Serine palmitoyltransferase; structural and mechanistic studies of a key enzyme in sphingolipid biosynthesis

Marine C.C. Raman

PhD

The University of Edinburgh

2010

To my family

ABSTRACT

Sphingolipids are a large family of ubiquitous molecules involved in many cellular events such as proliferation, differentiation and apoptosis. In all organisms the first and rate-limiting step of the sphingolipid biosynthetic pathway is catalysed by the pyridoxal 5'-phosphate (PLP)-dependent enzyme Serine Palmitoyltransferase (SPT). Many human diseases are caused by abnormalities in sphingolipid catabolism, and it was discovered that the human neurologic disease Hereditary Sensory and Autonomic Neuropathy type 1 (HSAN1) is caused by mutations of SPT.

In this study recombinant SPT from the sphingolipid-producing bacterium *Sphingomonas paucimobilis* was used as a model for the human enzyme in order to provide an insight into the effects of HSAN1 mutations on the structure and mechanism of the human isoform. A second *Sphingomonas* strain (*S. wittichii*) has been used to begin a study of the bacterial glycosphingolipid pathway since its whole genome sequence is available.

Three bacterial SPT HSAN1 mimic mutants N100W, N100Y and N100C have been successfully cloned, expressed and purified. Their ability to produce 3-ketodihydrosphingosine (KDS) was measured by radiochemical assay and a full kinetic evaluation carried out using a new spectrophotometric assay. Characteristic spectroscopic properties displayed by SPT provided useful information on substrate binding, specificity and on the formation of intermediates during the catalytic cycle. All three HSAN1 mutants were successfully crystallised in the holo-(PLP bound) form. Four high-resolution structures were solved in total including three in the holo-form (N100C: 1.25Å, N100W: 1.43Å and N100Y: 1.5Å) as well as the PLP: L-serine external aldimine complex of SPT N100Y (2.15Å), which was trapped by soaking the crystal with the L-serine substrate.

A putative SPT homologue was identified in *S. wittichii* and a recombinant form of the enzyme was cloned and characterised. After further analysis of the *S. wittichii* genome, a small ORF encoding a putative acyl carrier protein (ACP) of ~8.8 kDa was identified upstream of the SPT gene. A plausible hypothesis is that this ACP could be the acyl donor for its SPT partner, together forming a novel link between sphingolipid and fatty acid biosynthesis. The ACP and SPT genes were

successfully cloned and co-expressed in *E. coli*. Recombinant SPT *S. wittichii* was purified and shown to be enzymatically active. Crystals that diffracted to 2.1 Å were prepared and the *S. wittichii* SPT structure was solved by molecular replacement. Recombinant ACP was expressed in its apo-form form that required further *in vitro* modification to its active holo-form using a phosphopantetheinyl transferase (PPTase). An *S. wittichii* acyl carrier protein synthase (ACPS) gene was identified, the corresponding clone was designed and the protein successfully expressed and purified. ACPS *S. wittichii* was active and converted the apo-ACP into its holo-form. Transfer of acyl chains of increasing length (C10, C12, C14 and C16) onto ACP was investigated using several PPTases (ACPS *S. coelicolor*, Sfp *B. subtilis* and Svp *S. verticillus*). The holo-ACP was formed with two of the PPTs tested, Sfp and Svp. Only Sfp successfully transferred short acyl-chains onto ACP.

ACKNOWLEDGEMENTS

I wish to thank Dr. Dominic Campopiano for his supervision, advice, support and everlasting enthusiasm during these last three years. I am grateful to Prof. Jim Naismith for letting me use the SSPF facilities, sharing his knowledge and offering me help and guidance on structure solving. I thank the Chemistry Research School of Edinburgh and St. Andrews (EastCHEM) for funding my PhD.

I am indebted to Dr. Ken Johnson who allowed the chemist that I am to learn the secrets of protein crystallography. He is the one who taught me everything I know in this discipline.

Many thanks to Beverley Yard, Lester Carter and Stephen McMahon who all at some point participated and offered me help on the SPT project. I also wish to thank Prof. Bob Baxter for his helpful questions/discussions on SPT.

A special thank to Prof. Teresa Dunn who often received us in her lab in Washington and for all the exciting and invaluable discussions on our favourite protein.

I would like to thank Dave Clarke for his help and advice during my first year of research and for having the patience to run so many ACP samples during my last months in the chemistry building.

Of course, I thank everybody in the lab 229. A special mention to Scott who was undoubtedly the most helpful person in the lab.

Many thanks to all the project students who worked alongside me on the SPT project: Scott, Meike, George, Sarah, Kaj and Rosannah.

I am indebted to Dr. Isabelle Parrot for all her advice and support and without whom I would certainly not be here.

Thanks to Peter who took some of his time to read and correct parts of this thesis, for his friendship and for all the great meals shared together.

I am very grateful to my parents who supported me in all the ways someone can be supported.

And at last but definitely not the least, all my gratitude and love goes towards Jonathan, for all he did for me; patiently reading and correcting my thesis, facing stoically my susceptibility and especially for his constant support and encouragements which prevent me to fall in the dark side.

AUTHOR'S DECLARATION

I, Marine Raman, hereby certify that this thesis has been composed by myself, and that it is a record of my work, and that it has not been accepted in partial or complete fulfilment of any other degree or professional qualification.

PAGE OF CONTENTS

ABBREVIATIONS

AT	Acyltransferase
ACP	Acyl Carrier Protein
ACPS	Acyl Carrier Protein Synthase
AKB	2-amino-3-ketobutyrate
ALA	5-aminolevulinate
Ala	Alanine
ALAS	5-Aminolevulinate
AON	8-Amino-7-oxononanoate
AONS	8-Amino-7-oxononanoate synthase
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
BCA	Bicinchoninic acid
B. subtilis/*BS*	*Bacillus subtilis*
Bp	Base-pairs
BSA	Bovine serum albumine
CAI-1	*Cholerae* autoinducer-1
CoA/CoASH	Coenzyme A
CqsA	*Cholera* quorum sensing autoinducer
Da	Dalton
DH	Dehydratase
DHS	Dihydrosphingosine
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DoxSa	1-Deoxysphinganine
DTNB	5,5'-Dithiobis-2-nitrobenzoic acid
DTT	Dithiothreitol
E. coli	*Escherichia coli*
ER	Enoyl reductase
ESI	Electrospray ionisation

FAS	Fatty acid synthase
FT-ICR	Fourier Transform Ion Cylotron Resonnance
Fum8	Fumonisin 8
Gal/GalA	Galactose/ Galacturonic Acid
GlcA	Glucuronic Acid
GlcN	Glucosamine
Gln	Glutamine
Gly	Glycine
GSL	Glycosphingolipids
HBM	4-Hydroxy-2,2'-bipyrrole-5-methanol
H-bond	Hydrogen bond
HEPES	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonate
His	Histidine
HPLC	High performance liquid chromatography
HSAN1	Hereditary sensory and autonomic neuropathy type 1
IMAC	Immobilised metal affinity chromatography
IPTG	Isopropyl-β-D-galactoside
KBL	2-Amino-3-ketobutyrate-CoA ligase
k_{cat}	Catalytic constant
K_d	Dissociation constant
KDS	3-Ketodihydrosphingosine
K_m	Michaelis-Menten constant
KPhos	Potassium phosphate buffer
KR	Ketoreductase
KS	Ketosynthase
LAI-1	*Legionella* autoinducer-1
LB	Luria-Bertani Medium
LC	Liquid chromatography
LCB1/LCB2	Long chain base 1 and 2
Leu	Leucine
LqsA	*Legionella* quorum sensing autoinducer
Lys	Lysine

Man	Mannose
MAT	Malonyl/acetyl transferase
Met	Methionine
MPT	Malonyl/palmitoyl-transacylase
MS	Mass spectrometry
Ni-NTA	Nickel-nitrilotriacetic acid
NKT	Natural killer cell
NMR	Nuclear magnetic resonance
NRPS	Non-ribosomal peptide synthase
OD_{600}	Optical density at 600 nm
ORF	Open reading frame
PCP	Peptide carrier protein
PCR	Polymerase chain reaction
PEG	Polyethyleneglycol
Phe	Phenylalanine
PKS	Polyketide synthase
PLP	Pyridoxal-5'-phosphate
AOS	PLP-dependent α-oxoamines synthase
PP	Phosphopentetheinyl
PPT	Phosphopentetheinyl transferase
Pro	Proline
r.m.s.	Root mean square
rpm	Revolution per minute
S. coelicolor/SC	*Streptomyces coelicolor*
S. multivorum/SM	*Sphingobacterium multivorum*
S. paucimobilis/SP	*Sphingomonas paucimobilis*
S. verticillu/SV	*Streptomyces verticillus*
S. wittichii/SW	*Sphingomonas wittichii*
SDS-PAGE	Sodium dodecyl sulphate - polyacrylamide gel electrophoresis
Ser	Serine
SerT	Seryl transferase
SL	Sphingolipid

SOC	Super Optimal Culture
SPT	Serine palmitoyltransferase
TAE	Tris-acetate-ethylenediamine tetraacetate
TCEP	Tris(2-carboxyethyl)phosphine
TE	Thioesterase
TFA	Trifluoroacetic acid
Thr	Threonine
TLC	Thin Layer Chromatography
TM	Transmembrane
Tris	Tris(hydroxymethyl)aminomethane
Tyr	Tyrosine
WT	Wild type

CHAPTER I: INTRODUCTION

I.1 INTRODUCTION TO SPHINGOLIPIDS

I.1.1 Structure, Function, Pathways

It is now 125 years since the concept of sphingolipid-like molecules and their role in brain metabolism was first introduced by Johann Thudichum[1]. Intrigued by their structures and biological functions, he described them as "Sphinx-like" molecules, as these new compounds reminded him of the Sphinx riddles. A half–century later, Herbert Carter[2] determined the chemical structure of the family and named them sphingolipids. All sphingolipids contain a common 'sphingosine' component comprised of an unsaturated hydrocarbon chain and a polar head group arising from serine (Figure I-1). Diversity in the sphingolipid family arises from derivatisation of the sphingosine skeleton via acylation of the amino group with various fatty acids to form ceramides (Figure I-2). The addition of diverse polar head-groups on the ceramide hydroxyl leads to either sphingomyelins or glycosphingolipids (Figure I-2).

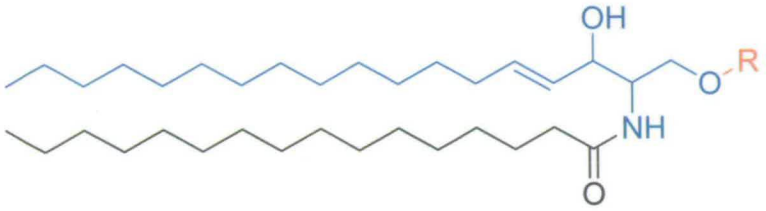


Figure I–1 General structure of the sphingolipids. Representation of their 3 distinctive parts, the sphingosine skeleton is coloured blue, the fatty acid group is black and the variable polar head group is red.

Sphingolipids are ubiquitous constituents of eukaryotic membranes and are also found in some prokaryotic organisms. Sphingolipids are known to be highly bioactive molecules and they are involved in many important functions in the body. Their metabolites act as second messengers which are involved in proliferation, differentiation and apoptosis[3, 4]. In association with cholesterol, they form lipid-rafts which are implicated in signal transduction and membrane trafficking[5]. Furthermore, sphingolipid metabolites play crucial roles in some pathological processes[6]. In

addition, malfunction of enzymes within the sphingolipid biosynthetic pathway can also lead to severe disease. For example the human neurological disease Hereditary Sensory and Neuropathy Type 1 (HSAN1) is caused by mutation of serine palmitoyltanserase[7] (SPT), the first enzyme in the pathway.

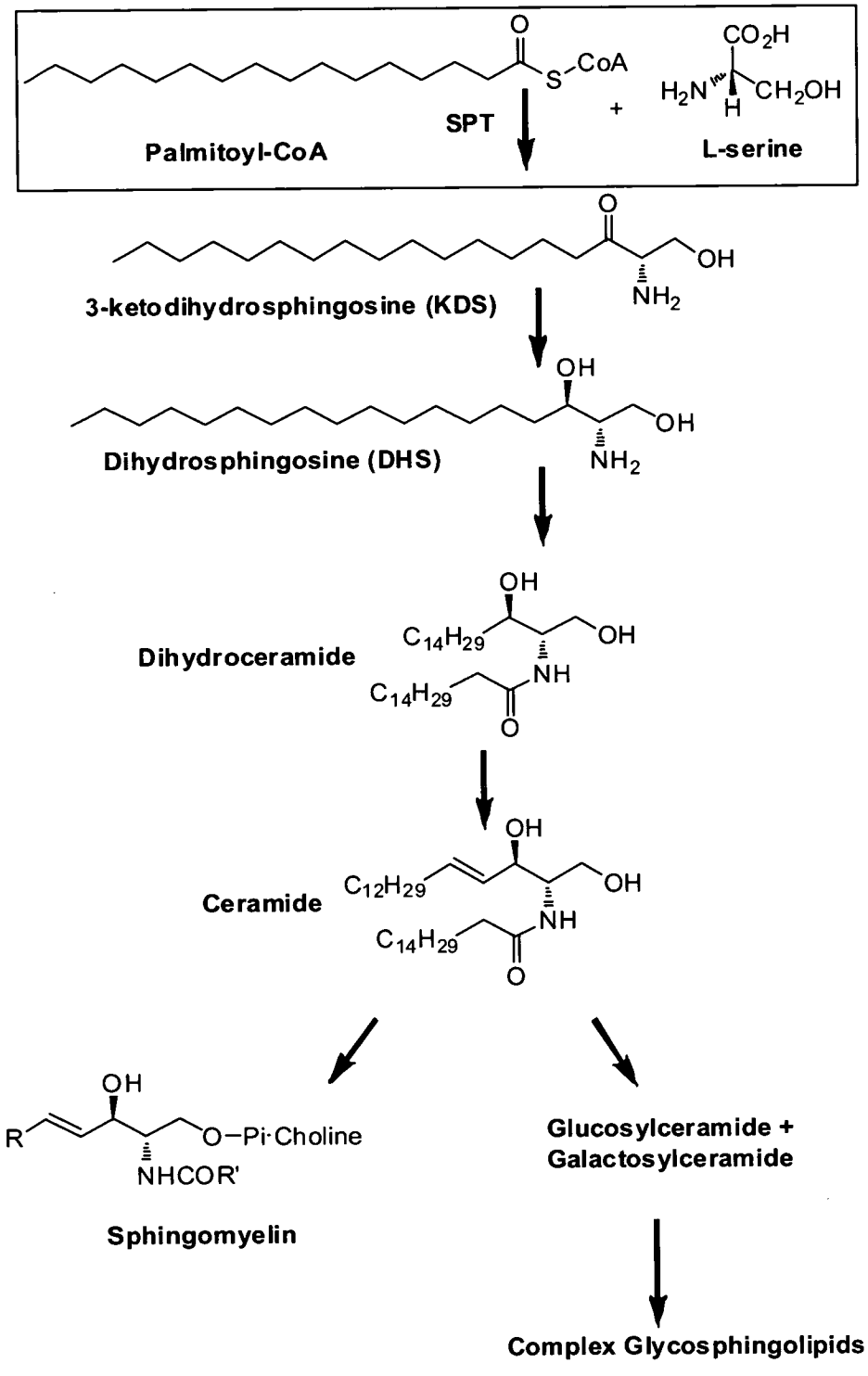


Figure I-2 Abbreviated sphingolipid biosynthetic pathway showing the early steps.

Sphingolipids (SLs) are found in a wide variety of species and some have their own sphingolipid biosynthetic pathways (Figure I-2). Eukaryotic SL pathways from organisms such as mammals, plants and fungi have been extensively studied and most of the intermediates and enzymes have been identified and characterised. Comprehensive reviews of the yeast and plant pathways were recently published[8, 9]. The first steps of SL biosynthesis, namely Claisen-condensation and ketone reduction are highly conserved. SLs and glycosphingolipids (GSLs) diversity occurs throughout the organisms with the attachment of various N-acyl chains, hydroxylation at specific carbons and addition of complex sugars. Interest about the bacterial SL pathway arose with the discovery of an atypical family of Gram-negative bacteria that have a membrane composed of GSLs (Figure I-3) instead of the usual lipopolysaccharides (LPS). These bacteria nammed "Sphingomonads" include organisms such as *Sphingomonas paucimobilis*, *Sphingobacterium multivorum, Sphingobium yanoikuyae and Sphingomonas wittichii*[10]. Since bacterial GSLs have similar core structures to their eukaryotic homologs it was assumed that the pathway would begin with the production of the common intermediate, 3-ketodihydrosphingosine (KDS) from L-serine and palmitoyl-CoA. This reaction is catalysed by the enzyme serine palmitoyltransferase (SPT; EC2.3.1.50). Indeed, several bacterial SPTs were isolated and characterised by Ikushiro *et al.*[11, 12].

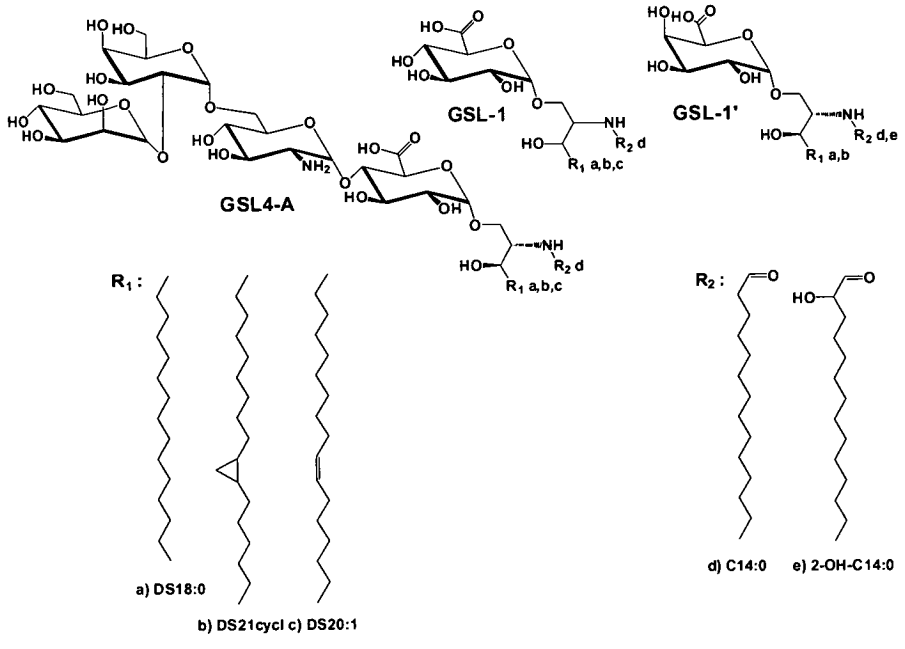


Figure I–3 Chemical structure of GSL-4A, GSL-1 and GSL-1' isolated from *S.paucimobilis* and *S.wittichii.*

I.1.2 Bacterial Glycosphingolipids

The structure and composition of GSLs isolated from "Sphingomonads" and in particular *S. paucimobilis* and *S. wittichii* has been extensively studied by Kawara *et al.*[13-17]. All bacterial GSLs have a similar fatty acid composition but the saccharidic moiety differs with species (Figure I-3). The GSLs of *S. paucimobilis* were shown to contain the three following dihydrosphingosines; *erythro*-1,3-dihydroxy-2-amino-octadecane (DS18:0), *erythro*-1,3-dihydroxy-2-amino-cis13,14-methylene-eicosane (DS21cycl) in comparable amounts and *erythro*-1,3-dihydroxy-2-amino-cis13,14-eicosene (DS20:1) as a minor component (Figure I-3 R_1 a,b,c). The fatty acid was shown to be 2-hydroxymyristic acid (2-0H-C14:0) (Figure I-3 R_2 d). They are also composed of two types of saccharide moiety, GSL-1 (Figure I-3) which contains only glucuronic acid (GlcA) and GSL-4A (Figure I-3) which is composed of mannose (Man), galactose (Gal), glucosamine (GlcN) and glucuronic acid (GlcA). *S. wittichii* GSLs contain only two dihydrosphingosine chains, 1,3-dihydroxy-2-amino-octadecane (DS18:0) and 1,3-dihydroxy-2-amino-cis13,14-methylene-eicosane (DS21cycl)(Figure I-3 R_1 a,b). The ceramide portion is composed of myristic acid (C14:0) or 2-hydroxy-myristic acid (2-0H-C14:0) (Figure I-3 R_2 d, e). One of the GSLs is GSL-1 present in *S. paucimobilis* and common to all *Sphingomonas* spp and the other is GSL-1' which contains a galacturonic acid (GalA) instead of the GlcA present in GSL-1. A first study on the configuration of the sugar-lipid linkage reports that the Galacto-ceramide of *S. yanoikuyae* was β[18]. After repeating these experiments and finding a different configuration, Kawara *et al.* extended their investigation and proved that all sugar-lipid linkages in bacterial GSLs were of the α-configuration[16, 17]. This result was further confirmed by the fact that the α-configurations of GlcA- and GalA-ceramides have recently been shown to be potent activators of the mammalian immune system via specific antigen receptors (CD1d) on Natural Killer T-cells (NKTs). The β-anomer present in eukaryotic GSL has no effect on NKT cell stimulation. This subject has been reviewed in a book chapter[19] (See appendices section) and will not be describe in further detail here.

I.2 SERINE PALMITOYLTRANSFERASE (SPT)

SPT catalyzes the first and rate-limiting step of the sphingolipid biosynthetic pathway in all organisms studied to date[20]. The reaction is a pyridoxal-5'-phosphate (PLP)-dependent, decarboxylative, Claisen condensation of the amino acid L-serine and the long chain (C16) fatty acid palmitoyl-CoA, which produces the sphingolipid precursor, 3-ketodihydrosphingosine (KDS, Figure I-2).

I.2.1 PLP-dependent Enzymes

I.2.1.1 Chemistry

PLP is the biologically active form of vitamin B6 and it is a cofactor in numerous enzymatic reactions involving various amino acid and amine substrates[21]. PLP-depedent enzymes are found to play important roles in all species of life and are involved in essential biochemical pathways. In spite of this diversity, the mechanisms of PLP-dependent enzymes share several common steps, the first one involving the initial Schiff's base exchange: the transimination step. The cofactor PLP is covalently bound in the catalytic site by formation of a Schiff's base (imine) with the amino group of a lysine residue, named the internal aldimine (Figure I-4). The next common step consists of the binding of the amino acid substrate and displacement of the lysine residue to form the external aldimine (transaldimination reaction, Figure I-4).

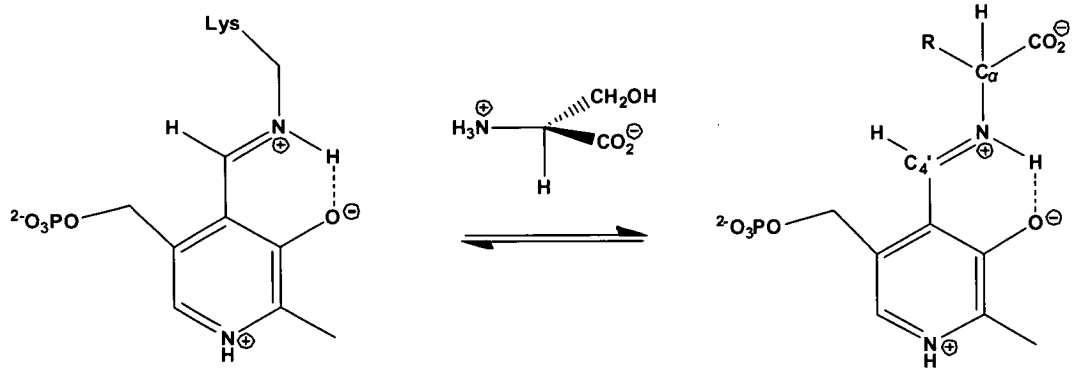


Figure I-4 Internal and external aldimine.

The variation in PLP-dependent enzymatic reactions arises from this common external aldimine intermediate and by the reaction at one of the amino acid carbons (α, β or γ). The PLP enzymes can be divided into three major families depending of which carbon modification they catalyse (Figure I-5).

1) α-carbon reactions: transamination, racemisation and decarboxylation (in green), elimination (aldol cleavage).
2) β-carbon reactions: elimination and replacement (in blue).
3) γ-carbon reactions: elimination and replacement (in pink).

All of these reactions have as a starting point the breakage of one three bonds the attached to Cα of the PLP:amino acid external aldimine and are described below (Figure I-5). The diversity of PLP-catalysed reactions has been extensively reviewed[22-24].

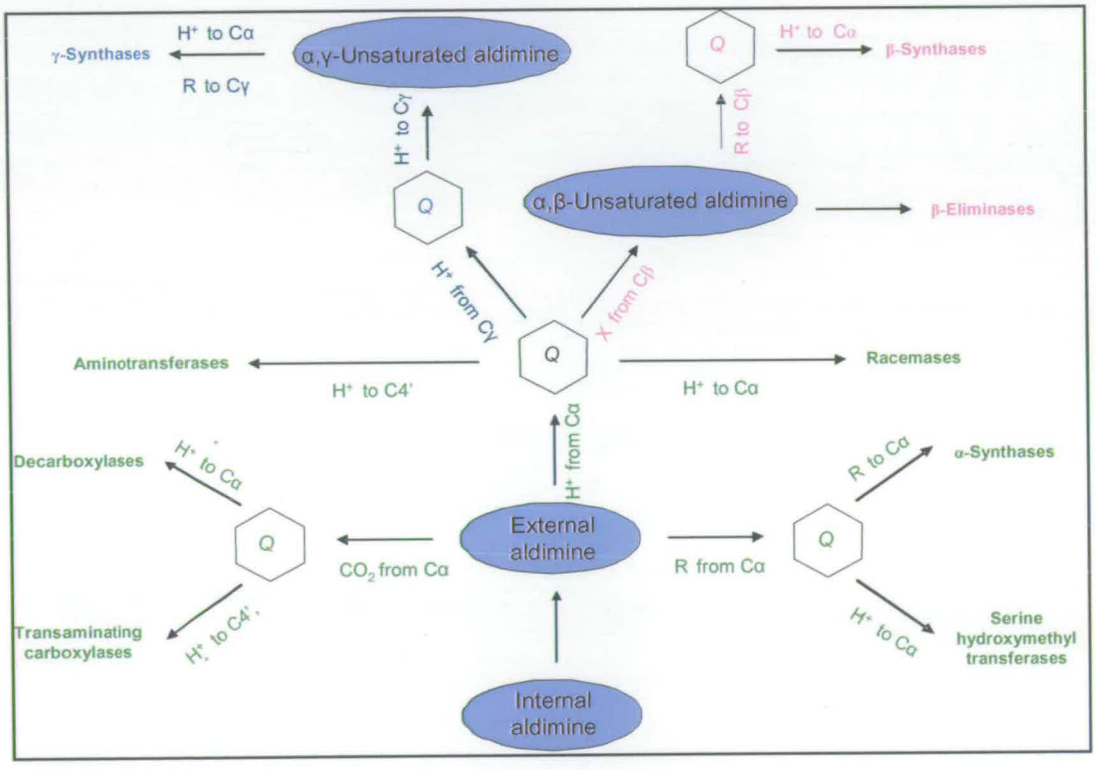


Figure I–5 The diversity of reactions catalysed by PLP dependent enzymes. Q indicates a quinonoid intermediate. Adapted from reference 24.

The reaction specificity is determined by the position of the three possible "leaving groups" relative to the plane of the pyridine ring in the external aldimine-protein

conjugate. In 1966 Dunathan [25] proposed that the transition state of the reaction must place the bond to be broken in a plane perpendicular to the pyridoxal imine π-system (Figure I-6A), whereby its p orbital is aligned with those of the conjugated Schiff base/pyridine ring π-system. The bond so oriented would be selectively labilized. The three conformations described below (Figure I-6B) will yield different types of reactions and reaction products. Conformation (I) activates the Cα-H bond and leads to transamination, α,β-elimination, β,γ-elimination and β-decarboxylation. The second conformation (II) is favourable for the loss of R group and conformation (III) favours α-decarboxylation.

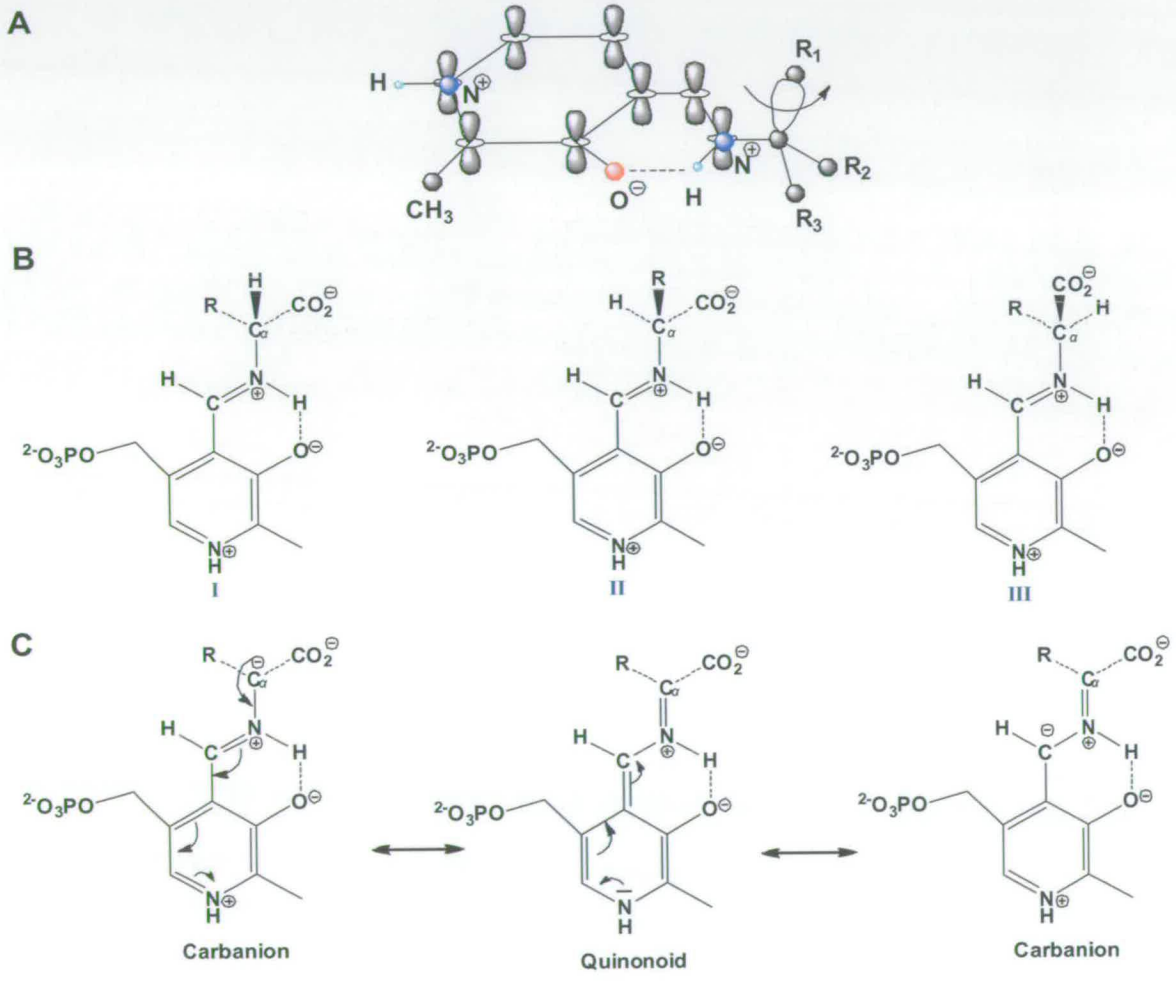


Figure I–6 Illustration of Dunathan hypothesis (A) Stereoelectronic control of reaction specificity in PLP-catalyzed reactions (B) The 3 possible conformations of the external aldimine Cα (C) Principal resonance forms obtained after external aldimine deprotonation (carbanion and quinonoid). Adapted from references 25 and 26.

The bond breaking steps are facilitated by the formation of a carbanion that is stabilised by the electron sink property of the pyridine ring of the PLP cofactor. The stabilised carbanion intermediate is named a quinonoid in that particular resonance form (Figure I-6C). The exact roles played by the protonated/deprotonated states of the Schiff's base and the pyridine nitrogen on the carbanion stabilisation process has been discussed recently by Toney[26, 27] and Bach[28].

Each PLP-dependent enzyme works using this common mechanism. The only exceptions are the PLP-dependent phosphorylases. The chemistry of these enzymes is completely different and the phosphate group of the PLP is used in catalysis but at the present its role is unclear [29].

I.2.1.2 Classification

Due to the large number of PLP-dependent enzymes and the vast chemical diversity they exhibit, a classification was proposed by Christen *et al.*[30] based on the chemical characteristics of the catalysed reactions. This classification divides the PLP-enzymes into three classes: α, β, γ depending on the carbon atom of the substrate involved in the chemistry. Later, another classification was proposed by Goldsmith *et al.*[31] that further separated this superfamily into five different structural fold types. These five classes are discussed in two reviews by Jansonius[23] and Schneider[24]. Enzymes were classified according to their amino acid sequences, secondary structure predictions and available three-dimensional structures. The five classes are often named by the first enzyme characterised for the group. This classification includes; the aspartate aminotransferase family (fold type I), the tryptophan synthase β family (fold type II), the alanine racemase family (fold type III), the D-amino acid aminotransferase family (fold type IV) and the glycogen phosphorylase family (fold type V). This last family is sometimes considered as separate from the others due to its utilisation of PLP-phosphate group in the catalysis of its reactions.

The Fold type I or aspartate aminotransferase family

Of the five fold type families, the Fold-type I family of enzymes has been the best characterised and it is further divided into six subclasses. Aminotransferases catalyse the amino group transfer between an amino acid and an oxoacid. The three-dimensional structures of this group reveal them to be homodimers containing two actives sites. Each subunit of this dimer contains a large and a small domain. The small domain contains the C and N termini. The PLP is attached to the large domain by a covalent bond with a conserved lysine residue and residues from both subunits are necessary for the formation of an active catalytic site.

The Fold type II or tryptophan synthase β family

This family is characterised by three different three-dimensional structures. The first is represented by the tryptophan synthase β from *Salmonella typhimurium* and forms a $\alpha_2\beta_2$ tetramer where β is the PLP catalytic subunit and α the regulatory subunit. Threonine deaminase is an α_4 tetramer which is a dimer of a "catalytic" dimer. Each monomer consists of a PLP-binding site in the N-terminal domain, which form the catalytic part, and a regulatory domain situated at the C-terminus. Finally, O-acetylserine sulfhydrylase is an active α_2 dimer. Each subunit contains a PLP-binding domain but lacks a regulatory domain.

The Fold type III or alanine racemase family

Enzymes of this group are homodimers. Each monomer contains two domains, an α/β barrel and a domain mainly comprising β strands. These enzymes contain only one molecule of PLP which is bound at the mouth of the barrel between the two subunits. The active-site lysine is situated at the C-terminal end of the first strand.

The Fold type IV or D-amino acid aminotransferase family

Here, the enzyme is homodimeric and contains a unique active site at the interface of the two subunits. The monomers are composed of a large C-terminal domain and a small N-terminal domain. The PLP binds to a lysine residue situated on the C-terminus but residues of both subunits contribute to the catalytic site. The active sites

of D-amino acid transferases are virtually mirror images of the active site of the fold type I aminotransferase.

The Fold type V or glycogen phosphorylase family

These enzymes are multidomain proteins and contain three distinctive domains, the C- and N-terminal domains and a glycogen-binding domain. The C-terminal domain binds the PLP and also contains a dinucleotide-binding fold. The action of PLP in this family is totally different from those mentioned above and instead of playing a role as electrophilic catalyst; the phosphate group is used for proton transfer.

I.2.2 PLP-dependent α-oxoamines Synthase Family (AOS)

SPT belongs to the α-oxoamine synthase (AOS) subfamily of PLP-dependent enzymes, which contains three other well characterized members: 8-amino-7-oxononanoate synthase (AONS)[32, 33], 5-aminolevulinate synthase (ALAS)[34-36], and 2-amino-3-ketobutyrate-CoA ligase (KBL)[37]. These enzymes catalyze the Claisen-like condensation between a specific pair of amino acid and an acyl-coenzyme A thioester[21] substrates to produce α-oxoamines (Figure I-7). Three of the four enzymes catalyze decarboxylation of the amino acid (SPT, AONS and ALAS) but KBL is unusual as it produces a carboxylated product, 2-aminoketobutyric acid. Each enzyme is involved in a different, important biosynthetic pathway (Table I-1).

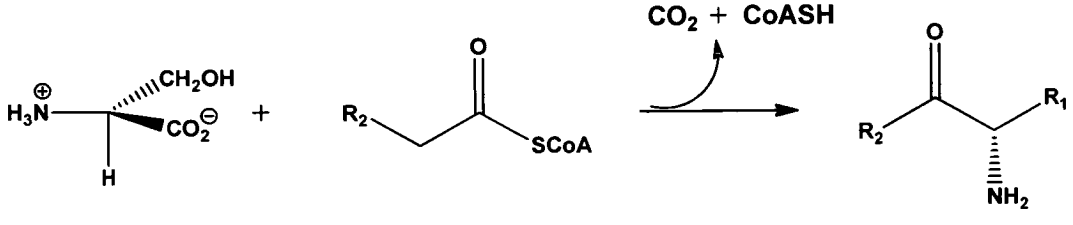


Figure I–7 The decarboxylative Claisen-like condensation between an amino acid and thioester substrates catalysed by enzymes in the AOS family.

Structural studies have shown AONS, ALAS, and KBL isoforms from various species to be homodimers[32, 36, 37]. In marked contrast, eukaryotic SPTs are heterodimeric, membrane-bound proteins consisting of two subunits, SPT1 and SPT2

(encoded by the *lcb1* and *lcb2* genes, respectively)[38]. However, a water soluble homodimeric enzyme has been isolated from the cytoplasm of the bacterium *Sphingomonas paucimobilis*[11] and sequence alignment data has shown it displays ~30% sequence homology with the other members of α-oxoamine synthase family. The first high-resolution x-ray crystal structure of the *S. paucimobilis* holo-SPT was recently solved[39] in our group in collaboration with Prof. Jim Naismith (Figure I-8D). The four AOS structures closely resemble each other and they are all intimate, interlocked dimers with each subunit comprising three domains (N-terminal, central catalytic and C-terminal, see Figure I-8D). The two PLP-containing active sites are situated at the dimer interface. Key residues (Lys, His, Asp) are responsible for the stabilisation of the cofactor and are highly conserved throughout the family (Figure I-8E).

Table I-1 Members of the PLP-dependent α-oxoamines synthase family (AOS)

Amino acid	Thioester	Enzyme	Product	Pathway	PDB files
L-Alanine	Pimeloyl-CoA	8-amino-7-oxononanoate synthase (AONS)	8-amino-7-oxononanoate (AON)	Biotin	*E. coli* apo AONS **1BSO** *E. coli* AONS complexed with PLP and 8-amino-7-oxononanoate **1DJ9**
Glycine	Acetyl-CoA	2-amino-3-ketobutyrate-CoA ligase (KBL)	2-amino-3-ketobutyrate (AKB)	Threonine degradation salvage	*E. coli* holo-KBL **1FC4**
Glycine	Succinyl-CoA	5-aminolevulinate acid synthase (ALAS)	5-aminolevulinate (ALA)	Tetrapyrrole	*R. capsulatus* holo-ALAS **2BWN** *R. capsulatus* ALAS complexed with glycine **2BWP** *R. capsulatus* ALAS complexed with succinyl-CoA **2BWO**
L-Serine	Palmitoyl-CoA	Serine palmitoyltransferase (SPT)	3-ketodihydrosphingosine (KDS)	Sphingolipids	*S. paucimobilis* apo-SPT **2JGT** *S. paucimobilis* holo-SPT **2JG2**

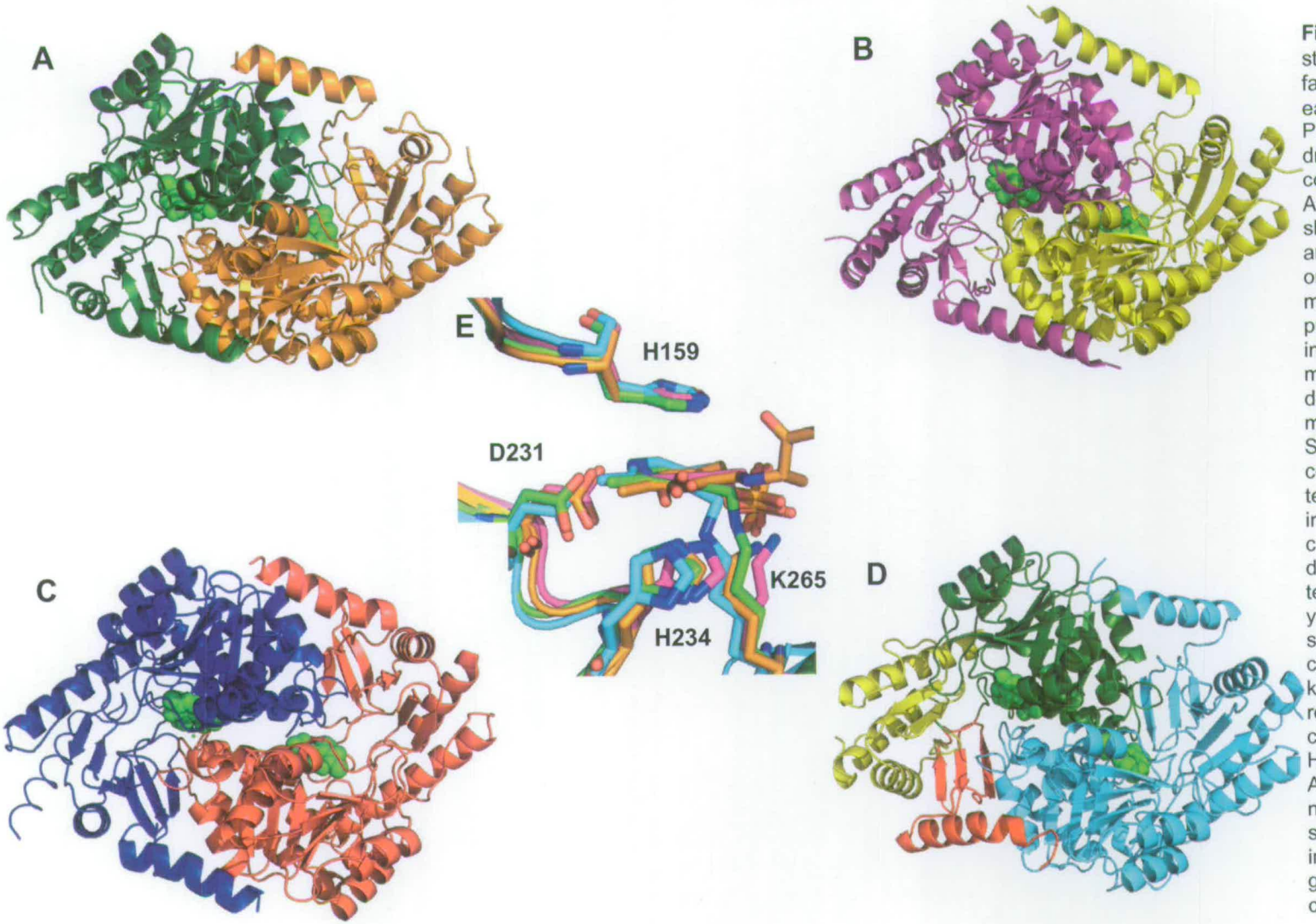


Figure I–8 Overall structure of the AOS family homodimers. In each enzyme, the two PLP molecules are drawn in spheres and coloured green. (A) AONS monomer A is shown in dark green and monomer B in orange (B) KBL monomer A is shown in purple and monomer B in yellow (C) ALAS monomer A is shown in dark blue and monomer B in red (D) SPT monomer A is coloured with the N-terminal domain shown in red, the central catalytic domain in dark green, and the C-terminal domain in yellow. The second subunit is shown in cyan. (E) Overlay of key conserved residues involved in cofactor binding: His234, His159, Asp231 Lys265 (SPT numbering). AONS shown in orange, KBL in magenta, ALAS in green and SPT in cyan.

Since all four of the AOS enzymes catalyse similar type of condensation reaction, structural and functional studies have led to the proposition of a common mechanism[20, 33, 40-43] comprising the following steps: formation of an external aldimine via displacement of the lysine-PLP internal aldimine (holo-SPT) by the incoming amino acid substrate; formation of a quinonoid intermediate by abstraction of the α-proton from the PLP-amino acid external aldimine; a Claisen-like condensation with the fatty acid-CoA thioester substrate, followed by displacement of the CoASH to form a α-ketoacid intermediate; decarboxylation of this species to form a product quinonoid; protonation of this quinonoid to form the product external aldimine; and finally release of the α-oxoamine product and regeneration of the enzyme PLP-internal aldimine (Figure I-9).

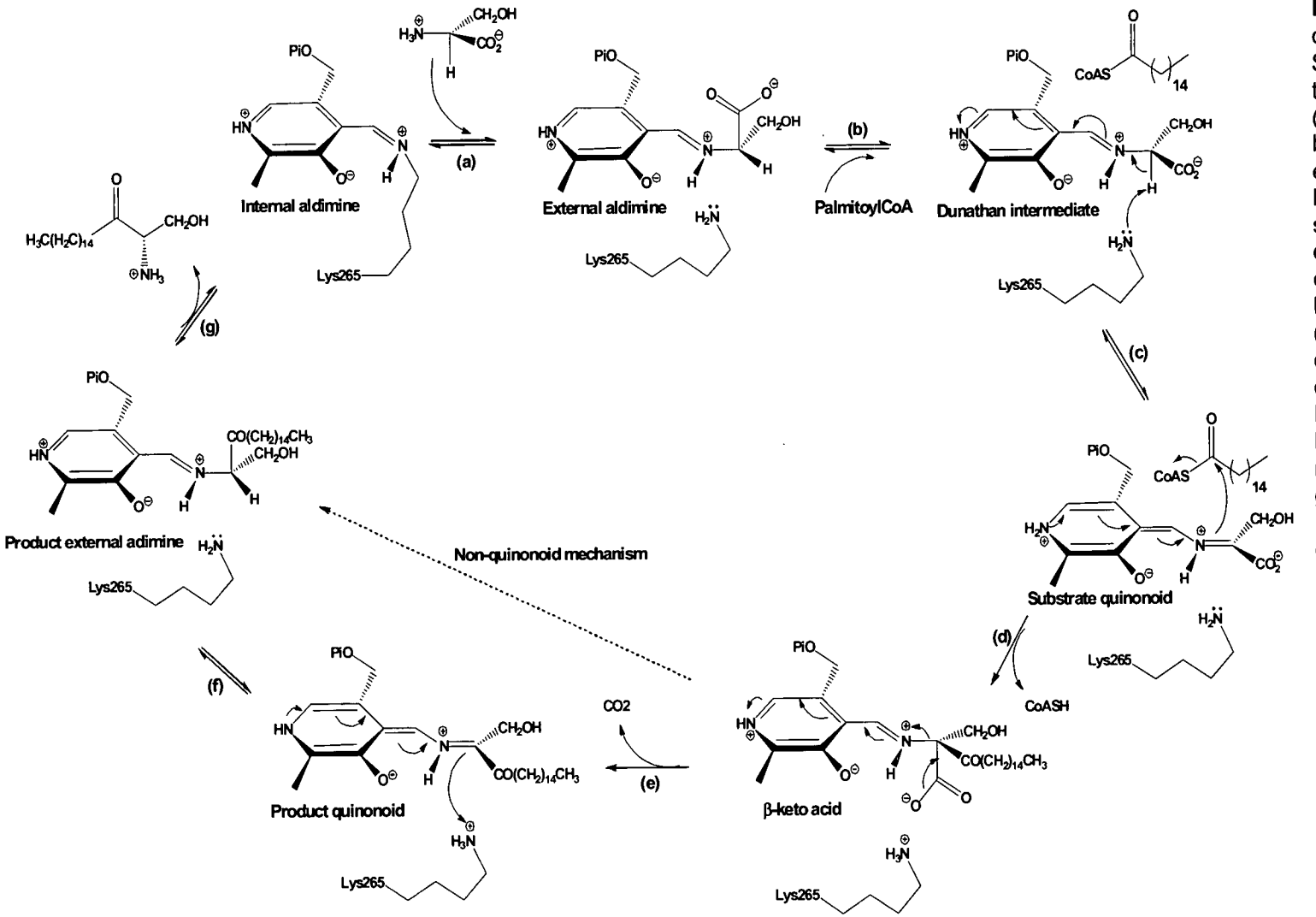


Figure I–9 Proposed catalytic mechanism of SPT and AOS family. (a) the internal aldimine (holo-form) is displaced by L-serine to form the external aldimine; (b), binding of second substrate palmitoyl-CoA causes conformational change to give the Dunathan intermediate; (c) formation of the quinonoid by deprotonation of Cα hydrogen; (d) thioester bond hydrolysis and release of CoASH to give β-keto acid intermediate; (e) decarboxylation to form KDS product quinonoid; (f) reprotonation to form KDS product external aldimine; (g) KDS product release and reformation of the holo-form. The *dotted line* shows that steps *e* and *f* may be bypassed by an alternative mechanism that does not proceed via a product quinonoid. Adapted from reference 44.[44]

Although PLP-dependent AOS enzymes have been extensively studied periodically newly characterized enzymes appear in the literature that display interesting twists and surprises. The AOS family has recently expanded with the characterisation of four new members (CqsA, LqsA, Fum8 and SerT). CqsA[45, 46] catalyses the condensation between L-threonine and/or L-2-aminobutyric acid and decanoyl-CoA and is involved in the biosynthesis of the cholerae autoinducer-1 (CAI-1) or its intermediate amino-(CAI-1)(Figure I-10A 1 and 2). LqsA[47] is part of the gene cluster producing the Legionella autoinducer-1 (LAI-1) (Figure I-10B) and its substrates remain undetermined at the moment. CqsA and Lqsa are involved in quorum sensing by production of the volatile auto-inducers which are able to "escape" their original cell to go on to activate receptors of neighbouring cells. The structure of CqsA has been recently solved by two independent research groups[45, 46] and both these structures revealed that CqsA displayed the same characteristics as the others members of the AOS family. CqsA is a homodimeric enzyme with two active sites and both monomers are needed to form an active enzyme. The structure of the decanoyl-CoA complex revealed a hydrophobic tunnel to which this substrate binds. Fum8p[48] is involved in chain release in the biosynthesis of the fungal natural product fumonisin and catalyses the condensation between L-alanine and steroyl-ACP to produce the corresponding sphingosine (Figure I-10C). Finally, SerT[49], subunit of a multidomain NRPS-PKS protein, catalyses the condensation between L-serine and pyrrolyl-β-ketoacyl-S-ACP to form 4-hydroxy-2,2'-bipyrrole-5-methanol (HBM) (Figure I-10D) and is involved in prodigiosin biosynthesis.

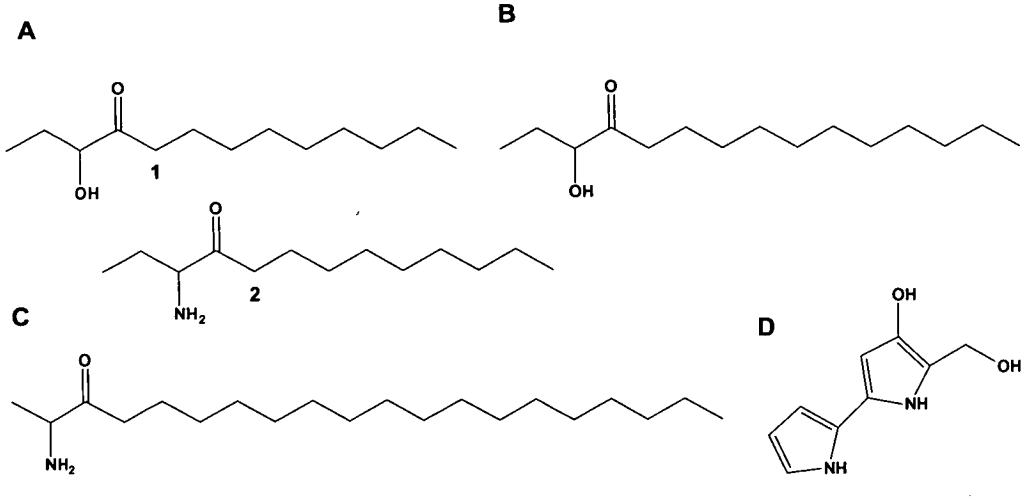


Figure I–10 Condensation products of (A) CqsA (B) LqsA (C) Fum8 (D) SerT.

I.2.3 Serine palmitoylransferase isozymes

I.2.3.1 Eukaryotic SPTs

SPT was first isolated from *S. cerevisiae* by Pinto *et al.* in their work on mutant strains which required an external supply of phytosphingosine for growth[50, 51]. Subsequently, mammalian homologs were identified in mice[52], hamster CHO cells[53] and humans[54]. Sequence alignment of these proteins has shown that there is ~ 40% amino acid sequence identity between mammalian and yeast and ~ 95% homology between mammalians types. It was shown[51, 53] that mammalian and yeast SPTs comprise two non-identical subunits each encoded by two different genes LCB1 and LCB2. In the human genome, the LCB1 gene is located in the chromosome 9q21-q22 region. It comprises 15 exons and contains around 85kbp. The second gene, LCB2, is located in the chromosome 14q24.3-q31 region and is formed by 12 exons and its size is 110kbp. Weiss *et al.*[54] and Hanada *et al.*[55] have both successfully purified the two protein subunits and characterised the isolated SPT enzyme but only in microgram amounts. Based on their encoded gene sequences the hSPT1 and hSPT2 monomers are proteins of 53kDa and 63kDa respectively and sequencing alignment shows ~ 20% identity between the two subunits. The recent identification of a hLCB3 gene expressed mainly in placenta (with homology to hLCB2) suggests greater complexity of the enzyme than previously thought.[56] A highly hydrophobic domain situated at the N-terminus of hSPT1 is thought to play the role of a membrane anchor.[52, 54] Studies by Mandon *et al.*[57] showed that SPT is localised in the ER with membrane spanning domains in both the C- and N-termini. The major part of the enzyme faces the cytosol allowing the substrates free access to their active sites.

In yeast SPT1, three other hydrophobic domains have been identified.[58] The first domain is located between residues 50-84. It is not required for stability, membrane association, interaction with ySPT2 protein or enzyme activity. The two other domains are situated between 342-371 and 425-457 and are required for protein stability. Another subunit of ~10kDa encoded by the Tsc3 gene has been identified by Teresa Dunn and her colleagues[59] This protein is associated with the yeast

SPT1/SPT2 heterodimer but is not essential for SPT activity.[60] It appears to play a role in regulation and is required for optimum activity of SPT. The same research group very recently identified small subunits (ssSPTa and ssSPTb) of mammalian SPT that increase SPT activity over 10 fold and confer distinct acyl-CoA substrate specificities.[61] These two small subunits possess a central hydrophobic domain predicted to reside in the membrane. Each small subunit interacts with both hSPT1 and hSPT2.

For both human and yeast SPT, residues constituting the catalytic site are highly conserved (Figure I-11). These residues are situated on the SPT2 subunit around the catalytic lysine residue involved in the PLP binding. Although SPT2 is considered as the "catalytic" unit and SPT1 the "regulatory" one both of the subunits are required to obtain an active enzyme.

Figure I–11 The catalytic site, a highly conserved domain with the active site lysine highlighted (h, human; y, yeast; b, bacterial; v, viral).

More recent studies showed that eukaryotic SPTs are even more complex proteins. Two research groups[62, 63] identified the conserved endoplamic reticulum (ER) membranes proteins Orm1 and Orm2 as negative regulators of sphingolipids biosynthesis. Breslow *et al.*[62] showed that Orm proteins and yeast SPT formed a high molecular weight complex (the SPOTS complex) constituted of Orm1, Orm2, SPT1, SPT2, Tsc3 and the phosphatase Sac1. The exact role of these other subunits in the regulation of SPT activity awaits further analysis.

I.2.3.2 Viral SPT

Recently, a novel single chain SPT has been identified in Coccolithovirus.[64] This enzyme contains two different domains comparable to the eukaryotic SPT but these are linked together, in a head-to-tail fashion (Figure I-12). The N-terminal domain is related to the SPT2 subunit and the C-terminus to the SPT1 and both

display high overall amino acid sequence homology to other members of the family. The catalytic site is also highly conserved (Figure I-11). However, the viral SPT lacks the N-terminal membrane binding region that exists in the mammalian and yeast enzymes. It is noteworthy that this hydrophobic domain is not present in other AOS enzymes and is not essential for enzymatic activity. Substrate specificity has been observed for this enzyme which has a preferred affinity for myristoyl-CoA instead of palmitoyl-CoA.

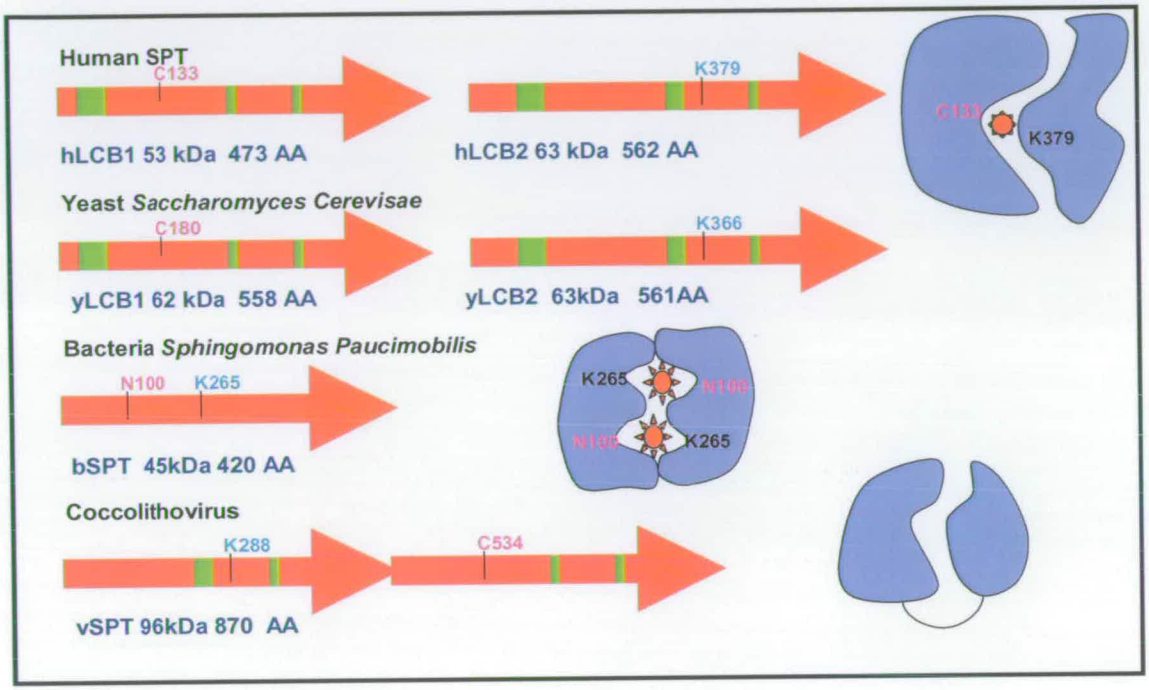


Figure I–12 Subunit composition of SPT isozymes. Green bars represent hydrophobic transmembrane domains.

I.2.3.3 Bacterial SPTs

A water soluble homodimeric SPT was isolated from *S. paucimobilis* by Ikushiro *et al.*[11] in 2001. This homodimer is composed of two monomers of 45kDa and contains two active sites (Figure I-12). Each monomer is able to bind a molecule of PLP but the two subunits are required to obtain the active enzyme. Recently, three novel bacterial SPTs from *Sphingobacterium multivorum, Sphingobacterium spiritivorum* and *Bdellovibrio stolpii*, have been isolated and purified by Ikushiro and colleagues.[12] Each enzyme is a water-soluble active homodimer that displays ~30% homology with the two human subunits. Interestingly the *S. multivorum* and

Bdellovibrio stolpii enzymes are bound to the inner membrane of cells which suggests that these enzymes could be a prokaryotic model of the membrane bound eukaryotic SPT.

I.3 HEREDITARY SENSORY AND AUTONOMIC NEUROPATHY TYPE 1 (HSAN1)

HSAN1 is the most common inherited disorder of peripheral sensory neurons and involves an autosomal dominant progressive degeneration of dorsal root ganglia and motor neurons.[7] In the early stages of disease a symmetric sensory loss is observed, whereafter symptoms progress slowly and affect the lower limbs disproportionately. This loss of sensation leads to painless injuries, predisposition to tissue/bone infections and mutilation of the feet. Deafness has been reported in some families.[65] As the disease progresses, there is also degeneration of motor neurons with secondary denervation atrophy and weakness of distal limb muscles.[66, 67] The first symptoms start to appear between 20 and 30 years old and the expression of the disease is different by gender within the same family.[68] Symptoms start later in women and they are, in general, less severe than in men. It is interesting to note that HSAN1 specifically affects the peripheral neurons; no adverse effect on the function of other neuronal tissues such as the CNS was reported. HSAN1 is a genetically heterogeneous disease for which 4 variants genes have been reported.[69-71] The HSAN1 locus was initially mapped to chromosome 9q22.1-22.3[72], then more specifically by Dawkins *et al.*[69] who assigned it to the SPTLC1 gene, *lcb1*, encoding the SPT1 subunit. Three missense mutations concerning two amino acids have been identified.[69] Two of these are located in exon 5 and result in the replacement of a cysteine residue by a tryptophan or a tyrosine (C133W or C133Y, Figure I-13). The third mutation is that of a valine residue to an aspartate (V144D, Figure I-13). These three mutations are located within a twelve amino acid segment encoded by exons 5 and 6. In initial studies[69], it was believed that these mutations lead to an increase of SPT enzymatic activity that resulted in high levels of ceramides and their metabolites present in organisms carrying this mutation. However in 2002, Bejaoui *et al.*[70] showed that HSAN1 leads to a decrease of SPT activity and that mutations confer dominant negative effects on the SPT enzyme. A fourth mutation linked to HSAN1

has been identified by Verhoeven *et al.*[71] This mutation is located in exon 13 and concerns a missense mutation of a glycine residue to an alanine (G387A, Figure I-13). Unlike the two other amino acids, this glycine is not involved in the same amino acid segment. Although patients had a typical HSAN1 phenotype, it has been recently shown that the G387A mutation is not disease associated.[73] It is possible that G387A has a bystander effect by increasing the risk to develop a HSAN phenotype in conjunction with another, not as yet identified, mutation.[73]

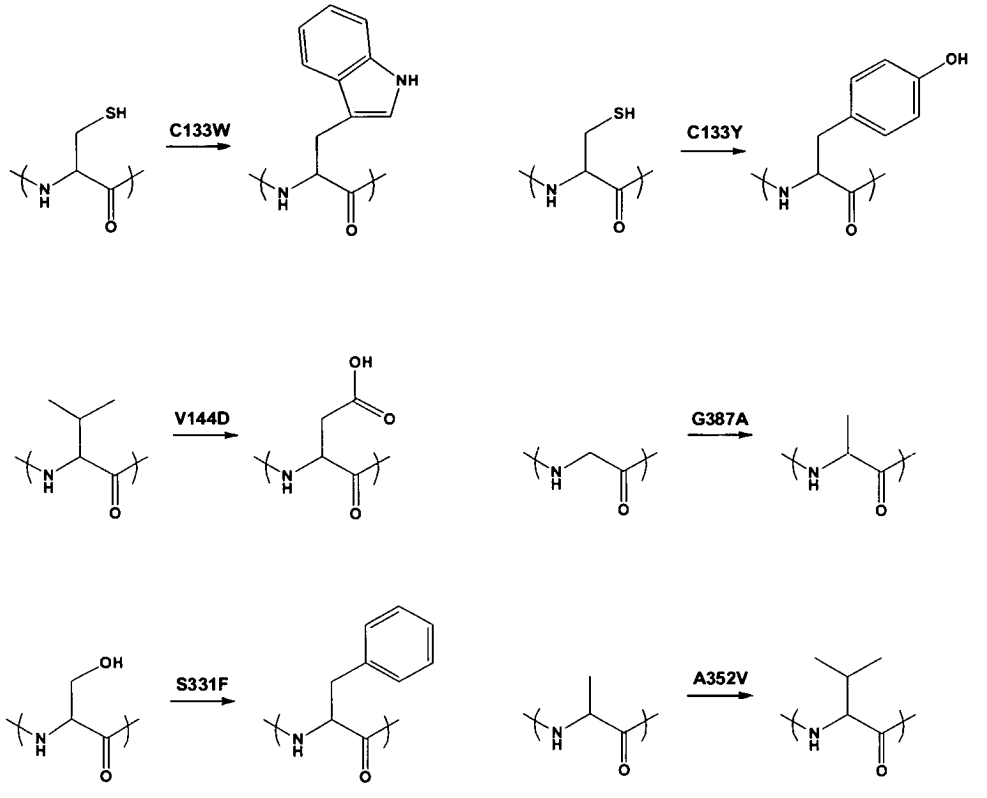


Figure I–13 HSAN1 mutations on SPT1.

More recently, two new mutations that cause HSAN1 (S331F and A352V, Figure I-14) have been identified downstream of the previously characterised HSAN1 mutations[74]. The S331F mutation changes a serine residue that is highly conserved in eukaryotic SPTs (Figure I-14). Interestingly this mutation was not carried by the parents of the patient and must have arisen by spontaneous mutation. The resulting phenotype is very severe and is characterised by a congenital onset and mental retardation. In an independent study Dunn and colleagues found several LCB1 mutants that complemented mammalian cells lacking the small subunits (ssSPTa and ssSPTb); incredibly, the LCB1 mutant S331F was the best at complementing the

cells and also at increasing basal SPT activity in the absence of small subunits. This mutant was not activated by ssSPTa but was activated by ssSPTb. According to the most recent topological study of LCB1[58], the S331F mutation is predicted to be located on a loop between TM2 and TM3 in the lumen of the ER membrane. In contrast, the most common HSAN1 mutations C133Y, C133W and V144D are all located between TM1 and TM2 of LCB1; this section of the protein is predicted to face the cytosol and contains the residues that interface with LCB2 at the enzyme active site. Therefore it is likely that S331F affects interaction between LCB1 and ssSPTa and results in a highly compromised enzyme function causing symptoms more severe than those associated with the most common HSAN1 mutations. The pathology of the second mutation A352V could not be confirmed and the affected amino acid is not well-conserved among the species. Furthermore the amino acid change is mild considering the chemical properties of alanine and valine.

Figure I–14 SPT1 alignment-HSAN 1 mutations. Identic amino acids are highlighted in black. Chemically similar amino acids are highlighted in grey. Amino acid involved in HSAN 1 mutations are highlighted in magenta.

The changes created in SPT and their effect on the enzyme are not well understood but the mutated cysteine residue involved in both C133W and C133Y mutants is thought to lie in close proximity to the PLP binding site. Therefore, it has been postulated that the mutation affects PLP binding and catalysis[70, 75]. Hornemann *et al.*[76-78] propose that HSAN1 mutations alter the amino acid specificity of the mutant SPT enzyme leading to the formation of two toxic side-products, 1-deoxysphinganine (DoxSA) and 1-amino-2-deoxy-*n*-heptadecane (ADHD) (Figure I-15), formed by the condensation of palmitoyl-CoA with L-alanine and glycine. The hypothesis is that these lipids are not metabolised through normal SL catabolism and build up slowly over time to a toxic level. Support for this comes from the fact the the DoxSA and ADHD lipids do show toxicity towards human cells.

