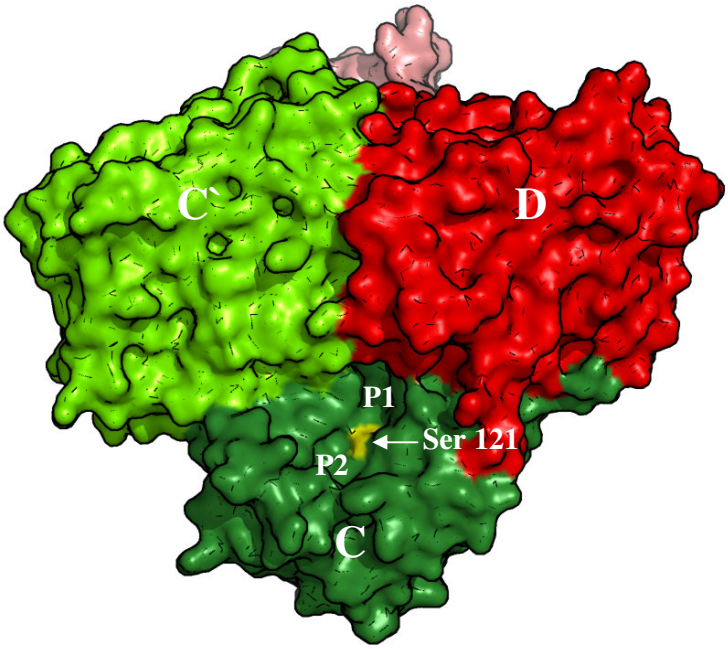


Figure 2.9 *The surface of an EstA tetramer, made up of C and D subunits. The catalytic serine (Ser 121) is shown in yellow. The monomers are labelled, except monomer D` which is just visible at the top. Pockets P1 and P2 are indicated.*

2.3.5.1 The P1 Binding Pocket

The P1 pocket (Figure 2.10, upper panel) is narrow, and enclosed, with its base formed by the monomer of which the pocket is part. However, the pocket extends into the tetramerisation interface, with its top being formed from both monomers of the other pair in the tetramer. For example, the top of the P1 pocket of Chain C is formed by residues from D and C` of the tetramer. The P1 pocket itself is lined by residues with both hydrophobic and polar sidechains, and, excluding the Met 46 mainchain amide that is involved in oxyanion hole stabilisation, no mainchain atoms are involved in its formation. There are also no hydrogen bonds or ion pairs between any of the sidechains lining the pocket, resulting in a plethora of possible interactions between enzyme and substrate. Because the sidechains are not locked into position, the substrate pocket also provides a reasonably fluid environment, in which sidechains can move to promote more favourable interactions with incoming substrate molecules. The size and the shape of the P1 pocket place certain steric constraints on the substrates that can enter it. A single saturated acyl chain would appear to be more likely to bind in the P1 site, as it is not wide enough to accommodate multiple or unsaturated chains. There is also a lack of suitable hydrophobic binding regions running up the sides of the D and C` monomers for an outward projecting glyceride *sn*-2 chain to bind. The shape and the properties of the pocket would thus suggest that the scissile acyl chain, rather than the glyceride backbone of the substrate, binds in it.

2.3.5.2 The P2 Binding Pocket

The P2 binding pocket, located opposite the P1 pocket (relative to the catalytic serine), is more open, and is shaped like a basin, lined with aromatic and hydrophobic sidechains (Figure 2.10, lower panel). Like the P1 site, potential substrate–enzyme interactions in the P2 pocket would probably involve the sidechain, rather than the mainchain, atoms. The open structure of the P2 binding site suggests that it is more likely than the narrow P1 site to be involved in binding the glyceride backbone and associated acyl chains and is characterised by the large number of aromatic sidechains. Such a concentration of aromatic sidechains in a binding site is often associated with the binding of carbohydrates (Vyas 1991).

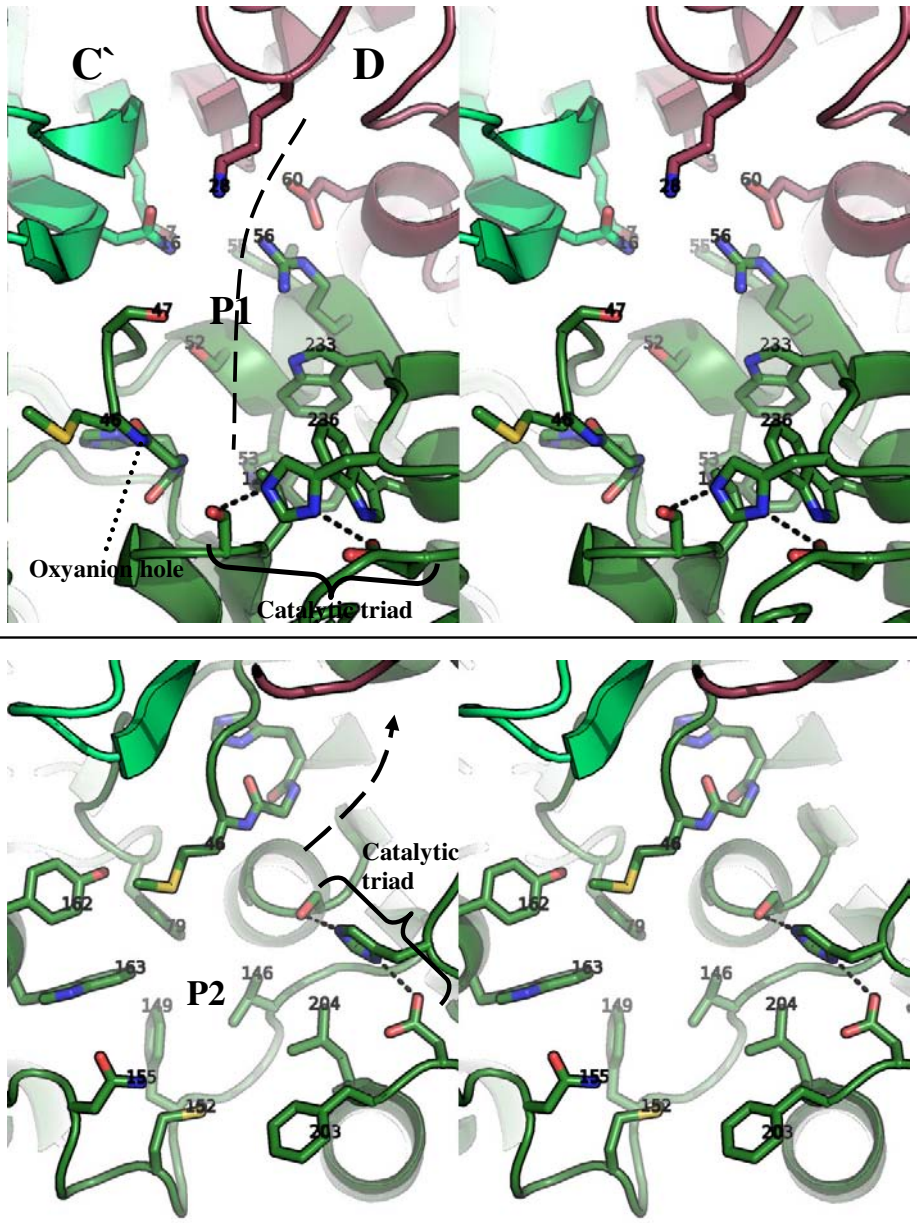


Figure 2.10 *Stereoviews of the P1 and P2 binding pockets. Upper panel: P1. Lower panel: P2. Residues that make up the pocket are shown as sticks and numbered according to primary sequence, with the protein backbone shown in cartoon form. Residues shown at the top belong to the D monomer (red) and the C` monomer (lime). All other residues belong to the C monomer (green). The catalytic triad residues and oxyanion hole amides are indicated, with inferred hydrogen bonding between the residues shown. The dashed curve shows the approximate position of the base of the P1 cleft.*

In the open form, an additional cavity at the end of the P2 pocket away from Ser 121 is formed by the sidechains of phenylalanine residues 79, 149, 156, 167 and 173 and tryptophan residues 163 and 170, and faces towards the catalytic site. Although relatively small compared with the P1 and P2 binding sites, this pocket might be able to hold other molecules important to catalysis. For example, this cavity might be a binding site for an acceptor molecule of a transferase reaction. The inside of the cavity is composed almost exclusively of aromatic sidechains. As discussed earlier, interactions between oligosaccharides and aromatic residues appear to be common in carbohydrate binding, and it is interesting to note that Phe 149, 167, 173 and Trp 163 have faces in this cavity that could be involved in such interactions with an oligosaccharide or other acceptor molecule. It is equally possible that this highly hydrophobic cavity could be involved in binding some other hydrophobic substrate.

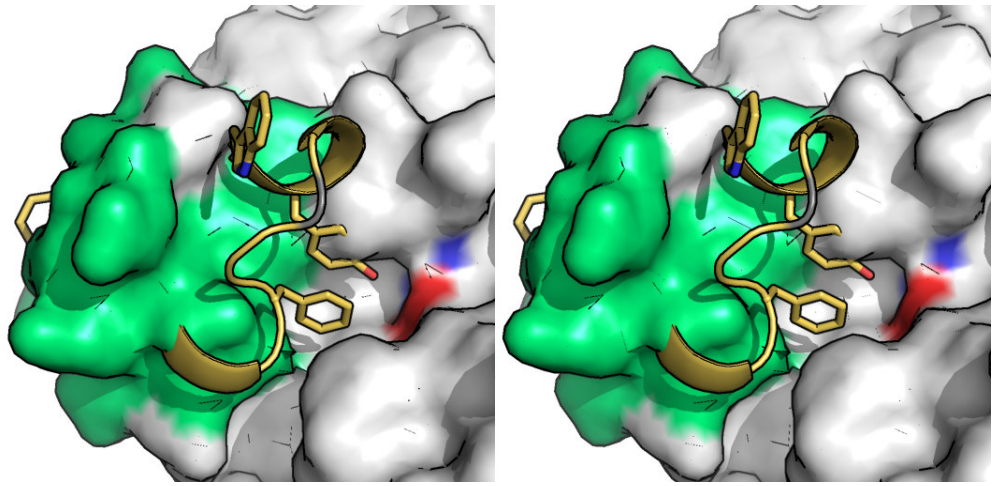


Figure 2.11 *Stereo diagram of changes to the surface of the P2 pocket during the conformational change from open (green) to closed (yellow). Key residues are shown as sticks, notably Tyr 162 and Phe 156 at the centre. The oxygen of Ser 121 (red) and the oxyanion hole amides (blue) are shown to indicate their positions on the surface of the enzyme.*

2.3.5.3 P2 Cavity Conformational Change

The P2 cavity becomes closed during the conformational change from the open to the closed state, as shown in Figure 2.11. The movement would result in any acceptor or product being physically pushed out of the cavity and towards the catalytic machinery, where it would be primed to undergo the second nucleophilic attack of the catalytic cycle, to yield a transferase product. The slight bias of the catalytic histidine towards the P2 pocket also supports the hypothesis that the donor and acceptor substrates are bound in this cavity during catalysis. A proposed mechanism for this sequence of events is outlined in Section 3.3.8.

2.3.6 Conformational Change in Catalytic Triad

In the closed state, some changes in the positioning of certain sidechains in and around the catalytic triad that may be linked to the catalytic mechanism are observed (Figure 2.12). During the change from the open form to the closed form of the enzyme, the sidechain of Ser 121 appears to change rotamer from –177.1^{o} to –58.6^{o}, bringing the Oγ to within hydrogen-bonding distance of the Nε of His 231 (2.89 Å from 3.30 Å). This change, although interesting, needs to be confirmed in a higher resolution structure before being accepted as real. The positions of the other two catalytic triad residues (His 231 and Asp 202) appear to remain fixed during the change from the open state to the closed state. The sidechain of Met 122, the residue after the nucleophile, moves towards the serine. Although the reason for this movement is not obvious, it might result in slightly altering the orientation of the mainchain amide of this residue, which, as it occurs at the top of the helix of the nucleophile elbow, is important in stabilising the oxyanion of the transition state during catalysis (Hakansson 2002). The sidechain of Met 46, which follows Gly 45 (the amide of which forms part of the oxyanion hole), is found in two conformations. In one of these conformations, the sidechain is close to Ser 121; in the other conformation, it has rotated almost 90^{o} away from it. The difference between these alternative conformations is more exaggerated in the closed state of the enzyme, as, during the change from the open state to the closed state, Tyr 162 is able to move closer to the active site, as the Met 46 sidechain moves away. This movement appears to perturb the direction of the oxyanion hole amide preceding it, which appears to become angled more downwards, away from an optimal position for oxyanion

stabilisation. Whether this movement is significant or not is difficult to ascertain because of the limited resolution of the data.

The most significant conformational change seen is the movement of Tyr 162 into the active site during the change from the open state to the closed state, where it is positioned in such a way that the phenolic oxygen points towards the Oγ of the catalytic serine. Its position has similarities to those of tyrosines seen in prolyl oligopeptidase (Fulop, Bocskei et al. 1998), *Rhodococcus* cocaine esterase (Larsen, Turner et al. 2002) and *Lactococcus lactis* PepX (Rigolet, Mechin et al. 2002), which have been shown to be involved in the oxyanion hole (Szeltner, Rea et al. 2002; Turner, Larsen et al. 2002). Sequence alignment, using CLUSTALX (Thompson, Gibson et al. 1997), with other enzymes that, on the basis of sequence similarity, are likely to be EstA homologues shows that Tyr 162 and the tryptophan (163) that follows it are highly conserved.

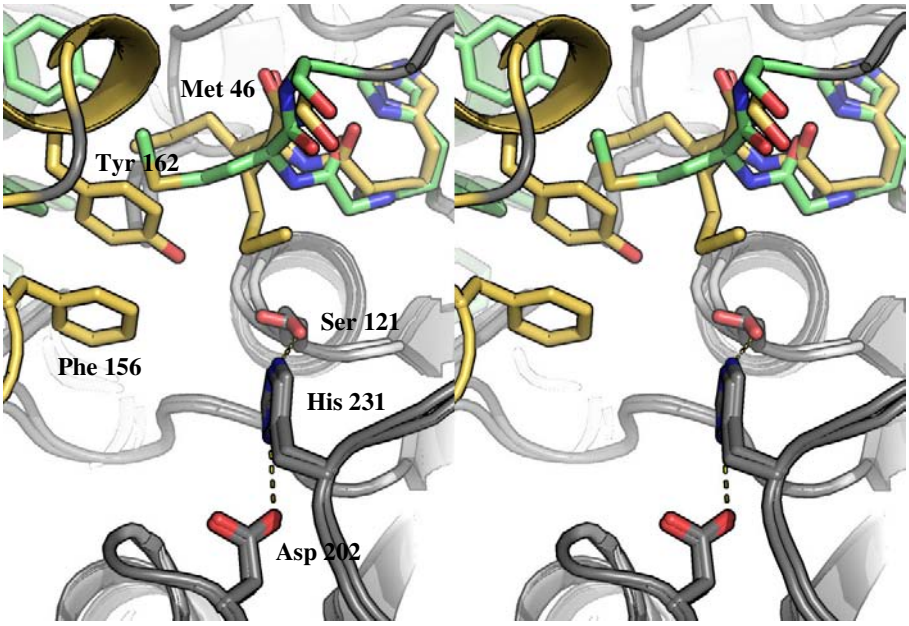


Figure 2.12 *Stereoview of conformational changes around the catalytic site during the open-to-closed state change. Open state residues are shown in green and closed state residues are shown in yellow, with residues that do not undergo significant change shown in grey regardless of state. Although the catalytic serine is shown in two rotamers, this may represent limitations on fitting the residue because of the resolution of the dataset.*

2.3.7 Contrast with Other Enzymes

In cocaine esterase, the tyrosine comes from a loop that usually projects a mainchain amide into the oxyanion hole in more prototypical αβ hydrolases. In cocaine esterase and its homologues, this mainchain amide interaction is absent, and the tyrosine is positioned so that its sidechain OH can stabilise the oxyanion. In contrast, in EstA, which has a permanent fully formed typical αβ hydrolase oxyanion hole, the tyrosine comes from the mobile D` region, and the tyrosine OH is found in the catalytic site only in the closed form of the enzyme. These differences are illustrated in Figure 2.13. The architecture of the catalytic site in cocaine esterase and PepX means that the tetrahedral intermediate formed during catalysis by these enzymes must be situated on the opposite side of the catalytic serine to where it is formed in other αβ hydrolases. As a result, the scissile bond must be bound in a position close enough to the catalytic machinery for this to be possible, effectively making the orientation of the substrate for these enzymes the opposite of that seen in the typical αβ hydrolases, where the scissile bond is positioned over the top of the nucleophile elbow's helix, to allow stabilisation of the oxyanion transition state by association with the helix dipole (Hakansson 2002). In cocaine-esterase-like enzymes, the scissile bond must be positioned in the void to one side of the nucleophile elbow's helix, to allow stabilisation to be performed by the tyrosine. However, such positioning of the substrate would mean that the dipole of the helix would be unable to be so strongly involved in stabilising the oxyanion. Why EstA has both a traditional oxyanion hole and a tyrosine sidechain OH present in the active site, which seems to be a combination of both a typical αβ hydrolase architecture and a cocaine esterase oxyanion architecture, may be reflected in its activity. The presence of both oxyanion-stabilising features suggests that a hybrid mechanism might exist in this enzyme, perhaps making it a more efficient transferase.

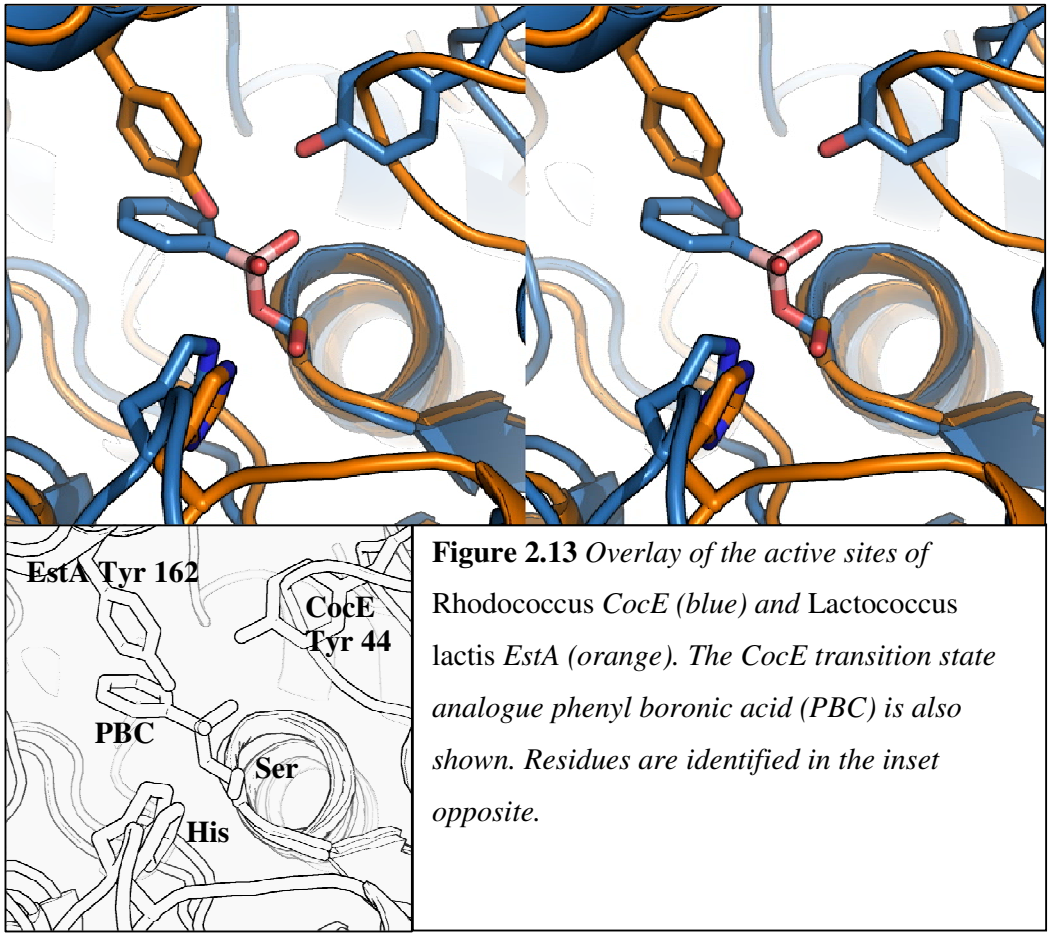


Figure 2.13 *Overlay of the active sites of* Rhodococcus *CocE (blue) and* Lactococcus lactis *EstA (orange). The CocE transition state analogue phenyl boronic acid (PBC) is also shown. Residues are identified in the inset opposite.*

2.3.8 Proposed Transferase Mechanism

In the mechanism proposed for the enzyme (Figure 2.14), the formation of an acyl-enzyme intermediate would proceed by the same mechanism that has been proposed for other αβ hydrolases. The catalytic serine (Ser 121 in EstA) is primed to carry out a nucleophilic attack on the scissile bond of the triacylglyceride substrate by the transfer of its proton on to His 231, with the oxyanion transition state being stabilised by the dipole of the nucleophile elbow helix and the mainchain amide of Met 46. This would occur with the enzyme in the "open" state. The first reaction product, a partially substituted glyceride, would be free to leave the enzyme at this point, as it would no longer be covalently bound to carboxylate of the substrate. At some point during catalysis, an acceptor molecule (in the case of EstA, ethanol is one such acceptor) would bind in the P2 cavity of the enzyme. This might occur before the initial nucleophilic attack or after the formation of the acyl-enzyme intermediate. The enzyme would then shift to the "closed" conformation, completing dissociation of the

first product, closing the P2 cavity and forcing the acceptor towards the catalytic site. This would bring the acceptor within hydrogen-bonding distance of His 231, allowing it to become activated to attack the acyl-enzyme intermediate. This second transition state would be stabilised by the sidechain OH of Tyr 162, using a mechanism similar to that proposed for the stabilisation of both the tetrahedral transition states (acylation and deacylation) of prolyl oligopeptidase and cocaine esterase (Larsen, Turner et al. 2002; Szeltner, Rea et al. 2002). The collapse of the oxyanion transition state back to a trigonal state would cause the breakage of the acyl intermediate–Ser 121 bond, liberating the product from the enzyme.

The presence of this second oxyanion stabilisation site might be essential in EstA, because the catalytic histidine is located closer to the P2 binding site than to the oxyanion hole, and because the close proximity of the other molecules making up the tetramer prevents the access of an acceptor from the direction of the P1 pocket. It is interesting to note that *Mycobacterium tuberculosis* mycolyl transferase Antigen 85B, the closest sequence homologue to EstA whose structure has been reported (Anderson, Harth et al. 2001), has the catalytic histidine located in a similar position relative to the orientation of the substrate-binding cleft, but has no oxyanion-stabilising tyrosine sidechain, as seen for either EstA or PepX. This shows that the presence of the tyrosine sidechain is not necessary for the enzyme to have a transferase activity. It might be that the orientation of the catalytic histidine relative to the substrate-binding sites is more important in allowing access and subsequent activation of larger acceptors than water during transferase reactions. It should also be noted that the acceptors used in the transferase activities of EstA and Antigen 85B are very different. The cell-wall sugar components utilised by Antigen 85B are much larger than ethanol, the only characterised acceptor known so far for EstA. Sugars would be much better than small straight chain alcohols at shielding the catalytic site from water molecules, which would otherwise be activated to complete deacylation of the enzyme, producing a carboxylate product rather than an acyl ester. If a similar mechanism for positioning sugar molecules during a transferase reaction exists in EstA, it might also be serendipitously utilised for alcoholysis by ethanol in the reactions that have been characterised for EstA.

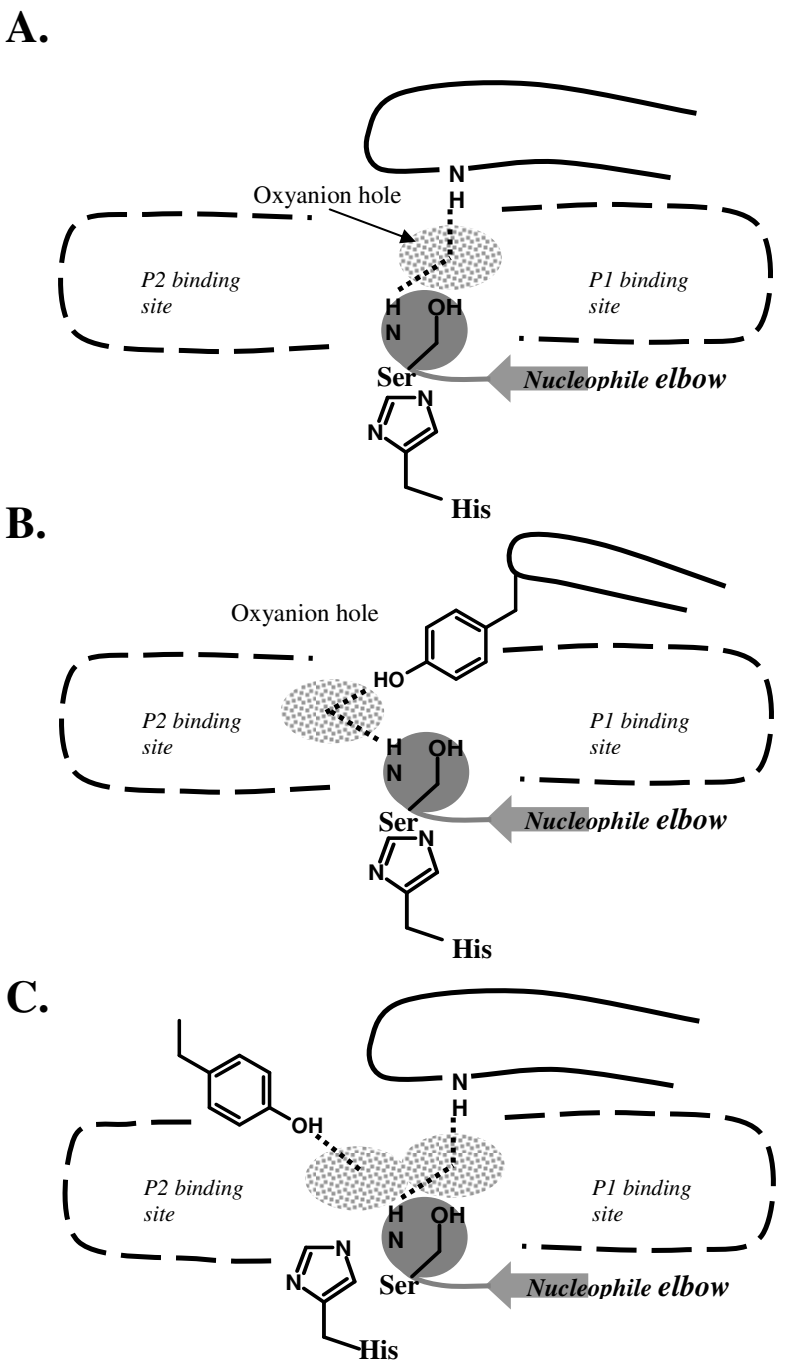


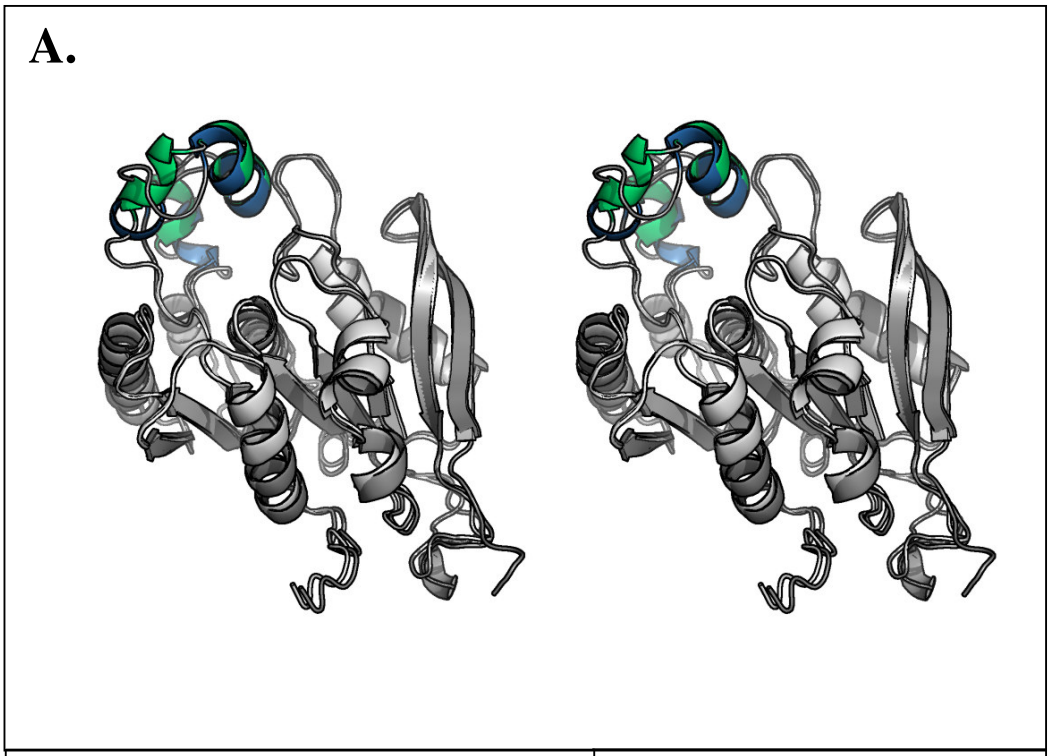
Figure 2.14 *Schematic diagram comparing αβ hydrolase fold active sites. A: Typical site. B: Cocaine esterase. C: EstA. The helix of the nucleophile elbow is represented as a filled circle and hydrogen bonds are indicated by dotted lines.*

2.4 Comparison with *Streptococcus pyogenes* EstA

Because the solution of *L. lactis* EstA required the solution of *S. pyogenes* EstA, structures of both enzymes are available for comparison.

TOP was used to calculate the RMS deviation between the structures (RMSD = 0.9Å). An SSM superpose overlay of both structures shows that the similarity between them is high (Figure 2.15). The only significant difference is seen in the D` helix region. The D`$_1$ helix of PyoEstA appears to be tilted slightly relative to the equivalent helix from *Lactococcus* EstA. There is a two-residue insertion in the short loop that connects this helix to the D`$_2$ helix, making this loop extend slightly more over the P2 cleft than in the *Lactococcus* enzyme. A further short insertion in the loop between the D`$_2$ and D`$_3$ helices appears to displace the D`$_3$ helix slightly in PyoEstA when compared with *Lactococcus* EstA.

The differences seen in this area of the protein, which forms part of the substrate-binding cleft, may provide an insight into the identity of the enzyme's natural substrate, which is still unknown. Likewise, the similarities in this region may indicate which residues are essential for substrate binding and which residues are superfluous.



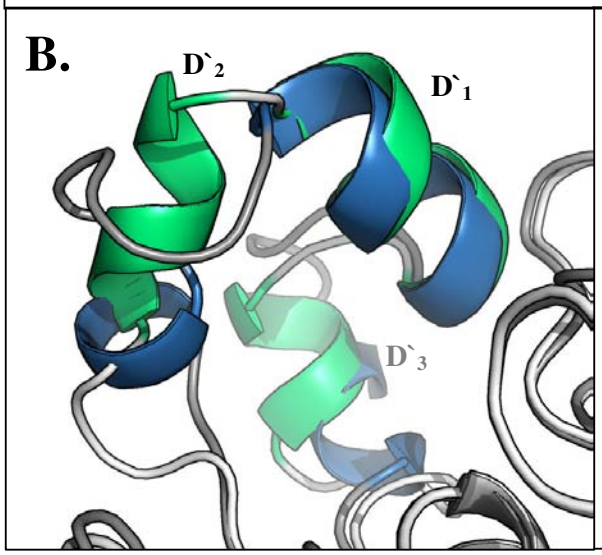


Figure 2.15 *SSM superposition of EstA from* Lactococcus lactis *and EstA from* Streptococcus pyogenes. *The enzyme core is shown in grey, with the D` helices highlighted in colour (*Lactococcus *green*, Streptococcus *blue). Panel A: Stereoview of overall monomer. Panel B: Close-up of D` helices.*

2.5 Further Work

2.5.1 Crystallisation

A notable problem associated with EstA has been the establishment of reproducible conditions for crystallisation. Despite several attempts at crystallising protein from native and recombinant sources, and with an inhibitor present, no reliable method for producing diffracting crystals has been established. The solution of the EstA structure, albeit by a roundabout method, has provided an insight into how further crystallisation studies might be undertaken. The most notable observation is that the N-terminus of the enzyme is involved in a secondary structural element (strand 1), which is needed for tetramerisation. The N-terminus also projects directly at the substrate-binding cleft of the adjacent monomer. This potential for steric interference between monomers might explain why N-terminally tagged expression constructs could not be crystallised in a form that was suitable for diffraction studies (although this was done successfully with hexahistidine-tagged *S. pyogenes* EstA). Therefore, a C-terminally tagged construct of EstA might be more amenable to crystallisation.

2.5.2 Mutagenesis

The presence of Tyr 162 in a location that might be involved in oxyanion stabilisation makes it an ideal target for site-directed mutagenesis. Similar experiments have been performed for the oxyanion hole tyrosines in both prolyl oligopeptidase (Szeltner, Rea et al. 2002) and cocaine esterase (Turner, Larsen et al. 2002), demonstrating the requirement for this residue in catalysis. Producing a Y162F mutant of EstA would allow the influence of this residue on the kinetics of the enzyme to be ascertained. As mentioned earlier, with EstA apparently possessing features associated with two oxyanion holes, mutagenesis of this residue might affect only the transferase mode of activity, leaving the enzyme able to hydrolyse substrate by using the Met 46 amide/nucleophile helix dipole to stabilise the oxyanion.

Mutagenesis aimed at altering the size, shape and chemical properties of the substrate-binding site could also be undertaken, although, without a structure of an enzyme–inhibitor complex, designing such mutations might be difficult. Care would also be

required so as not to disturb the tetramerisation interface, which might result in complete loss of activity.

2.6 Summary

A crystal structure of *Lacotoccus lactis* subsp. *lactis* EstA was solved by molecular replacement, using the structure of EstA from *Streptococcus pyogenes* as a model. The overall fold of EstA was found, as expected from sequence analysis, to be an αβ hydrolase. The anticipated Ser–Asp–His catalytic triad was found, bisecting a relatively deep cleft, producing two binding pockets defined as P1 and P2, with P1 being identified as being the most likely binding site for the acyl chain of the scissile bond.

The quaternary structure of EstA consists of a dimer of dimers, each dimer forming an extended twisted β sheet linked through a β strand from each monomer, making an antiparallel strand with the same strand in the opposite monomer. Each dimer then joins another dimer to form a type of separated β sandwich. However, the residues from the sheets of each dimer do not directly interact with those of the opposite dimer. The nature of this association results in the top of the P1 substrate-binding pocket for any monomer being made up partially from residues of the opposite dimer pair, showing why the enzyme is a tetramer.

EstA monomers were found in two distinct conformations, “open” and “closed”. When assembled into a biological unit, tetramers are composed either of fully open monomers or of two open and two closed monomers, one of each per dimer.

The conformational change between open and closed occurs only in the D` region of each monomer, which is made up of three short helices that are inserted into the canonical αβ hydrolase fold. It causes the P2 pocket to close up, and would appear to force any compound bound in the P2 pocket towards the catalytic site. It also brings Tyr 162 towards the catalytic site, where it is postulated to have a role in stabilising

the second oxyanion transition state of catalysis that occurs during deacylation. This hypothesis could be supported by site-directed mutagenesis on the enzyme to generate a Y162F mutant that would lack this hydroxyl, but, to gain structural support for this mutation, a reliable crystallisation strategy must also be devised, possibly involving the use of a C-terminally tagged fusion construct that would not cause the steric interference that appears to cause crystallisation difficulties in N-terminally tagged construct.

Chapter 3

The Structure of AA7 Esterase from *Lactobacillus rhamnosus*

Chapter 3

The Structure of AA7 Esterase from *Lactobacillus rhamnosus*

3.1 Introduction

AA7 is an esterase from *Lactobacillus rhamnosus*, a lactic acid species that occurs adventitiously in the non starter flora of cheese and has been developed as a flavouring adjunct in cheese manufacture (Crow, Curry *et al.* 2002). It has been biochemically characterised to a limited extent, and its preference for the chain length of acylglyceride substrates has been determined (Holland, Liu *et al.* 2005). Like the other esterases examined in this project, no physiological role for AA7 has been established; it does not appear to fit in any known biochemical pathway.

3.1.1 AA7 Sequence Analysis

Sequence analysis of AA7 indicated that it was a member of the αβ hydrolase family. Its sequence contains the G-X-S-X-G serine nucleophile motif, as well as conserved aspartic acid and histidine residues of the catalytic triad when aligned with other enzymes of similar sequence. According to the most recent classification of bacterial carboxylesterases, AA7 is a member of Family IV, which contains bacterial lipases/esterases that have sequence similarity to the mammalian hormone sensitive lipases (Arpigny and Jaeger 1999). AA7 shares 90% protein residue identity with EstB from *Lactobacillus casei* LILA, which has been biochemically characterised (Fenster, Parkin *et al.* 2003).

3.1.2 AA7 Activities

As well as being capable of ester hydrolysis, AA7 can synthesise esters using a transferase mechanism (Delabre, Ng *et al.* 2006). Whereas hydrolysis in acid–base–nucleophile enzymes such as chymotrypsin is well understood, the mechanisms

employed by transferases are somewhat less understood. Two mechanisms are proposed for transferase reactions in hydrolase-like enzymes (Milkowski and Strack 2004). In the first, and most accepted, mechanism, there is an initial nucleophilic attack on the substrate by the enzyme nucleophile (usually a serine), to form an acyl-enzyme intermediate, which is then subjected to a second attack by a nucleophile generated by the enzyme in the acceptor molecule. This mechanism, proposed for some αβ hydrolases, such as *Mycobacterium tuberculosis* Antigen 85B (Anderson, Harth *et al.* 2001), relies on the nucleophile activation site being somehow shielded from the bulk solvent.

The second possible mechanism involves the enzyme directly activating the acceptor as a nucleophile, with the reaction proceeding *via* a *gem*-diol intermediate, not an acyl-enzyme intermediate (Fleming, Robertson *et al.* 2000). These authors proposed a mechanism whereby the catalytic serine (in concert with the other triad members) acts as a base, removing a proton from a water molecule or an alternative acceptor to produce a hydroxyl, a powerful nucleophile and base. This mechanism requires only a slight repositioning of the serine away from the scissile bond, to allow the nucleophile to sit between it and the substrate.

A possible third mechanism could involve the direct activation of the acceptor molecule by the base and the acid of the catalytic triad, effectively skipping the first nucleophilic attack on the substrate by the catalytic serine. Because residues with potentially nucleophilic sidechains are conserved at the top of the nucleophile elbow in all members of the αβ hydrolase family, such a mechanism lacks credibility as there would be no pressure to conserve the nucleophile, and its loss in at least some members of the family would be expected.

It is currently presumed that AA7 uses the first mechanism described, as, in the current project, a tetrahedral intermediate analogue was observed to be covalently bound to the Ser 151 sidechain, indicating that the serine is induced to become a nucleophile.

3.1.3 Mutant AA7

Initial experiments carried out with AA7 were inadvertently performed on a polymerase chain reaction (PCR)-induced double-point mutant A223V/T179N. Although this error was subsequently rectified, diffracting crystals of the A223V/T179N mutant had already been obtained; thus all work was done on the mutant protein. Once the structure had been solved, the mutations were mapped to the surface of the enzyme; as the mutations do not appear to be within any significant distance of the active site, they are unlikely to have any effect on the activity of the enzyme. This was confirmed by some early experiments with synthetic substrates, which showed no difference in rates of hydrolysis between the mutant and WT enzyme.

3.2 Enzyme Inhibitors

To determine how an enzyme binds a substrate, the use of inhibitors that chemically resemble the substrate to form inhibitor–enzyme complexes has become commonplace for structural studies. These substrate analogues can take many forms depending on the chemistry of the enzymatic reaction. One form of substrate analogue, known as a “suicide inhibitor”, binds to the active site of the enzyme, where it is partially processed by the enzyme to form a stable intermediate complex that resembles the Michaelis enzyme–substrate complex. In these inhibitors, the scissile bond is generally replaced by a chemical group that resembles a transition state of the substrate.

3.2.1 Phosphonate Inhibitors

In the case of enzymes that carry out a nucleophilic attack on a carbonyl carbon, such as the serine proteases and αβ hydrolases, the scissile bond can be replaced with a phosphate group, to form an organophosphate compound. This phosphate group is often coupled with a good leaving group that is released by the enzyme during the formation of the tetrahedral state, resulting in the formation of a stable enzyme–inhibitor complex that not only is resistant to further hydrolysis but also resembles the tetrahedral transition state of the reaction. Such suicide inhibitors can take various forms in relation to the substrate they are intended to mimic; they can be small compounds containing only a limited number of additional chemical groups, or they can be larger inhibitors that closely resemble the enzyme substrate. Small molecule inhibitors, consisting of just a phosphate centre with a few attached specificity groups, can be useful, particularly where the substrate specificity of a particular enzyme is unknown. For example, in some cases, a larger inhibitor may have the inappropriate chirality to fit into the enzyme's substrate-binding pocket, and, if it does not bind, could give a false impression of the enzyme's substrate preferences. Small molecule inhibitors, such as paraoxon, have been used to determine the identity of active residues, as they will bind covalently to the nucleophile, producing extra density adjacent to the active residue in electron density maps. However, disadvantages from the use of small inhibitors can arise, because it is difficult to determine binding mechanisms. For example, a small inhibitor might not induce all conformational changes associated with the binding of a physiological substrate, if it is not big enough to make the appropriate interactions.

Larger inhibitors with structures that more closely resemble a known substrate are therefore often preferred, as they are more likely to mimic what really happens when a substrate binds, and hence provide a more complex model of the catalytic mechanism. These inhibitors will generally bind in a location that is similar (if not identical) to the substrate itself, providing insight into the groups that are required for substrate recognition and binding on both the enzyme and the substrate. However, in some cases, the use of "small molecule" inhibitors can still give useful information about how an enzyme might bind larger substrates. An example of this is seen in the structure of *Candida rugosa* lipase complexed with a monoacylglyceride mimetic. In this special case, two molecules of the inhibitor bound, one to the serine and one to

the histidine of the catalytic triad. The authors noted that, although this was not physiologically relevant, the position that each inhibitor molecule occupied was an indication of where the acyl chains of a triacylglyceride might bind. On the basis of these results, they were able to produce a model of the substrate in the active site (Grochulski, Bouthillier *et al.* 1994).

3.2.2 Inhibitor Preparation

(2R)-1,2-di-O-butyl-*sn*-glycerol-3-O-*p*-nitrophenyl-*n*-propylphosphonate (Figure 3.1), a structural analogue of tributyrin, was a kind gift from D. Lun (Massey University). Its synthesis was based on a prior methodology that was devised to produce inhibitors for investigations into lamb pregastric lipase (Taylor, O'Connor *et al.* 2001). The synthesis produced two stereoisomers (about the phosphate group) that were analogues of tributyrin, with a phosphonate-based scissile bond incorporated on to the *sn*-3 chain of the acylglyceride. Ester bonds, which normally link the acyl chains to the glyceride backbone in triacylglycerides, were replaced by more stable ether bonds. This reduced the possibility of enzymatic and spontaneous hydrolysis at these positions, and ensured that the entire triglyceride analogue would be irreversibly bound to the enzyme.

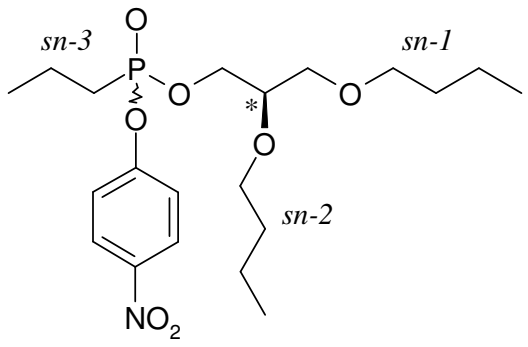


Figure 3.1 *(2R)-1,2-di-O-butyl*-sn-*glycerol-3-O*-p-*nitrophenyl*-n-*propylphosphonate.*
The three chains are numbered to indicate their positions relative to the stereocentre, which is designated with an asterisk.
There is a second chiral centre about the phosphorus atom.

3.3 Methods

3.3.1 AA7 Preparation

The gene for AA7 was obtained as a clone hosted in the plasmid expression vector pGEX-6P3 (GE Healthcare, kind gift from M.-L. Delabre, Fonterra, Palmerston North). This vector expresses the product of the inserted gene as a fusion protein with an N-terminal glutathione-S-transferase purification tag. The purification protocol used had already been determined (M.-L. Delabre, personal communication, 2005). *Escherichia coli* BL-21 hosting the AA7-containing plasmid was grown in LB medium, first as a small starter culture of 20 mL grown overnight at 37°C, which was then used to inoculate 2 L of medium. The larger culture was incubated at 25°C for 5 h. Expression of AA7 was induced when the culture reached an optical density at 600 nm (OD_{600}) of ≈ 0.5 by the addition of isopropyl-beta-D-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. The culture was grown overnight at 25°C. Cells were harvested by centrifugation (4000 x *g* for 25 min), resuspended in phosphate-buffered saline (PBS) and then lysed by French Press (American Instrument Co., two passes at 6000 psi). Cellular debris was removed by centrifugation at 26 000 x *g* for 40 min, after which two x 40 mL of supernatant was added to two x 2 mL of pre-equilibrated glutathione agarose (Sigma). The tubes containing resin and supernatant were incubated with gentle turning at room temperature for 90 min. The resin was collected by centrifugation, washed five times with 40 mL of PBS, then suspended in cleavage buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM dithiolthreitol (DTT), pH 7.5) along with 8 µL of PreScission protease (GE Healthcare), and incubated at 4°C overnight with gentle agitation. The resin was then pelleted by centrifugation at 500 x *g* for 2 min, before being washed three times with 2 mL of PBS to separate the cleaved protein. The partially purified protein was subjected either to final polishing (detailed in Section 4.3.4) or to incubation with inhibitor.

3.3.2 Preparation of Selenomethionine-substituted AA7

After several unsuccessful attempts to find suitable heavy atom derivatives (G. Norris, personal communication, 2005) as well as unsuccessful attempts to find a suitable model for molecular replacement, protein suitable for multiple anomalous dispersion (MAD), with methionine replaced with selenomethionine, was produced using the method of Hendrikson et al. (1990). The pGEX-AA7 plasmid was transformed into electrocompetent *E. coli* DL-21, using electroporation (Sambrook, Fritsch *et al.* 1989). Cells were grown on LB Amp plates, and an initial starter culture was taken from a single colony that was grown overnight in 20 mL of LB Amp. This culture was used to make glycerol stocks of DL-21 pGEX-AA7 cells. A 20 mL culture of this recombinant strain was grown in methionine-containing LeMaster medium, to an OD_{600} of 0.5, then induced with IPTG and analysed for protein expression using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) and a chromogenic activity assay (Shaw 1999). As it was evident that there was expression at acceptable levels, the preparation was scaled up to a 1 L culture with selenomethionine-containing medium. The selenomethionine-substituted enzyme was purified as described above, and polished using size exclusion chromatography (SEC) as described in Section 3.3.4.

3.3.3 Binding of Inhibitors to AA7

For co-crystallisation experiments, a 10-fold molar excess of (2R)-1,2-di-O-butyl-*sn*-glycerol-3-O-*p*-nitrophenyl-*n*-propylphosphonate dissolved in dimethyl sulphoxide (DMSO) was added to AA7 and incubated at 4^{o}C for up to 8 days. The binding of inhibitor to the enzyme was followed by taking small aliquots of the enzyme solution, measuring the residual activity using *p*-nitrophenyl butyrate as a substrate and plotting the loss of activity against time. A control experiment with enzyme only was incubated under the same conditions. Previous experiments had indicated that AA7 with a low occupancy of inhibitor would crystallise as a complex. Therefore, crystallisation experiments were carried out once the activity of the enzyme had decreased by 90%. This was judged to be a suitable compromise between optimal inhibitor occupancy and protein age.

3.3.4 Final Polishing

As a final step, the AA7 protein was subjected to SEC. Protein was concentrated to 30 mg/mL using a 10 kDa molecular weight cutoff centrifugal ultrafiltration device (Sartorius) and exchanged into 20 mM HEPES, 150 mM NaCl pH 7.0. Aliquots (250 µL) were then injected on to a pre-equilibrated G-200 Sephadex tricorn column (GE Healthcare). The column was eluted at 0.4 mL/min, and 0.8 mL fractions were collected for the duration of the run (1.5 column volumes per run ≈ 25 mL). One major peak was observed and, when analysed using SDS-PAGE, showed that a single protein species with a molecular weight of approximately 30 kDa was present in all fractions (Figure 3.2). Fractions from this peak were pooled, concentrated to 25 mg/mL and exchanged into 20 mM HEPES pH 7.0 using ultrafiltration as previously described.

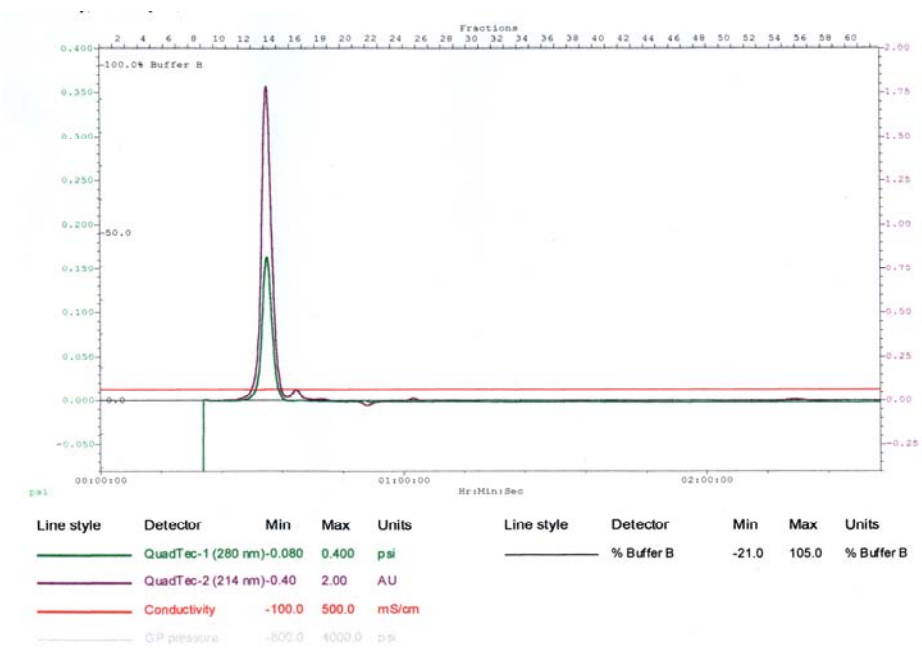


Figure 3.2 *Size Exclusion Run of AA7 on G-200 column. This shows that a single major peak of protein was eluted, which was well separated from some minor peaks that are likely to represent contaminanting proteins.*

3.3.5 Crystallisation

For selenomethionine-labelled protein, crystallisation conditions that had been previously determined for native protein were used (T. Loo, personal communication, 2005). Crystallisation trials to identify lead conditions for the inhibited protein were carried out using a 96-well CrystalQuick sitting drop plate (Greiner BioOne) and a sparse matrix screen (Crystallisation Screen I and II, Molecular Dimensions). Conditions that produced needle-like crystals were optimised using hanging drops suspended over a VDX plate (Hampton Research). Cubic and rectangular crystals, with dimensions of up to 0.2 mm x 0.3 mm, were produced after an overnight incubation at 21^{o}C. Optimal crystal growth was produced in 100 mM sodium acetate pH 4.6, 200 mM $CaCl_2$ and 20% isopropyl alcohol.

3.3.6 Crystallographic Methods

Data were collected and processed as described in Appendix A. Procedures specific to this protein are detailed below.

Several datasets were collected from both native crystals and selenomethionine derivatives, as well as from protein that had been co-crystallised with inhibitor. The data were collected either in-house or at a synchrotron source, as detailed below.

3.3.6.1 Native Structure

A native dataset to resolution 1.7 Å was collected in-house from crystals of mutant AA7 (A223V/T179N) obtained earlier by another investigator (G.E. Norris) using the methods described in Appendix A (Table 3.1). The Matthews coefficient for this was calculated as 2.46 for a single molecule in the asymmetric unit, with a solvent content of 50.1%.

3.3.6.2 MAD Structural Solution

Three datasets from a single selenomethionine-labelled crystal were collected at the Stanford Synchrotron Radiation Laboratory, with the MAD solution being computed on-site on returned electronically (Clyde Smith). Dataset statistics are shown in Table 3.2. Initial phases were calculated using SOLVE/RESOLVE (Terwilliger and Berendzen 1999) to produce an initial solution and then an improved density map. Of the expected eight selenium sites, six were located.

An initial model was built into this map using TURBO-FRODO (Roussel and Cambillau 1989). Limited continuity in the density map, together with the limited resolution of the dataset, meant that the initial model was only about 90% complete. However, the model was adequate to use for molecular replacement to solve the 1.7 Å native dataset provided by T.Loo (Massey Univeristy) using MOLREP (Vagin and Teplyakov 1997). This resulted in a much improved electron density map, and allowed the structure to be completed and refined. Several rounds of rebuilding and refinement were carried out to produce the final model. The statistics of the structure solution and refinement are given in Table 3.3.

	Native AA7	**Inhibited AA7**
Resolution Range (Å)	31.10 – 1.71	38.84 – 1.96
Outer Shell	(1.71 – 1.81)	(2.15 – 1.96)
Space Group	I4	I4
Cell parameters a, b, c (Å)	109.50 109.50 59.43	109.85 109.85 59.34
α, β, γ (°)	90 90 90	90 90 90
Unique Reflections	38046	25554
Total Reflections	149269	110597
Completeness % (Outer Shell)	99.1 (97.5)	99.5 (100)
Redundancy	3.93 (3.83)	4.33 (4.21)
Mean I/σ	17.4 (5.7)	11.4 (4.9)
R_{merge} (%)*	0.040 (0.182)	0.066 (0.207)

Table 3.1 *Data collection statistics for native and inhibited structures.*
$(\Sigma|I-\{I\}|/ \Sigma\{I\})$ Where I = intensity of a given reflection and {I} = mean of the intensity for multiple measurements of that reflection

MAD Data Collection and Scaling			
Dataset	**Peak**	**Remote**	**Inflection**
Wavelength	0.97929 Å	0.89194 Å	0.97953 Å
Total Measurements	186532	129645	123976
Unique Reflections	12302	12257	12298
Resolution Range (Outer Shell)	50–2.5 Å (2.59–2.50)	50–2.5 Å (2.59–2.50)	50–2.5 Å (2.59–2.50)
Average Redundancy (Outer)	3.7 (3.6)	3.6 (3.6)	3.6 (3.7)
Average Completeness (Outer)	99.8 (99.7)	99.7 (99.8)	99.7 (100)

MAD Solution (SOLVE/RESOLVE)						
Resolution Range Used	19.433–2.5 Å					
Reflections	12141					
Unit Cell Parameters	109.3 109.3 59.2 (90 90 90)					
Space Group	I4 (# 79)					
Solution Quality (After SOLVE)	50.133					
Heavy Atom Site	1	2	3	4	5	6
Occupancy	0.639	0.496	0.349	0.509	0.693	0.282
B-factor	44.2	30.4	4.8	29.3	60.0	33.6
Height/Sigma	16.3	15.5	14.9	14.1	12.9	7.4
Mean Figure of Merit (After RESOLVE)	0.63					

Table 3.2 *MAD datasets and solution.*

	Native AA7	**Inhibited AA7**
No. Residues Built (Total)	308 (317)	303 (317)
No. Water Molecules	267	117
Average B-factors, Å^2		
Mainchain	21.95	38.02
Sidechain	23.49	37.90
Water molecules	35.86	39.66
Overall	24.00	38.21
Refinement		
Reflections used	35974	24175
Reflections R_{free}	1894 (5%)	1249 (5%)
RMSD bond lengths (Å)	0.011	0.003
RMSD bond angles (^{o})	1.28	0.695
R-factor	0.156	0.213
R-free	0.178	0.237
Ramachandran Analysis		
Most favoured	249 (91.9%)	244 (92.1%)
Additionally allowed	20 (7.4%)	19 (7.2%)
Generously allowed	1 (0.4%)	1 (0.4%)
Disallowed	1 (0.4%)	1 (0.4%)

Table 3.3 *Structural refinement and model statistics.*

3.3.6.3 Inhibitor-bound Structure

A single dataset was collected in-house, as summarised in Tables 3.1 and 3.3. The model of the native structure described above was used to solve the inhibitor-bound structure directly using REFMAC5 (Murshudov, Vagin *et al.* 1997). This initial rigid body refinement gave an R_{work} and R_{free} of 24% and was also found to produce an interpretable density map. This direct refinement approach was used as the native and inhibited structures had virtually identical unit cells, and were of the same space

group, so performing a translation and rotation function to locate the asymmetric unit in the cell was deemed unnecessary.

After examination of the initial density map using COOT (Emsley 2004), regions of the model that coincided with continuous negative density in the F_o–F_c difference map were set to an occupancy of zero. After 10 cycles of refinement using REFMAC, the structure of the peptide chain was clarified for areas where conformational change between the native and inhibitor-bound structures had occurred. Density representing a tetrahedral ligand with acyl groups attached to the serine nucleophile sidechain was seen. Strong density was obvious for the phosphonate-associated propyl chain, and for the glyceride backbone and *sn*-1 acyl chain, and these features were easily modelled into the structure. No density associated with the base of the *sn*-2 chain was seen, but density was found near the lip of the P2 cleft.

A molecular model of the inhibitor was generated using Sketcher (CCP4i), with bond restraints generated using REFMAC (Murshudov, Vagin *et al.* 1997).

In the final model, the inhibitor was modelled with the *sn*-2 chain projecting upwards, associated with the density at the side of the cleft, as this was the only sterically reasonable position that the *sn*-2 chain could occupy (discussed more fully in Sections 4.5.2 and 4.5.3).

Additionally, the changes in conformation of the mainchain between residues 198 and 208 made it difficult to model all the sidechain atoms unambiguously into the density. Although this area had continuous, albeit weak, density that was representative of the mainchain of the peptide, sidechains could not be unambiguously placed during modelling for some residues. Because of weak density, residues 200–205 were left unbuilt in the final model, although, because of the short distance between the last visible points of mainchain density, the predicted position of the mainchain has been shown for reference in some diagrams in this chapter. It is likely that the peptide chain is mobile in this region, which would fit with the fact that it is the region that undergoes the greatest movement during inhibitor binding, and is not restrained by any contacts with the rest of the protein.

3.4 Native Structure of AA7 Esterase

3.4.1 Overall Structure

The native structure of AA7 esterase, shown in Figure 3.3, conforms to the αβ hydrolase fold (Ollis, Cheah *et al.* 1992), with a few additional features. The enzyme has an eight-stranded β sheet, surrounded by nine helices, plus three short helical twists.

The N-terminus of the protein consists of a long α helix that stretches from residue 10 to residue 33. This helix is additional to the original description of αβ hydrolases, as the first helix in the prototype fold, helix A, is found between strands 3 and 4, and is therefore designated helix N. The traditional system of naming inserted helices in a more internal position, for example A`, to indicate a relationship to the neighbouring prototype helix does not seem to be appropriate here, as this helix is separated from helix A by two β strands and 66 residues, rather than being an immediately adjacent insertion.

There is a short helix inserted after helix C, which has been termed helix C` (residues 167–173). The largest addition to the standard fold occurs between strands 6 and 7, which is the most common place for additional helices to be found (Ollis, Cheah *et al.* 1992). In AA7, this insertion consists of two short helical regions (residues 220–226 and 228–232) in an ordered loop, a longer helix (203–216), which is partially rearranged during inhibitor binding, and a small helix consisting of only three residues (194–197). It is not entirely clear which out of these four helices is actually helix D, which is normally found between strands 6 and 7 in the αβ fold. The only large helix (203–216) is unusually distant from the core of the enzyme and is involved in inhibitor binding, its two N-terminal residues unwinding to form an ordered loop when the inhibitor binds (discussed in Section 3.5.1). For this reason, it may not be part of the core αβ fold, but may be rather better described as an insertion. As helices (220–226) and (228–232) are closer to the core of the fold, and immediately adjacent to each other, they are more likely to represent two halves of helix D with a short loop between them. The helices of the structure are therefore defined as follows. The large helix (203–216) involved in inhibitor binding is helix D`, the three-residue helix (194–197) is helix D`` and the two helices that make up the split helix D are helix D_1

(residues 220–226) and helix D_2 (residues 228–232). There is also a short helix adjacent to helix C, termed C` (167–173), and two short helices, F` (275–277) and F`` (285–288), adjacent to the large F helix (293–316), at the C-terminus of the protein.

Figure 3.3 *Stereoview of the native structure of AA7 esterase.*
The residues that comprise the catalytic triad are shown as sticks. Helices that are insertions into the prototype αβ hydrolase fold are shown in yellow.
Upper panel: Looking into the active-site region. Lower panel: Side view showing the central β sheet, and emphasising how the inserted helices are positioned around the active-site cleft.

3.4.2 Catalytic Triad

Like most αβ hydrolases, AA7 contains a catalytic triad made up of a serine nucleophile, a histidine base and an aspartic acid residue, shown in Figure 3.4. As discussed in Chapter 1, these residues are connected by a hydrogen-bonding network (Section 1.6). The catalytic serine (Ser 151) is found in a five-residue motif (G-D-S-A-G) that conforms to the G-X-S-X-G motif that is typical for αβ hydrolases (Ollis, Cheah *et al.* 1992). The serine mainchain torsion angles are also found in a disallowed region of the Ramachandran plot, again characteristic for hydrolases with this fold. The catalytic histidine, His 283, is found near the C-terminus of the protein, on a loop between helices F` and F``, which projects over the catalytic serine. The acid of the triad, Asp 248, is located on a loop between strand 7 and helix E. The identity of Ser 151 as the nucleophile was confirmed by the observation of electron density representing a tetrahedral species linked to the serine Oγ in the structure with inhibitor bound.

In addition to the three catalytic triad residues, AA7 also has a conserved hydrogen-bond donor at the top of strand 6, believed in other αβ hydrolases to anchor the acid of the catalytic triad (Stok, Goloshchapov *et al.* 2004). This residue is most commonly found to be serine, but in AA7 it is a cysteine, Cys 189, one of three free cysteine residues in the sequence. It is interesting to note that in a prior study on rat esterase mutating the equivalent serine to cysteine resulted in low levels of recombinant enzyme expression (Stok, Goloshchapov *et al.* 2004). The relative stability of AA7 esterase may make it suitable for studies on the effects of mutating this cysteine to other residues, such as serine, to investigate its role in catalysis more fully.

3.4.3 Oxyanion Hole

The oxyanion hole in AA7 is composed of backbone amides contributed by three residues (distances to the oxyanion indicated in brackets): Gly 80 (2.94 Å), Gly 81 (2.82 Å) and the top residue of the nucleophile helix Ala 152 (3.11 Å), as shown in Figure 3.4. The first two glycines are found in the sequence motif HGGG (residues 78–81), a motif that is conserved in all members of the HSL (hormone sensitive lipase) family.

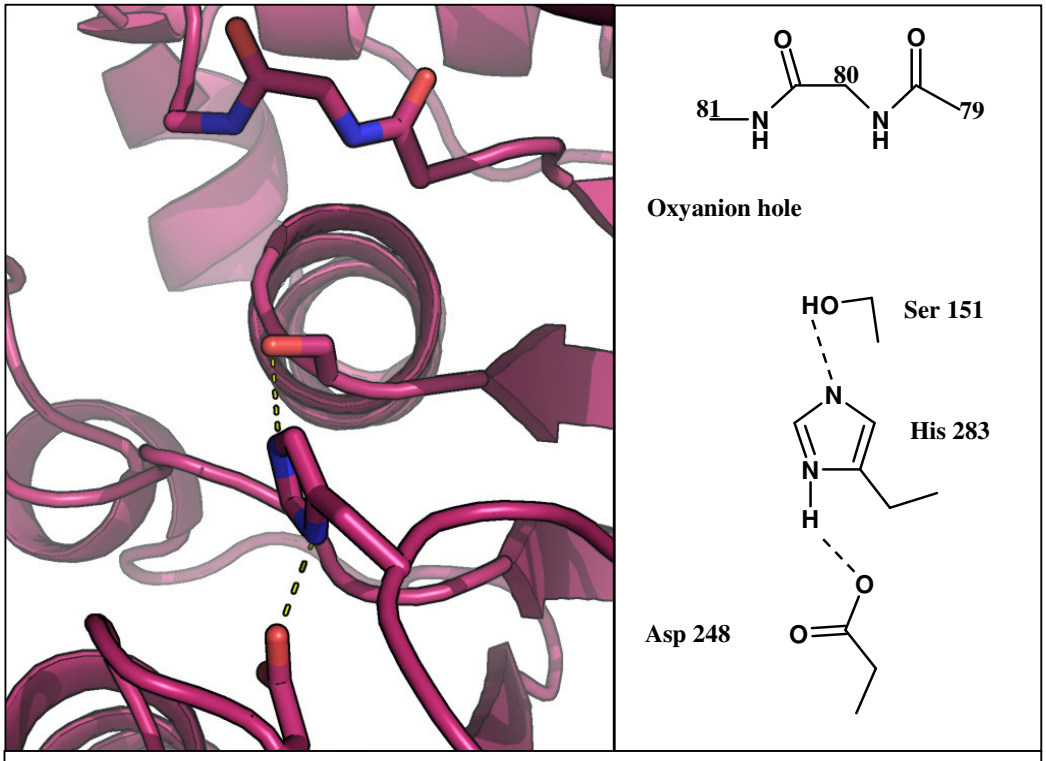


Figure 3.4 *Catalytic triad and oxyanion hole in AA7.*
In this view, the substrate-binding cleft bisects the frame horizontally, with the P1 tunnel on the right and the P2 cleft on the left (see Section 4.4.5), in the same plane as the catalytic residues. The amides of the nucleophile elbow helix, which would point up out of the plane of the paper, have been omitted for clarity, but also contribute to stabilising the oxyanion.

This motif allows the amide of Gly 80 to point towards the catalytic site, constrained by a hydrogen bond between the mainchain carbonyl oxygen of Gly 79 and the Nδ of His 78 (2.62 Å) (De Simone, Galdiero *et al.* 2000). Although the interaction with the third mainchain amide, that of Ala 152, is long for a hydrogen bond, it is a special case, as a general interaction with the N-terminal dipole of the helix is believed to assist the stabilisation of the oxyanion (Hakansson 2002). Like other bacterial hormone-specific lipase orthologues, such as Est2 from *Alicyclobacillus acidocaldarius*, the active-site histidine is within reach of both sides of the substrate-binding cleft. The oxyanion hole preserves this symmetry, as the two mainchain glycine amides (Gly 80 and Gly 81) surround the tetrahedral oxyanion analogue in the inhibited structure.

It is interesting to note that, for Est2 from *A. acidocaldarius*, structural analysis using inhibitors has shown that the enzyme can act on substrates that are bound in either orientation relative to the catalytic triad in the catalytic cleft (De Simone, Galdiero *et al.* 2000; De Simone, Mandrich *et al.* 2004). Work on trypsin has also shown that it can hydrolyse "inverse substrates", which are ester compounds that are designed to bind in the active site in the reverse orientation to the usual (Tanizawa, Kanaoka *et al.* 1987). These findings are a precedent for the variability in substrate-binding directionality that has been observed for AA7, which is discussed later in Section 4.6.2.

3.4.4 Mechanism

Although the structures solved do not completely elucidate the mechanism of catalysis, the arrangement of the catalytic triad residues appears to be identical to that seen in other αβ hydrolases. The mechanism of AA7 is therefore assumed to be similar, consisting of an initial nucleophilic attack by Ser 151 on the scissile bond of the substrate, to form a transient tetrahedral oxyanion state. The tetrahedral adduct rapidly decomposes, breaking the scissile bond, to produce an acyl-enzyme intermediate. In hydrolysis, a water molecule is then activated by His 283, and attacks the reaction centre to break the acyl-enzyme bond, producing a free fatty acid. In alcoholysis, an alcohol instead of a water is activated to carry out the nucleophilic attack, yielding an ester rather than a free fatty acid.

As it appears that AA7 can operate as both a transferase and a hydrolase, it may use either water or ethanol as an acceptor, depending on which one is available in the catalytic site at the time. In support of this, there appears to be a specific binding site for a non-water acceptor (Section 4.6.3).

3.4.5 Substrate-binding Sites

The surface geography of AA7 is dominated by a large cleft that is open to solvent and contains the catalytic serine. Immediately adjacent to the cleft is a deep cavity, which extends under the surface of the enzyme to form a closed tunnel, as shown in Figure 3.5. Upon binding of the inhibitor to the enzyme, the cleft becomes slightly enlarged at the end opposite the tunnel by a conformational change in the protein, as

described in Section 4.5.1. Like the other esterases studied in this project, the cleft can be divided into two binding pockets, P1 (the tunnel) and P2 (the cleft), which are located on either side of the nucleophile.