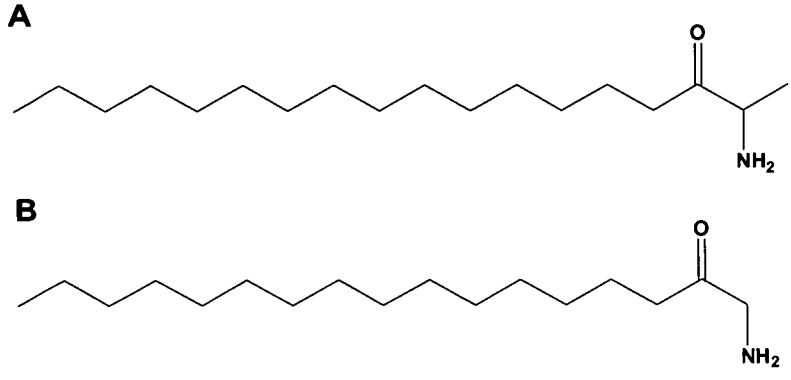


Figure I–15 Deoxysphingosine bases formed by condensation of palmitoyl-CoA and (A) L-alanine, (B) glycine.

As the human SPT is a membrane bound enzyme, mechanistic and structural studies of both wild type and mutants SPTs would be extremely difficult (less than 1% of available structures in the PDB are membrane proteins).[79] To gain insight into the nature of these mutations, Yard *et al.*[39] used bioinformatic analysis to map the cysteine residue (C133) of the human SPT onto the asparagine (N100) on *S. paucimobilis* SPT. Analysis of the tertiary structure of the wild type SPT showed that the side-chain of this asparagine residue is located near to the catalytic site and is involved in a hydrogen-bond with the backbone chain supporting the lysine residue (Lys 267) binding the PLP (Figure I-16). The impact of HSAN1-like mutations at this N100 residue on the structure, function and enzymatic activity of the bacterial SPT is the focus of part of this thesis.

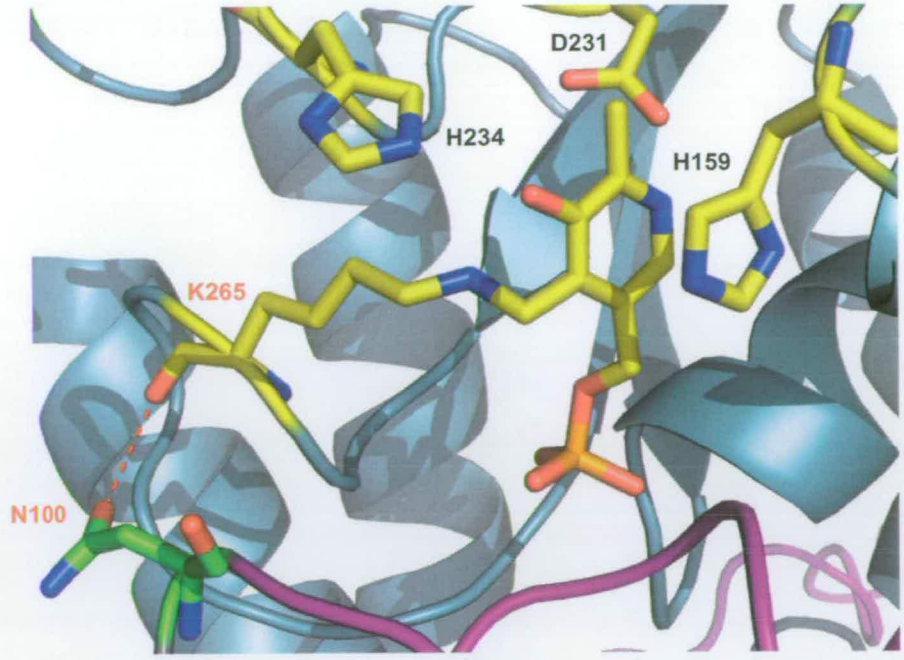


Figure I–16 Location of Asn100 in the active site of *S. paucimobilis* SPT. Structure of the SPT-L-serine external aldimine complex. Catalytic site of SPT dimer showing one monomer in teal and the other in purple. The internal aldimine and the residues His234, His159, Asp23 from the teal monomer are shown in stick form and are coloured yellow. The residues N100 from the purple monomer is coloured green. This clearly shows that the Asn100 from one monomer interacts with the active site of the other monomer.

I.4 ACYL CARRIER PROTEIN (ACP)

Acyl carrier proteins (ACPs) are highly conserved small acidic proteins made up of approximately 70-80 amino acids. ACPs require post-translational activation by addition of a 4'-phosphopanthetheinyl (4'-PP) group from coenzyme A onto a specific serine residue; this modification is carried out by phosphopantheinyl transferase enzymes (PPTase). ACPs are multifunctional and have been shown to be used in several essential primary metabolic pathways. For example, they are an indispensable component of the fatty acid synthases (FAS), where they are part of a multi-domain protein (type I) or are an independent subunit (type II). In bacteria, ACPs or PCPs (peptide carrier proteins) are involved in polyketide synthase (PKS) and non-ribosomal peptide synthase (NRPS) which produce a large diversity of natural products such as polyketides and peptide antibiotics.[80]

I.4.1 Fatty Acid Synthase (FAS)

I.4.1.1 Type I FAS

Eukaryotic ACPs are small components of large multifunctional proteins called fatty acid synthases (FAS). Mammalian FAS is an homodimeric enzyme of ~250 kDa[81, 82]. Each subunit comprises three N-terminal domains separated from four C-terminal domains by a central "core" domain devoid of any catalytic activity. The order of functional domains was established as β-ketoacyl synthase (KS), malonyl/acetyl transferase (MAT), dehydratase (DH), enoyl reductase (ER), β-ketoacyl reductase (KR), ACP and thioesterase (TE)(Figure I-17A).

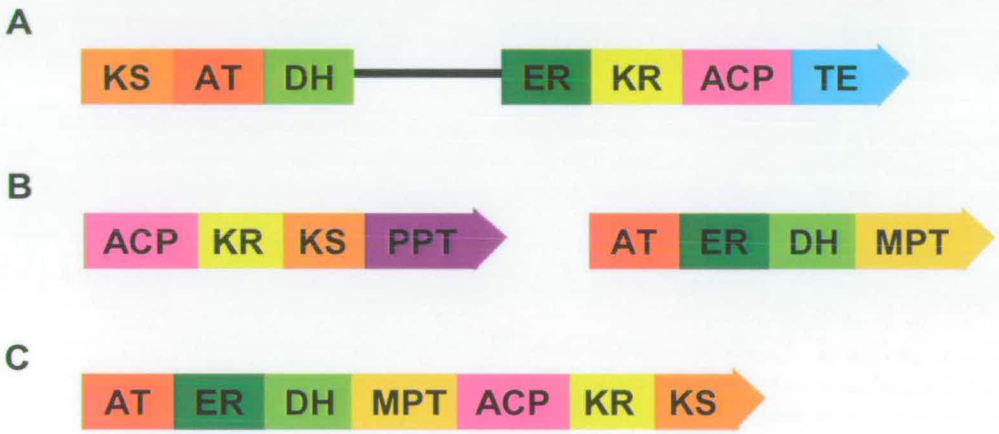


Figure I-17 FAS type I systems (A) Mammalian FAS (B) Yeast FAS (C) Mycobacterium FAS KS: ketoacyl synthase, AT: acyltransferase, DH: dehydratase, ER: enoyl reductase, KR: ketoreductase, ACP: acyl carrier protein, TE: thioesterase, PPT: phosphopantetheine transferase, MPT: malonyl/palmitoyl-transacylase.

The ACP domain has spatial access to all FAS components. Recently, the porcine FAS crystal structure was solved[83] and this revealed an asymmetric X-shaped conformation with two reaction chamber (Figure I-18). With a structural resolution of 4.5 Å each catalytic domain was identified in the complex apart from the ACP and TE domains which could not be assigned, presumably because of their inherent flexibility.

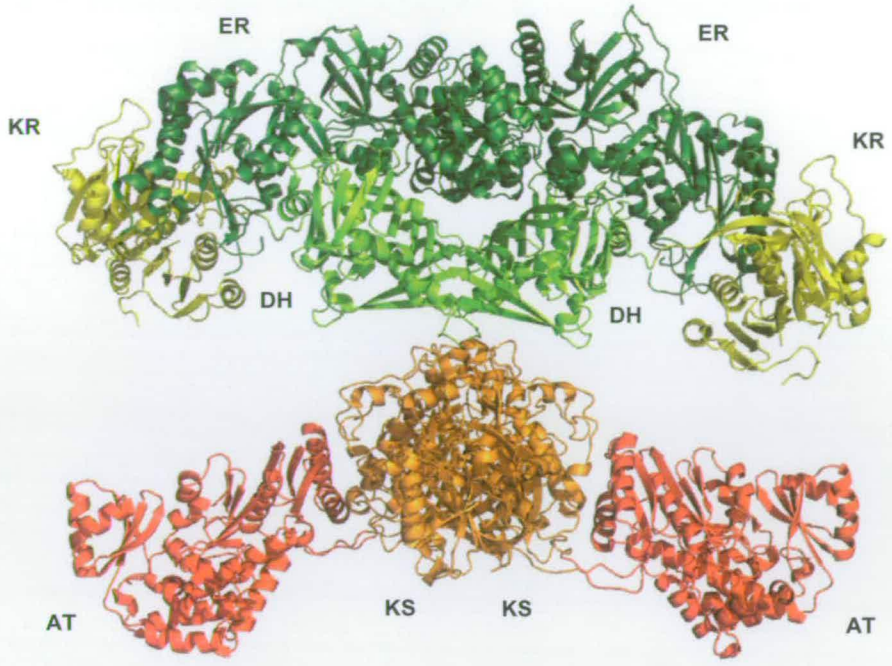


Figure I-18 Mammalian FAS type I (PDB: 2VZ8). Each domain is represented by a different color. The ketoreductase (KR) domains are coloured yellow. The enoyl reductase (ER) domains are represented in dark green. The dehydratase (DH) domains are drawn in pale green. The acyltransferase (AT) domains are red and the ketoacyl synthase (KS) domains are orange.

The structure of a fungal FAS has been determined by Jenni *et al.*[84] Like the mammalian FAS, the ACP domain was not "visible" in this structure. In contrast to mammalian FAS, the fungal and yeast enzymes comprise two different polypeptides (~185 and 180 kDa) which aggregate into a multimeric protein ($\alpha_6\beta_6$) resulting in the formation of six reaction chambers (Figure I-19). Each α and β subunit contains four functional domains (Figure I-17B). Recently, two research groups[85, 86] crystallised and solved the yeast multidomain protein structure including the ACP and PPT domains (Figure I-19). Interestingly, the PPT domain is situated on the outside of the assembly line suggesting modification of apo-ACP before the complete formation of the complex.

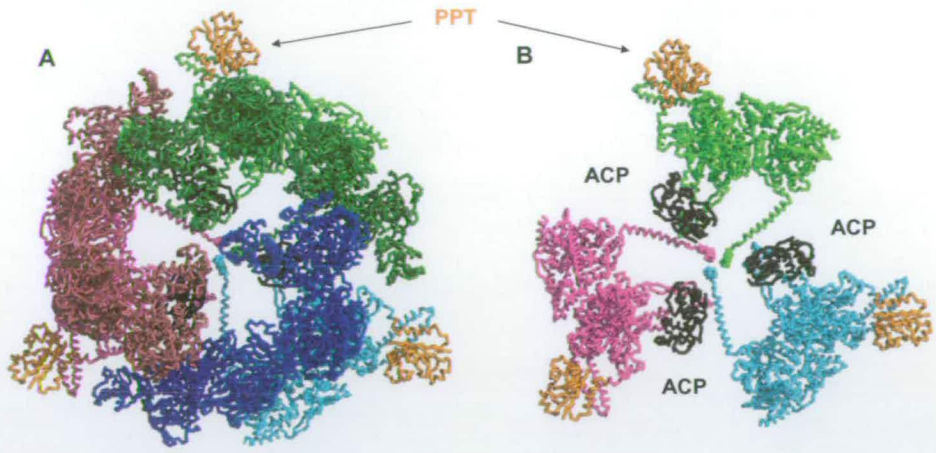


Figure I-19: Yeast FAS type I (PDB: 2CDH). (A) Multimer α3β3 top view (B) Multimer α3 top view. β subunits are coloured in dark purple, blue and green, α subunits are coloured in pale purple, blue and green. ACP subunits are coloured dark grey and the PPT subunits are coloured orange.

It is interesting to note that as well as higher organisms certain bacteria such as *Mycobacterium tuberculosis* possess a type I FAS.[87] These enzymes are α6 hexamers and each subunit contains seven domain, acyltransferase (AC), enoyl reductase (ER), dehydratase (DH), malonyl/palmitoyl-transacylase (MPT), acyl carrier protein (ACP), ketoacyl reductase (KR) and ketoacyl synthase (KS) (Figure I-17C).

I.4.1.2 Type II FAS

Type II FAS are found in bacteria, eukaryotic plastids and mitochondria[81]. Each FAS component is an individual protein encoded by a discrete gene. The growing acyl intermediates remain covalently bound to an independent ACP which shuttles them along the synthetic pathway. The FAS type II system from *E. coli* has been extensively studied and the structures of all of the individual FAS enzymes are now available[88]. Previous X-ray and NMR studies on bacterial and plant ACPs had shown that they share a high degree of structural similarity[88]. Indeed, ACPs from different organisms are interchangeable.[88, 89] The typical, highly-conserved 3-D structure of an ACP is characterised by 4 α-helices and a long structured loop between helix 1 and 2 (Figure I-20).

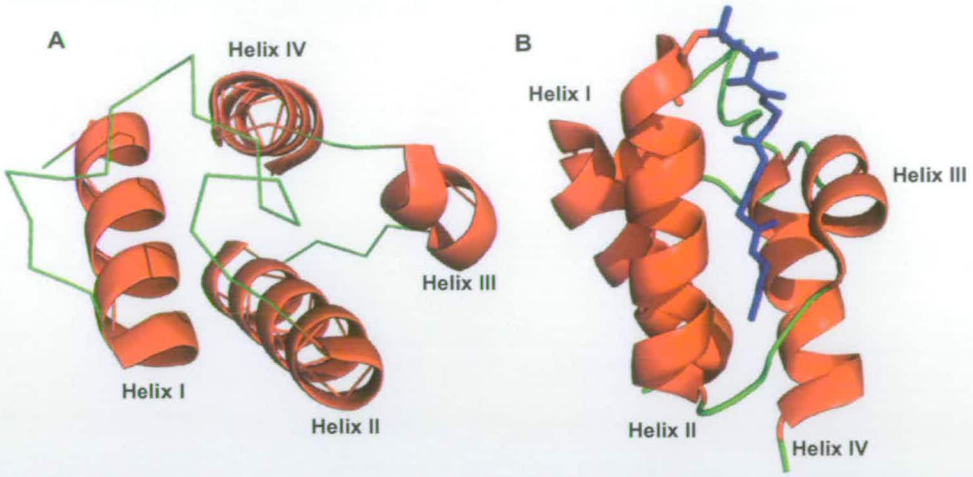


Figure I-20 ACP type II structure from *E. coli* (PDB: 1LOH) (A) Top view (B) Side view. The prosthetic and C8 acyl groups are represented in stick form and coloured blue.

High sequence conservation is observed in the second helix[90] (Figure I-21) and it is thought to be a "recognition" helix toward others enzymes of the FAS complex[91-93]. The unique motif DSL, situated at the end of helix 2, is present in all ACP sequences (Figure I-21). The 4'-PP prosthetic group derived from coenzyme A is attached to the serine residue of the DSL motif in the ACP holoform. The acyl chain binding site was determined by NMR and crystallographic characterisation of the *E. coli* butyryl-ACP intermediate revealing a hydrophobic cavity formed by helices 2 and 3 (Figure I-20)[94]. Further X-ray studies on longer acyl chains (C6, C7 and C10) showed that the hydrophobic cavity can expand in order to accommodate the growing acyl chain.[95]

Figure I-21 Sequence alignment of ACPs highlighting secondary structure elements. Identic amino acids are highlighted in black. Chemically similar amino acids are highlighted in grey.

Mitochondrial FAS type II are largely present in eukaryotes and their principal function is to synthesise octanoyl-ACP, a precursor of lipoic acid which is a key cofactor for pyruvate dehydrogenase and others mitochondrial enzymes.[96]

An interesting atypical group of type II ACPs is found in *Mycobacterium tuberculosis*.[97] These proteins possess a longer C-terminus than the other ACPs and

are involved in the synthesis of very long chain fatty acids (C50, C60) from shorter acyl chains released by their FAS type I system.

I.4.2 Polyketide synthase (PKS) and non-ribosomal peptide synthetase (NRPS)

Polyketides and non-ribosomal peptides are a large class of natural products synthesised by bacteria using acyl-CoA and amino acids as building block.[81, 98, 99] Although these two classes of compounds are chemically and functionally different, the enzymatic systems used for their synthesis share many similarities. Both molecules are synthesised via megasynthases (~200-2000kDa) comprising repeated functional units named "modules". Each module has a different enzymatic activity. Different types of PKS exist: type I is the multidomain protein, the type II system is an assembly of discrete proteins and type III uses acyl-CoA instead of acyl carrier protein (by analogy with FASs).[100, 101] Only megasynthase systems are known for the NRPS. It is interesting to note that some ACPs have been found as free-standing proteins or incomplete modules.[80, 102] The overall structure and organisation is similar in PKS and NRPS. They are both made up of 3 "core" modules; two catalytic and one carrier (ACP for PKS and PCP, peptide carrier protein, for NRPS), some auxiliary domains (reduction, dehydratation, epimerisation) and two domains for initiation and termination (Figure I-22).

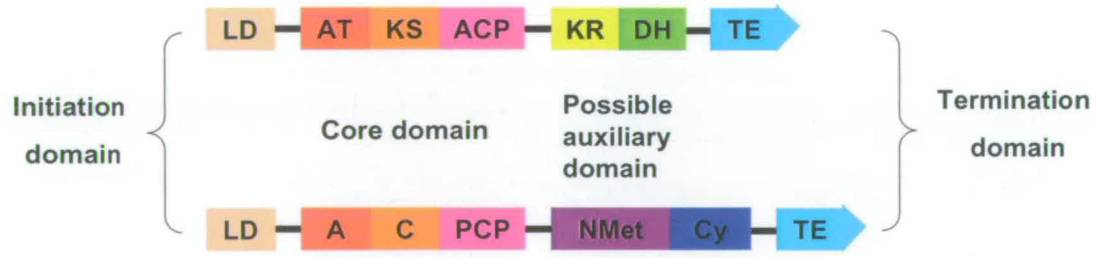


Figure I-22 PKS and NRPS systems. Top: PKS domains; AT: acyltransferase, KS: ketosynthase, ACP: acyl carrier protein, KR: ketoreduction, DH: dehydratation. Bottom: NRPS domains; A: adenylation, C: condensation, PCP: peptide carrier protein, NMet: N-methylation, Cy: cyclisation. LD: loading, TE: thioesterase.

Despite the different chemistry, formation of C-C bond for polyketides and C-N bond formation for non-ribosomal peptides, the chain-elongation strategy is the same; the required building block (acyl- or AA) is transferred from an ACP or PCP domain to the growing chain by a condensation reaction (Figure I-23).

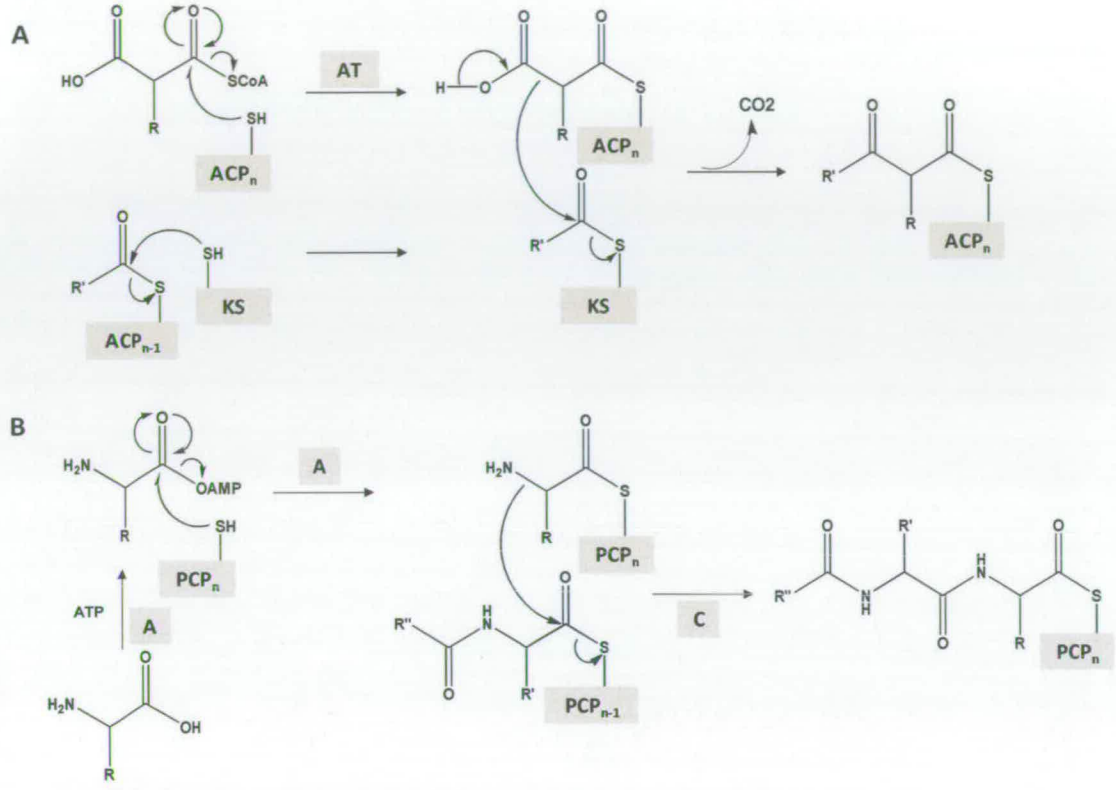


Figure I-23 Condensation reactions with (A) PKS (B) NRPS.

I.4.3 Acyl Carrier Protein Synthase (ACPS)

As was already mentioned, to become catalytically active, an ACP must be converted from the inactive apo-form to the active holo-ACP by specific PPTase [81, 103] (Figure I-24). These enzymes transfer the 4'-PP moiety of coenzyme A (CoA) to the side chain hydroxyl of a serine residue invariant in all ACPs. This post-translational phosphopantetheinylation introduces a covalently-attached nucleophilic thiol on a long tether that becomes the site of all the initiation and acyl transfer events involved in the assembly of the broad array of natural products synthesized by these enzymes.

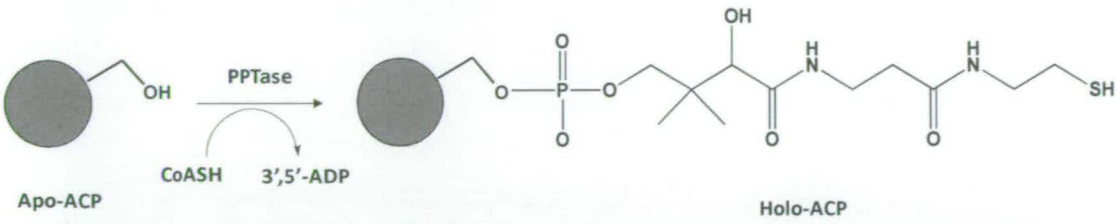


Figure I-24 Conversion of apo-ACP to holo-ACP.

The PPTase superfamily is composed of two distinct subgroups.[103] The first subgroup (Type I PPT) includes the PPTases that modify peptide carrier protein in NRPSs. These enzymes are generally ~230 amino acids long. For example the Sfp from *B. subtilis* involved in surfactin synthetase, belongs to this group. Its structure has been solved by Reuter *et al.*[104] and showed that at least two monomers were needed to form the active enzyme. The second subgroup (type II PPT) includes acyl carrier protein synthase (ACPS) involved in fatty acid synthase. They are smaller proteins comprising ~120 residues. Three crystal structures have been obtained with ACPS from *S. pneumoniae*[105] and *B. subtilis* ACPS[106] in complex with coenzyme A and ACP. The latest structure showed that these PPTases are actually trimeric enzymes which form three active sites at the subunit interfaces (Figure I-25A). These two subgroups share low sequence similarities although some residues are highly conserved throughout the species (Figure I-25B). [103, 104, 106]

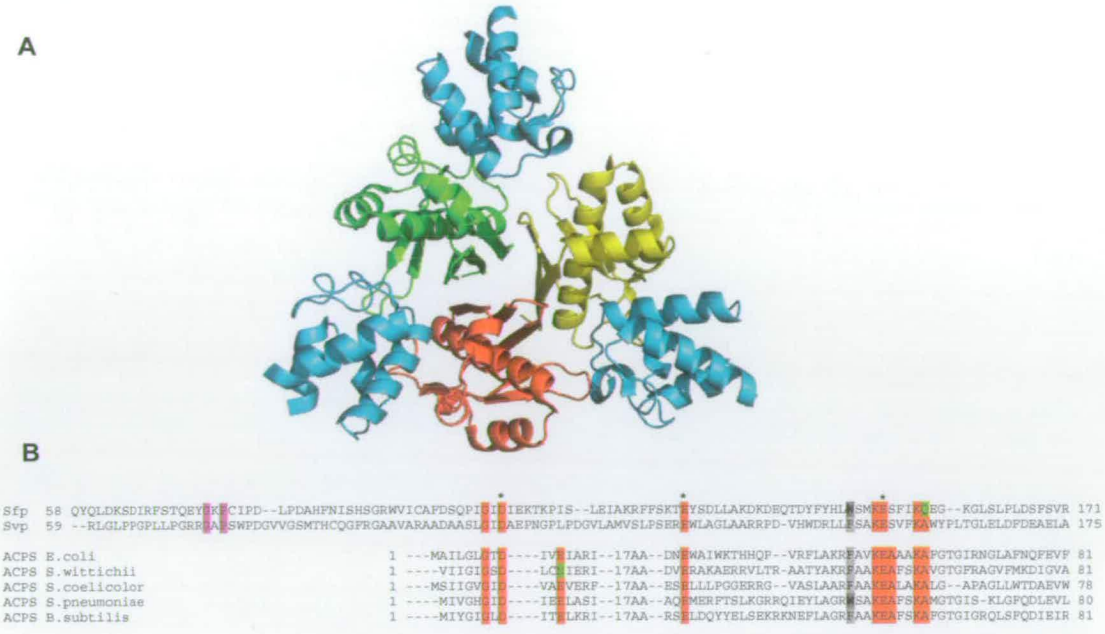


Figure I-25 (A) ACPS/ACP *B. subtilis* trimer overall structure (PDB: 1F80). The three monomers of the ACPS protein are represented in green, yellow and red. (B) Partial sequence alignment of 45 putative and proven phosphopantetheinyl transferases of the Sfp-type (top) and ACPS-type (bottom). Invariant residues are shaded in pink (Sfp-type) and red (both types), highly conserved residues are shade in light and dark grey. Residues occupying conserved positions but deviating from the consensus amino acid are shaded in green . The three acidic residues of Sfp involved in Mg^{2+} complexation are marked by an asterisk (*). The sequence alignment was performed using CLUSTALW (Higgins *et al.*, 1996) .Adapted from[104].

It has been shown that the conserved residues are involved in ACP/ACPS interface, coenzyme A and Mg^{2+} binding[106]. One particular residue, Lys 62 (*B. subtilis* ACPS numerotation), is believed to be involved in the mechanism by activation of the ACP serine residue for the nucleophile attack of the pyrophosphate moiety of coenzyme A (Figure I-26).

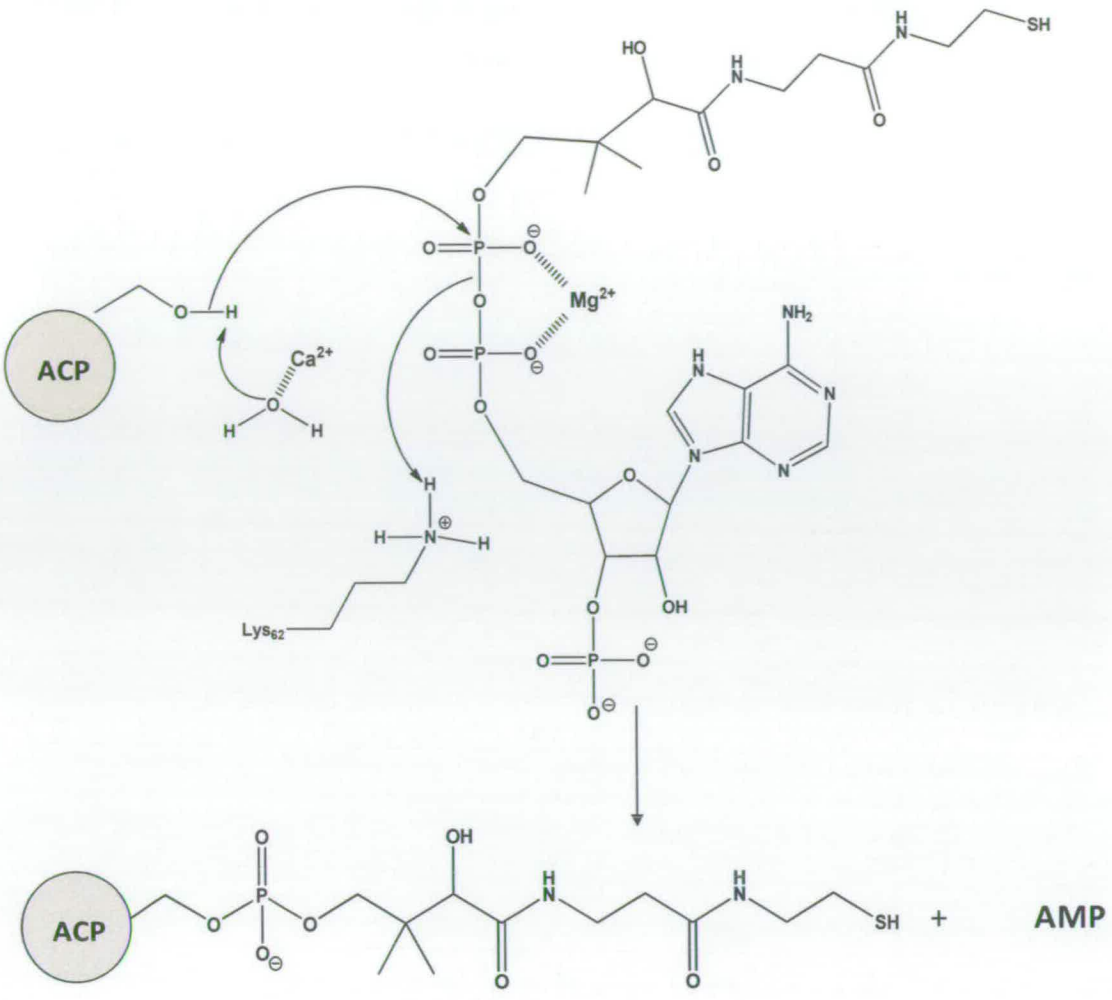


Figure I-26 Mechanism of holo-ACP formation.

It appears that each ACP has a corresponding PPT that has been optimized for activity. It has been shown that ACPSs from the type II PPTases family are able to modify 4'-PP-dependent enzymes from the same class but are not able to modify type II 4'-PP-dependent ACPs.[106] On the contrary, type I PPTases proteins are able to modify both type of carrier proteins. A logical explanation has been proposed by Parris *et al.* when they looked at the superposition of both protein active sites[106]. Type I PPTases such as Sfp have a shallow and wide active site which can easily accept ACPs from both large domain synthases and smaller, independent ACPs. It is understandable that the small, independent ACP can fit in an ACPS active site but a larger ACP domain cannot access it. It is interesting to note that PPTases are not only limited to the transfer of 4'-PP group. It has been shown in numerous studies that Sfp[49, 107] or other PPTases such as Svp[48, 108] and ACPS from *S. coelicolor*[109], are able

to directly transfer the acyl chain in addition to the 4'-PP group from an acyl-CoA (Figure I-27). This is useful since it allows the conversion of an apo-ACP into its acyl-ACP form in a single step rather than the two-step conversion where the holo-form is formed first then acylated specifically on the 4'-PP afterwards.

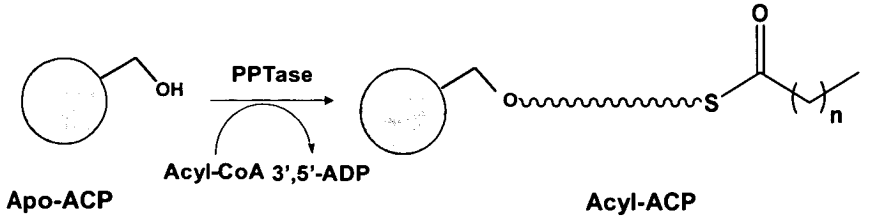


Figure I-27 Direct conversion of apo-ACP to acyl-ACP.

1.4.4 Aims

The overall goal of this work is to investigate the effect of SPT mutations that cause the human genetic disease HSAN1 in a bacterial mimic (SPT from *Sphingomonas paucimobilis*). The main aims are as follows:

- Characterisation of wild-type *SP* SPT and development of a DTNB activity assay for kinetic analysis.
- Mapping of the HSAN1 mutations onto the bacterial enzyme and subsequent preparation of the mutant SPT mimic clones by site-directed mutagenesis.
- Characterisation of HSAN1 mutants: activity, substrate binding, mass-spectrometry.
- Structural biology of wild-type and mutants, co-crystallisation with substrates.

The second part of the thesis investigates the possibility that a bacterial SPT from *Sphingomonas wittichii* utilises an acyl carrier protein (ACP) to deliver the acyl chain moiety to the enzyme instead of a palmitoyl-CoA substrate.

- Amplification of ACP and SPT genes from *SW* genomic DNA, followed by cloning into pET vectors.
- Characterisation of SPT: activity assays, mass-spectrometry, substrate binding, kinetic analysis.
- Modification of apo-ACP to holo- and acyl-ACP with several PPTase. Characterisation by mass spectrometry.

CHAPTER II: MATERIALS AND METHODS

II.1 MATERIALS AND REAGENTS

All reagents, chemicals and media were purchased from Sigma-Aldrich, Fisher or Biorad unless otherwise stated. S-(2-Oxoheptadecyl) CoA was gifted from Hiroko Ikushiro, Department of Biochemistry, Osaka Medical College. All primers were purchased from Sigma Genosys. All competent cells and pET plasmids were purchased from Novagen and all chromatography columns were from GE Healthcare unless otherwise indicated. The pET28a-SPT plasmid was constructed in the lab, prior to this work, by Dr. Beverley Yard. The plasmids encoding the *Streptomyces coelicolor* ACPS (pET15b-ACPS) and *Bacillus subtilis* Sfp (pET28a/Sfp) were generous gifts from Dr. Matt Crump (University of Bristol) and Prof. Jason Micklefield (University of Manchester) respectively. The synthetic DNA construct (encoding *Streptomyces verticillus* Svp) was purchased from Mr. Gene, Regensburg, Germany.

II.1.1Competent *E. coli* Cell Lines

Table II-1

Species	Cell line	Genotype	Application
E. coli	BL21 DE3 (Novagen)	F⁻ *ompT hsdSB* ($r_B^- m_B^-$) *gal dcm* (DE3)	Transformation, expression host.
E. coli	DH5α™ (Invitrogen)	F⁻ φ80d*lacZ*ΔM15 Δ(*lac*2YA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 ($r_K^- m_K^-$) *pho*A *sup*E44λ⁻ *thi*-1 *gyr*A96 *rel*A1	Transformation, plasmid storage.
E. coli	HMS 174 DE3 (Novagen)	F⁻ *rec*A1 *hsd*R(r_{K12}^- m_{K12}^+) (DE3) (Rif^R)	Transformation, expression host.
E. coli	One Shot® Top 10 (Invitrogen)	F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80d*lacZ*ΔM15 Δ*lac*x74 *rec*A1 araΔ139 Δ(*ara*-leu)7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G	Transformation of ligation, plasmid storage.
E. coli	JM109 (Promega)	*end*A1, *rec*A1, *gyr*A96, *thi*, *hsd*R17 (r_K^- m_K^+), *rel*A1, *sup*E44, Δ(*lac-pro*AB), [F' *tra*D36, *pro*AB, *laq*I^q ZΔM15].	Transformation of ligation, plasmid storage.

II.1.2 *Sphingomonas wittichii* Cell stocks

The *S. wittichii* RW1 strain was plated onto LB agar and grown for 2 days at 37 °C. A single colony was used to inoculate LB (5mL) under sterile conditions. The inoculant was grown overnight at 37 °C with agitation. Glycerol was added to a final concentration of 20%. The cell stock was stored at -80 °C.

II.1.3 Growth Media

All media were dissolved in distilled water and the pH adjusted to pH 7.5. All were autoclaved for 20 minutes at 120 °C, 15 psi before use unless otherwise stated and if required to be stored, were kept at 4 °C.

2YT: Bacto- tryptone (16 g/L), bacto-yeast extract (10 g/L), sodium chloride (5 g/L)

Luria Bertani (LB): Tryptone (10 g/L), yeast extract (5 g/L), sodium chloride (10 g/L)

SOC: Tryptone (20 g/L), yeast extract (5 g/L), sodium chloride (10 mM), potassium chloride (2.5 mM), magnesium chloride (10 mM), magnesium sulphate (10 mM), glucose (2 % w/v)

II.1.4 Agar

All agars were dissolved in distilled water and were autoclaved for 20 minutes at 120 °C, 15 psi. Antibiotic was added on cooling (approximately 50 °C) to the required concentration. Plates were allowed to dry before being stored at 4 °C.

LB: Tryptone (10 g/L), yeast extract (5 g/L), sodium chloride (5 g/L), Agar (15 g/L)

S-Gal/Agar: Tryptone (10 g/L), yeast extract (5 g/L), sodium chloride (10 g/L), Agar (12 g/L), S-Gal (0.300 g/L), Ferric ammonium citrate (0.500 g/L), IPTG (0.030 g/L)

II.1.5 Antibiotic Stock Solutions

All antibiotic stocks were sterile filtered (0.2 μm) and stored at 4 °C.

Ampicillin: A 100 mg/mL stock solution was made and the required volume added to media to give a final concentration of 100 μg/mL.

Kanamycin: A 30 mg/mL stock solution was made and the required volume added to media to give a final concentration of 30 μg/mL.

II.1.6 Solutions and Buffers

Each buffer was made by dissolving the components described below in sterile water. All buffers were 0.2 μm filtered and degassed prior to chromatography. All buffer stocks were stored at 4 °C.

Stock buffer (0.5 M Potassium Phosphate, KPhos pH 7.5) : Dibasic potassium phosphate, K_2HPO_4 (0.470 moles/L), monobasic potassium phosphate, KH_2PO_4 (0.0299 moles/L).

Buffer A: 20 mM KPhos pH 7.5, 25 μM PLP, 10 mM imidazole, 150 mM NaCl

Buffer B: 20 mM KPhos pH 7.5, 25 μM PLP, 50 mM, 150 mM NaCl

Buffer C: 20 mM KPhos pH 7.5, 25 μM PLP, 100 mM imidazole, 150 mM NaCl

Buffer D: 20 mM KPhos pH 7.5, 25 μM PLP, 150 mM NaCl

Buffer E: 20 mM KPhos pH 7.5, 150 mM NaCl

Buffer F: 20 mM KPhos pH 7.5, 0 M KCl

Buffer G: 20 mM KPhos pH 7.5, 1 M KCl

Buffer H: 20 mM KPhos pH 7.5, 150 mM NaCl. Addition of 10% glycerol in ACPS SW buffer

Buffer I: 20 mM KPhos pH 7.9, 0.5 mM NaCl, 5 mM imidazole. Addition of 10% glycerol in Sfp buffer

Buffer J: 20 mM KPhos pH 7.9, 0.5 mM NaCl, 500 mM imidazole. Addition of 10% glycerol in Sfp and ACPS SW buffers

Buffer K: 20 mM KPhos pH 7.5, 150 mM NaCl, 20 mM Imidazole. Addition of 10% glycerol in ACPS SW buffer

Buffer L: 50 mM Tris pH 7.8

Buffer M: 75 mM MES pH 6.0

Buffer N : 20mM KPhos pH 7.5

II.2 METHODS

II.2.1 DNA Manipulation

II.2.1.1 Cloning and Expression Plasmids

Table II-2

Plasmid	Cloning / Expression	Application
pGem®-T Easy (Promega)	Cloning in *E. coli*	Provides stock of DNA available for manipulation into various expression vector
pET-28a	Expression in *E. coli*	Expression of C-terminal His_6-tagged protein
pET-22b	Expression in *E. coli*	Expression of C-terminal His_6-tagged protein

II.2.1.2 Oligonucleotide Primers

Table II-3

Primer	Sequence (5'-3')
SPT Sp N100W for	GGGTCGGGCACCGGCAGCCGGATG
SPT Sp N100W rev	CATCCGGCTGCC[CCA]GGTGCCCGACCC
SPT Sp N100Y for	GGGTCGGGCACC[TAT]GGCAGCCGGATG
SPT Sp N100Y rev	CATCCGGCTGCC[ATA]GGTGCCCGACCC
SPT Sp N100C for	GGGTCGGGCACC[TGT]GGCAGCCGGATG
SPT Sp N100C rev	CATCCGGCTGCC[ACA]GGTGCCCGACCC
SPT Sp R378A for	CTCTACGTCAACATGGCG[GCC]CCGCCC
SPT Sp R378A rev	GGGCGG[GGC]CGGCATGTTGTCGTAGTG
SPT Sp R378N for	CTCTACGTCAACATGGCG[AAC]CCGCCC
SPT Sp R378N rev	GGGCGG [GTT] CGGCATGTTGTCGTAGTG
ACP-SPT Sw BspHI for	CATGTCATGAGCAGCAGAGAAGACATCTTCACC
ACP-SPT Sw HindIII rev	GGAAAAGCTTGGGGATTACTCCCGTGGCGCGCC
SPT Sw NcoI for	TGCGGCCGCAAGCTTGGGGATTAC
SPT Sw HindIII rev	AGCGGGGAACCCATATGATCATCGGCATC
ACP Sw HindIII rev	AAGCTTGGAGAGGAGGTCGGCCACC
ACPS Sw NdeI for	AGCGGGGAACCCATATGATCATCGGCATC
ACPS Sw XhoI rev	ATCCCCCTCGAGCAGCGGGCGCGCATACA
F2SPT BamHI for	AGCCATATGGCTGGATCCGAGGCTGCCGCTCAG
F2SPT NheI rev	GCAGCCTCTGTCATGCTAGCACCAATAACCCCG

II.2.1.3 Transformation of Competent *E. coli*

Single Transformation

Plasmid DNA (2 μL) was added to competent *E. coli* cells (25 μL) and incubated for 20 minutes on ice. The cells were then heat shocked for 30 seconds at 42 °C before being incubated on ice for a further 2 minutes. SOC medium (80 μL) was added and the cells were agitated (250 rpm) for 1 hour at 37 °C. Subsequently,

the cells (50 µL) were spread on a LB kanamycin plate (30 µg/mL) and incubated overnight at 37 °C.

Co-Transformation of *S. wittichii* ACP pET28a and ACPS pET22b

The two plasmids DNA (2 µL) were added to competent *E. coli* cells (25 µL) and incubated for 20 minutes on ice. The cells were then heat shocked for 30 seconds at 42 °C before being incubated on ice for a further 2 minutes. SOC medium (80 µL) was added and the cells were agitated (250 rpm) for 1 hour at 37 °C. Subsequently, the cells (50 µL) were spread on a LB Agar kanamycin (30 µg/mL) and ampicillin (100 µg/mL) plate and incubated overnight at 37 °C.

II.2.1.4 Polymerase Chain Reaction: QuickChange® Site-Directed Mutagenesis (Stratagene)

Generally, each reaction contained forward primer (125 ng), reverse primer (125 ng), dNTP (1 µL) reaction buffer (5 µL, 10x Buffer) double stranded template DNA (50 ng) and PfuTurbo DNA polymerase (1 µL, 2.5 U). The reaction was made up to 50 µL with sterile water. The general thermal cycling procedure consisted of the following steps; step 1, initial denaturation (95 °C, 30 seconds), steps 2-19, denaturation (95 °C, 30 seconds), annealing (55 °C, 1 minute) and elongation (68 °C, 1 minute/kbp). Following thermal cycling, the reaction mixture was cooled to 37 °C before adding the restriction endonuclease DpnI (1 µL, 10 U) to remove any dam methylated/parental DNA. This solution was mixed using a pipette before being incubated for 1 hour at 37 °C. After digestion with DpnI, the DNA mixture (2 µL) was used to transform TOP10 competent cells as mentioned in section II.2.1.

II.2.1.5 Polymerase Chain Reaction (PCR): Taq Polymerase Beads (GE Healthcare)

Generally, each PCR reaction contained; template DNA (50 ng), forward primer (125 ng), reverse primer (125 ng) and puRe Taq Ready-To-Go PCR beads (x2,

GE Healthcare). Each PCR bead was reconstituted to a final volume of 25 µL to yield the following concentration of reactants; dNTPs (200 µM in 10 mM Tris-HCl, pH 9 at room temperature), potassium chloride (50 mM) and magnesium chloride (1.5 mM). The final reaction volume (50 µL) was made with sterile water. A general PCR reaction consisted of 30 temperature cycles; denaturation (95 °C, 1 minute), annealing (55 °C, 1 minute) and elongation (72 °C, 2 minutes) followed by an additional extension period (72 °C, 10 minutes). The reaction was terminated by the addition of DNA loading dye (6x, Novagen) and the mixture was separated by electrophoresis. The purified PCR product could then be cloned using pGEM T-Easy vector (Promega) according to manufacturer's instructions. All cloned DNA was used to transform JM109 competent cells (50 µL) and cultures were grown on S-Gal agar plates containing ampicillin (100 µg/mL) thus providing blue/white screening for the identification of viable clone.

II.2.1.6 Ligation Into An Expression Vector

Ligation were carried out using the gene of interest (3 µL), the restricted plasmid (5 µL), T4 Ligase (1 µL, $1x10^6$ ligations /unit, New England Bio-labs), T4 ligase buffer (1 µL, 10x, New England Bio-labs). The ligation was incubated overnight at room temperature and subsequently, 2 µL was used to transform JM109 competent cells (50 µL) as described in section II.2.1.

II.2.2 DNA Purification

II.2.2.1 Mini-prep: QIAprep® Spin Miniprep Kit (QIAGEN)

This procedure was carried out according to the manufacturer's instructions, provided with the kit and is summarised below. An overnight culture (5 mL) was centrifuged at 3,500 rpm for 15 minutes (Denley BR 401). The supernatant was discarded and the pellet was resuspended in P1 buffer (250 µL) and transferred to a sterile eppendorf (1.5 mL). To this, P2 buffer (250 µL) was added and the eppendorf was inverted 4-6 times. Subsequently, N3 buffer (350 µL) was added and the mixture

inverted immediately (4-6 times). The resultant cloudy mixture was centrifuged (13,000 rpm, 10 minutes) after which, the supernatant was transferred to a QIAprep spin column and centrifuged (13,000 rpm, 1 minute). The flow through was discarded and the spin column washed with PE buffer (750 μL) and centrifuged (13,000 rpm, 1 minute). The column was centrifuged for a further minute to remove any residual wash buffer before the purified DNA was eluted with sterile water (100 μL) and stored at -20 °C until required.

II.2.2.2 Gel Extraction-QIAquick® Gel extraction Kit (QIAGEN)

DNA was excised from an agarose gel (1%) and purified according to the manufacturer's instructions, which are summarised below. Agarose containing the DNA was covered in sufficient QG buffer (300 μL) and incubated at 50 °C until the agarose had completely dissolved. This solution was transferred to a spin column and centrifuged (13,000 rpm, 1 minute). The supernatant was discarded and QG buffer (500 μL) was added to the column and centrifuged (13,000 rpm, 1 minute). The supernatant was discarded and PE buffer (750 μL) added to the column and centrifuged (13,000 rpm, 1 minute). The supernatant was discarded and the column centrifuged again to remove any residual buffer. The purified DNA was eluted with sterile water (10-15 μL) and stored at -20 °C until required.

II.2.3 DNA Analysis

II.2.3.1 Digestion by Restriction Endonuclease

Generally, a single restriction digest contained the following; template DNA (8 μL), restriction buffer (1 μL, 10x, New England Biolabs) and restriction endonuclease (1 μL, 10x, New England Biolabs). This mixture was incubated at 37 °C for 3 hours and the reaction was terminated by the addition of DNA loading dye (6x, Novagen). The digestion mixture was loaded onto an agarose gel (1%) and separated by electrophoresis.

II.2.3.2 Electrophoresis

Agarose (1g) was added to TAE buffer (100 mL, Bio-Rad) and boiled until the agarose had completely dissolved. On cooling, to approximately 50 °C, ethidium bromide was added to a final concentration of 0.5 μg/mL and the solution swirled to ensure complete mixing. The molten agarose solution was transferred to the casting mould and allowed to set at room temperature. DNA loading dye (6x, Novagen) was added to the DNA samples (2.5 μL to small sample of approximately 10 μL and 12.5 μL to larger samples of approximately 50 μL). The gel was immersed in TAE buffer and run at a constant voltage of 100 volts until good separation between bands was observed.

II.2.3.3 Sequencing PCR

To prepare a DNA sample for sequencing, the DNA source (3 μL) was added to sequencing buffer (2 μL, 5x, Applied Bioscience) containing Big Dye version 3.1 (2 μL) and sequencing primer (1 μL forward or reverse, 10 μM stock). Primers (Table II-4) were purchased from Sigma-Aldrich. Only one primer, forward or reverse was required for each sequencing reaction. The general sequencing program consisted of 25 cycles of the following steps; 95 °C for 30 seconds, 50 °C for 20 seconds and 60 °C for 4 minutes.

Table II-4 Sequencing primers

Primer	Plasmid	Sequence (5'- 3')
pET forward	pET	TTAATACGACTCACTATAGGG
pET reverse	pET	CTAGTTATTGCTCAGCGGT
pGem T7 forward	pGem	TAATACGACTCACTATAGGG
pGem T7 reverse	pGem	ATTTAGGTGACACTATAGAA

II.2.4 Protein Production and Isolation

II.2.4.1 Large Scale Over-Expression in *E. coli*

Single Expression

One colony was used to inoculate 2YT or LB broth (500ml) containing the appropriate antibiotic and was grown overnight at 37 °C with shacking (250 rpm). This inoculant was sub-cultured into 2YT or LB broth (4 litres containing antibiotic) and grown to an OD_{600} of 0.6, at 37 °C before being induced with IPTG (0.1 mM). Induction was carried out at 30 °C for 3 to 5 hours depending on the protein.

Co-Expression of *S. wittichii* ACP and ACPS

One colony was used to inoculate 2YT broth (500ml) containing kanamycin (30 μg/mL) and ampicillin (100 μg/mL) and was grown overnight at 37 °C with shacking (250 rpm). This inoculant was sub-cultured into 2YT broth (4 litres containing antibiotic) and grown to an OD_{600} of 0.6, at 37 °C before being induced with IPTG (0.1 mM, final concentration). Induction was carried out at 30 °C for 5h.

II.2.4.2 Cell Lysis by Sonication

All steps were performed on ice. The cell pellet was resuspended in resuspension buffer (40 mL, 4:1 v/w) containing 1 protease inhibitor cocktail tablet (Roche) and sonicated for 15 cycles (30 seconds on, 30 second off) using a soniprep 150. The cell debris was harvested by centrifugation (14,000 rpm, 30 minutes, 4 °C) and the supernatant cell free extract (CFE) retained for purification.

II.2.5 Protein Purification

II.2.5.1 Purification of *S. wittichii* and *S. paucimobilis* SPT wild-type and mutants

Ni-NTA Agarose

The CFE from sonication (approximately 40 mL) was incubated with pre-equilibrated Ni-NTA agarose (2 mL, 5-10 mg protein/ml beads) at 4 °C for 1 hour prior to packing the beads into an empty polyprep column (5 mL, QIAGEN) by gentle suction. The beads were packed and washed with 20 column volumes of Ni-NTA binding buffer (buffer A), under gentle suction, and the protein eluted with 10 column volumes of 50 mM and 100 mM imidazole Ni-NTA elution buffer (buffer B and C).

Size Exclusion Chromatography

A HiPrep™ 26/60 Sephacryl™ S-200 High Resolution (318 mL) size exclusion column was pre-equilibrated with 1 column volume of buffer D. Before loading, the protein sample was 0.45 μm filtered and concentrated to 3-5 % of the total column volume. It was eluted over 1 column volume and 10 mL fractions were collected. The approximate mass of purified protein was calculated using a calibration curve (Figure II-1) obtained from known molecular weight standards (GE Healthcare).

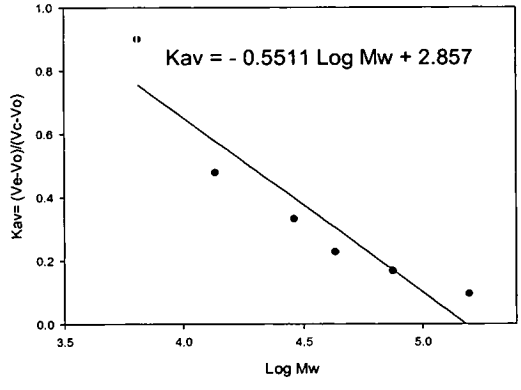


Figure II-1 Sephacryl™ S-200 molecular weight calibration curve.

II.2.5.2 Purification of *S. wittichii* ACP

Anion Exchange Chromatography

A HiLoad™ 26/10 Q Sepharose™ High Performance (53 mL) anion exchange column was washed with 2 column volumes of buffer F and recharged with 1 column volume of buffer G. The column was then equilibrated with 4 column volumes of buffer F. The cell free extract was diluted to 100mL with buffer F and 0.45 μm filtered prior being loaded onto the column. After loading, the column was washed with buffer F (4-6 column volumes, 4 mL/minute) and the protein eluted over 20 column volumes using a gradient of 0-1M KCl.

Size Exclusion Chromatography

A HiLoad™ 16/60 Superdex™ 75 prep grade (120 mL) size exclusion column was pre-equilibrated with 1 column volume of buffer H. Before loading, the protein sample was 0.45 μm filtered and concentrated to 3-5 % of the total column volume. It was eluted over 1 column volume and 5 mL fractions were collected. The approximate mass of purified protein was calculated using a calibration curve (Figure II-2) obtained from known molecular weight standards (GE Healthcare).

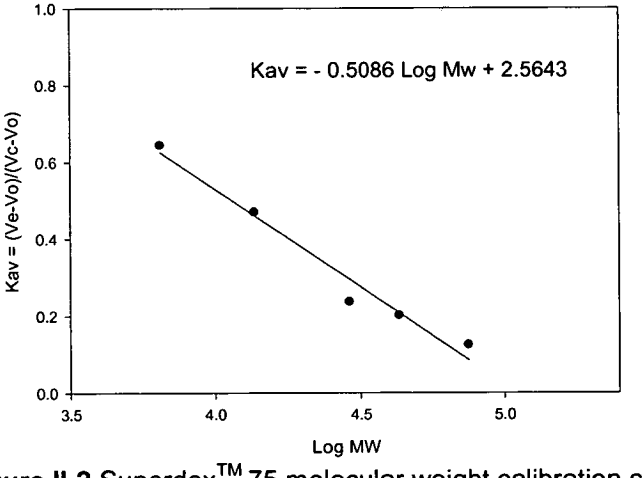


Figure II-2 Superdex™ 75 molecular weight calibration curve.

II.2.5.3 Purification of *S. coelicolor* ACPS and *B. subtilis* Sfp

His-trap purification

A HisTrap™ HP (5 mL) column was washed with 10 column volumes of buffer I and recharged with 5 column volumes of buffer J. The column was then equilibrated with 10 column volumes of buffer I. The cell free extract was filtered (0.45 μm filter) prior being loaded onto the column. After loading, the column was washed with buffer I (20-25 column volumes) and the protein eluted over 30 column volumes using a linear gradient of 5-500 mM imidazole. For the purification of Sfp, all buffers contained 10% glycerol.

II.2.5.4 Purification of *S. verticillus* Svp

His trap chelating purification

The protein in buffer K was loaded onto a HiTrap Chelating HP (1 mL) column was washed and pre-equilibrated as described above using buffer K and J. The filtered cell free extract filtered prior being loaded onto the column. After loading, the column was washed with buffer K (20-25 column volumes) and the protein eluted with a linear gradient [20 mM to 500 mM imidazole] over 10 column volumes.

Size Exclusion Chromatography

Fractions containing the protein of interest were pooled and loaded onto a previously calibrated HiLoad™ 16/60 Superdex™ 75 prep grade (120 mL) size exclusion column. Purification was carried out as described above for ACP gel extraction step.

II.2.6 Protein Analysis

II.2.6.1 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

An average gel (10 mL) consisted of a 15% running gel (3.5 mL H_20, 3.75 mL of 40 % acrylamide, 2.5 mL of 1.5 M Tris pH 8.8, 100 µL of 10 % w/v SDS, 200 µL of 50 mg/mL ammonium peroxodisulfate and 15 µL TEMED) and a 4% stacking gel (6 mL H_20, 1 mL of 40 % acrylamide, 2.5 mL of 0.5 M Tris pH 6.8, 100 µL of 10 % w/v SDS, 200 µL of 50 mg/mL ammonium peroxodisulfate and 15 µL TEMED). All protein samples were denatured by the addition of SDS_loading buffer (1.5 M Tris pH 8.8, 5 % v/v glycerol, 10 % w/v SDS, 0.4 mL β-mercaptoethanol and 0.05 % w/v bromophenol blue) and boiled for 10 minutes. The mixture was centrifuged (13,000 rpm, 5 minutes) prior to loading the gel. The gel was run for approximately 50 minutes at a constant voltage of 200 volts. Subsequently it was immersed in Coomassie stain (H_20, 0.1 % w/v Coomassie brilliant blue R250, 40 %v/v methanol and 10 % v/v acetic acid) and incubated for 20 minutes at 37 °C. The gel was destained (H_20, 40 % v/v methanol and 10 % v/v acetic acid) until all excess stain had been removed. An approximate molecular mass was obtained using pre-stained protein molecular weight markers (SeeBlue Plus 2, Invitrogen or Low Molecular Weight Marker, GE Healthcare).

II.2.6.2 BCA™ Protein Assay (PIERCE)

The BCA Protein Assay combines the reduction of Cu^{2+} to Cu^{+} by protein with the highly sensitive and selective colorimetric detection of the cuprous cation (Cu^{+}) by bicinchoninic acid (BCA).[110, 111] The first step is the chelation of copper with protein in an alkaline environment containing sodium potassium tartrate to form a blue colored complex (λ = 540 nm). The intensity of the color produced is proportional to the number of peptide bonds participating in the reaction. In the second step of the color development reaction, BCA reacts with the cuprous cation that was formed in step 1. The purple-colored reaction product is formed by the

chelation of two molecule of BCA with one cuprous cation. The BCA/copper complex exhibits a strong linear absorbance at 562 nm with increasing protein concentrations.

1mL of working reagent (50:1, Reagent A: Reagent B) was added to a clean, dry eppendorf tube, containing the protein sample (50uL) and incubated for 30 minutes at 37 °C. A standard curve was produced from serial dilutions (Table II-5) of a known stock concentration of bovine serum albumin (BSA, 2 mg/mL) and the concentration of protein calculated according to this standard curve.

BCA Reagent A contains sodium carbonate, sodium bicarbonate, bicinchoninic acid and sodium tartarate in 0.1 M sodium hydroxide

BCA Reagent B contains 4% cupric sulphate

Table II-5 Standard Curve Concentrations

STANDARD	CONCENTRATION (mg/ml)
A	2.00
B	1.50
C	1.00
D	0.75
E	0.5
F	0.25
G	0.125
H	0.025
I	0.00

II.2.6.3 Protein and Small Molecules Mass Spectrometry Analysis

Liquid Chromatography Electrospray Ionisation Mass Spectrometry (LC-ESI-MS) of proteins

Mass spectrometry was performed on a Micromass Platform II quadrupole mass equipped with an electrospray ion source. The spectrometer cone voltage was ramped from 40 to 70 volts and source temperature set to 140 °C. Protein samples were separated with a Waters HPLC 2690 with a Jupiter C5 reverse phase column (5 μm, 250 x 4.6 mm, Phenomenex) directly connected to the spectrophotometer. Proteins were eluted from the column with 5-95 % acetonitrile gradient using an acetonitrile/water/TFA (0.01 % v/v) solvents system (flow rate 0.4 ml/min). The total ion count in the range 500 to 2000 m/z was scanned at 0.1 s intervals. The scans were accumulated, the spectra combined and the molecular mass was determined by the MaxEnt and Transform algorithms of the Mass Lynx software (MicroMass).

Electrospray Ionisation Mass Spectrometry (ESI-MS) of Small Molecules

Direct infusion mass spectrometry was performed on a Micromass ZMD mass spectrometer equipped with an electrospray ion source. Samples were filtered (0.22 μm) and directly infused at a flow rate of 10 μL/min. Source parameters were kept constant except for the cone voltage and desolvation temperature. Cone voltage was typically set at 40-60 V and desolvation temperature set at 200 °C. All electrospray ionization was in positive mode and data acquired in continuum mode. Scans were accumulated, spectra combined and molecular mass calculated using the transform algorithm of the Mass Lynx software (Micromass).

Liquid Chromatography Fourier Transform Ion Cylotron Resonnance Mass Spectrometry (LC-FT-ICR)

This analysis was carried out by Dr. David J. Clarke (University of Edinburgh). For high resolution LC-FT-ICR MS, an Ultimate 3000 HPLC system (Dionex Corporation, Sunnyvale, CA), equipped with a monolithic PS-DVB (500

μM x 5 mm) analytical column (Dionex Corporation), was used in conjunction with Fourier Transform-Inductively coupled mass spectrometry (FT-ICR MS). Samples containing ~1 μg of ACP were centrifuged (16,100 *g* for 2 min) immediately prior to injection onto the column. Solutions A and B were prepared comprising of 2:97.95 and 80:19.95 actonitrile:water with 0.05 % formic acid respectively. Samples were injected onto the analytical column, washed with buffer A for 5 min, followed by a 20 min linear gradient elution (20 μl/min) into buffer B. MS data was acquired on a Bruker 12 Tesla Apex Qe FT-ICR (Bruker Daltonics, Billerica, MA) equipped with an electrospray ionization source. Desolvated ions were transmitted to a 6 cm Infinity cell® penning trap. Trapped ions were excited (frequency chirp 48-500 kHZ at 100 steps of 25 μs) and detected between *m/z* 600 and 2000 for 0.5 s to yield a broadband 512 K word time-domain data. Fast Fourier transforms and subsequent analyses were performed using DataAnalysis (Bruker Daltonics) software.

Isotopic Fitting

Isotope distributions of specific charge states were predicted using IsotopePattern software (Bruker Daltonics) from theoretical empirical formulae. These were overlaid upon the recorded experimental data as scatter plots, with the theoretical apex of each isotope peak designated by a circle.

Top-Down Fragmentation.

Top-down fragmentation was performed on the 12T Qh-FT-ICR. Two specific ions, m/z 1521 and 1304, were sequentially isolated using the mass resolving quadrupole, and MS/MS was performed using collision induced dissociation. For CID, the collision voltage was typically set between 20-35 V. Fragmentation data was the sum of 200 acquisitions and data analyses were performed using DataAnalysis (Bruker Daltonics). The SNAP 2.0 algorithm was used for automated peak picking. The resulting top-down fragment mass lists were combined and searched against the primary sequence of ACP using Prosight-PTM.[112] Mass error tolerances were set for all searches at 10 ppm.

II.2.7 Protein Chemistry

All UV-visible spectra were recorded on a Varian's Cary 50 UV-Vis spectrophotometer and analysed using Cary WinUV software (Varian). All assays were carried out on a 500 μL scale at 25 °C unless otherwise stated.

II.2.7.1 Spectroscopic Measurements

All UV-visible spectra were recorded on a Varian's Cary 50 UV-Vis spectrophotometer and analysed using Cary WinUV software (Varian). To convert the apo-SPT to the holo-SPT the enzyme was dialysed for 1h at 4 °C against 20 mM potassium phosphate (pH 7.5) containing 150 mM NaCl and 25 μM PLP. Excess PLP was removed by passing the protein through a PD-10 (Sephadex G-25M) desalting column (GE Healthcare) before any spectrophotometric measurements were taken. The concentration of recombinant SPT was 10 μM and the spectrophotometer was blanked with 20 mM potassium phosphate (pH 7.5) containing 150 mM NaCl.

II.2.7.2 Determination of dissociation constants

Assays were carried out in 0.5 mL quartz cuvettes (1 cm path length) and typically contained 10 μM SPT in 20 mM potassium phosphate (pH 7.5). Varying amounts of L-serine (0-80 mM) were used in each assay. After addition of substrate, the reactants were mixed and allowed to equilibrate for 15 mins at 25 °C. Small baseline changes were corrected using Sigma Plot software. Changes in absorbance at 425 nm (SPT WT, N100C, R378A and R378N) and 420 nm (SPT N100W and N100Y) and were plotted against L-serine concentrations, and data points were fitted to a hyperbolic saturation curve (equation 1) using Sigma Plot software:

$$\Delta A_{obs} = \frac{\Delta A_{max}\,[\text{serine}]}{K_d + [\text{serine}]} \qquad (1)$$

The ΔA_{obs} is the observed change in absorbance at 425 nm and the ΔA_{max} is the maximal absorbance change, [serine] is the L-serine concentration, and K_d is the dissociation constant.

II.2.7.3 SPT WT, N100Y and R378A quinonoid formation

Assays were carried out in 0.5 mL quartz cuvettes and typically contained 5 μM SPT in 20 mM potassium phosphate (pH 7.5). To this 45 mM of L-serine and 1.45 mM of S-(2-oxoheptadecyl)-CoA were added in the assay. After addition of the substrate, the reactants were mixed and allowed to equilibrate for 20 mins at 25 °C. The spectroscopic data were recorded between 200 and 800 nm. Small baseline changes were corrected using Sigma Plot software.

II.2.7.4 Detection of 3-ketodihydrosphingosine

The activity of SPT was measured by monitoring the formation of ^{14}C-KDS. A final enzyme concentration of 10 μM SPT in 20 mM potassium phosphate buffer, pH = 7.5, 150 mM NaCl) was incubated with 20 mM U-^{14}C L-Serine (9250 Bq, 0.250 μCi, GE Healthcare) and 1.6 mM palmitoyl-CoA in a final volume of 125 μL. The reaction was incubated at 37 °C for 20 minutes then the reaction was quenched by the addition of 30 μL of 2.0 M NH_4OH (final concentration 0.4 M). This was then extracted with an equal volume of $CHCl_3$:CH_3OH (2:1, v:v). The sample was centrifuged at 13,000 rpm for 5 minutes and the aqueous phase was discarded before the organic phase was allowed to evaporate overnight. The resulting lipid residue was re-suspended in 15 μL of $CHCl_3$:CH_3OH (2:1, v:v) and spotted onto a Silica Gel 60 TLC plate. Separation was carried out with a mobile phase of $CHCl_3$:CH_3OH:NH_4OH (40:10:1, v:v:v) using a KDS reference standard (Matreya). The TLC was developed with a Storage Phosphor Imager (GE Healthcare) for 4 days at room temperature and the Phosphor screen was visualised using ImageJ software.