3.4.5.1 The P1 Binding Pocket

The P1 binding pocket is closest to Ser 151 and is a narrow closed cavity that modelling studies using COOT (Emsley 2004) and Pymol (DeLano 2002) suggest could accommodate an acyl chain of up to 18 carbons, if it was somewhat kinked. The tunnel opens near the catalytic serine, while leaving some room for groups to project up the side of the active-site cleft, before the tunnel becomes fully closed over.

3.4.5.2 The P2 Binding Pocket

The P2 cleft is open to the solvent, and contains a high proportion of hydrophobic and aromatic residues. The pocket opens when the enzyme binds a substrate analogue, which is detailed in Section 4.5. In the open conformation, P2 can accommodate the *sn*-2 chain in an upward-extending conformation, as well as the glyceride backbone and the other primary acyl chain. The size of the secondary (*sn*-2) acyl chain is not limited by any steric factors, as it can project upwards and out of the cleft. The primary acyl chain size may be limited by the length of the P2 cleft, but there is a "shelf" above Val 197 that may accommodate longer acyl chains in the primary positions by hydrophobic association.

Structural and kinetic data suggest that the arrangement and the size of the binding pockets are a major determinant of substrate specificity in the enzyme. This is discussed in Section 4.6.

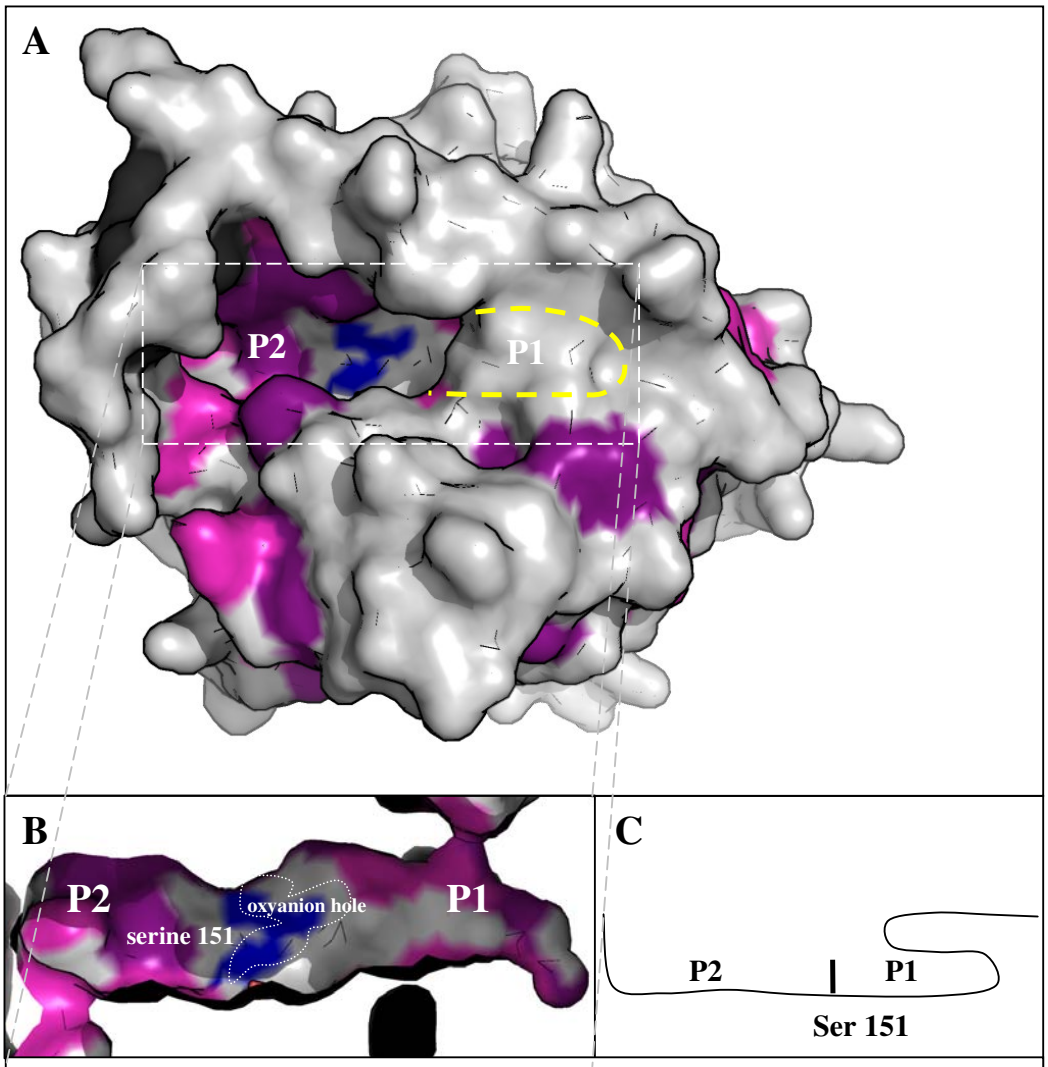


Figure 3.5 *AA7 substrate binding.*

A: Overview of the binding cleft in the open (inhibitor-bound) form of the enzyme, with the substrate-binding clefts indicated. The inhibitor has been removed for clarity.

B: Cutaway of the surface showing the extent of the P1 tunnel, with its hydrophobic lining, shown in pink (hydrophobic sidechains) and purple (aromatic sidechains).

3.5 Inhibitor-bound Structure

3.5.1 Structural Changes upon Inhibitor Binding

The structure solved with (2R)-1,2-di-O-butyl-*sn*-glycerol-3-*n*-propylphosphonate bound in the active site showed one major conformational change. When the inhibitor binds to the enzyme, the kinked terminal of helix D` becomes straightened. This causes the movement of several sidechains belonging to residues of this helix, most notably Met 204, which moves out of the substrate-binding cleft. Although no clear density for this residue's sidechain is visible, the movement of the mainchain means that the sidechain must move to open the binding cleft (Figure 3.6).

The conformational change appears to involve residues 196–208 and results in an enlargement of the P2 binding cleft. The residues that make up the rest of the cleft lining come from the more canonical regions of the fold and do not appear to undergo any conformational changes, supporting the observation that αβ hydrolases provide a basic scaffold for the positioning of different specificity determinants (Heikinheimo, Goldman *et al.* 1999). A comparison of the native and inhibitor-bound structures shows that, when the sidechain of Met 204 moves, its position is replaced by the *sn*-1 acyl chain of the inhibitor (Figure 3.7). Although it appears that the sidechain of Met 204 becomes exposed to the solvent as a result of this movement, the hydrophobic interactions it made with residues in the P2 pocket are replaced by interactions between the inhibitor acyl chain and the pocket residues, in what is effectively an exchange of the hydrophobic sidechain of Met 204 with a hydrophobic acyl chain from a substrate. The displacement of the α carbons that occurs during this movement is shown in Figure 3.8

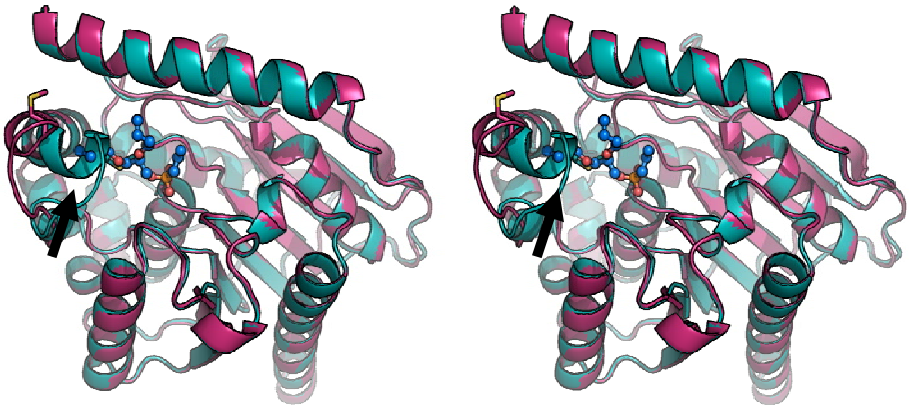


Figure 3.6 *Stereo diagram of AA7 with (magenta) and without (aqua) inhibitor bound. The inhibitor is shown in blue, with helix D` indicated by the arrow. Met 204 is shown as sticks.*

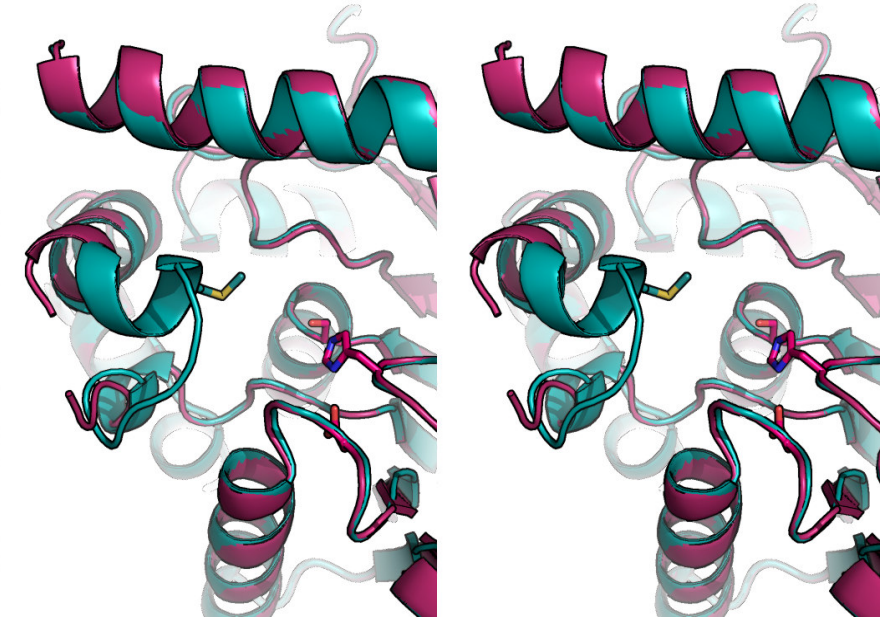


Figure 3.7 *Stereoview of the movement of Met 204 that occurs upon substrate binding. Molecules are coloured as in Figure 4.4, with Met 204 shown in stick representation. The inhibitor has been omitted for clarity.*

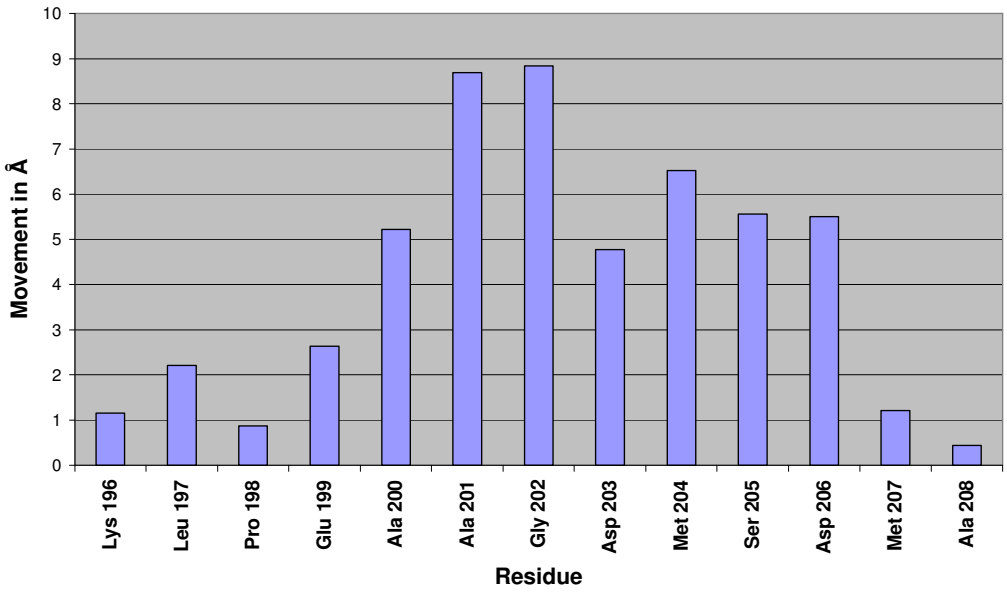


Figure 3.8 *Movement of α carbons during conformational change of the mobile loop.*

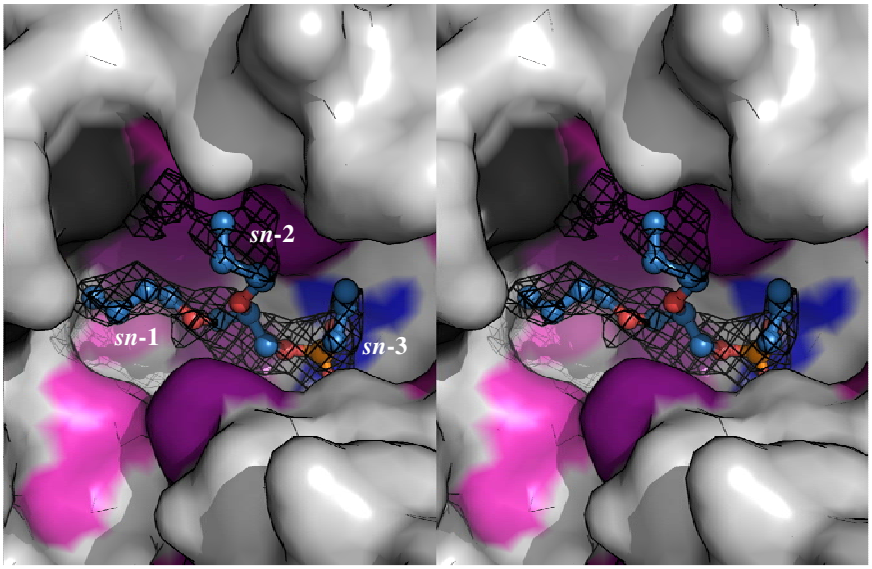


Figure 3.9 *Omit F_o–F_c density map contoured at σ = 0.50 for the region about the inhibitor. Hydrophobic residues are shown in pink, aromatics in purple, methionine dark grey, and the oxyanion hole in blue.*

3.5.2 Inhibitor-binding Studies

Only the *sn*-3 chain, glyceride backbone and *sn*-1 chain of the inhibitor were found to have visible electron density, but the steric constraints on the inhibitor allowed a possible conformation of the *sn*-2 chain to be modelled. The acyl chain associated with the scissile bond, in this case the *sn*-3 chain, points upwards near the entrance to the P1 tunnel with the glyceride backbone and the *sn*-1 chain lying flat along the cleft. From knowledge of the structure of triglycerides, the *sn*-2 chain was modelled in an outward-facing conformation, as there is no other realistic conformation that this chain could adopt (Figure 3.9). Some density that could be associated with the end of the *sn*-2 acyl chain when it was placed in this position was seen. Interestingly, other bacterial HSL family members can bind similar substrates in both orientations in the catalytic cleft (De Simone, Galdiero *et al.* 2000; De Simone, Mandrich *et al.* 2004). Recent studies suggest that this is the case for AA7 (D. Colbert, unpublished data, 2007), which might explain the difficulties experienced in modelling the inhibitor into the available density, because, if binding was occurring in both orientations, the *sn*-2 chain would be in a different position for each orientation, and might not be visible because of disorder.

3.5.3 Inhibitor Binding by Chain

As already mentioned, it is possible that the inhibitor used for this experiment binds in either direction along the substrate-binding cleft. The description below is based on the assumption that the inhibitor binds in the direction shown in Figure 3.8, although, as the interactions shown are hydrophobic in nature and therefore relatively non-specific, they could also apply if the inhibitor was bound in the opposite orientation. These interactions, detailed in the following sections, are depicted in Figure 3.10.

3.5.3.1 The sn-*3 Acyl Chain*

The *sn*-3 chain, which contains the reactive phosphonate group, binds near the entrance to the cavity, but projects upwards, rather than into the P1 tunnel. The main contact between the *sn*-3 chain and the enzyme appears to be mediated by a hydrophobic interaction with Trp 20, although, if the *sn*-3 chain was larger (e.g. C6), it might be able to interact with Val 284, found on the opposite side of the binding

cleft. Recent structural data obtained in our laboratory indicate that the tunnel entrance might be occupied by an acyl chain during the binding of a substrate analogue, but not completely, leaving room for an acceptor to bind either in the tunnel or below the glyceride backbone. In this structure, the substrate glyceride backbone lies in a similar position to that of the *sn*-3 chain as modelled in the structure reported here (D. Colbert, personal communication, 2007).

3.5.3.2 The sn*-2 Acyl Chain*

As mentioned previously, although the whole *sn*-2 chain could not be seen in the structure, there was enough density to suggest that it projected upwards from the P2 cleft where it could form a hydrophobic association with Trp 20. This puts the chain near a hydrophobic patch on the side of the cleft, which could potentially accept acyl chains that are longer than the groups of the inhibitor used in this experiment. There are two possible explanations for the lack of electron density for part of the *sn*-2 chain. One is disorder or mobility of the *sn*-2 acyl chain, possibly caused by the inhibitor binding in two different orientations relative to the binding cleft. If the *sn*-2 chain was itself also mobile, then the occupancy of the chain might be too low for it to be seen in the density maps. The other explanation is that the *sn*-2 chain has been partially lost in the inhibited enzyme complex because of a hydrolytic process. As ether bonds are chemically very stable, this scenario is unlikely. It should be noted that ether bonds can be hydrolysed slowly in the presence of halogen acids, such as HCl, but, as the incubation of AA7 with the inhibitor was carried out at pH 7.0 in HEPES buffer, acid hydrolysis of the *sn*-2 chain is extremely unlikely. Enzymatic hydrolysis is another possibility, although the usual ester linkages in the *sn*-1 and *sn*-2 positions of the inhibitor were replaced with ether bonds to prevent this occurring. Enzymes with an etherase activity appear to be rare, although they have been characterised in *Pseudomonas paucimobilis* and *E. coli* respectively (Masai, Katayama *et al.* 1991; Jaeger, Arsic *et al.* 2005).

The possibility of an aging reaction removing the sidechain was considered. This has been seen in structures such as that of the lipase from *Pseudomonas cepacia*, where the diacylglyceride group was slowly hydrolysed from a similar phosphonate triglyceride analogue over time, leaving only a butyl phosphonate bound to the enzyme (Lang, Mannesse *et al.* 1998). It should be noted that “aging” in bound

adducts is usually used to refer to a process that is catalysed by the enzyme itself, in which a reaction involving the catalytic base and acid causes partial deacylation of the bound adduct. In the case of AA7, His 283 is 6.34 Å from the ether oxygen of *sn*-2, and is thus too far away to catalyse any aging process, making this explanation of the lack of density unlikely.

3.5.3.3 Glyceride Backbone

The main interactions between the glyceride backbone and the enzyme cleft are between Trp 20 and Trp 82, Phe 250 and Val 191. Both the backbone and the *sn*-1 chain of the inhibitor can interact with Trp 20 as the inhibitor is bent about the phosphonate group into almost a semicircle. Trp 82 and Val 191 form part of the base of the cleft below the glyceride, with Phe 250 opposing Trp 20 to form the other side of the cleft at this point.

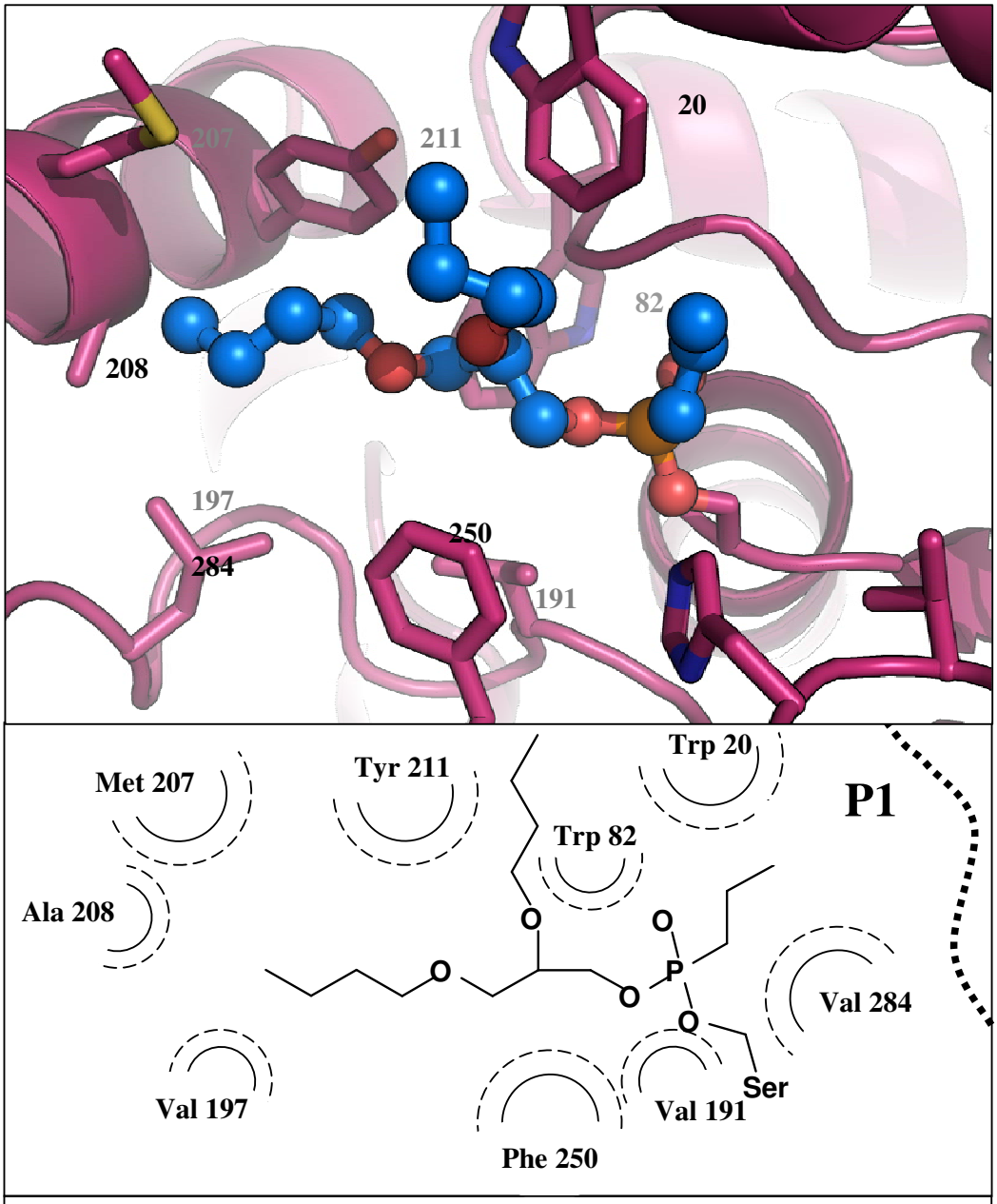


Figure 3.10 *Hydrophobic and aromatic interactions in the substrate-binding cleft.*

*Upper panel: Cartoon showing the interactions between the adduct formed between (2R)-1,2-di-O-butyl-*sn*-glycerol-3-*n*-propylphosphonate and AA7. Hydrophobic and aromatic residues are shown as sticks with other sidechains omitted for clarity.*

Lower panel: Schematic diagram of binding site interactions. Residues on the bottom of the cleft are represented by smaller semicircles, whereas upper residues are represented by larger semicircles.

3.5.3.4 The sn-*1 Acyl Chain*

Interactions between the *sn*-1 chain of the inhibitor and the enzyme are also hydrophobic in nature, and occur between the acyl chain and the sidechains of Tyr 211, Met 207, Ala 208, Val 197 and Phe 250.

These interactions can occur only after the conformational change in the enzyme is induced by the inhibitor binding, and involve similar interactions to those that occur with the sidechain of Met 204 in the unbound conformation of the enzyme. The exception to this is the sidechain of Ala 208, which does not interact with Met 204 in the closed state, but can interact with the inhibitor after change in conformation of helix D`, of which it is part.

3.5.4 Triglyceride Binding

The key difference between the inhibitor and a triglyceride substrate is that a triglyceride has scissile bonds linking all three acyl chains to the glyceride backbone, whereas, in the inhibitor, only the *sn*-3 bond can be attacked by the enzyme to form the phosphonate enzyme intermediate. This means that, whereas the inhibitor is forced to bind with the *sn*-3 chain associated with the catalytic serine, a prochiral triglyceride such as tributyrin would not be so restricted, as all three ester bonds are, in theory, able to be hydrolysed. Chains *sn*-1 and *sn*-3 are effectively interchangeable, if the stereo configuration about the glyceride stereocentre is ignored. In AA7, as there appear to be no polar binding groups in the substrate-binding cleft, other than those that form the oxyanion hole, any specificity of the enzyme would be restricted solely by the size and the conformation of the substrate and by the shape and the size of the binding cleft.

3.6 Substrate/Inhibitor Binding and Kinetics

As well as the structural data obtained in this investigation, some data on the glyceride preference of AA7 have previously been reported (Holland, Liu *et al.* 2005), which allows some inferences to be made about the substrate specificity with respect to the structural features observed. The data from this publication are reproduced in Table 3.4.

Because these data do not include specific activities, a direct comparison between hydrolysis and alcoholysis cannot be performed in absolute terms. However, the relative activities can be used to try to understand the data in terms of the structure.

3.6.1 Differences between Mono-, Di- and Triglyceride Activities

From the data, it is obvious that the rates of activity are greater on monoglycerides than diglycerides, and that diglycerides produce a higher activity rate than triglycerides. This may be due to monoglyceride association with the enzyme being easier because of fewer steric constraints on binding, compared with a triglyceride. However, it must be remembered that the solubility of glycerides in general is inversely proportional to the number of substituents present on the glyceride backbone, meaning that, as all glycerides were provided at an identical concentration (66 mM) during the experiment (Holland and Coolbear 1996; Liu, Holland *et al.* 2003), there would be a higher availability of soluble substrate to the enzyme for mono- and diglycerides than for triglycerides. The greater availability for a given concentration of a mono- or diglyceride over the same concentration of a triglyceride may explain why higher activity is seen in these compounds in both the hydrolysis experiment set and the alcoholysis experiment set.

		Relative alcoholytic activity	Relative hydrolytic activity
Tributyrin	C4	0	54
Tricaproin	C6	29	**100**
Tricaprylin	C8	**100**	1
Dicaproin	C6	1973	5540
Dicaprylin	C8	35	119
Dicaprin	C10	57	0
Monocaprylin	C8	5805	31748
Monocaprin	C10	109976	93720
Monolaurin	C12	518	27556
Monomyristin	C14	582	8471
Monopalmitin	C16	0	2580

Table 3.4 *Relative amounts of hydrolysis/alcoholysis of glycerides by AA7. Glycerides are grouped by number of substituents, and the normalisation standard is indicated in bold in each case. Table reproduced from (Holland, Liu et al. 2005).*

3.6.2 Difference in Acyl Chain Length

Table 3.4 shows an interesting difference between hydrolytic and alcoholytic modes of catalysis in AA7, in that the preference for acyl chain length appears to differ depending on the mode. From analysis of the positioning of the catalytic triad relative to the substrate-binding cleft, and from what is known about related enzymes, the kinetic results can be explained as a combination of typoselectivity for acyl chain size, and how this influences both the rate of substrate binding and the orientation in which a substrate binds relative to the substrate-binding cleft.

Because of the symmetry of the catalytic triad and the oxyanion hole in relation to the substrate-binding sites, and the ease of access of water to the catalytic machinery, AA7 esterase should be able to hydrolyse substrates that are bound in either direction along the substrate-binding cleft. This is because water, being small and in high abundance, can enter the site easily, needing to be bound in position only transiently to be activated as a nucleophile and for complete deacylation of the enzyme. Because there appears to be only one site for the binding of an acceptor other than water (either in or near the entrance to the P1 tunnel), alcoholysis can occur only if the substrate binds so as to leave the scissile acyl chain in the P2 cleft. The directionality of this binding is likely to be affected by typoselectivity, as has been shown for the

bidirectional substrate binding in Est2 from *A. acidocaldarius* (De Simone, Galdiero *et al.* 2000; De Simone, Mandrich *et al.* 2004).

This means that, despite the catalytic triad arrangement producing a symmetrical chemical environment in the active site, the asymmetry of the substrate-binding sites directs the differences in typoselectivity and deacylation substrate seen in the kinetic results.

3.6.3 Hydrolysis and Transferase Substrate Preference

The differences in preference for chain size of triglycerides shown in Table 3.4 can be explained if glycerides with smaller substituents are preferentially bound in a different direction from glycerides with larger substituents, and if there is only one site for a non-water acceptor to bind. In this model, summarised in Figure 3.11, the smaller glycerides tributyrin and tricaproin would bind with scissile chains towards the P1 binding pocket, preventing the access of any acceptor to the catalytic machinery from this direction. However, the C8 triglyceride tricaprylin binds in the opposite direction along the cleft. After acylation of the enzyme and dissociation of the diglyceride, an acceptor bound either in or near the entrance to the P1 tunnel could move forward into position to deacylate the enzyme and produce an ester product. It is tempting to make assumptions from the reported results, such as that tricaproin must bind to the enzyme in the correct direction for alcoholysis to occur 29% of the time, because of the relative numbers shown. However, tricaproin is more soluble than tricaprylin, although it is less hydrophobic, and, because these assays were conducted at saturation concentration, it is extremely difficult to compare the activities of different acylglycerides in this way. Qualitatively, however, this model appears to fit the observed differences in activity rates. This model also fits the observed peak of ester production seen for C10 monoglycerides, because the P2 pocket appears to be able to fit an acyl chain of up to 10 atoms in length. This would explain the sudden drop in ester synthesis for larger monoglycerides.

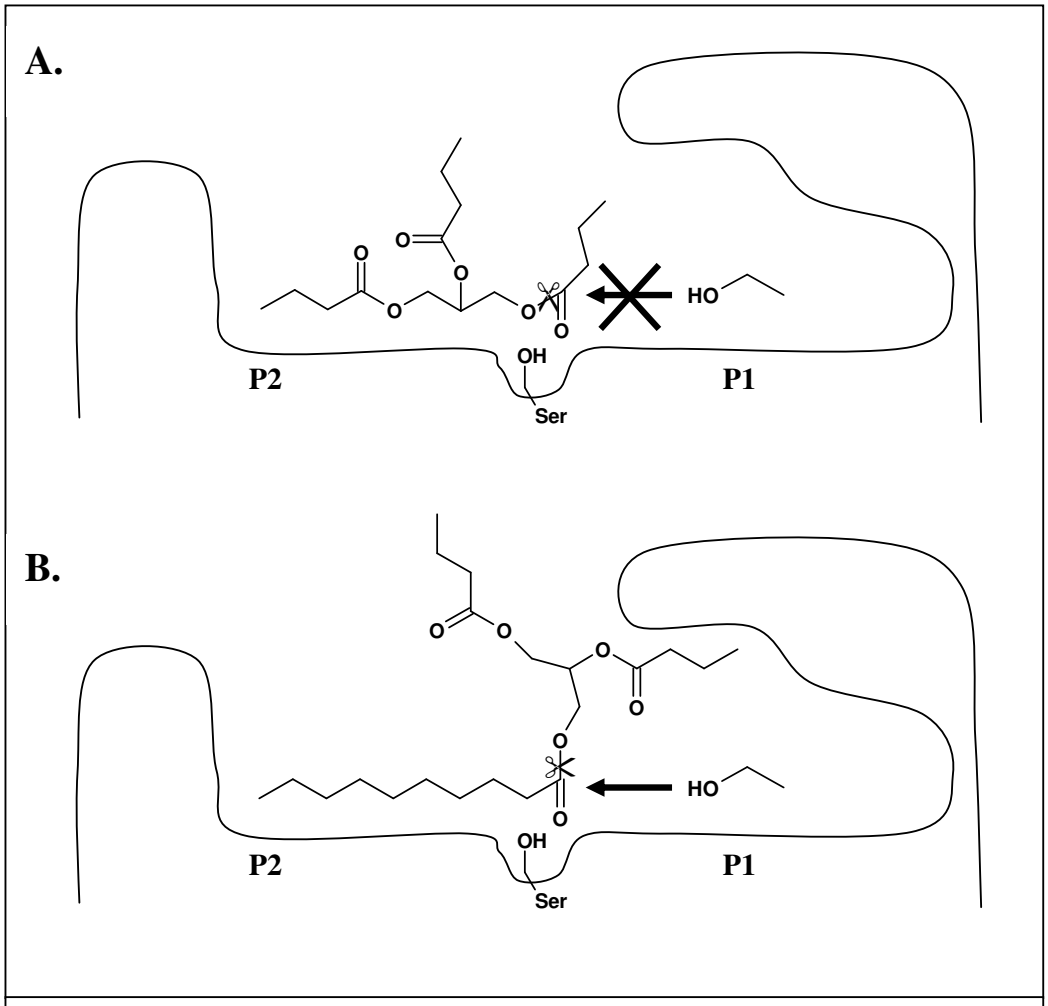


Figure 3.11 *Schematic representation of substrate binding in AA7.*

A model showing how substrates bind to AA7 relative to the two binding sites and the catalytic machinery. The scissile bond is indicated by scissors, and the P1 and P2 binding sites are labelled. Panel A illustrates how, in a tributyrin analogue where the scissile acyl chain binds in the P1 pocket, any acceptor bound in the P1 tunnel (represented here by an ethanol molecule) is prevented from accessing the catalytic machinery because of the bound acyl chain. Panel B shows how a substrate with a C8 scissile chain binds in the opposite direction to the C4 substrate, allowing an acceptor in the P1 tunnel to access the catalytic site (D. Colbert, unpublished data, 2007). The lack of solid density for the tributyrin analogue suggests that tributyrin might bind in either of these orientations, but the greater rate of hydrolytic activity on shorter triglycerides might mean that the orientation illustrated in Panel A is favoured.

3.6.4 Specificity

The model of the enzyme–triglyceride complex suggests that the enzyme may have a preference for binding the primary acyl chains in a position in which they can be hydrolysed. This is because access to the catalytic machinery would be difficult for a secondary chain due to steric constraints caused by the glyceride backbone.

The main difference in binding the two alternative primary positions of an acylglyceride is the direction in which the *sn*-2 chain points relative to the second glyceride carbon. As shown in Figures 3.8 and 3.9, the active site of AA7 appears to have hydrophobic patches on both sides of the substrate-binding cleft, mainly contributed by Trp 20 (at the top in the diagram) and Phe 250 (at the bottom), which could stabilise the *sn*-2 chain regardless of which primary chain is bound in the active site. This hydrophobic patch also appears to be able to stabilise the *sn*-2 chain of a glyceride bound in the opposite conformation, with the scissile chain in the P2 pocket and the backbone groups extending upwards out of the binding cleft, meaning that there are no obvious molecular features that could cause stereospecificity in AA7.