II.2.7.5 DTNB assay

SPT activity was also measured using a continuous spectrophotometric assay by monitoring the release of CoASH from acyl-CoA substrates and reaction with 5,5'-dithiobis-2-nitrobenzoic acid (DTNB)[113]. Assays were carried out in 0.5 ml cuvettes in a Varian Cary-50 UV-Visible Spectrophotometer. The enzyme was incubated with L-serine in a buffered solution containing DTNB and the assay was started by addition of the second substrate palmitoyl-CoA. The CoASH thiol product was monitored by observation of the TNB^- anion at 412 nm (Δ_{max} = 14150 $M^{-1}cm^{-1}$).[114] Initial rates were measured at increasing concentrations of L-serine (0.12 – 50 mM) while maintaining palmitoyl-CoA in excess. A typical experiment to determine the K_m value for L-serine contained 0.16 µM SPT, 0.1 – 50 mM L-serine, 250 µM palmitoyl CoA and 0.2 mM DTNB in 50 mM potassium phosphate buffer pH 7.5. Kinetic constants (k_{cat} and K_m values) were calculated from Michaelis-Menten plots using Sigma Plot. Kinetic constant values were calculated in a similar way for CoA substrates (decanoyl, lauroyl, myristoyl, palmitoyl, stearoyl, arachidoyl) but by maintaining L-serine (100 mM) in excess. As a check, the k_{cat}/K_m parameters for CoA substrates were also determined by the method of complete condensation when [CoA]<<K_m. In this case k_{cat}/K_m is calculated at a single substrate concentration as an observed rate constant (k_{obs}) by monitoring disappearance of the substrate. The data was fitted to the equation $A = A_{lim}*(1 - \exp(-k_{obs}*t))$ where A is the absorbance at 412 nm, A_{lim} is the limiting absorbance value when the substrate has completely disappeared, k_{obs} is the observed rate constant and t is the time in seconds.

II.2.7.6 Conversion of Apo-ACP to Holo- or Acyl-ACP

ACP was converted into holo- or acyl- form by incubation with a specific PPTase (ACPS, Sfp, Svp) and free CoASH or acyl-CoA (C10, C12, C14 and C16 derivatives). For ease of reading, a common protocol is described using the name PPTase referring to ACPS *S. coelicolor*, Sfp, Svp or ACPS *S. wittichii.*

A final enzyme concentration of 60 µM ACP in reaction buffer was incubated with 1 µM PPT, 1 mM $MgCl_2$ and 250 µM to 1 mM CoASH/acyl-CoA in a final volume of

250 µL. The reaction was incubated at 37 °C overnight. The reaction mixture was then directly injected onto a LC/MS platform for analysis.

For each PPT, a different reaction buffer was used. Buffer L for ACPS *S. coelicolor* and Svp, buffer M for Sfp and buffer N for ACPS *S. wittichii*. 5 mM of DTT was added to the reaction mixture when use of free CoASH.

It is important to note that a concentration greater than 1 mM of $MgCl_2$ could not be use for reaction with acyl-CoA as this would lead to precipitation of these compounds and the reaction would not take place.

II.2.8 Structural Biology

All proteins were screened for suitable crystallisation conditions at the Scottish Structural Proteomics Facility (SSPF) in collaboration with Prof. Jim Naismith. The proteins were dialysed in the presence of excess PLP to ensure complete reloading prior to crystallisation. Data and structures have been deposited in the Protein Data Bank. They have the following codes.

2W8T: SPT WITH PLP, N100C
2W8U: SPT WITH PLP, N100Y
2W8V: SPT WITH PLP, N100W
2W8W: N100Y SPT WITH PLP-SER
2W8J: SPT WITH PLP-SER
2X8U: *SW* SPT WITH PLP

II.2.8.1 *S. paucimobilis* holoform SPTs N100Y, N100W and N100C

Crystallisation conditions that gave diffraction-quality crystals of the wild-type/PLP protein[39] were used as a starting point to formulate stochastic optimization screens using software developed in-house. The screens were built on a Hamilton Microstar liquid handling robot controlled by the Rhombix system software (Thermo). Crystallisation trials were set up in hanging drop plates (EasyXtal DG-CrystalSupport; Qiagen) using 2 – 3 µl of the protein solution and 1 µl of well

solution in the hanging drop. Crystals of the mutant SPTs were obtained in the form of yellow plates corresponding to the high-resolution form of the holo wild-type SPT. The SPT N100W mutant was crystallized at 20 mg/ml in 10 mM Tris (pH 7.5), 150 mM NaCl, 25 μM PLP using a well solution of 0.1 M $MgCl_2$, 22% (w/v) PEG 3350, and 0.10 M Hepes (pH 6.5). The N100Y mutant was crystallized at 20 mg/ml in 10 mM Tris (pH 7.5), 150 mM NaCl, 250 μM PLP using a well solution of 0.15M $MgCl_2$, 22% (w/v) PEG 3350, and 0.10 M Hepes (pH 6.5). Finally the N100C mutant was crystallized at 20 mg/ml in 10 mM Tris (pH 7.5), 150 mM NaCl, 250 μM PLP using a well solution of 0.1 M $MgCl_2$, 27% (w/v) PEG 3350, and 0.10 M Hepes (pH 6.5). The solutions used for equilibrating harvested crystals varied slightly according to the crystallization conditions and consisted of 22-29 % PEG3350 (w/v), 0.1 M HEPES (pH6.5), 0.08 – 0.12 M $MgCl_2$. All crystals used for data collection were soaked for a few minutes in the respective equilibrating solution plus 1-2 mM PLP. Analysis of the density revealed that the PLP was covalently bound to Lys265 in every mutant, confirming we had obtained the holo-form of the proteins. The data for the N100C, N100Y, N100W structures were processed using XDS and scaled using XSCALE (http://www.mpimf-heidelberg.mpg.de/~kabsch/xds/)[115]. All mutant models were refined using Refmac5[116] and manually adjusted including addition of water molecules with WinCoot[117]. Density was clearly visible to the N100Y mutation but we could not satisfactorily locate the tryptophan side chain, there is disordered electron density which indicates multiple conformations.

II.2.8.2 SPT WT and N100Y L-serine External Aldimines

The external aldimine form of the SPT wild-type was produced by soaking a crystal of the holo-SPT in the equilibrating solution plus 5mM L-serine for 10 minutes and that of the N100Y mutant by soaking the crystal in the equilibrating solution plus 50 mM L-serine for 15 mins at room temperature. The crystals were mounted in a cryo-loop (Molecular Dimensions) and cryo-protected in solutions varying slightly for each crystal and containing the equilibrating solutions plus 15-20 % PEG 400 (v/v). The crystals were then frozen by plunging them into liquid nitrogen and carried in a dry cryogenic Dewar to the European Synchrotron

Radiation Facility (ESRF, Grenoble, France) for data collection. The data sets were collected at 100 K to varying resolutions using 3 different beamlines. The data for the wild-type serine external aldimine were process with Mosflm and scaled with Scala from the CCP4 suite of program[118]. The data for the N100Y external aldimine structure were processed using XDS and scaled using XSCALE (http://www.mpimf-heidelberg.mpg.de/~kabsch/xds/)[115]. All models were refined using Refmac5[116] and manually adjusted including addition of water molecules with WinCoot.[117]

II.2.8.3 *S. wittichii* holoform SPT

The screens were built on a Hamilton Microstar liquid-handling robot controlled by Rhombix system software (Thermo). The *S. wittichii* SPT was crystallized at 20 mg/ml in 10 mM Tris (pH 7.5), 150 mM NaCl, 25 µM PLP using 30% PEG MME 2000, 0.1M KSCN, (NEXTAL JCSG+ screen condition 81). Crystals used for data collection were soaked for a few minutes in a solution of 27 % PEGMME2k, 0.22 M K-Thiocyanate, 18 % PEG 400 plus 1–2 mM PLP. The crystal was mounted in a cryo-loop (Molecular Dimensions) and then frozen by plunging it into liquid nitrogen and carried in a dry cryogenic Dewar to the European Synchrotron Radiation Facility (Grenoble, France) for data collection. Multiple crystals were screen at beam line BM14 .The data sets were collected at 100 K to 2.1

using beam line 14 (supplementary Table 1). The data for SPT *S. wittichii* were processed with Mosflm and scaled with XDS[115]. The density revealed that the PLP was covalently bound to Lys245 confirming that we had obtained the internal aldimine, holo-form of the protein. The SPT model was refined using Refmac5[116] and manually adjusted, including the addition of water molecules with WinCoot.[117].

CHAPTER III: RESULTS AND DISCUSSION

III.1 *Sphingomonas paucimobilis* - SPT WILD TYPE AND MUTANTS

Plasmid pET28a containing the gene for wild-type SPT was prepared previously in this lab. The expression and purification of the enzyme was first described by Yard *et al.*[39]

III.1.1 Wild Type SPT

III.1.1.1 Expression, Purification and Mass Spectrometry Analysis

The *S. paucimobilis* wild-type SPT gene cloned in plasmid pET28a (Figure III-1) was available at the start of this study. The pET28a plasmid contains a kanamycin resistance gene for selection. SPT WT gene originally cloned into *Nco*I and *Xho*I restriction site yields a protein with an 8 residue extension at the C-terminus, namely LEHHHHHH, containing the 6 histidine tag which enables purification of the protein via Immobilised Metal Affinity Chromatography (IMAC) on a nickel resin.

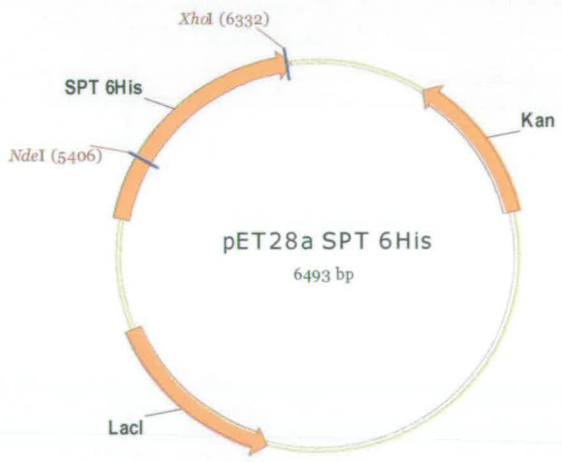


Figure III-1 pET28a/SPT WT plasmid map.

pET28a/SPT WT was used to transform HMS174 (DE3) competent cells and protein expression was induced with 0.1 mM of IPTG in 2YT broth at 30 °C for five hours SPT. The protein of interest was isolated from the cell free extract by nickel affinity chromatography. Analysis of the resulting fractions under reducing conditions by SDS-PAGE clearly shows a band at ~45 kDa due to SPT monomer

(Figure III-2A). The fractions containing SPT were further purified on a size exclusion Sephadex S-200 column (GE Healthcare). SPT eluted in a single peak (Figure III-2B) with a retention time corresponding to homodimeric SPT of ~92 kDa. From ~6 grams cell pellet this two-step purification process allowed the isolation of ~50 mg of pure and homogenous homodimeric SPT (size ~92 kDa) (Figure III-2).

Figure III-2 Expression and purification of wild-type SPT (A) SDS-PAGE Lane 1: Low molecular weight marker, Lane 2: Non-induced cells, Lane 3: Induced cells, Lane 4: 10 mM imidazole imidazole wash fraction (40 ml) from IMAC, Lane 5: 50 mM imidazole elution fraction (20 ml) from IMAC, Lane 6: 100 mM imidazole elution fraction (20 ml) from IMAC, Lanes 7-10: size exclusion chromatography, fractions 11-14. (B) Elution profile of size exclusion column (blue line A_{280} nm).

After purification, the recombinant wild-type SPT was analysed by Liquid Chromatography-Electrospray-Mass spectrometry (LC-ESI-MS). The theoretical mass of the apo-form of SPT is 46104 Da (the reversibly-bound PLP cofactor is lost during analysis by mass spectrometry). The purified protein mass was 45985 Da which corresponds to the SPT mass minus the N-terminal methionine which is cleaved post-translationally (Figure III-3A). In order to capture the holo-SPT form, sodium borohydride was used to reduce the reversible PLP internal aldimine to form enzyme with PLP irreversibly bound to the sidechain of Lys.[265] A mass of 46213 Da was observed which corresponds to an increase of 228 Da (protein-PLP+2H, Figure III-3B).

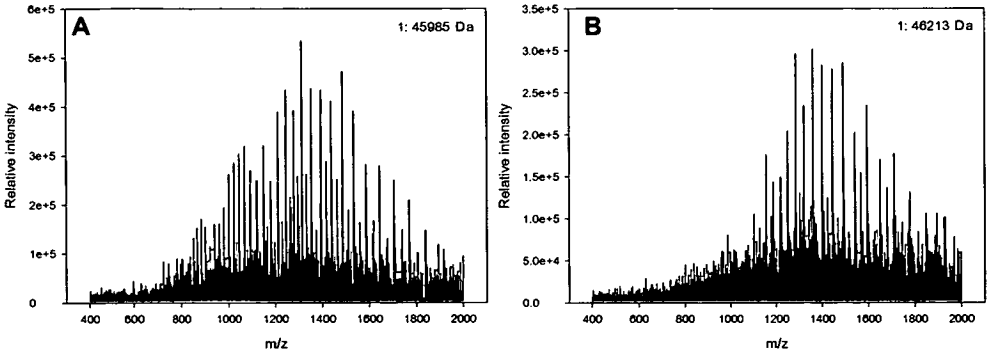


Figure III-3 LC-ESI-MS of wild-type SPT showing the ion envelopes of (A) Apo-form (B) the $NaBH_4$-reduced, holo form. Both masses were calculated by deconvolution.

III.1.1.2 Characterisation, Activity Assay and Substrate Selectivity

Although one of the two natural substrates for SPT is the amino acid L-serine, bacterial SPT has been shown to bind many L-serine analogues such as D-serine and L-serine methyl ester and other amino acids such as L-threonine.[20, 119] Other reactive L-serine derivatives such as L-cycloserine and β-chloro-L-alanine (Figure III-4) are potent inhibitiors of both bacterial and mammalian SPTs and are commonly used as inhibitors of *de novo* sphingolipid biosynthesis.[20, 119] The most potent and selective SPT inhibitor is the fungal natural product myriocin[20, 119] (Figure III-4) which is thought to mimic the putative β-keto acid intermediate during normal catalysis. The second of the two natural substrates for SPT is an acyl-CoA thioester. Ikushiro *et al.*[11] have shown that the bacterial *SP* SPT can utilise saturated and unsturated acyl-CoA substrates of varying carbon chain length (C12 to C20) but the best substrate is palmitoyl-CoA (C16). However this is not the case in all organisms, for example in a viral fusion SPT, the best acyl-CoA substrate is myristoyl-CoA (C14).[64] Here the binding affinity (UV-Vis spectrometry), the product formation (radiochemical assay) and the kinetics (DTNB assay) of several amino acids and acyl-CoAs were tested.

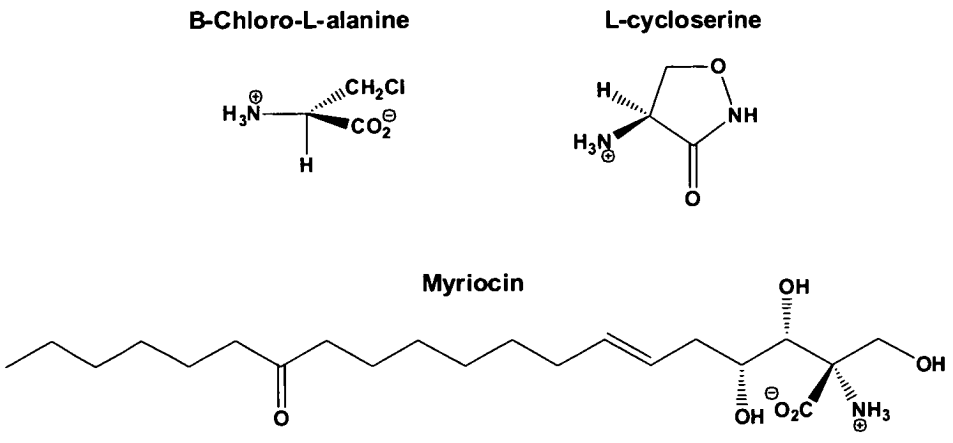


Figure III-4 Chemical structure of some SPT's inhibitors.

Spectroscopic Properties of SPT WT - Amino acids Selectivity

A solution of the vitamin B6 cofactor pyridoxal 5'-phosphate (PLP) is yellow in colour and has a characteristic UV-visible spectrum. In holo-SPT, the cofactor bound to Lys^{265} in the active site via an internal aldimine or Schiff's base. The internal alimine is an equilibrium between two species, each having a specific absorbance maximum (enolimine form λ= 335 nm and ketoenamine form λ= 422 nm, Figure III-4).

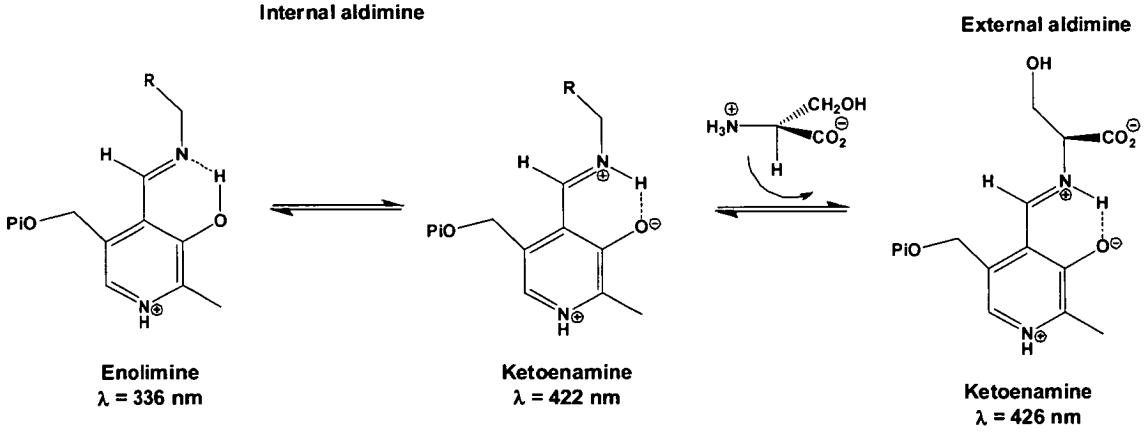


Figure III-5 Enolimine and ketoenamine equilibrium. Conversion of apo-SPT/internal aldimine to holo-SPT/external aldimine.

When L-serine binds to SPT an aldimine linkage or Schiff's base is formed between the cofactor and the amino group of the L-serine substrate (Figure III-5). This is

thought to occur via a geminal diamine intermediate[21]. In this way the PLP:Lys^{265} internal aldimine is displaced by a PLP:L-Ser external aldimine. Because the active form of the external aldimine is the ketoenamine species it is possible to follow binding of the L-serine substrate by monitoring an increase at 425 nm in the UV-vis spectrum (Figure III-6A). By measuring the change in absorbance at 425 nm, the apparent serine dissociation constant ($K_{d(\text{L-Ser})}$) has been determined (Figure III-6D, Table III-1). The wild-type *SP* SPT $K_{d(\text{L-Ser})}$ is 1.1 mM, in good agreement with the value reported by Ikushiro *et al.*[119] (1.4 mM).

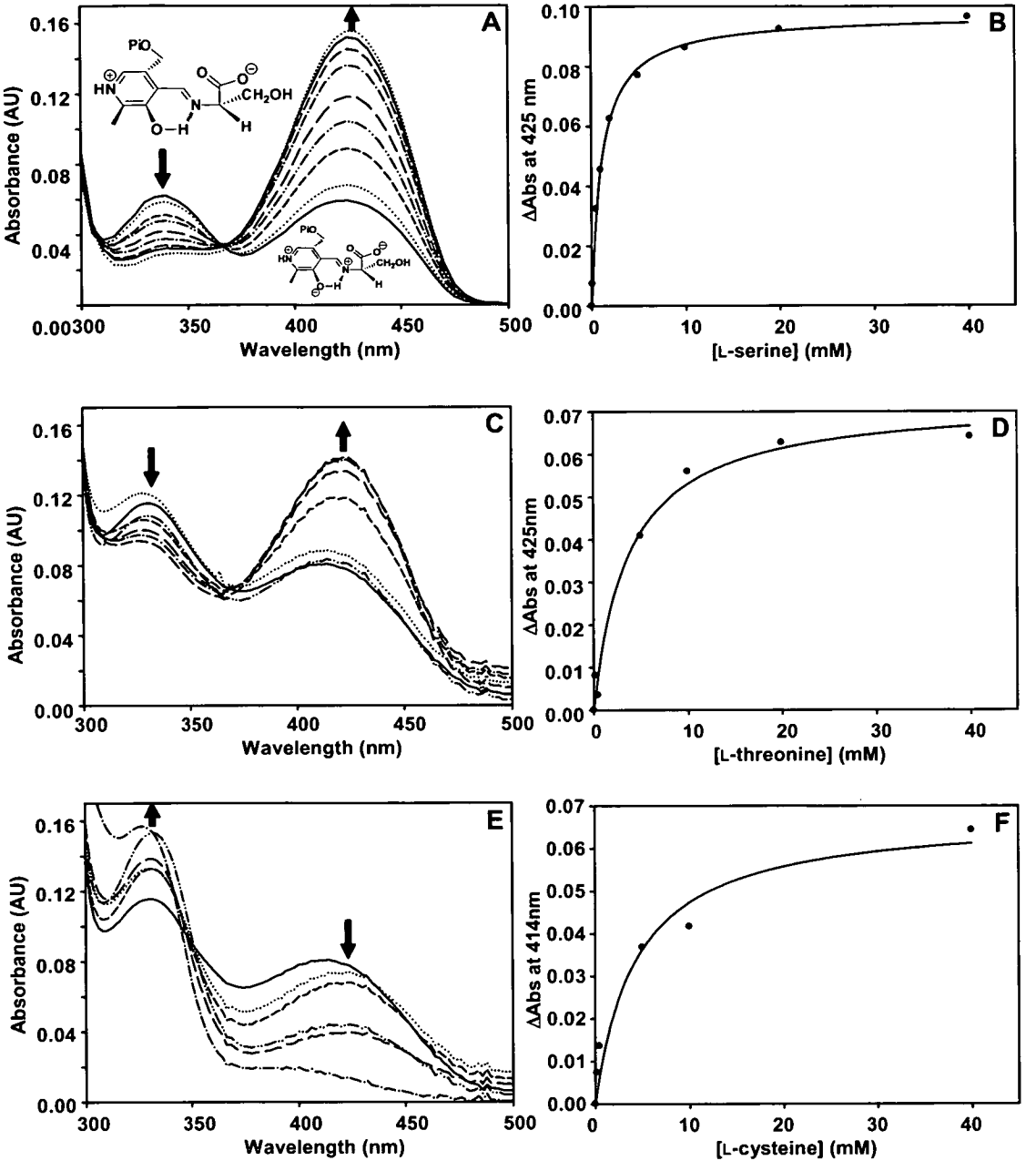


Figure III-6 UV-visible analysis of wild-type SPT. Absorbance spectra of *S. paucimobilis* SPT wild-type in presence of (A) L-serine (C) L-threonine and (E) L-cysteine. The structures of the enolimine (335 nm) and ketoenamine (420 nm) forms of the external aldimine are drawn in spectrum A. The *solid line* in each spectrum is the holo-form of the enzyme (10 µM SPT, 20mM potassium phosphate buffer (pH 7.5), 150mM NaCl, 25 °C). Increasing concentrations of L-amino acids were added (0, 0.1, 0.5, 1, 2, 5, 10, 20, and 40 mM; *dotted* and *dashed lines*), and the spectrum was recorded after 15 min. Dissociation constants were measured at 425 nm for (B) L-serine (D) L-threonine and (F) L-cysteine. *AU*, absorbance units.

Table III-1 *SP* wild-type SPT dissociation constants and kinetics table.

Enzyme	k_{cat} x 10^3 (s^{-1})	K_m (mM)	K_d (mM)	k_{cat} / K_m (mM)
L-serine	1150.0 ± 30.0	1.4 ± 0.1	1.1 ± 0.1	821.4
L-threonine	ND	ND	3.6 ± 0.9	ND
L-cysteine	ND	ND	5.0 ± 1.7	ND

The amino acid specificity of SPT can be explored by its ability to form external aldimines. By monitoring changes in the 425 nm absorbance maximum, due to formation of the ketoenamine external aldimine, binding of two amino acids, L-threonine and L-cysteine, was investigated. The amino acids were added to holo-SPT at increasing concentrations (0 and 40 mM) (Figure III-6C and 6E, respectively). L-threonine, which contains an extra methyl group on C_β, formed a stable external aldimine and like L-serine the absorbance maximum at 420 nm increased in the presence of increasing concentrations of amino acid (Figure III-6C). The calculated dissociation constant (K_d = 3.6 ± 0.9 mM, Figure III-6D) was comparable to that previously determined by Ikushiro *et al.*[119] (K_d = 3.8 mM). When L-cysteine was added to SPT the absorption maximum due to the external aldimine was at 412 nm and not 425 nm (Figure III-6E). In contrast to L-serine and L-threonine the 412 nm peak decreased upon addition of increasing concentrations of the amino acid and, interestingly, there was concomitant growth of a new peak at 330 nm. It is likely that the external aldimine formed with L-cysteine is not stable and reacts quickly to form another species. The "dissociation constant" of L-cysteine was evaluated by measuring the decrease in absorbance at 412 nm (K_d = 5 ± 1.7 mM, Figure III-6F,). Terzuoli *et al.* studied the condensation of aminothiols, in particular L- and D-cysteine with PLP, and showed that these reactions yield thiazolidine compounds.[120] This study suggests that the unknown compound observed at 330 nm could be a thiazolidine ring formed by nucleophilic attack of the external aldimine Schiff base by the thiol group of L-cysteine (Figure III-7).

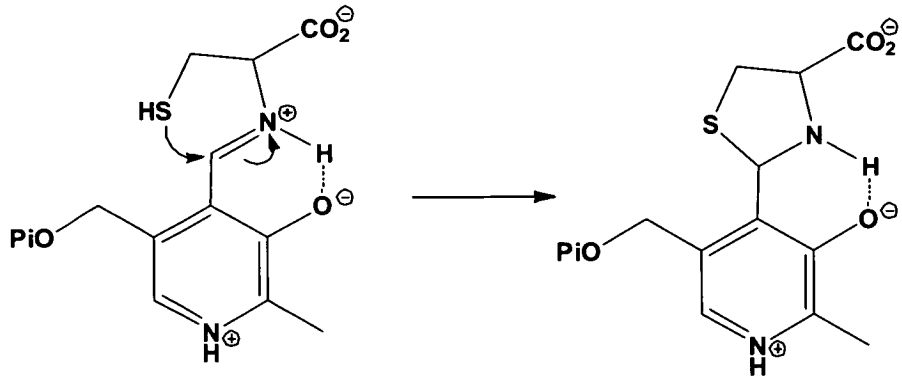


Figure III-7 Formation of thiazolidone.

3-Ketodihydrosphingosine (KDS) Formation

A radiochemical assay of SPT activity, based on the method previously reported by Ikushiro and colleagues[11], allows measurement of the 3-ketodihydrosphingosine (KDS) product formed. Recombinant holo-SPT was incubated with U-[^{14}C]-labelled L-serine and palmitoyl-CoA for 20 min at 37 °C. The assay was stopped by the addition of 30 µL of 2.0 M NH_4OH (final concentration 0.4 M) and the product extracted with an equal volume of chloroform/methanol (2:1, v/v). Organic phases were removed, dried, separated by thin-layer chromatography (TLC), and subsequently developed on an autoradiograph using a Storage Phosphor Imager. A commercial KDS standard was used as a marker to help identify the enzymatically-produced KDS (Figure III-8A). In this case KDS was identified in the organic phase after extraction of the SPT reaction mixture by direct injection into ESI-MS. The resulting mass spectrum showed a large peak at 300 Da which is consistent with a protonated KDS product (KDS+1H, Figure III-8B).

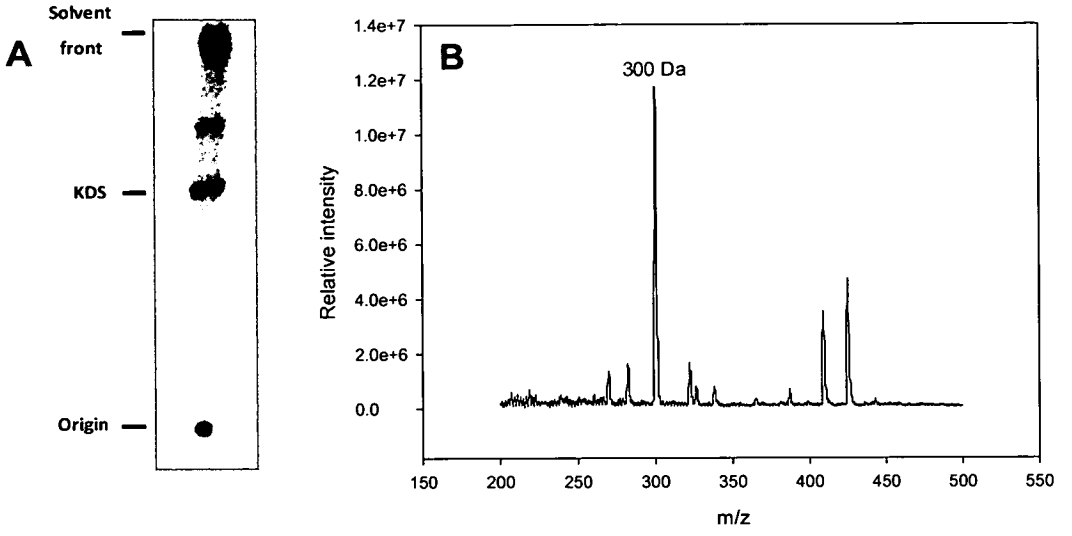


Figure III-8 Formation of 3-ketodihydrosphingosine (KDS) (A) TLC of radiolabeled products obtained by SPT-catalyzed conversion of of [^{14}C]L-serine L-serine and palmitoyl-CoA to [^{14}C]KDS. The black lines mark the positions of the origin, commercial standard KDS and the solvent front. (B) Mass spectrum of SPT-produced KDS.

Since L-threonine can form an external aldimine with SPT (Figure III-6C) samples were analysed for product formation when L-threonine was used as substrate instead of L-serine. Because [^{14}C]-L-threonine is not commercially available, it was not possible to utilise a labelled substrate and subsequently develop TLC plates by phosphor-imaging. Instead unlabelled L-threonine was used as substrate and ninhydrin, a reagent that forms a purple colour upon reaction with free amines, was used to develop TLC plates. Colour development was observed for both commercial and SPT-produced KDS but the corresponding ketohydrosphingosine product derived from an L-threonine substrate was not observed (data not shown). Since SPT can form an external aldimine with L-threonine, but cannot form product upon addition of the second substrate palmitoyl-CoA, this suggests that the reaction pathway may be blocked at one of the intermediates between the external aldimine and the final product. Intermediates in the catalytic cycle of several α-oxoamine synthases have been observed spectrophotometrically[33, 43, 46, 119, 121-123](Figure III-9). To investigate whether an intermediate accumulates when L-threonine is the substrate, the reaction was followed spectrophotometrically but no further changes occurred to the UV-visible spectrum (not shown) when palmitoyl CoA was added to

the SPT:L-Thr external aldimine (425 nm peak). It is possible that the SPT:L-Thr external aldimine cannot deprotonate to form the next quinonoid intermediate that reacts with the palmitoyl CoA thioester (Figure III-9) and therefore remains predominantly in the external aldimine form even in the presence of the second substrate.

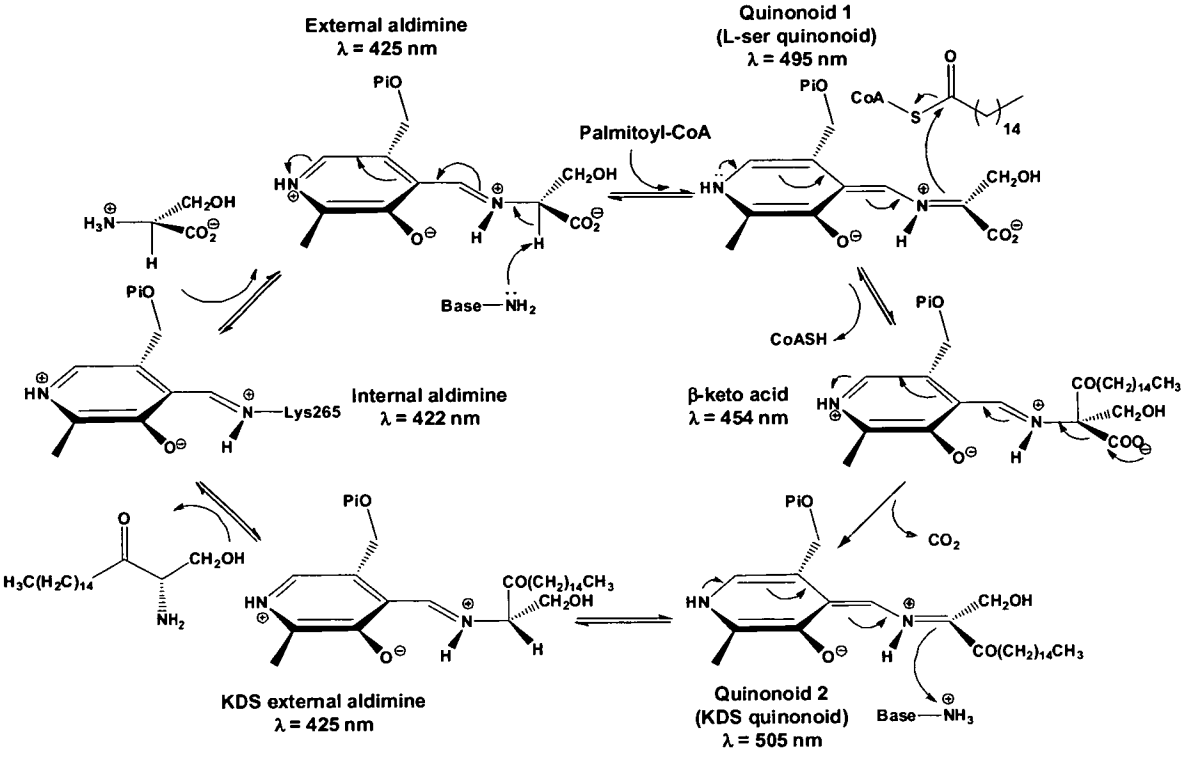


Figure III-9 SPT mechanism, highlight on UV-Vis absorbing intermediates.

SPT Activity – Kinetic Analysis and Acyl-CoA Selectivity

Although the autoradiography-based SPT assay is useful for checking SPT activity by identification of the KDS product, it is not a good method for determining catalytic constants. Quantitation of KDS is crude and error-prone since it relies on the measurement of the intensity of developed spots on the TLC plate. Also, the method is time-consuming as it can take up to several days to develop the TLC plate. To overcome these disadvantages a continuous and stoichiometric spectrophotometric SPT activity assay was developed whereby the release of free CoASH from acyl-CoA substrates was monitored by reaction with 5,5'-dithiobis-2-nitrobenzoic acid (DTNB)[113]. Free thiols react with DTNB to form a mixed

disulphide and a 2-nitro-5-thiobenzoate (NTB^-) anion that has an extinction coefficient of 14150 M^{-1} cm^{-1} at 412 nm[114]. Initial rates of CoASH release were measured spectrophotometrically at 412 nm and at increasing concentrations of L-serine in presence of recombinant SPT, palmitoyl-CoA, and DTNB. Kinetics values for SPT were determined using this method providing the substrates affinities ($K_{m(L\text{-}ser)}$ 1.4 mM and $K_{m(PCoA)}$ 35.4 µM, Table III-1 and 2), as well as turnover and catalytic efficiency of the enzyme (k_{cat} = 1.150 ± 0.030 s^{-1} and $k_{cat}/ K_{m(L\text{-}ser)}$ = 821.4 M^{-1} s^{-1}, Table III-1). In the same way, the SPT affinity for diverse acyl-chains was investigated. As expected, palmitoyl-CoA (C16:0) had the fastest turnover (k_{cat} = 1.150 s^{-1}) of all of the acyl-thioesters (Table III-2). Although, when looking at the efficiency of the substrates, the longer chain stearoyl-CoA (C18:0) was the most efficient substrate (k_{cat}/ K_m 65,547 M^{-1} s^{-1}). Accurate rates could not be determined for longer substrates such as arachidoyl-CoA (C20:0) due to the limit of detection with the DTNB reagent. For example, release of 1 µM CoASH will result in an absorbance change at 412 nm of only 0.014 units. It was clear enough that substrates with a chain length shorter than 16 carbons were poor substrate in terms of binding and turnover. A similar trend was observed by Ikushiro *et al.*[11] in a previous specificity study of *SP* SPT.

Table III-2 Acyl-CoA chain length specificity of *SP* SPT.

Enzyme	k_{cat} x 10^3 (s^{-1})	K_m (mM)	k_{cat} / K_m (mM)
Decanoyl (10:0)	0.045 ± 0.002	2324.9 ± 175.1	19.4
Lauroyl (12:0)	0.262 ± 0.006	822.2 ± 32.6	318.7
Myristoyl (14:0)	0.601 ± 0.012	97.1 ± 5.1	6189.5
Palmitoyl (16:0)	1.150 ± 0.030	35.4 ± 2.0	32485.9
Stearoyl (18:0)	0.898 ± 0.012	13.7 ± 0.8	65547.4
Arachidoyl (20:0)	0.327 ± 0.014	<10	>32700

III.1.1.3 Structural biology

Structure of the SPT:PLP:L-Ser external aldimine complex

The optimisation of growth of diffraction-quality crystals, the building of structural models and the solution of x-ray structures of wild-type *SP* SPT:L-Ser external aldimine complexes were the result of collaborative work between myself, Berverley Yard, Ken Johnson and Lester Carter.

The L-Ser external aldimine SPT complex (1.50Å) was trapped by soaking SPT crystals with 5 mM of L-serine (Figure III-10A). In the external aldimine form, the bond between Lys265 and the co-factor has been broken (Figure III-11) and a new Schiff base between L-serine and PLP is clearly observed (Figure III-10B). The structure of holo-SPT shows that the PLP binding site lies at the interface between the two SPT monomers and is composed of residues from both subunits. An overlay of the structures of the internal aldimine (holo-form) and SPT:L-Ser external aldimine complex reveals that the overall shape of the enzyme (not shown) and the relative orientation of the two monomers does not change upon formation of this first intermediate (r.m.s. deviation of 0.3 Å with the internal aldimine structure).

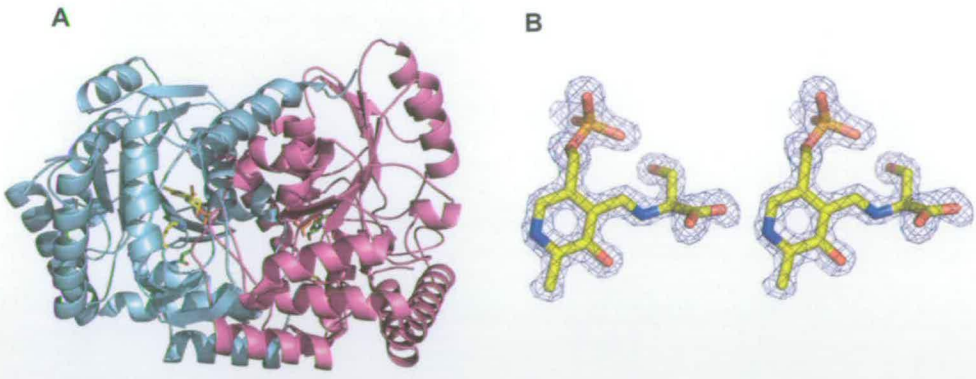


Figure III-10 Structure of the SPT:L-Ser external aldimine complex. (A) Structure of the entire SPT dimer showing one monomer in *teal* and the other monomer in *purple*. The residues Lys^{265} and Asn^{100} from the *teal* monomer are shown in *stick form* and are *coloured yellow*. The equivalent residues from the *purple monomer* are *coloured green*. This clearly shows that the Asn^{100} from one monomer interacts with the active site of the other monomer. The PLP:L-Ser external aldimine is also shown in *stick form*. (B) A *stereo image* of an *Fo _ Fc* electron density map contoured at 3 σ around the PLP:L-Ser external aldimine bound at the active site of SPT. The phases were calculated from a model that did not include included the PLP or aldimine.

The main interactions between the enzyme and the PLP cofactor that were observed previously in the holo-SPT internal aldimine are conserved in the SPT:L-Ser external aldimine complex. In both structures the PLP cofactor makes polar contacts with the side-chains of Asn^{138}, Asp^{231}, His^{234} and Thr^{262}, with the main chain of Gly^{134}, Tyr^{135} and π-stacks with His^{159}. In the wild-type SPT external aldimine complex, others interactions appear to help stabilise the L-serine moiety. The L-serine hydroxyl group is within H-bonding distance of the amino side-chain of Lys^{265} (2.78 Å). The carboxylate is involved in two interactions; with the Nε atom of His^{159} and an electrostatic interaction with guanidinium group of Arg^{378}. The Arg^{378} residue has undergone a large and striking conformational change (Figure III-10A). In the holo-form this residue is situated outwith of the catalytic site where it interacts with Gln^{357} (at a distance of 2.42 Å) which is on helix 13 (Figure III-10B). This Arg^{378} residue undergoes a large swing to bring it into contact with the L-serine carboxylate, an overall movement of ~14 Å (measured between the guanidinium -Nω- of the Arg residue and the oxygen atom of the PLP:L-Ser carboxylate,). Immediately following Arg^{378} are the residues PPATP (numbered 379-383). These residues form a

conserved loop (the "PPATP" loop) which has significantly altered its conformation and position in the external aldimine structure compared to the internal aldimine structure (*e.g.* the C_α atom of Pro^{379} has moved over 10 Å). The reordering of the structure around Arg^{378} seems to be linked to other movements at Phe^{47} (1.2-Å C_α shift). These residues are contained within the same monomer. The phosphate of PLP binds to the side chain of Thr^{294} from one subunit and to the main chain of Ala^{294} from the other subunit. The hydroxyl of L-Ser makes a polar contact to the main chain of Ala^{294}. In the other monomer, the side chain Met^{104} adjusts its conformation, and there is a slight rigid body shift of Leu^{105} to Asn^{106}. The side chain of Asn^{100} from the other subunit has altered its position, but the hydrogen bond to Lys^{265} is maintained.

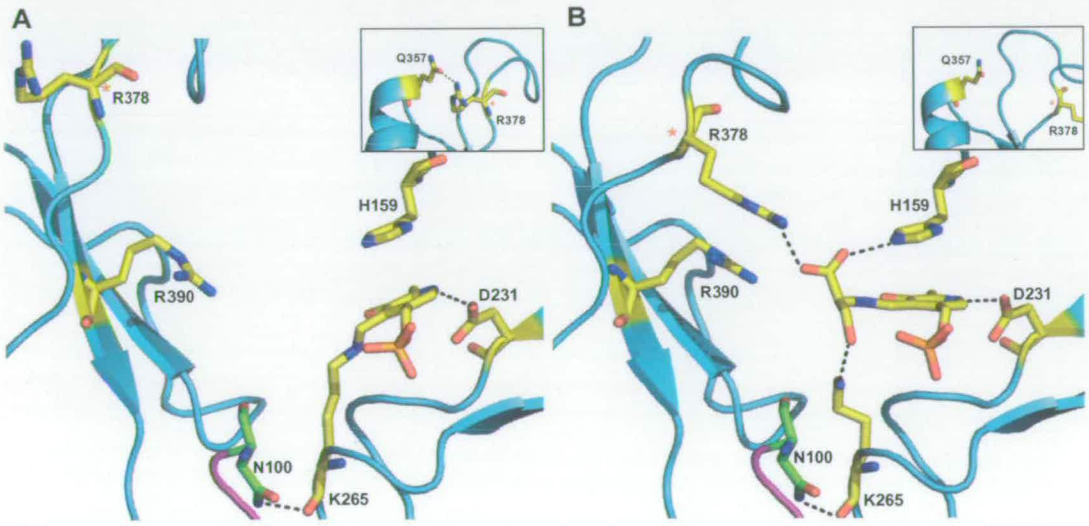


Figure III-10 Comparison of (A) internal and (B) external aldimine forms of wild-type SPT *S. paucimobilis*. In both images, one monomer is coloured in *teal* and the other in *purple*. All residues from the *teal* monomer are shown in *stick form* and are *coloured yellow*. The residue Asn^{100} from the *purple monomer* is *coloured green*. Black dashed lines represent the interactions existing between the internal/external aldimines complexes with active site residues. In A and B, the top right corner box shows the different locations of Gln^{357} and Arg^{378} when the external aldimine forms.

Shortly after the publication of the *SP* SPT external aldimine structure by Raman *et al.*[44](see appendices section), Ikushiro *et al.*[124] published another SPT:L-Ser external aldimine structure of the protein from the bacterium *Sphingobacterium multivorum (SM)*. While the overall protein structures are fairly similar to each other, sequence

alignment reveals some interesting differences. For example, Arg^{378} of *SP* SPT is not conserved in *SM* SPT indicating that there are differences in how the PLP:L-Ser complex is bound by each enzyme. Indeed, analysis of the *SM* SPT external aldimine structure reveals that the carboxylate of the bound L-serine interacts only with His^{138} (His^{159} in *SP* SPT numbering) and further stabilisation of the carboxylate is assured by H-bonding with two water molecules. Another obvious difference between the two SPTs is the UV-visible spectrum of *SM* SPT lacking the enolimine form (λ=336 nm) of the internal and external aldimines.

Substrate selection

The structure of the SPT:L-Ser external aldimine was used to investigate why an SPT:L-Thr complex could not proceed past the external aldimine intermediate (see Section III.1.1.2). There are no residues in the active site of the SPT:L-Ser external aldimine that indicate an L-threonine substrate cannot bind, and this is keeping with the observed SPT:L-Thr external aldimine by UV-visible spectroscopy (Figure III-6C). To form the putative quinonoid intermediate there must first be rotation around C_α to bring the proton into the correct orientation for deprotonation by Lys.[265] A model of the putative SPT:L-Ser quinonoid intermediate was made by mimicking the C_α rotation on the structure of the external aldimine. In this model it was found that Ser^{102} (Figure III-12A) could spatially restrain C_α rotation of other amino acid external aldimines (Figure III-12B). Rotation around C_α in the L-serine external aldimine reduced the distance betweenC_β of the L-serine substrate and C_β of the Ser^{102} sidechain reduced from 4.78 Å to 3.68 Å (Figure III-12A and B). While this rotation is tolerated in a SPT:L-Ser complex it is likely that the presence of a bulky methyl group in a SPT:L-Thr external aldimine would cause a steric clash with SPT residues. Future experiments are required to investigate whether a Ser^{102} mutant of bacterial SPT could turnover an L-threonine substrate. This "selective" function appears to be conserved throughout the AOS members. Sequence alignment shows that Ser^{102} is conserved in AONS, while a threonine and a valine are present in ALAS and KBL respectively. It is possible that ALAS and KBL, which both utilise glycine as a substrate, require a larger side-chain to prevent condensation of other

amino acids. In the bacterial SPT, Ser^{102} of monomer A controls the selectivity in the active site of monomer B, affirming that the active site is comprised of residues from both subunits. Sequence alignment of bacterial SPT with LCB1 of human SPT reveals a proline residue (Pro^{135}) at the equivalent position in the LCB1 subunit. It is possible that the bulky proline residue could carry out the same function as bacterial Ser^{102} by causing a steric 'block' on amino acids other than the L-serine substrate. However the cyclic structure of proline means that it is often found in distinctive secondary structural elements such as 'turns'. It is possible that Pro^{135} could have an effect on substrate specificity in human SPT. Recently a study on substrate specificity of human SPT in the presence of the activity-enhancing small subunits (ssSPTa and ssSPTb) revealed that, in contrast to bacterial SPT, L-alanine is a good substrate for wild-type human SPT.[125] Future experiments are required to investigate whether Pro^{135} at this position is the cause of a relaxed substrate specificity in human SPT. There is increasing evidence that the HSAN1 disorder is caused by a 'gain of function' in SPT resulting from a change in substrate specificity. Since mutation of Cys^{133} is the most prevalent mutation found in HSAN1 patients, it is interesting to note that Pro^{135} is situated so close to this residue.

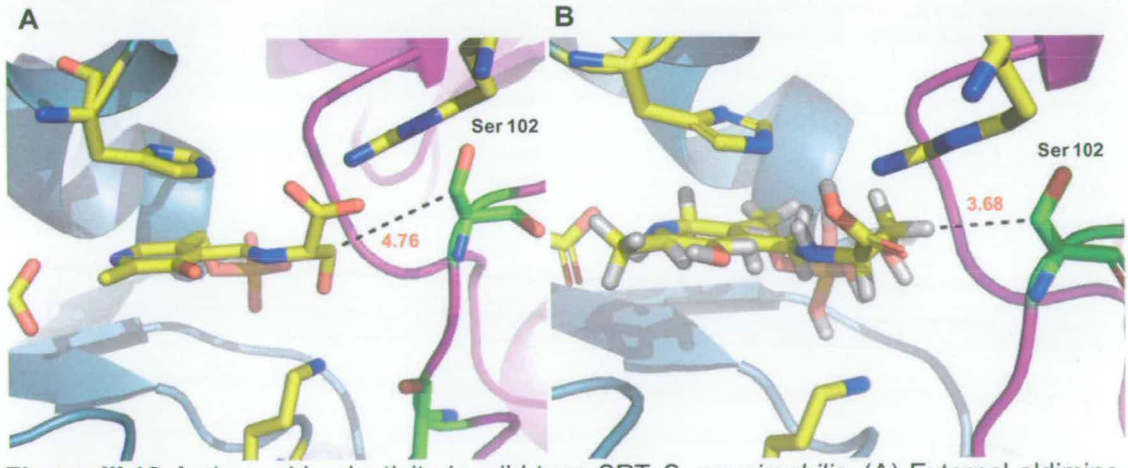


Figure III-12 Amino acid selectivity in wild-type SPT *S. paucimobilis*. (A) External aldimine. (B) Model of the putative SPT:L-Ser quinonoid intermediate. In both images, one monomer is coloured in *teal* and the other in *purple*. All residues from the *teal* monomer are shown in *stick form* and are *coloured yellow*. Residues from the *purple monomer* are *coloured green*. Black dashed lines represent the distance between C_β of the L-Serine substrate and the C_β of Ser^{102}.

III.1.2 Hereditary Sensory and Autonomic Neuropathy Type I SPT mutants

Genetic studies have shown that the human neurological disease HSAN1 is caused by several single point mutations on the LCB1 gene of SPT[69, 72; 126] (see Chapter 1). The most common mutation results in replacement of Cys^{133} by a tryptophan (Trp) or a tyrosine (Tyr). These two mutations have been previously studied and contradictory results were observed by studies in different laboratories. While Dawkins *et al.*[69] reported that the mutations lead to an increase in SPT activity, several other groups reported the opposite effect[70, 75, 127] and it is generally accepted that HSAN1 mutations have a dominant negative effect on SPT activity *in vivo*. Interestingly the total sphingolipid composition in HSAN1 patients carrying these mutations did not change[127]. Yard *et al.*[39] used the sequence of human LCB1 and LCB2 along with the solved structure of bacterial SPT to generate a heterodimeric SPT1/SPT2 model of the human enzyme. They showed that Cys^{133} on human SPT1 mapped onto Asn^{100} of the bacterial SPT. In the current study two bacterial SPT mutants N100W and N100Y were prepared. These are both "bacterial mimic models" of the C133W and C133Y human mutations that occur in HSAN1 patients. It was hoped that a spectroscopic, kinetic and structural analysis of these bacterial mimics would shed light on the effect of HSAN1 mutations in human SPT. A third N100C mutant was constructed as a control, since cysteine occurs at this position in wild-type human SPT.

III.1.2.1Cloning of pET-28a/SPT N100W, N100Y and N100C plasmids

The three HSAN1 mutations were introduced into SPT gene by site-directed mutagenesis (Stratagene) using the plasmid pET-28a/SPT WT as template.
After transformation in TOP10 cells and purification of DNA samples, each mutant plasmid was analysed by enzymatic restriction (Figure III-13) and Big Dye sequencing.

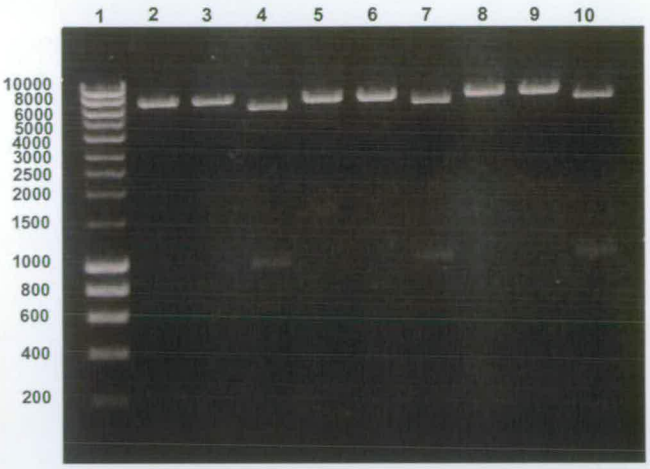


Figure III-13 Restriction enzyme analysis of pET28a/N100W, pET28a/N100Y and pET28a/N100C. Lane 1: Hyperladder I, Lanes 2-4: pET28a/N100W, Lanes 5-7: pET28a/N100Y, Lanes 8-10: pET28a/N100C. Lanes 2, 5 and 8: plasmid digested with *Nde*I, Lanes 3, 6 and 9: plasmid digested with *Xho*I, Lanes 4, 7 and 10: plasmid digested with *Nde*I and *Xho*I.

III.1.2.2 Expression, Purification and Mass Spectrometry Analysis

Several induction conditions were tested for each mutant and different parameters such as the cell type, temperature, time and concentration of IPTG inducer were varied. For all mutants the optimised conditions for expression of large soluble quantities of protein were found to be: transformation into BL21 (DE3) and induction at 30 °C, for 5h with 0.1 mM of IPTG. Because SPT N100W, N100Y and N100C were all cloned in the pET28a vector the resulting protein of each construct contained a LEHHHHHH C-terminus extension. Each SPT mutant was purified following the same procedure as wild-type SPT. All four proteins were indistinguishable when analysed under reducing conditions by SDS-PAGE giving rise to a band at ~45 kDa and identical elution profiles were observed for each protein when purified by gel filtration on a Sephacryl S-200 column (data not shown). Protein identity and integrity were confirmed by LC-ESI-MS. The experimentally determined masses of recombinant *SP* SPT mutants were all within 0.1% of the theoretical masses, which includes the C-terminal LEHHHHHH fusion affinity tag with the enzyme lacking the N-terminal methionine.

III.1.2.3 Characterisation and Activity of HSAN1 Mutants

Spectroscopic Properties of SPT mutants

In the reported structure of bacterial SPT Yard *et al.*[39] showed that residue N100 lies close to the catalytic site (Figure I-14), where it forms a contact with Lys^{265} on the opposite monomer. Because the sidechain of Lys^{265} forms a Schiff's base with the PLP cofactor it is possible that a mutation at position N100 may influence binding of cofactor and/or substrates. Here, binding of both PLP and L-serine were investigated for each of the mutants. The UV-visible spectra of the SPT N100W and N100Y mutants have maximum absorbance values at 340 and 415 nm (Figure II-14, A and B) confirming their ability to bind PLP. However both peaks have shifted and the ketoenamine become the dominant form in the spectrum of each of these enzymes. These differences suggest a common perturbation in the PLP binding site of the mutants compared with the wild type (Figure III-14, A and B, *solid lines*). The SPT N100C mutant has absorbance maxima at 335 and 425 nm which are comparable to the wild-type SPT values of 335 and 425 nm (Figure III-14, C and D). In both native and N100C, the enolimine peak at 335 nm is dominant, indicating that the PLP cofactor is in a very similar chemical environment in both proteins. To investigate the ability of the mutant SPT's to bind L-serine and form a stabilised external aldimine, L-serine was added to each enzyme at increasing concentrations. This resulted in an increase of the absorbance at 425nm for the N100C mutant which is similar to wild-type enzyme (Figure III-14, C and D, *broken lines*). Addition of L-serine to SPT N100W and N100Y also resulted in an increase of the ketoenamine peak but there was a wavelength shift to ~425 nm upon substrate saturation. Also, the overall ΔAbs was less for N100W and N100Y compared to the wild-type enzyme. Taken together these results suggest that the SPT:L-Ser external aldimine could be bound differently in the bacterial mimics due to perturbation of the PLP cofactor (Figure III-14, A-D).

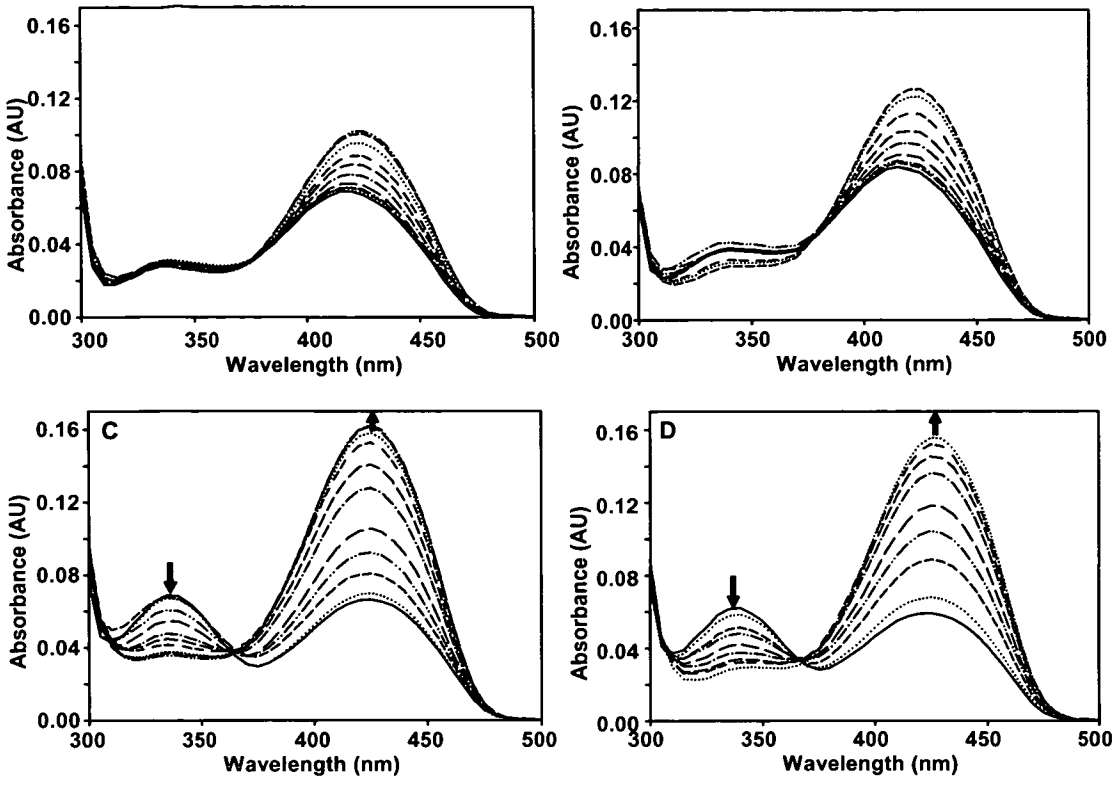


Figure III-14 UV-visible analysis of SPT wild-type and mutants. Absorbance spectra of S. paucimobilis (A) SPT N100W, (B) N100Y, (C) wild type and (D) N100C. The *solid line* in each spectrum is the holo-form of the enzyme (10 μM SPT, 20mM potassium phosphate buffer (pH 7.5), 150mM NaCl, 25 °C). Increasing concentrations of L-serine were added (0, 0.1, 0.5, 1, 2, 5, 10, 20, and 40 mM; *dotted* and *dashed lines*), and the spectrum was recorded after 15 min. For the SPT N100C, N100W, and N100Y spectra the final concentrations of L-serine were 0, 0.1, 0.5, 1, 2, 5, 10, 20, 40, 60, and 80 mM. AU, absorbance units.

The apparent serine dissociation constants ($K_{d(L\text{-}Ser)}$) were determined for all three mutants (Table III-3). A $K_{d(L\text{-}Ser)}$ similar to the wild-type enzyme was obtained for the N100C "control" mutant (2.7 mM). Higher $K_{d(L\text{-}Ser)}$ values were obtained for the two HSAN1 mutants (N100W, 16.2 mM, a 15-fold increase, and N100Y, 7.5 mM, a 7-fold increase, compared with the wild-type SPT) which reinforce the hypothesis of active site/substrates binding perturbations created by HSAN1 mutations.

Table III-3 Wild-type and SPT mutants kinetic analysis table.

Enzyme	k_{cat} x 10^3 (s^{-1})	K_m^{Ser} (mM)	K_m^{PCoA} (μM)	k_{cat}/K_m^{Ser} ($M^{-1}s^{-1}$)	k_{cat}/K_m^{PCoA} ($M^{-1}s^{-1}$)	K_d^{Ser} (mM)
SPT WT	1150.0 ± 30.0	1.4 ± 0.1	35.4 ± 2.0	821.4	32486	1.1 ± 0.1
SPT N100C	250.0 ± 4.0	7.0 ± 0.4	60.0 ± 9.8	35.7	4167	2.7 ± 0.2
SPT N100W	9.0 ± 0.4	1.6 ± 0.4	19.3 ± 2.2	5.6	466.3	16.2 ± 3.1
SPT N100Y	5.0 ± 0.1	2.5 ± 0.4	31.4 ± 3.3	2.0	159.2	7.5 ± 0.5
SPT R378A	78.3 ± 3.2	3.8 ± 1.0	39.2 ± 3.1	20.5	1997	1.1 ± 0.1
SPT R378N	33.0 ± 1.0	2.4 ± 0.2	31.0 ± 6.0	13.8	1571	3.1 ± 0.2

3-Ketodihydrosphingosine (KDS) Formation

The ability of the mutant SPT's to form the KDS product was investigated using the radioactivity assay whereby incorporation of [^{14}C]-L-serine into product was measured by autoradiography (see Section III.1.1.2). The SPT N100C, N100W and N100Y were all able to form KDS but the total amount of radioactive products was ~84% lower for the HSAN1 mimics compared to wild-type and N100C control enzymes (Figure III-15). This indicates that activity of the HSAN1 mutant mimics is compromised to a large extent even in substrate saturating conditions.

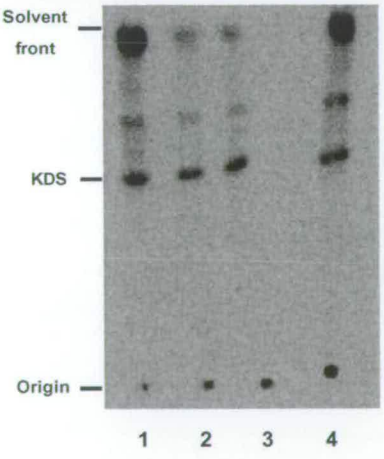


Figure III-15 Formation of 3-ketodihydrosphingosine (KDS) - TLC plate showing radiolabeled products obtained by SPT-catalyzed conversion of [^{14}C]-L-Ser and palmitoyl-CoA to [^{14}C]KDS. The *arrow* marks the position of the standard KDS. Lane 1: SPT N100C, Lane 2: SPT N100W, Lane 3: SPT N100Y, Lane 4: Wild-type SPT.

SPT Activity – Kinetic Analysis

To confirm the loss in activity of the HSAN1 mimics a more thorough kinetic analysis of these mutants was carried out using the spectrophotometric DTNB assay described earlier (see Section III.1.1.2). The k_{cat}/K_m values for the N100W and N100Y mutants (5.6, and 2.0 M^{-1} s^{-1}, respectively) were 147- and 410-fold lower, respectively, compared to wild-type enzyme (k_{cat}/K_m = 821.4 $M^{-1}s^{-1}$) (Table II-2). This confirmed that the HSAN1 mimics were catalytically much poorer enzymes. Surprisingly the k_{cat}/K_m value for the N100C mutant was ~23-fold lower than wild-type but can be explained by the reactivity of the Cys^{100} thiol with the DTNB reagent. Formation of a mixed disulfide with the reactive thiol group at the active site could potentially hinder binding of the L-serine substrate. An increased K_m value (K_m = 7.0 ± 0.4 mM) for L-serine binding but a k_{cat} value (k_{cat} = 250.0 ± 4.0) similar to wild-type is in keeping with this explanation. In addition the N100C mutant could form KDS to the same extent as wild-type under saturating conditions (Figure III-15). The K_m values for palmitoyl-CoA were comparable for mutant and wild-type enzymes suggesting that acyl-CoA affinity is not affected in the HSAN1 mimics. However a full investigation into acyl-CoA specificity for the HSAN1 mimics was not carried out because their extremely low enzyme activity would have made it difficult to measure rates for shorter acyl-CoA substrates. In addition, commercial acyl-CoA substrates are expensive.

III.1.2.4 Structural biology

Structure of N100C, N100W and N100Y holo-forms

High resolution structural studies of the *S. paucimobilis* SPT N100C, N100W and N100Y internal aldimines were the result of collaborative work between Ken Johnson and myself.

The three SPT mutants were crystallised as large yellow crystals (Figure II-16A). Four high-resolution structures were solved: three holo-form structures (N100C: 1.25Å, N100W: 1.43Å and N100Y: 1.5Å) as well as the structure of the

SPT:L-Ser external aldimine for the N100Y HSAN1 mimic (2.15Å). The enzyme:substrate complex was trapped by soaking the fully formed crystal in a solution containing 50mM of L-serine. An overlay of the wild-type holo-SPT structure with N100C, N100W and N100Y reveals that N100C is quasi identical to the wild-type with a r.m.s deviation of 0.2 Å along the 396 residues. N100W and N100Y are very similar to each other (0.1 Å r.m.s. deviation) but display noticeable differences with both wild-type and N100C (r.m.s. deviation of 0.6 Å). For clarity, only the differences between N100Y and the wild-type enzyme are discussed here. Certain parts of the SPT structure do not fully align, perceived as small shifts, when the SPT HSAN1 mutant structures are overlaid with the structure of the wild-type enzyme. These shifts are largely conserved regardless of the combination of structures aligned (N100W *versus* native, N100W *versus* N100C, N100Y *versus* N100C, and N100Y *versus* native), suggesting that they are real and not simply a crystallographic artefact. The largest of these 'shifts' (1.7Å at C_α level) occurs at the N-terminal helix (Asp^{23}–Gly^{42}), which forms an important part of the dimeric interface of SPT[39].

It might be expected that mutation of Asn^{100} to an aromatic Tyr residue could cause structural perturbations (Figure III-16B). Closer examination of the N100Y structure shows that the tyrosine side chain has "flipped" out leading to a main chain displacement of 3.4 Å of and a 180° rotation of the Ramachandran angle of the neighbouring Thr^{99} residue. As a result, the hydrogen bond observed in wild-type enzyme between the side chain of Asn^{100} and the carbonyl of Lys^{265} is not present in the mutant enzyme. Rotation of the Cα atom at position 100 has a ripple effect on close residues and large positional movements are undergone by Leu^{105}–Gly^{107} (over 4 Å for the Cα positions of Gly^{107}). The Tyr^{100} now occupies a new pocket at the dimer interface formed by hydrophobic residues from both monomers (Phe^{109} from the same monomer; Met^{51}, Thr^{72}, and Ile^{69} from the other monomer) (Figure III-16B). The movement of Y100 to a new location leads to further modification of the protein structure and in particular results in the perturbation of several residues of α-helix 3[39]. Further shifts are observed with residues His^{234} and Asp^{231} which directly interact with the cofactor. Indeed, in the native holo-SPT both His^{234} and Asp^{231} stabilises PLP[39] (Figure III-10B). Measurements of specific distances in the N100Y

mutant showed that distances between the pyridine ring and theses two residues are now too long to form hydrogen bonds (3.44 Å and 3.64 Å, Figure III-16B).

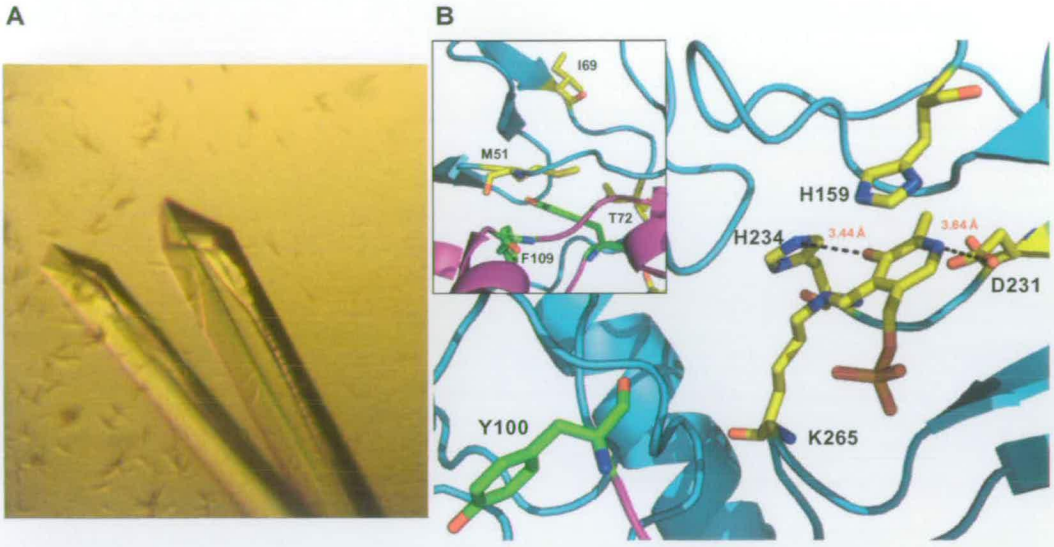


Figure III-16 SPT mutant internal aldimine (A) SPTs crystals (B) N100Y active site. One monomer is coloured in teal and the other in purple. All residues from the teal monomer are shown in stick form and are coloured yellow. Residues from the purple monomer are coloured green. Black dashed lines shows the distances between the internal aldimine, His^{234} and Asp^{231}. In B, the top left corner box shows the hydrophobic pocked occupied by Tyr^{100}.

Structural Biology of Wild Type and N100Y L-Ser External Aldimines

Crystallisation and solution of the *S. paucimobilis* N100Y SPT:L-ser external aldimine structure was the result of collaborative work between myself and Ken Johnson.

An overlay of external aldimine structures, wild-type and N100Y SPT give a smaller r.m.s deviation (0.4 Å) than the value obtained for the internal aldimine structures (0.6 Å). Thus, it appears that formation of the external aldimine drives N100Y toward a more "native" conformation. The most striking difference between the two external aldimine structures is that in the N100Y structure, the "PPATP loop" remains in its internal aldimine conformation, and consequently Arg^{378} does not enter the active site and does not make a salt link with the carboxylate of L-serine (Figure III-17). It is interesting to note that in the wild-type holo-SPT structure the side-

chains of Arg^{378} and Gln^{357} interact directly and this interaction is broken when the external aldimine is formed. In contrast, in the N100Y holo-form these two residues are too far apart to interact (distance 4.77Å). However, in the N100Y L-Ser:external aldimine form the Arg^{378}/Gln^{357} interaction is clearly observed with a shorter distance between the two residues (2.92Å).

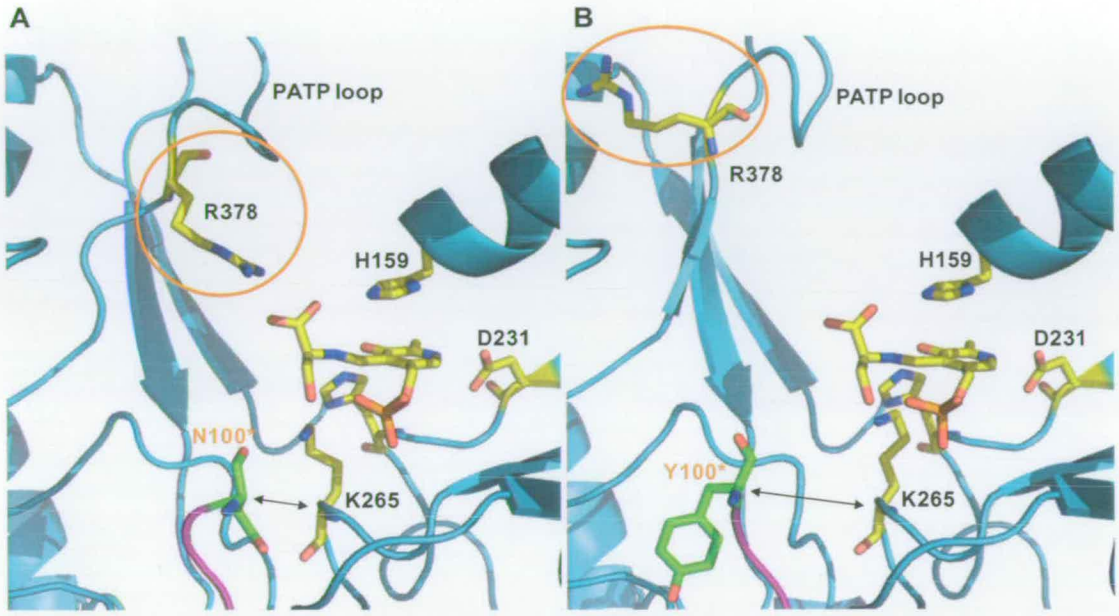


Figure III-17 Comparison of wild-type (A) and N100Y (B) SPT external aldimine. One monomer is coloured in teal and the other in purple. All residues from the teal monomer are shown in stick form and are coloured yellow. Residues from the purple monomer are coloured green.

UV and kinetics analysis suggest changes in PLP environment and in the active site in general. These results are confirmed by the several perturbations observed in the crystal structures and principally at the dimer interface. The most interesting change concerns the "non-movement" of Arg^{378} in N100Y external aldimine. These observations combined with high $K_{d(L\text{-}Ser)}$ values for both HSAN1 mutants suggest a role for the arginine residue in amino acid binding/recognition. In order to investigate further the role of this residue, two new mutants (R378A and R378N) were made and studied.

III.1.3 R378 a singular residue – SPT R378A and R378N mutants

III.1.3.1 Cloning

The two arginine mutations were introduced into plasmid pET-28a/SPT by the same method used with HSAN1 mutant.

After transformation of TOP10 cells with the mutagenesis mixture and purification of DNA samples, each mutant plasmid was analysed by restriction enzyme digest (Figure III-18) and Big Dye sequencing.

Figure III-18 Restriction enzyme analysis of pET28a/R378A, pET28a/R378N. Lane 1: Hyperladder I, Lanes 2-4: pET28a/R378A, Lanes 5-7: pET28a/R378N. Lanes 2 and 5: plasmid digested with *Nde*I, Lanes 3 and 6: plasmid digested with *Xho*I, Lanes 4 and 7: plasmid digested with *Nde*I and *Xho*I.

III.1.3.2 Expression, Purification and Mass Spectrometry Analysis

After SPT induction trials it was found that the optimal conditions for expression of SPT R378A and R378N were obtained with BL21 (DE3) competent cells at 30 °C for 5h with 0.1 mM IPTG. The protein of interest was isolated from the cell free extract by nickel affinity chromatography and the fractions containing SPT were further purified on a size exclusion column. This two-step purification process

allowed the isolation of ~50mg of pure and homogenous homodimeric SPT mutants (size ~92 kDa). Protein identity and integrity were confirmed by LC-MS.

III.1.3.3 Characterisation and Activity of R378 Mutants

Spectroscopic Properties

As seen in the previous chapter, structural analysis of wild-type and N100Y SPT reveals a specific role for R378 in external aldimine formation/stabilisation. In order to confirm and/or understand the role of this residue, binding of PLP and L-serine was analysed. UV-analysis was carried out as described with wild-type SPT (see Section III.1.1.2). SPT R378A and R378N bound PLP in a manner similar to wild-type with absorbance maxima at 335 and 425 nm (Figure III-19). The dissociation constants obtained with L-serine were similar to that of the wild-type SPT ($K_{d(\text{L-Ser})}$ values of 1.1 mM and 3.1mM for the R378A and R378N, respectively).

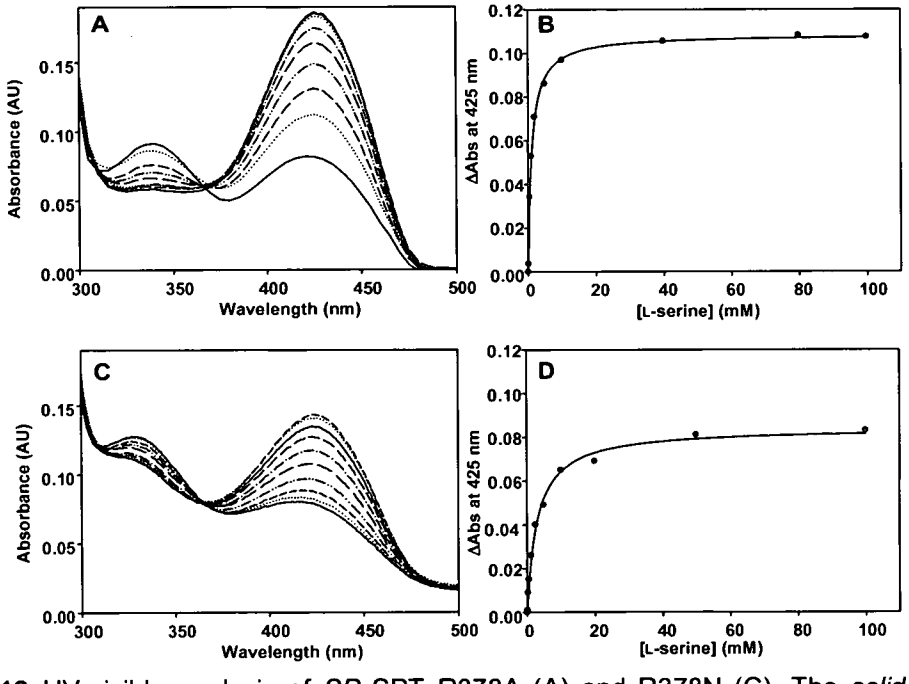


Figure III-19 UV-visible analysis of *SP* SPT R378A (A) and R378N (C). The *solid line* in each spectrum is the holo-form of the enzyme (10 μM SPT, 20mM potassium phosphate buffer (pH 7.5), 150mM NaCl, 25 °C). Increasing concentrations of L-serine were added (0, 0.1, 0.5, 1, 2, 5, 10, 20, and 40 mM; *dotted* and *dashed lines*), and the spectrum was recorded after 15 min. Dissociation constants for SPT R378A (B) and R378N (D) were measured at 425 nm. *AU*, absorbance units.

3-Ketodihydrosphingosine (KDS) Formation

The ability of each enzyme to convert [^{14}C]L-serine and palmitoyl-CoA to the product [^{14}C]KDS was examined. Only 50 and 40% of KDS was formed by R378N and R378A respectively when compared with the wild-type enzyme (at saturating substrate concentrations, Figure III-20).

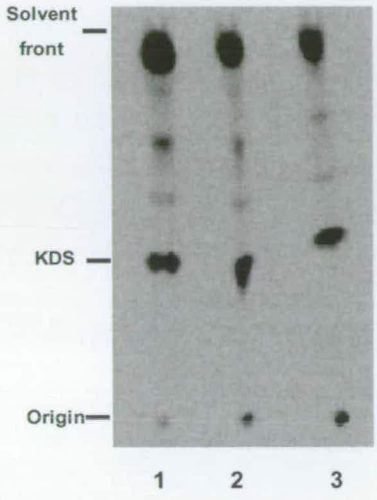


Figure III-20 Formation of 3-ketodihydrosphingosine (KDS) - TLC of radiolabeled products obtained by SPT-catalyzed conversion of ^{14}C]L-serine and palmitoyl-CoA to [^{14}C]KDS. The *arrow* marks the position of the standard KDS. Lane 1: Wild type SPT, Lane 2: SPT R378A, Lane 3: SPT R378N.

Kinetic Analysis

Using the DTNB spectrophotometric assay, K_m values for L-serine and palmitoyl-CoA with both mutants were determined (Table III-3). The $K_{m(\mathrm{PCoA})}$ values determined were very similar to the wild-type value (39.2 μM, 31 μM and 35.4 μM for R378A, R378N and WT respectively). Slightly higher $K_{m(L\text{-Ser})}$ values were calculated as 3.8 mM (R378A) and 2.4 mM (R378N) compare to 1.1 mM for wild-type SPT. The overall catalytic efficiency (k_{cat}/K_m) of the R378A and R378N mutants were 60- and 40-fold lower compared to wild-type (Table III-3).

The fact that the $K_{d(\mathrm{L\text{-}Ser})}$ values are similar for the R378 mutants and wild-type enzymes suggest that R378 is not important in binding the L-serine substrate. If this

is the case, then the interaction observed between R378 guanidinium and the L-serine carboxylate does not contribute hugely to stabilisation of the external aldimine complex. Another possibility is that another residue(s) can help stabilise the external aldimine in the absence of R378. However, it can be concluded that R378 is necessary for a fully functioning SPT since an increase in $K_{m(\text{L-Ser})}$ and large decrease in SPT activity (k_{cat}) indicates that this residue must play a role in one of the mechanistic steps. These data suggest that mutation of the R378 residue slows catalysis possibly by preventing an intermediate, probably formed after CoA binding, from achieving its optimal orientation within the active site. The step after external aldimine formation is deprotonation but this cannot occur until the second substrate binds. Binding of palmitoyl-CoA triggers rotation around the external aldimine C_α. Deprotonation of the resulting Dunathan intermediate gives rise to a quinonoid intermediate (Figure III-9 and 21). This intermediate can instantaneously react with palmitoyl-CoA to form a β-ketoacid (Figure III-9). Because of its high reactivity, the transient quinonoid intermediate cannot be observed during normal catalysis.To overcome this problem, Ikushiro *et al.*[121] synthesised an analogue of palmitoyl-CoA (*S*-(2-oxoheptadecyl)-CoA) where the reactive thioester bond is replaced by a keto/thioether group yielding an unreactive molecule (Figure III-21A). This palmitoyl-CoA mimic can bind to the enzyme and the deprotonation at C_α can occur. The resulting quinonoid cannot condensate with the analogue and its accumulation can be observed spectroscopically at 495 nm. To investigate whether R378 plays a role in quinonoid formation, stabilisation of this intermediate was monitored spectroscopically for wild-type SPT, N100Y and R378A using the thioether analogue.

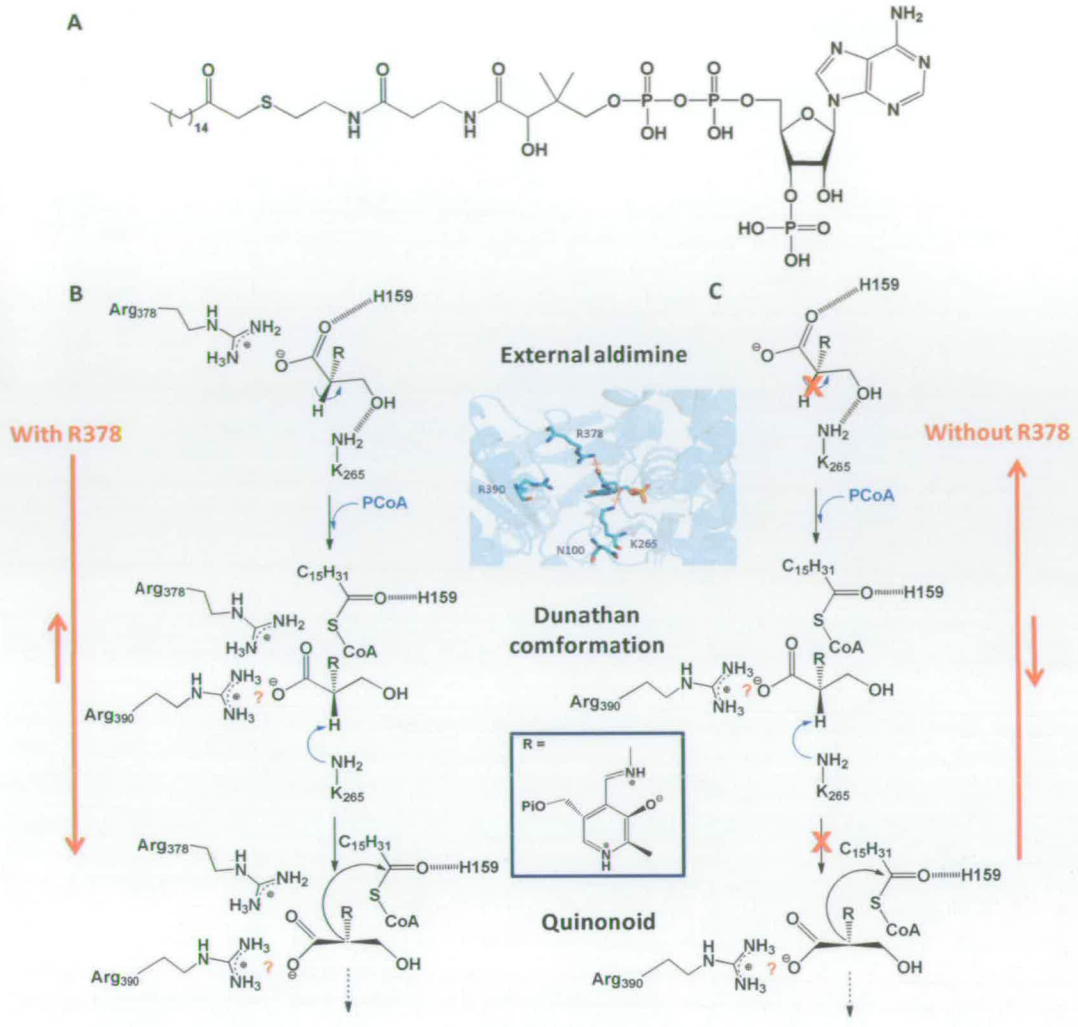


Figure III-21 (A) Chemical structure of palmitoyl-CoA analogue. Schematic representation of the control of the external aldimine conformation in (B) Wild type SPT and (C) R378A.

Quinonoid formation

The thioether analogue was added to a fully formed SPT:L-Ser external aldimine in wild-type enzyme, as well as the external aldimines of the N100Y and R378A mutants, and quinonoid formation was monitored spectroscopically. For the wild-type enzyme, a clear peak at 495 nm was observed corresponding to the formation of the quinonoid intermediate (maximal absorbance 0.06) (Figure III-22A). In contrast, the SPT N100Y mutant, under the same conditions, produced a small, broad shouldered 490–500 nm peak (maximal absorbance 0.01 absorbance units; Figure III-22B). Interestingly, no formation of quinonoid species was observed with the R378A mutant (Figure III-22C).

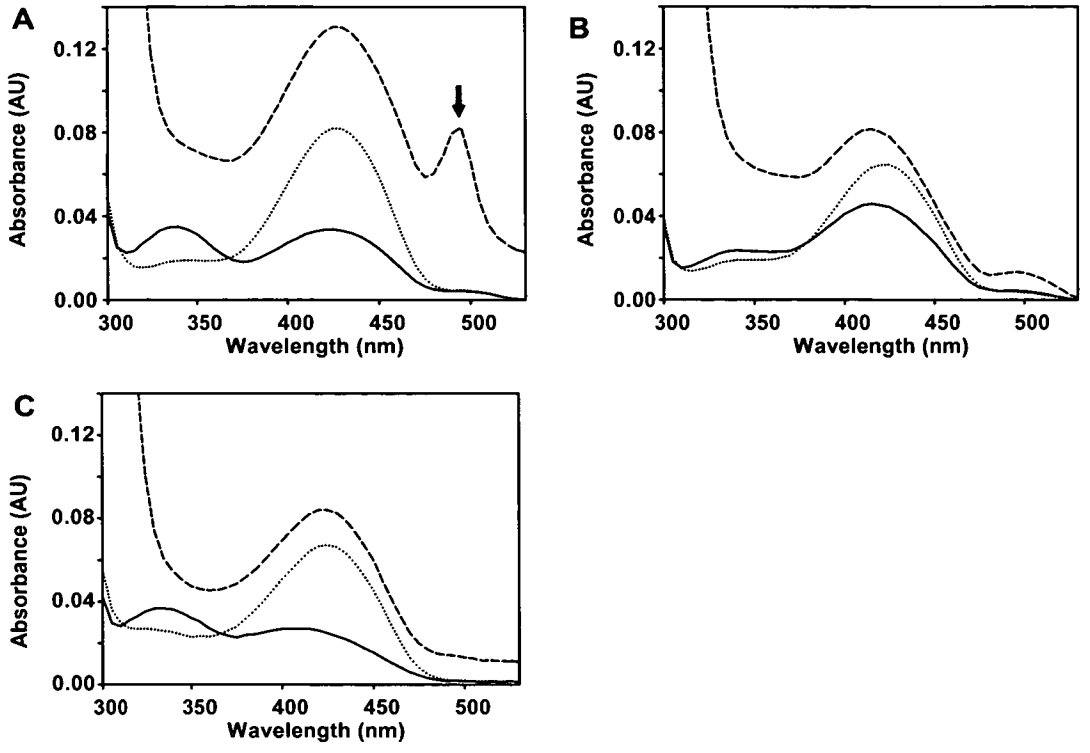


Figure III-22 Quinonoid formation in the wild-type SPT, N100Y, and R378A mutants. UV-visible spectra of the SPT wild type (A), N100Y (B), and R378A (C) external aldimine forms after titration palmitoyl-CoA thioether analog, (*S*-(2-oxoheptadecyl)-CoA). Incubations were carried out with 5 μM enzyme, 45 mM L-serine, 1.4 mM analog, 20 mM potassium phosphate buffer (pH 7.5), 150 mM NaCl for 15 min at 25 °C. The *solid line* in each spectrum represents the holo-form; the *dotted line* is the external aldimine form after the addition of L-serine; and the *dashed line* is after the addition of the analogue. An absorbance maximum at 495 nm (marked with an arrow), attributed to quinonoid species[121], is clearly observed in the wild-type enzyme, but only a broad shoulder is produced in the N100Y mutant, and no peak is observed for the R378A mutant. *AU*, absorbance units.

III.1.4 Conclusion

III.1.4.1 Role of R378 and Mechanistic Implications

In view of the fact that in the external aldimine structure Arg^{378} undergoes a large swing into the active site to make a salt link with the carboxylate of the bound L-serine, it was surprising that the R378 mutants easily formed an external aldimine with L-serine when analysed by UV-visible spectroscopy. In addition, the

dissociation constant for L-serine was almost identical for the R378 mutants and wild-type enzyme indicating that binding of L-serine is not hindered in the absence of R378. It was interesting then that the R378 mutant enzymes were not able to form the subsequent quinonoid intermediate upon titration with a palmitoyl CoA thioether analogue. This suggests that the Arg^{378} may play a role in either formation or stabilisation of the quinonoid intermediate.

Analysis of the wild-type SPT external aldimine structure shows sp^3-hybridization at C_α of the PLP:L-Ser complex (the angles are 111.5, 106.4, and 109.6°, which are close to the theoretical values for a tetrahedral carbon). The angles between each functional group and the N-imine (which is co-planar with the PLP ring) were measured as follows: H-C_α-N, 107.7°; CO_2-C_α-N, 108.4°; CH_2-C_α-N, 113.3°. None of these are perpendicular to the plane of the PLP ring and as expected the external aldimine is not in the optimal "Dunathan conformation" suitable for deprotonation. As well as the salt link with Arg^{378}, other interactions occur from the bound PLP:L-Ser complex to His^{159} and Lys^{265} that prevent rotation to the Dunathan in the absence of palmitoyl-CoA (Figure III-21B). When palmitoyl-CoA enters the active site, the hydrogen bond between the external aldimine carboxylate and His^{159} is broken[121] and leads to rotation of the Cα. Arg^{378} is still present and escort the CO_2^- to its new position. At this stage it is possible that another arginine (possibly the highly-conserved R390) make contacts with the external aldimine carboxylate to tighten/ fix the new conformation (Figure III-21B). In a final step, Lys^{265} now facing the proton, deprotonates it and the quinonoid is formed. A double C-C bond is created and flattens the Cα conformation (Figure III-21B). In the R378A mutant the guanidium side chain is absent and cannot interact with the carboxylate in the external aldimine. However, the UV absorbance maxima of the enolimine and ketoenamine forms and dissociation constant of L-serine, the formation, conformation and general environment of the external aldimine does not seem to suffer due to the absence of this residue. However, when the second substrate is added the enzyme is not able to form or stabilise a quinonoid (Figure III-22C) but deprotonation is still possible as SPT is active and able to form the product KDS (Figure III-20). This confirms the role of R378 in the control of conformation and suggests that it also plays an important role in quinonoid stabilisation. The passage

between external aldimine and quinonoid intermediates is an equilibrium controlled by R378 which when present will favours formation and stabilisation of the quinonoid intermediate (Figure III-21B) and when absent will favour stabilisation of the external aldimine and this leads to reduced SPT turnover (Figure III-21C). The role of R390 still needs to be confirmed and outstanding questions about the role of this residue arise. Does R390 assists R378 in the stabilisation of quinonoid 1, formed after deprotonation (Figure III-9) or is it involved in the stabilsation of quinonoid 2, formed after decarboxylation of the β-ketoacid (Figure III-9)?

Another possible explanation of the non-formation of the quinoinoid intermediate in the R378 mutant is that in the absence of R378, it is R390 itself that forms the $-CO_2^-$ salt link in the external aldimine. Because of this, there is no transfer of the CO_2^- when palmitoylCoA binds, and the external aldimine remains 'stuck'. It follows C_α rotation and subsequent deprotonation cannot occur, resulting in substantially decreased enzyme activity (Figure III-23). In support of this, R390 is highly conserved among AOS members and forms a salt link with external aldimines in KBL (Arg^{368}) and ALAS (Arg^{374}).

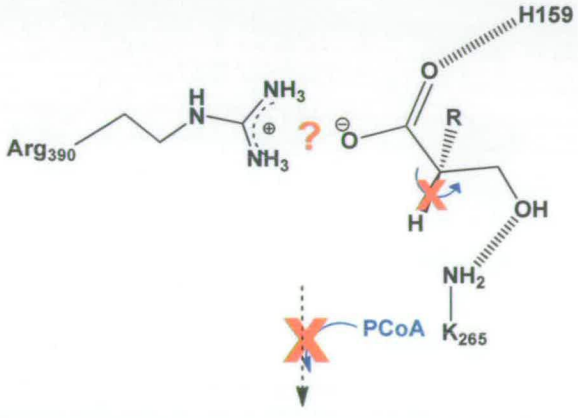


Figure III-23 Schematic representation of the possible control of the external aldimine by Arg^{390} in SPT R378A mutant.

III.1.4.2 Mimics of the HSAN1 Disease Causing Mutations

While the SPT N100C "control" mutant had similar characteristics (UV-visible spectroscopic, kinetic, and high resolution structural analyses) to the wild-type bacterial enzyme, the HSAN1 mutants showed perturbations in the cofactor environment. Noticeable changes were observed in the UV spectra (Figure III-14, *A*–

D) caused by slight shifts in the PLP-stabilising residues. In the wild-type SPT (and N100C), the side chains of N100 and C100 in one monomer make a direct contact with the amide backbone of the conserved PLP Schiff base Lys^{265} of the other monomer. The N100W/Y structures reveal that this hydrogen bond is broken (Figure III-17) creating a more mobile lysine residue and consequently more mobile internal and external aldimines. Secondly, the "swung-out" position of R378 in the N100Y:L-Ser structure means that a key interaction is missing in its external aldimine form that could also prevents formation of the quinonoid intermediate (Figure III-22B). Radioactivity assays along with kinetic analysis confirm that HSAN1 mutants can form the final KDS product but activity of the enzyme is severely compromised. It appears that local perturbations caused by a single mutation (N100Y or N100W) can 'ripple' across the interface to affect the chemical environment of the PLP cofactor. This in turn impacts upon catalysis. The fact that N100Y could not form a quinonoid upon titration with the thioether analogue suggests that the HSAN1 mimics are unable to form or stabilise essential intermediates on the catalytic pathway.

The structures of the HSAN1 mimics presented here allow an insight for the first time as to how HSAN1 mutations can result in compromised SPT activity, previously observed by Bejaoui *et al.*, Gable *et al.*[75] and McCampbell *et al.*[127] However many questions remain unanswered. Do the other disease-causing mutations (V144D, A352V, S331F) also cause deleterious structural changes that result in compromised activity? The recently identified HSAN1-causing mutation S331F seems to affect SPT in a different way to C133W and C133Y since it can increase basal activity in the absence of the small subunits. In addition the most recent topology studies place S331 in a different region of SPT to the other affected residues so it is intriguing to think how its mutation results in a similar phenotype. More and more evidence suggests that HSAN1 is caused by a build-up of 'deoxy'-metabolites arising from an SPT 'gain of function' due to a change in amino acid specificity[76-78]. The studies presented here do not shed light on this hypothesis since the HSAN1 bacterial mimics did not have an altered substrate specificity but rather severely compromised catalytic activity. Perhaps this is not surprising since alanine and glycine are relatively good substrates for wild-type human SPT ($K_{m(L\text{-}Ala)}$ ~10 mM, $K_{m(Gly)}$ ~20 mM[125] but very poor substrates for bacterial SPT ($K_{m(L\text{-}Ala)}$ ~1 M,

$K_{m(Gly)}$ ~2 M, data not shown). Although the SPT crystal structure was useful in identifying Ser^{102} as a possible residue that confers specificity on SPT for an L-serine substrate rather than L-threonine, there is little to indicate why L-alanine and glycine are such poor substrates compared to L-serine. Therefore it is difficult to speculate how HSAN1 mutations can change SPT specificity to an enzyme with preference for these substrates.

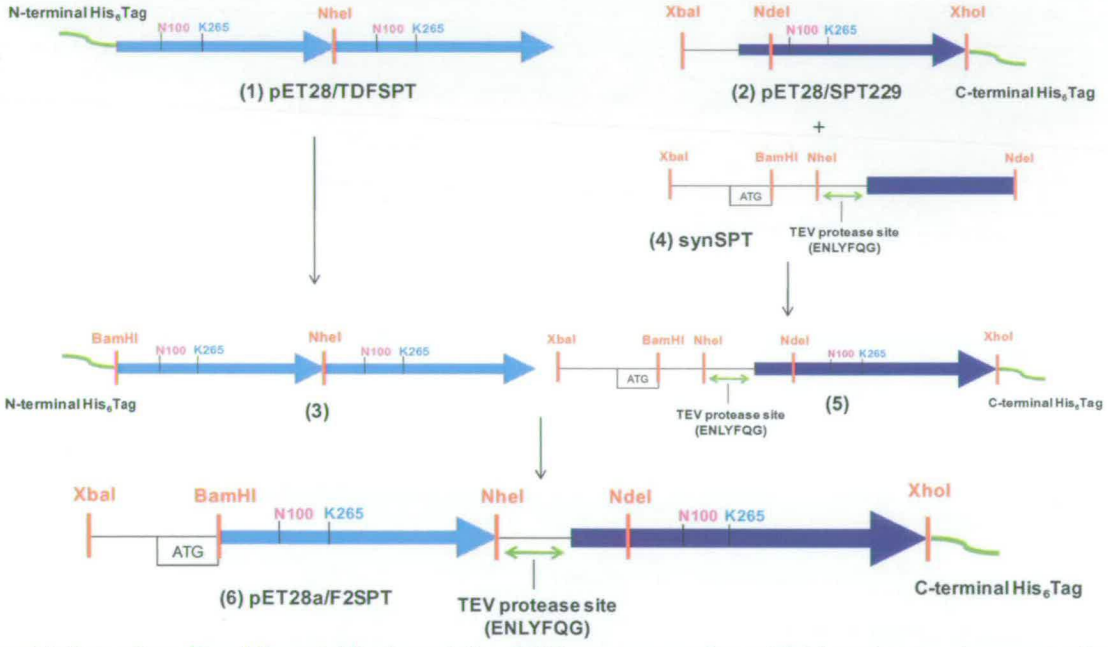
III.1.5 Construction of a bacterial fusion as an improved HSAN1 human SPT mimic

There is an obvious disadvantage when using a homodimeric bacterial SPT as a mimic for the heterodimeric human SPT since mutations generated on the bacterial SPT gene give rise to mutations on both subunits. In the case of HSAN1, mutations on human SPT occur only on the LCB1 subunit. One way around this, is to create a bacterial fusion protein comprising two identical subunits fused together by a small amino acid linker. Taking advantage of the degenerate genetic code the oligonucleotide encoding the "SPT-SPT fusion" comprises two distinct genes each coding for an identical SPT subunit. In this way it is possible to create mutations on one subunit while preserving wild-type character in the other. Successive mutations and/or insertions on one gene could give rise to a subunit that more closely resembles the inactive LCB1 subunit of human SPT. Therefore a SPT 'fusion' gene is a useful tool for creating a bacterial heterodimeric enzyme that more closely resembles human SPT.

A fusion SPT clone comprising two codon-optimised sequences was previously prepared by Teresa Dunn research group (pET28/TDF-SPT, TDF-SPT for Teresa Dunn Fusion-SPT, Figure III-24 (1)). Even though it gave rise to an active fusion SPT homodimer, it was not possible to make mutations specific to an individual subunit since the clone consisted of identical genes. However the codon-optimised sequence is distinct from the native DNA sequence in the SPT clone used by our group (pET28a/SPT229, Figure III-24 (2)). In fact there is only 37% identity between them. Therefore a hybrid fusion clone comprising one codon-optimised gene from TDFSPT and one native gene from SPT229 was prepared. First, a *Bam*HI

restriction site was cloned by site-directed mutagenesis at the C-terminus end of the pET28a/TDFSPT clone so that the first gene of this fusion could be extracted (Figure III-24 (3)). Next, a small chemically-synthesised DNA fragment comprising BamHI and *Nhe*I restriction sites, as well as the DNA sequence coding for a TEV protease site (synSPT, Figure III-24 (4)), was ligated at the C-terminus of the SPT gene in pET28a/229SPT (Figure III-24 (5)). Finally, the codon-optimised gene was ligated between these *Bam*HI and *Nhe*I sites to give rise to the required hybrid SPT fusion clone pET28a/F2SPT (Figure III-24 (6)). Further studies on this fusion have been done by a member of the group. Dr. Jonathan Lowther has subsequently expressed and purified the SPT-SPT fusion and found that it is soluble and active (personnal communication).

Figure II-24 Construction of F2SPT clone. (1) Map of the pET28a/TDFSPT clone, (2) Map of the pET28a/229SPT clone, (3) PCR product after insertion of *Bam*HI into clone pET28a/TDFSPT/(1), (4) DNA fragment chemically synthesised containing two new



restriction sites *Bam*HI and *Nhe*I, and the DNA sequence for a TEV protease cleavage site, (5) Clone obtained by ligation of synSPT/(3) into pET28a/229SPT/(2), (6) Final pET28a/F2SPT clone resulting from ligation of *Bam*HI-*Nhe*I DNA fragment from (4) into plamid (5).

III.2 SPHINGOLIPIDS BIOSYNTHESIS IN *Sphingomonas wittichii* – A MODEL SL-PRODUCING BACTERIUM.

Sphingomonas wittichii is a non-pathogenic bacteria (Figure III-25A) isolated from water of the river Elbe by RM Wittich[128]. This bacterium has stimulated much interest because of its environmental action on organic compounds[129]. Indeed, *S. wittichii* is known to metabolise dibenzo-p-dioxin (Figure III-25B) and its derivatives which are highly toxic compounds and persistent organic pollutants.

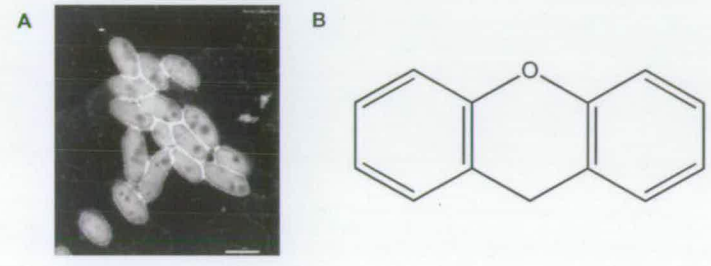


Figure III-25 (A) *Sphingomonas wittichii* strain. Credit: David Colquhoun, white scale = 1 micron from website http://genome.jgi-psf.org/sphwi/sphwi.home.html (B) Chemical structure of Dibenzo-p-dioxin.

The entire genome of *Sphingomonas wittichii* has been sequenced by the USA Department of Energy (http://www.jgi.doe.gov). Like all sphingomonas strains, the cell wall of *S. wittichii* comprises glycosphingolipids (GSL, see Section I.1.2)[16]. Since the GSL *de novo* biosynthetic pathway begins with condensation of palmitoyl-CoA and L-serine catalysed by SPT, it follows that the genome of *S. wittichii* should contain a gene encoding an SPT homologue. Indeed, a putative SPT was annotated on the genome. Interestingly a small gene encoding an 80 amino acid protein is located immediately upstream of SPT and has been annotated as an acyl carrier protein (ACP) (Figure III-26). Some SPT homologues are expressed with an accompanying ACP that brings the acyl chain to the enzyme. Examples include a seryl transferase (SerT) that along with an ACP didomain make up the PigH gene involved in prodigiosin biosynthesis[49]. On the biosynthetic pathway of the fungal natural product fumonisin, Fum8p catalyses the condensation between L-alanine and steroyl-ACP[48]. Taking these examples into consideration a possible hypothesis for *SW* SPT is that instead of using palmitoyl-CoA as a substrate it uses palmitoyl-ACP.

Figure III-26 ACP and SPT gene cluster.

III.2.1 Cloning of *Sphingomonas wittichii* ACP and SPT

The *S. wittichii* RW1 strain was a kind gift from Dr. Rolf Halden, John Hopkins University, USA. The SPT and ACP genes were amplified from the *S. wittichii* genomic DNA by PCR and expressing plasmids were engineered. ACP and SPT were cloned separately and a third clone was engineered co-expressing both proteins (Figure III-27). Three DNA sequences encoding for ACP-SPT, ACP and SPT were amplified from *S. wittichii* strain RW1 genomic DNA and introduce with the appropriate restriction sites into pET28a (ACP and ACP-SPT) and pET30c (SPT) resulting in a His_6 Tag at the C-terminal of the SPT gene and an untagged ACP gene.

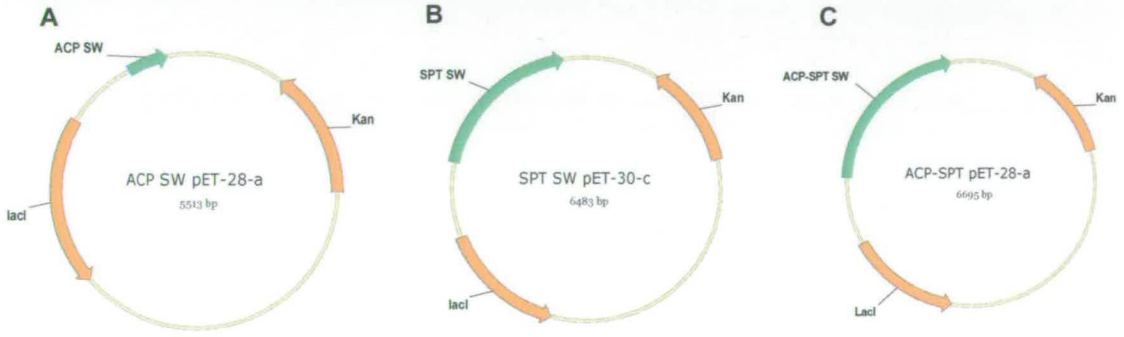


Figure III-27 Plasmid maps of *S. wittichii* expressing clones (A) ACP (B) SPT (C) ACP-SPT.

Note: After purification and characterisation of SPT from pET28a ACP-SPT plasmid (see section III.2.2 below), it was discovered that the bioinformatics analysis/annotated gene in the genome (http://img.jgi.doe.gov) contains an N-terminal extension not present in the isolated protein (at the time of the experiment). After further analysis of the SPT genomic sequence, it has been deduced that the original chosen start codon was wrong and that the correct one was located 10 amino acids further down stream. Since this observation, the genome database has been

updated and the SPT sequence contains now 400 amino acids instead of the 410 annotated previously. The correct SPT sequence was amplified form the previous ACP-SPT pET28a plasmid and cloned back into pET28a.

The SPT start codon was a valine (GTG) and has been changed to a methionine (ATG) in the new clone.

III.2.2 Expression and purification of *S. wittichii* ACP and SPT

Expression of the three plasmids described above was investigated and optimised by small-scale induction. It is interesting to note that the ACP protein could not express by itself but required co-expression of the SPT gene. The expression of SPT alone, as well as the co-expression of ACP-SPT in pET28a, was carried out in BL21 (DE3) at 30 °C for 5h with 0.1 mM of IPTG. SPT was purified by nickel affinity using its C-terminal His_6 Tag. Similar one-step purification on Ni-NTA agarose was carried out to investigate whether the untagged ACP co-purified with SPT in an ACP/SPT complex. All buffers were prepared without salt to prevent potential disruption of ionic interactions between the two enzymes. Unfortunately only SPT bound to the resin while ACP remained in the cell free extract indicating that only weak interactions, if any, form between the two proteins (Figure III-28).

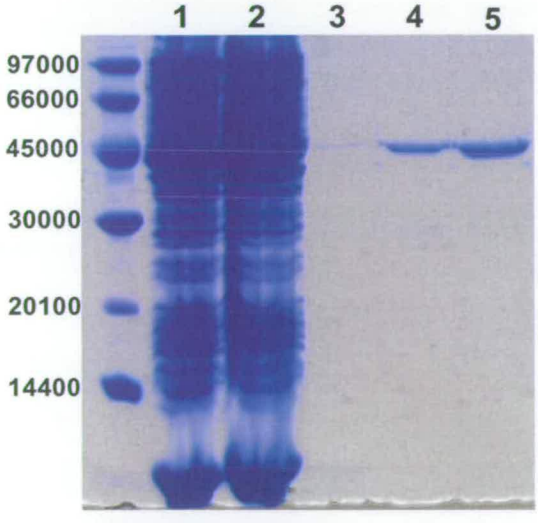


Figure III-28 SDS-PAGE S. wittichii SPT and ACP cell free extract and IMAC purification. Lane 1: cell free extract, lane 2: flow-trough, lane 3: 10 mM imidazole, lane 4: 50 mM imidazole, lane 5: 100 mM imidazole.

The *SW* SPT was then further purified to homogeneity by gel filtration. The experimentally determined mass of recombinant *SW* SPT was within 0.1% of the theoretical mass, which includes the C-terminal KLAAALEHHHHHH fusion affinity tag but lacks the N-terminal methionine. The ACP containing cell free extract was loaded on an anion exchange column and the fractions containing the protein of interest were pooled, concentrated and re-loaded on a size exclusion column. This two-step purification process allowed the isolation of ~30mg of pure and homogenous monomeric ACP from 5L culture (Figure III-29).

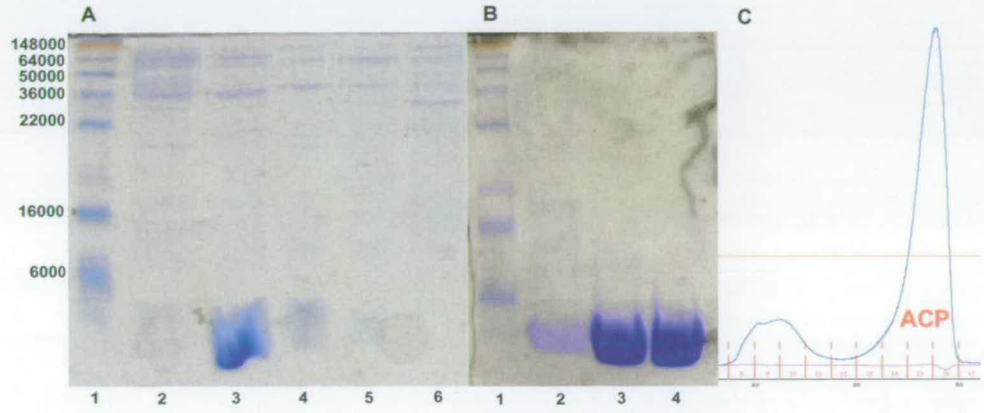


Figure III-29 Expression and purification of *S.wittichii* ACP (A) Anion exchange purification step, lane 1: See-blue marker, Lane 2-6: Q-sepharose fractions (B) Size exclusion purification step, lane 1: See-blue marker, Lane 2-4: Superdex 75 fractions (C) Elution profile of size exclusion column (blue line A_{280} nm).

III.2.3 Characterisation of *SW* SPT

Spectroscopic Properties

Using the same method as previously described for *SP* SPT the ability of *SW* SPT to bind PLP and L-serine was investigated. The UV-visible spectrum of *SW* SPT was very similar to *SP* SPT. The internal aldimine had maxima at 340 and 425 nm. Addition of increasing concentrations of L-serine gave rise to formation of the external aldimine (λ= 425 nm) and allowed measurement of the L-serine dissociation constant ($K_{d(L\text{-}Ser)}$ = 0.80 ± 0.08 mM, Figure III-30, Table III-4).

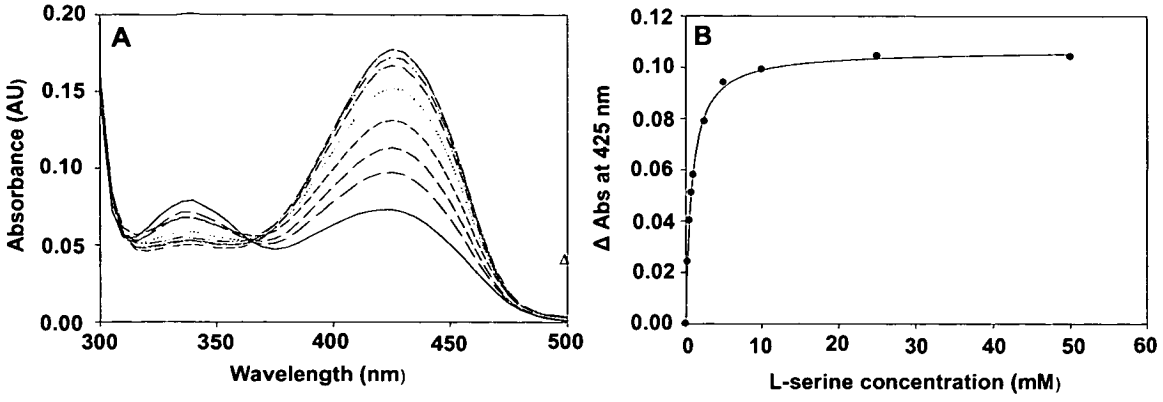


Figure III-30 UV-visible spectroscopic analysis of *S. wittichii* SPT (A) in presence of L-serine. The *solid line* in the spectrum is the holo-form of the enzyme (10 µM SPT, 20 mM potassium phosphate buffer (pH 7.5), 150 mM NaCl, 25 °C). Increasing concentrations of L-serine were added (0, 0.1, 0.5, 1, 2, 5, 10, 20, and 40 mM; *dotted* and *dashed lines*), and the spectrum was recorded after 15 min. (B) The L-serine dissociation constant was extracted from changes in absorbance at 425 nm. *AU*, absorbance units.

Table III-4 Kinetic parameters for the purified *SP* and *SW* SPTs.

Enzyme	k_{cat} x 10^3 (s^{-1})	K_m^{Ser} (mM)	K_m^{PCoA} (µM)	k_{cat}/K_m^{Ser} ($M^{-1}s^{-1}$)	k_{cat}/K_m^{PCoA} ($M^{-1}s^{-1}$)	K_d^{Ser} (mM)
SPT *S. wittichii*	68.7 +/- 1.5	0.78 +/- 0.10	23.4 +/- 4.5	88.1	2936	0.80 +/- 0.1
SPT *S.paucimobilis*	1150.0 +/- 30.0	1.40 +/- 0.10	35.4 ± 2.0	821.4	32486	1.1 +/- 0.1

KDS Formation – Acyl-CoA selectivity

The ability of *SW* SPT to convert L-serine and palmitoyl-CoA to the product KDS was examined by the radiochemical method and also by mass spectrometry (Figure III-31A, B and C). Furthermore, the ability of SPT to use other substrates than palmitoyl-CoA was investigated. It was shown that *SW* SPT was active and able to form KDS and can also use several acyl-CoA substrates from C12 to C20. The mass of each KDS derivatives was measured by mass spectrometry (C14-KDS M+H= 244 Da, C16-KDS M+H= 272 Da, C18-KDS M+H= 300 Da, C20-KDS M+H= 328 Da). The C22-KDS compound could not be observed, probably due to ionisation problems.

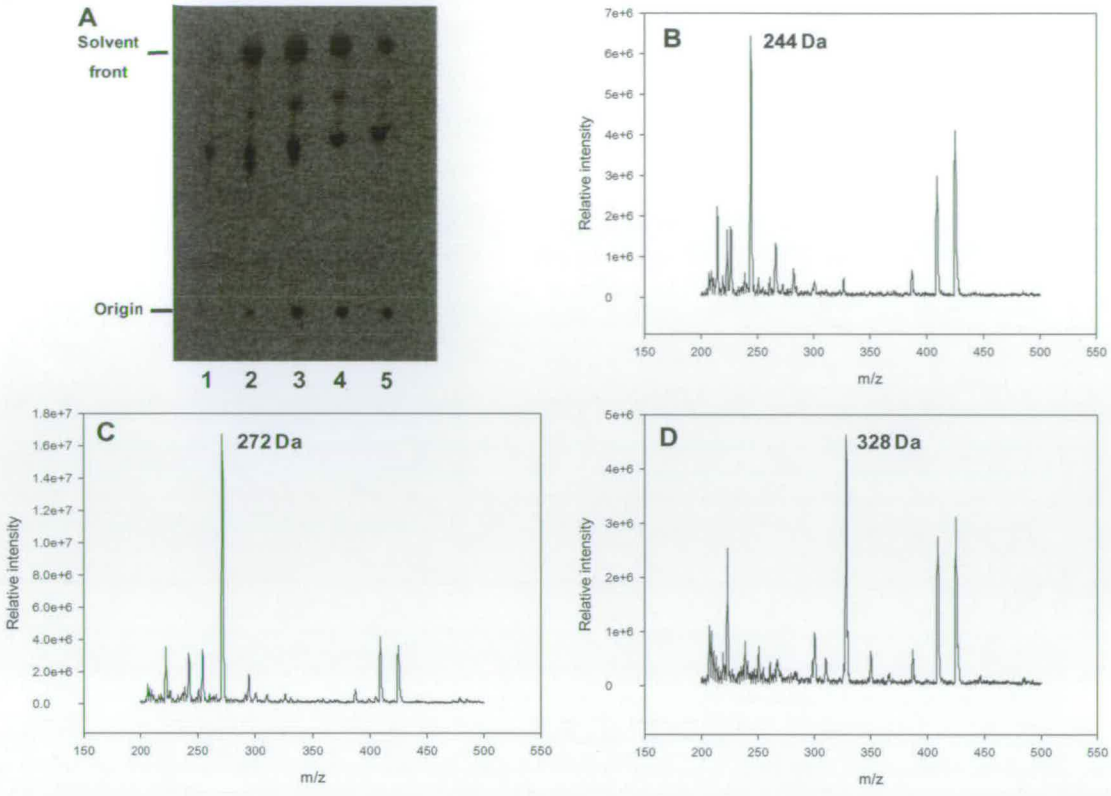


Figure III-31 Formation KDS and different acyl length derivatives (A) TLC of radiolabeled products obtained by SPT-catalyzed conversion of [^{14}C]L-serine and acyl-CoA to [^{14}C]KDS. Lane 1: 1-hydroxy-2-aminotetradecan-3-one(C14KDS), Lane 2: 1-hydroxy-2-aminohexan-3-one (C16KDS), Lane 3: 3-ketodihydrosphingosine (KDS), lane 4: 1-hydroxy-2-aminoteicosan-3-one (C20KDS), Lane 5: 1-hydroxy-2-aminodocosan-3-one (C22KDS). ESI-MS of (B) 1-hydroxy-2-aminotetradecan-3-one (C14KDS) (C) 1-hydroxy-2-aminohexan-3-one (C16KDS) (D) 1-hydroxy-2-aminoteicosan-3-one (C20KDS).

Kinetic analysis

SW SPT bound L-serine and palmitoyl-CoA with K_m values of 0.78 mM and 23.4 μM, respectively (Table III-4).The enzyme turned over with a k_{cat} of 0.069 s^{-1} and an efficiency (k_{cat}/K_m) of 87.7 M^{-1} s^{-1} for L-serine. The comparison of these values with those of *SP* SPT, show that *SW* SPT have similar binding affinities for its substrates but has a much lower turnover and efficiency (~16 fold and 10 fold respectively). This could indicate that instead of using palmitoyl-CoA, *SW* SPT utilises palmitoyl-ACP. The acyl-CoA chain length specificity was investigated. Steroyl-CoA (C18:0) displayed the fastest turnover ($k$cat = 0.083 s^{-1}) of all of the

acyl-thioesters and a K_m of 12.92 μM (Table III-5). When comparing the k_{cat}/K_m values, stearoyl-CoA (C18:0) was also the most efficient substrate (6,416 M^{-1} s^{-1}). The shorter substrates, decanoyl-CoA (C10:0) and lauroyl-CoA (C12:0), were the poorest substrates. Their binding constants and turnover could not been measured due to limit of detection of the DTNB assay.

Table III-5 Acyl-CoA chain length specificity of *SW* SPT.

Acyl CoA	k_{cat} (s^{-1})	K_m^{CoA} (μM)	k_{cat} / K_m^{CoA} ($M^{-1}s^{-1}$)
Decanoyl (10:0)	ND	ND	ND
Lauroyl (12:0)	ND	ND	ND
Myristoyl (14:0)	0.0077 ± 0.0005	39.0 ± 7.7	198
Palmitoyl (16:0)	0.0687 ± 0.0037	23.4 ± 4.5	2938
Stearoyl (18:0)	0.0829 ± 0.0015	12.9 ± 0.9	6416
Arachidoyl (20:0)	0.0530 ± 0.0019	<10	>5300

III.2.4 Structure biology

The growth of *SW* SPT crystals and the solution of the crystal structure were carried out in collaboration with Dr. Ken Johnson and Prof. Jim Naismith.

Structure of *SW* holo-SPT

SW SPT crystallised is a dimer in the asymmetric unit (Figure III-32), in contrast to the *SP* SPT where the dimer is generated by crystal symmetry.[39, 44, 124] There was no density for the disordered C-terminal histidine tag on each monomer and therefore our model contains 399 of the 413 amino acids from Ala2 to Pro399 inclusive. A PLP cofactor is bound at the active site of each monomer and at the interface with the opposite monomer.

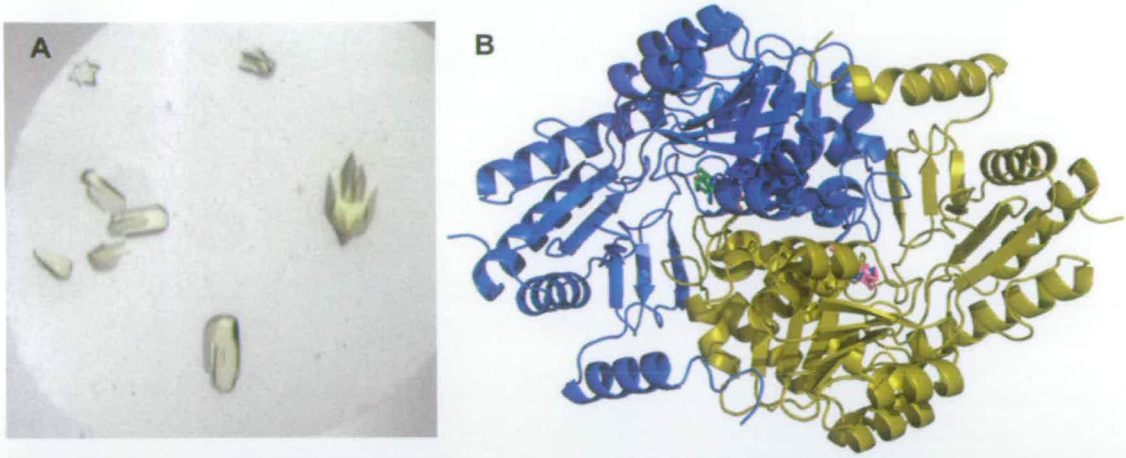


Figure III-32 Structure of *SW* SPT (A) Crystals (B) Overall structure of the SPT dimer showing one monomer in *blue* and the other in *olive*. The cofactor PLP bound to the residue Lys265 the *blue* monomer is shown in *stick form* and is *colored green*. The equivalent residue from the *olive monomer* is *colored magenta SW* SPT.

Several interactions occur between active site residues and the PLP cofactor that are present in both active sites. There are hydrogen bonds to the side chains of Asn^{118}, Thr^{242}, Thr^{274} and to main chains of Gly^{114}, Tyr^{115}, Ala^{275} (Figure III-33A). In addition His^{139} π-stacks with the pyridine ring of PLP (Figure III-33A). Two important interactions that occur in the internal aldimine of *SP* SPT and other members of the AOS family seem not to be formed or weakened in this structure. Both Asp^{211} and His^{214} appear to be too far away (3.2 and 3.1 Å respectively) to form the hydrogen bonds to the PLP that were observed in other structures (Figure III-33B). These two residues are absolutely conserved amongst all SPTs and AOS family members and are known to stabilise and anchor the PLP in the active site. Although these residues seem too far to interact with PLP in the internal aldimine, they must be involved in solution during stabilisation of other intermediates (eg the external aldimine). The conserved Arg^{358}, equivalent to Arg^{378} in *SP* SPT, is in the "swung-out" position characteristic of the enzyme holo-form[39].

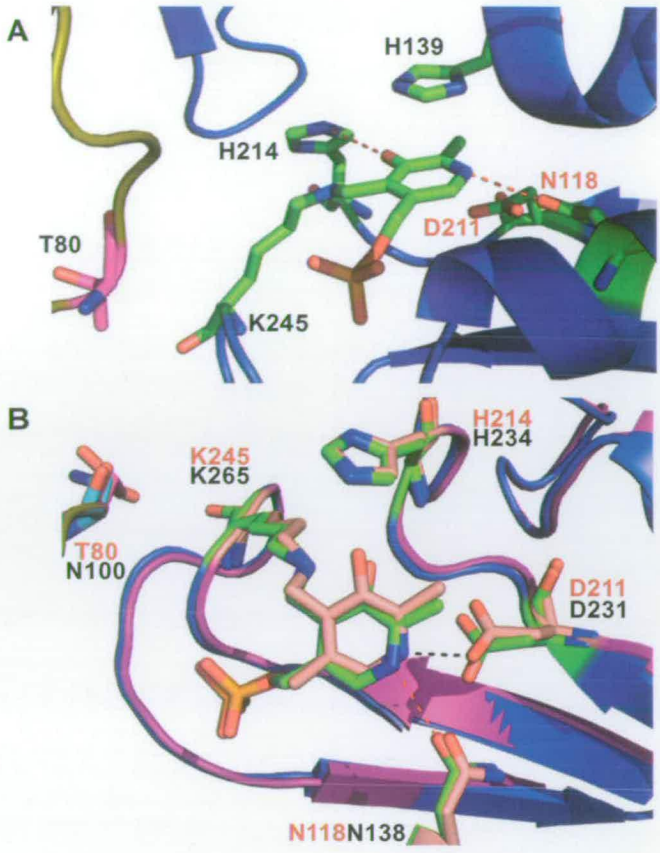


Figure III-33 (A) *SW* SPT active site. View showing the cofactor PLP of monomer A and the residues H139, H214 and N118 involved in its stabilisation. Monomer A is drawn in *blue* and monomer B is drawn in olive. The black dashed line represents the H-bond usually present between PLP and equivalent residues of H214 in others AOS enzyme structures. The red dashed line represents the "new" H-bond between PLP and N118. (B) Active sites overlay of *SW* and *SP* SPT. The black dashed line represents the H-bond between PLP and D231 of *SP* SPT. The red dashed line represents the H-bond between PLP and N118 of *SW* SPT.

The entrance to the active site situated at the "top" of the protein surface is much wider than observed in *SP* SPT, which could indicate the binding of a much wider substrate (ACP) instead of palmitoyl-CoA (Figure III-34). The entrance to the active site is delimited by four conserved residues. Vertically we can measure the distance between SWD136/SPD156 and SWL10/SPL30 (distance at the 2 C_γ) and horizontally between SWP360/SPP380 and SWF273/SPF290 (distance at ProC_β and PheC_{para}). The vertical distances in *SW* SPT are longer of 0.8 Å vertically and 1.4 Å horizontally. Both the dimer and monomer structures are similar (with a few differences) between the proteins of the Sphingomonas strains: the root mean square

(rms) deviation for the dimer is 0.65 Å (796 C alpha backbone) and 0.61 Å for the monomer (398 C alpha backbone).

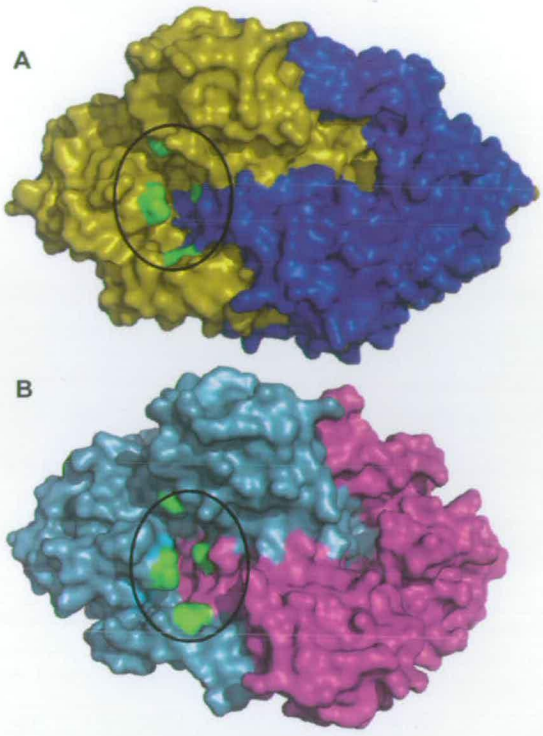


Figure III-34 Surface structures of SPTs showing their active site entrances. (A) *SW* SPT. Monomer A is show in blue and monomer B in olive. The four residue delimiting the active site entrance: D136, L10, P360, F273 are coloured green. (B) *SP* SPT. Monomer A is show in steal blue and monomer B in purple. The four residue delimiting the active site entrance: D156, L30, P380, F290 are coloured green.

Palmitoyl-CoA hydrophobic channel

The structure of the holo-form of CqsA, another PLP-dependent AOS family member, has recently been solved along with structures of the enzyme co-crystallised with various substrates.[45] Of particular interest, the product external aldimine from the condensation of L-threonine and decanoyl-CoA was trapped in the active site. A clear channel or groove was delineated where the C_{10} acyl chain binds. To define a potential channel in SPT for binding of the acyl chain of palmitoyl-CoA, the *SW* SPT and CqsA structures (PDB: 2WKA) were overlaid (Figure III-35 A and B). The protein backbone of CqsA was hidden and only the substrate decanoyl was kept into *SW* SPT structure to identify the potential residues creating the hydrophobic

channel. The channel is formed at the interface of the two monomers (Figure III-35 A and B) and comprises side-chains or backbone-chains of the following residues: Monomer A (L10, R15, 17-ALL-19, 21-TGVRDPY-27, V341, 358-RPP-361), Monomer B (84-VL-85, L266, V272). The side chains from two valines (V23 and V341 from the same monomer) appear to delimit the end of the channel.

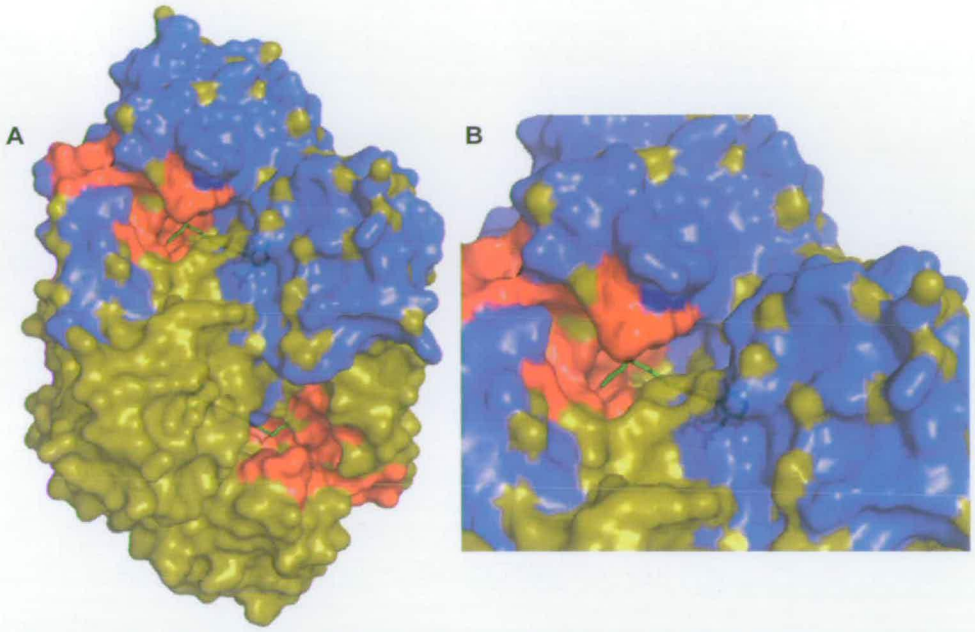


Figure III-35 Surface structure of *SW* SPT showing palmitoyl-CoA binding cavity. (A) View from the top. (B) Top view zoom. Monomer A is show in blue and monomer B in olive. The two hydrophobic cavities are coloured in red. Decanoyl-CoA is drawn in green stick form.

III.2.5 Characterisation of *SW* acyl carrier protein

Acyl carrier proteins (ACP) require post-translational modification to be active[81] (see section I.4). The apo-protein is converted to the holo-form by addition of a 4'-PP arm from coenzyme A[103]. As discussed in the introduction (section I.4) ACP is used in fatty acid and complex natural compound biosyntheses. Varying fatty acid synthase (FAS) and polyketide synthases (PKS) systems exists in mammalian, plants and bacteria but some elements of a similar system like ACP or ACPS issued from different organisms can be interchangeable. Indeed, the high degree of structural similarity within the ACP family means that ACPs from other bacteria, plants and mammals can function as an acyl group carrier with the enzymes of *E. coli* FAS and can be subsequently modified in their holo- and acyl-forms when over-

expressed in *E.coli*.[88, 89, 130] After expression and purification of *SW* ACP, the protein was analysed by LC-ESI-MS to determine the post-translational status of the protein. A mass of 8781 Da was measured, corresponding to the apo-form theoretical mass (M+H) lacking the *N*-terminal methionine (Figure III-36). This result suggests that *SW* ACP must not be recognised by *E. coli* ACPS enzymes and therefore not converted into the active holo-form required for activity. To this end, PPTases more promiscuous than *E. coli* ACPS, were tested to modify *SW* ACP in vitro.