It should be noted that it is not uncommon for lipases and esterases to be capable of hydrolysis at both of the primary positions of a glyceride; for example, human pancreatic lipase is active on both *sn*-1 and *sn*-3 positions of lipids (Jensen, deJong *et al.* 1983). The binding site of human pancreatic lipase is essentially symmetrical about the catalytic site (Egloff, Marguet *et al.* 1995), and, whereas the AA7 binding site is asymmetric, this is only because of the presence of the acceptor-binding P1 tunnel.

Selectivity in AA7 might be limited to regio- and typoselectivity dictating the direction in which a glyceride can bind. With the entire binding-site region being relatively hydrophobic, it would appear that either of the primary chains could be accommodated as long as they are less than the 10 carbon maximum size permitted by the P2 cleft. Further experiments with chiral inhibitors are under way to confirm if any stereoselectivity exists, and to explore the effects of chain size on binding direction and rates.

3.6.4.1 Determining Stereospecificity

Although not within the scope of the current work, several methods exist for investigating the stereospecificity of the enzyme, and how different-sized acyl groups at different positions are able to bias this.

The first method involves analysis of the activity of the enzyme on different glyceride substrates, with different-sized substituents on each position of the glyceride. From this experiment, the liberation of the different-sized glycerides can be linked to the position from which they came. This would require some caution in drawing conclusions, however, as the stereospecificity of some enzymes can be biased depending on the size of acyl groups at other positions. For example, certain bacterial lipases exhibit changes in *sn*-1 and *sn*-3 stereospecificity when the constituents present on the *sn*-2 position are varied (Stadler, Kovac *et al.* 1995).

An alternative involves using chirally pure glyceride-based inhibitors, such as that used in this study. By synthesising inhibitors with the phosphonate group at the three different chain positions, each inhibitor can be assayed for the rate at which it inhibits the enzyme. The inhibitor that matches the stereopreference of the enzyme will inhibit most efficiently. This work is currently being undertaken in a separate project (D. Colbert, personal communication, 2007).

3.7 Summary

The crystal structure of AA7 esterase from *Lactobacillus rhamnosus* was solved using MAD methods to gain a low resolution (2.5 Å) structure, with a molecular replacement solution on a native crystal providing a structure to a resolution of 1.7 Å. This showed that, as anticipated from sequence analysis, AA7 is a single-domain αβ hydrolase. In addition to the basic αβ hydrolase fold, the enzyme contains six inserted helices, which are largely distributed around the substrate-binding regions of the enzyme. The enzyme contains a recognisable catalytic triad, composed of Ser–His–Asp, as is typical for these enzymes, with two distinct binding pockets that are bisected by the catalytic machinery. The structure of the enzyme in complex with a phosphonate-based tributyrin analogue suicide inhibitor was also solved by molecular replacement, showing that, with an inhibitor bound, the enzyme undergoes some conformational changes. These are primarily localised to helix D`, which, although bent into the substrate-binding site at the N-terminus in the native structure, becomes straightened upon binding of the substrate analogue in the inhibited structure, which removes the sidechain of Met 204 from its native position where it blocks most of the P2 binding pocket. With the catalytic machinery bisecting the binding cleft perpendicularly, the catalytic histidine appears to be positioned to activate the catalytic serine, as well as to activate a hydrolytic water that could enter the active site from either the P1 side or the P2 side. This bidirectional substrate binding is known to occur in Est2 from *A. acidocaldarius*, with the direction dependent on the substrate size, and recent data obtained in our group also appear to support this bidirectionality for AA7. The tunnel adjacent to the P1 binding site is postulated to be a binding site for an acceptor for the alcoholysis reaction. If this is the case, then alcoholysis could occur only when the substrate is oriented in such a way that the leaving group is in the P1 pocket, which, upon departure, would allow the acceptor to move forward from its binding tunnel to complete the deacylation of the enzyme and yield an ester product. This hypothesis, along with bidirectional substrate binding dependent on typoselectivity, appears to provide an explanation for the differences in hydrolytic and alcoholytic rates of the enzyme on substrates of different sizes. More experiments to prove this hypothesis are currently being undertaken in our group.

Chapter 4

The Structure and Mechanism of AZ4 Esterase

Chapter 4

The Structure and Mechanism of AZ4 Esterase

4.1 Introduction

AZ4 is a carboxylesterase from *Lactobacillus rhamnosus*, identified primarily by its sequence similarity to the αβ carboxylesterase Est30 from *Geobacillus stearothermophilus* (30% similarity) (Ewis, Abdelal *et al.* 2004). Prior to this work, AZ4 was cloned into a plasmid (pProEXHT-B, Invitrogen), with initial tests showing that it had hydrolytic activity on tributyrin agar, but not ester synthetic activity using tributyrin and an ethanol acceptor (M.-L. Delabre, personal communication, 2006). As the original scope of this project was to investigate the structural features of dairy lactic acid bacteria (LAB) that lead to substrate size preference, as well as hydrolysis/alcoholysis preference, AZ4 was chosen for study as an example of a hydrolase-only enzyme.

4.2 Methods

4.2.1 Expression of AZ4

AZ4 was obtained as a hexahistidine-tagged fusion construct using the plasmid vector pProEXHT-B (Invitrogen-Lifetech). This vector introduces a six-histidine purification tag to the N-terminus of the expressed protein, followed by a TEV protease cleavage site and a short (seven-residue) linker peptide. This allows the cleavage tag to be easily and specifically removed if required. The vector was transformed into *Escherichia coli* BL-21 DE3 cells, and transformant colonies were identified by their plasmid-conferred ampicillin (Amp) resistance. A single colony was used to inoculate 20 mL of Luria-Bertani (LB) Amp medium, which, after overnight growth at 37^{o}C, was used to inoculate 2 L of LB Amp medium (2% inoculum). This large culture was incubated at 25^{o}C, and protein expression was induced at an optical density at 600 nm (OD_{600}) of 0.5 with the addition of 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG). After shaking incubation overnight at 25^{o}C, cells were harvested by centrifugation (4000 x *g* for 25 min), resuspended in phosphate-buffered saline (PBS) and then lysed by French Press (American Instrument Co., two passes at 6000 psi).

4.2.2 Purification of AZ4

4.2.2.1 Immobilised Metal Affinity Chromatography

Cell debris was cleared by centrifugation (26000 x *g* for 40 min) and then the cell-free lysate was loaded on to a nickel-charged chelating sepharose column. Unbound protein was eluted using PBS, and then a wash of 50 mM phosphate buffer (pH 7.5) containing 20 mM imidazole was used to remove non-specifically bound proteins. 50 mM phosphate buffer (pH 7.5) containing 250 mM imidazole was then used to elute AZ4, along with some minor contaminants. The protein was dialysed overnight into 20 mM HEPES pH 7.1 to remove imidazole and change buffers.

4.2.2.2 Hydrophobic Interaction Chromatography

An additional step of hydrophobic interaction chromatography using phenyl superose (GE Healthcare) medium was carried out to remove contaminants that could not be removed by either immobilised metal ion chromatography (IMAC) or size exclusion chromatography (SEC) as shown by initial trials, as this was a highly effective step in earlier LAB esterase purifications (Holland and Coolbear 1996).

Finely ground NaCl was added slowly to the protein with gentle stirring, to a final concentration of 2 M. The protein was loaded on to an equilibrated column at 0.5 mL/min. A stepwise gradient from the loading buffer into Milli-Q water was applied, with a large protein peak that contained AZ4 protein being observed at a 5% salt buffer concentration (100 mM salt concentration). During subsequent preparations of the enzyme, protein elution was consistently observed between 75 and 150 mM salt concentration.

4.2.2.3 Removal of Fusion Tag and Size Exclusion Final Polishing

It is generally a matter of personal preference whether proteins are crystallised with or without small (for example, hexahistidine) fusion tags present (Derewenda 2004). In the case of AZ4, two attempts at crystallisation were made, with and without a fusion tag present. To remove the fusion tag, the protein was incubated with AcTEV protease (Invitrogen) according to the manufacturer's directions. Removal of uncleaved protein was attempted by a second pass through a nickel-charged chelating sepharose column. As cleaved protein was found to still bind to the column, this clean-up step was abandoned. Instead, protein was incubated until the removal of the fusion tag was complete (as judged by Coomassie-stained sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) analysis), and then concentrated and subjected to SEC using a Superdex G-200 10/30 Tricorn column (GE Healthcare). Injections of 400 μL of concentrated protein were used and the column was run at 0.5 mL/min for a run volume of 1.5 column volumes (25 mL). 0.8 mL fractions were collected. A single major peak was found to elute from the size exclusion column, with no indication of shoulder or minor peaks (Figure 4.1). Comparison of the retention times of AZ4 against calibration standards showed that AZ4 (Mwt 27000) was monomeric under the conditions used to run the column, as it co-eluted with Carbonic Anhydrase (MWt 29000).

Protein that was not subjected to fusion tag removal was concentrated after the hydrophobic interaction chromatography step, and then directly subjected to SEC. Protein that was to be co-crystallised with an inhibitor bound was reacted with the inhibitor prior to SEC. The chromogenic substrate assay (Shaw 1999) was used to monitor the inhibition of enzyme activity.

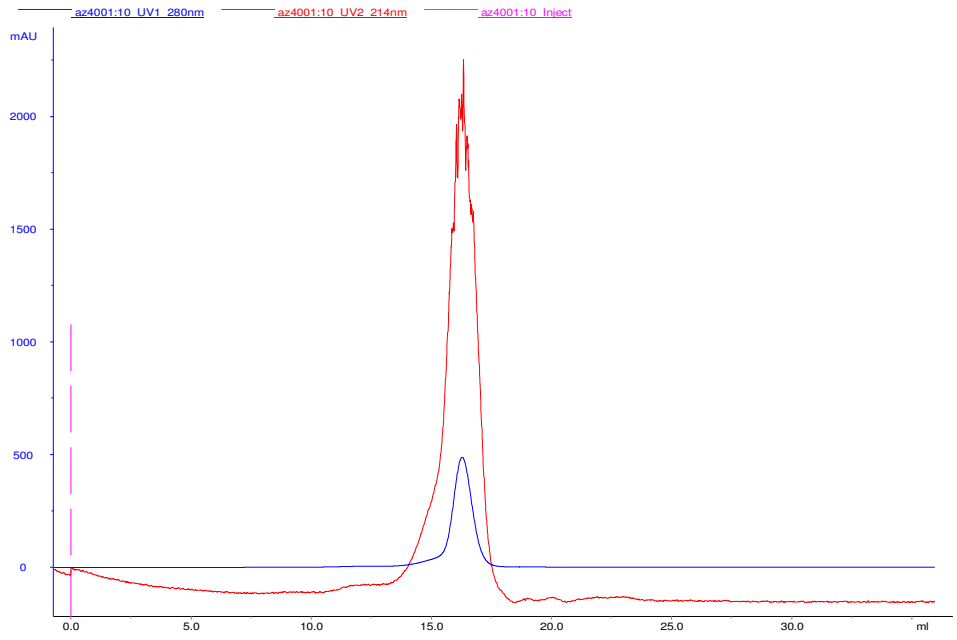


Figure 4.1 *Size Exclusion Chromatography trace of AZ4*

4.2.3 Crystallisation of AZ4

Crystallisation trials to identify initial lead conditions were carried out using Structure Screen I and Structure Screen II (Molecular Dimensions) (Jancarik and Kim 1991), in a round-well sitting-drop format, using Crystal Quick 96-position plates (Greiner Bio-one). A protein concentration of 10 mg/mL was used. Initial crystallisation trials were carried out using AZ4 that had extra N-terminal residues that included the hexahistidine purification tag and the TEV protease cleavage site conferred by the expression system. After optimisation of lead conditions, single crystals were obtained in several conditions. However, either these crystals did not diffract or the diffraction pattern indicated that the crystals were twinned.

A second crystallisation condition screening was performed using protein with the N-terminal hexahistidine tag cleaved off. After an initial overnight incubation, several crystal morphologies were observed, including crystals that were blade-like and cubic

crystals. A two-dimensional refinement screen was set up around these conditions, by varying the pH and the precipitant concentration, using 24-well VDX plates and a hanging-drop format. Two drops of protein were suspended over each well, using protein concentrations of 5 and 10 mg/mL. After a further 24 h, cubic crystals were observed, and these were tested for diffraction. Diffraction to 1.8 Å was observed. After another 5 days, further crystals were observed in a drop containing 5 mg/mL of protein. These crystals were significantly larger than but of a similar habit to the faster forming crystals, and were found to diffract to 1.4 Å. Conditions for growth were 100 mM sodium cacodylate pH 6.0, 15% PEG 8000 with 200 mM sodium acetate, 5 mg/mL protein.

4.2.4 Crystal Soaking

To bind inhibitors and other compounds to the enzyme, the crystals were soaked for between 1 h and several days, to allow time for the enzyme to interact with the compound of interest.

4.2.5 Crystallography

Crystallographic methods were carried out as described in Appendix A. Procedures specific to this chapter are detailed below.

4.2.5.1 Crystal Mounting and Cryoprotection

Solutions containing mother liquor with increasing concentrations of glycerol were trialed. At 22% glycerol, the mother liquor was observed to freeze as a clear glass. Crystals were picked up from the cover slip in a cryoloop, soaked in a drop of the cryoprotectant for about 10 s, and then mounted directly into position on the goniometer in the cold stream from the low temperature device.

4.2.5.2 Data Collection

Data for the initial structure solution were collected from a single crystal. Weak diffraction was observed to high resolution (1.4 Å). The data were integrated and truncated to 1.6 Å, which gave an outer shell R_{merge} of 0.204. The Matthews coeffcient

was calculated as 2.10 for a single molecule in the asymmetric unit, giving a solvent content of 41.4%.

4.2.5.3 AZ4 Molecular Replacement

BLAST (Altschul, Madden *et al.* 1997) was used to search the Protein Databank to identify structures with a high sequence similarity to AZ4. Est30 from *G. stearothermophilus* (pdb code 1TQH) was identified as having significant sequence similarity, at 30% (Liu, Wang *et al.* 2004). Several approaches to gaining a molecular replacement solution were attempted. Initially the 1TQH coordinates, with waters removed, were used as a model for the program MOLREP (Vagin and Teplyakov 1997), with a search resolution maximum of 2.4 Å. No distinct solutions were found, and the highest scoring molecular orientation gave an uninterpretable density map. The SWISS-MODEL server (Schwede, Kopp *et al.* 2003) was used to build a model using the AZ4 sequence and the Est30 structure (1TQH). Attempts at solution using MOLREP were again unsuccessful.

The automated model building and molecular replacement program MRBUMP (Bahar, Ballard *et al.* 2006) was then used to generate a model, and to find a molecular replacement solution. Although the software did not output a solution, examination of the log file showed that a solution based on the co-ordinates 1R1D (Cuff, Zhou *et al.*, 2004, unpublished data) had been found that had an R_{free} 46%. This solution was considered marginal, but examination of the electron density maps generated from it showed that they were highly interpretable. Automated model building using ARP/wARP (Perrakis, Morris *et al.* 1999) generated a model that fitted the density apart from a single gap of five residues, which was built using Coot (Emsley 2004). A final model was produced after several rounds of rebuilding with Coot and refinement using REFMAC 5 (Murshudov, Vagin *et al.* 1997). Two regions of trigonal-shaped density were modelled as acetate molecules, because the mother liquor for crystallisation contained 200 mM sodium acetate. During refinement, the occupancies of the catalytic triad residues were set to zero to ensure that possible changes in their conformation could be identified and modelled correctly. It should be noted that this measure was taken with all the subsequent models built based on this starting structure.

4.2.6 Model Quality

The quality of the model was assessed using PROCHECK (Laskowski, MacArthur *et al.* 1993) and Baverage (Dodson 1991) (Table 4.1). For the acetate and butyrate structures, all statistics were found to be within acceptable parameters. Interestingly, although the nucleophile of αβ hydrolases is usually found in a disallowed Ramachandran conformation, it was found in a helical conformation in these structures.

In the remainder of the structures, the structural statistics were also acceptable overall. However, in these structures, the nucleophile was found in the disallowed region of the Ramachandran plot that is typical of αβ hydrolases. Additionally, Ile 120 was found in a similar disallowed conformation. This residue is part of a loop that undergoes a conformational change during catalysis, which is likely to be the reason for its high energy mainchain torsion angles. This is discussed more fully in Section 4.4.

Data Collection Statistics					
Ligand	**Acetate**	**Butyrate**	**NaCl**	**NaCl/tributyrin**	**Paraoxon**
Resolution (Outer shell)	40.80–1.62 (1.71–1.62)	40.57–1.90 (2.01–1.90)	33.54–1.72 (1.80–1.72)	33.40–1.81 (1.95–1.81)	33.52–2.22 (2.54–2.22)
Space Group	$P3_221$ (No. 154) (angles 90° 90° 120°)				
Cell Dimensions a	47.11	46.85	47.44	47.15	47.32
b	47.11	46.85	47.44	47.15	47.32
c	174.48	173.95	174.18	174.23	174.74
Unique Reflections	70731	18460	23985	21223	11933
Total Reflections	377121	121639	106336	121354	106337
Completeness (Outer)	90.0 (100)	100 (100)	92.8 (95.9)	100 (99.8)	100 (100)
Redundancy	5.33 (1.45)	6.59 (6.75)	4.48 (4.44)	5.72 (5.99)	8.91 (9.0)
Mean I/σ	15 (2.6)	7.3 (2.5)	16.3 (5.8)	16.2 (7.4)	9 .0 (4.0)
R_{merge} (%)	0.057 (0.233)	0.079 (0.438)	0.047 (0.239)	0.049 (0.145)	0.099 (0.324)
Model and Refinement Statistics					
No. Residues Built	241	239	240	238	241
No. Water Molecules	258	107	243	168	93
Average B-factors, Å²					
Mainchain	21.731	41.851	24.158	26.710	38.284
Sidechain	22.578	42.438	25.171	27.417	38.477
Water molecules	32.461	45.765	34.745	34.279	41.470
Overall	23.391	42.354	25.826	27.672	38.550
Refinement					
Reflections used	29186	17408	22856	20296	11366
Reflections R_{free}	1528	928	1241	1078	561
RMSD bond lengths(Å)	0.006	0.013	0.007	0.006	0.005
RMSD bond angles (°)	1.035	1.361	1.030	0.981	0.865
R_{factor}	0.167	0.207	0.178	0.204	0.244
R_{free}	0.191	0.254	0.200	0.230	0.277
Ramachandran Analysis					
Most favoured	188 (91.3)	187 (91.7)	186 (91.2)	184 (90.6)	189 (91.7)
Additionally allowed	17 (8.3)	16 (7.8)	16 (7.8)	17 (8.4)	15 (7.3)
Generously allowed	1 (0.5)	1 (0.5)	0	0	0
Disallowed	0	0	2 (1.0)	2 (1.0)	2 (1.0)

Table 4.1 *Data collection, refinement and structural statistics for AZ4 structures.*

4.3 The Crystal Structure of AZ4

4.3.1 AZ4 Structures

In total, five structures of AZ4 were obtained: the native structure, as described above, and four derivatives that were obtained by soaking the crystals in mother liquor plus various compounds. These structures were: AZ4 with a chloride ion (from NaCl) bound in the active site; AZ4 complexed with acetate (the native crystals) and butyrate; AZ4 complexed with a phosphonate inhibitor (paraoxon); and AZ4 with a chloride ion plus tributyrin. Statistics from these data collections are summarised in Table 4.1. These derivatives were chosen as it was hoped that they would produce structures that were representative of the catalytic cycle of the enzyme.

4.3.2 The Overall Fold of AZ4

AZ4 has the αβ hydrolase fold that is typical for esterases, described by Ollis, Cheah et al. (1992). In this case, the enzyme has a seven-strand twisted β sheet, which is surrounded by seven α helices. It appears to form a homodimer in the crystal form, the dimeric form being generated by crystallographic symmetry from the single molecule in the asymmetric unit, as shown in Figure 4.2.

A standard αβ hydrolase fold has an eight-stranded β sheet, but the first strand, strand 1, is absent in AZ4, and strand 2 is short, spanning only three residues. This truncation or complete loss of β strands 1 and 2 is common in bacterial lipases (Heikinheimo, Goldman *et al.* 1999). There is an additional short helix bracketed by an ordered loop region between sheet 4 and helix B. The most significant change in the topology of AZ4, compared with the canonical αβ hydrolase, is a relatively large insertion after strand 6, consisting of an additional helix (helix D`) comprising residues 127–142. Helix D, which follows helix D`, is significantly larger than the other helices in the structure, spanning 27 residues (148–175), as well as being much longer than that found in the prototype fold. There is also a distinct kink towards the N-terminal end of this helix, suggesting that it is due to an additional helix that has been inserted at the end of the usually shorter helix D.

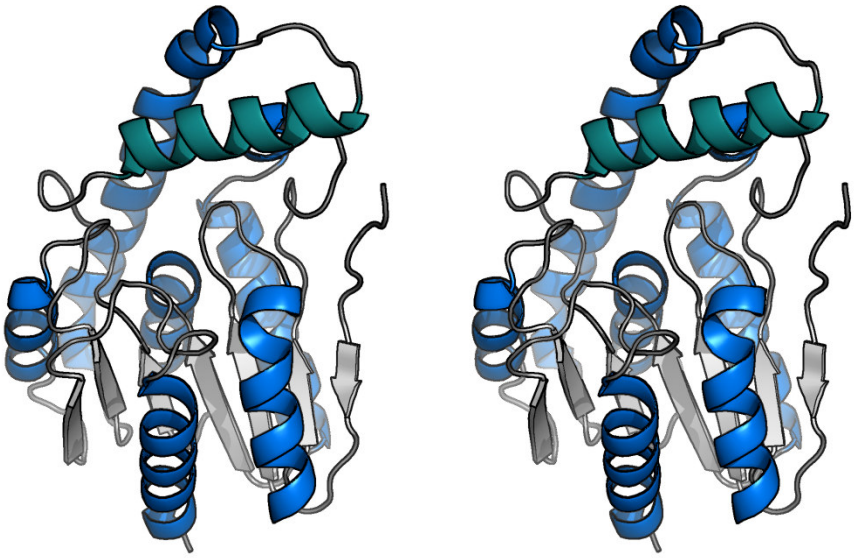


Figure 4.2 *Stereoview of the overall fold of AZ4 monomer. The helix that is additional to those found in the basic αβ hydrolase fold, helix D`, is shown in teal.*

Together, these produce a helical "decoration" on the top of the main αβ body of the enzyme. Insertions between strand 6 and helix D are common in αβ hydrolases (Ollis, Cheah *et al.* 1992). In this case, the inserted helix and the extended D helix are involved in forming the dimerisation interface between AZ4 monomers. The loop linking strand 6 with helix D` undergoes a conformational change during the catalytic cycle, implying that the insertion may be important for enzyme function, as well as dimerisation.

4.3.3 Catalytic Triad and Oxyanion Hole

Like other αβ hydrolases, AZ4 possesses a catalytic triad, containing a serine nucleophile (Ser 94), a histidine base (His 219) and an aspartic acid (Asp 188). Although well separated in terms of the protein's primary sequence, they are in close proximity in the folded enzyme. The order of catalytic residues in the primary sequence is nucleophile–acid–base, the normal order for these residues in all αβ hydrolases. In addition to the catalytic triad residues, AZ4 also contains an additional serine at the end of β strand 6 that is heavily conserved in αβ hydrolases (Thomas,

Church *et al.* 1999) as well as serine proteases (McGrath, Vasquez *et al.* 1992). This serine is postulated to have a role in maintaining the correct orientation of the acidic residue in the catalytic triad (Stok, Goloshchapov *et al.* 2004).

The catalytic serine is found in a tight γ-like turn, to allow the correct orientation of the nucleophile with respect to the substrate scissile bond. The other two triad members are also in the appropriate positions to promote the proton transfer that is necessary for catalysis to occur. In the structure with paraoxon bound, a tetrahedral species that is covalently attached to the sidechain of Ser 94 is observed, confirming that this is the reactive residue. In the acetate and butyrate bound forms of AZ4, Ser 94 assumes a helical conformation because its mainchain amide is involved in an interaction with the functional group of the carboxylate ligand, rather than the strained mainchain torsion, a change that has not been reported for other αβ hydrolases. This unusual state is discussed later in this chapter.

Adjacent to the catalytic triad, opposite the histidine, are two backbone amides that make up the oxyanion hole. One amide is contributed by Leu 95 (found immediately after the catalytic serine) and the other is contributed by Tyr 24 (located in a loop region at the top of ß strand 2). These amides are oriented to form hydrogen bonds with the tetrahedral transition state of the reaction, allowing the enzyme to bind preferentially to the transition state of the substrate, promoting catalysis. Such an arrangement of hydrogen donors in the oxyanion hole (where one amide is contributed from an adjacent mainchain residue and one is contributed from the top of the α helix immediately after the nucleophile elbow) is seen in a number of αβ hydrolases, and contrasts with an alternative arrangement found in AA7 esterase (Chapter 4) and other hormone-sensitive lipase (HSL) family members, where the adjacent mainchain β strand contributes two hydrogen-bond mainchain amide donors. Although AZ4 does not contain the HGGG oxyanion hole motif seen in the HSL family, it still has a hydrogen bond between the mainchain carbonyl of Ala 23 and the Nδ1 of His 22 (2.84 Å) to help to orient the oxyanion hole amide of Tyr 24 towards the catalytic site.

4.3.4 Binding Cleft

The catalytic serine of AZ4 is located at the entrance to a deep cleft, bounded by the main enzyme body on one side and by the D` and D helices on the other side. Along the cleft past the serine, buried in the main body of the enzyme, is a significant cavity. This cavity is in an "open state" in structures in which acetate has been displaced by paraoxon or a chloride ion, but becomes closed during the formation of a carboxylate product, represented in the structures with acetate and butyrate bound.

There are a significant number of hydrophobic sidechains in and around the active site of the enzyme, as was found in the other esterases studied. Some residues that line both the cleft and the cavity in the "open state" of the enzyme also have the potential to be involved in hydrogen bonds with the substrate. The residues lining the open cleft and the cavity are shown in Table 4.2.

Hydrophobic (Sidechains)	**Potential Hydrogen Bonds**	
Tyr 24	Ser 94	O
Leu 95	Gly 97	NH
Ile 98	Ile 98	NH
Leu 127	Ser 117	O
Leu 159	Ser 118	O
Ile 162	Ile 120	NH
Ala 166	Gly 123	NH
Val 191	Asp 163	Oδ1

Table 4.2 *Potential residues for interactions in AZ4 cavity.*

4.3.5 Quaternary Structure

Whilst gel filtration studies shown that AZ4 appears to be a monomer in solution, AZ4 does form a dimer-like association in crystalline form, as shown in Figure 4.3. The crystal interface is formed between the N-terminal ends of helices D` and D, and the corresponding helices on the other molecule of the dimer. The molecular interactions in the interface, which are both hydrophobic and electrostatic in nature, are summarised in Table 4.3.

Molecule A	Molecule B	Type of Interaction
Glu 136	His 126	*Salt bridge*
Pro 129	Pro 129	*van der Waals'*
Leu 152 Val 128	Val 128	*Hydrophobic*
Leu 132	Val 128	*Hydrophobic*
Leu 152	Pro 156	*van der Waals'*

Table 4.3 *Crystal psuedodimer interactions in AZ4 by residue. Molecule A represents any molecule and Molecule B is the other molecule. As the dimerisation interface is symmetrical, interactions between molecules A and B of the dimer are reciprocated between molecules.*

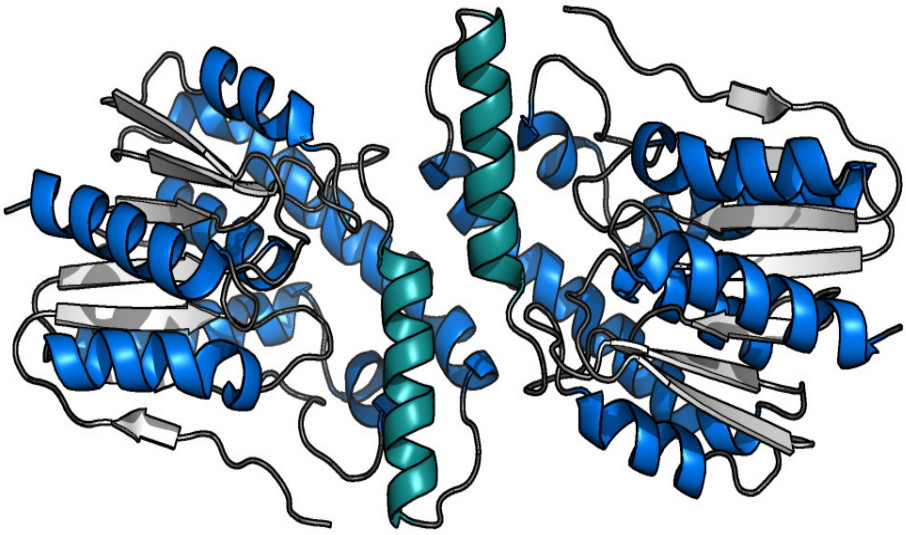


Figure 4.3 *Overall crystalline pseudodimer, showing the interactions between the C-terminus of helix D` (front in aqua) and the extended helix D (at rear) in dimerisation.*

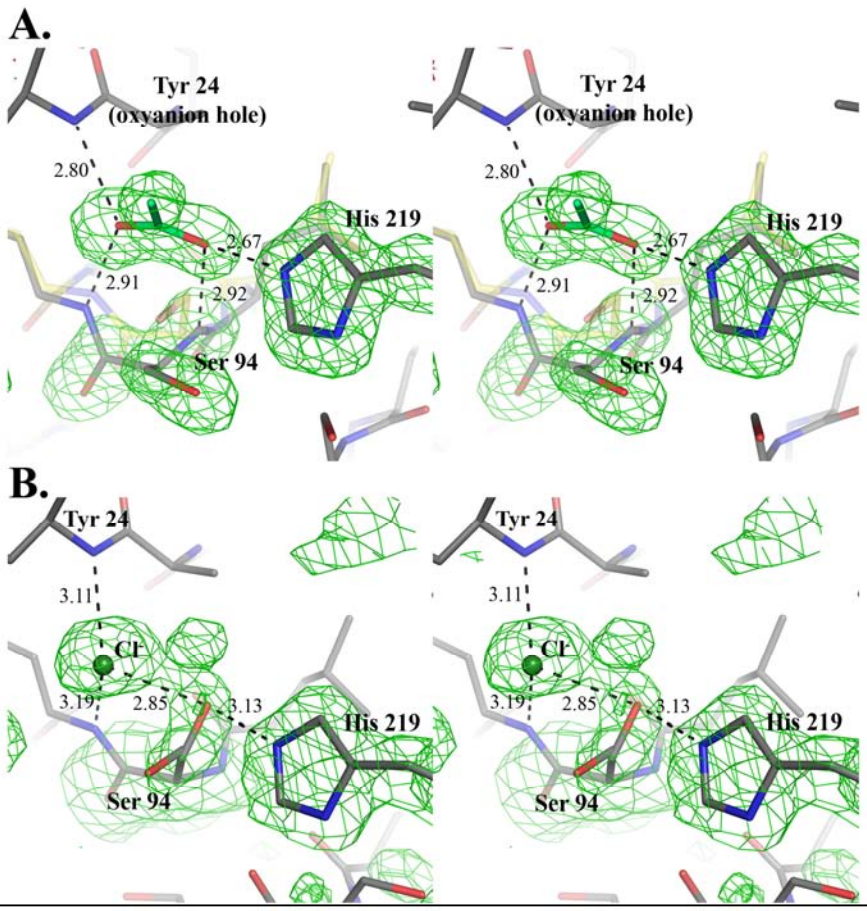


Figure 4.4 *Stereoview of the catalytic site of AZ4. Fo-Fc maps are shown, and were computed with the occupancy of the active site residues and any ligands set to zero. All distances are shown in Å.*

Panel A: With acetate bound (σ = 3.0). The nucleophile elbow of the NaCl bound structure has been shown in transparent yellow, to demonstrate the different cinfomration found with a carboxylate bound.

Panel B: With Cl^- bound in the presence of tributyrin (σ = 1.8). The split conformations of Ser 94 are shown, along with the weak hydrogen bond with His 219.

4.3.6 AZ4 with Acetate Bound

The initial structure of AZ4 was obtained at a resolution of 1.4 Å with all regions of the polypeptide chain clearly defined, other than a few surface residues with sidechains that could be modelled into several different conformations. Two regions

of density representing trigonal molecules were observed. These were modelled as acetates because the mother liquor used in crystallisation contained sodium acetate. Whereas one acetate is sandwiched between the two molecules of the dimer at the interface, the second acetate is positioned in the active site, with one acetate oxygen 2.65 Å from the epsilon nitrogen of the catalytic histidine and the second acetate oxygen positioned in the oxyanion hole, 2.89 Å from the mainchain NH of Leu 95 and 2.80 Å from the mainchain NH of Tyr 24. Because of its location, and its similarity to a product of AZ4 (butyric acid), it is possible that this structure mimics the state of the enzyme at the end of the catalytic cycle, i.e. with a product molecule bound.

4.3.7 AZ4 with Butyrate Bound

A structure of AZ4 with butyrate substituted for acetate was obtained by soaking the original crystals (with acetate bound) in mother liquor containing sodium butyrate (200 mM). This structure had clear electron density for a four-carbon butyrate ligand, when compared with the density for the acetate ligand. The density map remained unambiguous, even when refined with the occupancy of the butyrate set to zero (Figure 4.4). Other than the presence of the larger ligand, there appears to be no significant conformational change between the butyrate bound and acetate bound forms of the enzyme.

4.3.7.1 Conformation of AZ4 with Butyrate

The butyrate binds in the active site with the butyl chain pointing upwards, but slightly into the substrate-binding cavity. This suggests that the enzyme might bind its substrate in a similar way, with the acyl chain associated with the scissile bond in the cavity and the remainder of the glyceride outside, past the cavity opening.

It is of interest to note that the binding cleft and the cavity of AZ4 superficially somewhat resemble those of *Candida rugosa* high molecular weight lipase (CRL) (Grochulski, Bouthillier *et al.* 1994). In both enzymes, there is a large cleft on the enzyme surface, linked to a cavity inside the enzyme. The cavity of CRL is narrow and extends well into the interior of the enzyme, compared with the cavity of AZ4, reflecting the preference of CRL for larger glyceride substrates. An examination of

the surface of the catalytic site cleft in both enzymes shows some interesting similarities (Figure 4.5). Although the hydrophobic surface on the outside of CRL is more extensive than that of AZ4, both enzymes have two distinct "platforms" on either side of the substrate-binding cleft. Modelling studies of substrate-binding interactions in CRL suggest that triacylglyceride binds in a "tuning fork" conformation (Grochulski, Bouthillier *et al.* 1994), and, given the similarities in the cleft and the cavity of these enzymes, it is possible that AZ4 binds substrates in a similar fashion.

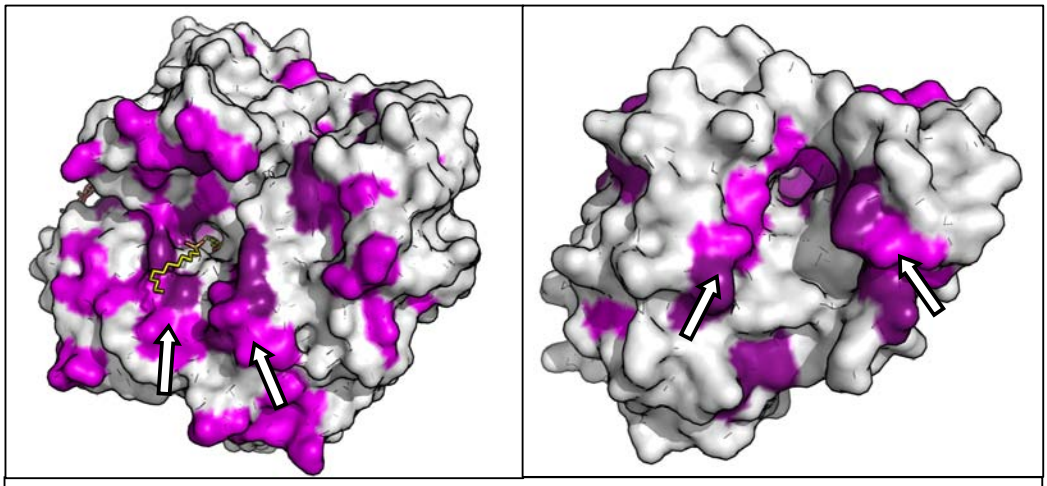


Figure 4.5 C. Rugosa *lipase and AZ4 esterase.*
Surface diagrams of CRL (left) and AZ4 (right). Water molecules have been removed from the structures. Hydrophobic residues (Leu, Val and Ile) are shown in pink and aromatic residues (Phe, Trp and Tyr) are shown in purple. CRL also has two molecules of a single-chain phosphonate-based acylglyceride mimic present (in grey). Arrows indicate the hydrophobic "platforms" in the cleft that are implicated in substrate binding (Grochulski, Bouthillier et al. 1994).

4.3.8 AZ4 with Paraoxon Bound

A structure of AZ4 complexed with the irreversible inhibitor paraoxon was solved. Paraoxon is an organophosphate-based inhibitor that mimics the tetrahedral transition state of the enzyme reaction because of the tetrahedral geometry of its phosphonate centre (Figure 4.6). Crystals of AZ4 were soaked in a solution of mother liquor containing 1% paraoxon (w/v). During initial soaking trials, a yellow colour developed around the crystals, indicating that the reaction product *p*-nitrophenol had been liberated. Data were collected from these crystals and the structure was solved

by molecular replacement using the AZ4 acetate structure as a model (Phaser; McCoy, Grosse-Kunstleve et al. 2005). Residues that produced negative density in the difference map were set to an occupancy of zero and re-refined, which gave a clear indication that some parts of the structure of the polypeptide chain had changed conformation, necessitating rebuilding.

4.3.9 Structure with Paraoxon Bound

The structure of AZ4 in complex with paraoxon shows clear density for a tetrahedral compound attached to Ser 94 in the active-site cleft of the enzyme, indicating the formation of a paraoxon adduct. Some density is present for the carbons that are attached to the phosphorus atom, but not for the carbons adjacent to these. This is likely to be due either to partial occupancy of the compound or to these carbons being mobile.

A large movement of an 11-residue loop occurs, which has the overall effect of opening up a deep hole immediately adjacent to the catalytic triad. This hole is lined with hydrophobic residues and appears to open to allow substrate to be accommodated by the enzyme. This movement is discussed in Section 5.4.2.

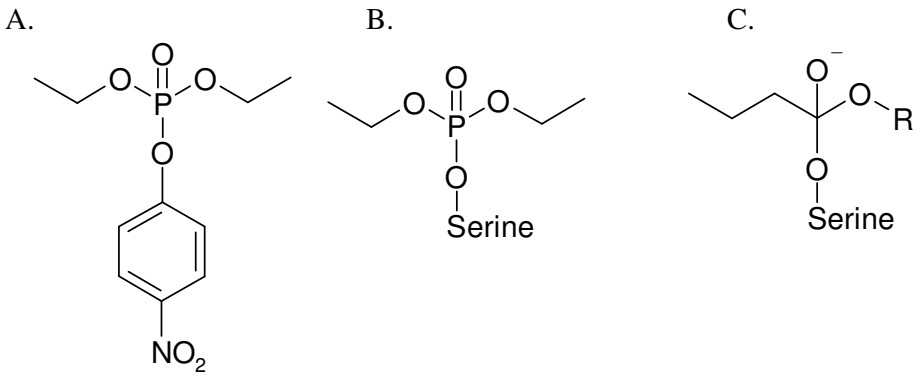


Figure 4.6 *(A) Paraoxon, (B) paraoxon acyl-enzyme complex and (C) serine butyl tetrahedral intermediate. B is generated by the nucleophilic attack of the enzyme on paraoxon (A) (Derewenda, Brzozowski et al. 1992). (C), the tetrahedral intermediate of a nucleophilic attack on a glyceride substrate, is shown for comparison.*

4.3.10 AZ4 with Chloride Bound

To remove acetate from AZ4, crystals of AZ4 were soaked in mother liquor in which the sodium acetate was replaced with the same concentration of sodium chloride (200 mM). Crystals were first dipped into this solution to remove any acetate mother liquor left in the loop during transfer, then soaked overnight in a fresh drop and suspended over a well of mother liquor that also contained NaCl instead of $NaCH_3COO$.

4.3.10.1 Conformation of AZ4 with Chloride

The structure of AZ4 with the acetate removed is virtually identical to that of AZ4 with paraoxon bound, implying that the structural conformation of the enzyme with paraoxon bound or with the acetate displaced by a chloride ion represents an "open" state of the enzyme that occurs when the enzyme is in an inactive state and persists through until product formation. This suggests that the observed "closed" state of the enzyme is induced by the formation of product – ions that have a carboxylate functional group. There was no electron density that could accommodate a trigonal ligand (i.e. acetate) in the electron density map of the NaCl-soaked crystal; however, there was a region of spherical density in the oxyanion hole that could accommodate a chloride ion. The chloride ion fits into the oxyanion hole where it forms hydrogen bonds to Leu 95 NH (2.31 Å) and Tyr 24 NH (3.17 Å). Halogen ions have been observed in this position in other αβ hydrolases, such as Antigen 85C from *Mycobacterium tuberculosis*, which has a F^- ion bound (Ronning, Klabunde *et al.* 2000).

This length is typical of N–H---Cl^- hydrogen bonds, which have been shown to have a mean length of 3.18 Å. This increased length compared with a non-halide bond has been shown to be due to the lengthening of the covalent N–H bond within the N–H---Cl^- bond (Steiner 1998). Unlike a carboxylate, Cl^- is not able to bind in both the oxyanion hole and to Ser 94 NH to induce the movement of the mobile loop that is seen in the acetate complex.

4.3.10.2 Replacement of Chloride in AZ4 with Acetate

In order to test the above hypothesis, the chloride ion was removed from a second crystal by soaking it in the original acetate-containing mother liquor overnight. The data collected from the crystal were solved using molecular replacement (Phaser;

(McCoy, Grosse-Kunstleve *et al.* 2005)). Density that could be fitted with a trigonal acetate molecule was observed in the same position as before, and there was a conformational change in the mobile loop, placing it back in the carboxylate bound position. To check for model-induced bias, the structure was solved by the same method, but using the model with chloride bound. In the $F_o - F_c$ difference map, this gave positive peaks that corresponded to the position of the mobile loop in the acetate bound structure, and negative peaks that corresponded to the position of the loop in the chloride bound structure, showing that the movement was genuine, and reversible.

4.3.11 AZ4 with Chloride and Tributyrin Bound

In order to further test the proposed mechanism, tributyrin, a known substrate, was soaked into the crystals. To do this, a solution containing 5% tributyrin plus 200 mM NaCl, to ensure that the acetate was displaced from the binding site, was prepared, and a crystal was soaked in 2 µL of this solution for 12 h. Data were collected as for the native AZ4 crystal and the structure was solved using Phaser.

The structure of AZ4 with acetate displaced by chloride in the presence of tributyrin is different from both the chloride/paraoxon conformations and the carboxylate bound structure, in both the conformation of the mobile loop and the conformation of the catalytic serine sidechain.

The mobile loop appeared to be in the "open" conformation. However, some density that was suggestive of the "closed" conformation was evident, particularly for Pro 119 and the sidechain of Ile 120, which supports the high mobility of these residues that is seen during the conformational change of the enzyme between the open and closed states.

However, Pro 119, which is positioned near the start of the mobile loop, has positive density in the difference map, but this positive density does not superimpose with the density representing proline in a closed position. Instead, it appears to be shifted slightly within the plane of the imino ring, rather than being flipped into the perpendicular position that it assumes in the closed structure. The density, at less than 1 Å from the mainchain density, is too close to the mainchain to be ordered water molecules. It is possible that it may represent an intermediate conformation of the proline that occurs during the switching of this loop from an open to a closed position.

As mentioned above, there is density that, on initial inspection, appeared to be associated with the sidechain of Ile 120, the residue that has the greatest lateral movement during the conformational change of the loop. However, no interpretable density is associated with the mainchain of Ile 120 or the residues adjacent to it in the polypeptide chain. It is therefore possible either that this density represents ordered water or that it is part of a reaction intermediate.

In this structure, the catalytic serine is found in two mainchain conformations. One of these is the "open" state conformation, in which the serine mainchain atoms are in an unfavourable conformation as judged by a Ramachandran plot. The second serine conformation is the same as that seen in the acetate bound closed state enzyme, which was seen in the $F_o - F_c$ difference map and only at a very low level of occupancy. Because of this low occupancy, this conformation was not built as an alternative conformation in this structure. This "closed" state conformation of the catalytic serine was not present in the structure containing the chloride ion, showing that this conformation is likely to have been induced by the presence of a tributyrin-derived product, rather than any "left over" acetate remaining after the crystals had been soaked in chloride-containing buffer. This shows that the structures with carboxylate ions bound do represent stages of catalysis, because they are also able to be induced by the presence of a substrate, as well as the enzyme product.

In the "open" state conformation, two rotamers of the Ser 94 sidechain are seen, the -155.33^{o} rotamer seen in the Cl^- structure, where the Oγ points in the direction of the substrate binding cavity, and a different conformation with a rotameric torsion of 31.96^{o}, which places the Oγ of the serine within hydrogen-bonding distance (2.99 Å) of the Nε of the catalytic triad histidine, pointing almost 180^{o} opposite from the chloride rotamer. The switching of the serine between these two conformations during catalysis may have a role both in activation of the enzyme (Kawasaki, Kondo *et al.* 2002), by bringing the nucleophile to within hydrogen-bonding range of the catalytic histidine during the start of the cycle, and in preventing the collapse of the tetrahedral transition state back on to the histidine NH, by breaking this hydrogen bond during establishment of the acylated enzyme intermediate (Radisky, Lee *et al.* 2006). The relative positions of the different Ser 94 rotamers are shown in Figure 4.7.

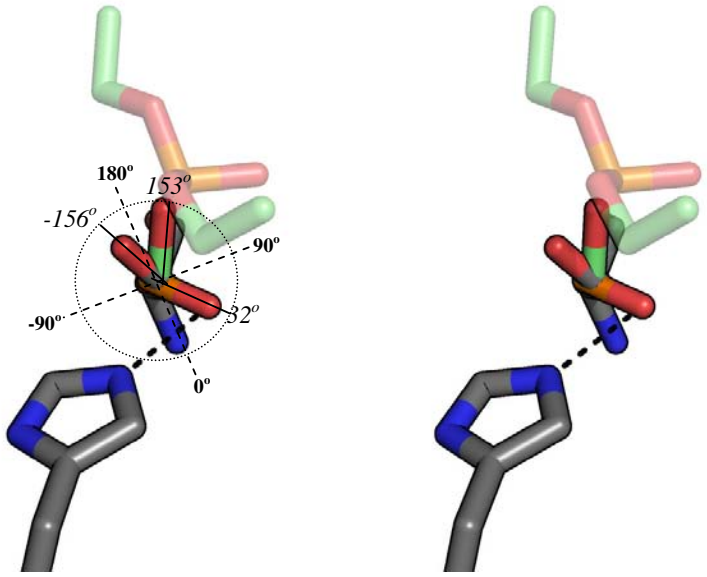


Figure 4.7 *Stereoview of rotameric conformations of Ser 94. The NaCl (grey), NaCl/tributyrin (orange) and paraoxon (green) structures have been superimposed and positioned to look directly along the Cα–Cβ bond. A key to the rotamer angles has been overlaid to show how they are constructed. The paraoxon molecule has been made transparent for clarity.*

4.4 Conformational Changes in AZ4 during Catalysis

Several conformational changes are apparent both in the catalytic triad and in a mobile loop stretching between residues 116 and 126 during catalysis. The catalytic serine, Ser 94, undergoes a small movement, which appears to be caused by a change in the Φ and Ψ angles of its mainchain atoms to an α-helical conformation, in contrast to the less favourable conformation that is normally assumed by the serine in the "open" enzyme. This is accompanied by a larger conformational change in a loop (residues 116–126) adjacent to the strand–loop–helix motif containing the nucleophile, and is responsible for the two major conformations of the enzyme, an

"open" state, in which the hydrophobic cavity next to the active serine is open to the solvent, and a "closed" form, in which the cavity disappears as a result of the movement of the mobile loop. The details of these movements are discussed in detail in the following sections.

4.4.1 Conformational Change in Catalytic Site

Logically, the enzyme would be expected to be in its open state without substrate bound to ensure unimpeded binding of the substrate. However, it is possible that this opening may be induced by the presence of a substrate, as occurs in αβ fold lipases during interfacial activation (Brzozowski, Savage *et al.* 2000). The current models of AZ4 and the similar enzyme Est30 (Liu, Wang *et al.* 2004) do not provide any conclusive evidence about which of these possibilities is correct, because neither structure was obtained without any ligand or ion in its respective active site. As AZ4 was found in the open state in all structures that did not have a carboxylate in the oxyanion hole, it is reasonable to assume that, in the absence of a product, the structure remains open, and that this conformation represents the start of the catalytic cycle.

The unfavourable conformation of Ser 94 in a disallowed region of the Ramachandran plot ($\Phi = 51.2$, $\Psi = -112.9$) is typical of the nucleophile conformation in αβ hydrolases (Ollis, Cheah *et al.* 1992), as well as in a number of serine proteases (Hakansson 2002). The sidechain of the serine in this position projects from the enzyme to allow it easy access to the substrate, and also allows the amide from the residue at the nuc + 1 position to be aligned so that it forms part of the oxyanion hole (Hakansson 2002). When a chloride ion is present, serine is observed in a single sidechain conformation, but, upon the addition of tributyrin, its sidechain is found in a second rotameric torsion, bringing the Oγ into close enough proximity to the Nε of His 219 to form a hydrogen bond (2.87 Å), thus increasing the nucleophilicity of the sidechain. These two distinct conformations may be a mechanism that has been adopted by the enzyme to ensure that catalysis is driven in a productive cycle, as detailed in Section 5.5.2.

As described earlier, the hydrolysis reaction catalysed by αβ hydrolases is believed to proceed in two steps. Firstly, there is a nucleophilic attack by the enzyme on a susceptible bond in the substrate to form a tetrahedral transition state that rapidly

degrades to an acyl-enzyme complex. A second nucleophilic attack by a water molecule forms the second tetrahedral transition state, which then degrades to form the carboxylate product and regenerate the enzyme. The details of this cycle that can be deduced from the different structures of AZ4 are discussed in Section 5.5.

The oxygen atom of the carboxylate product that is placed in the oxyanion hole after forming the oxyanion forms two hydrogen bonds, one with the amide of Tyr 24 (2.80 Å) and the other with the amide of Leu 95 (2.89 Å) (Figure 4.4).

The second carboxylate oxygen makes a hydrogen bond with the Nε of the catalytic histidine, but, more unusually, also hydrogen bonds to the mainchain amide of the catalytic Ser 94. In the open state of the enzyme, this amide is hydrogen bonded in a β-strand arrangement to the carbonyl oxygen of Phe 116 (3.06 Å). The Phe 116–Ser 94 hydrogen bond is broken as the serine mainchain amide flips almost 90^{o} to form a hydrogen bond with the carboxylate carbonyl oxygen. This results in the mainchain torsions Φ and Ψ of the serine nucleophile assuming a more favourable helical conformation. This movement of the catalytic serine into a helical arrangement from its usual disallowed Ramachandran conformation has not been previously reported in the αβ hydrolase family. The breakage of the Ser 94–Phe 116 hydrogen bond appears to be the event that initiates the conformational change of the mobile loop between Phe 116 and His 126 (Figure 4.9, upper panel).

4.4.2 Conformational Change in Mobile Loop

Upon breakage of the hydrogen bond between the amide nitrogen of Ser 94 and the carbonyl oxygen of Phe 116, the peptide between Phe 116 and Ser 117 rotates 109^{o}, breaking a β-sheet hydrogen bond between the NH of Ser 117 and the CO of Gly 183. This drives the sidechain of Ser 117 into an alternative conformation, with the result that the hydrogen bond between Ser 117 Oγ and Ala 185 NH is replaced by a hydrogen bond between the serine sidechain and an ordered water molecule. The peptide between Ser 117 and Ser 118 rotates by 90^{o}, whereas the sidechain of Ser 118 rotates 165^{o}, with the α carbon moving 2.08 Å away from the catalytic site. As a result, two residues, Pro 119 and Ile 120, rotate towards the binding pocket, with the α carbon of Ile 120 moving 5.31 Å. The best demonstration of the magnitude of this movement is that the δ carbon of the Ile 120 sidechain undergoes a movement of 8.79 Å to fill the active-site cavity, the movement involving a rotation of 160^{o} relative to

its initial position. The displacement of α carbons is summarised in Figure 4.8, with a structural comparison of the two positions of the loop shown in Figure 4.9 (lower panel).

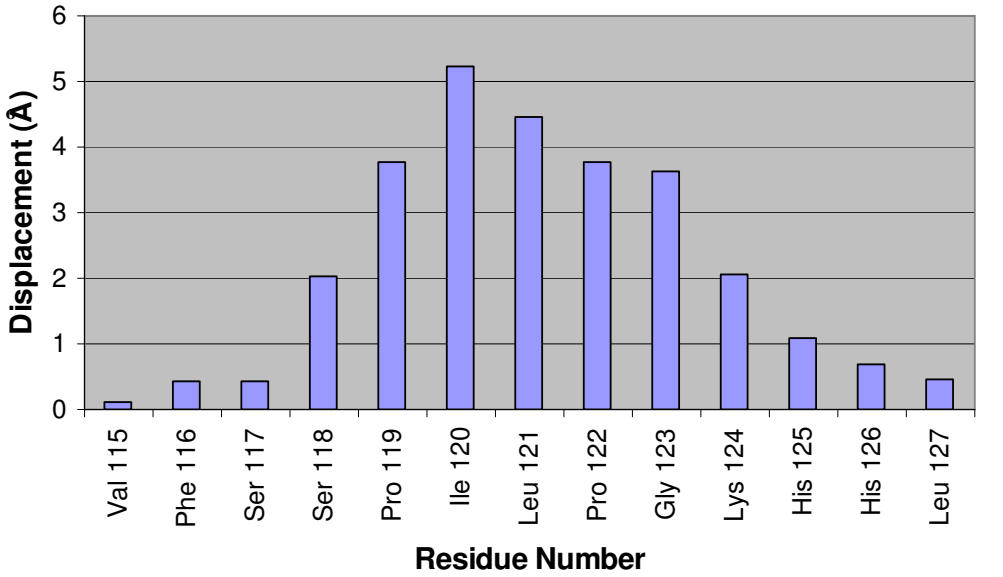


Figure 4.8 *Displacement in Å of α carbons in the mobile loop according to residue number.*

As well as breaking and reorganising hydrogen bonds, the conformational change in this loop results in changes in the hydrophobic interactions associated with Ile 120. In the open state of the enzyme, Ile 120 is in a pocket lined by Leu 199, Val 115, Met 101, Leu 173, Ile 98, Ala 166 and Ala 170. The conformational change forces Ile 120 out of this pocket and into the substrate-binding cleft. There is no apparent movement of these "lining" sidechains. As a result, the hydrophobic pocket previously occupied by Ile 120 in the open state becomes open to the external environment, but on the side of the protein opposite to the catalytic site. The α carbon of Leu 121 also moves 4.48 Å, pulling its sidechain after it, to expose the cavity vacated by Ile 120 to bulk solvent. In the closed position, Ile 120 associates with Ile 98, Leu 127 and Ile 162, completely filling the substrate-binding cleft.

Residues 121–126 are also affected by the movement of Ile 120, being dragged into new positions by the movement of the loop but without any conformational changes. Leu 127 is in an identical position and conformation in both the open and closed states of the enzyme, and so defines the end of the mobile loop.

A number of hydrogen bonds are broken during this conformational change but are not re-formed in the closed form. However, the closed conformation may be stabilised by the hydrophobic interactions that are formed between Ile 120 and the hydrophobic sidechains that surround the substrate-binding cleft. These may be adequate to temporarily stabilise the closed position, prior to product dissociation and reversion of the catalytic cycle to its initial conformation. The brief existence of the closed state is supported by the fact that Ser 94 in the tributyrin-soaked crystals is observed at a low occupancy in the closed conformation. The changes in hydrogen bonding during the transition from the open to the closed state are summarised in Table 4.4.

Open Conformation Bonds					**Bonds in Closed State**	
Loop Residue	**Atom**	**Bonding Partner**	**Atom**	**Length, Å**	**Residue**	**Length, Å**
Phe 116	NH	Gly 92	CO	2.69	No change	2.76
	CO	Ser 94	NH	3.06	Broken	
Ser 117	N	Gly 183	CO	2.70	Broken	
	Oγ	Ala 185	NH	3.43	Broken (a water bonds to Ala 185)	3.06
Ser 118	NH	Water 18	OH	2.88	Gln 184 Oε1	2.83
	Oγ	Water 42	OH	2.88	Broken	
Ile 120	NH	Water 42	OH	3.00	Broken	
Lys 124	CO	Water 69	OH	2.64	Asp 192 NH	3.00

Table 4.4 *Changes in hydrogen bonds in mobile loop during conformational change.*

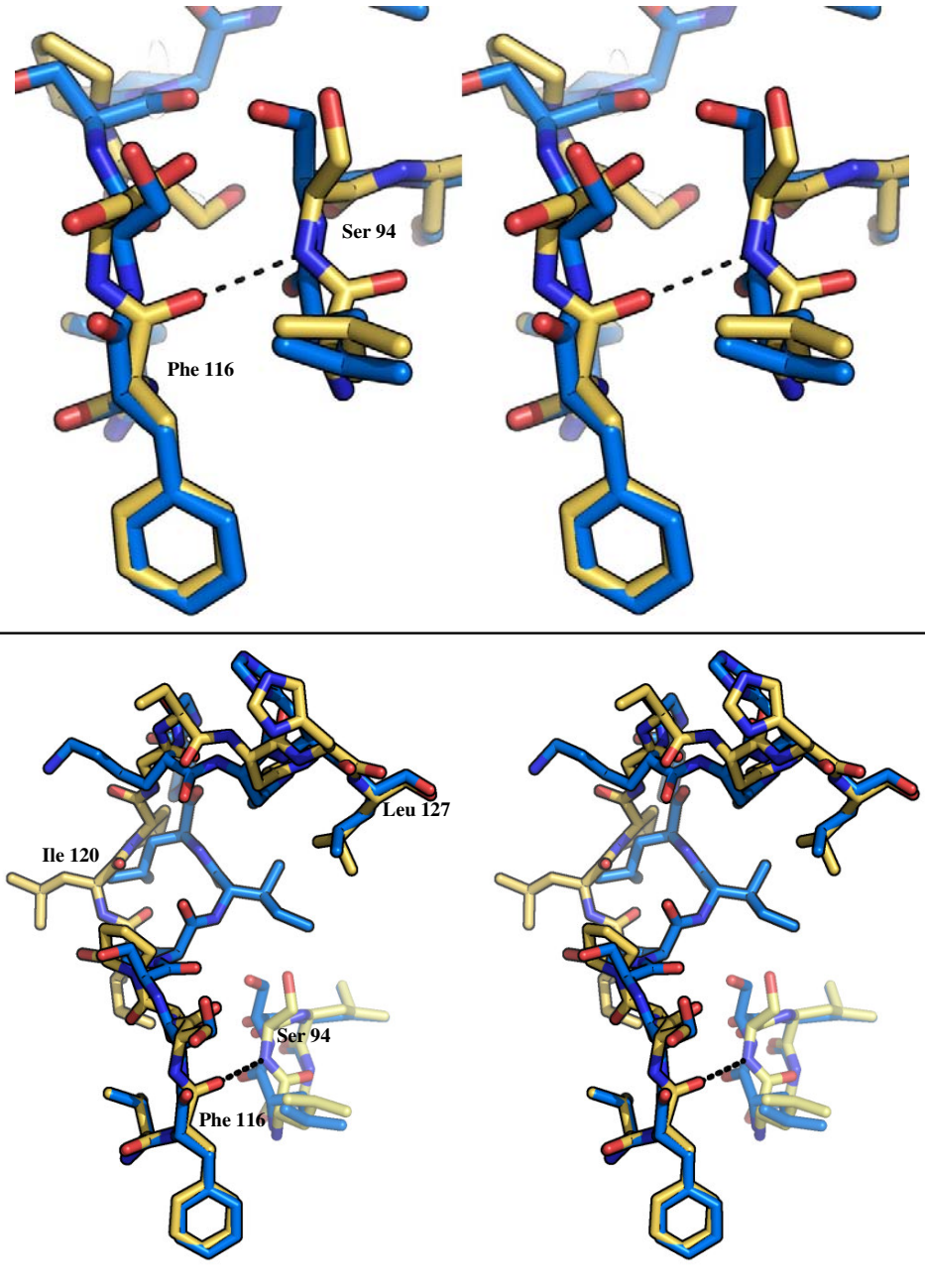


Figure 4.9 *Stereo comparison of the mobile loop in chloride bound open (yellow) and acetate bound closed (blue) forms of the enzyme.* *Key residues are numbered. The hydrogen bond between Phe 116 and Ser 94 in the open conformation is shown, with the adjacent nucleophile elbow. This also shows the conformational change undergone by the serine during the change from the open to the closed state, which appears to be driven by the breakage of this bond. Upper panel: Close-up of Phe 116 and Ser 94. Lower panel: Overview of loop conformational change.*

4.4.3 A Role for the Conformational Change in the 116–126 Loop

The net result of the conformational change observed with both acetate and butyrate bound in the catalytic site is the closure of the substrate-binding cavity, via the movement of the loop containing residues 116–126, as shown in Figure 4.10. The role of this movement is not yet clear, but the data obtained suggest that the formation of the product causes the enzyme to move into the closed formation, which would displace any compounds from the substrate-binding cavity. It is therefore possible that this movement may have a role in ejecting the reaction product from the hydrophobic binding cavity at the end of catalysis.

It is noted that, in the structure containing butyrate, the acyl chain of the product is positioned across the mobile loop cleft. In products with a larger acyl group than the acetate and butyrate used in these structures, it is conceivable that the acyl group might project into the cleft, which would mean that the product would be displaced as the cleft closes. The movement may therefore be an "auto-eject" mechanism, in which the product induces its own displacement from the enzyme. Interestingly, there is room for a longer carboxylate up to approximately eight carbons in length to be bound in the pocket in the closed state. Modelling studies indicate that a longer substrate would have to project either into the open substrate-binding pocket or outwards on to the surface of the cleft, suggesting that only substrates larger than C8 might utilise this auto-eject mechanism. In the structural study of the lipase from *C. rugosa* (Grochulski, Bouthillier *et al.* 1994), the two acidic residues near the mouth of the substrate-binding cavity were predicted to assist in the removal of the fatty acid product from the hydrophobic tunnel. AZ4 does not have acidic residues in the equivalent position. In AZ4, the only positionally equivalent residue to either of the CRL acidic pair is Leu 93. Therefore, a physical displacement mechanism such as the mechanism hypothesised may be necessary to assist the dissociation of the carboxylate reaction product.

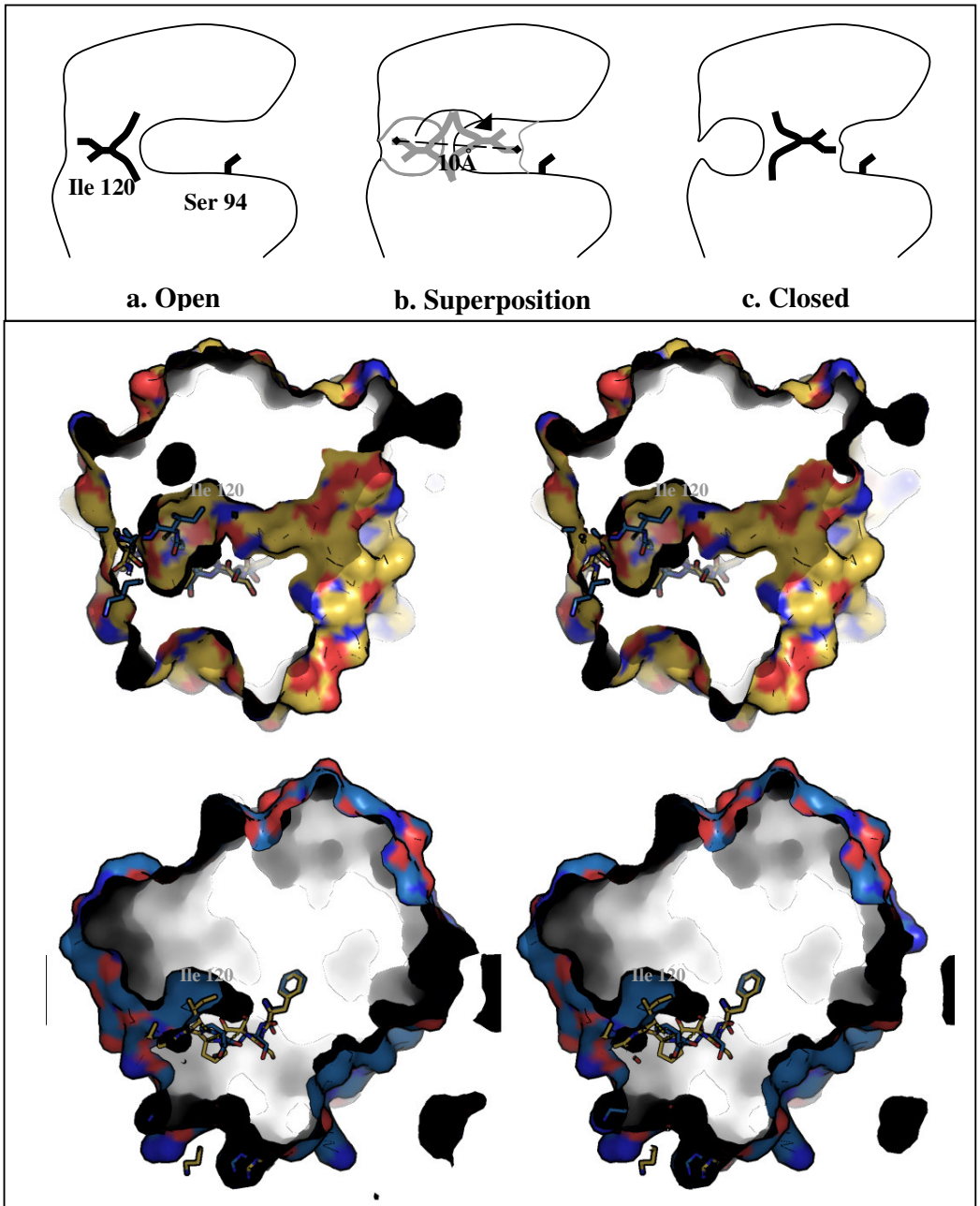


Figure 4.10 *Ile 120 displacement during the conformational change from the open to the closed state.*

Top: Schematic diagram of the movement. Bottom: Stereodiagram cutaways of the movement. The acetate bound (closed) structure is shown in yellow, and the NaCl bound (open) structure is shown in blue. Sticks are shown to illustrate the different position that Ile 120 adopts in each state. The two cutaways are shown from slightly different perspectives to demonstrate how the Ile 120 sidechain alternates between two voids in the protein in each conformation.

4.5 Catalytic Cycle of AZ4

Analysis of the various structures obtained with different ligands bound suggests a model for the catalytic cycle of AZ4. From the observations, evidence of the different conformations of the serine nucleophile, and the movement of a mobile loop associated with these different conformations, has arisen.

4.5.1 Acylation of Enzyme

The enzyme begins the catalytic cycle in what has been termed the open state, represented by the NaCl-soaked crystal structure (Figure 4.11, upper panel). An "induced-fit" mechanism whereby this conformation is induced by association with a substrate can be ruled out because of the observation of the enzyme in this open conformation when a chloride ion is present in the active site. This chloride, as a single ion, does not have any of the hydrophobic properties of an acyl chain of an ester or triglyceride. It binds into the oxyanion hole, but, because of its size, cannot promote the hydrogen-bond formation with the mainchain atoms that leads to the change in conformation to the open state. It is therefore reasonable to assume that the open state is the resting state of the enzyme, as this is the conformation that is found when no hydrogen-bond acceptor for the serine mainchain amide is available. In the resting state, Ser 94 assumes an uncommon conformation (the Cα–Cβ 155.33^{o} rotamer respective to the bond formed from NH–Cα–Cβ–Oγ). This results in the serine sidechain pointing towards the open cavity and is close to the least common rotamer of serine (179.00^{o}, found in 23.54% of cases). In this conformation, the Oγ of the serine is too far from Nε2 of His 219 to form a hydrogen bond and therefore is unactivated. Nevertheless, this conformation must be the most favourable for the serine sidechain to adopt in this particular state of the enzyme, probably because the other rotamers would bring the serine sidechain close to the mainchain amide of the following residue.