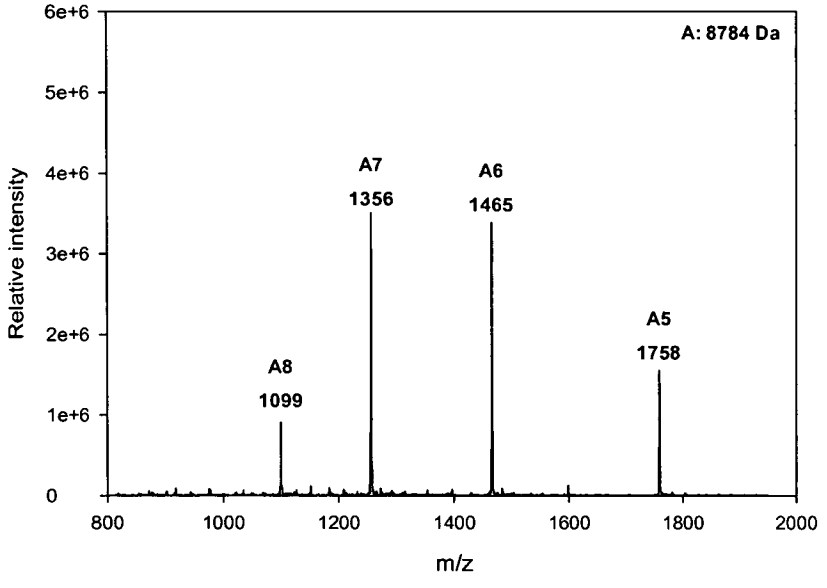


Figure III-36 ESI-MS of *SW apo*-ACP. Values A8, A7, A6 and A5 represent the different protonation states of the protein.

III.2.6 Modification of *SW* ACP by PPTases

In general, ACPs are modified by PPTases by addidtion of the 4'-PP group and subsequently the fatty acid chain is built up step-by-step by a fatty acid synthase. However, some PPTases are able to catalyse the direct transfer of an acyl group to ACP from an acyl-CoA.[48, 49, 107-109] In this study, three PPTases were investigated: *S. coelicolor* ACPS, *B. Subtilis* Sfp and *S. verticillus* Svp. These were expressed, purified and incubated with ACP, coenzyme A and several acyl-CoA in an attempt to transfer the corresponding group to ACP. Furthermore, an ACPS gene was identified in the *S. wittichii* genome and the corresponding DNA sequence was amplified and cloned in order to investigate the function and activity of the protein.

III.2.6.1 Cloning of *SW* ACPS and *Streptomyces verticillus* Svp

The ACPS DNA sequence was amplified from *S. wittichii* strain RW1 genomic DNA and was introduced using appropriate restriction sites into pET22b and to add a His_6 Tag at the C-terminal. *SW* ACPS original start codon was Val (GTG). This has been changed to Met (ATG). The *Streptomyces verticillus* (*SV*) Svp gene has been chemically synthesised from the sequence described in Sanchez *et al.*[108] The synthetic gene contained the restriction site *Nco*I and *Xho*I to ligate the gene into pET22b and to obtain a His_6 Tag at the *C*-terminal.

III.2.6.2 Expression, Purification and Characterisation

BL21 (DE3) competent cells transformed with the relevant plasmid were induced with 0.1 mM IPTG and incubated at 30 °C for five hours. These were the optimised conditions for the over-expression of soluble *SC* ACPS, *SW* ACPS and *B. subtilis* (*BS*) Sfp. Under these conditions, expression of Svp was too high leading to protein aggregation. To solve this problem, Svp expression was slowed-down/reduced by using HMS174 (DE3) competent cells and stopping the induction after three hours. Co-expression of *SW* ACP and ACPS was performed in BL21 (DE3) competent cells with 0.1 mM IPTG at 30 °C for four hours (Figure III-37).

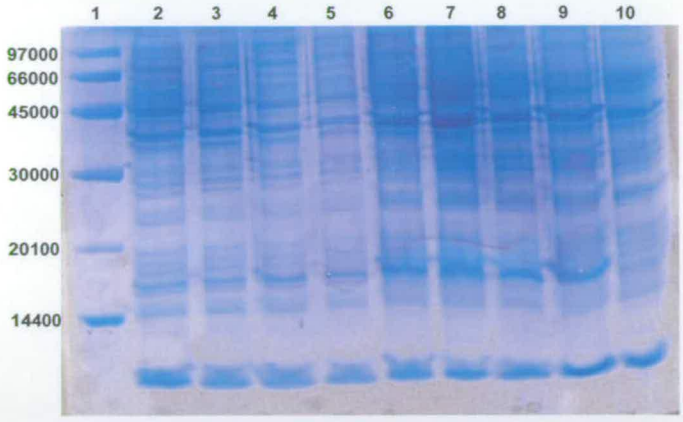


Figure III-37 SDS-PAGE of the induction trial of S. wittichii ACPS and ACP co-expression. Induction tested in BL21 (DE3) competent cells at 30 °C. Lane 1: Low molecular weight protein marker, lane 2: induction with 0.1 mM IPTG for 3h, lane 3: induction with 0.25 mM IPTG for 3h, lane 4: induction with 0.5 mM IPTG for 3h, lane 5: induction with 1 mM IPTG for 3h, lane 6: induction with 0.1 mM IPTG for 4h, lane 7: induction with 0.25 mM IPTG for 4h, lane 8: induction with 0.5 mM IPTG for 4h, lane 9: induction with 1 mM IPTG for 4h, lane 10: non-induced cells.

All proteins were purified by nickel affinity chromatography (Figure III-38A to D).

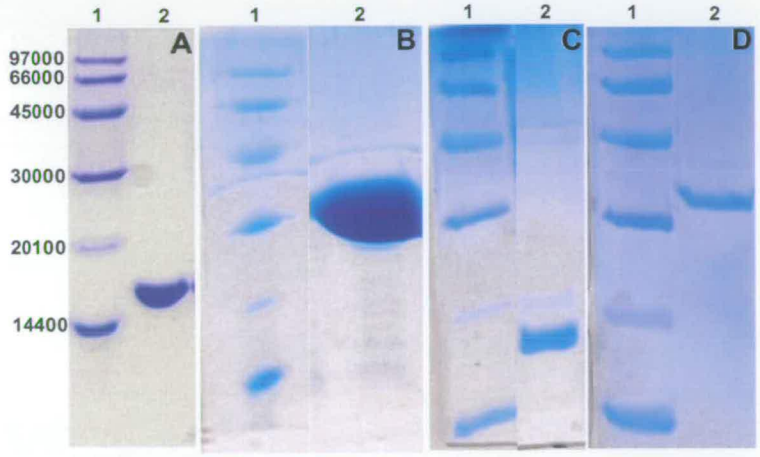


Figure III-38 SDS page of purified PPTases. (A) *S.coelicolor* ACPS (Mw = 14729 Da) (B) *BS* Sfp (Mw = 28373 Da) (C) *SW* ACPS (Mw = 15548 Da) (D) *SV* Svp (Mw = 26682 Da).

A second purification step by gel filtration was performed in some cases (Figure III-39A, B and C). *SC* ACPS was obtained in the trimeric form and Svp as a mixture of dimeric and hexameric forms (Figure III-39A and B). Svp expression and isolation were not easily reproducible and each purification gave different results and a different ratio of dimer versus hexamer (Figure III-39B).

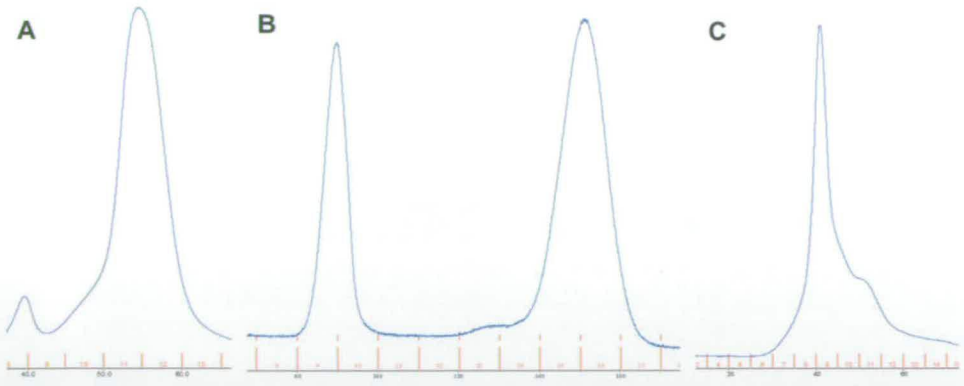


Figure III-39 PPTase size exclusion elution profiles (blue line A_{280} nm) (A) *SC* ACPS (B) *SV* Svp (C) *SW* ACPS.

III.2.6.3 Modification of ACP

The protocol used to modify the *SW* apo-ACP with the different PPTases was modified from previously described methods[107-109]. For brevity, the protocol is described using the generic name PPTase that refers to each of the recombinant ACPS enzymes used; *SC* ACPS, *BS* Sfp, *SV* Svp and *SW* ACPS. ACP in reaction buffer was incubated with PPTases, $MgCl_2$ and CoASH/acyl-CoA at 37 °C overnight. The reaction mixture was then injected directly onto an LC-ESI-MS platform for analysis. For each PPTase a different reaction buffer was used. For *SC* ACPS and S*V* Svp, 50 mM Tris buffer, pH 7.8 was used. For *BS* Sfp the reaction was performed in 75 mM MES buffer at pH 6.0 and for *SW* ACPS this was carried out in 20 mM potassium phosphate buffer (pH 7.5). DTT (5 mM) was added to the reaction mixture when using free CoASH as a substrate. It is important to note that a concentration greater than 1 mM of $MgCl_2$ could not be used for reaction with acyl-CoA as this lead to precipitation of these compounds and the reaction did not take place. Concentrations of enzymes and reactants were varied until optimal conversion conditions were achieved.

Several lengths of fatty acid chain (decanoyl C10, lauroyl C12, myristoyl C14, palmitoyl C16 and steroyl C18) CoA derivatives were tested as substrates that could modify the ACP. The first attempt to modify *SW* ACP with *SC* ACPS was unsuccessful. *SC* ACPS is well characterised and was previoulsy used to modified several ACPs with a large diversity of substrates[109, 131]. This negative result suggests

that *SC* ACPS does not recognise *SW* ACP. Sfp successfully modified ACP into the holo-form (Figure III-40) and the C10 (Figure III-40B), C12 (Figure III-40C) and C14 (Figure III-40D) acyl-forms. However, Sfp failed to transfer a palmitoyl (C16) or a steroyl (C18) chain onto the ACP. Palmitoyl chain transfer by Sfp was further explored by testing several reaction conditions (enzymes, substrate, co-factor concentrations, temperature, time) without success. In the literature Sfp has been successfully used to modify a large number of ACP and PCP with various substrates but never with long fatty acid chains ($\geq$ C16)[49, 107, 132].

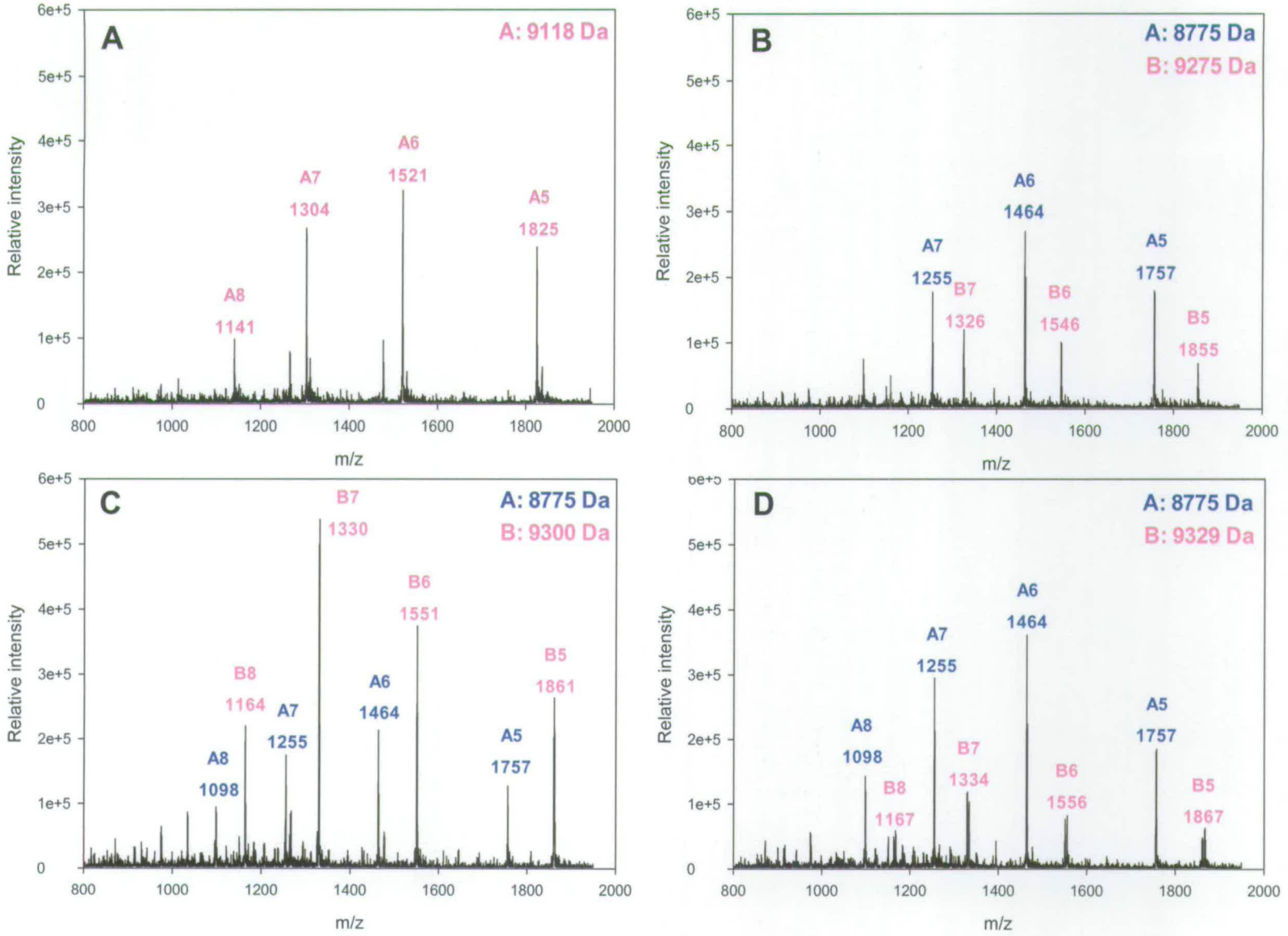


Figure III-40 Mass spectometry of modified ACP.
(A) Holo-ACP
(B) Decanoyl-ACP
(C) Lauroyl-ACP
(D) Myristoyl-ACP

Non-converted apo-forms of ACP are shown in blue. Holo and acyl-ACP forms are shown in magenta.

Another interesting PPTase, Svp, with similar characteristics to Sfp, has been shown to have a preference for longer carbon chains[48]. Incubation of Svp with ACP successfully produced the holo-form of ACP but failed to transfer any acyl groups. Hence Svp can accept *SW* ACP as a substrate (producing the holo-form) but is unable to tranfer an acyl group onto it. The problem with the long chain transfer must come from the ACP itself.

SW ACPS was successfully expressed and purified and used to obtain the holo-form of the *SW* ACP. Furthermore, *SW* ACP and ACPS have been co-expressed in *E. coli* to see if, after modification *in vivo* by *SW* ACPS the holo-ACP would be recognised by the FAS enzyme of *E. coli* and build a long chain fatty acid on *SW* ACP. After co-expression, ACP was isolated from the cell free extract, purified and analysed by high resolution Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR-MS). The ACP fraction analysed reveals a mixture of several compounds (Figure III-41), including the apo and the holo-forms of ACP. The most abundant compound corresponds to ACP modified with phosphopantetheine and acetyl group (+ 381.2 Da). Another pair of peaks was also observed, the first corresponding to an unidentified modification of apo-ACP (apo-ACP + X) and the second corresponding to further addition of phosphopantetheine and acetyl again (+ 381.2 Da). It appears that the fatty acid chain cannot extend past the acetyl-form and this further confirms that *SW* ACP is not recognised by *E. coli*.

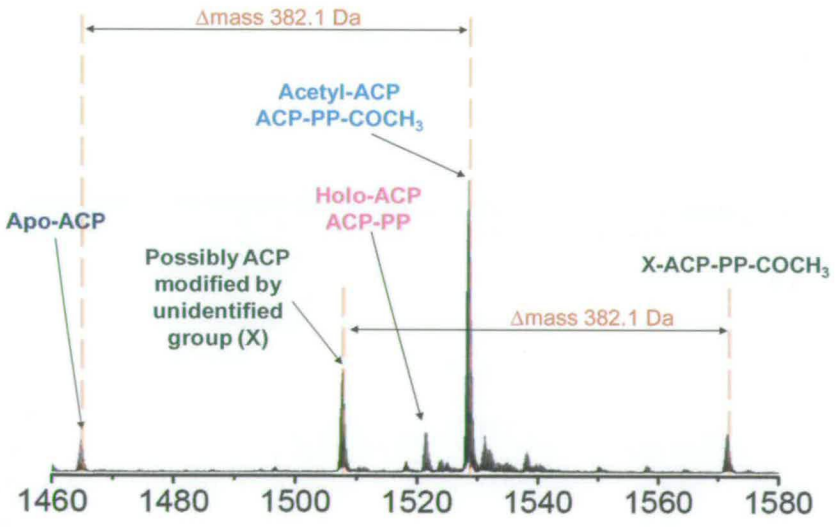


Figure III-41 High-resolution FT-ICR MS analysis of *SW* ACP co-expressed with *SW* ACPS.

High resolution mass spectrometry coupled with fragmentation techniques were used on both apo- and holo-ACP samples (Figure III-42A, B and C) to confirm that the sidechain of Ser[39] (within the conserved DSL motif residue, Figure III-43) is the site of 4'-PP attachment on the *SW* holo-ACP.

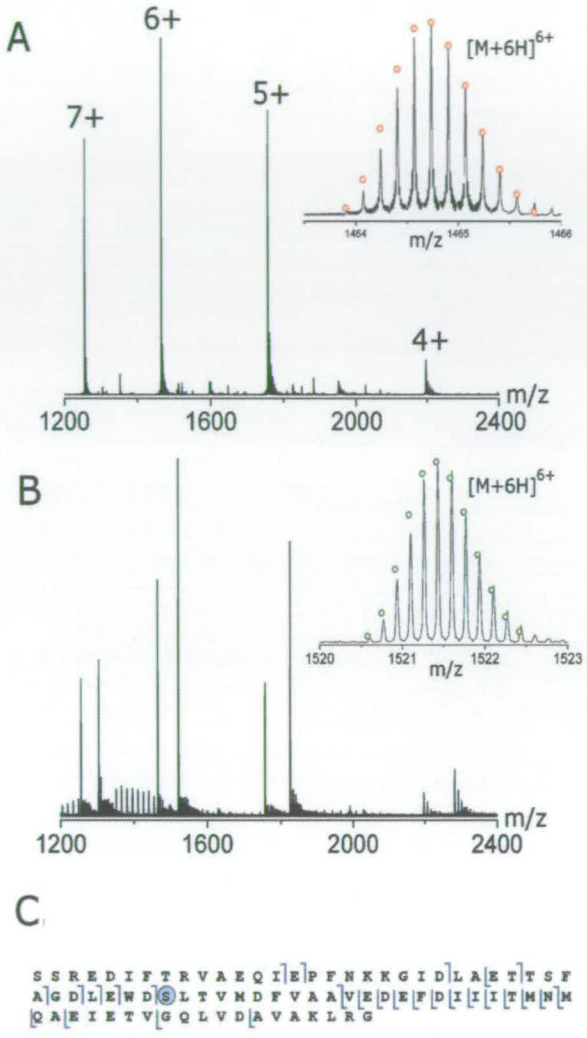


Figure III-42 (A) FT-ICR mass spectrum of apo-ACP. 4 charge states were observed corresponding to $[M+4H]^{4+}$ to $[M+7H]^{7+}$ forms of apo-ACP. *Insert*, the isotopic distribution of the $[M+6H]^{6+}$ is consistent with that predicted for ACP (red circles, [C387 H617 N97 O129 S3]$^{6+}$; error 2ppm). (B) FT-ICR mass spectrum of apo-ACP and holo-ACP. A second species is observed with charge states were corresponding to the $[M+4H]^{4+}$ to $[M+7H]^{7+}$ forms of apo-ACP (nominal mass +339 Da). *Insert*, the isotopic distribution of the $[M+6H]^{6+}$ species is consistent with that predicted for holo-ACP (green circles, [C398 H638 N99 O135 S4 P1]$^{6+}$; error 2ppm). (C) Top-down fragmentation map of holo-ACP. A single *m/z* species was isolated and subject to CID. The resulting *b* and *y* ion fragments, shown in blue, were mapped onto the primary sequence of ACP using Prosight-PTM. Fragmentation allowed the assignment of the phosphopantetheinylation to a 10 amino acid section of the protein containing Ser39 which is within the DSLT motif. Raw data in appendice.

III.2.6.4 ACP structural model

Although protein sequence alignment revealed that *SW* ACP resembles FAS II ACPs (Figure III-43), there are some differences that could be responsible for non-recognition by *E. coli* ACPS and also the inability of Sfp to transfer the 4'-PP arm and acyl chain of palmitoyl-CoA to the ACP. Of the six ACPs aligned in Figure III-42 only *SW* ACP contains a tryptophan residue positioned immediately before the DSL catalytic sequence motif. The bulky sidechain of tryptophan (Trp^{37}) is likely to have structural consequences that would be characteristic of *SW* ACP (Figure III-43), perhaps indicating a specific role for this residue.

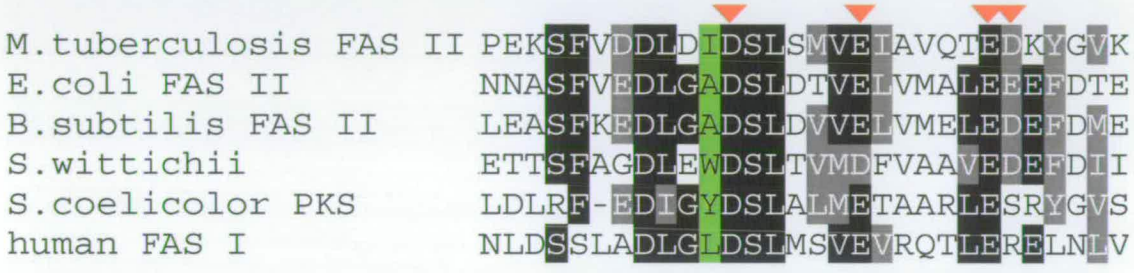


Figure III-43 ACPs alignment. Zoom on recognition residues: D38, D44, E50 and D51 (*SW* numbering).

Numerous ACP structures are available and despite the diversity of ACP in nature, their secondary structure and overall folding are very similar[88]. Using 3D-JIGSAW server[133-135] and Swiss-model[136-138], 3D-models of *SW* ACP have been generated in order to understand the function of the Trp^{37} residue (Figure III-44A). Superposition of these models shows two possible conformations for the Trp^{37} residue. With the residue positioned on a flexible loop, it is plausible to imagine the Trp side-chain moving between open and closed conformations thereby selecting the entrance of different substrates. Superposition of the 3D-JIGSAW model with the ACP *E. coli* structure (Figure III-44B) shows the tryptophan residue blocking the entrance to the active site and this supports the idea that this tryptophan could control binding and/or release of phosphopantetheinyl or acyl groups. A possible explanation for the non-modification of our ACP with C16- or C18-CoA esters is that the tryptophan acts as a gate which requires a specific signal or recognition to allow longer chains to be transferred.

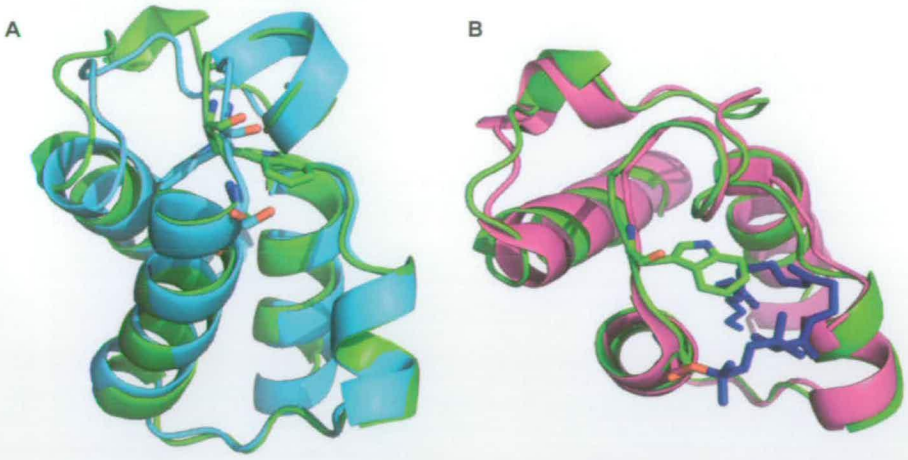


Figure III-44 3D-models of *SW* ACP (A) Superposition of Swiss and JIGSAW ACP 3D-model. The ACP Swiss-model is coloured cyan and the JIGSAW-model is coloured green. In each model, Trp^{37} is shown in stick form. (B) Superposition of JIGSAW model and *E.coli* ACP structure (PDB: 1LOH). JIGSAW-model is coloured green and *E. coli* structure in magenta with the acyl group shown in stick form and coloured in dark blue.

III.2.7 Conclusion

SW SPT was expressed, purified and characterised using a number of techniques which showed that the enzyme binds its co-factor PLP and its substrate L-serine. *SW* SPT was enzymatically active but kinetic analysis revealed that the enzyme was ten-fold slower than its homologue *SP* SPT. One explanation for such a large discrepancy is that *SW* SPT utilises an acyl-ACP instead of acyl-CoA substrates. This hypothesis is tempting given that the sequenced *S. wittichii* genome reveals a gene encoding an ACP lying immediately prior to the SPT gene. In order to investigate this relationship, a recombinant form of the ACP was expressed and subsequently purified. ACP was not modified to the acylated active form by the FAS system in *E. coli* and was isolated in its inactive apo-form. Four phosphopantetheinyl tranferase (*SW* ACPS, *SC* ACPS, *BS* Sfp and *SV* Svp) were expressed, purified, characterised and subsequently used to modify the ACP to its holo- and acyl-forms. Three of these enzymes successfully placed a 4'-PP arm on ACP. Only Sfp managed to transfer acyl-groups. Transfer of a palmitoyl group (C16) on *SW* ACP was unsuccessful and is possibly due to a recognition problem between *SW* ACP and the

PPTases used. Arthur *et al.*[139] recently described the different interactions created between *SC* ACP and two PPTases from FAS or PKS systems. ACP interacts with FAS PPTs by ionic interaction while the recognition between PKS PPTases and ACP is made *via* a hydrophobic interface. *SW* ACP is closely related to other FAS II ACPs and despite possessing the four necessary residues for recognition by ACPS (D38, D44, E50 and D51, Figure III-43), was not always recognised as a substrate by the ACP synthases used in this study. Modelling studies reveal that these problems could be caused by an uncharacteristic tryptophan positioned at the beginning of helix II. Further investigation by site-directed mutagenesis is required to confirm whether removal of the sidechain from this Trp could improve modification to the acylated active form by ACPS. If the tryptophan is effectively involved in the recognition of phosphopanthetenyl transferase and/or regulation of acyl-goup length/entrance, the corresponding mutant should be recognised by a larger number of PPTases and/or accept longer substrates. Design, expression, isolation and characterisation of this mutant are the aim of a new project. Another solution would be to chemically modify the ACP using N-acylimidazole chemistry.[140]

III.3 GENERAL CONCLUSIONS AND FUTURE WORK

The PLP-dependent enzyme serine palmitoyltransferase (SPT) is a fascinating enzyme to study for many reasons. Firstly, it catalyses the first and rate-limiting step in *de novo* biosynthesis of sphingolipids, a family of natural products fundamental to membrane composition in mammals that have gained increased importance as potent signalling molecules in numerous cellular processes. Secondly, eukaryotic SPT is unusual among the AOS family (which together catalyse similar chemical reactions) because it is a heterodimer. While all other members of the AOS family reported to date are homodimers, with each subunit binding a PLP cofactor and being catalytically competent, human SPT comprises only one catalytic monomer (*lcb2* encoding SPT2) and an other which lacks the necessary residues to bind PLP or carry out catalysis (*lcb1* encoding SPT1). The discovery of small regulatory subunits (ssSPTs) in various eukaryotes as well as novel phosphorylating Orm proteins, reported during the course of this work, have added to the elaborate organisation of an increasingly large membrane-bound human SPT complex. Uncovering how all these components come together to regulate SPT activity is an exciting prospect. Lastly, puzzling questions remain regarding why single point mutations in human SPT specifically on the "inactive" SPT1 subunit that appears not to bind PLP leads to the neurological disorder HSAN1. Structural information on the human enzyme remains elusive so that at the moment we have to rely on models from soluble cytosolic bacterial SPT's. At the start of the project the first crystal structures of an apo-SPT and a holo-SPT (PLP-bound form) from *Spingomonas paucimobilis* had been reported by this laboratory.

The work presented here encompasses structural, spectroscopic and kinetic analyses of recombinant forms of SPT from two bacterial strains of *Sphingomonas*. The crystal structure of SPT from *S. paucimobilis* with its amino acid substrate L-serine bound at the active site is presented here and has provided the next step. Capturing this enzyme-bound PLP:L-Ser external aldimine intermediate by x-ray crystallography along with the identification of quinonoid intermediates by UV-vis spectroscopy for wild-type and mutant SPT enzymes has provided an insight into how bacterial SPT controls catalysis of the Claisen-like condensation between L-

serine and palmitoyl CoA. An unexpected finding was the importance of an arginine residue (R378) located on a mobile loop that swings into the active site to form a contact with the carboxylate of the PLP:L-Ser complex. This residue is necessary for a fully functioning enzyme and R378A and R378N mutants had severely reduced activity. Interestingly, this residue is not conserved among the AOS family indicating that although all members of the family catalyse the condensation of amino acid and acyl-CoA substrates, the way that each particular enzyme controls this reaction is specific to that AOS member. A second arginine residue (R390) conserved in all other SPT homologues, also points into the active site of the external aldimine structure but does not make any contacts. Future studies utilising an R390 mutant will be useful to investigate whether this residue is necessary to stabilise quinonoid intermediates subsequent to the external aldimine. The co-crystallisation of SPT with various substrates, inhibitors and analogues would shed further light on how SPT binds its intermediates during normal catalysis. For instance, the first structure of a quinonoid intermediate is attainable by capturing SPT in complex with the palmitoyl CoA thioether analogue. This structure would also identify the binding site of the acyl chain moiety of the palmitoyl-CoA substrate. The determination of SPT structures in complex with other ligands such as D-serine, L-threonine, and the product KDS, could give valuable information on what active site residues confer specificity on SPT for its amino acid substrate and product binding. The inhibitor myriocin is also a valuable ligand since it mimics the β-keto acid intermediate formed during normal catalysis. One other feature that has not been probed in any AOS enzyme is the nature of the decarboxylation step. What triggers loss of CO_2 from the β-keto acid intermediate and by what route does this product leave the active site? It may be that the conserved R390 residue is the last SPT side-chain to make contact with the carboxylate before it is expelled from the intermediate.

A characterisation of bacterial mutant SPT's that mimic mutations found in human SPT of HSAN1 patients is presented here. These mimics have helped our understanding of how a single point mutation can cause such a debilitating disease. Structural analysis of these mimics demonstrates how a mutation on one monomer can ripple across the protein interface to cause structural changes on the opposite monomer. Although these structural changes were subtle, binding of the essential

PLP cofactor was perturbed, catalytic efficiency was decreased, and an overall more rigid structure did not allow the L-serine substrate to be bound in the same way as in wild-type enzyme. There is growing evidence in other laboratories that HSAN1 mutations result in a 'gain of function' that could arise from relaxed specificity for the amino acid substrate in the mutated SPT. These promiscuous mutants generate deoxysphingolipids that are thought to build up over time to a toxic level that damages tissues in HSAN1 patients. Until the structure of the human SPT is available it will be difficult to answer why five different single point mutations all result in a HSAN1 phenotype. Although study of membrane proteins is still challenging, methods and techniques in this area are improving. Future studies could first be directed towards preparing an improved bacterial SPT mimic. Successive mutations on the 'fusion' SPT presented here may result in a soluble SPT mimic that more closely resembles the heterodimeric human homologue. Excitingly, Teresa Dunn and colleagues have constructed a clone which enables them to express a recombinant, soluble and active "triple" fusion (lcb2/ssSPTa/lcb1) of human SPT.

Whilst great progress has been made on delineating the genes, enzymes and intermediates involved in sphingolipid biosynthesis in plants, yeast, fungi and mammals our knowledge of bacterial synthesis has surprisingly lagged behind. Since *Shingomonas paucimobilis* is a pathogenic organism we used a safer strain whose complete genome sequence was recently completed by the US DoE. This work reports the structure and characterisation of holo-SPT from *Sphingomonas wittichii*. It also describes the isolation, characterisation and modification of an unusual ACP and proposes a link between fatty acid biosynthesis and sphingolipid biosynthesis. While it was difficult to prepare the C16 holo-ACP, this problem could possibly be overcome by mutation of an unusual tryptophan residue (W37) which is predicted to block the active site. A second alternative is to prepare the holo-ACP by chemical acylation. Only when the acyl-ACP is prepared will it be possible to show whether the acyl-ACP is a substrate for SPT. A kinetic analysis to compare the efficiency of acyl-ACPs with acyl-CoA substrates will be interesting to investigate whether the specificity trend remains the same.

REFERENCES

1. Thudichum, J. L. A Treatise on the Chemical Constitution of the Brain. (Bailliere, Tindall and Cox, London, 1884).
2. Carter, H. E., Haines, H. W., Ledyard, W. E. & Norris, W. P. Biochemistry of the sphingolipides. J. Biol. Chem. 169, 77-82 (1947).
3. Futerman, A. H. & Hannun, Y. A. The complex life of simple sphingolipids. EMBO reports 5, 777-782 (2004).
4. Hannun, Y. A. & Obeid, L. M. Principles of bioactive lipid signalling: lessons from sphingolipids. Nat Rev Mol Cell Biol 9, 139-50 (2008).
5. Simons, K. & Ikonen, E. Functionnal rafts in cell membranes. Nature 387, 569-572 (1997).
6. Merrill, A. H. *De novo* Sphingolipid Biosynthesis: A Necessary, but dangerous, Pathway. J. Biol. Chem. 277, 25843-25846 (2002).
7. Houlden, H. et al. Clinical, pathological and genetic characterization of hereditary sensory and autonomic neuropathy type 1 (HSAN I). Brain 129, 411-25 (2006).
8. Dickson, R. C., Sumanasekera, C. & Lester, R. L. Functions and metabolism of sphingolipids in Saccharomyces cerevisiae. Prog Lipid Res 45, 447-65 (2006).
9. Lynch, D. V. & Dunn, T. M. An introduction to plant sphingolipids and a review of recent advances in understanding their metabolism and function. New Phytologist, 677-702 (2004).
10. White, D. C., Sutton, S. D. & Ringelberg, D. B. The genus Sphingomonas: physiology and ecology. Curr Opin Biotechnol 7, 301-6 (1996).
11. Ikushiro, H., Hayashi, H. & Kagamiyama, H. A water-soluble homodimeric serine palmitoyltransferase from *Sphingomonas paucimobilis* EY2395T strain. Purification, characterization, cloning, and overproduction. J Biol Chem 276, 18249-56 (2001).
12. Ikushiro, H., Islam, M. M., Tojo, H. & Hayashi, H. Molecular Characterization of Membrane-Associated Soluble Serine Palmitoyltransferases from *Sphingobacterium multivorum* and *Bdellovibrio stolpii*. J Bacteriol 189, 5749-61 (2007).
13. Kawahara, K. et al. Chemical structure of glycosphingolipids isolated from *Sphingomonas paucimobilis*. FEBS Lett. 292, 107-110 (1991).
14. Kawasaki, S. et al. The cell envelope structure of the lipopolysaccharide-lacking gram-negative bacterium Sphingomonas paucimobilis. J Bacteriol 176, 284-90 (1994).
15. Kawahara, K., Kuraishi, H. & Zahringer, U. Chemical structure and function of glycosphingolipids of Sphingomonas spp and their distribution among members of the alpha-4 subclass of Proteobacteria. J Ind Microbiol Biotechnol 23, 408-413 (1999).
16. Kawahara, K., Kubota, M., Sato, N., Tsuge, K. & Seto, Y. Occurrence of an alpha-galacturonosyl-ceramide in the dioxin-degrading bacterium Sphingomonas wittichii. FEMS Microbiol Lett 214, 289-294 (2002).

17. Kawahara, K., Sato, N., Tsuge, K. & Seto, Y. Confirmation of the anomeric structure of galacturonic acid in the galacturonosyl-ceramide of Sphingomonas yanoikuyae. Microbiol Immunol 50, 67-71 (2006).
18. Naka, T. et al. A novel sphingoglycolipid containing galacturonic acid and 2-hydroxy fatty acid in cellular lipids of Sphingomonas yanoikuyae. J Bacteriol 182, 2660-3 (2000).
19. Cellular and Biomolecular Recognition (Wiley-VCH, Weinheim, 2009).
20. Hanada, K. Serine palmitoyltransferase, a key enzyme of sphingolipid metabolism. Biochim Biophys Acta 1632, 16-30 (2003).
21. Eliot, A. C. & Kirsch, J. F. Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. . Annu Rev Biochem 73, 383-415 (2004).
22. John, R. A. Pyridoxal phosphate-dependent enzymes. Biochim Biophys Acta 1248, 81-96 (1995).
23. Jansonius, J. N. Structure, evolution and action of vitamin B6-dependent enzymes. Curr Opin Struct Biol 8, 759-69 (1998).
24. Schneider , G., Kack, H. & Lindqvist, Y. The manifold of vitamin B6 dependent enzymes. Struct. Fold. Des. 8, R1-R6 (2000).
25. Dunathan, H. C. Conformation and reaction specificity in pyridoxal phosphate enzymes. Proc Natl Acad Sci U S A 55, 712-716 (1966).
26. Toney, M. D. Computational studies on nonenzymatic and enzymatic pyridoxal phosphate catalysed decarboxylation of 2-aminoisobutyrate. Biochemistry 40, 1378-1384 (2001).
27. Toney, M. D. Reaction specificity in pyridoxal phosphate enzymes. Arch Biochem Biophys 433, 279-287 (2005).
28. Bach, R. D., Canepa, C. & Glukhovtsev, M. N. Influence of Electrostatic Effects on Activation Barriers in Enzymatic Reactions: Pyridoxal-5'-Phosphate-Dependent Decarboxylation of α-Amino Acids. J. Am. Chem. Soc. 121, 6542-6555 (1999).
29. Sprang, S. R., Madsen, N. B. & Withers, S. G. Multiple phosphate positions in the catalytic site of glycogen phosphorylase: structure of the pyridoxal-5'-pyrophosphate coenzyme-substrate analog. Protein Sci 1, 1100-11 (1992).
30. Alexander, F. W., Sandmeier, E., Mehta, P. K. & Christen, P. Evolutionary relationships among pyridoxal-5'-phosphate-dependent enzymes. Regio-specific alpha, beta and gamma families. Eur J Biochem 219, 953-60 (1994).
31. Grishin, N. V., Phillips, M. A. & Goldsmith, E. J. Modeling of the spatial structure of eukaryotic ornithine decarboxylases. Protein Sci 4, 1291-304 (1995).
32. Alexeev, D. et al. The crystal structure of 8-amino-7-oxononanoate synthase: a bacterial PLP-dependent, acyl-CoA-condensing enzyme. J Mol Biol 284, 401-19 (1998).
33. Webster, S. P. et al. Mechanism of 8-amino-7-oxononanoate synthase: spectroscopic, kinetic, and crystallographic studies. Biochemistry 39, 516-28 (2000).
34. Jordan, P. M. & Shemin, D. in The Enzymes (ed. Boyer, P. D.) 339-356 (Academic Press, London, New York, 1972).
35. Ferreira, G. C. & Gong, J. 5-Aminolevulinate synthase and the first step of heme biosynthesis. J Bioenerg Biomembr 27, 151-9 (1995).

36. Astner, A. et al. Crystal structure of 5-aminolevulinate synthase, the first enzyme of heme biosynthesis, and its link to XLSA in humans. Embo J 24, 3166-3177 (2005).
37. Schmidt, A. et al. Three-dimensional structure of 2-amino-3-ketobutyrate CoA ligase from *Escherichia coli* complexed with a PLP-substrate intermediate: inferred reaction mechanism. Biochemistry 40, 5151-5160 (2001).
38. Hanada, K. et al. Mammalian cell mutants resistant to a sphingomyelin-directed cytolysin. Genetic and biochemical evidence for complex formation of the LCB1 protein with the LCB2 protein for serine palmitoyltransferase. J Biol Chem 273, 33787-94 (1998).
39. Yard, B. A. et al. The structure of serine palmitoyltransferase; gateway to sphingolipid biosynthesis. J Mol Biol 370, 870-86 (2007).
40. Hunter, G. A. & Ferreira, G. C. Lysine-313 of 5-aminolevulinate synthase acts as a general base during formation of the quinonoid reaction intermediates. Biochemistry 38, 3711-8 (1999).
41. Zaman, Z., Jordan, P. M. & Akhtar, M. Mechanism and stereochemistry of the 5-aminolaevulinate synthetase reaction. Biochem. J. 135, 257-263 (1973).
42. Alexeev, D. et al. Suicide inhibition of alpha-oxamine synthases: structures of the covalent adducts of 8-amino-7-oxononanoate synthase with trifluoroalanine. Org. Biomol. Chem. 4, 1209-1212 (2006).
43. Kerbarh, O., Campopiano, D. J. & Baxter, R. L. Mechanism of alpha-oxoamine synthases: identification of the intermediate Claisen product in the 8-amino-7-oxononanoate synthase reaction. Chem. Commun., 60-62 (2006).
44. Raman, M. C. et al. The external aldimine form of serine palmitoyltransferase: structural, kinetic, and spectroscopic analysis of the wild-type enzyme and HSAN1 mutant mimics. J Biol Chem 284, 17328-39 (2009).
45. Jahan, N. et al. Insights into the biosynthesis of the Vibrio cholerae major autoinducer CAI-1 from the crystal structure of the PLP-dependent enzyme CqsA. J Mol Biol 392, 763-73 (2009).
46. Kelly, R. C. et al. The Vibrio cholerae quorum-sensing autoinducer CAI-1: analysis of the biosynthetic enzyme CqsA. Nat Chem Biol 5, 891-5 (2009).
47. Spirig, T. et al. The Legionella autoinducer synthase LqsA produces an alpha-hydroxyketone signaling molecule. J Biol Chem 283, 18113-23 (2008).
48. Gerber, R., Lou, L. & Du, L. A PLP-Dependent Polyketide Chain Releasing Mechanism in the Biosynthesis of Mycotoxin Fumonisins in Fusarium verticillioides. J Am Chem Soc (2009).
49. Garneau-Tsodikova, S., Dorrestein, P. C., Kelleher, N. L. & Walsh, C. T. Protein assembly line components in prodigiosin biosynthesis: characterization of PigA,G,H,I,J. J Am Chem Soc 128, 12600-1 (2006).
50. Pinto, W. J., Wells, G. W. & Lester, R. L. Characterization of enzymatic synthesis of sphingolipid long-chain bases in *Saccharomyces cerevisiae*: mutant strains exhibiting long-chain-base auxotrophy are deficient in serine palmitoyltransferase activity. J Bacteriol 174, 2575-81 (1992).
51. Pinto, W. J. et al. Sphingolipid long-chain-base auxotrophs of *Saccharomyces cerevisiae*: genetics, physiology, and a method for their selection. J Bacteriol 174, 2565-74 (1992).

52. Hanada, K. et al. A mammalian homolog of the yeast LCB1 encodes a component of serine palmitoyltransferase, the enzyme catalyzing the first step in sphingolipid synthesis. J Biol Chem 272, 32108-14 (1997).
53. Hanada, K. et al. Sphingolipids are essential for the growth of Chinese hamster ovary cells. Restoration of the growth of a mutant defective in sphingoid base biosynthesis by exogenous sphingolipids. J Biol Chem 267, 23527-33 (1992).
54. Weiss, B. & Stoffel, W. Human and murine serine-palmitoyl-CoA transferase--cloning, expression and characterization of the key enzyme in sphingolipid synthesis. Eur J Biochem 249, 239-47 (1997).
55. Hanada, K., Hara, T. & Nishijima, M. Purification of the serine palmitoyltransferase complex responsible for sphingoid base synthesis by using affinity peptide chromatography techniques. J Biol Chem 275, 8409-15 (2000).
56. Hornemann, T., Richard, S., Rutti, M. F., Wei, Y. & von Eckardstein, A. Mammalian serine-palmitoyltransferase - cloning and initial characterisation of a new subunit. J Biol Chem (2006).
57. Mandon, E. C., Ehses, I., Rother, J., van Echten, G. & Sandhoff, K. Subcellular localization and membrane topology of serine palmitoyltransferase, 3-dehydrosphinganine reductase, and sphinganine N-acyltransferase in mouse liver. J Biol Chem 267, 11144-8 (1992).
58. Han, G. et al. The topology of the Lcb1p subunit of yeast serine palmitoyltransferase. J Biol Chem 279, 53707-16 (2004).
59. Gable, K., Slife, H., Bacikova, D., Monaghan, E. & Dunn, T. M. Tsc3p is an 80-amino acid protein associated with serine palmitoyltransferase and required for optimal enzyme activity. J Biol Chem 275, 7597-603 (2000).
60. Monaghan, E., Gable, K. & Dunn, T. Mutations in the Lcb2p subunit of serine palmitoyltransferase eliminate the requirement for the TSC3 gene in *Saccharomyces cerevisiae*. Yeast 19, 659-70 (2002).
61. Han, G. et al. Identification of small subunits of mammalian serine palmitoyltransferase that confer distinct acyl-CoA substrate specificities. Proc Natl Acad Sci U S A 106, 8186-91 (2009).
62. Breslow, D. K. et al. Orm family proteins mediate sphingolipid homeostasis. Nature 463, 1048-53.
63. Han, S., Lone, M. A., Schneiter, R. & Chang, A. Orm1 and Orm2 are conserved endoplasmic reticulum membrane proteins regulating lipid homeostasis and protein quality control. Proc Natl Acad Sci U S A (2010).
64. Han, G. et al. Expression of a novel marine viral single-chain serine palmitoyltransferase and construction of yeast and mammalian single-chain chimera. J Biol Chem 281, 39935-42 (2006).
65. Dyck, P. J., Thomas, P. K., Griffin, J. W., Low, P. A. & Poduslo, J. F. Neuronal atrophy and degeneration predominantly affecting peripheral sensory and autonomic neurons. (W.B. Saunders Co, Philadelphia, 1993).
66. Denny-Brown, D. Hereditary sensory radicular neuropathy. J Neurol Neurosurg Psychiatry 14, 237-52 (1951).
67. Campbell, A. M. & Hoffman, H. L. Sensory Radicular Neuropathy Associated with Muscle Wasting in Two Cases. Brain 87, 67-74 (1964).

68. Wallace, D. C. Observations upon a predominantly sensory hereditary neuropathy. Proc Aust Assoc Neurol 3, 101-9 (1965).
69. Dawkins, J. L., Hulme, D. J., Brahmbhatt, S. B., Auer-Grumbach, M. & Nicholson, G. A. Mutations in SPTLC1, encoding serine palmitoyltransferase, long chain base subunit-1, cause hereditary sensory neuropathy type I. Nat Genet 27, 309-12 (2001).
70. Bejaoui, K. et al. Hereditary sensory neuropathy type 1 mutations confer dominant negative effects on serine palmitoyltransferase, critical for sphingolipid synthesis. J Clin Invest 110, 1301-8 (2002).
71. Verhoeven, K. et al. SPTLC1 mutation in twin sisters with hereditary sensory neuropathy type I. Neurology 62, 1001-2 (2004).
72. Nicholson, G. A. et al. The gene for hereditary sensory neuropathy type I (HSN-I) maps to chromosome 9q22.1-q22.3. Nat Genet 13, 101-4 (1996).
73. Hornemann, T. et al. A systematic comparison of all mutations in hereditary sensory neuropathy type I (HSAN I) reveals that the G387A mutation is not disease associated. Neurogenetics 10, 135-43 (2009).
74. Rotthier, A. et al. Genes for hereditary sensory and autonomic neuropathies: a genotype-phenotype correlation. Brain 132, 2699-711 (2009).
75. Gable, K. et al. Mutations in the yeast LCB1 and LCB2 genes, including those corresponding to the hereditary sensory neuropathy type I mutations, dominantly inactivate serine palmitoyltransferase. J Biol Chem 277, 10194-200 (2002).
76. Hornemann, T., Penno, A. & von Eckardstein, A. The accumulation of two atypical sphingolipids cause hereditary sensory neuropathy type 1. Chem. Phys. Lipids 154S, S63, PO111 (2008).
77. Eichler, F. S. et al. Overexpression of the wild-type SPT1 subunit lowers Desoxysphingolipid levels and rescues the phenotype of HSAN1. J Neurosci 29, 14646-51 (2009).
78. Penno, A. et al. Hereditary sensory neuropathy type 1 is caused by the accumulation of two neurotoxic sphingolipids. J Biol Chem.
79. Carpenter, E. P., Beis, K., Cameron, A. D. & Iwata, S. Overcoming the challenges of membrane protein crystallography. Curr Opin Struct Biol 18, 581-6 (2008).
80. Fischbach, M. A. & Walsh, C. T. Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. Chem Rev 106, 3468-96 (2006).
81. Byers, D. M. & Gong, H. Acyl carrier protein: structure-function relationships in a conserved multifunctional protein family. Biochem Cell Biol 85, 649-62 (2007).
82. Smith, S., Witkowski, A. & Joshi, A. K. Structural and functional organization of the animal fatty acid synthase. Prog Lipid Res 42, 289-317 (2003).
83. Maier, T., Jenni, S. & Ban, N. Architecture of mammalian fatty acid synthase at 4.5 A resolution. Science 311, 1258-62 (2006).
84. Jenni, S., Leibundgut, M., Maier, T. & Ban, N. Architecture of a fungal fatty acid synthase at 5 A resolution. Science 311, 1263-7 (2006).

85. Lomakin, I. B., Xiong, Y. & Steitz, T. A. The crystal structure of yeast fatty acid synthase, a cellular machine with eight active sites working together. Cell 129, 319-32 (2007).
86. Leibundgut, M., Jenni, S., Frick, C. & Ban, N. Structural basis for substrate delivery by acyl carrier protein in the yeast fatty acid synthase. Science 316, 288-90 (2007).
87. Schweizer, E. & Hofmann, J. Microbial type I fatty acid synthases (FAS): major players in a network of cellular FAS systems. Microbiol Mol Biol Rev 68, 501-17 (2004).
88. White, S. W., Zheng, J., Zhang, Y. M. & Rock. The structural biology of type II fatty acid biosynthesis. Annu Rev Biochem 74, 791-831 (2005).
89. Ohlrogge, J., Savage, L., Jaworski, J., Voelker, T. & Post-Beittenmiller, D. Alteration of acyl-acyl carrier protein pools and acetyl-CoA carboxylase expression in Escherichia coli by a plant medium chain acyl-acyl carrier protein thioesterase. Arch Biochem Biophys 317, 185-90 (1995).
90. Zhang, Y. M., Marrakchi, H., White, S. W. & Rock, C. O. The application of computational methods to explore the diversity and structure of bacterial fatty acid synthase. J Lipid Res 44, 1-10 (2003).
91. Zhang, Y. M. et al. Identification and analysis of the acyl carrier protein (ACP) docking site on beta-ketoacyl-ACP synthase III. J Biol Chem 276, 8231-8 (2001).
92. Zhang, Y. M., Wu, B., Zheng, J. & Rock, C. O. Key residues responsible for acyl carrier protein and beta-ketoacyl-acyl carrier protein reductase (FabG) interaction. J Biol Chem 278, 52935-43 (2003).
93. Showalter, A. D. et al. Differential conservation of transcriptional domains of mammalian Prophet of Pit-1 proteins revealed by structural studies of the bovine gene and comparative functional analysis of the protein. Gene 291, 211-21 (2002).
94. Roujeinikova, A. et al. X-ray crystallographic studies on butyryl-ACP reveal flexibility of the structure around a putative acyl chain binding site. Structure 10, 825-35 (2002).
95. Roujeinikova, A. et al. Structural studies of fatty acyl-(acyl carrier protein) thioesters reveal a hydrophobic binding cavity that can expand to fit longer substrates. J Mol Biol 365, 135-45 (2007).
96. Brody, S., Oh, C., Hoja, U. & Schweizer, E. Mitochondrial acyl carrier protein is involved in lipoic acid synthesis in Saccharomyces cerevisiae. FEBS Lett 408, 217-20 (1997).
97. Wong, H. C., Liu, G., Zhang, Y. M., Rock, C. O. & Zheng, J. The solution structure of acyl carrier protein from Mycobacterium tuberculosis. J Biol Chem 277, 15874-80 (2002).
98. Cane, D. E. & Walsh, C. T. The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases. Chem Biol 6, R319-25 (1999).
99. Lai, J. R., Koglin, A. & Walsh, C. T. Carrier protein structure and recognition in polyketide and nonribosomal peptide biosynthesis. Biochemistry 45, 14869-79 (2006).
100. Cane, D. E. Introduction: Polyketide and Nonribosomal Polypeptide Biosynthesis. From Collie to Coli. Chem Rev 97, 2463-2464 (1997).

101. Moore, B. S. & Hopke, J. N. Discovery of a new bacterial polyketide biosynthetic pathway. Chembiochem 2, 35-8 (2001).
102. Garneau, S., Dorrestein, P. C., Kelleher, N. L. & Walsh, C. T. Characterization of the formation of the pyrrole moiety during clorobiocin and coumermycin A1 biosynthesis. Biochemistry 44, 2770-80 (2005).
103. Lambalot, R. H. et al. A new enzyme superfamily - the phosphopantetheinyl transferases. Chem Biol 3, 923-36 (1996).
104. Reuter, K., Mofid, M. R., Marahiel, M. A. & Ficner, R. Crystal structure of the surfactin synthetase-activating enzyme sfp: a prototype of the 4'-phosphopantetheinyl transferase superfamily. Embo J 18, 6823-31 (1999).
105. Chirgadze, N. Y., Briggs, S. L., McAllister, K. A., Fischl, A. S. & Zhao, G. Crystal structure of Streptococcus pneumoniae acyl carrier protein synthase: an essential enzyme in bacterial fatty acid biosynthesis. Embo J 19, 5281-7 (2000).
106. Parris, K. D. et al. Crystal structures of substrate binding to Bacillus subtilis holo-(acyl carrier protein) synthase reveal a novel trimeric arrangement of molecules resulting in three active sites. Structure 8, 883-95 (2000).
107. Quadri, L. E. et al. Characterization of Sfp, a Bacillus subtilis phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. Biochemistry 37, 1585-95 (1998).
108. Sanchez, C., Du, L., Edwards, D. J., Toney, M. D. & Shen, B. Cloning and characterization of a phosphopantetheinyl transferase from Streptomyces verticillus ATCC15003, the producer of the hybrid peptide-polyketide antitumor drug bleomycin. Chem Biol 8, 725-38 (2001).
109. Cox, R. J. et al. *Streptomyces coelicolor* phosphopantetheinyl transferase: a promiscuous activator of polyketide and fatty acid synthase acyl carrier proteins. J. Chem. Soc., Perkin Trans. 1, 1644-1649 (2002).
110. Smith, P. K. et al. Measurement of protein using bicinchoninic acid. Anal Biochem 150, 76-85 (1985).
111. Kessler, R. J. & Fanestil, D. D. Interference by lipids in the determination of protein using bicinchoninic acid. Anal Biochem 159, 138-42 (1986).
112. LeDuc, R. D. et al. ProSight PTM: an integrated environment for protein identification and characterization by top-down mass spectrometry. Nucleic Acids Res 32, W340-5 (2004).
113. Ellman, G. L. Tissue sulfhydryl groups. Arch Biochem Biophys 82, 70-7 (1959).
114. Riddles, P. W., Blakeley, R. L. & Zerner, B. Ellman's reagent: 5,5'-dithiobis(2-nitrobenzoic acid)--a reexamination. Anal Biochem 94, 75-81 (1979).
115. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. J. Appl. Cryst. 26, 795-800 (1993).
116. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53, 240-55 (1997).
117. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallographica Section D-Biological Crystallography. 60, 2126-2132 (2004).

118. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 50, 760-3 (1994).
119. Ikushiro, H., Hayashi, H. & Kagamiyama, H. Reactions of serine palmitoyltransferase with serine and molecular mechanisms of the actions of serine derivatives as inhibitors. Biochemistry 43, 1082-92 (2004).
120. Terzuoli, L. et al. Some chemical properties and biological role of thiazolidine compounds. Life Sci 63, 1251-67 (1998).
121. Ikushiro, H., Fujii, S., Shiraiwa, Y. & Hayashi, H. Acceleration of the substrate Calpha deprotonation by an analogue of the second substrate palmitoyl-CoA in Serine Palmitoyltransferase. J Biol Chem 283, 7542-53 (2008).
122. Zhang, J., Cheltsov, A. V. & Ferreira, G. C. Conversion of 5-aminolevulinate synthase into a more active enzyme by linking the two subunits: spectroscopic and kinetic properties. Protein Sci 14, 1190-200 (2005).
123. Ikushiro, H., Hayashi, H. & Kagamiyama, H. Bacterial serine palmitoyltransferase: a water-soluble homodimeric prototype of the eukaryotic enzyme. Biochim Biophys Acta 1647, 116-20 (2003).
124. Ikushiro, H. et al. Structural insights into the enzymatic mechanism of serine palmitoyltransferase from Sphingobacterium multivorum. J Biochem 146, 549-62 (2009).
125. Gable, K. et al. A disease-causing mutation in the active site of serine palmitoyltransferase causes catalytic promiscuity. J Biol Chem 285, 22846-52.
126. Bejaoui, K. et al. SPTLC1 is mutated in hereditary sensory neuropathy, type 1. Nat Genet 27, 261-2 (2001).
127. McCampbell, A. et al. Mutant SPTLC1 dominantly inhibits serine palmitoyltransferase activity in vivo and confers an age-dependent neuropathy. Human Molecular Genetics 14, 3507-3521 (2005).
128. Wittich, R. M., Wilkes, H., Sinnwell, V., Francke, W. & Fortnagel, P. Metabolism of dibenzo-p-dioxin by Sphingomonas sp. strain RW1. Appl Environ Microbiol 58, 1005-10 (1992).
129. Yabuuchi, E. et al. Proposal of Sphingomonas wittichii sp. nov. for strain RW1T, known as a dibenzo-p-dioxin metabolizer. Int J Syst Evol Microbiol 51, 281-92 (2001).
130. Tropf, S., Revill, W. P., Bibb, M. J., Hopwood, D. A. & Schweizer, M. Heterologously expressed acyl carrier protein domain of rat fatty acid synthase functions in Escherichia coli fatty acid synthase and Streptomyces coelicolor polyketide synthase systems. Chem Biol 5, 135-46 (1998).
131. Ploskon, E. et al. A mammalian type I fatty acid synthase acyl carrier protein domain does not sequester acyl chains. J Biol Chem 283, 518-28 (2008).
132. Yin, J. et al. Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. Proc Natl Acad Sci U S A 102, 15815-20 (2005).
133. Bates, P. A. & Sternberg, M. J. Model building by comparison at CASP3: using expert knowledge and computer automation. Proteins Suppl 3, 47-54 (1999).
134. Bates, P. A., Kelley, L. A., MacCallum, R. M. & Sternberg, M. J. Enhancement of protein modeling by human intervention in applying the

automatic programs 3D-JIGSAW and 3D-PSSM. Proteins Suppl 5, 39-46 (2001).
135. Contreras-Moreira, B. & Bates, P. A. Domain fishing: a first step in protein comparative modelling. Bioinformatics 18, 1141-2 (2002).
136. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18, 2714-23 (1997).
137. Schwede, T., Kopp, J., Guex, N. & Peitsch, M. C. SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Res 31, 3381-5 (2003).
138. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22, 195-201 (2006).
139. Arthur, C. J. et al. Structure and malonyl CoA-ACP transacylase binding of streptomyces coelicolor fatty acid synthase acyl carrier protein. ACS Chem Biol 4, 625-36 (2009).
140. Cronan, J. E., Jr. & Klages, A. L. Chemical synthesis of acyl thioesters of acyl carrier protein with native structure. Proc Natl Acad Sci U S A 78, 5440-4 (1981).
141. Leduc, R. D. & Kelleher, N. L. Using ProSight PTM and related tools for targeted protein identification and characterization with high mass accuracy tandem MS data. Curr Protoc Bioinformatics Chapter 13, Unit 13 6 (2007).