In the model produced from data collected from crystals that were soaked in both NaCl and tributyrin, the serine is present in two conformations, both of which are different from the conformation seen in the resting state (Figure 4.11, lower panel).

One conformation represents the favourable helical mainchain conformation seen in the structure with carboxylate present, although this is represented at a low occupancy. The other conformation has a higher occupancy and represents the serine with the same mainchain angles as the closed state, but with a split sidechain conformation, one with a rotameric torsion of 31.96^{o} and one the same as that seen in the resting state of the enzyme. This new conformation places the sidechain above the peptide bond preceding the serine α carbon. Because the peptide bonds with mainchain atoms at the top of β strand 6, this peptide bond is horizontal to the serine sidechain, minimising any potential steric clashes with the amide of Ser 94. This second conformation brings the serine oxygen to within 3.16 Å of the Nε of His 219, which is close enough to form a hydrogen bond. This bond will be weak because its angle is 94.09^{o}, but is not dissimilar to the hydrogen-bonding distance observed upon nucleophile rotation in other αβ hydrolases (Kawasaki, Kondo *et al.* 2002). It is also possible that, during activation of the nucleophile, the serine transiently rotates even closer to His 219, forming a stronger bond that allows proton transfer to occur.

4.5.1.1 Possible Hydrolytic Water-binding Site

It is also possible that a nearby water molecule (water 113), located 2.99 Å from the Nε of His 219 and 2.62 Å from the Oγ of Ser 94, may participate in this hydrogen-bond network by forming a bifurcated hydrogen bond.

4.5.1.2 Effects of Mainchain Torsions on Nucleophile Activation

In addition to allowing the nucleophile sidechain to project towards the substrate, the high energy mainchain conformation of the nucleophile contributes by allowing the serine sidechain to easily adopt rotamers that might be energetically unfavourable at normal mainchain angles. This would appear to be particularly important in the activation stage of catalysis, which requires the rotation of the serine sidechain into a position that is closer to the histidine than in the resting state. Interestingly, because it was seen only in the tributyrin-soaked structure, the rotamer that the nucleophile adopts to become activated must be higher in energy that the resting state rotamer. This observed change in sidechain conformation is probably induced by the binding of substrate into the active site, and, although a definitive mechanism by which this

occurs is unclear from substrate-modelling studies, it may be a result of steric interference between the serine sidechain and the bound substrate. However, if the nucleophile was in the helical mainchain conformation seen in the acetate bound state of AZ4, its mainchain amide hydrogen would project upwards and would sterically clash with the sidechain, preventing rotation to this position entirely. This suggests that, as well as allowing easy access of the nucleophilic sidechain to the substrate during catalysis, the disallowed mainchain conformation angles of αβ hydrolase nucleophiles are probably involved in promoting activation of the nucleophile during catalysis by allowing a reasonable freedom of rotation. This theory has support from the structure of lipase from *Bacillus subtilis* (Kawasaki, Kondo *et al.* 2002), the nucleophile of which is in a similar split conformation, with only one in hydrogen-bonding range of the catalytic histidine.

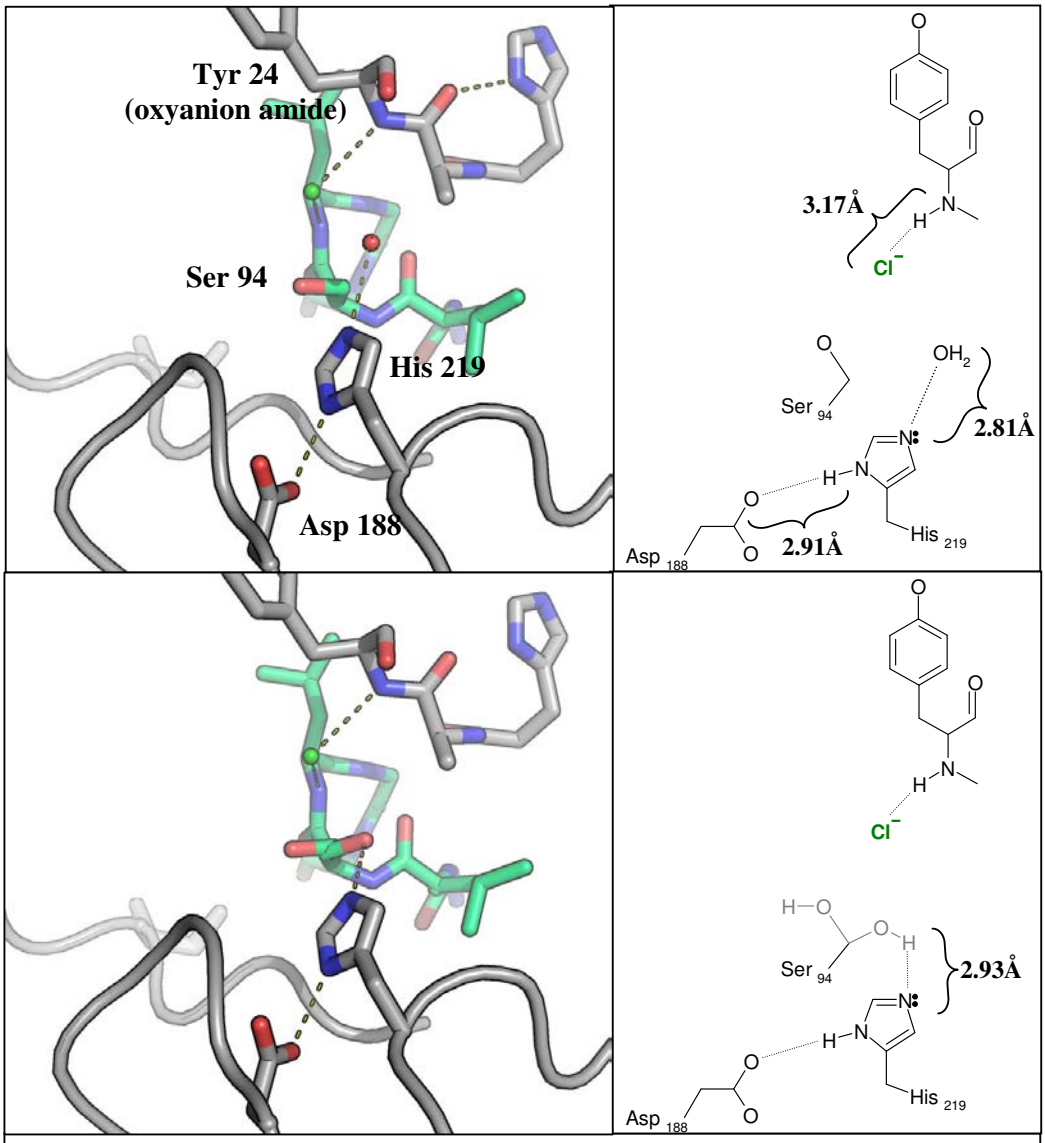


Figure 4.11 *Resting and activation states of AZ4. Hydrogen bonds are represented by dashed lines. Water 113 is shown as a red sphere and Cl^- is shown as a green sphere. Top panels: AZ4 in the resting state structure, obtained from crystals soaked in sodium chloride. Bottom panels: AZ4 with a split occupancy serine nucleophile sidechain (shown in grey in the chemical representation), obtained from crystals soaked in sodium chloride and tributyrin. Hydrogen-bond distances are shown in the chemical diagram. Distances that do not change significantly are not shown in subsequent diagrams.*

4.5.2 Tetrahedral Transition State

As the tetrahedral intermediate is likely to be unstable and short lived during the hydrolysis of an authentic substrate, it is not surprising that it was not seen in the tributyrin-treated crystal structure. The paraoxon bound structure provides a representation of this stage of catalysis, because its tetrahedral phosphate forms a covalent bond to Ser 94 Oγ (Figure 4.12, upper panel). In this structure, the serine sidechain has rotated approximately 50^{o} from its position in the resting state of the enzyme (to the -153.34^{o} rotamer). As expected, it is also covalently bound to the phosphonate group of the paraoxon, mimicking the tetrahedral intermediate of the catalytic cycle. The other unbonded oxygen of the organophosphate sits in the oxyanion hole, whereas the two two-carbon acyl chains of the inhibitor are bound along the substrate-binding cleft. As well as this being a more preferred conformation for the serine sidechain to adopt (inferred from the fact that this rotamer is observed only when Cl^{-} and paraoxon are present in the binding cleft), the rotation of the serine sidechain about the Cα–Cβ bond may play two roles in the catalytic mechanism. Firstly, it moves the serine out of hydrogen-bonding range from the catalytic histidine, leaving the "nucleophile activation" pocket unoccupied. As a result, a water molecule can enter the pocket and become activated to perform the second nucleophilic attack necessary to deacylate the acyl-enzyme complex and liberate the second hydrolytic reaction product. Secondly, the movement of the serine back to this rotamer is probably necessary to push the equilibrium of the hydrolysis reaction towards completion. If the serine were to remain within hydrogen-bonding distance of the catalytic histidine, there would be nothing to prevent the tetrahedral transition state from reverting back to substrate plus enzyme, resulting in a futile cycle. Some catalytic triad enzymes, for example trypsin, prevent this situation by a subtle movement of the nucleophile of less than 1 Å away from the histidine, after the initial nucleophilic attack, to prevent hydrogen bonding (Radisky, Lee *et al.* 2006). It would appear that, in AZ4, where the distance changes from 3.09 to 3.63 Å (and to a less favourable bond angle) during the formation of the tetrahedral intermediate, the rotation of the serine sidechain may be used to break this essential hydrogen bond after the initial nucleophilic attack has occurred, forcing the equilibrium of the reaction towards the formation of product. This also removes the serine from being in close proximity to the histidine, allowing the proton that was removed from the serine by the histidine to be transferred on to the departing alcohol product. This proton

transfer between the catalytic nucleophile, base and product occurs in a cyclical nature, ensuring reaction directionality by a different mechanism from that observed in trypsin, and providing an interesting point of contrast to these otherwise very similar catalytic triads.

4.5.3 Acyl-enzyme Intermediate

After the collapse of the tetrahedral intermediate (represented in the paraoxon structure) the alcohol product of the enzyme is able to diffuse away from the binding site as it is no longer covalently bonded to the remainder of the substrate, which remains covalently bound to form an acyl-enzyme intermediate. In the case of tributyrin as a substrate, this alcohol would be glycerol, or a mono- or dibutyrin molecule.

A water molecule then enters the unoccupied nucleophile activation site, and forms a hydrogen bond with His 219 to become activated for a nucleophilic attack on the bound substrate. In the structure with NaCl bound, water 113 is located 2.81 Å from the Nε2 of His 219, in an ideal position to be activated to carry out the deacylation nucleophilic attack. However, this water is not seen in structures that are further along the catalytic cycle, because, during catalysis, this position would normally be expected to be occupied by the leaving group of the substrate, which must be able to hydrogen bond with the histidine to allow proton transfer at the end of the acylation step of catalysis. The leaving group would therefore be expected to displace this water, which is likely to be why it is not seen in the tributyrin-soaked and paraoxon bound structures. Although there is no structure in the study that represents an acylated enzyme, the presence of this ordered water in the NaCl “resting state” structure suggests that water could bind in this same position later in the catalytic cycle to allow its activation to perform deacylation of the enzyme. As in the first part of the reaction, this intermediate would decompose, breaking the acyl–enzyme bond and forming a carboxylic acid. It is not clear how the proton that is abstracted from the hydrolytic water is returned to the serine at the end of the catalytic cycle, or how this part of the reaction is driven in a productive direction. It may be that the acyl–enzyme bond is unstable enough that this process occurs without the movements that ensure the productive establishment of the acylated enzyme state.

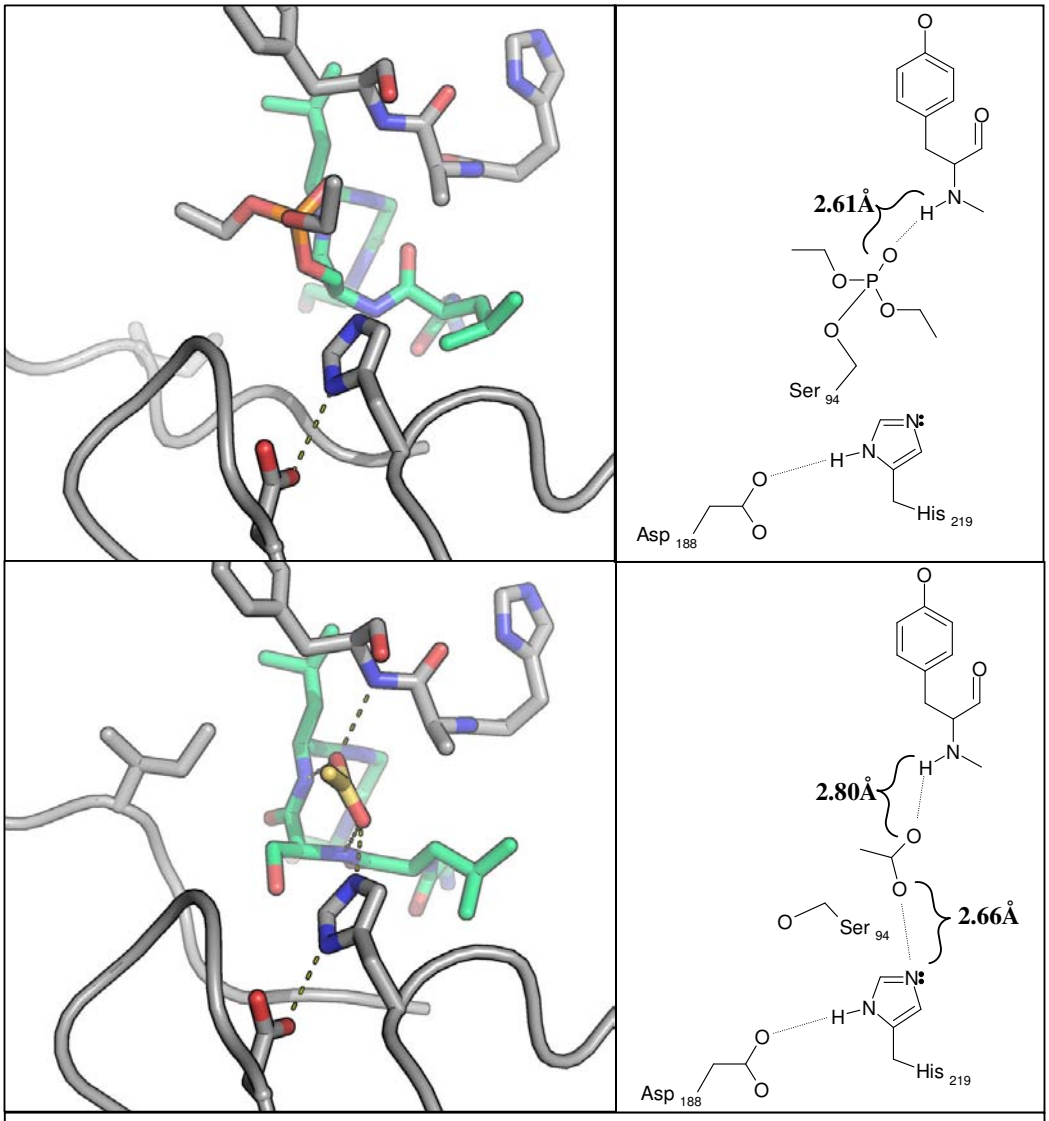


Figure 4.12 *Paraoxon and acetate bound structures of AZ4. Hydrogen bonds are represented by dashed lines. The second hydrogen bond of the oxyanion hole has been omitted from the chemical representations for clarity, but is shown in the structural representations. Top panels: Paraoxon bound structure of AZ4, obtained from soaked crystals. Bottom panels: Acetate structure of AZ4 obtained directly from crystals. The hydrogen bond between the acetate and the mainchain amide of Ser 94 has been omitted for clarity in the chemical representation.*

4.5.4 Product Bound State

In the structure with acetate bound in the reaction cavity, one oxygen of the acetate remains close to the oxyanion hole. The other oxygen is situated in a position that a water carrying out the second nucleophilic attack of the mechanism might be expected to occupy, at 2.66 Å from the Nε2 of His 219, although not the position in which water 113 is found in the Cl^- bound structure (Figure 4.12, lower panel). It is possible that the hydrolytic water is activated in the position represented by water 113 in the Cl^- structure, and moves to the position represented by the carboxylate oxygen during the collapse of the second tetrahedral state, because both these positions are slightly out of the plane of the His 219 sidechain, although on opposite sides of it. This shows that, during the reaction, some, although minimal, movement of atoms in the catalytic site is required, which is in agreement with the overall observations made about atomic movements during catalysis in trypsin (Radisky, Lee *et al.* 2006). However, in this publication, and in an earlier general study (Schramm and Shi 2001), it was noted that the nucleophile was generated in a fixed position, and then the substrate was brought into the required proximity with the nucleophile. In AZ4, the opposite appears to be the case, as the serine adopts one conformation where it becomes activated, and then adopts the second conformation during the formation of the tetrahedral intermediate, moving it closer to the substrate. Although there was no density for any substrate in the density maps calculated for these structures, it is clear that there is no need for the substrate to move in tandem with the motion of the serine. This provides a second point of distinction between the mechanism of action that has been established for serine proteases and the mechanism proposed here for αβ hydrolases.

4.5.5 Displacement of Product

As mentioned above, the presence of the carboxylate product causes the hydrogen bond from Ser 94 NH to Phe 116 CO to be broken, and replaced by a hydrogen bond to the carbonyl oxygen of the carboxylate product, triggering the movement of the mobile loop (residues 116–126). This change in conformation of the large loop, resulting from the small change in conformation of the serine, may result in the product of the enzyme reaction being ejected from the enzyme as the cavity closes (Section 5.4.3), although why this would be necessary rather than relying on natural

diffusion away from the binding site is unclear. It could be that the hydrophobic interactions between the product and the enzyme are too strong to be overcome by diffusion forces. Without knowledge as to the physiological substrate for this enzyme, it is difficult to estimate how strong these forces might be.

4.6 Comparison of the AZ4 Structure with Other Structures

4.6.1 DALI Structural Comparison

As well as the basis for the MRBUMP-generated initial molecular replacement model Est30 (Liu, Wang *et al.* 2004), identified using sequence similarity searching with BLAST (Altschul, Madden *et al.* 1997), other similar structures were identified using DALI (Holm and Sander 1996), which is a structural-similarity-based searching tool. Structures with significant similarity to AZ4 are summarised in Table 4.5. The models of AZ4 in complex with paraoxon or acetate were used as search targets.

PDB	Z score	RMSD	LALI	LSEQ2	%IDE	FUNCTION	SPECIES
1r1d-A	29.3	2.1	227	234	32	carboxylesterase	*Geobacillus stearothermophilus*
1c4x-A	17.5	3	195	281	14	2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase	*Rhodococcus* sp.
2hdw-A	17	3.5	199	321	13	hypothetical protein pa2218	*Pseudomonas aureginosa*
1brt	16.6	2.9	189	277	19	bromoperoxidase a2	*Streptomyces aureofaciens*
1ufo-A	16.4	3.2	187	234	16	hypothetical protein tt1662 (hydrolase)	*Thermus thermophilus*
1auo-A	16.3	2.2	164	218	16	carboxylesterase	*Pseudomonas fluorescens*
1q0r-A	16	3.5	199	297	18	aclacinomycin methylesterase	*Streptomyces purpurascens*
2fuk-A	15.9	2.6	172	218	19	xc6422 protein (hydrolase)	*Xanthomonas campestris*
1jfr-A	15.8	2.2	168	260	15	lipase	*Streptomyces exfoliatus*
1cr6-B	15.7	3.7	198	541	12	epoxide hydrolase	*Mus musculus*

Table 4.5 *DALI server summary for AZ4 acetate structure. Similar hits at ranks 1–3 were obtained for AZ4 paraoxon, but hits at ranks 4–10 were 1ufo, 1brt, 2fuk, 1auo, 2j2o and 1cr6 respectively.*

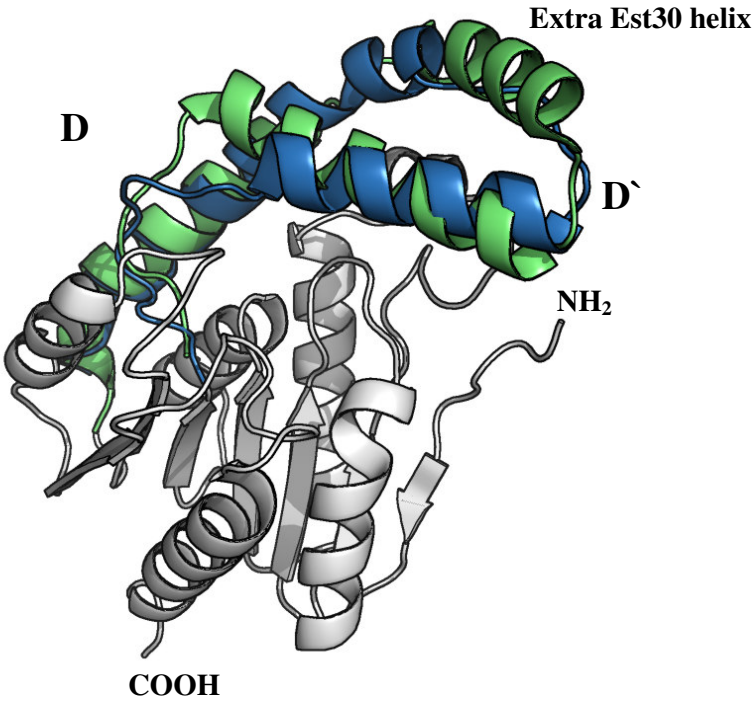
The coordinates for the top 10 hits for both the paraoxon structure and the acetate structure were loaded into Coot (Emsley 2004) and the SSM superpose function was used to overlay the structures shown in Table 4.5, which were then examined for regions of similarity.

As expected from the low RMS deviation, the overall structures were very similar. The structures all had the similar connectivity of the secondary structural elements that are characteristic of members of the αβ hydrolase family. Also, in all the structures examined, the nucleophilic serine was found to have the unfavourable mainchain torsion angles typical of the catalytic serine in these enzymes.

The most notable difference between all the structures was in the region of AZ4 from residue 116 to residue 126, which is located immediately after strand 6 in the structure, and forms the loop that moves during the catalytic cycle. All the structures appear to be different in this region because of different insertions, with the exception of the Est30 carboxylesterase from *G. stearothermophilus* (Liu, Wang *et al.* 2004), the apparent orthologue of AZ4 in this bacterial species. As mentioned before, as the region between strand 6 and helix D is reported as the most common location for insertions in the αβ hydrolase family (Ollis, Cheah *et al.* 1992), it is unsurprising that such a variability was seen.

Est30 has a similar mobile loop to AZ4 and a "cap" over the top of the substrate-binding cleft, and has a D` helix inserted between strand 6 and helix D. The lengthened helix D in AZ4 is split into two helices in Est30. In terms of three-dimensional space, the enzyme caps in both species occupy much the same area, because the loop between helix D` and helix D is shortened and replaced by the additional helix in Est30. Both the Est30 structures published have ligands bound to, or very near, the active site. One structure (pdb code 1TQH) contains a propyl acetate ligand, covalently bonded to the catalytic serine (Liu, Wang *et al.* 2004), and the alternative structure (pdb code 1R1D) contains a HEPES molecule, derived from the crystallisation mother liquor, in the active site (Cuff, Zhou *et al.* 2004). Both of these structures represent the open conformation of the enzyme. In the case of 1TQH, this was expected as the enzyme had been trapped in a state mid way through the catalytic cycle, and therefore must have been open to allow the binding of substrate. In 1R1D, the HEPES molecule is not covalently bonded to the enzyme, but rather is weakly hydrogen bonded (3.14 Å) between one of its phosphate oxygens and the Nε2 of histidine. However, there is no hydrogen bond between the HEPES molecule and the

NH of the serine nucleophile, the interaction that appears to be important in triggering the change in conformation from open to closed in AZ4.




```
                  2          3                A          4                B`
AZ4     6   PQPFEYEGTDTGVVLLHAYTGSPNDMNFMARALQRSGYGVYVPLFSGHGTVEPLDILTKG  65
            P+PF +E   +  V+LLH +TG+   D+   + R L+  GY   + P++ GHG    P + L
Est30   7   PKPFFFEAGERAVLLLHGFTGNSADVRMLGRFLESKGYTCHAPIYKGHGV--PPEELVHT  64

                     B              5          C          6
AZ4     66  NPDIWWAESSAAVAHMTAK-YAKVFVFGLSLGGIFAMKALETLPGITAGGVFSSPILPGK  124
             PD WW +        +  K Y K+ V GLSLGG+F++K   T+P      + +   + +
Est30   65  GPDDWWQDVMNGYEFLKNKGYEKIAVAGLSLGGVFSLKLGYTVPIEGIVTMCAPMYIKSE  124

                     D`                                  D
AZ4     125 HHLVPGFLKYAEYMNRLAGKSDESTQILAYLPGQ-----LAAIDQFATTVAADLNLVKQP  179
              +  G L+YA     +  GKS+E  +        Q     L A+ +     V   L+L+  P
Est30   125 ETMYEGVLEYAREYKKREGKSEEQIEQEMEKFKQTPMKTLKALQELIADVRDHLDLIYAP  184

            7               E                8                    F
AZ4     180 TFIGQAGQDELVDGRLAYQLRDALINAARVDFHWYDDAKHVITVNSAHHALEEDVIAFMQ  239
            TF+ QA  DE+++   A  + + + +  +    WY+ + HVIT++      L ED+ AF++
Est30   185 TFVVQARHDEMINPDSANIIYNEIESPVK-QIKWYEQSGHVITLDQEKDQLHEDIYAFLE  243
```

Figure 4.13 *Superposition of AZ4 with Est30 from* G. stearothermophilus. *Upper panel: The AZ4 core is shown in white, with the post strand 6 insertion shown in blue. The equivalent region from Est30 is shown in green. The αβ parts of the two enzymes were virtually identical when superimposed. Note how the bent N-terminal end of helix D in AZ4 corresponds to the additional helix found in Est30.*

Lower Panel: Pairwise alignment of AZ4 with Est30. Secondary structural elements have been highlighted, β-strands in black, and α helices in grey.

4.6.2 Structural Comparison with *Actinidia eriantha* Carboxylesterase

Recently, the crystal structure of AeCXE1, an HSL class carboxylesterase from *Actinidia eriantha* (kiwifruit), was published (Ileperuma, Marshall *et al.* 2007). Despite no significant similarity in sequence, as expected from the large phylogenetic divergence of the source species, and the fact that the enzymes are members of different lipase classes, this structure has some remarkably similar features to that of AZ4. Both enzymes have a "cap" that forms the top of the substrate-binding site. In AZ4, this cap is made up of helix D` and extended helix D, whereas, in AeCXE1, the cap is made partially from inserted elements in the region of helix D and partially from an extended loop contributed from the N-terminus of the protein. Despite these differences in the formation of the cap, the end result is a very similar surface topology in this region, as shown in Figure 4.14. This suggests that the shape of the binding sites in αβ hydrolases might be subjected to conservation more than the structural elements from which they are is composed.

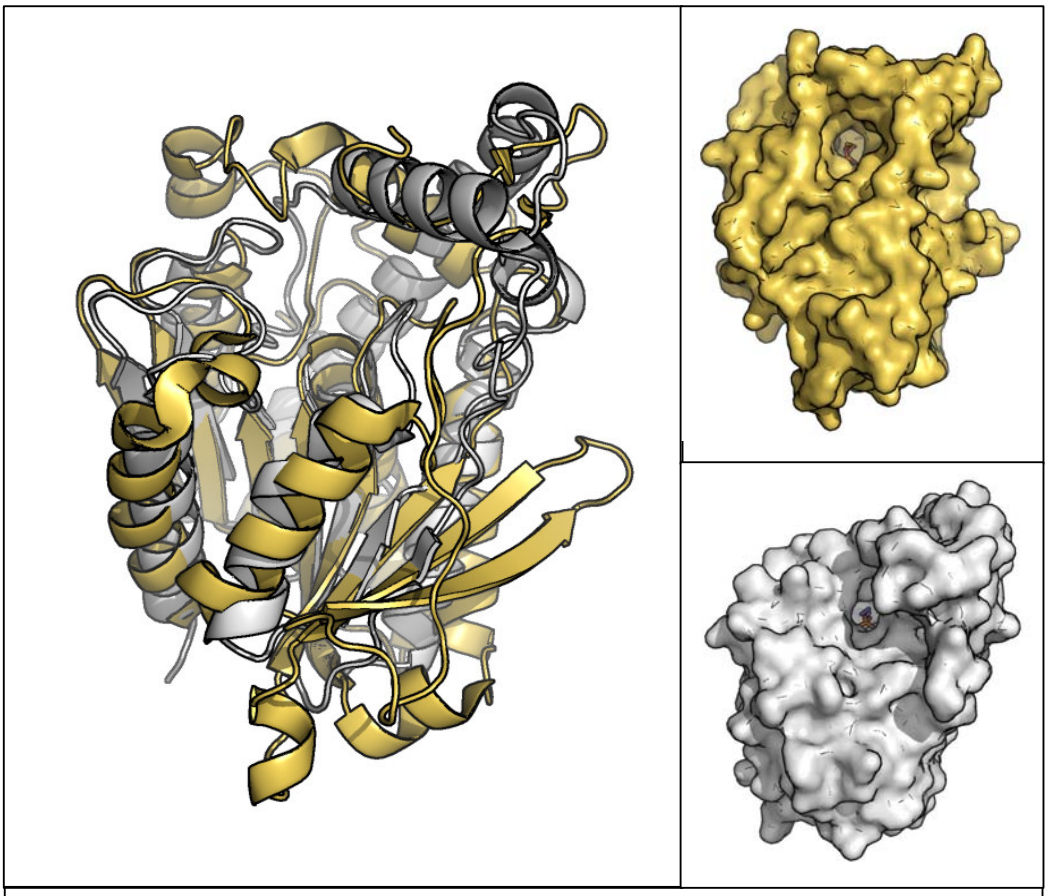


Figure 4.14 *Paraoxon bound AZ4 (grey) and paraoxon bound AeCXE1 from* A. eriantha *(yellow). Left panel: Cartoon comparison contrasting secondary structural differences, seen mainly at the top of the molecule. Right panels: Surface map of individual structures, showing the similar surface topography. The bound paraxon residues are shown as sticks, visible in the catalytic-site cleft. The colouring is the same as for the main panel.*

4.7 Summary

From the structures obtained, much information about the mechanism of catalysis and substrate binding of AZ4 has been ascertained. Like all other αβ hydrolases, AZ4 uses a catalytic triad made of a nucleophile, a base and an acid, with the nucleophile located in an unfavourable Ramachandran conformation in a nucleophile elbow.

The resting state structure of the enzyme was solved, and showed that the substrate-binding cavity and the cleft are preformed, rather than being induced by substrate association with the enzyme. Comparison of this structure with the structure of the high molecular weight lipase from *C. rugosa* (Grochulski, Bouthillier *et al.* 1994) showed that there were similarities in the architecture of the cleft, suggesting that AZ4 utilises a similar "tuning fork" substrate-binding conformation, with the scissile acyl chain projecting into the cavity and the other two chains of the triglyceride accommodated in the cleft on the enzyme surface.

In the resting state, the serine nucleophile is in a conformation that places it too far from the catalytic His 219 to form a hydrogen bond with it, a prerequisite for activation of the nucleophile. In the structure solved from crystals soaked in NaCl and tributyrin, the serine nucleophile was seen in a split conformation, one of which was close enough and in an appropriate geometry to form a hydrogen bond with the Nε of His 283.

Analysis of the structure obtained from crystals soaked in paraoxon provided information about the formation of a tetrahedral transition state and details of the possible mechanism. Most importantly, the change in conformation of the serine nucleophile shows that, either during the nucleophilic attack or after it, the serine sidechain rotates back to a conformation close to that seen in the resting state. This movement takes the serine out of hydrogen-bonding range of the histidine and appears to ensure that the reaction proceeds in a productive direction, by preventing the passage of the proton from His 219 back to Ser 94.

The structures with carboxylate groups bound in the oxyanion hole give information about the final stages of the reaction. In these structures, one of the carboxylate oxygens is involved in hydrogen bonding with the mainchain amides of the oxyanion hole, whereas the second carboxylate oxygen forms hydrogen bonds with Nε of His

219 and to the mainchain amide of the catalytic serine. Such product binding has not been previously observed in αβ hydrolases. It is postulated that it is this interaction that induces the serine to change conformation from its usual unfavourable mainchain configuration into a helical conformation. As well as inducing the change in the serine, the formation of the hydrogen bond between the serine amide and the carboxylate oxygen breaks the bond that was previously present between the serine amide and the carbonyl oxygen of Phe 116. This causes a conformational change in the loop region between Phe 116 and His 126, which results in the rotation of Ile 120 into the binding cavity, effectively filling it completely. Although this movement is not sufficient to cause an acetate or butyrate product to become dissociated from the enzyme (under the conditions used for crystal soaking), it may have a role in assisting the ejection of larger molecules from the enzyme.

Chapter 5

Concluding Remarks

Chapter 5

Concluding Remarks

5.1 Summary

This study has answered many of the long-standing questions about the structure of tributyrin esterases from dairy lactic acid bacteria, showing that, as anticipated from sequence analyses, they are all members of the αβ hydrolase family. Despite this similarity, however, all three enzymes examined have unique features that will help to advance understanding of how each enzyme functions in an industrial context, as well as to increase knowledge of the αβ hydrolase family in general.

5.1.1 Structural Elucidation

The crystal structures of three different esterases, EstA from *L. lactis*, and AA7 and Az4 from *Lb. rhamnosus*, were solved, and complexes of AA7 and AZ4 with suicide inhibitors or ligands bound were also determined.

Overall, all three enzymes conform to the basic αβ hydrolase fold that is typical of many esterases and lipases. Compared with the prototype fold, they all contain inserted helices in locations that are typical for deviations from the basic fold. No additional β strands are seen in the main sheet, a phenomenon that tends to be limited to lipases.

The substrate-binding elements of all three enzymes are found either on the loops between the secondary structural elements of the basic fold or on the inserted secondary elements, in line with the observation that the αβ hydrolase enzyme acts as a framework that allows modification for specific catalytic roles (Ollis, Cheah et al. 1992).

Although the enzymes catalyse reactions on similar substrates, analyses of the structures have shown that they are sufficiently different that a detailed comparison of all three structures is virtually impossible.

The solutions of the structures of these three enzymes nonetheless provide additional insights into the relationship between esterase structure and function. All three proteins share a similar fold, yet each has a differing arrangement of substrate-binding pockets, showing that different approaches can be used to catalyse different reactions. All three structures also show unique features that add to the understanding of the αβ hydrolase family in general.

One of the key differences in the activities of the three enzymes is that, whereas AA7 and EstA have been shown to be capable of hydrolysing ester bonds and also of synthesising esters through a transferase reaction, AZ4 is believed to be capable only of hydrolysis. This difference in activity is reflected in the arrangement of the substrate-binding sites of each enzyme.

5.2 Structural Summaries

5.2.1 AZ4

AZ4 was found to have an αβ hydrolase fold, although strand 1 of the prototype fold is absent. It has several insertions to the basic fold, with the most significant being between strand 6 and helix D, consisting of a loop that forms part of the substrate-binding cavity and a large cap over the top of the active site. The cap is also involved in forming a crystallographic homodimer through a symmetry operation. The native crystals and the crystal-soaking experiments produced structures that appear to represent the enzyme in resting, activated, intermediate and product-bound states. From these structures, it was found that the nucleophilic serine sidechain undergoes changes in conformation during the catalytic cycle. In the resting state, the serine is located too far from the histidine to become deprotonated. It appears to rotate into hydrogen-bonding range to become activated, and then rotates out of range again during the formation of the tetrahedral intermediate, preventing the histidine from

donating the proton back to the serine. A different mainchain conformation was observed in the product-bound state, where, because of rotation of the preceding peptide bond, the serine is in a helical conformation, which is uncharacteristic of αβ hydrolases. This change appears to trigger a further conformational change in the loop that is inserted between strand 6 and helix D, which closes the substrate-binding site. It is easy to imagine that such a conformational change might have a role in assisting the dissociation of a substrate at the end of the catalytic cycle, providing a novel "self-ejection" mechanism. The conformational changes observed have not been seen in any other αβ hydrolase family members to date, although the closest relative of AZ4 with a structure solved, Est30 from *Bacillus stearothermophilus* (Liu, Wang *et al.* 2004), has a very similar inserted loop, which is likely to be capable of undergoing a similar change.

5.2.1.1 AZ4 Binding-site Topology

There appears to be no acceptor-binding site in AZ4. The enzyme has a single binding cleft, with the catalytic histidine exposed to bulk solvent, allowing easy access of water during hydrolysis (Figure 5.1). This likely reflects the fact that AZ4 has been shown to be strictly a hydrolytic enzyme, and therefore does not require a specific binding site for an acceptor molecule, with hydrolytic water able to directly enter the active site from solution.

Interestingly, the arrangement of the binding site of AZ4 superficially resembles that of *Actinidia eriantha* AeCXE1 carboxylesterase, despite the enzymes being apparently unrelated (other than both being αβ hydrolases), demonstrating convergent evolution of binding-site topology for a similar function.

5.2.2 The Transferases

The studies of the structures of AA7 and, to a lesser extent, EstA give insights into how these enzymes are able to catalyse acyltransferase reactions as well as simple hydrolyses.

All esterases utilise a substituted-enzyme mechanism to transfer a group from a donor to an acceptor. In the substituted-enzyme mechanism, the transferred group is moved from the donor on to a reactive group of the enzyme. The donor then leaves the

enzyme as the first product, and the acceptor binds into the same, or an overlapping, site. Finally, the transferred group is transferred from the enzyme on to the acceptor to form the second reaction product. The acceptor is water during hydrolysis, and is an alcohol during alcoholysis (transferase activity).

In hydrolysis, the acceptor is a water molecule that enters the catalytic site from solution, where it is present at a high concentration, so binding of it as an acceptor is not an issue. However, for a transfer on to an acceptor that is present at a much lower concentration, there must be an alternative mechanism that eliminates the availability of water to act as an acceptor.

From kinetic studies on EstA and AA7 (Liu, Baker *et al.* 2004; Holland, Liu *et al.* 2005; Delabre, Ng *et al.* 2006), it is known that, in the absence of an alcohol acceptor, the enzymes will act as hydrolases. Therefore, for the transferase reaction to occur, the enzymes must bind an acceptor prior to the donor substrate. Consequently, the substrate binding sites of both these enzymes appear to have features that may be involved in the binding of an acceptor. In EstA this is the small pocket located off the P2 pocket, and in AA7 the deep closed over tunnel adjacent to the active site appears to be likely location for the binding of the acceptor.

The structures of AA7 and EstA suggest that it would be sterically difficult for an acceptor to enter its binding site after the donor has bound, meaning that a pseudo-compulsory substrate-binding order mechanism exists in these enzymes. This is reminiscent of the mechanism seen in compulsory-order ternary complex enzymes, which typically catalyse a single-step group transfer. However, it should be noted that, in most catalytic triad enzymes, the mechanistic requirements mean that the acceptor cannot enter the catalytic site itself until the first product has departed, as the transfer of the group on to the acceptor uses the same machinery as has been used to remove the transferable group from the donor. Ordered binding is typically found in substituted-enzyme mechanism catalysis where the sites for donor and acceptor usually overlap, because of the usage of the same catalytic machinery for acylation and deacylation of the enzyme. The positioning of the acceptor prior to the commencement of catalysis ensures that it is poised to enter the catalytic site to complete the reaction and displace any bound water, preventing hydrolysis. The proposed mechanism has not been kinetically established and is beyond the scope of this project, but is an interesting field for future studies.

Simple kinetic studies to establish substrate preferences for both hydrolysis activity and alcoholysis activity have shown that there is nothing to prevent hydrolysis from occurring in the absence of an alcohol acceptor. This fits with there being no apparent structural mechanism such as a conformational change to allow the donor to be bound only after an acceptor, which would cause the enzyme to perform alcoholysis exclusively. Intriguingly, there is no evidence that either enzyme will hydrolyse the ester products of the transferase mechanism.

5.2.2.1 EstA

The structures of EstA from *Streptococcus pyogenes* and *Lactococcus lactis* revealed a conserved tetrameric structure, with dimer association between the β strand 1 of each monomer, to produce an extended twisted β sheet as the main dimer interface. These dimers then pair together to form a tetramer, made of two dimer pairs.

With four EstA monomers in the asymmetric unit, and one of these monomers found in a different conformation from the other three, inferences about possible conformational changes during catalysis could be made. No obvious cause for the different states was found, as no density that could be attributed to a ligand was found in the active site of either enzyme. The enzyme has a well-defined, open substrate-binding cleft, bisected by the catalytic machinery and lined with mainly hydrophobic sidechains. It would appear that the enzyme can adopt either an open state or a closed state. It is probable that the enzyme adopts the open state during initial substrate binding, which would allow the binding of both a triglyceride substrate and an acceptor molecule. It is envisaged that conformational change to the closed state could then push the acceptor molecule out of its binding site, towards the catalytic machinery, allowing it to become activated in order to complete deacylation of the acyl-enzyme intermediate.

Although not all of the protein chain is visible in the closed state, the entire protein is ordered in the open state (excluding some N-terminal and C-terminal residues), showing that an interaction between monomers from the different dimers in the tetramer occurs. It is proposed that this communication mechanism between the two dimers could act as a means for cooperative binding to occur, whereby the opening during substrate binding to one monomer induces and stabilises the opening of the

other monomer to allow it to bind substrate also. Again, this hypothesis needs to be tested by a detailed kinetic analysis. Future work would also include analysis of enzyme-inhibitor complexes as has been done for AA7 (D. Colbert, personal communication, 2007).

5.2.2.2 AA7

The structure of AA7 showed that the enzyme was monomeric and had a typical αβ hydrolase fold, containing several inserted helices and loops, the most notable of which is an insertion involved in substrate binding between strand 6 and helix D. AA7 contains a trench-like binding site, bisected by the catalytic machinery, which is superficially similar in topology to that of EstA.

The structure solved with a tributyrin transition state analogue bound showed that the inserted region is involved in a conformational change that accommodates the inhibitor. A methionine sidechain from the top of the inserted helix is displaced from the active-site cleft, and is replaced by one of the butyl chains from the inhibitor. It is unclear whether the orientation of the inhibitor relative to the catalytic machinery is representative of how the donor substrate involved in a transferase reaction binds. This is because (i) the orientation of the inhibitor in the active site leaves no room for an acceptor molecule to bind, (ii) kinetic results suggest that substrates with longer acyl chains are preferred for transferase reactions (Holland, Liu *et al.* 2005), and (iii) studies of a similar esterase, Est2 (De Simone, Galdiero *et al.* 2000; De Simone, Mandrich *et al.* 2004), and more recent studies on AA7 carried out by another member of our group (Damon Colbert, unpublished data, 2007) suggest that the substrate orientation in the binding site is dependent on the length of substituents on the substrate glyceride backbone.

5.2.2.3 EstA and AA7 Binding-site Topology

Both EstA and AA7 have a cleft running across the enzyme that is bisected by the catalytic machinery (Figure 5.1). Both enzymes also have binding sites that are possibly designed to bind an acceptor. As the catalytic site is exposed to solvent, prebinding of the acceptor with expulsion of water is likely to be necessary for a transfer reaction, rather than hydrolysis, to occur.

EstA and AA7 differ in the location of these binding sites. The hypothesised acceptor-binding site is located off the P1 pocket in AA7, whereas it is found at the side of the P2 pocket in EstA. This is supported by the arrangement of the catalytic triad relative to the binding sites in each enzyme. The substrate-binding clefts of both AA7 and EstA are exposed to the solvent, which is an interesting contrast to the binding site of the acyltransferase MsAcT from *Mycobacterium smegmatis*, which forms a tight octamer, internalising the catalytic site to limit access of solvent water molecules (Mathews, Soltis *et al.* 2007).

5.2.2.4 AA7 Mechanism

In AA7, the catalytic histidine appears to be placed in a position where it could activate an acceptor entering from either the P1 side or the P2 side of the binding cleft (Figure 5.2). Such bidirectionality in substrate binding appears to be possible in other members of the hormone-specific lipase family, for example in Est2 from *Alicyclobacillus acidocaldarius* (De Simone, Galdiero *et al.* 2000; De Simone, Mandrich *et al.* 2004). In AA7, the tributyrin-analogue-bound structure (Section 4.5) appears to have the donor group in the P2 pocket, whereas, in a similar structure recently solved using a larger glyceride, substrate binding appears to be reversed (D. Colbert, unpublished data, 2007), implying that the direction of substrate binding down the catalytic cleft is dependent on the size of the acyl substituent of the substrate, a feature also shared with Est2 from *A. acidocaldarius*.

However, it should be noted that, with the P1 tunnel postulated to be a binding site for an acceptor, alcoholysis can occur only with the substrate bound so that the donor group is in the P1 pocket. This directionality appears to explain the differences in the rates of hydrolysis and alcoholysis on different sized glyceride substrates.

5.2.2.5 EstA Mechanism

In EstA, the catalytic histidine is located closer to and in a more favourable geometry to the P2 pocket than to the P1 pocket, suggesting that both donor and acceptor molecules both bind in the P2 pocket (at different stages of the catalytic cycle), close to the histidine. This is supported by the movement of Tyr 177 into the P2 side of the catalytic site as the enzyme adopts a closed formation, positioning the sidechain

oxygen so that it can act as a hydrogen-bond donor, perhaps assisting the more traditional mainchain amide oxyanion hole residues to stabilise the two intermediates of the transferase reaction (Figure 5.2).

Although the usage of tyrosine sidechains in the oxyanion hole has been seen previously in prolyl oligopeptidase and cocaine esterase (Fulop, Bocskei *et al.* 1998; Larsen, Turner *et al.* 2002), the tyrosine sidechain in both these examples is contributed from the loop that contributes the mainchain amide oxyanion hole groups in more conventional αβ hydrolases. Why tyrosine is present in this position is not clear. It may be required for the deacylation step because the acceptor in the transferase reaction is a much poorer nucleophile than the serine sidechain hydroxyl. Alternatively, if the acceptor was placed in a less optimal position for the histidine to activate it, perhaps by being further from it, tyrosine, which provides a stronger hydrogen bond than a mainchain amide, would better favour the formation of a tetrahedral intermediate during the deacylation step, inducing catalysis to occur under these less favourable conditions. The temporary formation of the oxyanion hole from a different part of the protein appears to be unique, and has not been seen in any published structure to date.

5.2.2.6 Acceptor Displacement Mechanisms

As the substituted-enzyme mechanism requires that the donor and the acceptor occupy a similar position in the catalytic site at different stages in catalysis, there must be a means of bringing this bound acceptor closer to the catalytic site when appropriate. If the acceptor binds in the P1 tunnel of AA7, as hypothesised, it would have to move forward to occupy the position vacated by the departing donor group. As no obvious conformational change is associated with this movement, it is postulated that it might occur by vibrational motion of the acceptor into close enough proximity to the catalytic site to become activated. This motion would be limited in its direction by the tunnel enclosing the acceptor, allowing it to move only forwards or backwards relative to the catalytic triad.

In EstA, the acceptor appears to be physically displaced towards the catalytic machinery by the change of the enzyme state from open to closed. This conformational change closes the postulated acceptor-binding site, pushing any

molecules present towards the catalytic histidine. This is the same motion that brings the sidechain of Tyr 177 into the active site to assist with oxyanion stabilisation.

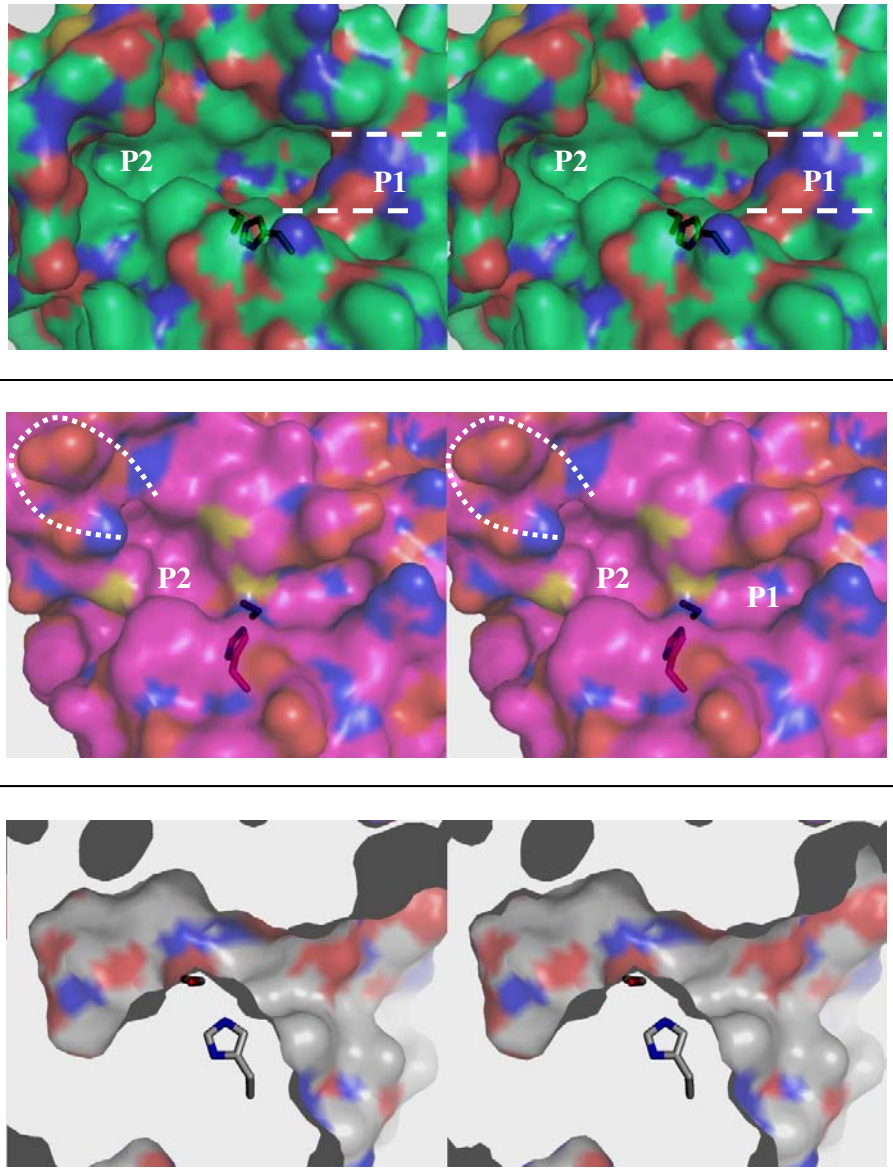


Figure 5.1 *Stereoview of arrangement of catalytic machinery and binding sites in esterases.*

The molecules have been aligned so that the binding sites lie horizontal on the page. Dotted lines indicate the extent of any buried pockets, and pockets have been labelled where appropriate. The catalytic serine and histidine have been shown as sticks, with the molecular surface also shown. Top: AA7; middle: EstA; bottom: AZ4. Note the differences in the relationship of the catalytic triad members, particularly histidine, to the binding pockets.

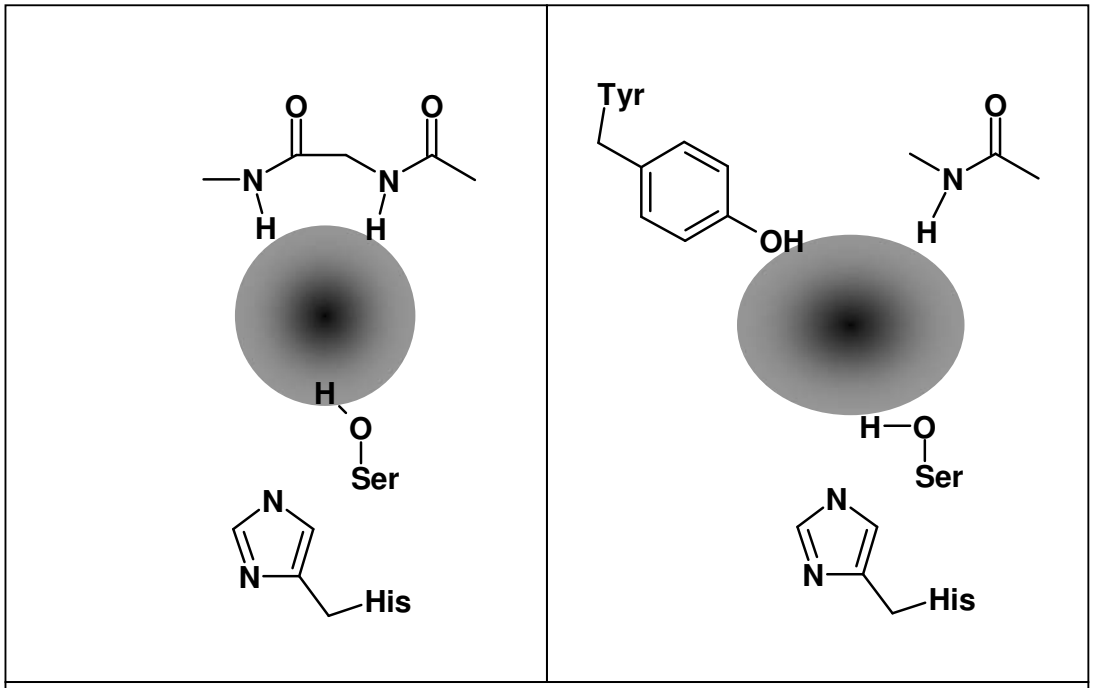


Figure 5.2 *Geometry of active site oxyanion holes.*
Left: AA7; right: EstA in closed state. Not shown is the dipole formed by the amides of the nucleophile helix, which would point out of the page, also contributing to the hole. Note the offset to the left of the oxyanion hole in EstA relative to AA7, because Tyr 177 is to the left of the nucleophile.

5.3 Structures and the Eight Class System

Although the bacterial lipase classification system devised by Arpigny and Jaeger (Arpigny and Jaeger 1999) is based purely on phylogeny and biochemical evidence, some of the structural information gained could be used to expand the definitions of the different classes.

EstA, a member of the class III extracellular lipases, is clearly an esterase, and all evidence suggests it is intracellular, which contrasts to its location in the classification system. It is of interest to note that some esterase activity is found to be cell wall associated however, suggesting the enzyme might be associating with cell wall components (Shaw, 1999). Bacterial lipases also often have a flap which folds over

the active site and leads to interfacial activation on insoluble lipid substrates. However, the EstA structure has no such flap.

AA7, a member of the Hormone Sensitive Lipase family was found to have the structural features typical of this family, in that it has the HGGG sequence motif involved in the stabilisation of the oxyanion holes main chain amides.

AZ4 falls outside the classification system, and, like the related Est30, suggests that the system requires expansion to allow the inclusion of the large number of enzymes that have been identified and characterised in the last decade.

5.4 Original Project Aims

One of the overall goals of this study was to ascertain if any information about the specificity of the chosen enzymes could be gained in regard to dairy product manufacture. In addition as a secondary goal, it was hoped that the layout of the substrate binding sites might allow the design of point mutants with more desirable industrial properties.

As a result of this study, the structural characteristics of esterases that act as transferases have been established. This information can be used to identify other potential orthologues in dairy starter species. However, as with most variable structural features in the αβ hydrolase fold, the regions that facilitate transferase activity are contributed largely from the variable loop regions that link the enzymes core elements. With αβ hydrolase enzymes being very widespread and diverse in sequence, it is unlikely that this feature could be reliably detected by sequence analysis. This would make it a necessity to gain structural information to determine if an enzyme has transferase activity or not. Such information that would be much more easily gained by screening for the desired activities.

Analysis of the structures show that it is potentially very difficult to rationally engineer the enzyme to alter its properties. There is the possibility of shortening the substrate binding cleft in EstA and AA7, although whether this would affect activity in a desirable way is very difficult to predict. Currently, the trend with enzyme design appears to be away from rationally designed mutations, and towards random

mutagenesis methods, followed by screening or selection for desirable activities using techniques such as directed evolution, and phage/ribosome display. This approach has been used successfully already to modify *B. subtilis* lipase (Droge, Boersma *et al.* 2006).

5.5 Dairy Systems and Enzyme Activity

Although two of the three enzymes studied (EstA and AA7) have been characterised extensively in terms of their activity in aqueous solution, and AZ4 to a lesser extent, little is known about how they might function in an environment where water is limited, such as that found in the maturation stage of cheese production.

The use of lowered water availability to alter enzyme activities is used extensively in an industrial context for the manufacture of various chemical compounds, especially where enantiomeric specific syntheses are required (Schoemaker, Mink *et al.* 2003). It seems intuitive that removal of water from a system might cause the activity of the enzyme to favour a transferase activity, as by lowering water concentration the effective concentration of other acceptors is increased. However, the situation in reality is more complicated. Removal of water reduces the hydration of the enzyme, which can change its overall size, and can alter the shape of the substrate binding regions in unpredictable ways. In turn this can lead to the enzyme activity being changed in unexpected ways. Even with the structural information now available, it is difficult to speculate as to how these enzymes might function during dairy product maturation, without more information as to how each of the enzymes operates in a reduced water environment.

5.6 Further Work

The structural elucidation of all three enzymes has answered many questions about the catalytic mechanisms used in their activities, yet has also reinforced the fact that many questions about them remain to be answered. Perhaps most interesting would be the elucidation of a physiological role of these enzymes in lactic acid bacteria. This is a potentially large scale task, as there is no indication of what the natural substrates of the enzymes might be. Deletion studies might also be of some benefit. However, EstA null mutant cells were found to be viable, whereas a null mutant of AA7 has not been successfully produced to date (M.-L. Delabre, personal communication, 2007); this might indicate that AA7 is required for cell viability. Nevertheless, this project has provided a good basic understanding of all three enzymes, and has uncovered some interesting and unique features in all three enzymes.

5.6.1 AZ4

AZ4 has proven to be easy to crystallise, and to be resilient to crystal-soaking experiments. This makes it an ideal candidate for further studies of αβ hydrolase mechanisms in general, to further explore the differences found between them and other catalytic-triad-based enzymes. We now have a range of triglyceride analogues (made available since this work was done), and these could be used to examine the stereo- and typo-specificities of the enzyme.

5.6.2 AA7

AA7 has been characterised for stereo- and typo-specificity using inhibitors. However, crystallography experiments to identify the acceptor-binding site have not been carried out. These would probably involve either co-crystallisation or crystal-soaking studies. Either larger alcohols or larger carboxylic acids could be used, although the solubility of these compounds is likely to be a complicating factor. Brominated forms of these compounds are available; they have a higher solubility, and may be more suited for this kind of experiment.

According to our knowledge of the catalytic cycle, the alcohol acceptor ought to bind to the enzyme in the absence of a donor substrate, but it will not be moved into position to accept the acyl group until halfway through catalysis. This means that it should also bind when a triglyceride analogue is bound, because this step represents the first intermediate of the reaction. The successful structure determination of such a complex would show the location of the acceptor-binding pocket, and would give some indication of the movements that it undergoes during the catalytic cycle.

5.6.3 EstA

Of the three enzymes studied in this project, EstA is the most enigmatic, because of the difficulties experienced in crystallising the protein, and therefore the reliance on previously obtained data. This makes it an ideal candidate for further studies. From analysis of the structure, it is apparent that any purification tag added to the N-terminal of the protein is likely to interfere with crystallisation. This means that, prior to undertaking studies involving inhibitor-bound structures, it will be necessary to produce a new construct with C-terminal purification tags, and to identify conditions under which these can be crystallised. Once the protein can be reproducibly crystallised, ligand-binding studies will be able to be undertaken. Although, fortuitously, an open state and a closed state of the protein were observed in the native crystals, it is not known whether substrate mimetics will induce the same movement, or whether ligand binding will cause a different conformational change in the enzyme. Site-directed mutagenesis studies would also be of interest to determine the roles of key residues in the catalytic site in terms of substrate binding. Also of interest as a mutagenesis target is Tyr 162, which might have a role in the formation of the oxyanion catalytic intermediate.

References

Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller and D. J. Lipman (1997). "Gapped Blast and Psi-Blast: A New Generation of Protein Database Search Programs." Nucleic Acids Research **25**(17).

Anderson, D. H., G. Harth, M. A. Horwitz and D. Eisenberg (2001). "An Interfacial Mechanism and a Class of Inhibitors Inferred from Two Crystal Structures of the Mycobacterium Tuberculosis 30 Kda Major Secretory Protein (Antigen 85b), a Mycolyl Transferase." J Mol Biol **307**(2): 671-81.

Arpigny, J. L. and K. E. Jaeger (1999). "Bacterial Lipolytic Enzymes: Classification and Properties." Biochem J **343 Pt 1**: 177-83.

Augustinsson, K.-B. and D. Nachmansohn (1949). "Studies on Cholinesterase. Vi. Kinetics of the Inhibition of Acetylcholinesterase." Journal of Biological Chemistry **179**(2): 543-559.

Bachovchin, W. W. (1986). "15n Nmr Spectroscopy of Hydrogen-Bonding Interactions in the Active Site of Serine Proteases: Evidence for a Moving Histidine Mechanism." Biochemistry **25**(23): 7751-9.

Bahar, M., C. C. Ballard, S. Cohen, K. Cowtan, E. J. Dodson, P. Emsley, R. M. Esnouf, R. M. Keegan, V. Lamzin, G. Langer, V. Levdikov, F. Long, C. J. Meier, A. Muller, G. Murshudov, A. Perrakis, C. Siebold, N. D. Stein, M. G. W. Turkenburg, A. Vagin, M. D. Winn, G. Winter and K. S. Wilson (2006). "Spine Workshop on Automated X-Ray Analysis: A Progress Report." Acta Crystallographica Section D - Biological Crystallography **62**: 1170-1183

Barth, A., M. Wahab, W. Brandt and K. Frost (1993). "Classification of Serine Proteases Derived from Steric Comparisons of Their Active Sites." Drug Des Discov **10**(4): 297-317.

Besancon, X., R. Ratomahenina and P. Galzy (1995). "Isolation and Partial Characterisation of an Esterase (Ec 3.1.1.1) from a Debaryomyces Hansenii Strain." Netherlands Milk and Dairy Journal **49**(2-3): 97-110.

Bills, D., M. Morgan, L. Libbey and E. Day (1965). "Identification of Compounds Responsible for Fruity Flavour Defect of Experimental Cheddar Cheese." Journal of Dairy Science **48**(9): 1168-1173.

Birk, Y. (1983). Proteinase Inhibitors. Hydrolytic Enzymes. A. Neuberger and K. Brocklehurst. Amsterdam, Elsevier. **16:** 265.

Blake, C. C. F., L. N. Johnson, G. A. Mair, A. C. T. North, D. C. Phillips and V. R. Sarma (1967). "Crystallographic Studies of the Activity of Hen Egg-White Lysozyme." Proceedings of the Royal Society of London. Series B, Biological Sciences **167**: 378-388.

Blow, D. M., J. J. Birktoft and B. S. Hartley (1969). "Role of a Buried Acid Group in the Mechanism of Action of Chymotrypsin." Nature **221**(5178): 337-40.

Bode, W., Z. Chen, K. Bartels, C. Kutzbach, G. Schmidt-Kastner and H. Bartunik (1983). "Refined 2 a X-Ray Crystal Structure of Porcine Pancreatic Kallikrein a, a Specific Trypsin-Like Serine Proteinase. Crystallization, Structure Determination, Crystallographic Refinement, Structure and Its Comparison with Bovine Trypsin." Journal of Molecular Biology **164**(2): 237-82.

Brady, L., A. M. Brzozowski, Z. S. Derewenda, E. Dodson, G. Dodson, S. Tolley, J. P. Turkenburg, L. Christiansen, B. Huge-Jensen, L. Norskov and et al. (1990). "A Serine Protease Triad Forms the Catalytic Centre of a Triacylglycerol Lipase." Nature **343**(6260): 767-70.

Brockerhoff, H. (1973). "A Model of Pancreatic Lipase and the Orientation of Enzymes at Interfaces." Chemistry and Physics of Lipids **10**(215-222).

Brockerhoff, H. and R. G. Jensen (1974). Lipolytic Enzymes. New York, Academic Press**:** 32-34.

Brockerhoff, H. and R. G. Jensen (1974). Lipolytic Enzymes. New York, Academic Press**:** 57.

Brzozowski, A. M., U. Derewenda, Z. S. Derewenda, G. G. Dodson, D. M. Lawson, J. P. Turkenburg, F. Bjorkling, B. Huge-Jensen, S. A. Patkar and L. Thim (1991). "A Model for Interfacial Activation in Lipases from the Structure of a Fungal Lipase-Inhibitor Complex." Nature **351**(6326): 491-4.

Brzozowski, A. M., H. Savage, C. S. Verma, J. P. Turkenburg, D. M. Lawson, A. Svendsen and S. Patkar (2000). "Structural Origins of the Interfacial Activation in Thermomyces (Humicola) Lanuginosa Lipase." Biochemistry **39**(49): 15071-82.

Cohen, G. H., E. W. Silverton and D. R. Davies (1981). "Refined Crystal Structure of Gamma-Chymotrypsin at 1.9 a Resolution. Comparison with Other Pancreatic Serine Proteases." Journal of Molecular Biology **148**(4): 449-79.

Collins, Y. F. (2003). "Evidence of a Relationship between Autolysis of Starter Bacteria and Lipolyis in Cheddar Cheese During Ripening." Journal of Dairy Research **70**: 103-113.

Cornish-Bowden, A. (2004). Fundamentals of Enzyme Kinetics. London, Portland Press**:** 157-163.

Crow, V. L., B. Curry, M. Christison, K. Hellier, R. Holland and S. Q. Liu (2002). "Raw Milk Flora and Non-Starter Lactic Acid Bacteria as Adjuncts." Australian Journal of Dairy Technology **57**: 99-105.

Cruz, H., C. Perez, E. Wellington, C. Castro and L. Servin-Gonzalez (1994). "Sequence of the Streptomyces Albus G Lipase-Encoding Gene Reveals the Presence of a Prokaryotic Lipase Family." Gene **144**(1): 141-2.

Cuff, M. E., M. Zhou, F. Collart and A. Joachimiak (2004). "Structure of a Carboxylesterase from Bacillus Stearothermophilus." To be Published.

De Simone, G., S. Galdiero, G. Manco, D. Lang, M. Rossi and C. Pedone (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**(5): 761-71.

De Simone, G., L. Mandrich, V. Menchise, V. Giordano, F. Febbraio, M. Rossi, C. Pedone and G. Manco (2004). "A Substrate-Induced Switch in the Reaction Mechanism of a Thermophilic Esterase: Kinetic Evidences and Structural Basis." J Biol Chem **279**(8): 6815-23.

Delabre, M.-L., J. Ng, Wingate, S., S.-Q. Liu, E. Chen, T. Wang, R. Holland and M. W. Lubbers (2006). Cloning and Characterization of the Main Intracellular Esterase from Lactobacillus Rhamnosus Hn001. Flavour Science: Recent Advances and Trends. W. L. P. Bredie, Petersen, M. A.**:** 69-72.

DeLano, W. L. (2002). The Pymol Molecular Graphics System. Palo Alto, CA, USA, DeLano Scientific.

Derewenda, U., A. M. Brzozowski, D. M. Lawson and Z. S. Derewenda (1992). "Catalysis at the Interface: The Anatomy of a Conformational Change in a Triglyceride Lipase." Biochemistry **31**(5): 1532-41.

Derewenda, Z. S. (2004). "The Use of Recombinant Methods and Molecular Engineering in Protein Crystallization " Methods **34**(3): 354-363

Dodson, E. (1991). Baverage. CCP4i Software Package.

Dodson, G. and A. Wlodawer (1998). "Catalytic Triads and Their Relatives." Trends Biochem Sci **23**(9): 347-52.

Droge, M. J., Y. L. Boersma, G. van Pouderoyen, T. E. Vrenken, C. J. Ruggeberg, M. T. Reetz, B. W. Dijkstra and W. J. Quax (2006). "Directed Evolution of Bacillus Subtilis Lipase a by Use of Enantiomeric Phosphonate Inhibitors: Crystal Structures and Phage Display Selection." Chembiochem **7**(1): 149-57.

Egloff, M. P., F. Marguet, G. Buono, R. Verger, C. Cambillau and H. van Tilbeurgh (1995). "The 2.46 a Resolution Structure of the Pancreatic Lipase-Colipase Complex Inhibited by a C11 Alkyl Phosphonate." Biochemistry **34**(9): 2751-62.

Ekstrom, F., C. Akfur, A. K. Tunemalm and S. Lundberg (2006). "Structural Changes of Phenylalanine 338 and Histidine 447 Revealed by the Crystal Structures of Tabun-Inhibited Murine Acetylcholinesterase." Biochemistry **45**(1): 74-81.

Emsley, P., Cowtan, K. (2004). "Coot: Model-Building Tools for Molecular Graphics." Acta Crystallographica Section D - Biological Crystallography **60**: 2126-2132.

Entressangles, B. and P. Desnuelle (1968). "Action of Pancreatic Lipase on Aggregated Glyceride Molecules in an Isotropic System." Biochim Biophys Acta **159**(2): 285-295.