APPENDICES

GENE AND PROTEIN SEQUENCES

Sphingomonas paucimobilis

Gene sequence pET28a/SPT WT

```
   1 ATGACCGAAG CCGCCGCTCA GCCCCACGCC CTCCCCGCCG ACGCGCCCGA CATCGCGCCG
  61 GAACGCGACC TGCTCTCCAA GTTCGACGGC CTGATCGCCG AGCGGCAGAA GCTGCTCGAC
 121 TCCGGCGTCA CCGATCCCTT CGCGATCGTG ATGGAACAGG TGAAGTCGCC GACCGAGGCC
 181 GTGATCCGTG GCAAGGACAC GATCCTGCTC GGCACGTACA ACTATATGGG CATGACCTTC
 241 GATCCGGACG TGATCGCAGC GGGCAAGGAA GCGCTGGAGA AGTTCGGGTC GGGCACCAAT
 301 GGCAGCCGGA TGCTCAACGG CACCTTCCAC GACCATATGG AAGTCGAACA GGCGCTGCGC
 361 GATTTCTACG GCACGACCGG CGCGATCGTC TTTTCGACCG GTTACATGGC CAATCTCGGC
 421 ATCATCTCAA CGCTGGCGGG CAAGGGTGAG TATGTCATCC TCGACGCCGA CAGCCATGCG
 481 TCGATCTATG ACGGCTGCCA GCAGGGCAAT GCCGAGATCG TCCGCTTCCG CCACAATTCG
 541 GTCGAGGATC TCGACAAGCG GCTGGGCCGT CTGCCCAAGG AACCTGCCAA GCTGGTCGTG
 601 CTGGAGGGCG TCTATTCGAT GCTCGGCGAC ATCGCTCCGC TGAAGGAGAT GGTCGCGGTC
 661 GCCAAGAAGC ATGGCGCAAT GGTCTTGGTC GACGAAGCGC ATTCGATGGG CTTTTTCGGC
 721 CCCAACGGGC GCGGCGTGTA CGAGGCGCAA GGGTTGGAAG GCCAGATCGA TTTCGTCGTC
 781 GGCACCTTCT CCAAATCGGT CGGCACAGTC GGCGGCTTCG TCGTGTCCAA TCATCCGAAG
 841 TTCGAGGCGG TCCGCCTCGC CTGCCGTCCG TACATCTTCA CCGCCTCGCT GCCGCCCTCG
 901 GTGGTAGCGA CCGCGACGAC GTCGATCCGC AAGCTGATGA CCGCGCATGA AAAGCGTGAG
 961 CGGCTGTGGT CGAATGCCCG CGCGTTGCAT GGCGGGCTGA AGGCGATGGG CTTCAGGCTC
1021 GGCACCGAGA CCTGCGACAG CGCGATCGTC GCGGTCATGC TGGAGGATCA GGAACAGGCC
1081 GCGATGATGT GGCAGGCGCT GCTCGACGGC GGGCTCTACG TCAACATGGC GCGCCCGCCC
1141 GCGACCCCGG CCGGCACCTT CCTGCTGCGC TGCTCCATCT GTGCCGAGCA CACGCCGGCG
1201 CAGATCCAGA CCGTGCTGGG CATGTTCCAG GCCGCGGGCC GCGCGGTCGG CGTCATCGGC
1261 CTCGAGCACC ACCACCACCA CCACTGA
```

SP SPT wild-type protein sequence (translated from the sequence above)

```
   1    MTEAAAQPHA LPADAPDIAP ERDLLSKFDG LIAERQKLLD SGVTDPFAIV
  51    MEQVKSPTEA VIRGKDTILL GTYNYMGMTF DPDVIAAGKE ALEKFGSGTN
 101    GSRMLNGTFH DHMEVEQALR DFYGTTGAIV FSTGYMANLG IISTLAGKGE
 151    YVILDADSHA SIYDGCQQGN AEIVRFRHNS VEDLDKRLGR LPKEPAKLVV
 201    LEGVYSMLGD IAPLKEMVAV AKKHGAMVLV DEAHSMGFFG PNGRGVYEAQ
 251    GLEGQIDFVV GTFSKSVGTV GGFVVSNHPK FEAVRLACRP YIFTASLPPS
 301    VVATATTSIR KLMTAHEKRE RLWSNARALH GGLKAMGFRL GTETCDSAIV
 351    AVMLEDQEQA AMMWQALLDG GLYVNMARPP ATPAGTFLLR CSICAEHTPA
 401    QIQTVLGMFQ AAGRAVGVIG LEHHHHHH
```

Gene sequence pET28a/SPT N100W

```
   1ATGACCGAAG CCGCCGCTCA GCCCCACGCC CTCCCCGCCG ACGCGCCCGA CATCGCGCCG
  61 GAACGCGACC TGCTCTCCAA GTTCGACGGC CTGATCGCCG AGCGGCAGAA GCTGCTCGAC
 121 TCCGGCGTCA CCGATCCCTT CGCGATCGTG ATGGAACAGG TGAAGTCGCC GACCGAGGCC
 181 GTGATCCGTG GCAAGGACAC GATCCTGCTC GGCACGTACA ACTATATGGG CATGACCTTC
 241 GATCCGGACG TGATCGCAGC GGGCAAGGAA GCGCTGGAGA AGTTCGGGTC GGGCACCTGG
 301 GGCAGCCGGA TGCTCAACGG CACCTTCCAC GACCATATGG AAGTCGAACA GGCGCTGCGC
 361 GATTTCTACG GCACGACCGG CGCGATCGTC TTTTCGACCG GTTACATGGC CAATCTCGGC
 421 ATCATCTCAA CGCTGGCGGG CAAGGGTGAG TATGTCATCC TCGACGCCGA CAGCCATGCG
 481 TCGATCTATG ACGGCTGCCA GCAGGGCAAT GCCGAGATCG TCCGCTTCCG CCACAATTCG
 541 GTCGAGGATC TCGACAAGCG GCTGGGCCGT CTGCCCAAGG AACCTGCCAA GCTGGTCGTG
 601 CTGGAGGGCG TCTATTCGAT GCTCGGCGAC ATCGCTCCGC TGAAGGAGAT GGTCGCGGTC
 661 GCCAAGAAGC ATGGCGCAAT GGTCTTGGTC GACGAAGCGC ATTCGATGGG CTTTTTCGGC
 721 CCCAACGGGC GCGGCGTGTA CGAGGCGCAA GGGTTGGAAG GCCAGATCGA TTTCGTCGTC
 781 GGCACCTTCT CCAAATCGGT CGGCACAGTC GGCGGCTTCG TCGTGTCCAA TCATCCGAAG
 841 TTCGAGGCGG TCCGCCTCGC CTGCCGTCCG TACATCTTCA CCGCCTCGCT GCCGCCCTCG
 901 GTGGTAGCGA CCGCGACGAC GTCGATCCGC AAGCTGATGA CCGCGCATGA AAAGCGTGAG
 961 CGGCTGTGGT CGAATGCCCG CGCGTTGCAT GGCGGGCTGA AGGCGATGGG CTTCAGGCTC
1021 GGCACCGAGA CCTGCGACAG CGCGATCGTC GCGGTCATGC TGGAGGATCA GGAACAGGCC
1081 GCGATGATGT GGCAGGCGCT GCTCGACGGC GGGCTCTACG TCAACATGGC GCGCCCGCCC
1141 GCGACCCCGG CCGGCACCTT CCTGCTGCGC TGCTCCATCT GTGCCGAGCA CACGCCGGCG
1201 CAGATCCAGA CCGTGCTGGG CATGTTCCAG GCCGCGGGCC GCGCGGTCGG CGTCATCGGC
1261 CTCGAGCACC ACCACCACCA CCACTGA
```

SP SPT N100W protein sequence (translated from the sequence above)

```
   1    MTEAAAQPHA LPADAPDIAP ERDLLSKFDG LIAERQKLLD SGVTDPFAIV
  51    MEQVKSPTEA VIRGKDTILL GTYNYMGMTF DPDVIAAGKE ALEKFGSGTW
 101    GSRMLNGTFH DHMEVEQALR DFYGTTGAIV FSTGYMANLG IISTLAGKGE
 151    YVILDADSHA SIYDGCQQGN AEIVRFRHNS VEDLDKRLGR LPKEPAKLVV
 201    LEGVYSMLGD IAPLKEMVAV AKKHGAMVLV DEAHSMGFFG PNGRGVYEAQ
 251    GLEGQIDFVV GTFSKSVGTV GGFVVSNHPK FEAVRLACRP YIFTASLPPS
 301    VVATATTSIR KLMTAHEKRE RLWSNARALH GGLKAMGFRL GTETCDSAIV
 351    AVMLEDQEQA AMMWQALLDG GLYVNMARPP ATPAGTFLLR CSICAEHTPA
 401    QIQTVLGMFQ AAGRAVGVIG LEHHHHHH
```

Gene sequence pET28a/SPT N100Y

```
   1 ATGACCGAAG CCGCCGCTCA GCCCCACGCC CTCCCCGCCG ACGCGCCCGA CATCGCGCCG
  61 GAACGCGACC TGCTCTCCAA GTTCGACGGC CTGATCGCCG AGCGGCAGAA GCTGCTCGAC
 121 TCCGGCGTCA CCGATCCCTT CGCGATCGTG ATGGAACAGG TGAAGTCGCC GACCGAGGCC
 181 GTGATCCGTG GCAAGGACAC GATCCTGCTC GGCACGTACA ACTATATGGG CATGACCTTC
 241 GATCCGGACG TGATCGCAGC GGGCAAGGAA GCGCTGGAGA AGTTCGGGTC GGGCACCTAT
 301 GGCAGCCGGA TGCTCAACGG CACCTTCCAC GACCATATGG AAGTCGAACA GGCGCTGCGC
 361 GATTTCTACG GCACGACCGG CGCGATCGTC TTTTCGACCG GTTACATGGC CAATCTCGGC
 421 ATCATCTCAA CGCTGGCGGG CAAGGGTGAG TATGTCATCC TCGACGCCGA CAGCCATGCG
 481 TCGATCTATG ACGGCTGCCA GCAGGGCAAT GCCGAGATCG TCCGCTTCCG CCACAATTCG
 541 GTCGAGGATC TCGACAAGCG GCTGGGCCGT CTGCCCAAGG AACCTGCCAA GCTGGTCGTG
 601 CTGGAGGGCG TCTATTCGAT GCTCGGCGAC ATCGCTCCGC TGAAGGAGAT GGTCGCGGTC
 661 GCCAAGAAGC ATGGCGCAAT GGTCTTGGTC GACGAAGCGC ATTCGATGGG CTTTTTCGGC
 721 CCCAACGGGC GCGGCGTGTA CGAGGCGCAA GGGTTGGAAG GCCAGATCGA TTTCGTCGTC
 781 GGCACCTTCT CCAAATCGGT CGGCACAGTC GGCGGCTTCG TCGTGTCCAA TCATCCGAAG
 841 TTCGAGGCGG TCCGCCTCGC CTGCCGTCCG TACATCTTCA CCGCCTCGCT GCCGCCCTCG
 901 GTGGTAGCGA CCGCGACGAC GTCGATCCGC AAGCTGATGA CCGCGCATGA AAAGCGTGAG
 961 CGGCTGTGGT CGAATGCCCG CGCGTTGCAT GGCGGGCTGA AGGCGATGGG CTTCAGGCTC
1021 GGCACCGAGA CCTGCGACAG CGCGATCGTC GCGGTCATGC TGGAGGATCA GGAACAGGCC
1081 GCGATGATGT GGCAGGCGCT GCTCGACGGC GGGCTCTACG TCAACATGGC GCGCCCGCCC
1141 GCGACCCCGG CCGGCACCTT CCTGCTGCGC TGCTCCATCT GTGCCGAGCA CACGCCGGCG
1201 CAGATCCAGA CCGTGCTGGG CATGTTCCAG GCCGCGGGCC GCGCGGTCGG CGTCATCGGC
1261 CTCGAGCACC ACCACCACCA CCACTGA
```

SP SPT N100Y protein sequence (translated from the sequence above)

```
   1    MTEAAAQPHA LPADAPDIAP ERDLLSKFDG LIAERQKLLD SGVTDPFAIV
  51    MEQVKSPTEA VIRGKDTILL GTYNYMGMTF DPDVIAAGKE ALEKFGSGTY
 101    GSRMLNGTFH DHMEVEQALR DFYGTTGAIV FSTGYMANLG IISTLAGKGE
 151    YVILDADSHA SIYDGCQQGN AEIVRFRHNS VEDLDKRLGR LPKEPAKLVV
 201    LEGVYSMLGD IAPLKEMVAV AKKHGAMVLV DEAHSMGFFG PNGRGVYEAQ
 251    GLEGQIDFVV GTFSKSVGTV GGFVVSNHPK FEAVRLACRP YIFTASLPPS
 301    VVATATTSIR KLMTAHEKRE RLWSNARALH GGLKAMGFRL GTETCDSAIV
 351    AVMLEDQEQA AMMWQALLDG GLYVNMARPP ATPAGTFLLR CSICAEHTPA
 401    QIQTVLGMFQ AAGRAVGVIG LEHHHHHH
```

Gene sequence pET28a/SPT R378A

```
   1 ATGACCGAAG CCGCCGCTCA GCCCCACGCC CTCCCCGCCG ACGCGCCCGA CATCGCGCCG
  61 GAACGCGACC TGCTCTCCAA GTTCGACGGC CTGATCGCCG AGCGGCAGAA GCTGCTCGAC
 121 TCCGGCGTCA CCGATCCCTT CGCGATCGTG ATGGAACAGG TGAAGTCGCC GACCGAGGCC
 181 GTGATCCGTG GCAAGGACAC GATCCTGCTC GGCACGTACA ACTATATGGG CATGACCTTC
 241 GATCCGGACG TGATCGCAGC GGGCAAGGAA GCGCTGGAGA AGTTCGGGTC GGGCACCAAT
 301 GGCAGCCGGA TGCTCAACGG CACCTTCCAC GACCATATGG AAGTCGAACA GGCGCTGCGC
 361 GATTTCTACG GCACGACCGG CGCGATCGTC TTTTCGACCG GTTACATGGC CAATCTCGGC
 421 ATCATCTCAA CGCTGGCGGG CAAGGGTGAG TATGTCATCC TCGACGCCGA CAGCCATGCG
 481 TCGATCTATG ACGGCTGCCA GCAGGGCAAT GCCGAGATCG TCCGCTTCCG CCACAATTCG
 541 GTCGAGGATC TCGACAAGCG GCTGGGCCGT CTGCCCAAGG AACCTGCCAA GCTGGTCGTG
 601 CTGGAGGGCG TCTATTCGAT GCTCGGCGAC ATCGCTCCGC TGAAGGAGAT GGTCGCGGTC
 661 GCCAAGAAGC ATGGCGCAAT GGTCTTGGTC GACGAAGCGC ATTCGATGGG CTTTTTCGGC
 721 CCCAACGGGC GCGGCGTGTA CGAGGCGCAA GGGTTGGAAG GCCAGATCGA TTTCGTCGTC
 781 GGCACCTTCT CCAAATCGGT CGGCACAGTC GGCGGCTTCG TCGTGTCCAA TCATCCGAAG
 841 TTCGAGGCGG TCCGCCTCGC CTGCCGTCCG TACATCTTCA CCGCCTCGCT GCCGCCCTCG
 901 GTGGTAGCGA CCGCGACGAC GTCGATCCGC AAGCTGATGA CCGCGCATGA AAAGCGTGAG
 961 CGGCTGTGGT CGAATGCCCG CGCGTTGCAT GGCGGGCTGA AGGCGATGGG CTTCAGGCTC
1021 GGCACCGAGA CCTGCGACAG CGCGATCGTC GCGGTCATGC TGGAGGATCA GGAACAGGCC
1081 GCGATGATGT GGCAGGCGCT GCTCGACGGC GGGCTCTACG TCAACATGGC GGCCCCGCCC
1141 GCGACCCCGG CCGGCACCTT CCTGCTGCGC TGCTCCATCT GTGCCGAGCA CACGCCGGCG
1201 CAGATCCAGA CCGTGCTGGG CATGTTCCAG GCCGCGGGCC GCGCGGTCGG CGTCATCGGC
1261 CTCGAGCACC ACCACCACCA CCACTGA
```

SP SPT R378A protein sequence (translated from the sequence above)

```
   1    MTEAAAQPHA LPADAPDIAP ERDLLSKFDG LIAERQKLLD SGVTDPFAIV
  51    MEQVKSPTEA VIRGKDTILL GTYNYMGMTF DPDVIAAGKE ALEKFGSGTW
 101    GSRMLNGTFH DHMEVEQALR DFYGTTGAIV FSTGYMANLG IISTLAGKGE
 151    YVILDADSHA SIYDGCQQGN AEIVRFRHNS VEDLDKRLGR LPKEPAKLVV
 201    LEGVYSMLGD IAPLKEMVAV AKKHGAMVLV DEAHSMGFFG PNGRGVYEAQ
 251    GLEGQIDFVV GTFSKSVGTV GGFVVSNHPK FEAVRLACRP YIFTASLPPS
 301    VVATATTSIR KLMTAHEKRE RLWSNARALH GGLKAMGFRL GTETCDSAIV
 351    AVMLEDQEQA AMMWQALLDG GLYVNMAAPP ATPAGTFLLR CSICAEHTPA
 401    QIQTVLGMFQ AAGRAVGVIG LEHHHHHH
```

Gene sequence pET28a/SPT R378N

```
   1 ATGACCGAAG CCGCCGCTCA GCCCCACGCC CTCCCCGCCG ACGCGCCCGA CATCGCGCCG
  61 GAACGCGACC TGCTCTCCAA GTTCGACGGC CTGATCGCCG AGCGGCAGAA GCTGCTCGAC
 121 TCCGGCGTCA CCGATCCCTT CGCGATCGTG ATGGAACAGG TGAAGTCGCC GACCGAGGCC
 181 GTGATCCGTG GCAAGGACAC GATCCTGCTC GGCACGTACA ACTATATGGG CATGACCTTC
 241 GATCCGGACG TGATCGCAGC GGGCAAGGAA GCGCTGGAGA AGTTCGGGTC GGGCACCAAT
 301 GGCAGCCGGA TGCTCAACGG CACCTTCCAC GACCATATGG AAGTCGAACA GGCGCTGCGC
 361 GATTTCTACG GCACGACCGG CGCGATCGTC TTTTCGACCG GTTACATGGC CAATCTCGGC
 421 ATCATCTCAA CGCTGGCGGG CAAGGGTGAG TATGTCATCC TCGACGCCGA CAGCCATGCG
 481 TCGATCTATG ACGGCTGCCA GCAGGGCAAT GCCGAGATCG TCCGCTTCCG CCACAATTCG
 541 GTCGAGGATC TCGACAAGCG GCTGGGCCGT CTGCCCAAGG AACCTGCCAA GCTGGTCGTG
 601 CTGGAGGGCG TCTATTCGAT GCTCGGCGAC ATCGCTCCGC TGAAGGAGAT GGTCGCGGTC
 661 GCCAAGAAGC ATGGCGCAAT GGTCTTGGTC GACGAAGCGC ATTCGATGGG CTTTTTCGGC
 721 CCCAACGGGC GCGGCGTGTA CGAGGCGCAA GGGTTGGAAG GCCAGATCGA TTTCGTCGTC
 781 GGCACCTTCT CCAAATCGGT CGGCACAGTC GGCGGCTTCG TCGTGTCCAA TCATCCGAAG
 841 TTCGAGGCGG TCCGCCTCGC CTGCCGTCCG TACATCTTCA CCGCCTCGCT GCCGCCCTCG
 901 GTGGTAGCGA CCGCGACGAC GTCGATCCGC AAGCTGATGA CCGCGCATGA AAAGCGTGAG
 961 CGGCTGTGGT CGAATGCCCG CGCGTTGCAT GGCGGGCTGA AGGCGATGGG CTTCAGGCTC
1021 GGCACCGAGA CCTGCGACAG CGCGATCGTC GCGGTCATGC TGGAGGATCA GGAACAGGCC
1081 GCGATGATGT GGCAGGCGCT GCTCGACGGC GGGCTCTACG TCAACATGGC GAACCCGCCC
1141 GCGACCCCGG CCGGCACCTT CCTGCTGCGC TGCTCCATCT GTGCCGAGCA CACGCCGGCG
1201 CAGATCCAGA CCGTGCTGGG CATGTTCCAG GCCGCGGGCC GCGCGGTCGG CGTCATCGGC
1261 CTCGAGCACC ACCACCACCA CCACTGA
```

SP SPT R378N protein sequence (translated from the sequence above)

```
   1    MTEAAAQPHA LPADAPDIAP ERDLLSKFDG LIAERQKLLD SGVTDPFAIV
  51    MEQVKSPTEA VIRGKDTILL GTYNYMGMTF DPDVIAAGKE ALEKFGSGTW
 101    GSRMLNGTFH DHMEVEQALR DFYGTTGAIV FSTGYMANLG IISTLAGKGE
 151    YVILDADSHA SIYDGCQQGN AEIVRFRHNS VEDLDKRLGR LPKEPAKLVV
 201    LEGVYSMLGD IAPLKEMVAV AKKHGAMVLV DEAHSMGFFG PNGRGVYEAQ
 251    GLEGQIDFVV GTFSKSVGTV GGFVVSNHPK FEAVRLACRP YIFTASLPPS
 301    VVATATTSIR KLMTAHEKRE RLWSNARALH GGLKAMGFRL GTETCDSAIV
 351    AVMLEDQEQA AMMWQALLDG GLYVNMANPP ATPAGTFLLR CSICAEHTPA
 401    QIQTVLGMFQ AAGRAVGVIG LEHHHHHH
```

Sphingomonas wittichii

Gene sequence pET28a/ACP-SPT SW

```
   1 ATGAGCAGCA GAGAAGACAT CTTCACCCGC GTCGCCGAGC AGATCGAGCC GTTCAACAAG
  61 AAGGGCATCG ACCTGGCTGA GACGACGAGC TTCGCCGGCG ACCTCGAATG GGACAGCCTG
 121 ACGGTCATGG ATTTCGTCGC GGCGGTGGAG GATGAGTTCG ACATCATCAT CACGATGAAC
 181 ATGCAGGCCG AGATCGAGAC CGTCGGCCAG CTCGTCGATG CCGTCGCCAA GCTGCGGGGC
 241 TGACCCGGTG GCCGACCTCC TCTCCAAGTT CGACCCGCTG ATCGCCGAGC GCGAGGCGCT
 301 GCTCGCCACC GGCGTGCGCG ATCCCTACGC GATCGTCATG GACAAGGTGC TGTCGCCGAC
 361 CGAGGCGATG ATCAACGGCA GGAAGACGAT CCTGCTCGGC ACCTATAATT ACATGGGCAT
 421 GACCTTCGAC CCGGACGTCA TCGCCGCCGG CAAGCAGGCG CTCGACGAAT TCGGCTCGGG
 481 CACCACCGGC AGCCGGGTGC TCAACGGCAC CTATCAGGGT CACAAGGCGT GCGAGGACGC
 541 GCTCAAGGAA TTCTACGGGA CCGAGCACGC CATCGTCTTC TCGACCGGCT ACCAGGCCAA
 601 TCTCGGGATG ATCTCCACCC TGGCTGGCAA GGGCGACTAC ATCATCCTCG ACGCCGACAG
 661 CCACGCCTCG ATCTATGACG GCTGCTGGCT CGGCGACGCG GAGATCGTCC GCTTCCGCCA
 721 CAATTCGGTC GAGGACCTCG ACAAGCGGCT CGGCCGCCTG CCGGCCGAGG CGGGCAAGCT
 781 CGTCGTCCTC GAAGGCGTCT ATTCGATGAT GGGCGACATC GCCCCGCTCC AGGAGATGGT
 841 CGCCGTCTCG AAGAAGCACG GCGCGATGAT CCTCGTCGAC GAGGCGCACG GCATGGGCTT
 901 CTTCGGCGAG CATGGCCGCG GCGTGTTCGA GGAGGCCGGG GTGGAGGCCG ACGTCGACTT
 961 CGTCGTCGGC ACCTTCTCCA AGTCGGTCGG CACGGTCGGT GGCTTCTGCG TCTCCAACCA
1021 TCCGAAGTTC GAGGTGCTGC GCCTCGTCTG CCGGCCCTAT GTCTTCACCG CCTCGCTGCC
1081 GCCGTCGGTG GTCGCGACCG CCGCGACCTC GATCCGCAAG CTGATGCATG CGGGCGACAA
1141 GCGCGCGCAC CTCTGGAAGA ATTCGCGCCG CCTGCACCAG GGCCTGCGCG ACATGGGCTA
1201 CAAGCTCGGC ACCGAGACCG CCCAGTCGGC GATCATCGCG GTGATCCTGA CCGACATGGC
1261 CCAGGCGGTC GCGCTGTGGC AGGGCCTGCT CGAGGCGGGG CTCTATGTGA ACACCGCCCG
1321 TCCGCCCGCG ACGCCGGCCG GCATGTTCCT GCTGCGCTGC TCGCTGTGCG CCGAACATAG
1381 CGACGAACAG GTCGAGCAGA TCCTGGGCAT GTTCGAGAGC GCGGGGCGCG CCACGGGAGT
1441 AATCCCCAAG CTTGCGGCCG CACTCGAGCA CCACCACCAC CACCACTGA
```

SW ACP protein sequence (translated from the sequence above)

```
   1  MSSREDIFTR VAEQIEPFNK KGIDLAETTS FAGDLEWDSL TVMDFVAAVE
  51  DEFDIIITMN MQAEIETVGQ LVDAVAKLRG
```

SW SPT protein sequence (translated from the sequence above)

```
   1 VADLLSKFDP LIAEREALLA TGVRDPYAIV MDKVLSPTEA MINGRKTILL
  51 GTYNYMGMTF DPDVIAAGKQ ALDEFGSGTT GSRVLNGTYQ GHKACEDALK
 101 EFYGTEHAIV FSTGYQANLG MISTLAGKGD YIILDADSHA SIYDGCWLGD
 151 AEIVRFRHNS VEDLDKRLGR LPAEAGKLVV LEGVYSMMGD IAPLQEMVAV
 201 SKKHGAMILV DEAHGMGFFG EHGRGVFEEA GVEADVDFVV GTFSKSVGTV
 251 GGFCVSNHPK FEVLRLVCRP YVFTASLPPS VVATAATSIR KLMHAGDKRA
 301 HLWKNSRRLH QGLRDMGYKL GTETAQSAII AVILTDMAQA VALWQGLLEA
 351 GLYVNTARPP ATPAGMFLLR CSLCAEHSDE QVEQILGMFE SAGRATGVIP
 401 KLAAALEHHH HHH
```

Gene sequence pET28a/ SPT SW

```
   1 ATGGCCGACC TCCTCTCCAA GTTCGACCCG CTGATCGCCG AGCGCGAGGC GCTGCTCGCC
  61 ACCGGCGTGC GCGATCCCTA CGCGATCGTC ATGGACAAGG TGCTGTCGCC GACCGAGGCG
 121 ATGATCAACG GCAGGAAGAC GATCCTGCTC GGCACCTATA ATTACATGGG CATGACCTTC
 181 GACCCGGACG TCATCGCCGC CGGCAAGCAG GCGCTCGACG AATTCGGCTC GGGCACCACC
 241 GGCAGCCGGG TGCTCAACGG CACCTATCAG GGTCACAAGG CGTGCGAGGA CGCGCTCAAG
 301 GAATTCTACG GGACCGAGCA CGCCATCGTC TTCTCGACCG GCTACCAGGC CAATCTCGGG
 361 ATGATCTCCA CCCTGGCTGG CAAGGGCGAC TACATCATCC TCGACGCCGA CAGCCACGCC
 421 TCGATCTATG ACGGCTGCTG GCTCGGCGAC GCGGAGATCG TCCGCTTCCG CCACAATTCG
 481 GTCGAGGACC TCGACAAGCG GCTCGGCCGC CTGCCGGCCG AGGCGGGCAA GCTCGTCGTC
 541 CTCGAAGGCG TCTATTCGAT GATGGGCGAC ATCGCCCCGC TCCAGGAGAT GGTCGCCGTC
 601 TCGAAGAAGC ACGGCGCGAT GATCCTCGTC GACGAGGCGC ACGGCATGGG CTTCTTCGGC
 661 GAGCATGGCC GCGGCGTGTT CGAGGAGGCC GGGGTGGAGG CCGACGTCGA CTTCGTCGTC
 721 GGCACCTTCT CCAAGTCGGT CGGCACGGTC GGTGGCTTCT GCGTCTCCAA CCATCCGAAG
 781 TTCGAGGTGC TGCGCCTCGT CTGCCGGCCC TATGTCTTCA CCGCCTCGCT GCCGCCGTCG
 841 GTGGTCGCGA CCGCCGCGAC CTCGATCCGC AAGCTGATGC ATGCGGGCGA CAAGCGCGCG
 901 CACCTCTGGA AGAATTCGCG CCGCCTGCAC CAGGGCCTGC GCGACATGGG CTACAAGCTC
 961 GGCACCGAGA CCGCCCAGTC GGCGATCATC GCGGTGATCC TGACCGACAT GGCCCAGGCG
1021 GTCGCGCTGT GGCAGGGCCT GCTCGAGGCG GGGCTCTATG TGAACACCGC CCGTCCGCCC
1081 GCGACGCCGG CCGGCATGTT CCTGCTGCGC TGCTCGCTGT GCGCCGAACA TAGCGACGAA
1141 CAGGTCGAGC AGATCCTGGG CATGTTCGAG AGCGCGGGGC GCGCCACGGG AGTAATCCCC
1201 AAGCTTGCGG CCGCACTCGA GCACCACCAC CACCACCACT GA
```

SW SPT protein sequence (translated from the sequence above)

```
   1    MADLLSKFDP LIAEREALLA TGVRDPYAIV MDKVLSPTEA MINGRKTILL
  51    GTYNYMGMTF DPDVIAAGKQ ALDEFGSGTT GSRVLNGTYQ GHKACEDALK
 101    EFYGTEHAIV FSTGYQANLG MISTLAGKGD YIILDADSHA SIYDGCWLGD
 151    AEIVRFRHNS VEDLDKRLGR LPAEAGKLVV LEGVYSMMGD IAPLQEMVAV
 201    SKKHGAMILV DEAHGMGFFG EHGRGVFEEA GVEADVDFVV GTFSKSVGTV
 251    GGFCVSNHPK FEVLRLVCRP YVFTASLPPS VVATAATSIR KLMHAGDKRA
 301    HLWKNSRRLH QGLRDMGYKL GTETAQSAII AVILTDMAQA VALWQGLLEA
 351    GLYVNTARPP ATPAGMFLLR CSLCAEHSDE QVEQILGMFE SAGRATGVIP
 401    KLAAALEHHH HHH
```

Gene sequence pET22a/ ACPS SW

```
  1 ATGATCATCG GCATCGGCTC CGACCTCTGC AACATCGAGC GAATCCAGAA GTCGCTCGAC
 61 CGCTTCGGCG AGCGCTTCGT CGTCCGGGTG TTCACCGATG TCGAGCGCGC CAAGGCGGAG
121 CGCCGGGTGC TGACCCGCGC CGCCACCTAT GCGAAGCGAT TCGCGGCGAA GGAGGCCTTC
181 TCCAAGGCGG TGGGCACCGG CTTCAGGGCC GGCGTGTTCA TGAAGGACAT CGGCGTCGCC
241 AACCTGCCGA GCGGCGCCCC TACCCTGGTT TTGACCGGCG GGGCCAAGGC AAGGCTTGAC
301 GCGCTCACGC CGCCGGGCCA CGCGGTGGAT GTACATTTGA CGATGACGGA CGATCATCCC
361 TGGGCGCAGG CGTTCGTGAT CCTGTATGCG CGCCCGCTGC TCGAGCACCA CCACCACCAC
421 CACTGA
```

SW ACPS protein sequence (translated from the sequence above)

```
  1    MIIGIGSDLC NIERIQKSLD RFGERFVVRV FTDVERAKAE RRVLTRAATY
 51    AKRFAAKEAF SKAVGTGFRA GVFMKDIGVA NLPSGAPTLV LTGGAKARLD
101    ALTPPGHAVD VHLTMTDDHP WAQAFVILYA RPLLEHHHHH H
```

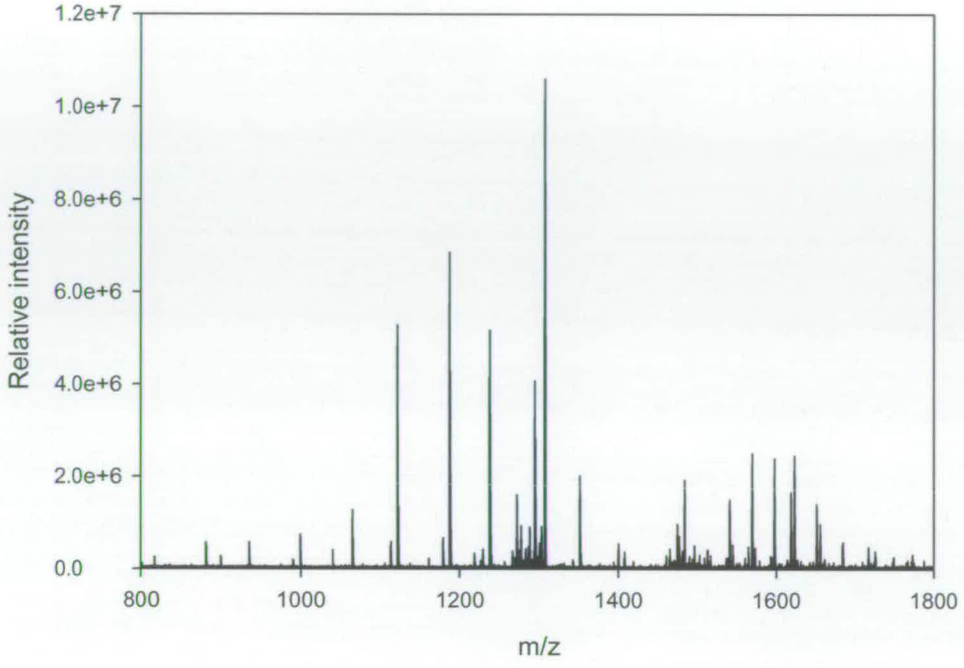
TOP-DOWN FT-ICR MASS SPECTROMETRY DATA

Method

Protein samples were buffer exchanged into 100 mM ammonium acetate, acidified with 2-fold H_2O:MeOH:HCOOH (v/v/v 50:48:2), and analysed by FT-ICR-MS at a final concentration of 5 μM. Mass spectrometry data was acquired on an Apex Ultra Qh-FT-ICR mass spectrometer equipped with a 12 Tesla superconducting magnet and an electrospray ion source (Bruker Daltonics, Billerica, MA). Nano-ESI was performed using a TriVersa Nanomate (Advion BioSciences, Ithaca, NY) running in infusion mode. Broadband data was typically acquired between *m/z* 600 and 4000 for 0.5 s and each spectrum was the sum of 32 scans. The mass spectra were externally calibrated using ES tuning mix (Agilent) and analysed using DataAnalysis software (Bruker Daltonics).

Top-down fragmentation was performed on the 12T Qh-FT-ICR. First, a specific ion species was isolated with the instrument's mass resolving quadrupole and MS/MS was performed using collision-induced dissociation (CID) The collision voltage was typically set between 20-35V. Fragmentation data was the sum of 200 scans and data analyses were performed using DataAnalysis (Bruker Daltonics). The SNAP 2.0 algorithm was used for automated peak picking and the resulting top-down fragment mass lists were searched against the primary sequence of ACP using Prosight-PTM software.[112, 141] Phosphopantetheinylation modification was calculated to produce a Δmass modification of 340.086 Da ($C_{11}H_{21}N_2O_6P_1S_1$).

Fragmentation spectra of holo-ACP



Top-down fragmentation map of holo-ACP

FRAGMENTATION DATA

Fragments list for peak at 1520 Da

#	m/z	I	Mol. Mass	MM-1	#	m/z	I	Mol. Mass	MM-1
1	714.483	347666	714.483	713.483	51	1576.185	78330	3151.363	3150.363
2	816.9463	99934	1632.885	1631.885	52	1577.308	110401	6306.21	6305.21
3	835.0118	77043	1669.016	1668.016	53	1578.36	203400	9465.124	9464.124
4	881.4701	203778	1761.933	1760.933	54	1579.867	170941	9474.165	9473.165
5	899.5476	353492	1798.088	1797.088	55	1582.703	141971	9491.18	9490.18
6	935.4276	563818	1869.848	1868.848	56	1584.869	129360	9504.174	9503.174
7	990.0848	158575	1979.162	1978.162	57	1586.036	222551	9511.177	9510.177
8	999.0892	1027422	1997.171	1996.171	58	1587.871	158598	9522.191	9521.191
9	1064.613	1431996	2128.219	2127.219	59	1589.381	258417	9531.248	9530.248
10	1113.124	691282	2225.241	2224.241	60	1591.86	78307	9546.123	9545.123
11	1121.639	6478008	2242.27	2241.27	61	1597.862	194960	6388.428	6387.428
12	1178.648	666168	2356.29	2355.29	62	1602.364	127680	6406.435	6405.435
13	1187.163	8192392	2373.318	2372.318	63	1605.582	1031437	6419.306	6418.306
14	1218.654	100166	3653.947	3652.947	64	1629.351	78044	6514.383	6513.383
15	1229.175	335044	2457.343	2456.343	65	1630.625	125197	6519.479	6518.479
16	1237.69	4825131	2474.372	2473.372	66	1633.867	2132571	6532.444	6531.444
17	1285.721	239197	2570.435	2569.435	67	1657.805	95550	6628.197	6627.197
18	1294.236	3426285	2587.465	2586.465	68	1662.334	2898583	6646.313	6645.313
19	1350.782	1650130	2700.556	2699.556	69	1682.89	1285845	6728.536	6727.536
20	1399.746	277234	4197.223	4196.223	70	1687.394	4274594	6746.555	6745.555
21	1407.329	171918	2813.651	2812.651	71	1715.654	1032469	6859.593	6858.593
22	1426.264	772760	1426.264	1425.264	72	1720.157	2807059	6877.605	6876.605
23	1463.798	85037	8777.751	8776.751	73	1748.669	932598	6991.654	6990.654
24	1464.845	214471	2928.683	2927.683	74	1756.566	315450	8778.801	8777.801
25	1470.289	522825	8816.697	8815.697	75	1764.159	224577	8816.765	8815.765
26	1474.119	70355	8839.677	8838.677	76	1765.48	101111	7058.899	7057.899
27	1476.616	425445	8854.657	8853.657	77	1772.359	265297	8857.768	8856.768
28	1480.279	122757	8876.638	8875.638	78	1773.961	120220	7092.823	7091.823
29	1483.109	577091	8893.617	8892.617	79	1779.344	131595	8892.692	8891.692
30	1486.605	154363	8914.591	8913.591	80	1781.433	735453	7122.708	7121.708
31	1489.268	414497	8930.571	8929.571	81	1784.722	138032	7135.865	7134.865
32	1493.261	62910	2985.515	2984.515	82	1787.976	117862	7148.882	7147.882
33	1495.762	304327	8969.537	8968.537	83	1790.492	213093	7158.948	7157.948
34	1499.258	197199	8990.512	8989.512	84	1794.728	107417	8969.611	8968.611
35	1502.088	580975	9007.492	9006.492	85	1813.45	392468	7250.778	7249.778
36	1505.751	108159	9029.469	9028.469	86	1817.488	79131	7266.928	7265.928
37	1508.415	356655	9045.453	9044.453	87	1820.74	81398	7279.937	7278.937
38	1514.741	215184	9083.407	9082.407	88	1823.007	106993	3645.007	3644.007
39	1514.81	151297	9083.822	9082.822	89	1824.59	659023	9118.922	9117.922
40	1515.518	762845	1515.518	1514.518	90	1831.211	297240	7321.821	7320.821
41	1517.811	1683973	9101.829	9100.829	91	2053.753	276250	6159.243	6158.243
42	1520.445	187816	3039.882	3038.882	92	2058.787	85978	6174.345	6173.345
43	1520.648	3397564	9118.854	9117.854	93	2091.453	495904	6272.345	6271.345
44	1522.916	603275	3044.824	3043.824	94	2096.489	117213	6287.452	6286.452
45	1522.978	153854	9132.833	9131.833	95	2129.152	590791	6385.44	6384.44
46	1524.311	493807	9140.83	9139.83	96	2157.168	138416	6469.489	6468.489
47	1526.807	1959378	9155.808	9154.808	97	2162.838	681459	6486.498	6485.498
48	1533.133	111449	9193.764	9192.764	98	2200.688	166307	6600.05	6599.05
49	1548.548	115662	6191.169	6190.169	99	2206.733	423066	6618.185	6617.185
50	1572.348	112741	6286.369	6285.369	100	2288.227	80263	6862.666	6861.666

FRAGMENTATION DATA

Fragments list for peak at 1305 Da

#	m/z	I	Mol. Mass	MM-1	#	m/z	I	Mol. Mass	MM-1
1	714.4831	680304	714.4831	713.4831	59	1483.451	1859264	8895.6689	8894.669
2	816.9468	204695	1632.8864	1631.886	60	1486.605	71404	8914.5961	8913.596
3	881.471	497034	1761.9346	1760.935	61	1487.777	73953	5948.084	5947.084
4	899.5482	233021	1798.0891	1797.089	62	1489.269	263585	8930.5763	8929.576
5	935.4277	484370	1869.8481	1868.848	63	1493.096	157897	8953.5374	8952.537
6	990.0846	145466	1979.1619	1978.162	64	1495.596	510371	8968.5367	8967.537
7	990.5747	122301	1980.1421	1979.142	65	1499.425	130003	8991.5129	8990.513
8	999.0893	677807	1997.1714	1996.171	66	1502.089	276662	9007.4962	9006.496
9	1040.22	330432	1040.22	1039.22	67	1508.415	206286	9045.4521	9044.452
10	1064.614	1180994	2128.2197	2127.22	68	1512.274	402251	6046.0755	6045.076
11	1113.124	560208	2225.2416	2224.242	69	1516.052	263797	6061.1861	6060.186
12	1121.639	5221105	2242.2701	2241.27	70	1536.044	213451	6141.1554	6140.155
13	1161.303	149317	3481.8945	3480.895	71	1538.387	205712	3075.7673	3074.767
14	1178.649	638422	2356.2907	2355.291	72	1540.548	1466809	6159.1684	6158.168
15	1187.163	6666894	2373.3183	2372.318	73	1544.325	430217	6174.2765	6173.277
16	1218.655	253705	3653.9517	3652.952	74	1548.668	75680	9286.9712	9285.971
17	1226.771	147424	1226.7706	1225.771	75	1552.182	95972	9308.0561	9307.056
18	1229.176	389841	2457.3441	2456.344	76	1559.826	162415	4677.4644	4676.464
19	1237.69	4997191	2474.3727	2473.373	77	1564.318	390922	6254.2482	6253.248
20	1256.352	107293	3767.0423	3766.042	78	1568.821	2511790	6272.2621	6271.262
21	1265.952	323188	8855.6185	8854.619	79	1572.599	422557	6287.3736	6286.374
22	1268.947	244639	8876.5881	8875.588	80	1577.809	130378	6308.2149	6307.215
23	1271.231	1600710	8892.5734	8891.573	81	1592.589	247499	6367.3346	6366.335
24	1274.371	301378	8914.5526	8913.553	82	1593.613	63238	6371.4308	6370.431
25	1276.653	885493	8930.53	8929.53	83	1597.094	2237073	6385.3541	6384.354
26	1279.793	180736	8952.509	8951.509	84	1597.134	115578	3193.2609	3192.261
27	1282.219	395519	8969.4895	8968.49	85	1602.914	67082	3204.8206	3203.821
28	1285.216	333222	8990.4659	8989.466	86	1611.103	71600	6441.391	6440.391
29	1286.222	204428	2571.4364	2570.436	87	1613.353	180231	6450.3906	6449.391
30	1287.642	850680	9007.4507	9006.451	88	1617.855	1557866	6468.396	6467.396
31	1290.638	183172	9028.4194	9027.419	89	1618.069	84045	3235.1305	3234.131
32	1292.921	392798	9044.4041	9043.404	90	1622.359	2444420	6486.4132	6485.413
33	1294.236	4004073	2587.4655	2586.466	91	1623.151	99712	1623.1511	1622.151
34	1298.487	277748	9083.367	9082.367	92	1626.123	178729	6501.4694	6500.469
35	1298.519	192249	5191.054	5190.054	93	1630.376	137963	6518.4813	6517.481
36	1299.371	262266	3896.0977	3895.098	94	1643.42	64304	6570.6582	6569.658
37	1300.128	530756	3898.3707	3897.371	95	1646.206	137911	6581.8039	6580.804
38	1301.817	131061	1301.8169	1300.817	96	1650.744	1291034	6599.9529	6598.953
39	1302.769	944566	2604.5303	2603.53	97	1655.278	849547	6618.0904	6617.09
40	1307.106	10875908	3919.3033	3918.303	98	1660.439	196863	3319.8703	3318.87
41	1331.241	85184	2661.4736	2660.474	99	1683.635	589363	6731.5171	6730.517
42	1342.267	159797	2683.5271	2682.527	100	1708.677	76779	6831.685	6830.685
43	1350.782	1983837	2700.5561	2699.556	101	1716.398	437571	6862.5707	6861.571
44	1393.475	100771	5570.8782	5569.878	102	1724.96	333608	3448.9127	3447.913
45	1399.746	481149	4197.2234	4196.223	103	1747.257	104161	5239.7568	5238.757
46	1407.329	310243	2813.6504	2812.65	104	1748.415	231038	6990.6395	6989.64
47	1407.412	166154	4220.221	4219.221	105	1764.558	68427	8818.7628	8817.763
48	1418.494	150444	5670.9543	5669.954	106	1766.175	154721	7061.6778	7060.678
49	1440.436	79899	4319.2946	4318.295	107	1771.27	117380	5311.7952	5310.795
50	1460.453	196862	4379.3447	4378.345	108	1772.362	289941	8857.7789	8856.779
51	1464.847	373917	2928.6865	2927.687	109	1774.499	95098	3547.9903	3546.99
52	1467.122	101902	8797.6933	8796.693	110	1786.937	109264	8930.6567	8929.657
53	1470.289	188772	8816.7001	8815.7	111	1804.299	163522	5410.8809	5409.881
54	1471.453	114809	8823.6829	8822.683	112	1845.043	73201	3689.0787	3688.079
55	1474.287	963061	8840.6845	8839.685	113	1885.659	84684	5654.9627	5653.963
56	1476.616	663874	8854.6602	8853.66	114	2016.052	99671	6046.1428	6045.143
57	1480.786	369587	8879.6772	8878.677	115	2053.754	178630	6159.2461	6158.246
58	1482.609	142416	8890.6194	8889.619	116	2091.455	91693	6272.3504	6271.35

FRAGMENTATION DATA

Prosight PTM Output

Ion	Observed Mass (Da)	Theoretical Mass (Da)	Mass Error (Da)	Mass Error (PPM)
Y32	3546.9903	3546.7968	0.1935	54.54
Y31	3447.9127	3447.7284	0.1843	53.44
Y30	3318.8703	3318.6858	0.1845	55.58
Y29	3203.8206	3203.6589	0.1617	50.47
Y28	3074.7673	3074.6163	0.151	49.11
Y27	2927.683	2927.5479	0.1351	46.15
	2927.6865	2927.5479	0.1386	47.34
Y26	2812.6508	2812.521	0.1298	46.16
	2812.6504	2812.521	0.1294	46.02
Y25	2699.5557	2699.4369	0.1188	44.01
	2699.5561	2699.4369	0.1192	44.16
Y24	2586.4649	2586.3528	0.1121	43.33
	2586.4655	2586.3528	0.1127	43.56
Y23	2473.3719	2473.2688	0.1031	41.69
	2473.3727	2473.2688	0.1039	42.02
Y22	2372.3178	2372.2211	0.0967	40.76
	2372.3183	2372.2211	0.0972	40.97
Y21	2241.2696	2241.1806	0.089	39.71
	2241.2701	2241.1806	0.0895	39.93
Y20	2127.2194	2127.1377	0.0817	38.42
	2127.2197	2127.1377	0.082	38.56
Y19	1996.171	1996.0972	0.0738	36.98
	1996.1714	1996.0972	0.0742	37.18
Y17	1797.0879	1797.0015	0.0864	48.08
	1797.0891	1797.0015	0.0876	48.75
Y16	1668.0163	1667.9589	0.0574	34.41
Y12	1225.7706	1225.7162	0.0544	44.4
Y7	713.483	713.4567	0.0263	36.83
	713.4831	713.4567	0.0264	36.97
B62	7320.8214	7320.4257	0.3957	54.05
B61	7249.7777	7249.3886	0.3891	53.67
B60	7121.7084	7121.3301	0.3783	53.13
B59	6990.6538	6990.2896	0.3642	52.1
B58	6876.605	6876.2466	0.3584	52.11
B57	6745.555	6745.2061	0.3489	51.72
B55	6531.4441	6531.0744	0.3697	56.6
B54	6418.306	6417.9903	0.3157	49.18
B53	6305.2101	6304.9063	0.3038	48.19
B52	6190.1693	6189.8793	0.29	46.84
B48	5669.9543	5669.6988	0.2555	45.06
B37	4196.2229	4196.0338	0.1891	45.08
	4196.2234	4196.0338	0.1896	45.2
B35	3895.0977	3894.9275	0.1702	43.7
B34	3766.0423	3765.8849	0.1574	41.79
B33	3652.9468	3652.8009	0.1459	39.95
	3652.9517	3652.8009	0.1508	41.3
B31	3480.8945	3480.7525	0.142	40.81
B27	3074.7673	3074.5672	0.2001	65.07
B23	2660.4736	2660.3558	0.1178	44.29
B15	1760.933	1760.8693	0.0637	36.17
	1760.9346	1760.8693	0.0653	37.08
B14	1631.8853	1631.8267	0.0586	35.9
	1631.8864	1631.8267	0.0597	36.58

PUBLICATIONS

6
Natural and Synthetic Stimulators of the Immune Response

Marine C. Raman and Dominic J. Campopiano

6.1
Introduction

Mammals have evolved sophisticated early-warning, self-defense systems to be ready to defend themselves against incoming pathogenic organisms. Once a pathogen has been detected, a vast battery of weaponry can be sent out to tackle the invader. This natural innate immunity is required to sense a range of pathogen-specific signals and set off an immunological response cascade that will lead to controlled killing and mediated cell death. Controlled inflammation is required and delicate balances must be maintained, since unnecessary stimulation and inflammation is highly dangerous and detrimental to the mammalian host. A range of stimuli are recognized by the innate immune system and in this chapter we concentrate on two that provide fascinating fields of study for microbiologists, immunologists, clinicians and chemists alike. Detailed molecular knowledge of "natural" stimulators and their protein receptors has begun to impact on the synthesis of chemical tools with highly desirable immunomodulatory properties which could have great clinical use.

6.2
Lipopolysaccharide Endotoxin – A Potent Immunostimulatory Molecule

Bacterial membranes are composed of lipids and proteins. The selective recognition of bacterial versus mammalian membrane components (nonself and self) is an important function of the immune system. Specific bacterial components can induce acute inflammatory responses via a number of sensitive cascades. One important molecule that has fascinated scientists for many years is lipopolysaccharide (LPS), also known as "endotoxin". An excellent review of the story of the discovery, characterization and biological analysis of this molecule can be found in [1]. Endotoxin is a highly abundant proportion of the outer membrane of the Gram-negative bacterial cell wall and can also be shed by growing as well as dying

Cellular and Biomolecular Recognition: Synthetic and Non-Biological Molecules. Edited by Raz Jelinek
Copyright © 2009 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim
ISBN: 978-3-527-32265-7

cells. The discovery and characterization of endotoxin dates back over 100 years. At the end of the 1800s, Richard Pfeiffer showed that cholera bacteria that had been killed by heat retained their toxic potential. This experiment proved that the poison was not a classical protein toxin and led him to conceive the idea that *Vibrio cholerae* harbored a heat-stable toxic substance that was associated with the insoluble part of the bacterial cell. The active substance was named endotoxin (taken from the Greek *endo* meaning "within"). Pfeiffer postulated that these endotoxins were fundamental constituents of bacterial species. Chemical characterization of the endotoxins involved contributions from many researchers who devised methods to extract, purify and analyze this complex natural product (Osborn, Nikaido, Lüderitz, Westphal, Boivin and Rietschel).

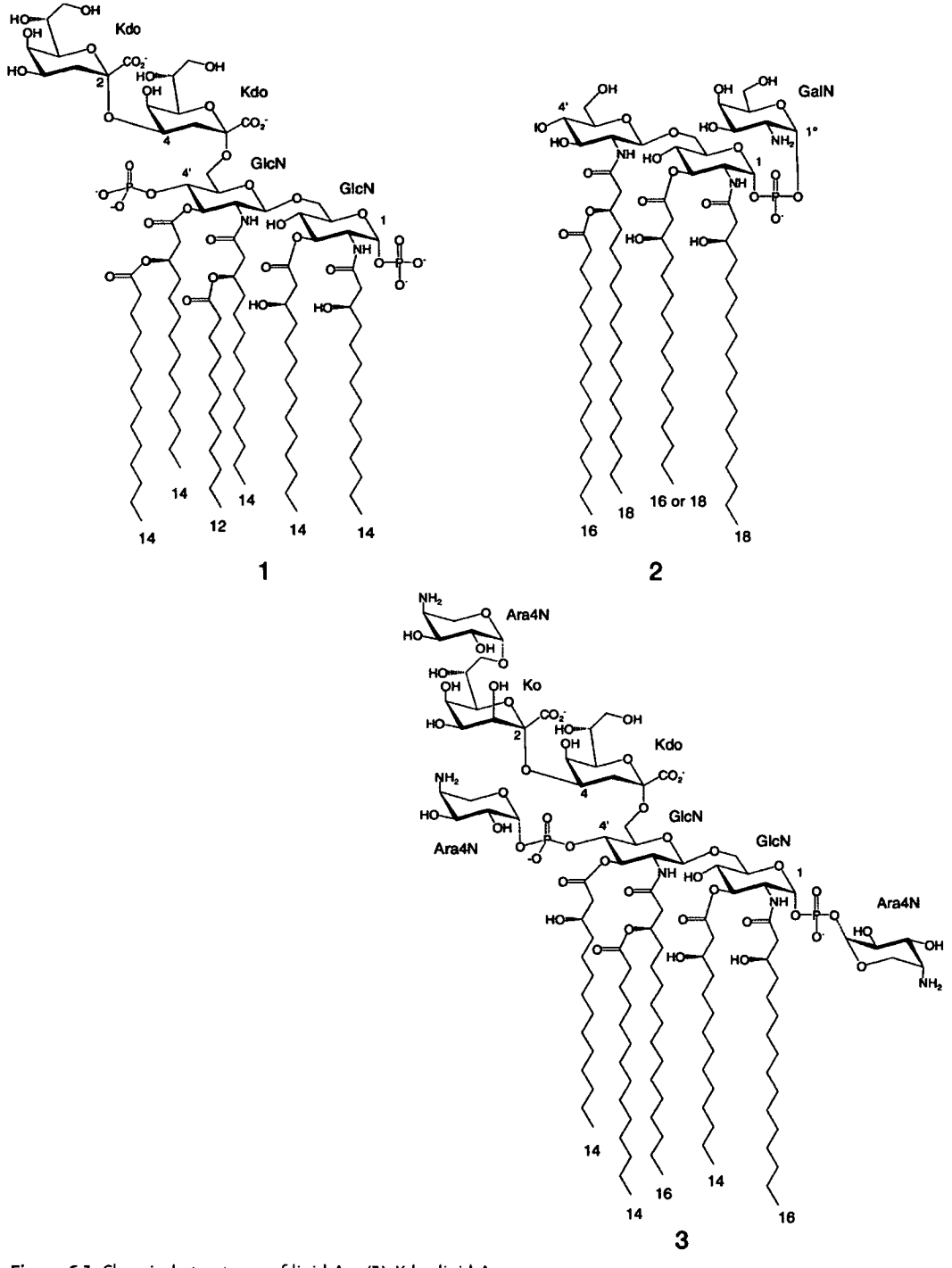
Detailed chemical analysis ultimately revealed LPS to be organized into three parts – lipid A, core and O-antigen. There are around 10^6 lipid A residues and 10^7 glycerophospholipids per *Escherichia coli* cell [2]. The LPS from *E. coli* has been studied extensively to provide a model endotoxin architecture for comparative studies with those from other species. The lipid A (Kdo_2-lipid A) from a deep rough *E. coli* strain is shown in Figure 6.1 (compound 1).

The lipid A is an usual glucosamine-based β1,6-linked disaccharide with long-chain fatty acids attached to the 2 and 2′ positions via amides. The 3 and 3′ positions are also derivatized as fatty acid esters. Monophosphate esters are found at the 1 and 4′ positions, and can be the site for further functionalization (see below). The 6′ position links the glucosamine disaccharide to an unusual 3-deoxy-D-manno-octulosonic acid (Kdo) disaccharide. To this Kdo is the site of attachment of heptose sugars that form part of the core region. Finally, the characteristic, bacterial-specific oligosaccharide O-specific chain ("O-antigen") is attached to this core.

A seminal breakthrough in the field of endotoxin research came in the 1980s when Kusumoto and Shiba synthesized *E. coli* lipid A, thereby confirming the structure of the isolated natural product. More importantly, this synthetic lipid A had identical endotoxic activity against mammalian cells as its bacterial counterpart [3]. Thereafter began a series of structure–functional relationships to explore the nature of the inflammatory potential of this important bacterial toxin [4]. A large number of synthetic lipid A derivatives were prepared, but none were as active as the parent molecule.

6.3 LPS Recognition

Various proteins have been sequentially discovered that generate a circuit that connects LPS recognition through cell surface receptor binding and intracellular signaling [5]. LPS-binding protein (LBP) is an approximately 60-kDa serum binding protein that belongs to the lipid transfer family. It is the LBP that binds the LPS in the blood and delivers it to the CD14 receptor that is present on many cell types, including macrophages and dendritic cells [6]. CD14 is a glycosyl phosphatidylinositol-anchored, high-affinity membrane protein that can also circulate in a soluble



12 **Figure 6.1** Chemical structures of lipid As. (**1**) Kdo_2-lipid A, from a deep rough strain of *E. coli*. (**2**) Lipid A from *F. novicida* U112. (**3**) Lipid A core from *B. cepacia* (dashed line represents non-stoichiometric substitution).

form. Another protein involved in passing this signal onto cell-surface receptors is myeloid differentiation-2 (MD-2) (see below). It is thought that CD14 concentrates LPS for binding to a Toll-like receptor (TLR) TLR4–MD-2 complex. How this complex recognizes lipid A and signals across the plasma membrane is still not completely understood [7]. Toll is a *Drosophila* fruit-fly receptor essential for the production of antimicrobial peptides in response to fungal pathogens that is also essential for fly development. This link in flies led to the isolation of a human homolog then the discovery of a family of mammalian transmembrane, Toll-like pattern recognition receptors. The Toll receptor has the characteristic modular structure of a type I transmembrane receptor: an ectodomain containing blocks of leucine-rich repeats (LRRs) and a conserved intracellular region, the Toll/interleukin (IL)-1 receptor domain, that is required for downstream signal transduction. In 1998, the direct link between LPS and the TLR4 receptor was made by Beutler *et al.* [who had previously isolated tumor necrosis factor (TNF)-α] [8] and the expanding TLR family has been intensively studied over the past decade [7–12]. Specific TLRs have been shown to interact with RNA, DNA and many other ligands. The TLR4 pathway activates the key transcriptional regulator NF-κB, resulting in the production of proinflammatory cytokines and the progression to adaptive immunity. A complicated series of intracellular proteins are then expressed including MyD88, although a MyD88-independent pathway has also been discovered. The complex gene expression networks are highly regulated, and are now the target of various anti-inflammatory and autoimmune disease drugs. The details of LPS binding and signaling, and discussions of the TLR family have been the subject of a number of recent excellent reviews [5, 9, 12–15].

6.4 Septic Shock

Sepsis is a systemic inflammatory response syndrome triggered by an infection and is the major cause of death in patients in intensive care units [16]. In severe cases it can lead to organ failure or dysfunction, and severe sepsis with hypotension unresponsive to fluid resuscitation defines septic shock. Sepsis syndrome results from the body's systemic inflammatory response to any of several infectious stimuli. Endotoxins, such as LPS from Gram-negative bacteria and other antigens from infectious agents (e.g., DNA, peptidoglycan and flagellan, lipotechoic acid, mannan) stimulate macrophages and monocytes to release TNF-α, resulting in a cascade of cytokine release [17]. Following the release of TNF-α, other proinflammatory cytokines, including IL-1 and IL-6, are released into the circulation that, in turn, trigger numerous additional proinflammatory events within endothelial cells and leukocytes. Although these proinflammatory responses are vital to the host defense against infection, TNF-α acts in conjunction with IL-1 to produce the clinical signs of the systemic inflammatory response syndrome, and their synergistic effects are probably responsible for the hypotension and resultant organ dysfunction seen early in the course of severe sepsis. Sepsis

syndrome afflicts almost 750000 patients in the United States each year, costing almost $17 billion and causing 210000 deaths annually. Despite increasing research and accrued knowledge about the pathophysiological pathways and processes involved in sepsis, morbidity and mortality remain unacceptably high. A large number of natural and synthetic immunomodulatory agents have been studied in experimental and clinical settings in an attempt to find an efficacious
4 anti-inflammatory drug that reduces mortality (see below [18]).

6.5 LPS Biosynthesis

The efforts of Raetz *et al.* have delineated the biosynthetic steps involved in the synthesis of the lipid A core using *E. coli* as a model (see excellent reviews [19, 20]). There are nine enzymes required to convert the UDP-Glc-NAc sugar nucleotide precursor into Kdo_2-lipid A. Each enzyme has been named, the gene cloned and characterized to some extent at a biochemical or genetic level: LpxA, LpxC, LpxD, LpxH, LpxB, LpxK, WaaA, LpxL and LpxM. In the first step, the UDP-Glc-NAc building block is acylated with an acyl-carrier protein substrate in a thermodynamically reversible step. Thereafter, an *N*-acyl deactylase irreversibly removes the acetate group which commits the UDP-sugar to the pathway. The *N*-acetyl glucosamine disaccharide is generated by cleavage of the UDP-pyrophosphate of 1 mol of bis-acylated UDP-2,3-diacylglucosamine to give 2,3-diacylglucosamine-1-phosphate (lipid X). Then another mole of bis-acylated UDP-2,3-diacylglucosamine is condensed with lipid X to give the disaccharide. Phosphorylation of the 4′ position is then followed by the addition of the Kdo residue(s). This step is bacteria-specific and Kdo-transferase isozymes can add one, two or three Kdo residues. The product in *E. coli*, Kdo_2-IVa, with four fatty acids is then converted to the hexa-acylated Kdo_2-lipid A by two acyl-transferases. Since the early genes in the pathway are essential they have been identified as good targets for the development of new lipid A-specific antibiotics. One such antibiotic, CHIR-90, has been shown to be active against *Pseudomonas aeruginosa* and *E. coli*. This is specific for the second enzyme, the deacetylase LpxC, and a recent structure of the LpxC/CHIR-90 complex has revealed the molecular details of the mechanism of inhibition thus paving the way for development of more potent agents [21].

6.6 Minimal, Modified Lipid A

Until recently, it was thought that the minimal LPS required for the growth of *E. coli* consisted of the lipid A core and disaccharide made up of two Kdo sugars (e.g.,
5 Figure 6.1a). However, recent work by Woodard *et al.* has challenged this dogma. Meredith *et al.* revealed that an *E. coli* K-12 nonconditional suppressor strain, KPM22, disrupted in D-arabinose-5-phosphate isomerase activity (through

manipulation of *kdsD* and *gutQ*) could not produce a LPS containing Kdo and thus generated the endotoxically inactive LPS precursor lipid IVa [22]. These cells grew and were viable, albeit that they were hypersensitive to large, hydrophobic antibiotics (e.g., rifampin, fusidic acid, novobiocin and erythromycin) and sodium dodecyl sulfate. The cell wall of the mutant strain also showed unusual altered morphology in the inner and outer membrane cell walls. Interestingly, inclusion of the inner membrane LPS transporter MsbA on a multicopy plasmid partially suppressed the lethal Kdo knockout phenotype directly in the auxotrophic parent strain, suggesting increased rates of nonglycosylated lipid A transport can, in part, compensate for Kdo depletion.

It appears that bacterial species use a range of biosynthetic strategies to build lipid A derivatives that allows them to escape detection by the mammalian innate immunity surveillance system. This is exemplified by a recent study that revealed *Francisella tularensis*, the causative agent of tularemia – a highly contagious disease of animals and humans, and the related mouse pathogen *Francisella novicida* synthesize unusual lipid A molecules lacking the 4′-monophosphate group typically found in the lipid A of Gram-negative bacteria (Figure 6.1, compound 2). During lipid A biosynthesis LpxF, a selective phosphatase located on the periplasmic surface of the inner membrane, removes the 4′-phosphate moiety in the late stages of *F. novicida* lipid A assembly. To evaluate the relevance of the LpxF 4′-phosphatase to pathogenesis, Raetz *et al.* constructed a *lpxF* deletion mutant and compared its virulence with the wild-type organism [23]. Intradermal injection of 1×10^6 wild-type *F. novicida* cells is lethal to mice, but 100-fold (1×10^8) mutant is not. The rapid clearance of the *lpxF* mutant was associated with a stronger local cytokine response and a greater influx of neutrophils compared with wild-type. Furthermore, the *lpxF* mutant was highly susceptible to the cationic antimicrobial peptide polymyxin. The authors conclude that LpxF represents a kind of virulence factor that confers a distinct lipid A phenotype, preventing *Francisella* from activating the host innate immune response and preventing the bactericidal actions of host cationic peptides. Here, the loss of a simple phosphate moiety allows the bacterial pathogen to survive. This emphasizes how chemically sensitive the mammalian LPS detection is, but also suggests it can be "blindfolded" by a small modification [24].

It is now apparent that there is an increasingly diverse range of covalent modifications that various bacteria use to decorate their lipid A core. This allows them to escape the mammalian immune system but also protects them from killing by lethal agents such as cationic antimicrobial peptide antibiotics (cAMPs) including mammalian defensins and bacterial polymyxins [polymixin B (PMB)] [25]. Moreover, bacteria can rapidly react to changes in their external environment (e.g., pH, NaCl concentration, as well as Mg^{2+}, Ca^{2+}, Fe^{3+}) and alter their metabolism to express lipid A-modifying enzymes. These enzymes have been shown to catalyze the transfer of, among other things, phosphoethanolamine (e.g., EptA and EptB) and the unusual sugar 4-amino-4-deoxy-L-arabinose (L-Ara4N; catalyzed by ArnT) onto the lipid A phosphate residues. Alteration of the fatty acid content of lipid A can also be achieved using acyl-transferases (e.g., PagP uses palmitate).

An excellent, in-depth review of this field has recently been published [26]. The sensing of environmental change that controls expression of the modification genes is complex, involving the interesting a two-component PhoP/PhoQ system. Detailed molecular, genetic and immunological analyses of this system and its role in *Salmonella* pathogenesis have been carried by out Miller *et al.* over a number of years [27]. Unusual anomalies in lipid A biosynthesis continue to spring up – we, in collaboration with others, have studied the pathogen *Burkholderia cepacia* and related species, and their unusually high resistance to cAMPs and other clinically useful antibiotics. These organisms cause a range of clinical problems, especially to Cystic Fibrosis sufferers, and infection with certain strains leads to chronic lung inflammation and increased mortality. *Burkholderia* LPS is also highly inflammatory and has unusual fatty acid and sugar content including L-Ara4N modification on the Ko octulosonic acid sugar (Figure 6.1, compound 3) [28, 29]. *Burkholderia* are able to modify both their lipid A core sugar phosphates and their inner core sugars by Ara4N presumably with two different transferases. It is interesting to also note that *Burkholderia* LPS cannot be neutralized by polymyxin and indeed the LPS/PMB complex has unusual properties [30]. Covalent L-Ara4N modification of LPS has been linked to increased cAMP resistance in *E. coli* and *Pseudomonas aeruginosa*, but the L-Ara4N biosynthetic pathway in these bacteria is not essential since deletion causes an increased cAMP sensitivity. In contrast, we found that the operon encoding the Ara4N biosynthetic genes in *B. cepacia* is constitutively expressed and that deletion of the L-Ara4N biosynthetic genes in strain K56-2 was lethal, suggesting that this pathogen has evolved to have an intrinsic requirement for this sugar [31]. We could not isolate sufficient quantities of LPS from the mutant strains to characterize the L-Ara4N content, but there appears to be no logical biochemical explanation as to why *Burkholderia* should need its lipid A to be covalently modified. It could be that L-Ara4N plays roles in other, as yet undiscovered, metabolic pathways in this highly resistant pathogen.

6.7
Isolation of "Natural" Kdo_2-Lipid A

Such is the importance of LPS in human biology that the National Institutes of Health and National Institute of General Medical Sciences awarded a 5-year, $35 million grant in 2003 to fund a Lipid Metabolites And Pathways Strategy (LIPID MAPS) consortium led by Edward A. Dennis which included 18 universities, medical research institutes and companies across the United States [32]. They have worked together in a detailed analysis of the structure and function of lipids, and their work is published in major journals and online (see www.lipidmaps.org and www.sphingomap.org). Their goal is to provide comprehensive procedures to the community for identifying all lipids produced by mammalian cells. They have begun with the macrophage, following activation by endotoxin, aiming to quantify temporal and spatial changes in lipids that occur with cellular metabolism, as well as to develop bioinformatic approaches that establish dynamic lipid networks. To

achieve these aims, a natural endotoxin of the highest possible analytical specification was crucial. The LPS used by researchers in the immunology field are typically chemically heterogeneous mixtures obtained from bacterial strains such as *E. coli* or *Salmonella typhimurium*. These preparations have high inflammatory properties, but display batch-to-batch variability. An extremely useful tool has recently become available from the LIPID MAPS consortium. They recently reported the isolation of Kdo_2-lipid A, a nearly homogeneous Re LPS substructure with endotoxin activity equal to LPS [33]. The LPS was extracted from 2 kg cell paste of a heptose-deficient *E. coli* mutant WBB06. This material is composed of six lipid A molecules with Kdo_2-lipid A present in 91% abundance (compound 1, Figure 6.1). The structure and purity were evaluated by electrospray ionization mass spectroscopy, liquid chromatography mass spectroscopy and ^{1}H-nuclear magnetic resonance (NMR), and its activity was compared with LPS in RAW 264.7 cells and bone marrow macrophages from wild-type and TLR4-deficient mice. Cytokine and eicosanoid production (e.g., TNF-α), in conjunction with gene expression profiling were employed as readouts, and it was found that this material was as potent as LPS. The possibility of generating pure ^{13}C-labeled Kdo_2-Lipid A should greatly facilitate NMR solution studies of complexes with MD-2 or CD14. This material is commercially available through Avanti Polar Lipids.

6.8
Synthetic LPS, Lipid A and Their Uses

As knowledge of the molecular details of the host/pathogen interaction is increased, it is hoped that this could be transferred into useful clinical applications. One such expanding area is the adjuvant field. An effective adjuvant should help modulate or enhance the effects of an antigen (vaccine) and promote an immune response, but have no toxic side-effects itself. A recent reviews of this area was recently published [34]. These adjuvants could act upon a number of cells and receptors to illicit a range of responses, and have taken the form of TLR agonists. Others such as liposomes and mineral salts (e.g., aluminum hydroxide, alum) have been used extensively where their exact mode of action is still unclear. Again, an understanding of what makes a pathogen able to escape detection as well as those with potent inflammatory properties could aid in adjuvant design.