Ewis, H. E., A. T. Abdelal and C. D. Lu (2004). "Molecular Cloning and Characterization of Two Thermostable Carboxyl Esterases from Geobacillus Stearothermophilus." Gene **329**: 187-95.

Fan, X. and M. Mattey (1999). "Small Enzymes with Esterase Activities from Two Thermophilic Fungi, Emericella Nidulans and Talaromyces Emersonii." Biotechnology Letters **21**(2): 1071-1076.

Fenster, K. M., K. L. Parkin and J. L. Steele (2003). "Nucleotide Sequencing, Purification, and Biochemical Properties of an Arylesterase from Lactobacillus Casei Lila." J Dairy Sci **86**(8): 2547-57.

Fernandez, L., M. M. Beerthuyzen, J. Brown, R. J. Siezen, T. Coolbear, R. Holland and O. P. Kuipers (2000). "Cloning, Characterization, Controlled Overexpression, and Inactivation of the Major Tributyrin Esterase Gene of Lactococcus Lactis." Appl Environ Microbiol **66**(4): 1360-8.

Ferretti, J. J., W. M. McShan, D. Ajdic, D. J. Savic, G. Savic, K. Lyon, C. Primeaux, S. Sezate, A. N. Suvorov, S. Kenton, H. S. Lai, S. P. Lin, Y. Qian, H. G. Jia, F. Z. Najar, Q. Ren, H. Zhu, L. Song, J. White, X. Yuan, S. W. Clifton, B. A. Roe and R. McLaughlin (2001). "Complete Genome Sequence of an M1 Strain of Streptococcus Pyogenes." Proc Natl Acad Sci U S A **98**(8): 4658-63.

Fleming, S. M., T. A. Robertson, G. J. Langley and T. D. Bugg (2000). "Catalytic Mechanism of a C-C Hydrolase Enzyme: Evidence for a Gem-Diol Intermediate, Not an Acyl Enzyme." Biochemistry **39**(6): 1522-31.

Forss, D. and S. Patton (1966). "Flavor of Cheddar Cheese." Journal of Dairy Science **49**(1): 89-91.

Franken, S. M., H. J. Rozeboom, K. H. Kalk and B. W. Dijkstra (1991). "Crystal Structure of Haloalkane Dehalogenase: An Enzyme to Detoxify Halogenated Alkanes." Embo J **10**(6): 1297-302.

Frenken, L. G., M. R. Egmond, A. M. Batenburg, J. W. Bos, C. Visser and C. T. Verrips (1992). "Cloning of the Pseudomonas Glumae Lipase Gene and Determination of the Active Site Residues." Appl Environ Microbiol **58**(12): 3787-91.

Frey, P. A., S. A. Whitt and J. B. Tobin (1994). "A Low-Barrier Hydrogen Bond in the Catalytic Triad of Serine Proteases." Science **264**(5167): 1927-30.

Fulop, V., Z. Bocskei and L. Polgar (1998). "Prolyl Oligopeptidase: An Unusual Beta-Propeller Domain Regulates Proteolysis." Cell **94**(2): 161-70.

Garavito, R. M., M. G. Rossmann, P. Argos and W. Eventoff (1977). "Convergence of Active Center Geometries." Biochemistry **16**(23): 5065-5071.

Gobbetti, M., P. F. Fox and L. Stepaniak (1997). "Isolation and Characterisation of a Tributyrin Esterase from *Lactobacillus Plantarum* 2739." Journal of Dairy Science **80**: 3099-3106.

Gobbetti, M., E. Smacchi and A. Corsetti (1997). "Purification and Characterization of a Cell Surface-Associated Esterase from Lactobacillus Fermentum Dt41." International Dairy Journal **7**(1): 13-21.

Grochulski, P., F. Bouthillier, R. J. Kazlauskas, A. N. Serreqi, J. D. Schrag, E. Ziomek and M. Cygler (1994). "Analogs of Reaction Intermediates Identify a Unique Substrate Binding Site in Candida Rugosa Lipase." Biochemistry **33**(12): 3494-500.

Grochulski, P., Y. Li, J. D. Schrag, F. Bouthillier, P. Smith, D. Harrison, B. Rubin and M. Cygler (1993). "Insights into Interfacial Activation from an Open Structure of Candida Rugosa Lipase." J Biol Chem **268**(17): 12843-7.

Grochulski, P., Y. Li, J. D. Schrag and M. Cygler (1994). "Two Conformational States of Candida Rugosa Lipase." Protein Sci **3**(1): 82-91.

Hakansson, K. (2002). "The Strand-Helix Motif Is a Recurring Theme in Biological Hydrolysis. Does the Conformation of the Ramachandran Outlier Enhance Its Electrophilicity?" Int J Biol Macromol **30**(5): 273-7.

Hedstrom, L. (2002). "Serine Protease Mechanism and Specificity." Chem Rev **102**(12): 4501-24.

Heikinheimo, P., A. Goldman, C. Jeffries and D. L. Ollis (1999). "Of Barn Owls and Bankers: A Lush Variety of Alpha/Beta Hydrolases." Structure **7**(6): R141-6.

Hemila, H., T. T. Koivula and I. Palva (1994). "Hormone-Sensitive Lipase Is Closely Related to Several Bacterial Proteins, and Distantly Related to Acetylcholinesterase and Lipoprotein Lipase: Identification of a Superfamily of Esterases and Lipases." Biochim Biophys Acta **1210**(2): 249-53.

Hendrickson, W. A., J. R. Horton and D. M. LeMaster (1990). "Selenomethionyl Proteins Produced for Analysis by Multiwavelength Anomalous Diffraction (Mad): A Vehicle for Direct Determination of Three-Dimensional Structure." EMBO Journal **9**(5): 1665-72.

Hirshmann, H. (1960). "The Nature of Substrate Asymmetry in Stereoselective Reactions." Journal of Biological Chemistry **235**(10): 2762-2767.

Hobson, A. H., C. M. Buckley, J. L. Aamand, S. T. Jorgensen, B. Diderichsen and D. J. McConnell (1993). "Activation of a Bacterial Lipase by Its Chaperone." Proc Natl Acad Sci U S A **90**(12): 5682-6.

Holland, R. and T. Coolbear (1996). "Purification of Tributyrin Esterase from Lactococcus Lactis Subsp Cremoris E8." Journal of Dairy Research **63**(1): 131-140.

Holland, R., S.-Q. Liu, V. L. Crow, M.-L. Delabre, M. Lubbers, M. Bennett and G. Norris (2005). "Esterases of Lactic Acid Bacteria and Cheese Flavour: Milk Fat Hydrolysis, Alcoholysis and Esterification." International Dairy Journal **15**: 711-718.

Holland, R., S. Q. Liu, T. Wang, M. Bennett, G. Norris, M.-L. Delabre, M. W. Lubbers, J. W. Dekker and V. L. Crow (2002). Esterases as Biocatalysts for Flavour in Cheese. Biocatalysis in the Food and Drinks Industry, University of Westminster.

Holm, L. and C. Sander (1996). "Mapping the Protein Universe." Science **273**(5275): 595-603.

Holmquist, M. (2000). "Alpha/Beta-Hydrolase Fold Enzymes: Structures, Functions and Mechanisms." Curr Protein Pept Sci **1**(2): 209-35.

Hotelier, T., L. Renault, X. Cousin, V. Negre, P. Marchot and A. Chatonnet (2004). "Esther, the Database of the Alpha/Beta-Hydrolase Fold Superfamily of Proteins." Nucleic Acids Res **32**(Database issue): D145-7.

Ileperuma, N. R., S. D. Marshall, C. J. Squire, H. M. Baker, J. G. Oakeshott, R. J. Russell, K. M. Plummer, R. D. Newcomb and E. N. Baker (2007). "High-Resolution Crystal Structure of Plant Carboxylesterase Aecxe1, from Actinidia Eriantha, and Its Complex with a High-Affinity Inhibitor Paraoxon." Biochemistry **46**(7): 1851-9.

Itabashi, Y. and T. Takagi (1987). "High Performance Liquid Chromatographic Separation of Triacylglycerol Enantiomers on a Chiral Stationary Phase." Journal of Chromatography **402**: 257-264.

IUPAC-IUB (1967). "Iupac-Iub Commission on Biochemical Nomenclature: The Nomenclature of Lipids." Journal of Biological Chemistry **212**(21): 1815-1819.

Jackson, M., C. Raynaud, M. A. Laneelle, C. Guilhot, C. Laurent-Winter, D. Ensergueix, B. Gicquel and M. Daffe (1999). "Inactivation of the Antigen 85c Gene Profoundly Affects the Mycolate Content and Alters the Permeability of the Mycobacterium Tuberculosis Cell Envelope." Mol Microbiol **31**(5): 1573-87.

Jaeger, T., M. Arsic and C. Mayer (2005). "Scission of the Lactyl Ether Bond of N-Acetylmuramic Acid by Escherichia Coli "Etherase"." J Biol Chem **280**(34): 30100-6.

James, M. N., A. R. Sielecki, G. D. Brayer, L. T. Delbaere and C. A. Bauer (1980). "Structures of Product and Inhibitor Complexes of Streptomyces Griseus Protease a at 1.8 a Resolution. A Model for Serine Protease Catalysis." Journal of Molecular Biology **144**(1): 43-88.

Jancarik, J. and S. H. Kim (1991). "Sparse Matrix Sampling: A Screening Method for the Crystallization of Macromolecules." Journal of Applied Crystallography **24**: 409-411.

Jensen, R. G., F. A. deJong and R. M. Clark (1983). "Determination of Lipase Specificity." Lipids **18**(3): 239-52.

Jorgensen, S., K. W. Skov and B. Diderichsen (1991). "Cloning, Sequence, and Expression of a Lipase Gene from Pseudomonas Cepacia: Lipase Production in Heterologous Hosts Requires Two Pseudomonas Genes." J Bacteriol **173**(2): 559-67.

Karnovsky, M. L. and D. Wolff (1960). Biochem. Lipids Proc. Int. Conf. Biochem. Probl. Lipids 5th. Pergamon, Oxford.

Katz, M., R. Medina, S. Gonzalez and G. Oliver (2002). "Esterolytic and Lipolytic Activities of Lactic Acid Bacteria Isolated from Ewe's Milk and Cheese." Journal of Food Protection **65**(12): 1997-2001.

Kawasaki, K., H. Kondo, M. Suzuki, S. Ohgiya and S. Tsuda (2002). "Alternate Conformations Observed in Catalytic Serine of Bacillus Subtilis Lipase Determined at 1.3 a Resolution." Acta Crystallogr D Biol Crystallogr **58**(Pt 7): 1168-74.

Kleywegt, G. J. (1996). "Use of Non-Crystallographic Symmetry in Protein Structure Refinement." Acta Crystallogr D Biol Crystallogr **52**(Pt 4): 842-57.

Klibanov, A. M. (1989). "Enzymatic Catalysis in Anhydrous Organic Solvents." Trends Biochem Sci **14**(4): 141-4.

Klibanov, A. M. (2001). "Improving Enzymes by Using Them in Organic Solvents." Nature (London) **409**(6817): 241-246.

Lang, D. A., M. L. Mannesse, G. H. de Haas, H. M. Verheij and B. W. Dijkstra (1998). "Structural Basis of the Chiral Selectivity of Pseudomonas Cepacia Lipase." Eur J Biochem **254**(2): 333-40.

Larsen, N. A., J. M. Turner, J. Stevens, S. J. Rosser, A. Basran, R. A. Lerner, N. C. Bruce and I. A. Wilson (2002). "Crystal Structure of a Bacterial Cocaine Esterase." Nat Struct Biol **9**(1): 17-21.

Laskowski, R. A., M. W. MacArthur, D. S. Moss and J. M. Thornton (1993). "Procheck: A Program to Check the Stereochemical Quality of Protein Structures." J. Appl. Cryst. **26**: 283-291.

Law, B. A., M. E. Sharpe and B. Reiter (1974). "The Release of Intracellular Dipeptidase from Starter Streptococci During Cheddar Cheese Ripening." Journal of Dairy Research **41**: 137-146.

Law, B. A. and A. S. Wigmore (1983). "Accelerated Ripening of Cheddar Cheese with a Commercial Proteinase and Intracellular Enzymes from Starter Streptococci." Journal of Dairy Research **50**: 519-525.

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**.

Lienhard, G. E., I. I. Secemski, K. A. Koehler and R. N. Lindquist (1971). "Enzymatic Catalysis and the Transition State Theory of Reaction Rates: Transition State Analogues." Cold Spring Harbor Symp. Quant. Biol. **36**(45): 45-51.

Liu, P., Y. F. Wang, H. E. Ewis, A. T. Abdelal, C. D. Lu, R. W. Harrison and I. T. Weber (2004). "Covalent Reaction Intermediate Revealed in Crystal Structure of the Geobacillus Stearothermophilus Carboxylesterase Est30." J Mol Biol **342**(2): 551-61.

Liu, S.-Q., K. Baker, M. Bennett, R. Holland, G. Norris and V. L. Crow (2004). "Characterisation of Esterases of Streptococcus Thermophilus St1 and Lactococcus Lactis Subsp. Cremoris B1079 as Alcohol Acyltransferases." International Dairy Journal **14**(10): 865-870.

Liu, S.-Q., R. Holland and V. L. Crow (2001). "Purification and Properties of Intracellular Esterases from Streptococcus Thermophilus." International Dairy Journal **11**: 27-35.

Liu, S. Q., R. Holland and V. L. Crow (2003). "Ester Synthesis in an Aqueous Environment by Streptococcus Thermophilus and Other Dairy Lactic Acid Bacteria." Appl Microbiol Biotechnol **63**(1): 81-8.

Mannesse, M. L., J. W. Boots, R. Dijkman, A. J. Slotboom, H. T. van der Hijden, M. R. Egmond, H. M. Verheij and G. H. de Haas (1995). "Phosphonate Analogues of Triacylglycerols Are Potent Inhibitors of Lipase." Biochim Biophys Acta **1259**(1): 56-64.

Mannesse, M. L., G. H. de Haas, H. T. van der Hijden, M. R. Egmond and H. M. Verheij (1997). "Chiral Preference of Cutinase in the Reaction with Phosphonate Inhibitors." Biochem Soc Trans **25**(1): 165-70.

Marguet, F., J.-F. Cavalier, R. Verger and G. Buono (1999). "New Synthesis of Sn-1,2- and Sn-2,3-O-Diacylglycerols - Application to the Synthesis of Enantiopure Phosphonates Analogous to Triglycerides: A New Class of Inhibitors of Lipases." Eur. J. Org. Chem. **1999**: 1671-1678.

Martinez, C., P. De Geus, M. Lauwereys, G. Matthyssens and C. Cambillau (1992). "Fusarium Solani Cutinase Is a Lipolytic Enzyme with a Catalytic Serine Accessible to Solvent." Nature **356**(6370): 615-8.

Martinez, C., A. Nicolas, H. van Tilbeurgh, M. P. Egloff, C. Cudrey, R. Verger and C. Cambillau (1994). "Cutinase, a Lipolytic Enzyme with a Preformed Oxyanion Hole." Biochemistry **33**(1): 83-9.

Masai, E., Y. Katayama, S. Kawai, S. Nishikawa, M. Yamasaki and N. Morohoshi (1991). "Cloning and Sequencing of the Gene for a Pseudomonas Paucimobilis Enzyme That Cleaves Beta-Aryl Ether." J Bacteriol **173**(24): 7950-5.

Mathews, I., M. Soltis, M. Saldajeno, G. Ganshaw, R. Sala, W. Weyler, M. A. Cervin, G. Whited and R. Bott (2007). "Structure of a Novel Enzyme That Catalyzes Acyl Transfer to Alcohols in Aqueous Conditions." Biochemistry **46**(31): 8969-79.

Mattey, M., D. Simoes, A. Brown and X. Fan (1998). "Enzymes with a Low Molecular Weight." Acta Chimica Slovenia **45**(1): 45-57.

Matthews, D. A., R. A. Alden, J. J. Birktoft, T. Freer and J. Kraut (1977). "Re-Examination of the Charge Relay System in Subtilisin Comparison with Other Serine Proteases." Journal of Biological Chemistry **252**(24): 8875-83.

Mattson, F. H. and L. W. Beck (1956). "The Specificity of Pancreatic Lipase for the Primary Hydroxyl Groups of Glycerides." Journal of Biological Chemistry **219**: 735.

Maylie, M. F., M. Charles and P. Desnuelle (1972). "Action of Organophosphates and Sulfonyl Halides on Porcine Pancreatic Lipase." Biochim Biophys Acta **276**: 162-175.

Maylie, M. F., M. Charles, L. Sarda and P. Desnuelle (1969). "Action of Organophosphates on Pancreatic Lipase." Biochim Biophys Acta **178**: 196-198.

McCoy, A. J., R. W. Grosse-Kunstleve, L. C. Storoni and R. J. Read (2005). "Likelihood-Enhanced Fast Translation Functions." Acta Crystallogr D Biol Crystallogr **61**(Pt 4): 458-64.

McGrath, M. E., J. R. Vasquez, C. S. Craik, A. S. Yang, B. Honig and R. J. Fletterick (1992). "Perturbing the Polar Environment of Asp102 in Trypsin: Consequences of Replacing Conserved Ser214." Biochemistry **31**(12): 3059-64.

Meyers, S. A., S. L. Cuppert and R. W. Hutkins (1996). "Lipase Production by Lactic Acid Bacteria and Activity on Butter Oil." Food Microbiology (London) **13**(5): 383-389.

Milkowski, C. and D. Strack (2004). "Serine Carboxypeptidase-Like Acyltransferases." Phytochemistry **65**(5): 517-24.

Muralidhar, R. V., R. Marchant and P. Nigam (2001). "Lipases in Racemic Resolutions." Journal of Chemical Technology & Biotechnology **76**(1): 3-8.

Murshudov, G. N., A. A. Vagin and E. J. Dodson (1997). "Refinement of Macromolecular Structures by the Maximum-Likelihood Method." Acta Crystallographica Section D - Biological Crystallography **53**: 240-255.

Nardini, M. and B. W. Dijkstra (1999). "Alpha/Beta Hydrolase Fold Enzymes: The Family Keeps Growing." Curr Opin Struct Biol **9**(6): 732-7.

Nicoll, A. J. and R. K. Allemann (2004). "Nucleophilic and General Acid Catalysis at Physiological Ph by a Designed Miniature Esterase." Org Biomol Chem **2**(15): 2175-80.

Nini, L., L. Sarda, L. C. Comeau, E. Boitard, J. P. Dubes and H. Chahinian (2001). "Lipase-Catalysed Hydrolysis of Short-Chain Substrates in Solution and in Emulsion: A Kinetic Study." Biochim Biophys Acta **1534**(1): 34-44.

Noble, M. E., A. Cleasby, L. N. Johnson, M. R. Egmond and L. G. Frenken (1993). "The Crystal Structure of Triacylglycerol Lipase from Pseudomonas Glumae Reveals a Partially Redundant Catalytic Aspartate." FEBS Lett **331**(1-2): 123-8.

Ollis, D. L., E. Cheah, M. Cygler, B. Dijkstra, F. Frolow, S. M. Franken, M. Harel, S. J. Remington, I. Silman, J. Schrag and et al. (1992). "The Alpha/Beta Hydrolase Fold." Protein Eng **5**(3): 197-211.

Oosterbaan, R. A., P. Kunst, J. van Rotterdam and J. A. Cohen (1958). "The Reaction of Chymotrypsin and Diisopropylphosphorofluoridate. Ii: The Structure of Two Dp-Substituted Peptides from Chymotrypsin-Dp." Biochim Biophys Acta **27**: 556-563.

Oterholm, A., Z. J. Ordal and L. D. Witter (1968). "Glycerol Ester Hydrolase Activity of Lactic Acid Bacteria." Appl Microbiol **16**(3): 524-7.

Paltauf, F., F. Esfandi and A. Holasek (1974). "Stereospecificity of Lipases. Enzymic Hydrolysis of Enantiomeric Alkyl Diacylglycerols by Lipoprotein Lipase, Lingual Lipase and Pancreatic Lipase." FEBS Lett **40**(1): 119-123.

Parodi, P. W. (1982). "Positional Distribution of Fatty Acids in Triglycerides from Milk of Several Species of Mammals." Lipids **17**(6): 437-42.

Pathak, D., K. L. Ngai and D. Ollis (1988). "X-Ray Crystallographic Structure of Dienelactone Hydrolase at 2.8 A." J Mol Biol **204**(2): 435-45.

Pathak, D. and D. Ollis (1990). "Refined Structure of Dienelactone Hydrolase at 1.8 A." J Mol Biol **214**(2): 497-525.

Pauwels, K., A. Lustig, L. Wyns, J. Tommassen, S. N. Savvides and P. Van Gelder (2006). "Structure of a Membrane-Based Steric Chaperone in Complex with Its Lipase Substrate." Nat Struct Mol Biol **13**(4): 374-5.

Perrakis, A., R. Morris and V. S. Lamzin (1999). "Automated Protein Model Building Combined with Iterative Structure Refinement." Nat Struct Biol **6**(5): 458-63.

Perry, K. D. (1961). "A Comparison of the Influence of *Streptococcus Lactis* and *Str. Cremoris* Starters on the Flavour of Cheddar Cheese." Journal of Dairy Research **28**: 221-229.

Potterton, E., P. Briggs, M. Turkenburg and E. Dodson (2003). "A Graphical User Interface to the Ccp4 Program Suite." Acta Crystallogr D Biol Crystallogr **59**(Pt 7): 1131-7.

Radisky, E. S., J. M. Lee, C.-J. K. Lu and D. E. Koshland, Jr. (2006). "Insights into the Serine Protease Mechanism from Atomic Resolution Structures of Trypsin Reaction Intermediates." Proceedings of the National Academy of Sciences of the United States of America **103**(18): 6835-6840.

Rahfeld, J., M. Schierhorn, B. Hartrodt, K. Neubert and J. Heins (1991). "Are Diprotin a (Ile-Pro-Ile) and Diprotin B (Val-Pro-Leu) Inhibitors or Substrates of Dipeptidyl Peptidase Iv?" Biochim Biophys Acta **1076**(2): 314-6.

Reiter, B., T. F. Fryer, M. E. Sharpe and R. C. Lawrence (1966). "Studies on Cheese Flavour." Journal of Applied Bacteriology **29**(2): 231-243.

Rigolet, P., I. Mechin, M. M. Delage and J. F. Chich (2002). "The Structural Basis for Catalysis and Specificity of the X-Prolyl Dipeptidyl Aminopeptidase from Lactococcus Lactis." Structure **10**(10): 1383-94.

Ro, H. S., H. P. Hong, B. H. Kho, S. Kim and B. H. Chung (2004). "Genome-Wide Cloning and Characterization of Microbial Esterases." FEMS Microbiol Lett **233**(1): 97-105.

Robillard, G. and R. G. Shulman (1974). "High Resolution Nuclear Magnetic Resonance Studies of the Active Site of Chymotrypsin. I. The Hydrogen Bonded Protons of The "Charge Relay" System." Journal of Molecular Biology **86**(3): 519-40.

Robinson, A., K. J. Edwards, P. D. Carr, J. D. Barton, G. D. Ewart and D. L. Ollis (2000). "Structure of the C123s Mutant of Dienelactone Hydrolase (Dlh) Bound with the Pms Moiety of the Protease Inhibitor Phenylmethylsulfonyl Fluoride (Pmsf)." Acta Crystallogr D Biol Crystallogr **56**(Pt 11): 1376-84.

Ronning, D. R., T. Klabunde, G. S. Besra, V. D. Vissa, J. T. Belisle and J. C. Sacchettini (2000). "Crystal Structure of the Secreted Form of Antigen 85c Reveals Potential Targets for Mycobacterial Drugs and Vaccines." Nat Struct Biol **7**(2): 141-6.

Roussel, A. and C. Cambillau (1989). Turbo-Frodo. Silicon Graphics Geometry Partner Directory. Moutain View, CA, Silicon Graphics**:** 77-88.

Sambrook, J., E. F. Fritsch and T. Maniatis (1989). Molecular Cloning: A Laboratory Manual. **1:** 1.74, 1.76.

Sarda, L. and P. Desnuelle (1958). "Action De La Lipase Pancreatique Sur Les Esters En Emulsion." Biochim Biophys Acta **30**: 513-521.

Schoemaker, H. E., D. Mink and M. G. Wubbolts (2003). "Dispelling the Myths--Biocatalysis in Industrial Synthesis." Science **299**(5613): 1694-7.

Schrag, J. D. and M. Cygler (1993). "1.8 a Refined Structure of the Lipase from Geotrichum Candidum." J Mol Biol **230**(2): 575-91.

Schrag, J. D., Y. G. Li, S. Wu and M. Cygler (1991). "Ser-His-Glu Triad Forms the Catalytic Site of the Lipase from Geotrichum Candidum." Nature **351**(6329): 761-4.

Schramm, V. L. and W. Shi (2001). "Atomic Motion in Enzymatic Reaction Coordinates." Curr Opin Struct Biol **11**(6): 657-65.

Schwede, T., J. Kopp, N. Guex and M. C. Peitsch (2003). "Swiss-Model: An Automated Protein Homology-Modeling Server." Nucleic Acids Res **31**(13): 3381-5.

Shah, D. S. and R. R. Russell (2004). "A Novel Glucan-Binding Protein with Lipase Activity from the Oral Pathogen Streptococcus Mutans." Microbiology **150**(Pt 6): 1947-56.

Shaw, E. J. (1999). The Characterisation of Tributyrin Esterase from *Lactococcus Lactis*. Institute of Molecular Biosciences. Palmerston North, Massey University. **Honours Thesis**.

Siek, T. J., I. A. Albin, L. A. Sather and R. C. Lindsay (1969). "Taste Thresholds of Butter Volatiles in Deodorised Butteroil Medium." Journal of Food Science **34**: 265-267.

Silverman, R. B. (2000). The Organic Chemistry of Enzyme-Catalysed Reactions, Academic Press**:** 19.

Spector, L. B. (1973). "Covalent Enzyme-Substrate Intermediates in Transferase Reactions." Bioorganic Chemistry **2**: 311-321.

Stadler, P., A. Kovac, L. Haalck, F. Spener and F. Paltauf (1995). "Stereoselectivity of Microbial Lipases. The Substitution at Position Sn-2 of Triacylglycerol Analogs Influences the Stereoselectivity of Different Microbial Lipases." Eur J Biochem **227**(1-2): 335-43.

Stadler, P., G. Zandonella, L. Haalck, F. Spener, A. Hermetter and F. Paltauf (1996). "Inhibition of Microbial Lipases with Stereoisomeric Triradylglycerol Analog Phosphonates." Biochim Biophys Acta **1304**(3): 229-44.

Steiner, T. (1998). "Hydrogen-Bond Distances to Halide Ion in Organic and Organometallic Crystal Structures: Up-to-Date Database Study." Acta Crystallographica Section B **54**: 456-463.

Stok, J. E., A. Goloshchapov, C. Song, C. E. Wheelock, M. B. Derbel, C. Morisseau and B. D. Hammock (2004). "Investigation of the Role of a Second Conserved Serine in Carboxylesterases Via Site-Directed Mutagenesis." Arch Biochem Biophys **430**(2): 247-55.

Sussman, J. L., M. Harel, F. Frolow, C. Oefner, A. Goldman, L. Toker and I. Silman (1991). "Atomic Structure of Acetylcholinesterase from Torpedo Californica: A Prototypic Acetylcholine-Binding Protein." Science **253**(5022): 872-9.

Suzuki, S. K., E. G. Hastings and E. B. Hart (1910). "The Production of Volatile Fatty Acids and Esters in Cheddar Cheese and Their Relation to the Development of Flavour." Journal of Biological Chemistry **7**: 431-458.

Szeltner, Z., D. Rea, V. Renner, V. Fulop and L. Polgar (2002). "Electrostatic Effects and Binding Determinants in the Catalysis of Prolyl Oligopeptidase. Site Specific Mutagenesis at the Oxyanion Binding Site." J Biol Chem **277**(45): 42613-22.

Szeltner, Z., V. Renner and L. Polgar (2000). "Substrate- and Ph-Dependent Contribution of Oxyanion Binding Site to the Catalysis of Prolyl Oligopeptidase, a Paradigm of the Serine Oligopeptidase Family." Protein Sci **9**(2): 353-60.

Tan, P. S., M. P. Chapot-Chartier, K. M. Pos, M. Rousseau, C. Y. Boquien, J. C. Gripon and W. N. Konings (1992). "Localization of Peptidases in Lactococci." Appl Environ Microbiol **58**(1): 285-290.

Tanizawa, K., Y. Kanaoka and W. B. Lawson (1987). "Inverse Substrates for Trypsin and Trypsin-Like Enzymes." Acc. Chem. Res. **20**: 337-343.

Taylor, C. M., C. J. O'Connor, J. L. Chamberlain and C. Q. Sun (2001). "Design, Synthesis and Evaluation of Two Diastereoisomeric Phosphonate Inhibitors of Lamb Pregastric Lipase." Journal of Molecular Catalysis B: Enzymatic **15**: 15-22.

Terwilliger, T. C. and J. Berendzen (1999). "Automated Mad and Mir Structure Solution." Acta Crystallogr D Biol Crystallogr **55**(Pt 4): 849-61.

Thomas, B. A., W. B. Church, T. R. Lane and B. D. Hammock (1999). "Homology Model of Juvenile Hormone Esterase from the Crop Pest, Heliothis Virescens." Proteins **34**(2): 184-196.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins (1997). "The Clustal_X Windows Interface: Flexible Strategies for Multiple Sequence Alignment Aided by Quality Analysis Tools." Nucleic Acids Res **25**(24): 4876-82.

Tsujita, T., A. Nakagawa, K. Shirai, Y. Saito and H. Okuda (1984). "Methyl Butyrate-Hydrolyzing Activity of Hepatic Triglyceride Lipase from Rat Post-Heparin Plasma." J Biol Chem **259**(18): 11215-20.

Tsujita, T., K. Shirai, Y. Saito and H. Okuda, Eds. (1990). Isozymes: Structure, Function, and Use in Biology and Medicine, Wiley-Liss.

Tsukada, H. and D. M. Blow (1985). "Structure of Alpha-Chymotrypsin Refined at 1.68 a Resolution." Journal of Molecular Biology **184**(4): 703-11.

Turner, J. M., N. A. Larsen, A. Basran, C. F. Barbas, 3rd, N. C. Bruce, I. A. Wilson and R. A. Lerner (2002). "Biochemical Characterization and Structural Analysis of a Highly Proficient Cocaine Esterase." Biochemistry **41**(41): 12297-307.

Vagin, A. and A. Teplyakov (1997). "Molrep: An Automated Program for Molecular Replacement." J. Appl. Cryst. **30**: 1022-1025.

Voet, D. and J. G. Voet (2004). Biochemistry. Hoboken NJ, John Wiley and Sons, Inc. **1:** 523.

Voet, D. and J. G. Voet 2004. Biochemistry. Hoboken NJ, John Wiley and Sons, Inc.

Vyas, N. K. (1991). "Atomic Features of Protein-Carbohydrate Interactions." Curr Opin Struct Biol **1**: 732-740.

Wang, J. H. (1970). "Directional Character of Proton Transfer in Enzyme Catalysis." Proceedings of the National Academy of Sciences of the United States of America **66**(3): 874-881.

Wei, Y., J. L. Schottel, U. Derewenda, L. Swenson, S. Patkar and Z. S. Derewenda (1995). "A Novel Variant of the Catalytic Triad in the Streptomyces Scabies Esterase." Nat Struct Biol **2**(3): 218-23.

Wolfenden, R. (1969). "Transition State Analogues for Enzyme Catalysis." Nature **223**: 704-705.

Zheng, J. (1998). Purification, Crystallization and Cloning of Tributyrin Esterase from Lactococcus Institute of Molecular Biosciences. Palmerston North, Massey University. **Master of Technology**

Appendices

Appendix A
General Crystallographic Methods

A.1 Data Collection

A1.1 Methodology and Instrument

Data were collected as oscillation photographs, using a Rigaku R-Axis IV++ imaging plate system, with a Rigaku RA-Micro7 Horizontal Rotaflex rotating anode generator fitted with a copper target. The potential of the generator was set at 40kV, with a current of 20mA. An Osmic Blue confocal optic producing a focal width of 0.3mm was used.

A.1.2 Initial Images

Initial exposures of 0.5^{o} width were taken at 0^{o} and 90^{o} (5 min exposure). Crystal Clear (Rigaku MSC) was used to index the images, generate an initial unit cell estimates, and to estimate mosaicity.

A.1.3 Collection Strategy

Crystal Clear was used to predict a dataset width that would produce a dataset with a completeness of at least 98% and a four fold redundancy. Image width was chosen based on mosaicity. Typically an image width of half the estimated mosaicity was chosen, with a minimum width of 0.25^{o} and a maximum of 0.5^{o}.

A.1.4 Crystal Mounting

To minimise the chance of producing inaccessible reflections, care was made to avoid mounting crystals with any major axis coincident with the axis of oscillation.

A.2 Cryocrystallography

Unless specified, data were collected at 120 K, using cryocrystallographic methods, with low temperature cooling provided by a Cryostream 700 low temperature device (Oxford Cryosystems).

A.3 Data Processing

Two approaches to data processing were used. The first was to use the Crystal Clear package, to collect images, integrate, scale and average the dataset. The second approach used MOSFLM (Leslie 1992) to integrate data, and SCALA from the CCP4i package to scale and average the dataset (Potterton, Briggs et al. 2003).

A.4 Model Building

Models were rebuilt using either TURBO-FRODO (AA7), or COOT (EstA and AZ4) (Emsley 2004).

A.5 Model Validation

Model validation was done during model building using the tools built in to COOT (density fit, rotamer, peptide omega, geometry and Ramachandran analyses). After model construction PROCHECK (Laskowski, MacArthur et al. 1993) was used to generate a report to identify potential problem areas. Baverage was also used for B-factor analyses (Dodson 1991).

A.6 Analysis and Graphics

Structural diagrams in this thesis were created using Pymol (DeLano 2002). Pymol was also used for structural analysis. COOT was also used for structural analysis, particularly for that involving electron density.

Appendix B
Common Names of Fatty Acids

This table summarises saturated selected fatty acids from C1 - C16, giving systematic and common names. Particular care should be taken with C6 - C10 acids, which have very similar common names.

Chain Length	**Systematic name**	**Common name**
C1	Methanoic	Formic
C2	Ethanoic	Acetic
C3	Propanoic	Propionic
C4	Butanoic	Butyric
C6	Hexanoic	Caproic
C8	Octanoic	Caprylic
C10	Decanoic	Capric
C12	Dodecanoic	Lauric
C14	Tetradecanoic	Myristic
C16	Hexadecanoic	Palmitic