Since natural preparations are inherently heterogeneous (even the *E. coli* Kdo_2-lipid A has five other lipid A derivatives), synthetic preparations of single compounds have proven useful tools to tease out the details of LPS recognition. A number of groups have developed synthetic routes to various lipid A derivatives and tested their activity against a range of cell types. It is around 20 years since Shiba *et al.* made synthetic lipid A derivatives and tested their activity [3]. A nontoxic lipid A derivative from *Salmonella Minnesota* R595 lacking the 1-phosphate and fatty acid esters on the first glucosamine was chemically synthesized and this monophosphoryl lipid A (MPL) was shown to have low cytokine stimulation activity (Figure 6.2, compound 4). This further emphasized that hexa-acylation is a

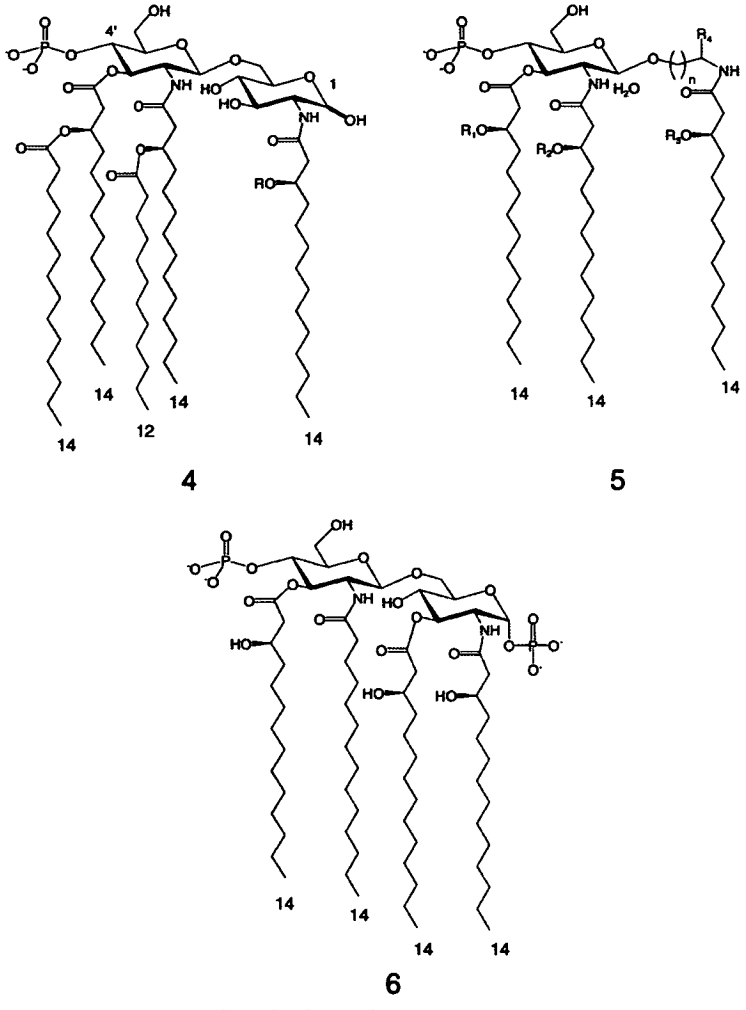


Figure 6.2 Various synthetic lipid A analogs. (**4**) MPL (R = H or C16). (**5**) Generic AGPs (R_1, R_2, R_3, R_4 range in change length, n = 1–7). In adjuvant RC529 $R_1 = R_2 = R_3 = n\text{-}C_{13}H_{27}CO$, R_4 = H, n = 1. (**6**) Synthetic lipid IVa.

prerequisite for endotoxin activity and that underacylated lipid As tend to be less immunologically active. The authors also probed fatty acid chain length, and used their strategy to prepare the MPL containing the all C14 hexa-acyl lipid A which had "intermediate" activity and toxicity. Stöver *et al.* used transcriptional profiling of human monocytes and responses of TLR4 transfectants to demonstrate a clear dependence of the length of the secondary acyl chain on receptor activation [35]. Also of interest is a MPL derivative with seven fatty acids that is being used as an adjuvant in therapeutic cancer vaccine trials [36].

A range of monosaccharide lipid A derivatives have also been prepared and tested for biological activity (reviewed in [18].). In some cases the glucosamine saccharide core has been retained, whereas in others it has been replaced altogether by an acylic derivative. Although penta-acylated monosaccharides tend to be inactive, some containing three acyl residues on a nonreducing monosaccharide are endotoxic, but still 100 times less active than *E. coli* lipid A in TNF-α production, possibly due to its biophysical properties. To overcome these problems, a series of synthetic aminoalkyl glucosaminide 4-phosphonates (AGPs) has been investigated (Figure 6.2, compound **5**). These glycolipid/LPS "hybrids" are based on a serine scaffold and it appears that the carboxyl group is a good isostere for the 1-phosphate of lipid A. A number of academic and industrial groups have explored the structure–activity relationships of these AGPs by changing esters for ethers, moderating chain length and optimization of the distances between functional groups. This has led to the clinical use of RC-529 as a safe and effective adjuvant in a hepatitis B vaccine.

Boons *et al.* have also developed elegant, highly convergent synthetic strategies to generate a range of useful lipid A analogs. They generate a "prototypical" lipid A from *E. coli* that is hexa-substituted in an asymmetrical fashion, as well as a derivative with several of its acyl groups having been shortened. They also synthesized hepta-acylated lipid As derived from *S. typhimurium* LPS that differ in lipid length and phosphorylation pattern. Their strategy employed an advanced disaccharide intermediate building block that could then be selectively modified with any lipid at the desired position [37]. The authors then exposed mouse macrophages to their synthetic lipid As and *E. coli* 055:B5 LPS, and the resulting supernatants were examined for mouse TNF-α, interferon (IFN)-β, IL-6, IP-10, RANTES (Regulated upon Activation, Normal T cell Expressed and Secreted) and IL-1β1. They found that particular modifications had different effects on the potencies and efficacies of induction of the various cytokines. However, no bias toward a MyD88- or TRIF (Toll/IL-1 receptor domain-containing adaptor)-dependent response was observed, which emphasizes the fact that lipid A derivatives modulate innate immune responses in a highly complex manner.

As well as tackling lipid A targets, Boons *et al.* have also examined the role of the Kdo moieties in enhancing the biological activities of LPS. They have also optimized the synthesis of a lipid A derivative containing Kdo and the strategy was employed for the synthesis of *Neisseria. meningitidis* lipid A containing Kdo on the nonreducing sugar [38]. Mouse macrophages were exposed to the synthetic compound and *E. coli* lipid A and a hybrid derivative that has the asymmetrical acylation pattern of *E. coli* lipid A, but the shorter lipids of meningococcal lipid A. The resulting supernatants were examined for TNF-α and IFN-β production. The lipid A derivative containing Kdo was much more active than lipid A alone and just slightly less active than its parent LPS, indicating that one Kdo moiety is sufficient for full activity of TNF-α and IFN-β induction. The lipid A of *N. meningitidis* was a significantly more potent inducer of TNF-α and IFN-β than *E. coli* lipid A, which is due to a number of shorter fatty acids. The compounds did not demonstrate a bias towards a MyD88- or TRIF-dependent response.

A number of Japanese groups have also contributed greatly to the development of synthetic tools for studying lipid A activity [4, 39]. Various lipid A analogs, as well as their radiolabeled derivatives and more complex partial structures of LPS have also synthesized. For example, a Kdo_2-lipid A tetrasaccharide (Re LPS) has been synthesized and its IL-6-inducing activity tested. The synthetic Kdo_2-lipid A clearly showed that the addition of the Kdo residues enhances the potency of lipid A to activate the innate immune system. Their strategy also allowed the synthesis of ^{3}H-labeled LPS to study the stoichiometry of binding to various receptors (e.g., MD-2, TLR4). These synthetic LPS derivatives including molecules such as lipid IVa have also been used in high resolution structural studies (Figure 6.2, compound 6, and below) [40].

The neutralization of LPS to prevent toxic shock has also been a goal of synthetic chemists. There is growing literature in this field but, a recent example took a lead from the fact that the bacterial peptide natural product PMB binds and neutralizes LPS toxicity and is commonly used by immunologists. PMB is too toxic to be used clinically; however, based on the NMR-derived model of a PMB–LPS complex the authors designed a small molecule (a linked mono-alkyl spermine derivative) that binds LPS and neutralizes its toxicity with a potency indistinguishable from that of PMB in a wide range of *in vitro* assays. It also affords complete protection in a murine model of LPS-induced lethality and is apparently nontoxic in vertebrate animal models [41, 42].

Finally, LPS has been an attractive target for synthetic chemists to target sensitive sensors to detect the presence of contaminating endotoxin in water and other fluids. For example, one group have prepared fluorescently labeled analogs of two peptide variants derived from the putative ligand-binding domain of the LBP CD14 that detect and discriminate LPS and lipids down to the submicromolar concentration range [43]. In a different sensor approach fluorescent quantum dots coated with a zinc complex can selectively stain a rough LPS *E. coli* mutant and permit optical detection in a living mouse leg infection model [44].

6.9
Structural Studies

6.9.1
LBP, CD14 and FhuA

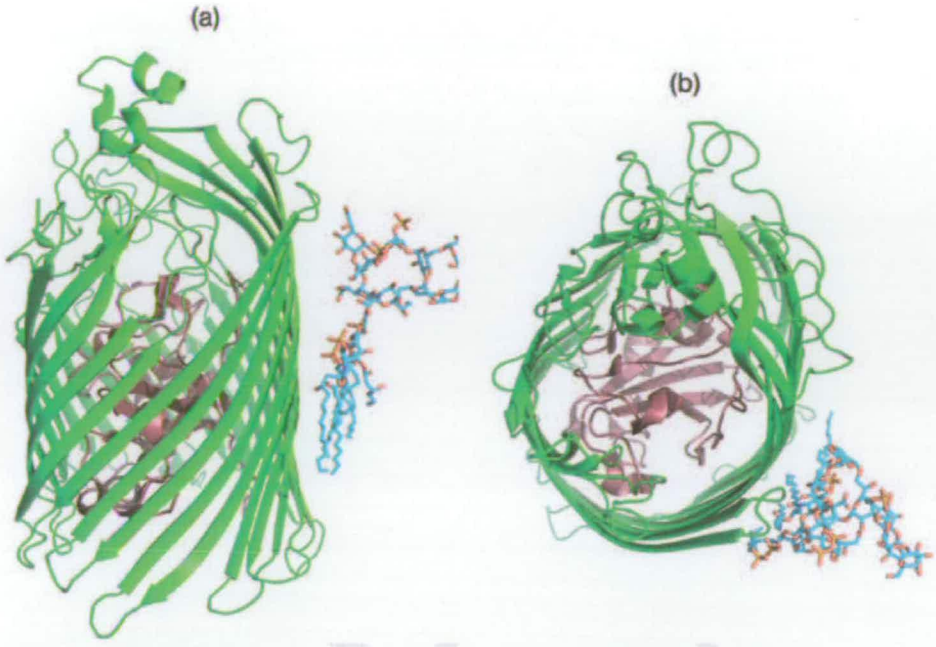
It is hoped that knowledge of the high-resolution three-dimensional structures of various LBPs (e.g., MD-2, CD14 and TLR4) in complex with synthetic LPS/lipid A ligands and their natural product ancestors will further guide the design of more specific, clinically useful molecules [45]. Furthermore, the structures of protein complexes will shed light on how these proteins signal to each other and generate the interlinked networks. A number of structural studies have been carried; these are too extensive to cover in this chapter, but we aim to pick out personal highlights. Efforts to reveal the molecular structures of LPSs from various organisms

and their interaction with bacterial and mammalian protein targets are time consuming due to the inherent heterogeneity of natural LPS and the fact that many of the protein receptors are membrane bound.

The importance of LBP as the initial LPS binder has meant it has undergone a range of structural studies, but a high-resolution LBP/LPS structure has not been realized to date. However, the structure of a related human bactericidal/permeability-increasing protein has been determined and has been used as a model for LBP [46]. A group of cationic residues at the N-terminus of LBP have been shown to be essential for LPS binding. Synthetic fragments of LBP (14 amino acids) have been used in NMR studies with LPS to reveal the LPS-binding epitope and aid design inhibitors [47].

The first structure of mouse CD14 (residues 5–313) was also recently published, albeit without a bound ligand (PDB code: 1WWL) [48]. The monomeric subunit of CD14 contains 13 β-strands and 11 of them overlap with conserved LRRs. The concave surface of the horseshoe-shaped structure consists of a large β-sheet of 11 parallel and two antiparallel β-strands. The binding sites for LPS in CD14 have been intensively studied by mutagenesis and by epitope mapping of antibodies that block LPS binding, and four regions have been identified within the N-terminal 65 residues of CD14. The most striking feature of the structure of CD14 is the N-terminal pocket that is located on the side of the horseshoe near the N-terminus and it is entirely hydrophobic except for the rim. The main pocket is both wide and deep with dimensions 8 Å wide × 13 Å long × 10 Å deep; thus, overall, the pocket has a total volume of around 820 $Å^3$ and is thus large enough to accommodate at least part of the lipid chains of LPS. The authors propose that LPS would induce only small changes to the CD14 structure and we await the structure of such a complex.

It was thought that the inherent conformational flexibility (and heterogeneity) would prevent crystallization of LPS, but over 10 years ago a breakthrough was made due to the serendipitous cocrystallization of LPS with a protein, FhuA, which is the receptor for ferrichrome-iron (Figure 6.3) [49]. FhuA is found in the outer membrane of *E. coli*, where, in addition to its crucial role in iron metabolism, it also functions as the primary receptor for the structurally related antibiotic albomycin and for several bacteriophages. The authors were primarily interested in the structure and mechanism of action of FhuA, and to facilitate this they overexpressed the FhuA gene in *E. coli* with a surfaced-exposed six His-tag and extracted the protein from the outer membrane C12-DAO as a solubilizing agent. It turned out that crystallization of FhuA is dependent on the presence of stoichiometric amounts of LPS. In fact, if LPS is completely removed from FhuA protein preparations or if an excess of LPS is present in such preparations, the growth of FhuA crystals is inhibited. They propose that LPS remained bound to FhuA throughout the process of purification and crystallization, and that it did not adsorb to FhuA during isolation. Since it is known that LPS is localized to the outer leaflet of the outer membrane, the location of bound LPS marks its position relative to the upper aromatic girdle of FhuA and to the outer membrane. The single LPS molecule that is noncovalently associated with the membrane-embedded outer surface of



Colour fig

Figure 6.3 (a) Side view of FhuA with bound LPS. (b) Top view of same complex.

FhuA exhibits the expected chemical structure for *E. coli* K-12 LPS (PDB code: 1FCP, Figure 6.3).

6.9.2 MD-2/Lipid IVa Complex

Recent structural analyses by Ohto *et al.* have begun to reveal the key features of LPS recognition by essential components of the mammalian innate system. They presented crystal structures of human MD-2 alone and its complex with *E. coli* lipid IVa [40]. Lipid IVa inhibits the effects of bacterial LPS since it is similar in structure to lipid A (it is a lipid A precursor), but lacks two of the six fatty acids present on lipid A that is found in the LPS of *E. coli* or *S. typhimurium*. LPS is recognized by the receptor complex of MD-2 and TLR4. MD-2 is a 160-amino-acid glycoprotein and member of the MD-2-related lipid recognition family that forms a complex with TLR4 on the cell surface [50]. It has been shown that MD-2 and the MD-2/TLR4 complex can bind LPS with nanomolar affinity although LPS can bind to TLR4 alone. The recombinant human MD-2 used in the study was isolated from *Pichia pastoris* and treated with a glycosidase to produce a homogenous, monomeric protein. The structure of MD-2 alone and in complex with lipid IVa revealed that MD-2 displays a deep hydrophobic cavity sandwiched by two β-sheets, in which four acyl chains of the lipid IVa ligand are fully confined (Figure 6.4a and b).

Colour fig

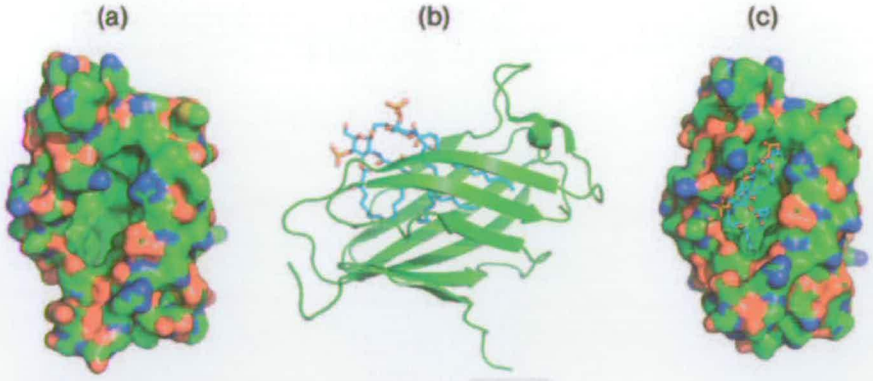


Figure 6.4 (a) Structure of ligand-free MD-2 rendered as a molecular surface showing cavity. (b) Structure of MD-2 with lipid IVa bound with MD-2 displaying β-sheet fold and lipid in chemical structure format. This shows penetration of fatty acids into cavity. (c) Structure of this complex rendered in molecular surface.

The phosphorylated glucosamine disaccharide parts are located at the entrance to the cavity. MD-2 is folded into a single domain consisting of two sets of β-sheets in a common immunoglobulin fold. The structures suggest that MD-2 plays a principal role in LPS recognition. These two sheets (made up of a three and six antiparallel strands) generate a large and deep hydrophobic cavity with a volume
8 big enough to accommodate the LPS molecule (dimensions of $15 \times 8 \times 10$ Å giving a volume of around 1710 Å^3). The ligand-free MD-2 cocrystallized with bound lipid molecules (which were presumed to be $3 \times$ C14 myristic acids). In the MD-2/lipid IVa complex electron densities bound in the cavity were assigned to glucosamine disaccharide, 1- and 4′-phosphates, and four fatty acid chains. Comparison of the complex structure with the "unbound" (remembering that it has myristic acid bound) revealed very little conformational change upon LPS binding apart from a lysine residue at the cavity entrance. Overall, the MD-2 is cationic (p*I* 8.7) and two positively charged residues are at the entrance where they interact with hydrophilic parts of the lipid IVa. This positively charged entrance leads to the hydrophobic cavity (Figure 6.4c) that accommodates the four fatty acid chains. The 1- and 4′-phosphates do not directly interact with MD-2 residues, which is a surprise since these ligands are essential for the inflammatory properties of LPS. MD-2 residues that are essential to the interaction with the TLR4 receptor are located at the cavity entrance. The lipid IVa-binding cavity of MD-2 shows some similarity to the larger hydrophobic cavity of CD14 [48], the protein that transfers LPS to MD-2. Of interest, it appears that CD14 can accommodate lipid A (with six fatty acid chains) due to this expanded cavity, whereas MD-2 cannot. Although lipid IVa is not a potent immunostimulatory molecule since it lacks the six fatty acids of lipid A, the MD-2/lipid IVa complex provides a useful insight into the molecular details of the immune response and begin to answer the question "How does the mammalian

immune response recognize such diverse LPSs from the thousands of bacterial species?".

In a complementary paper in the same issue of *Science*, Mata-Haro *et al.* also investigate the molecular basis of the interactions between LPS and its receptors [51]. The exact molecular details are not yet known as to why certain lipid molecules trigger an inflammatory response as opposed to others that do not. They report that, in mice, MPL interacts with the TLR4–MD-2 complex differently from LPS – they suggest that it is likely caused by the active suppression, rather than passive loss, of proinflammatory activity of this LPS derivative; it is thought that it only activates one part of the inflammatory signal cascade [TRIF and TRAM (TRIF-related adaptor molecule)] and results in T cell activation. They also suggest that MPL could be used as an adjuvant to stimulate the immune response but not cause adverse inflammation.

6.9.3 TLR4–MD-2 Complex

Another recent breakthrough in the understanding of the molecular interactions between the partners involved in LPS detection has recently been published. Eritoran is a synthetic lipid A analog based on the LPS analog from *Rhodobacter sphaerodies*. It is a strong antagonist of the TLR4–MD-2 complex and is currently in phase III clinical trails for severe septic shock. The structure of the full-length ectodomain of the mouse TLR4 alone and in complex with MD-2 was determined using protein expressed in insect cells [52]. The TLR4–MD-2 form a 1 : 1 complex in solution and in the crystals. The crystal structure shows that TLR4 is an unusual member of the "typical" subfamily of the LRR superfamily with a characteristic horseshoe-like structure whose concave surface is formed by parallel β-strands and whose convex surface is formed by loops and 3_{10} helices. Unlike other typical family members, analysis of the β-sheet conformation of TLR4 demonstrates that it can be divided into N-, central and C-terminal domains and undergoes sharp structural transitions at the domain boundaries. The MD-2 adopts a β-cup fold with two antiparallel β-sheets that contain three and six β-strands, respectively, similar to that found for the human MD-2. In comparison to the human protein the mouse MD-2 pocket is narrow and deep with a total surface area of around 1000 $Å^2$ that could accommodate LPS. The structure of the complex provided insight into the surface of TLR4 that interacts with MD-2 – it has a long and narrow shape with dimensions 40 × 20 Å. It can be divided into two chemically and evolutionarily distinct areas – the A and B patches. The A patch is negatively charged and evolutionarily conserved across other species, whereas the B patch is positively charged and located in a less-conserved area, although the residues directly interacting with MD-2 are strictly conserved. The A and B patches of TLR4 are composed of the residues in the concave surface derived from the "LxLxxN" part of the LRR modules in the N-terminal domain and of the central domain, respectively The interaction between TLR4 and MD-2 is also mediated by an extensive network of charge-enhanced hydrogen bonds.

Colour fig

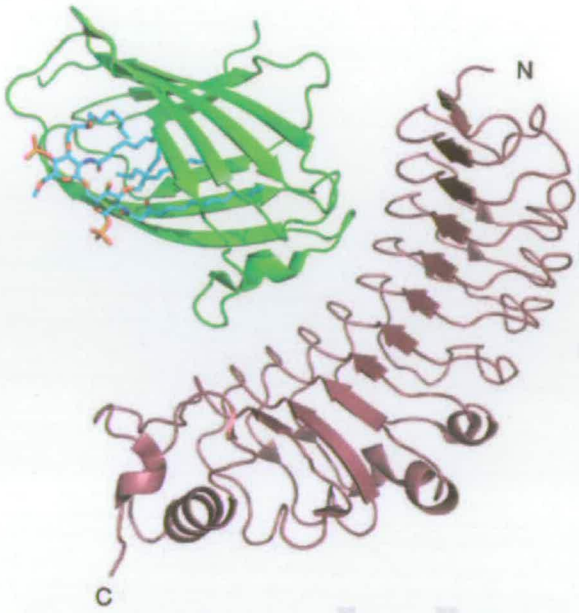


Figure 6.5 The TLR4–MD-2–Eritoran complex. The TLR4 is in red, the MD-2 in green, and Eritoran in ball and stick. The TLR4 contains a TV3 hybrid domain at the C-terminus.

To facilitate soluble expression and crystallization of the TLR4–MD-2 complex with bound ligands, the authors also developed a novel technique that they termed the "hybrid LRR technique". They produced fusion proteins – the variable lymphocyte receptor (VLR) proteins of hagfish were chosen as fusion partners because all VLR proteins have canonical LRR structures with sequence diversity in their variable regions maximized. Among the hybrids, MD-2-bound TV3 was successfully crystallized with Eritoran and their complex structure was solved. They found that Eritoran binds to the hydrophobic pocket in human MD-2, and that there is no direct interaction between Eritoran and TLR4. The structure formed by the four acyl chains of Eritoran complements the shape of the hydrophobic pocket and is a very good fit since it occupies almost 90% of the solvent-accessible volume of the pocket, leaving only a narrow groove near its opening (PDB codes: 2Z65 and 2Z64) (Figure 6.5). They used their fusion structures to also assemble a model of the human TLR4 protein and also used biochemical studies to study cross-linking of the complexes. This allowed them to propose three models for LPS-induced dimerization of the TLR4–MD-2 complex and favor one that fits with the biochemical data. Although Eritoran has a clear structural resemblance to LPS, it is an antagonist, whereas LPS is an agonist. Therefore, there must be significant differences in their modes of interaction. The molecular details of these differences await further structural analysis.

WILEY-VCH

6.10
Bacterial LPS-Binding Proteins

As well as components of mammalian LPS detection and signaling, another area of intensive research is to understand the molecular details of processing of the lipid A and LPS intermediates during the bacterial biosynthetic pathway described by Raetz. Recent breakthroughs have come in the understanding how the lipid A is transported from where it is synthesized in the cytoplasm through the inner to the outer membrane in bacteria. MsbA is a member of the ABC transporter superfamily and is a lipid flippase that transports lipid A and LPS from the cytoplasmic leaflet (inward-facing) to the periplasmic leaflet (outward-facing) of the inner membrane. The structures of the MsbA from three different bacteria (*E. coli*, *V. cholerae* and *S. typhimurium*) in four various open and closed conformational states has begun to reveal the how bacteria process the LPS intermediates [53]. More recently the structure of a bacterial periplasmic transporter LptA has been determined – this protein is believed to act as a LPS shuttle (with MsbA, LptB and Imp/RlpB) between the inner and outer membranes [54].

6.11
Sphingolipids – Essential Membrane Components of Mammals, Plants, Fungi, Yeast and Bacteria

The sphingolipids were discovered over 100 years ago by Johann Thudichum who described an enigmatic "Sphinx-like" molecule, and since then numerous studies of their structure, function and metabolism have been described in the literature [55]. It was not until the middle of the last century that the chemical structure of the family of sphingolipids was determined by Carter *et al.* [56]. They represent a large family of bioactive molecules, and are ubiquitous constituents of all eukaryotic and some prokaryotic membranes. In association with cholesterol, they form lipid rafts that are implicated in signal transduction and membrane trafficking [57]. Their metabolites (such as ceramides) are also known to be involved in several cellular events such as proliferation, differentiation and apoptosis [58, 59]. Furthermore, sphingolipid metabolites play a crucial action in various pathological processes including tumor cell angiogenesis [60]. For example, sphingosine-1-phosphate (S1P) is a potent signaling molecule in eukaryotic cells and a focus of intense research since it has have been linked to various human diseases [61, 62]. The sphingolipids consists of a long-chain base (LCB), fatty acid and a polar head group. This amino alcohol core is decorated with a large variety of functionality and hundreds of sphingolipids are now known.

Sphingolipids are found in several species and so assume that all of these species possess their own sphingolipid biosynthetic pathway (Figure 6.6). Mammalian, plant, yeast, fruit-fly, parasite and fungal sphingolipid pathways are well described, and many of the intermediates, genes and encoded enzymes have been

c06.indd 153 3/30/2009 3:47:17 PM

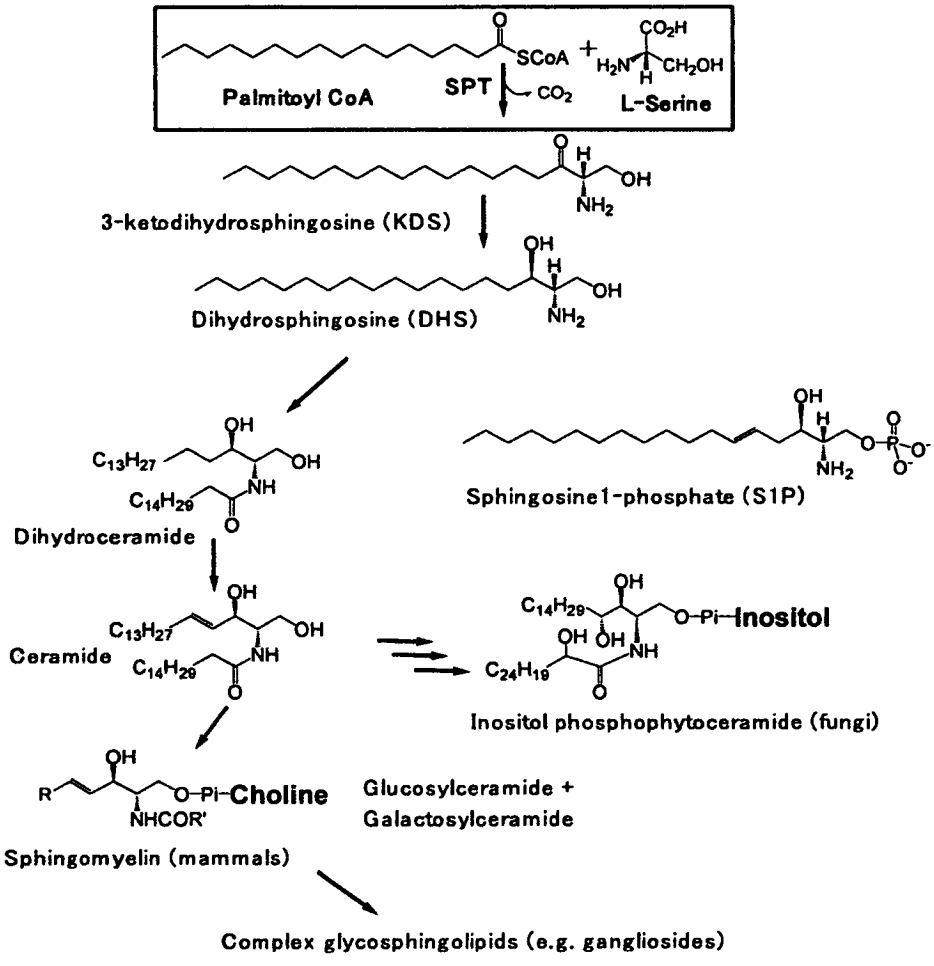


Figure 6.6 Abbreviated sphingolipid biosynthetic pathway showing the early steps and key molecules such as S1P.

identified. Comprehensive reviews of the yeast and plant pathways were recently published [63, 64]. Thematic reviews have recently been published in the *Journal of Lipid Research* (from May 2008 onwards) in honor of Professor Carter and also in *Nature Reviews Molecular Cell Biology* (February 2008).

In summary, the early steps of sphingolipid biosynthesis appear to be highly conserved and the species-specific features occur after attachment of various *N*-acyl chains, hydroxylation at specific carbons and addition of complex sugars. Of intense recent interest has been the characterization of an unusual, growing family of Gram-negative bacteria that have an unusual membrane composition – it was discovered that instead of the LPS described above for *E. coli*, their membranes consist of a variety of glycosphingolipids (GSLs; see below) [65].

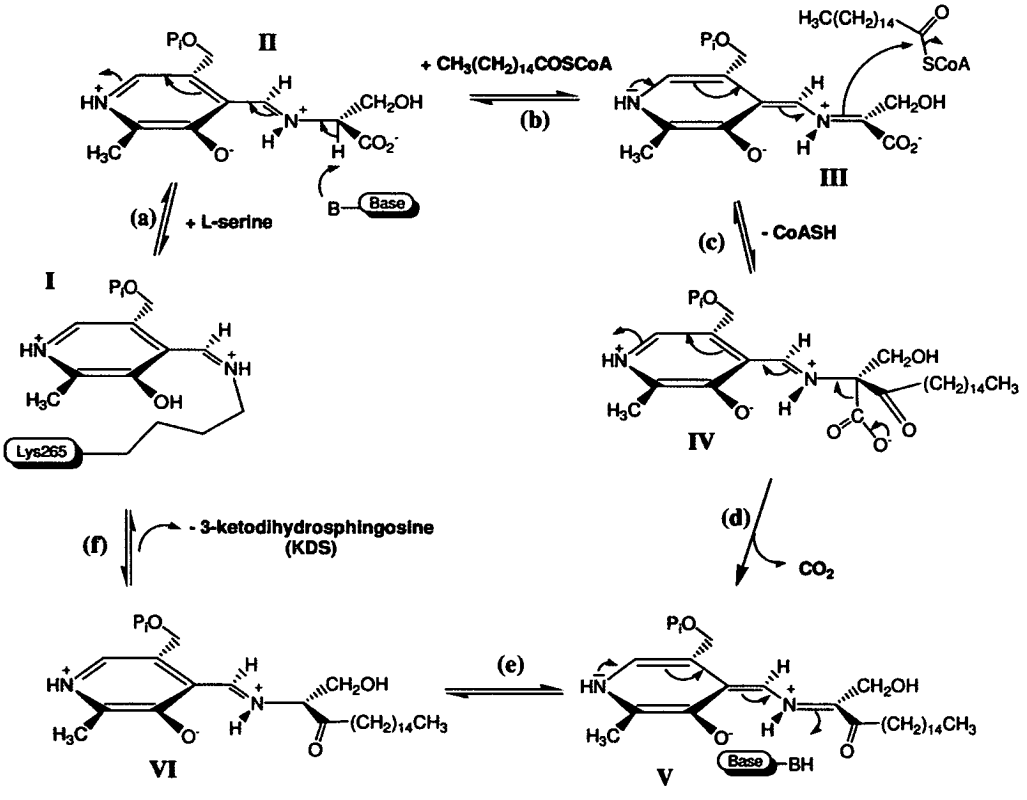


Figure 6.7 Abbreviated reaction mechanism of SPT.

The "Sphingomonads" include organisms such as *Sphingomonas paucimobilis*, *Sphingobacterium multivorum* and *Sphingobium yanoikuyae* [66]. Since bacterial GSLs have similar core structures to their eukaryotic homologs it was assumed that the pathway would begin with the production of the common intermediate, 3-ketodihydrosphingosine (KDS) from L-serine and palmitoyl-CoA. This reaction is catalyzed by the enzyme serine palmitoyltransferase (SPT; EC 2.3.1.50). Indeed, a bacterial SPT was isolated from *S. paucimobilis* by Ikushiro *et al.* in 2001 [67].

SPT catalyses the first and rate-limiting step of the sphingolipid biosynthetic pathway in all organisms studied to date [68]. The reaction is a pyridoxal-5′-phosphate (PLP)-dependent, decarboxylative, Claisen condensation of the amino acid L-serine and the long chain (C16) fatty acid palmitoyl-CoA, which produces KDS. SPT is a member of the α-oxoamine synthase subfamily of PLP-dependent enzymes, which contains three other enzymes: 8-amino-7-oxononanoate synthase, 5-aminolevulinate synthase and 2-amino-3-ketobutyrate-CoA ligase [69–74]. These enzymes generally catalyze the Claisen-like condensation between an amino acid and a CoA-thioester [75]. A common mechanism has been suggested [68, 74, 76–79] comprising the following steps (Figure 6.7): (a) formation of an external

aldimine (II) via displacement of the lysine-PLP internal aldimine (I, holo-SPT) by the incoming amino acid substrate; (b) formation of a quinonoid intermediate (III) by abstraction of α-proton from the PLP-amino acid external aldimine; (c) a Claisen condensation with the fatty acid-CoA substrate followed by displacement of the CoA to form a β-ketoacid (IV); (d) decarboxylation of this species to form a product quinonoid (V); (e) protonation of this quinonoid to form the product external aldimine (VI); and, finally (f) release of the α-oxoamine product and regeneration of the enzyme PLP-internal aldimine. The enzyme is assumed to go through a number of conformational changes as it proceeds through the reaction cycle of C–C formation, C–S bond cleavage and decarboxylation. More recent evidence for the reaction sequence shown was obtained using the elegant use of a CoA thioether analog, *S*-(2-oxoheptadecyl)-CoA. L-Serine deprotonation was accelerated more than 100-fold when this CoA substrate was added, suggesting remarkable substrate synergism [80].

SPT isozymes have been discovered in numerous organisms, such as mammals, fungi, yeast, bacteria and viruses, and sequence analysis of the encoded SPTs from various organisms suggests structural diversity across different isozymes [68]. For example, in humans and the yeast *Saccharomyces cerevisiae* the SPT is formed by two different subunits (SPT1 and SPT2, encoded by the genes *lcb*1 and *lcb*2, respectively), both of which are required for a functionally active enzyme. These two eukaryotic SPTs are predicted to contain large hydrophobic domains at their N- and C-termini, and are membrane-bound proteins, which has made their isolation in high yield and purity particularly difficult. Sequence analysis revealed the SPT1 subunit to be devoid of active site residues, suggesting the SPT1/SPT2 dimer has only one active site, with the SPT1 subunit playing a regulatory role. The recent identification of a LCB3 gene expressed mainly in placenta (with homology to LCB2) suggests greater complexity to this important first enzyme in the pathway [81]. A viral SPT was also recently discovered where it appears that the LCB2 and LCB1 genes are fused head-to-tail to give a large SPT enzyme [82, 83].

Fortunately, a soluble homodimeric enzyme was isolated from *S. paucimobilis* –the structure of this enzyme was determined in our group (the first high-resolution structure of an enzyme from the sphingolipid biosynthetic pathway) and provides a useful tool for structural and mechanistic studies of the SPTs [67, 84]. The bacterial SPT consists of two monomers of 45 kDa and possesses two active sites at the subunit interface (Figure 6.8a).The cofactor PLP is bound into the active site through the formation of a Schiff base with a specific lysine residue (Lys265). The internal aldimine and other intermediates are stabilized by a group of specific catalytic residues. These residues (His159, Asp231 and His234) interact with PLP via π–π stacking, hydrogen bonding or salt bridge (Figure 6.8b). Each monomer is able to bind a molecule of PLP, but the two subunits are required to obtain the active enzyme.

The structure of the bacterial enzyme allowed a model of the human SPT1/SPT2 heterodimer to be built using powerful bioinformatics methods. This model gave molecular insight into the cause of heredity sensory and autonomic neuropathy

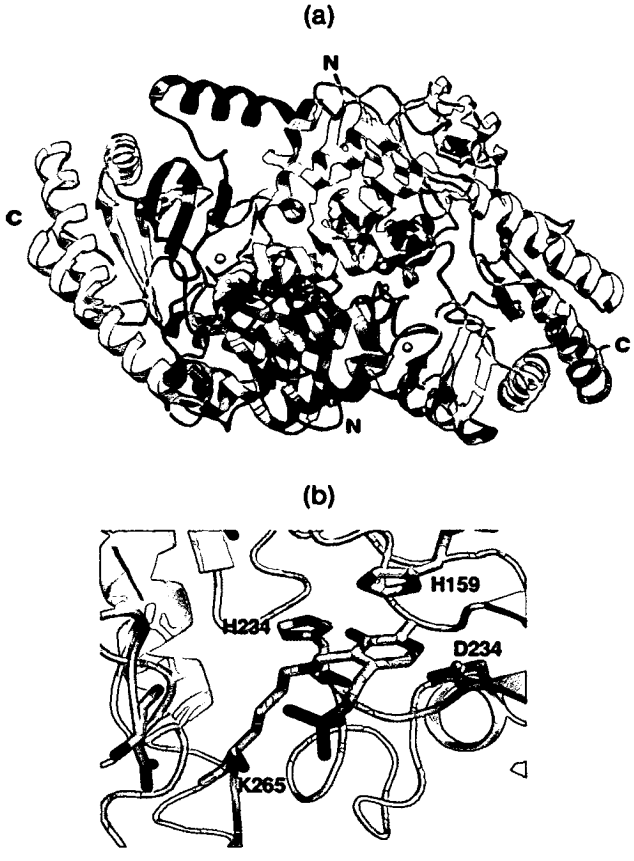


Figure 6.8 Structure of *S. paucimobilis* SPT. (a) Overall dimeric structure. (b) Catalytic residues at the PLP-binding site.

type I, a rare genetic disease that leads to progressive loss of peripheral nerve function. Mutations causing this disease are clustered on the *lcb1* gene and encoded SPT1 subunit, which is surprising since this subunit lacks catalytic residues [85, 86]. Studies of how these mutations impact SPT activity and sphingolipid metabolism are underway in a number of laboratories.

Sphingolipid biosynthesis is a multistep pathway, with many species-specific branch points catalyzed by specific enzymes, so inhibitors of each step have the potential to be powerful drugs. A range of natural product inhibitors are known to inhibit the sphingolipid biosynthetic pathway. Myriocin inhibits directly SPT since it is thought to perfectly mimic the intermediate in the catalytic reaction between L-serine and palmitoyl-CoA. L-Cycloserine has also been shown to inhibit SPTs prokaryotes and eukaryotes. Other enzymes in the pathway from other organisms are also inhibited by interesting natural products–for example

fumonisins inhibit ceramide synthase and inositol phosphorylceramide synthase (AUR1) is a target for the peptide aureobasidin in fungal cells. In fact, this pathway is a good antifungal target in general [87]. The regulation of S1P metabolism in mammals is also a target for a number of potential therapies.

6.12
Natural Killer T Cells and GSLs

As well as playing roles as membrane building blocks and signaling molecules, GSLs have recently been shown to be potent activators of the mammalian immune system via specific antigen receptors. The mechanism is somewhat different to the TLR4–LPS interaction. Natural killer T (NKT) cells are a sublineage of natural memory T cells that share properties of both T lymphocytes that were discovered over 10 years ago. Excellent, comprehensive reviews of the NKT field have been published [88, 89]. GSLs are presented by CD1d molecules on antigen-presenting cells (e.g., dendritic cells) to NKT T cell receptors (TCRs), leading to production of several cytokines including IFN-γ [T helper type (T_h1) response] and IL-4 [T helper type 2 (T_h2) response] (Figure 6.9). CD1d is part of the CD1 family, major histocompatibility complex-like molecules that evolved to capture lipid antigens for display at the surface of antigen presenting cells. CD1 genes encode five CD1 molecules in human (CD1a, CD1b, CD1c, CD1d and CD1e) and two homologs in mice (CD1d1 and CD1d2). CD1d, involved in GSL transport, associate with conserved semi-invariant human $V_\alpha24$–$J_\alpha18/V_\beta8$ (V = "variable" and J = "joining") or mouse $V_\alpha14$–$J_\alpha18/V_\beta8$ TCR.

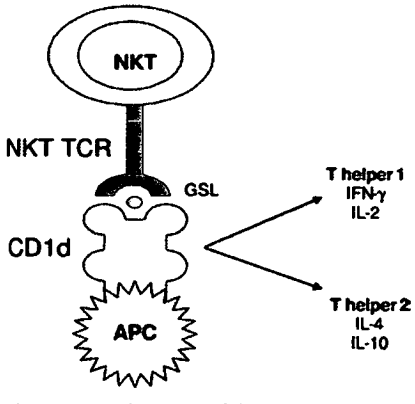


Figure 6.9 Schematic of the interaction between the CD1d receptor found on antigen-presenting cells (APC, e.g., dendritic cells), with presented antigen (GSL) and the NKT TCR, and examples of released inflammatory cytokines.

6.13
GSL Antigens

Screening for products with antitumor activities, Kirin Pharmaceuticals identified the first ligand for the NKT TCR. This was a natural product (Agelasphin 9b,
10 Figure 6.5a) extracted from the marine sponge *Agelas mauritianu* and structural studies of the molecule identified it as a GSL [90, 91]. This compound was unusual due to its α-glycosidic linkage instead of the β-linkage present in mammalian glycolipids. In order to increase its potency, structure–activity relationship studies were carried out on a range of agelasphin derivatives. They simplified the sphinganine chain and added two more carbons on the acyl chain. After modification of the saccharidic hydroxyl groups they found that the 4′-OH played a minor role and that the 3′-OH was necessary for optimal activity. Extension of the *N*-acyl fatty acid chain led to the final compound named KRN7000 or α-galactosylceramide (α-GalCer) (Figure 6.10) [92]. This GSL is still one of the most potent agonists known for NKT TCR. A number of synthetic GSL variants were also made (including both mono- and disaccharides) to begin to explore the nature of the receptor–ligand interaction [93]. Since then extensive synthetic
11 libraries have probed the molecular details to an even greater extent (see below). Very recently, Chung *et al.* explored the stereoselectivity involved by the synthesis of each of all eight stereoisomers of KRN7000 and presentation of their biological activities [94].

The fact that α-GalCer was extracted from a marine sponge raised doubts about its physiological significance and it seems illogical that the mammalian immune system would evolve a defense mechanism against a product from a marine species. It is interesting to speculate that since the sponge *Algea mauritianus* could have been contaminated by α-proteobacteria, the original glycolipid α-GalCer isolated could be in fact of bacterial origin from a symbiotic organism. Numerous studies searched for other natural product antigens for TCRs, which led to the

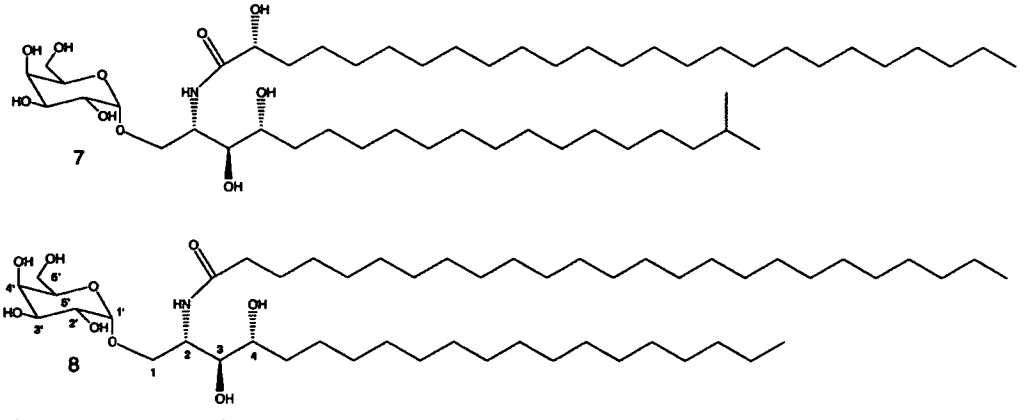


Figure 6.10 Structure of agelasphin 9b (**7**) and α-GalCer (KRN7000, **8**).

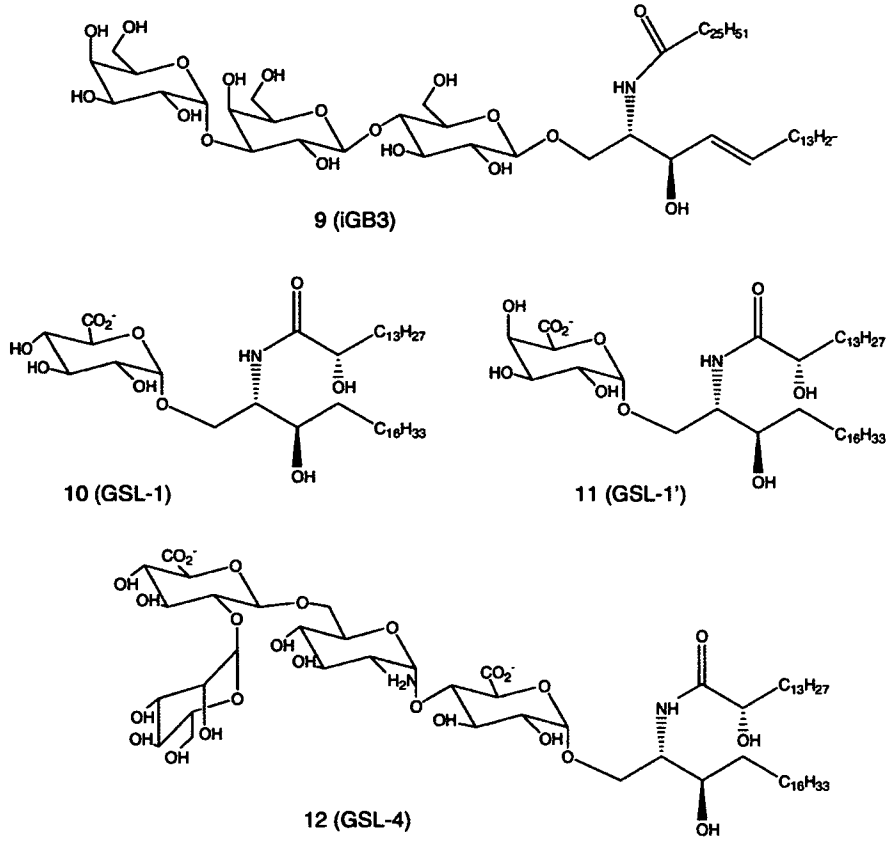


Figure 6.11 Structures of endogenous and exogenous ligands **9** (iGB3) and GSLs of bacterial origin (**10–12**).

discovery of an endogenous ligand, isoglobotrihexosyl ceramide (iGB3), as well as exogenous ligands from bacteria (Figure 6.11).

iGB3 was identified as the endogenous ligand by various studies on mutant mice. The first study by Stanic *et al.* concerned mice deficient in β-D-galactosylceramide (β-D-GalCer) or β-D-glucosylceramide (β-D-GlcCer) [95]. Surprisingly, β-D-GalCer-deficient mice (produced by knocking out β-D-GalCer synthase) developed normal NKT cells and were able to activate NKT hybridomas. In contrast, β-D-GlcCer-deficient cells failed to react with the same hybridomas and the addition of β-D-GlcCer synthase restored the recognition. However, direct addition of β-D-GlcCer to NKT TCR did not result in the activation of the receptor. Work by Zhou *et al.* showed that mice deficient in β-hexosaminidase B were also deficient in NKT cells. β-Hexosaminidases B are responsible for degradation of GB4 and iGB4 to GB3 and iGB3 respectively. NKT cell stimulatory properties were tested with the two last molecules and only iGB3 led to a positive stimulation [96].

These analyses of self/nonself stimulators greatly expanded using the fact that sphingomonad-derived GSLs and mycobacterial lipids had been shown to activate

the immune system via NKT cells. Lipids from various species and strains have been tested for example, the mycobacterial lipid phosphatidylinositol mannoside was reported by Fischer *et al.* to stimulate a modest cytokine release from NKT cells [97]. More efficient NKT activators isolated from *Sphingomonas* possesses GSLs with mono-, di-, tri- and tetrasaccharide units. They present an α-glycosidic linkage similar to the α-GalCer isolated in *A. mauritianus*. GSL-1 can contain galactose or a glucose sugar and unlike α-GalCer can also have a carboxylic acid functional group on the saccharidic C6 instead of a hydroxyl group. GSL-1 and GSL-4 (Figure 6.11, compounds **10** and **12**) have shown high stimulatory activity against NKT TCR.

6.14
Structure–Activity Relationships

Since it now known that natural endogenous and exogenous GSLs are able to activate the mammalian immune system, it makes these molecules attractive targets as therapy against numerous diseases. However, the caveat must be that care must be taken to generate a molecule having only therapeutic benefits and no stimulatory side-effects. NKT cells have been implicated in several immune responses, including pathogens, tumors, tissues grafts, allergens and autoantigens. From these observations, several groups have worked on the modification of the α-GalCer agonist in order to increase the affinity/activity of this compound. Furthermore, activation of NKT cells leads to production of two types of biological response. When α-GalCer binds to TCR, IFN-γ (T_h1) and IL-4 (T_h2) are released in the cell. Several studies have shown that specific modification of the GSL chemical structure can lead to specific activation of one or the other response.

The α-GalCer structure can be separated in two parts: the sugar moiety and the lipid chains. Both of them contain functional groups that can be modified and play a role in the interaction with CD receptors and NKT TCRs. Over the years, a multitude of modifications have been tested and the results obtained permit a better understanding of the structure–activity relationship between the GSL ligands and their receptor. An excellent recent review of this growing literature has recently appeared [98]. The α-GalCer designed by Kirin group possesses a galactose sugar moiety, a sphingosine chain ($C_{18}H_{34}$) and acyl chain ($C_{26}H_{51}$) as the lipid part. The hydroxyl groups as well as the glycosidic linkage are numerous points of modification for the polar head. Concerning, the lipid chains, the length, the degree of unsaturation, as well as the two hydroxyls carried by the sphingosine part allow several possibilities of modification.

6.15
Saccharide Modification

A relatively simple change replaces the galactose by a glucose or a mannose [99] (Figure 6.12, compounds **13** and **14**). The anomeric configuration was tested as well, by synthesis of an α-GalCer (compound **15**). It was observed that the glucose

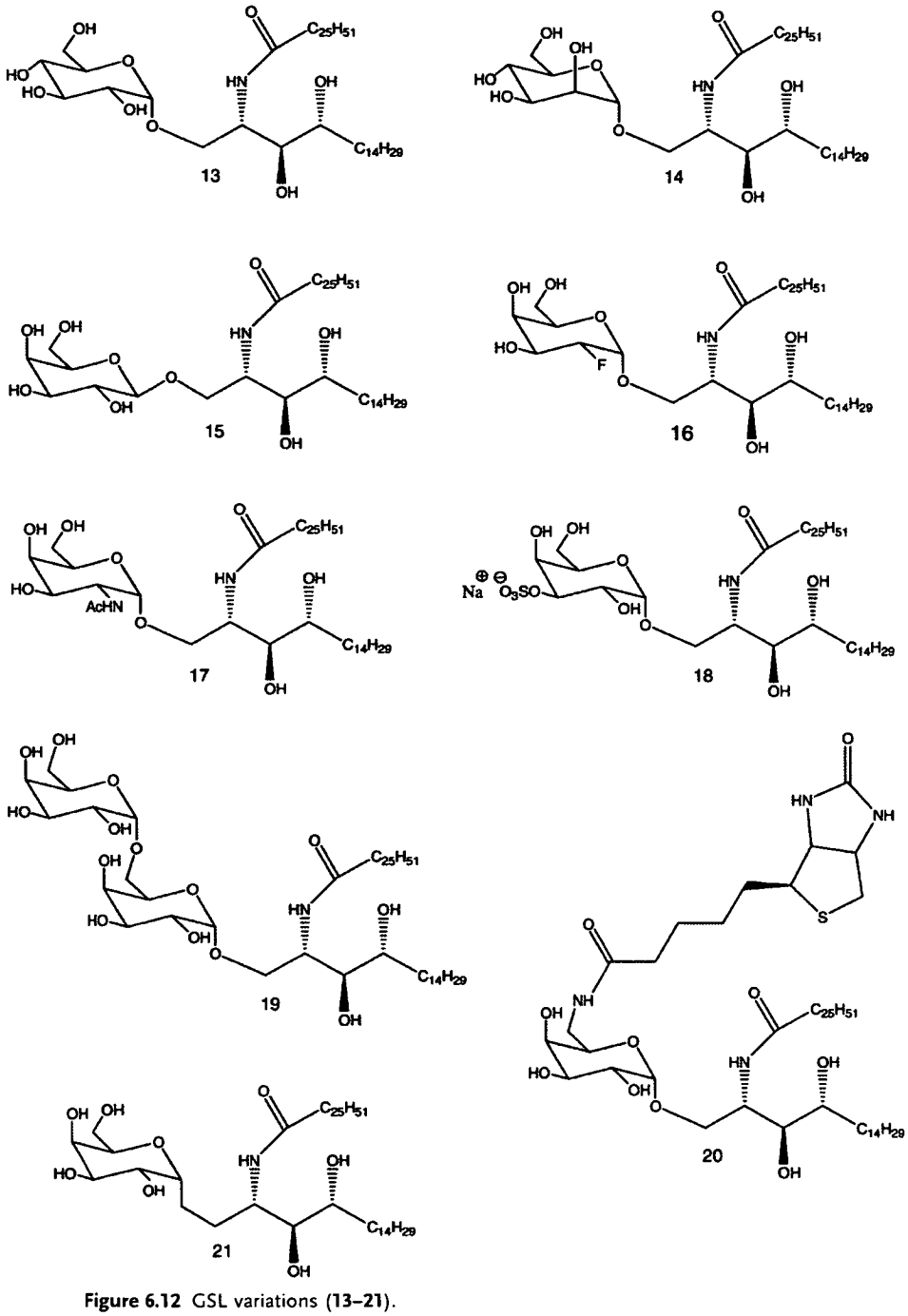


Figure 6.12 GSL variations (**13–21**).

analog has a weak agonist activity, and mannose and β-GalCer analogs have no activity at all. The different OH positions have been investigated and each of them gave rise to different results. The 2′-OH position has been substituted by a diverse range of functional groups (compounds **16** and **17**) and in every case the activity toward TCRs is lost [100, 101].

In contrast, the 3′-OH can accommodate changes such as the sulfate (compound **18**) [101]. The 6′-OH position is quite tolerant of other functional groups (COOH) (Figure 6.11, bacterial GSL-1, **10**), a second sugar (compound **19**) or affinity probes (biotin, compound **20**).The diverse modifications made on the sugar moiety described above affected the recognition with the CD1 receptor part of the complex. The degree of activity was measured against the α-GalCer "standard". No specificity toward T_h1 or T_h2 release was noticed. However, Schmieg *et al.* and Yang *et al.* synthesized two C-glycoside analogs and observed that these compounds were able to stimulate specifically a T_h1 response (compound **21**) [102, 103]. Furthermore, this compound showed an activity 1000 times superior compared with α-GalCer and this increase of potency was explained by the longer stability of the compound in the cells. Effectively, α-GalCer and other O-glycosides can be cleaved by glycosidase which is not the case for C-glycosides.

6.16
Ceramide Modification

In their library of synthetic GSLs, Sidobre *et al.* also probed the 3- and 4-sphingosine positions by substitution with H (Figure 6.13, compounds **22** and **23**) [99]. They reported that they were both involved in CD1d interaction and their substitution leads to destabilization of the complex, with a total loss of activity concerning the substitution of the 3-OH group. Miyamoto *et al.* first examined the effect of lipids chain truncation and synthesized three new compounds [100]. One of them, OCH (compound **24**), induced a T_h2 bias response of NKT cells. Several other truncated α-GalCer derivatives have been synthesized over the years with the addition of small modifications such as insertion of unsaturation [101, 104–106]. Two particular compounds (**25** and **26**) have shown similar behavior than OCH, leading to a specific T_h2 response. It has been suggested that this specificity could be explained by the fact that IFN-γ (T_h1) needs longer stimulation of NKT cells than IL-4 (T_h2) to be released [107]. Shorter chain analogs of α-GalCer or removal of OH groups leads to destabilization and shorter life time of the glycolipids–CD1d complexes, resulting in shorter stimulation of NKT cells.

6.17
High-Resolution Structural Analysis of Receptor–GSL Complexes

Recent breakthroughs in the determination of the structures of the receptor–GSLs complexes has greatly increased our understanding of the nature of the

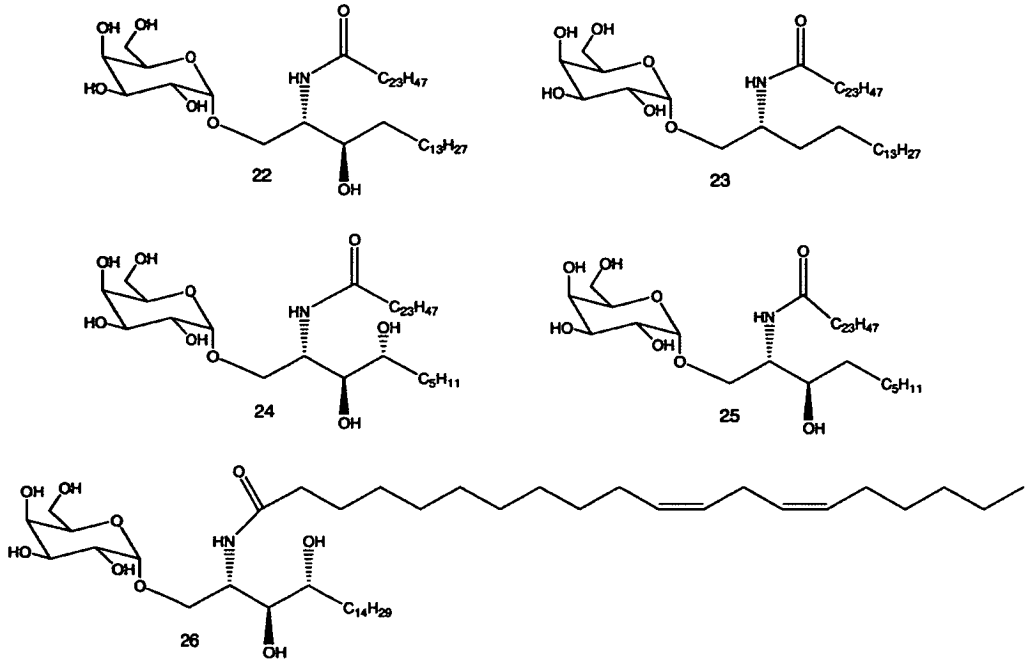
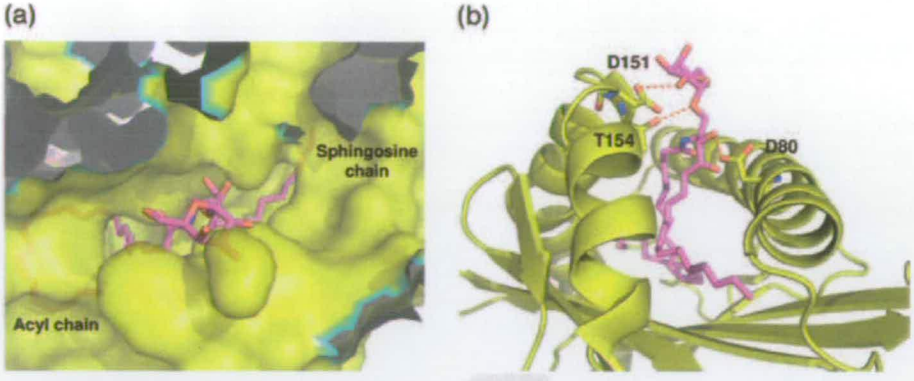


Figure 6.13 Ceramide modifications (22–26).

immunostimulatory properties of the various antigens. Numerous structural analyses have been performed on CD1d and NKT TCR. Both have been studied by alone, as well as with ligand-free, ligand-bound and in complex with each other. The different interactions and recognition systems between the CD1d–TCR and different glycolipids have been investigated by various groups [108–111]. Presented here are the structures of the human CD1d–α-GalCer complex (exogenous ligand, PDB code: 1ZT4) [109] and CD1d–α-GalCer–TCR [111], the mouse CD1d–iGB3 complex (endogenous ligand, PDB code: 2Q7Y) and a model of its interactions with TCR [112].

The human CD1d protein structure was reported in 2005 by Koch *et al.* They discuss the structural changes between the ligand-free and ligand-bound protein, as well as the similarities existing between the mouse and the human CD1d. The CD1d protein consists of two large domains (α and β_2M). Three helices α1, α2 and α3 form the heavy chain (α) of CD1d that is responsible for GSL binding. The antigen-binding site is composed of a large hydrophobic groove allowing the fitting of α-GalCer lipid chains. The acyl chain (C26) binds into pocket A′ in a counter-clockwise circular curve, filling the whole of the pocket. A second channel C′ (human) or F′ (mouse) allows the binding of the sphinganine chain (C18) in a more linear conformation (Figure 6.14a).

The galactose ring is located at the surface of the protein and stabilized by specific interactions. Three hydrogen bonds are involved in the stabilization of the



Colour fig

Figure 6.14 Human CD1d structure. (a) Binding of the GSL. (b) Direct interaction of the GSL 2′-OH with Asp151.

antigen. The 3′OH of the sphingosine chain interacts with Asp80 and the glycol linkage 1′-O forms a hydrogen bond with Thr154. Finally, the galactose ring is correctly oriented for TCR recognition by interaction of the 2′OH with Asp151 (Figure 6.14b). The "ligand-free" CD1d is highly similar to the α-GalCer-bound complex. The main difference concerns the binding groove which is noticeably wider (1 Å difference between the Cα chains) than in the lipid-bound form. It may be possible that CD1d adopts a unique "open" conformation to interact and load lipids. The comparison of human and mouse structures showed that they were highly similar, although some subtle differences in the mouse structure were observed, the ability to specifically bind α-GalCer and its recognition by TCR are not compromised. The ligand-binding pockets are similar in size and hydrophobicity, and residues involved in galactose and alkyl chain binding are conserved.

Recently, Rossjohn *et al.* successfully crystallized the human complex CD1d–α-GalCer–TCR (PDB code: 2PO6) giving a first insight into the GSL, TCR and CD1 interactions [111]. The TCR is a semi-invariant protein that comprises an invariant α-chain and a restricted TCR β-chain repertoire. The NKT TCR adopts an unusual parallel docking mode in its interactions with CD1d (Figure 6.15a). Contacts between the two proteins are made via the V_α chain and the V_β chain of TCR and α1-helix and α2-helix of CD1d. More specifically, two TCR loops, CDR3α and CDR2β, are responsible for the CD1d-restricted response. A number of arginine residues from both partners, Arg79 (CD1d), Arg103 and Arg95 (CDR3α), are involved in a complex hydrogen bond network permitting the specific recognition/activation of NKT TCR by CD1d–α-GalCer. The strong positive charged created by the group of arginines is dissipated by negative residues including Asp94 (CDR3α) and Asp80 (CD1d). Previous mutational studies of these residues on mouse CD1d led to the inactivation of NKT TCR [113–115].

α-GalCer protrudes outside of CD1d, and interacts only with the CDR1α and CDR3α loops of TCR. The galactose ring is sandwiched between Trp153 of CD1d and the aliphatic moiety of Arg95 (CDR3α). Moreover, the guanidinium group of

Colour fig

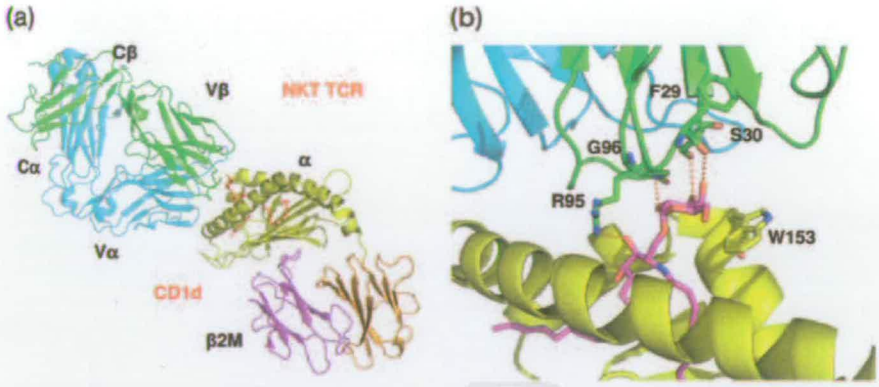


Figure 6.15 (a) Structure of the human NKT TCR–α-GalCer–CD1d complex (NKT TCR: cyan and green; CD1d: yellow, purple and orange; ligand: ball and stick. β2M = β_2-microglobulin. (b) Close-up view of the interface between the proteins and ligand [colors same as in (a)].

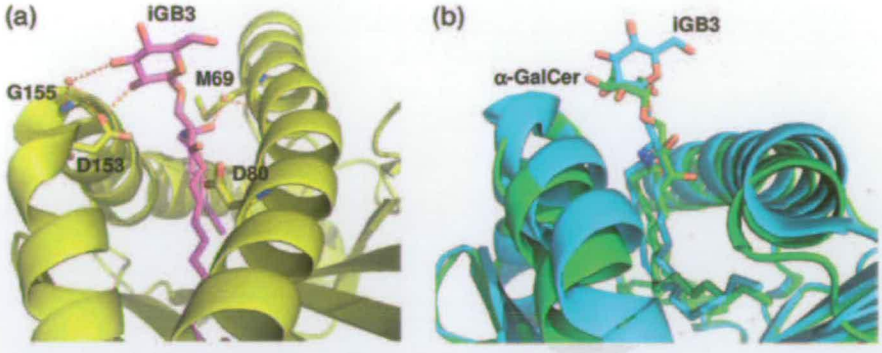
Arg95 is also involved in stabilization of the lipid part by forming a hydrogen bond with the 3′-OH group of the sphingosine chain. The three sugar hydroxyl groups are involved in specific hydrogen bonds with the α invariant α-chain of TCR. The 2′-OH interacts with Gly96α (CDR3α), 3′-OH and 4′-OH have contact with Ser30α and Phe29α (CDR1α). This specific hydrogen bond system is in correlation with α-GalCer high affinity for TCR (Figure 6.15b). The α-glycosidic linkage allows a strategic orientation of the sugar head and the presence of a β-linkage would lead to a more perpendicular orientation creating perturbations in CDR1α interactions.

Curiously, a cross-reactivity between human TCR and mouse CD1d (and the converse) have been observed. Indeed, it was proved that mouse CD1d–α-GalCer were able to bind and activate human NKT TCR and human CD1d–α-GalCer bind to mouse TCR. This cross-reactivity is easily understandable by comparison of both human and mouse α-GalCer–TCR interactions. All residues and interactions involved in antigen recognition are conserved.

Comparison of the TCR structure with and without CD1d–α-GalCer reveals that no large conformational changes takes place. The TCR protein appears to present a rigid "lock and key" conformation, attributable to strong intermolecular interactions between the CDR loops in the ligand-free TCR.

Most recently, Zajonc *et al.* have solved the structures of the mouse CD1d–iGB3 complex and the mouse TCR (PDB codes: 2Q7Y and 2Q86 for CD1d–iGb3 and TCR, respectively). By analogy with the human CD1d–α-GalCer–TCR, they proposed two models to explain the interactions between the endogenous ligand and TCR [112].

The stabilization of the lipid chains in CD1d–iGB3 complex is very similar to that observed in CD1d–α-GalCer. The most differences are, as expected, located at the interface between CD1 and the saccharidic moiety. The polar head group is still solvent exposed and only the proximal glucose is well ordered. The second



Colour fig

Figure 6.16 (a) CD1d and iGB3 specific interactions. (b) Superposition of human CD1d–α-GalCer (green) and mouse CD1d–iGB3 (cyan) structures.

glucose and the galactose are highly flexible, and do not interact with CD1d. The group of residues involved in the first glucose stabilization is totally different to those used for the galactose binding in CD1d–α-GalCer complex. The two hydroxyl groups, 2′-OH and 3′-OH, form H-bonds with Asp153 and Gly155 (water mediated), whereas the lipid chains interact with Asp80 and Met69 (Figure 6.16a).

The TCR crystal obtained by Zajonc *et al.* possesses two TCR molecules in the asymmetric unit. Superposition of the molecules reveals a change in the conformations of CDR loops, especially CDR3α and CDR3β. Furthermore, by alignment of mouse and human TCR, the same conformational differences appear, whereas CDR1 loops stay very similar. As in other TCRs, CDR3 domains are known to interact with the ligand – this flexibility in the unbound structure may reflect their capacity to recognize different ligands.

In order to understand the interaction between the CD1d–iGB3 complex and TCR, Zajonc *et al.* made a model where they positioned the TCR and CD1d in the exact same position as in the human structure with α-GalCer. However, this model reveals serious steric clashes between iGB3 and several TCR residues. This can be explained by the structural differences between iGB3 and α-GalCer (three versus one sugar). From these first results, they proposed two new models to obtain a better fit of iGB3 between the CD1d and TCR.

In the first, they described a "cavity" model where the mouse TCR presents a large and deep positive binding pocket. This cavity formed by the CDR3α and CDR3β loops would be able to fit the two "free" sugars of iGB3. In a second model, the three sugars of iGB3 are positioned between the CD1d and TCR. In this "squashed" model, the parallel orientation of CD1d along TCR as well as the interactions between CD1d and TCR described in the human structure are conserved. Each of these models allows us to probe the different ligand-binding constraints existing in the human structure; however, both of them have limitations. In the "cavity" model, the CDR3β flexibility permits TCR adaptability towards various ligands; however, it is commonly accepted that CDR3β is not

involved in ligand recognition. On the other hand, the "squashed" model requires bond breaking/reforming to orient the polar group properly, which is energetically unfavorable.

Recent studies have begun to shed light on the role of the tetrasaccharide and sphingosine units found in many naturally occurring GSLs derived from Sphingomonads. Kinjo *et al.* made synthetic GSL-4A derivatives varying the nature of the tetrasaccharide, as well as generating GSLs with different lipid chains [116]. Since the initial studies of the effect of GSLs on NKT cells used material isolated from bacteria (probably a heterogeneous mixture) it was therefore important to go back and test homogeneous, synthetic versions of these antigens. In contrast to earlier studies, GSL-4A showed weak activity and the GSL-4B was inactive. Long *et al.* had also used synthetic GSL-1, GSL-2, GSL-3 and GSL-4 (as well as various lipid hybrids) to show that only the monosaccharide GSLs are potent antigens. Furthermore, they probed whether these higher order GSLs could be metabolized by various cell types. It may be that Sphingomonads and other bacteria have evolved a sugar coat to avoid detection by the immune system.

6.18 Conclusions

In both the LPS/endotoxin and GSL fields combinations of natural product chemistry, synthetic chemistry, structural biology, microbiology, immunology and genetics have begun to reveal the molecular interactions involved in the innate recognition between colonizing bacterial cells and their mammalian host. The exact details of what makes one microbe a beneficial symbiotic partner and another a life-threatening pathogen are beginning to emerge. These details are also opening our understanding of the nature of the inflammatory response and its links to a range of autoimmune diseases.

Acknowledgments

M.C.R. is funded by an EaStCHEM PhD studentship. D.J.C.'s laboratory is supported by the Biotechnology and Biological Sciences Research Council and the Engineering and Physical Sciences Research Council.

References

1 Beutler, B. and Rietschel, E.T. (2003) Innate immune sensing and its roots: the story of endotoxin. *Nature Reviews. Immunology*, **3**, 169–76.

2 Galloway, S. and Raetz, C.R.H. (1990) A mutant of *Escherichia coli* defective in the first step of endotoxin biosynthesis. *Journal of Biological Chemistry*, **265**, 6394–402.

3 Tanamoto, K., Zähringer, U., McKenzie, G., Galanos, C., Rietschel, E., Lüderitz, O. *et al.* (1984) Biological activities of synthetic lipid A analogs: pyrogenicity, lethal toxicity, anticomplement activity, and induction of gelation of Limulus amoebocyte lysate. *Infection and Immunity*, **44**, 421–6.

4 Kusumoto, S. and Fukase, K. (2006) Synthesis of endotoxic principle of bacterial lipopolysaccharide and its recognition by the innate immune systems of hosts. *Chemical Record*, **6**, 333–43.

5 Medzhitov, R. (2007) Recognition of microorganisms and activation of the immune response. *Nature*, **819**, 819.

6 Wright, S.D., Ramos, R.A., Tobias, P.S., Ulevitch, R.J. and Mathison, J.C. (1990) CD14, a receptor for complexes of lipopolysaccharide (LPS) and LPS binding protein. *Science*, **249**, 1431–3.

7 Miller, S.I., Ernst, R.K. and Bader, M.W. (2005) LPS, TLR4 and infectious disease diversity. *Nature Reviews. Microbiology*, **3**, 36–46.

8 Poltorak, A., He, X., Smirnova, I., Liu, M.Y., Van Huffel, C., Du, X. *et al.* (1998) Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science*, **282**, 2085–8.

9 Aderem, A. and Ulevitch, R.J. (2000) Toll-like receptors in the induction of the innate immune response. *Nature*, **406**, 782–7.

10 Medzhitov, R., Preston-Hurlburt, P. and Janeway, C.A. (1997) A human homologue of the Drosophila Toll protein signals activation of adaptive immunity. *Nature*, **388**, 394–7.

11 Rock, F.L., Hardiman, G., Timans, J.C., Kastelein, R.A. and Bazan, J.F. (1998) A family of human receptors structurally related to Drosophila Toll. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 588–93.

12 Gay, N.J. and Gangloff, M. (2007) Structure and function of Toll receptors and their ligands. *Annual Review of Biochemistry*, **76**, 141–65.

13 Gay, N.J., Gangloff, M. and Weber, A.N. (2006) Toll-like receptors as molecular switches. *Nature Reviews. Immunology*, **6**, 693–8.

14 O'Neill, L.A. and Bowie, A.G. (2007) The family of five: TIR-domain-containing adaptors in Toll-like receptor signaling. *Nature Reviews. Immunology*, **7**, 353–64.

15 Shizuo Akira, S., Uematsu, S. and Takeuchi, O. (2006) Pathogen recognition and innate immunity. *Cell*, **124**, 783–801.

16 Rice, T.W. and Bernard, G.R. (2005) Therapeutic intervention and targets for sepsis. *Annual Review of Medicine*, **56**, 225–48.

17 Heine, H., Rietschel, E.T. and Ulmer, A.J. (2001) The biology of endotoxin. *Molecular Biotechnology*, **19**, 279–96.

18 Johnson, D. (2008) Synthetic TLR4-active glycolipids as vaccine adjuvants and stand-alone immunotherapeutics. *Current Topics in Medicinal Chemistry*, **8**, 64–79.

19 Raetz, C.R.H. and Whitfield, C. (2002) Lipopolysaccharide endotoxins. *Annual Review of Biochemistry*, **71**, 635–700.

20 Trent, M.S., Stead, C.M., Tran, A.X. and Hankins, J.V. (2006) Diversity of endotoxin and its impact on pathogenesis. *Journal of Endotoxin Research*, **12**, 205–23.

21 Barb, A.W., Jiang, L., Raetz, C.R.H. and Zhou, P. (2007) Structure of the deacetylase LpxC bound to the antibiotic CHIR-090: time-dependent inhibition and specificity in ligand binding. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 18433–8.

22 Meredith, T.C., Aggarwal, P., Mamat, U., Lindner, B. and Woodard, R.W. (2006) Redefining the requisite lipopolysaccharide structure in *Escherichia coli*. *ACS Chemical Biology*, **1**, 33–42.

23 Wang, X., Ribeiro, A., Guan, Z., Abraham, S. and Raetz, C. (2007) Attenuated virulence of a Francisella mutant lacking the lipid A 4′-phosphatase. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 4136–41.

24 Munford, R.S. (2008) Sensing Gram-negative bacterial lipopolysaccharides: a human disease determinant?. *Infection and Immunity*, **76**, 454–65.

25 Peschel, A. and Sahl, H.G. (2006) The co-evolution of host cationic antimicrobial peptides and microbial resistance. *Nature Reviews. Microbiology*, **4**, 529–36.

26 Raetz, C.R.H., Reynolds, C.M., Trent, M.S. and Bishop, R.E. (2007) Lipid A modification systems in Gram-negative bacteria. *Annual Review of Biochemistry*, **76**, 295–329.

27 Prost, L.R. and Miller, S.I. (2008) The *Salmonellae* PhoQ sensor: mechanisms of detection of phagosome signals. *Cellular Microbiology*, **10**, 576–82.

28 De Soyza, A., Silipo, A., Lanzetta, R., Govan, J. and Molinaro, A. (2008) Review: chemical and biological features of Burkholderia cepacia complex lipopolysaccharides. *Innate Immunity*, **14**, 127–44.

29 Silipo, A., Molinaro, A., Ieranò, T., De Soyza, A., Sturiale, L., Garozzo, D. *et al.* (2006) The complete structure and pro-inflammatory activity of the lipooligosaccharide of the highly epidemic and virulent Gram-negative bacterium Burkholderia cenocepacia ET-12 (strain J2315). *Chemistry–A European Journal*, **13**, 3501–11.

30 Shimomura, H., Matsuura, M., Saito, S., Hirai, Y., Isshiki, Y. and Kawahara, K. (2003) Unusual interaction of a lipopolysaccharide isolated from Burkholderia cepacia with polymyxin B. *Infection and Immunity*, **71**, 5225–30.

31 Ortega, X.P., Cardona, S.T., Brown, A.R., Loutet, S.A., Flannagan, R.S., Campopiano, D.J. *et al.* (2007) A putative gene cluster for aminoarabinose biosynthesis is essential for Burkholderia cenocepacia viability. *Journal of Bacteriology*, **189**, 3639–44.

32 Schmelzer, K., Fahy, E., Subramaniam, S. and Dennis, E. (2007) The lipid maps initiative in lipidomics. *Methods in Enzymology*, **432**, 171–83.

33 Raetz, C.R.H., Garrett, T.A., Reynolds, C.M., Shaw, W.A., Moore, J.D., Smith, D.C., Jr *et al.* (2006) Kdo2-Lipid A of Escherichia coli, a defined endotoxin that activates macrophages via TLR-4. *Journal of Lipid Research*, **47**, 1097–111.

34 Guy, B. (2007) The perfect mix: recent progress in adjuvant research. *Nature Reviews. Microbiology*, **5**, 505–17.

35 Stöver, A., Da Silva Correia, J., Evans, J., Cluff, C., Elliott, M., Jeffery, E. *et al.* (2003) Structure–activity relationship of synthetic toll-like receptor 4 agonists. *Journal of Biological Chemistry*, **279**, 4440–9.

36 North, S. and Butts, C. (2005) Vaccination with BLP25 liposome vaccine to treat non-small cell lung and prostate cancers. *Expert Review of Vaccines*, **4**, 249–57.

37 Zhang, Y., Gaekwad, J., Wolfert, M.A. and Boons, G.J. (2007) Modulation of innate immune responses with synthetic lipid A derivatives. *Journal of the American Chemical Society*, **129**, 5200–16.

38 Zhang, Y., Gaekwad, J., Wolfert, M.A. and Boons, G.J. (2008) Innate immune responses of synthetic lipid A derivatives of *Neisseria meningitidis*. *Chemistry–A European Journal*, **14**, 558–69.

39 Saitoh, S.I. and Miyake, K. (2006) Mechanism regulating cell surface expression and activation of Toll-like receptor 4. *Chemical Record*, **6**, 311–19.

40 Ohto, U., Fukase, K., Miyake, K. and Satow, Y. (2007) Crystal structures of human MD-2 and its complex with antiendotoxic lipid IVa. *Science*, **316**, 1632–4.

41 Sil, D., Shrestha, A., Kimbrell, M.R., Nguyen, T.B., Adisechan, A.K., Balakrishna, R. *et al.* (2007) Bound to shock: protection from lethal endotoxemic shock by a novel, nontoxic, alkylpolyamine lipopolysaccharide sequestrant. *Antimicrobial Agents and Chemotherapy*, **51**, 2811–19.

42 David, S.A. (2001) Towards a rational development of anti-endotoxin agents: novel approaches to sequestration of bacterial endotoxins with small

molecules. *Journal of Molecular Recognition*, **14**, 370–87.

43 Voss, S., Fischer, R., Jung, G., Wiesmüller, K.H. and Brock, R. (2007) A fluorescence-based synthetic LPS sensor. *Journal of the American Chemical Society*, **129**, 554–61.

44 Leevy, W.M., Lambert, T.N., Johnson, J.R., Morris, J. and Smith, B.D. (2008) Quantum dot probes for bacteria distinguish *Escherichia coli* mutants and permit *in vivo* imaging. *Chemical Communications*, **2008**, 2331–3.

45 Chaby, R. (2004) Lipopolysaccharide-binding molecules: transporters, blockers and sensors. *Cellular and Molecular Life Sciences*, **61**, 1697–713.

46 Beamer, L.J. (2003) Structure of human BPI (bactericidal/permeability-increasing protein) and implications for related proteins. *Biochemical Society Transactions*, **31**, 791–4.

47 Pristovsek, P., Simcic, S., Wraber, B. and Urleb, U. (2005) Structure of a synthetic fragment of the lipopolysaccharide (LPS) binding protein when bound to LPS and design of a peptidic LPS inhibitor. *Journal of Medicinal Chemistry*, **48**, 7911–14.

48 Kim, J., Lee, C., Jin, M., Lee, C., Paik, S., Lee, H. and Lee, J. (2005) Crystal structure of CD14 and its implications for lipopolysaccharide signaling. *Journal of Biological Chemistry*, **280**, 11347–51.

49 Ferguson, A.D., Hofmann, E., Coulton, J.W., Diederichs, K. and Welte, W. (1998) Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. *Science*, **282**, 2215–20.

50 Gangloff, M. and Gay, N.J. (2004) MD-2: the Toll "gatekeeper" in endotoxin signaling. *Trends in Biochemical Sciences*, **29**, 294–300.

51 Mata-Haro, V., Cekic, C., Martin, M., Chilton, P., Casella, C. and Mitchell, T. (2007) The vaccine adjuvant monophosphoryl lipid A as a TRIF-biased agonist of TLR4. *Science*, **316**, 1628–32.

52 Kim, H.M., Park, B.S., Kim, J.I., Kim, S.E., Lee, J., Oh, S.C. *et al.* (2007) Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell*, **130**, 906–17.

53 Ward, A., Reyes, C.L., Yu, J., Roth, C.B. and Chang, G. (2007) Flexibility in the ABC transporter MsbA: alternating access with a twist. *Proceedings of the National Academy of Sciences of the United States of America*, **104**, 19005–10.

54 Suits, M.D., Sperandeo, P., Dehò, G., Polissi, A. and Jia, Z. (2008) Novel structure of the conserved Gram-negative lipopolysaccharide transport protein A and mutagenesis analysis. *Journal of Biological Chemistry*, **380**, 476–88.

55 Thudichum, J.L. (1884). *A Treatise on the Chemical Constitution of the Brain*, Bailliere, Tindall and Cox, London.

56 Carter, H.E., Haines, H.W., Ledyard, W.E. and Norris, W.P. (1947) Biochemistry of the sphingolipids. *Journal of Biological Chemistry*, **169**, 77–82.

57 Simons, K. and Ikonen, E. (1997) Functional lipid rafts in cell membranes. *Nature*, **387**, 569–72.

58 Futerman, A.H. and Hannun, Y.A. (2004) The complex life of simple sphingolipids. *EMBO Reports*, **5**, 777–82.

59 Hannun, Y.A. and Obeid, L.M. (2008) Principles of bioactive lipid signaling: lessons from sphingolipids. *Nature Reviews. Molecular Cell Biology*, **9**, 139–50.

60 Merrill, A.H. (2002) De Novo sphingolipid biosynthesis: a necessary, but dangerous pathway? *Journal of Biological Chemistry*, **277**, 25843–6.

61 Chalfant, C.E. and Spiegel, S. (2005) Sphingosine 1-phosphate and ceramide 1-phosphate: expanding roles in cell signaling. *Journal of Cell Science*, **118**, 4605–412.

62 Spiegel, S. and Milstien, S. (2003) Sphingosine-1-phosphate: an enigmatic signaling lipid. *Molecular and Cellular Biology*, **4**, 397–407.

63 Dickson, R.C., Sumanasekera, C. and Lester, R.L. (2006) Functions and metabolism of sphingolipids in

Saccharomyces cerevisiae. Progress in Lipid Research, **45**, 447–65.

64 Lynch, D.V. and Dunn, T.M. (2004) An introduction to plant sphingolipids and a review of recent advances in understanding their metabolism and function. *New Phytologist*, **161**, 677–702.

65 Kawahara, K., Seydel, U., Matsuura, M., Danbara, H., Rietschel, E.T. and Zahringer, U. (1991) Chemical structure of glycosphingolipids isolated from *Sphingomonas paucimobilis*. *FEBS Letters*, **292**, 107–10.

66 White, D.C., Sutton, S.D. and Ringelberg, D.B. (1996) The genus *Sphingomonas:* physiology and ecology. *Current Opinion in Biotechnology*, **7**, 301–6.

67 Ikushiro, H., Hayashi, H. and Kagamiyama, H. (2001) A water-soluble homodimeric serine palmitoyltransferase from *Sphingomonas paucimobilis* EY2395T strain. Purification, characterization, cloning, and overproduction. *Journal of Biological Chemistry*, **276**, 18249–56.

68 Hanada, K. (2003) Serine palmitoyltransferase, a key enzyme of sphingolipid metabolism. *Biochimica et Biophysica Acta*, **1632**, 16–30.

69 Alexeev, D., Alexeeva, M., Baxter, R.L., Campopiano, D.J., Webster, S.P. and Sawyer, L. (1998) The crystal structure of 8-amino-7-oxononanoate synthase: a bacterial PLP-dependent, acyl-CoA-condensing enzyme. *Journal of Molecular Biology*, **284**, 401–19.

70 Astner, A., Schulze, J.O., van den Heuvel, J., Jahn, D., Schubert, W.D. and Heinz, D.W. (2005) Crystal structure of 5-aminolevulinate synthase, the first enzyme of heme biosynthesis, and its link to XLSA in humans. *EMBO Journal*, **24**, 3166–77.

71 Ferreira, G.C. and Gong, J. (1995) 5-Aminolevulinate synthase and the first step of heme biosynthesis. *Journal of Bioenergetics and Biomembranes*, **27**, 151–9.

72 Jordan, P.M. and Shemin, D. (1972) 5-Aminolevulinic acid synthase, in *The Enzymes*, Vol. **VII**, 3rd edn (ed. P.D. Boyer), Academic Press, New York, pp. 339–56.

73 Schmidt, A., Sivaraman, J., Li, Y., Larocque, R., Barbosa, J.A.R.G., Smith, C. *et al.* (2001) Three-dimensional structure of 2-amino-3-ketobutyrate CoA ligase from *Escherichia coli* complexed with a PLP-substrate intermediate: inferred reaction mechanism. *Biochemistry*, **40**, 5151–60.

74 Webster, S.P., Alexeev, D., Campopiano, D.J., Watt, R.M., Alexeeva, M., Sawyer, L. and Baxter, R.L. (2000) Mechanism of 8-amino-7-oxononanoate synthase: spectroscopic, kinetic, and crystallographic studies. *Biochemistry*, **39**, 516–28.

75 Eliot, A.C. and Kirsch, J.F. (2004) Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annual Review of Biochemistry*, **73**, 383–415.

76 Alexeev, D., Baxter, R.L., Campopiano, D.J., Kerbarh, O., Sawyer, L., Tomczyk, N. *et al.* (2006) Suicide inhibition of α-oxamine synthases: structures of the covalent adducts of 8-amino-7-oxononanoate synthase with trifluoroalanine. *Organic and Biomolecular Chemistry*, **4**, 1209–12.

77 Hunter, G.A. and Ferreira, G.C. (1999) Lysine-313 of 5-aminolevulinate synthase acts as a general base during formation of the quinonoid reaction intermediates. *Biochemistry*, **38**, 3711–18.

78 Kerbarh, O., Campopiano, D.J. and Baxter, R.L. (2006) Mechanism of α-oxoamine synthases: identification of the intermediate Claisen product in the 8-amino-7-oxononanoate synthase reaction. *Chemical Communications*, 60–2.

79 Zaman, Z., Jordan, P.M. and Akhtar, M. (1973) Mechanism and stereochemistry of the 5-aminolaevulinate synthetase reaction. *Biochemical Journal*, **135**, 257–63.

80 Ikushiro, H., Fujii, S., Shiraiwa, Y. and Hayashi, H. (2008) Acceleration of the substrate Cα deprotonation by an analogue of the second substrate palmitoyl-CoA in serine

palmitoyltransferase. *Journal of Biological Chemistry*, **283**, 7542–53.

81 Hornemann, T., Richard, S., Rütti, M.F., Wei, Y. and von Eckardstein, A. (2006) Cloning and initial characterization of a new subunit for mammalian serine-palmitoyltransferase. *Journal of Biological Chemistry*, **281**, 37275–81.

82 Wilson, W.H., Schroeder, D.C., Allen, M.J., Holden, M.T., Parkhill, J., Barrell, B.G. *et al.* (2005) Complete genome sequence and lytic phase transcription profile of a Coccolithovirus. *Science*, **309**, 1090–2.

83 Han, G., Gable, K., Yan, L., Allen, M.J., Wilson, W.H., Moitra, P. *et al.* (2006) Expression of a novel marine viral single-chain serine palmitoyltransferase and construction of yeast and mammalian single-chain chimera. *Journal of Biological Chemistry*, **281**, 39935–42.

84 Yard, B.A., Carter, L.G., Johnson, K.A., Overton, I.M., Dorward, M., Liu, H. *et al.* (2007) The structure of serine palmitoyltransferase; gateway to sphingolipid biosynthesis. *Journal of Molecular Biology*, **370**, 870–86.

85 Bejaoui, K., Wu, C., Scheffler, M.D., Haan, G., Ashby, P., Wu, L. *et al.* (2001) SPTLC1 is mutated in hereditary sensory neuropathy, type 1. *Nature Genetics*, **27**, 261–2.

86 Dawkins, J.L., Hulme, D.J., Brahmbhatt, S.B., Auer-Grumbach, M. and Nicholson, G.A. (2001) Mutations in SPTLC1, encoding serine palmitoyltransferase, long chain base subunit-1, cause hereditary sensory neuropathy type I. *Nature Genetics*, **27**, 309–12.

87 Thevissen, K., Francois, I.E., Aerts, A.M. and Cammue, B.P. (2005) Fungal sphingolipids as targets for the development of selective antifungal therapeutics. *Current Drug Targets*, **6**, 923–8.

88 Kronenberg, M. and Gapin, L. (2002) The unconventional lifestyle of NKT cells. *Nature Reviews. Immunology*, **2**, 557–68.

89 Bendelac, A., Savage, P.B. and Teyton, L. (2007) The biology of NKT cells. *Annual Review of Immunology*, **25**, 297–336.

90 Akimoto, K., Natori, T. and Morita, M. (1993) Synthesis and stereochemistry of Agelasphin-9b. *Tetrahedron Letters*, **34**, 5593–6.

91 Natori, T., Koezuka, Y. and Higa, T. (1993) Agelasphins, novel α-galactosylceramides from the marine sponge Agelas mauritianus. *Tetrahedron Letters*, **34**, 5591–2.

92 Kobayashi, E., Motoki, K., Uchida, T., Fukushima, H. and Koezuka, Y. (1995) KRN7000, a novel immunomodulator, and its antitumor activities. *Oncology Research*, **7**, 529–34.

93 Kawano, T., Cui, J., Koezuka, Y., Toura, I., Kaneko, Y., Motoki, K. *et al.* (1997) CD1d-restricted and TCR-mediated activation of $V_{\alpha}14$ NKT cells by glycosylceramides. *Science*, **278**, 1626–9.

94 Park, J.J., Lee, J.H., Ghosh, S.C., Bricard, G., Venkataswamy, M.M., Porcelli, S.A. and Chung, S.K. (2008) Synthesis of all stereoisomers of KRN7000, the CD1d-binding NKT cell ligand. *Bioorganic and Medicinal Chemistry Letters*, **18**, 3906–9.

95 Stanic, A.K., De Silva, A.D., Park, J.J., Sriram, V., Ichikawa, S., Hirabyashi, Y. *et al.* (2003) Defective presentation of the CD1d1-restricted natural $V_{\alpha}14J_{\alpha}18$ NKT lymphocyte antigen caused by β-D-glucosylceramide synthase deficiency. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 1849–54.

96 Zhou, D., Mattner, J., Cantu, C., 3rd, Schrantz, N., Yin, N., Gao, Y. *et al.* (2004) Lysosomal glycosphingolipid recognition by NKT cells. *Science*, **306**, 1786–9.

97 Fischer, K., Scotet, E., Niemeyer, M., Koebernick, H., Zerrahn, J., Maillet, S. *et al.* (2004) Mycobacterial phosphatidylinositol mannoside is a natural antigen for CD1d-restricted T cells. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 10685–90.

98 Savage, P.B., Teyton, L. and Bendelac, A. (2006) Glycolipids for natural killer T

cells. *Chemical Society Reviews*, **35**, 771–9.

99 Sidobre, S., Hammond, K.J., Benazet-Sidobre, L., Maltsev, S.D., Richardson, S.K., Ndonye, R.M. *et al.* (2004) The T cell antigen receptor expressed by $V_{\alpha}14i$ NKT cells has a unique mode of glycosphingolipid antigen recognition. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 12254–9.

100 Miyamoto, K., Miyake, S. and Yamamura, T. (2001) A synthetic glycolipid prevents autoimmune encephalomyelitis by inducing T_H2 bias of natural killer T cells. *Nature*, **413**, 531–4.

101 Wu, D., Xing, G.W., Poles, M.A., Horowitz, A., Kinjo, Y., Sullivan, B. *et al.* (2005) Bacterial glycolipids and analogs as antigens for CD1d-restricted NKT cells. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 1351–6.

102 Schmieg, J., Yang, G., Franck, R.W. and Tsuji, M. (2003) Superior protection against malaria and melanoma metastases by a C-glycoside analogue of the natural killer T cell ligand α-galactosylceramide. *Journal of Experimental Medicine*, **198**, 1631–41.

103 Yang, G., Schmieg, J., Tsuji, M. and Franck, R.W. (2004) The C-glycoside analogue of the immunostimulant α-galactosylceramide (KRN7000): synthesis and striking enhancement of activity. *Angewandte Chemie (International ed. in English)*, **43**, 3818–22.

104 Goff, R.D., Gao, Y., Mattner, J., Zhou, D., Yin, N., Cantu, C., 3rd *et al.* (2004) Effects of lipid chain lengths in α-galactosylceramides on cytokine release by natural killer T cells. *Journal of the American Chemical Society*, **126**, 13602–3.

105 Ndonye, R.M., Izmirian, D.P., Dunn, M.F., Yu, K.O., Porcelli, S.A., Khurana, A. *et al.* (2005) Synthesis and evaluation of sphinganine analogues of KRN7000 and OCH. *Journal of Organic Chemistry*, **70**, 10260–70.

106 Yu, K.O., Im, J.S., Molano, A., Dutronc, Y., Illarionov, P.A., Forestier, C. *et al.* (2005) Modulation of CD1d-restricted NKT cell responses by using *N*-acyl variants of α-galactosylceramides. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 3383–8.

107 Oki, S., Chiba, A., Yamamura, T. and Miyake, S. (2004) The clinical implication and molecular mechanism of preferential IL-4 production by modified glycolipid-stimulated NKT cells. *Journal of Clinical Investigation*, **113**, 1631–40.

108 Giabbai, B., Sidobre, S., Crispin, M.D., Sanchez-Ruiz, Y., Bachi, A., Kronenberg, M. *et al.* (2005) Crystal structure of mouse CD1d bound to the self ligand phosphatidylcholine: a molecular basis for NKT cell activation. *Journal of Immunology*, **175**, 977–84.

109 Koch, M., Stronge, V.S., Shepherd, D., Gadola, S.D., Mathew, B., Ritter, G. *et al.* (2005) The crystal structure of human CD1d with and without α-galactosylceramide. *Nature Immunology*, **6**, 819–26.

110 Zajonc, D.M., Maricic, I., Wu, D., Halder, R., Roy, K., Wong, C.H. *et al.* (2005) Structural basis for CD1d presentation of a sulfatide derived from myelin and its implications for autoimmunity. *Journal of Experimental Medicine*, **202**, 1517–26.

111 Borg, N.A., Wun, K.S., Kjer-Nielsen, L., Wilce, M.C., Pellicci, D.G., Koh, R. *et al.* (2007) CD1d-lipid-antigen recognition by the semi-invariant NKT T-cell receptor. *Nature*, **448**, 44–9.

112 Zajonc, D.M., Savage, P.B., Bendelac, A., Wilson, I.A. and Teyton, L. (2008) Crystal structures of mouse CD1d–iGb3 complex and its cognate $V_{\alpha}14$ T cell receptor suggest a model for dual recognition of foreign and self glycolipids. *Journal of Molecular Biology*, **377**, 1104–16.

113 Burdin, N., Brossay, L., Degano, M., Iijima, H., Gui, M., Wilson, I.A. and Kronenberg, M. (2000) Structural

requirements for antigen presentation by mouse CD1. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 10156–61.

114 Kamada, N., Iijima, H., Kimura, K., Harada, M., Shimizu, E., Motohashi, S. *et al.* (2001) Crucial amino acid residues of mouse CD1d for glycolipid ligand presentation to $V_{\alpha}14$ NKT cells. *International Immunology*, **13**, 853–61.

115 Sidobre, S., Naidenko, O.V., Sim, B.C., Gascoigne, N.R., Garcia, K.C. and Kronenberg, M. (2002) The $V_{\alpha}14$ NKT cell TCR exhibits high-affinity binding to a glycolipid/CD1d complex. *Journal of Immunology*, **169**, 1340–8.

116 Kinjo, Y., Pei, B., Bufali, S., Raju, R., Richardson, S.K., Imamura, M. *et al.* (2008) Natural sphingomonas glycolipids vary greatly in their ability to activate natural killer T cells. *Chemistry and Biology*, **15**, 654–64.

c06.indd 176 3/30/2009 3:47:33 PM

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2009/06/17/M109.008680.DC1.html

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 284, NO. 25, pp. 17328–17339, June 19, 2009
© 2009 by The American Society for Biochemistry and Molecular Biology, Inc. Printed in the U.S.A.

The External Aldimine Form of Serine Palmitoyltransferase

*STRUCTURAL, KINETIC, AND SPECTROSCOPIC ANALYSIS OF THE WILD-TYPE ENZYME AND HSAN1 MUTANT MIMICS**

Received for publication, March 16, 2009, and in revised form, April 15, 2009 Published, JBC Papers in Press, April 17, 2009, DOI 10.1074/jbc.M109.008680

Marine C. C. Raman[‡1,2], Kenneth A. Johnson[§1], Beverley A. Yard[‡2,3], Jonathan Lowther[‡], Lester G. Carter[§], James H. Naismith[§4], and Dominic J. Campopiano[‡2,3,5]

From [‡]EaStChem, School of Chemistry, University of Edinburgh, Edinburgh EH9 3JJ and [§]EaStChem, Scottish Structural Proteomics Facility, and Centre for Biomolecular Science, University of St. Andrews, Edinburgh KY16 9RH, Scotland, United Kingdom

Sphingolipid biosynthesis begins with the condensation of L-serine and palmitoyl-CoA catalyzed by the PLP-dependent enzyme serine palmitoyltransferase (SPT). Mutations in human SPT cause hereditary sensory autonomic neuropathy type 1, a disease characterized by loss of feeling in extremities and severe pain. The human enzyme is a membrane-bound hetereodimer, and the most common mutations are located in the enzymatically incompetent monomer, suggesting a "dominant" or regulatory effect. The molecular basis of how these mutations perturb SPT activity is subtle and is not simply loss of activity. To further explore the structure and mechanism of SPT, we have studied the homodimeric bacterial enzyme from *Sphingomonas paucimobilis*. We have analyzed two mutants (N100Y and N100W) engineered to mimic the mutations seen in hereditary sensory autonomic neuropathy type 1 as well as a third mutant N100C designed to mimic the wild-type human SPT. The N100C mutant appears fully active, whereas both N100Y and N100W are significantly compromised. The structures of the holoenzymes reveal differences around the active site and in neighboring secondary structure that transmit across the dimeric interface in both N100Y and N100W. Comparison of the L-Ser external aldimine structures of both native and N100Y reveals significant differences that hinder the movement of a catalytically important Arg^{378} residue into the active site. Spectroscopic analysis confirms that both N100Y and N100W mutants subtly affect the chemistry of the PLP. Furthermore, the N100Y and R378A mutants appear less able to stabilize a quinonoid intermediate. These data provide the first experimental insight into how the most common disease-associated mutations of human SPT may lead to perturbation of enzyme activity.

Sphingolipids are ubiquitous constituents of eukaryotic cells, where they play important roles in signaling, differentiation, and apoptosis (1–4). Defects in sphingolipid catabolism have been linked to several human diseases, such as hypertension, cancer, and disorders of the peripheral nervous system. The most common inherited peripheral neuropathy is hereditary sensory autonomic neuropathy type 1 (HSAN1).[6] The disease leads to progressive loss of sensation in extremities and is often associated with searing pain (5–8). Genetic studies by two independent groups mapped the disease-associated mutations to the *lcb1* (long chain base 1) gene on chromosome 9q22, which encodes the SPT1 subunit of serine palmitoyltransferase (SPT; EC 2.3.1.50) (9–11). SPT catalyzes the first and rate-limiting step of the sphingolipid biosynthetic pathway in all organisms studied to date (12). The reaction is a pyridoxal 5′-phosphate (PLP)-dependent, decarboxylative, Claisen condensation of the amino acid L-serine and the long chain (C16) fatty acid palmitoyl-CoA, which produces the sphingolipid precursor, 3-ketodihydrosphingosine (KDS).

SPT belongs to the α-oxoamine synthase subfamily of PLP-dependent enzymes, which contains three other well characterized members: 8-amino-7-oxononanoate synthase (AONS) (13, 14), 5-aminolevulinate synthase (ALAS) (15–17), and 2-amino-3-ketobutyrate-CoA ligase (KBL) (18). These enzymes catalyze the Claisen-like condensation between an amino acid and an acyl-coenzyme A thioester (19). A common mechanism has been proposed (12, 14, 20–23) comprising the following steps: formation of an external aldimine via displacement of the lysine-PLP internal aldimine (holo-SPT) by the incoming amino acid substrate; formation of a quinonoid intermediate by abstraction of the α-proton from the PLP-amino acid external aldimine; a Claisen condensation with the fatty acid-CoA thioester substrate, followed by displacement of the CoASH to form a β-ketoacid intermediate; decarboxylation of this species to form a product quinonoid; protonation of this quinonoid to form the product external aldimine; and finally release of the α-oxoamine product and regeneration of the enzyme PLP-internal aldimine (Fig. 1). Structural studies have shown AONS, ALAS, and KBL isoforms from various species to be

* This work was supported by Biotechnology and Biological Sciences Research Council (BBSRC) Grants BB/F009739/1 and BBS/B/14434 (to D. J. C. and J. H. N.) and the Scottish Structural Proteomics Facility (SSPF).

This paper is dedicated to the memory of Dr. Joe Spencer (University of Cambridge).

The atomic coordinates and structure factors (codes 2W8J, 2W8W, 2W8U, 2W8V, and 2W8T) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] Both of these authors contributed equally to this work.
[2] Supported by the University of Edinburgh, EaStCHEM.
[3] Supported by Syngenta.
[4] To whom correspondence may be addressed. Tel.: 44-133-446-3792; E-mail: naismith@st-andrews.ac.uk.
[5] To whom correspondence may addressed. Tel.: 44-131-650-4712; E-mail: Dominic.Campopiano@ed.ac.uk.

[6] The abbreviations used are: HSAN1, hereditary sensory autonomic neuropathy type 1; SPT, serine palmitoyltransferase; PLP, pyridoxal phosphate; KDS, 3-ketodihydrosphingosine; DTNB, 5,5′-dithiobis-2-nitrobenzoic acid; r.m.s., root mean square; AONS, 8-amino-7-oxononanoate synthase; ALAS, 5-aminolevulinate synthase; KBL, 2-amino-3-ketobutyrate-CoA ligase; PEG, polyethylene glycol.

Downloaded from www.jbc.org at Edinburgh University Library, on September 24, 2009

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2009/06/17/M109.008680.DC1.html

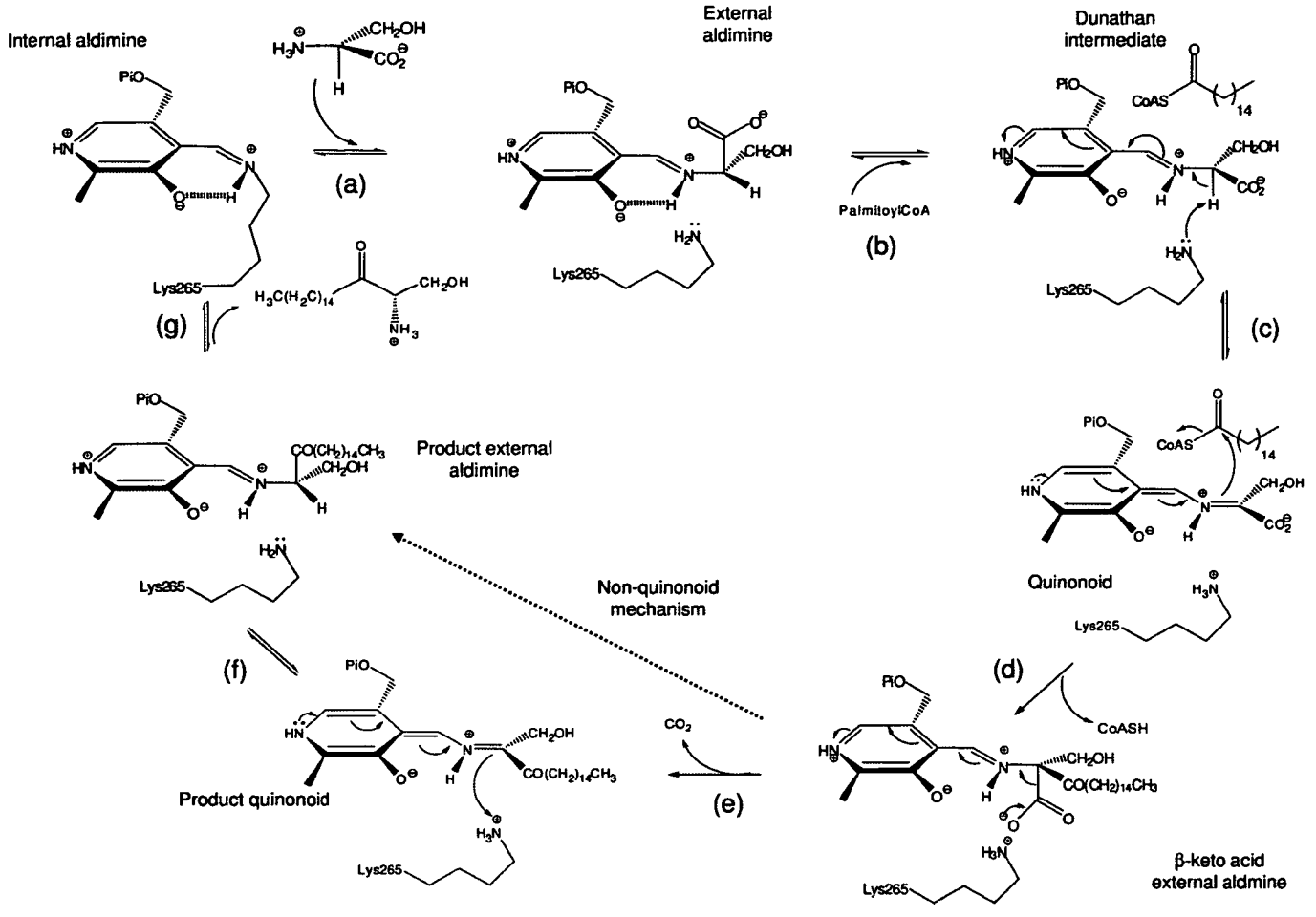


FIGURE 1. **Proposed catalytic mechanism of SPT.** *a*, the internal aldimine (holo-SPT) is displaced by L-serine to form the external adlimine; *b*, binding of second substrate palmitoyl-CoA causes conformational change to give the Dunathan intermediate; *c*, formation of the quinonoid by deprotonation of Cα hydrogen; *d*, thioester bond hydrolysis and release of CoASH to give β-keto acid intermediate; *e*, decarboxylation to form KDS product quinonoid; *f*, reprotonation to form KDS product external aldimine; *g*, KDS product release and reformation of the holo-form. The L-serine external aldimine has been trapped in the current study. The *dotted line* shows that steps *e* and *f* may be bypassed by an alternative mechanism that does not proceed via a product quinonoid (see Ref. 21).

homodimers (13, 17, 18). In marked contrast, eukaryotic SPTs are heterodimeric, membrane-bound proteins consisting of two subunits, SPT1 and SPT2 (encoded by the *lcb1* and *lcb2* gene, respectively) (24).

Analysis of a number of HSAN1 patients has revealed the four most common mutations to be C133W, C133Y, V144D, and G387A, with the cysteine mutations appearing to be the most prevalent in populations (10, 25). The impact that these mutations have on SPT activity and sphingolipid metabolism has been the focus of attention by a number of groups. Most surprisingly, these disease-associated mutations occur in the SPT1 monomer, a protein that must be inactive, since it lacks the key lysine, histidine, and aspartate residues necessary to bind and stabilize the PLP cofactor. It is the SPT2 protein that contains the conserved, active site lysine residue that forms a Schiff base internal aldimine with the PLP. However, both subunits are essential to produce functionally active SPT heterodimer (26–29), indicating that the inactive subunit has a crucial role in function. Dunn and co-workers (30) created several mutations in the yeast *lcb1* and *lcb2* genes, including those corresponding to the HSAN1 SPT1 mutations, and found that they dominantly inhibit SPT activity. A transgenic mouse model also revealed that the LCB1 C133W mutation led to mice with decreased tissue SPT activity and HSAN1 symptoms despite unaltered ceramide concentrations (31). A different study found that SPT activity was decreased in the tissues of HSAN1 patients, and in a Chinese hamster ovary model, hamster LCB1 C133Y and C133W mutations could not rescue cells lacking endogenous LCB1 (32). These combined studies revealed that the HSAN1 mutations act directly on the SPT enzyme but the pathological consequence on sphingolipid metabolism is unclear. Very recently, Dunn and colleagues discovered a small 80-amino acid protein (TSC3p) in yeast that stimulated yeast SPT activity, but its role is still unclear (33). Recently, a third eukaryotic subunit (SPT3) has been characterized and appears to be required for optimum SPT activity, although it is expressed most highly in placental tissue and human trophoblasts (34). It has been suggested that these three subunits could form a higher order SPT complex (35). The instability and hydrophobic nature of eukaryotic SPTs has made their isolation and therefore biophysical characterization particularly challenging (24, 36–38).

Downloaded from www.jbc.org at Edinburgh University Library, on September 24, 2009

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2009/06/17/M109.008680.DC1.html

We have targeted homodimeric SPT from *Sphingomonas paucimobilis* EY2395 to provide molecular insight into the enzymatic properties of SPT (39, 40). Recently, we reported the first high resolution x-ray crystal structure of the *S. paucimobilis* holo-SPT (41) and showed that the active site containing the PLP cofactor is at the dimer interface. We used this structure to model the human enzyme and map the human cysteine residue (Cys^{133}) of SPT1 onto Asn^{100} of the bacterial SPT (41). We found that Asn^{100} is proximal to the PLP binding site and lies at the dimer interface. To explore the effects of the HSAN1 mutations on human SPT activity and structure, we have studied mutants of the bacterial enzyme. We made the N100C mutant to better mimic native human SPT and N100W and N100Y to mimic the most common mutations in human SPT. We have characterized the mutations using kinetics, spectroscopy, and structural biology, including structures of the external L-Ser aldimine. These data provide molecular insights into the effects of HSAN1 mutations. Also, as part of these studies, we identify residues that play a role in the stabilization of intermediates in the SPT reaction and have been able to generate new insights into the substrate specificity and mechanism of the α-oxamine synthase family.

EXPERIMENTAL PROCEDURES

Chemicals—Plasmids and *Escherichia coli* competent cells were purchased from Novagen, and all chromatography columns were from GE Healthcare. All buffers and reagents were from Sigma. Palmitoyl-CoA was from Fluka and Avanti Lipids, and [U-^{14}C]L-serine (specific activity 1.85 MBq/ml) was from GE Healthcare. The QuikChange site-directed mutagenesis kit was purchased from Stratagene. The decanoyl-CoA (10:0), lauroyl-CoA (12:0), myristoyl-CoA (14:0), stearoyl-CoA (18:0), and arachidoyl-CoA (20:0) were from Sigma. Oligonucleotide primers were sourced from SigmaGenosys. PEG 3350 was from Fluka. *S*-(2-oxoheptadecyl)-CoA was a kind gift from Dr. Hiroko Ikushiro (Department of Biochemistry, Osaka Medical College, Japan).

Cloning and Expression of S. paucimobilis SPT N100W, N100Y, N100C, R378A, and R378N Mutants in E. coli—Plasmid pET-28a/SPT WT was used as the template for mutations (N100W, N100Y, N100C, R378A, and R378N) that were introduced into the SPT gene by site-directed mutagenesis using the following mutagenic primers (the mutagenic bases are shown in boldface type) and confirmed by big Dye sequencing: 5′-GGGTCGGGCACC**TGG**GGCAGCCGGATG-3′ (SPT N100W forward), 5′-CATCCGGCTGCC**CCA**GGTGCCCGACCC-3′ (SPT N100W reverse), 5′-GGGTCGGGCACC**TAT**GGCAGCCGGATG-3′ (SPT N100Y forward), 5′-CATCCGGCTGCC**ATA**GGTGCCCGACCC-3′ (SPT N100Y reverse), 5′-GGGTCGGGCACC**TGT**GGCAGCCGGATG-3′ (SPT N100C forward), 5′-CATCCGGCTGCC**ACA**GGTGCCCGACCC-3′ (SPT N100C reverse), 5′-CTCTACGTCAACATGGCG**GCC**CCGCCC-3′ (SPT R378A forward), 5′-GGGCGG**GGC**CGCCATGTTGACGTAGAG-3′ (SPT R378A reverse), 5′-CTCTACGTCAACATGGCG**AAC**CCGCCC-3′ (SPT R378N forward), and 5′-GGGCGG**GTT**CGCCATGTTGACGTAGAG-3′ (SPT R378N reverse).

The recombinant plasmids were used to transform *E. coli* BL21 (DE3) competent cells, and selection was carried out on LB agar containing 30 μg/ml kanamycin. For each mutant (and native), a single colony was used to inoculate 500 ml of 2YT broth (16 g/liter Bacto-tryptone, 10 g/liter Bacto-yeast extract, 5 g/liter sodium chloride (pH 7.5)), which was shaken at 250 rpm overnight at 37 °C. The overnight culture was added to 4 liters of 2YT broth supplemented with kanamycin and grown at 37 °C to an A_{600} of 0.6. Protein expression was induced by the addition of isopropyl 1-thio-β-D-galactopyranoside to give a final concentration of 0.1 mM, and growth was continued for 5 h at 30 °C. Cells were harvested (Sorvall RC5B centrifuge) by centrifugation at 3500 rpm for 20 min at 4 °C.

Purification of SPT WT, N100W, N100Y, N100C, R378A, and R378N—The proteins were purified essentially according to the previously described method, which is based polyhistidine-tagged protein binding to Ni^{2+}-nitrilotriacetic acid-agarose (Qiagen) (41). After elution, fractions were pooled and desalted by dialysis against 20 mM potassium phosphate (pH 7.5), 150 mM NaCl, and 25 μM PLP. A gel filtration step (Sephacryl S200 26/60 HR; GE Healthcare) was carried out in the same buffer. For enzymatic assays and UV-visible spectroscopy, this buffer is also used. For x-ray crystal trials, the protein was exchanged into 10 mM Tris (pH 7.5), 150 mM NaCl, and 25 μM or 250 μM PLP. For storage, the enzymes were transferred to the same buffer containing 20% glycerol (v/v) and stored at −80 °C until use. Protein identity and integrity were confirmed by high pressure liquid chromatography electrospray mass spectrometry on a MicroMass Platform II quadrupole mass spectrometer equipped with an electrospray ion source. The experimentally determined masses of recombinant *S. paucimobilis* SPT WT and each of the mutants were within 0.1% of the theoretical mass, which includes the C-terminal LEHHHHHH fusion affinity tag with the enzyme lacking the N-terminal methionine (41).

Spectroscopic Measurements—All UV-visible spectra were recorded on a Cary 50 UV-visible spectrophotometer (Varian) and analyzed using Cary WinUV software (Varian). Prior to enzymatic assays, SPT was converted to the holo-form by dialysis against freshly prepared 20 mM potassium phosphate (pH 7.5) containing 150 mM NaCl and 25 μM PLP for 1 h at 4 °C. Excess PLP was removed by passing the protein through a PD-10 (Sephadex G-25M) desalting column (GE Healthcare) before concentration to 10–20 mg/ml using a VivaSpin 30 kDa cut-off concentration filter. For UV-visible assays, the concentration of recombinant SPT was 10 μM, and the spectrophotometer was blanked with 20 mM potassium phosphate (pH 7.5) containing 150 mM NaCl.

Determination of Dissociation Constants—Assays were carried out in 0.5-ml quartz cuvettes (1-cm path length) and typically contained 10 μM SPT in 20 mM potassium phosphate (pH 7.5). Varying amounts of L-serine (0–80 mM) were used in each assay. After the addition of substrate, the reactants were mixed and allowed to equilibrate for 15 min at 25 °C. Small base-line changes were corrected using Sigma Plot software. Changes in absorbance at 425 nm (SPT WT, N100C, R378A, and R378N) and 420 nm (SPT N100W and N100Y) were plotted against

Downloaded from www.jbc.org at Edinburgh University Library, on September 24, 2009

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2009/06/17/M109.008680.DC1.html

L-serine concentrations, and data points were fitted to a hyperbolic saturation curve (Equation 1) using Sigma Plot software,

$$\Delta A_{obs} = \frac{\Delta A_{max}[serine]}{K_d + [serine]} \quad \text{(Eq. 1)}$$

where ΔA_{obs} represents the observed change in absorbance at 422 nm, and ΔA_{max} is the maximal absorbance change, [serine] is the L-serine concentration, and K_d is the dissociation constant (Table 1). The errors are asymptotic S.E. values of each parameter derived from the curve fitting procedure within Sigma Plot, and values were reproducible across a number of enzyme preparations.

SPT WT, N100Y, and R378A Quinonoid Formation—Assays were carried out in 0.5-ml quartz cuvettes and typically contained 5 μM SPT in 20 mM potassium phosphate (pH 7.5). To this 45 mM L-serine and 1.45 mM *S*-(2-oxoheptadecyl)-CoA were added in the assay. After the addition of the substrate, the reactants were mixed and allowed to equilibrate for 20 min at 25 °C. The spectroscopic data were recorded between 200 and 800 nm. Small base-line changes were corrected using Sigma Plot software.

Assay of Recombinant SPT Activity—The activity of SPT was measured by monitoring the formation of [^{14}C]KDS. A final enzyme concentration of 10 μM SPT in 20 mM potassium phosphate buffer, pH 7.5, 150 mM NaCl) was incubated with 20 mM [U-^{14}C]L-serine (9250 Bq, 0.250 μCi; GE Healthcare) and 1.6 mM palmitoyl-CoA in a final volume of 125 μl. The reaction was incubated at 37 °C for 20 min, and then the reaction was quenched by the addition of 30 μl of 2.0 M NH_4OH (final concentration 0.4 M). This was then extracted with an equal volume of $CHCl_3/CH_3OH$ (2:1, v/v). The sample was centrifuged at 13,000 rpm for 5 min, and the aqueous phase was discarded before the organic phase was allowed to evaporate overnight. The resulting lipid residue was resuspended in 15 μl of $CHCl_3/CH_3OH$ (2:1, v/v) and spotted onto a silica gel 60 TLC plate. Separation was carried out with a mobile phase of $CHCl_3/CH_3OH/NH_4OH$ (40:10:1, v/v/v), using a KDS reference standard (Matreya). The TLC was developed with a Storage PhosphorImager (GE Healthcare) for 4 days at room temperature, and the phosphor screen was visualized using ImageJ software.

SPT activity was also measured using a continuous spectrophotometric assay by monitoring the release of CoASH from acyl-CoA substrates and reaction with 5,5′-dithiobis-2-nitrobenzoic acid (DTNB) (42). Assays were carried out in 0.5-ml cuvettes in a Varian Cary-50 UV-visible spectrophotometer. The enzyme was incubated with L-serine in a buffered solution containing DTNB, and the assay was started by the addition of the second substrate, palmitoyl-CoA. The CoASH thiol product was monitored by observation of the TNB^- anion at 412 nm ϵ_{max} = 14,150 M^{-1} cm^{-1} (43). Initial rates were measured at increasing concentrations of L-serine (0.12–50 mM) while maintaining palmitoyl-CoA in excess. A typical experiment to determine the K_m value for L-serine contained 0.16 μM SPT, 0.1–50 mM L-serine, 250 μM palmitoyl-CoA, and 0.2 mM DTNB in 50 mM potassium phosphate buffer, pH 7.5. Kinetic constants (k_{cat} and K_m values; Table 1) were calculated from Michaelis-Menten plots using Sigma Plot. Kinetic constant values were calculated in a similar way for CoA substrates (decanoyl, lauroyl, myristoyl, palmitoyl, stearoyl, and arachidoyl) but by maintaining L-serine (100 mM) in excess (Table 2). As a check, the k_{cat}/K_m parameters for CoA substrates were also determined by the method of complete condensation when [CoA] was $\ll K_m$. In this case, k_{cat}/K_m is calculated at a single substrate concentration as an observed rate constant (k_{obs}) by monitoring disappearance of the substrate. The data were fitted to the equation, $A = A_{lim} \times (1 - \exp(-k_{obs} \times t))$, where A is the absorbance at 412 nm, A_{lim} is the limiting absorbance value when the substrate has completely disappeared, k_{obs} is the observed rate constant, and t is the time in seconds. The errors are asymptotic S.E. values of each parameter derived from the curve fitting procedure within Sigma Plot. The values obtained were reproducible over a number of enzyme preparations, and the errors can be used as a gauge of the fitted curve's accuracy and were generally less than 5% of the value of the parameter.

Structural Biology—The three mutant proteins were screened for suitable crystallization conditions at the Scottish Structural Proteomics Facility. The proteins were dialyzed in the presence of excess PLP to ensure complete reloading prior to crystallization. Conditions used to obtain crystals of the wild-type/PLP protein (41) were used as a starting point to formulate stochastic optimization screens using software developed in house. The screens were built on a Hamilton Microstar liquid-handling robot controlled by Rhombix system software (Thermo). Crystallization trials were set up in hanging drop plates (EasyXtal DG-CrystalSupport; Qiagen) using 2–3 μl of the protein solution and 1 μl of well solution in the hanging drop.

Crystals of the mutant SPTs were obtained in the form of yellow plates corresponding to the high resolution form of the holo wild-type SPT. The SPT N100W mutant was crystallized at 20 mg/ml in 10 mM Tris (pH 7.5), 150 mM NaCl, 25 μM PLP using a well solution of 0.1 M $MgCl_2$, 22% (w/v) PEG 3350, and 0.10 M Hepes (pH 6.5). The N100Y mutant was crystallized at 20 mg/ml in 10 mM Tris (pH 7.5), 150 mM NaCl, 250 μM PLP, using a well solution of 0.15 M $MgCl_2$, 22% (w/v) PEG 3350, and 0.10 M Hepes (pH 6.5). Finally, the N100C mutant was crystallized at 20 mg/ml in 10 mM Tris (pH 7.5), 150 mM NaCl, 250 μM PLP, using a well solution of 0.1 M $MgCl_2$, 27% (w/v) PEG 3350, and 0.10 M Hepes (pH 6.5).

The solutions used for equilibrating harvested crystals varied slightly according to the crystallization conditions and consisted of 22–29% PEG 3350 (w/v), 0.1 M HEPES (pH 6.5), 0.08 – 0.12 M $MgCl_2$. All crystals used for data collection were soaked for a few minutes in the respective equilibrating solution plus 1–2 mM PLP. Additionally, the external aldimine form of the SPT wild-type was produced by soaking the crystal in the equilibrating solution plus 5 mM L-serine for 10 min, and that of the N100Y mutant was produced by soaking the crystal in the equilibrating solution plus 50 mM L-serine for 15 min at room temperature. The crystals were mounted in a cryo-loop (Molecular Dimensions) and cryo-protected in solutions varying slightly for each crystal and containing the equilibrating solutions plus 15–20% PEG 400 (v/v). The crystals were then frozen by plunging them into liquid nitro-

Downloaded from www.jbc.org at Edinburgh University Library, on September 24, 2009

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2009/06/17/M109.008680.DC1.html

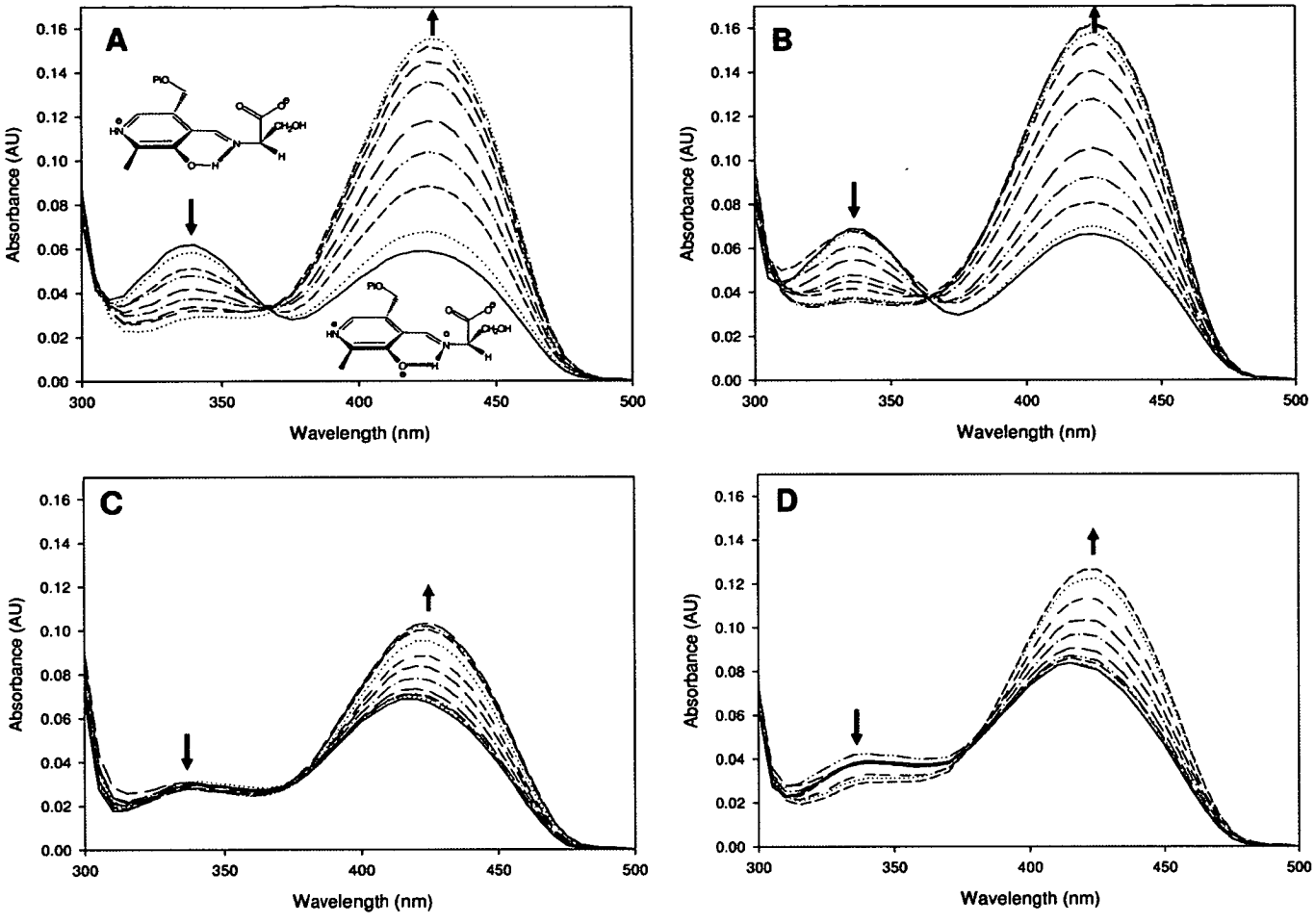


FIGURE 2. **UV-visible analysis of SPT wild-type and mutants.** Absorbance spectra of *S. paucimobilis* SPT wild-type (*A*) and mutants N100C (*B*), N100W (*C*), and N100Y (*D*). In *spectrum A* the enolimine (335 nm) and ketoenamine (420 nm) forms of the external aldimine are drawn. The *solid line* in each spectrum is the holo-form of the enzyme (10 μM SPT, 20 mM potassium phosphate buffer (pH 7.5), 150 mM NaCl, 25 °C). Increasing concentrations of L-serine were added (0, 0.1, 0.5, 1, 2, 5, 10, 20, and 40 mM; *dotted* and *dashed lines*), and the spectrum was recorded after 15 min. For the SPT N100C, N100W, and N100Y spectra the final concentrations of L-serine were 0, 0.1, 0.5, 1, 2, 5, 10, 20, 40, 60, and 80 mM. *AU*, absorbance units.

gen and carried in a dry cryogenic Dewar to the European Synchrotron Radiation Facility (Grenoble, France) for data collection. The data sets were collected at 100 K to varying resolutions using three different beam lines (supplemental Table 1). The data for the wild-type serine external aldimine were processed with Mosflm and scaled with Scala from the CCP4 suite of programs (44). The data for the N100C, N100Y, N100W, and N100Y serine external aldimine structures were processed using XDS and scaled using XSCALE (available on the World Wide Web) (45).

Analysis of the density revealed that the PLP was covalently bound to Lys^{265} in every mutant, confirming that we had obtained the holo-form of the proteins. All mutant models were refined using Refmac5 (46) and manually adjusted, including the addition of water molecules with WinCoot (47). Data and structures have been deposited in the Protein Data Bank. Density was clearly visible to the N100W mutation, but we could not satisfactorily locate the tryptophan side chain; there is disordered electron density, which indicates multiple conformations.

RESULTS

Spectroscopic Properties of SPT N100 Mutants—The three mutants were obtained in milligram quantities as the internal aldmine, holo-form, confirming their capacity to bind the PLP cofactor (Fig. 2). The SPT N100C mutant has absorbance maxima at 335 and 425 nm, akin to the wild-type SPT values of 340 and 425 nm (Fig. 2, *A* and *B*). In both native and N100C, the enolimine peak at 335 nm is dominant, indicating that the PLP cofactor is in a very similar chemical environment in both proteins. In contrast, the UV-visible spectra of the SPT N100W and N100Y mutants have maximum absorbance values at 340 and 415 nm (Fig. 2, *C* and *D*). We noted that the ketoenamine peak (415 nm) had not only blue-shifted; it had also become the dominant form in the spectrum in these enzymes. These differences suggest a common perturbation in the PLP binding site of the mutants compared with the wild type (Fig. 2, *C* and *D*, *solid lines*). The ability of the mutants to form the external aldimine form by binding L-serine was measured by the addition of increasing concentrations of the amino acid (ranging between 0 and 100 mM). This addition resulted in an increase of the

Downloaded from www.jbc.org at Edinburgh University Library, on September 24, 2009

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2009/06/17/M109.008680.DC1.html

absorbance at 425 nm for the wild-type and the N100C enzyme. Titration of L-serine into the SPT N100W and N100Y mutants led to a small shift of a few nm to a longer wavelength. Although all mutants form the external aldimine, the N100W and N100Y mutants have clear differences (Fig. 2, *A–D*). By measuring the change in absorbance at ketoenamine-specific wavelengths that are specific to the external aldimine, the apparent serine dissociation constants (K_d^{Ser}) have been determined for the wild type SPT and each of the mutants (Table 1). The wild-type SPT K_d^{Ser} is 1.1 mM, in good agreement with the K_d^{app} found by Ikushiro *et al.* (40) (1.4 mM). A similar K_d^{Ser} was obtained for the N100C mutant (2.7 mM; ~2.5-fold higher). Higher K_d^{Ser} values were obtained for two SPT mutants: N100W, 16.2 mM, a ~15-fold increase, and N100Y, 7.5 mM, a ~7-fold increase, compared with the wild-type SPT.

TABLE 1
Kinetic parameters for the purified SPT wild-type and mutant enzymes

Enzyme	$k_{cat} \times 10^3$	K_m^{Ser}	K_m^{pCoA}	k_{cat}/K_m^{Ser}	K_d^{Ser}
	s^{-1}	mM	μM	$M^{-1}\,s^{-1}$	mM
SPT WT	1150.0 ± 30.0	1.4 ± 0.1	35.4 ± 2.0	821.4	1.1 ± 0.1
SPT N100C	250.0 ± 4.0	7.0 ± 0.4	60.0 ± 9.8	35.7	2.7 ± 0.2
SPT N100W	9.0 ± 0.4	1.6 ± 0.4	19.3 ± 2.2	5.6	16.2 ± 3.1
SPT N100Y	5.0 ± 0.1	2.5 ± 0.4	31.4 ± 3.3	2.0	7.5 ± 0.5
SPT R378A	78.3 ± 3.2	3.8 ± 1.0	39.2 ± 3.1	20.5	1.1 ± 0.1
SPT R378N	33.0 ± 1.0	2.4 ± 0.2	31.0 ± 6.0	13.8	3.1 ± 0.2

KDS Formation by SPT WT and Mutants—The ability of each enzyme to convert [^{14}C]L-serine and palmitoyl-CoA to the product [^{14}C]KDS was examined (Fig. 3*A*), using the method previously described (48). The N100C mutant showed an activity similar to the wild-type SPT; however, both the N100W and N100Y mutants showed a large decrease in activity (~84%) compared with the wild-type enzyme (with saturating substrate concentrations) (41). The R378N and R378A produced 50 and 40% KDS, respectively, in comparison with the wild-type enzyme (Fig. 3*B*).

Kinetic Analyses of Wild-type, HSAN1 Mimics, and Arg^{378} Mutants—The raw data for wild-type SPT are shown in Fig. 3*C*. The SPT WT bound L-serine and palmitoyl-CoA with K_m values of 1.4 mM and 35.4 μM, respectively (Table 1). The enzyme turned over with a k_{cat} of 1.150 s^{-1} and an efficiency (k_{cat}/K_m) of 821.4 $M^{-1}\,s^{-1}$ for L-serine. We observed k_{cat}/K_m values of 35.7, 5.6, and 2.0 $M^{-1}\,s^{-1}$ for the mutant SPTs N100C, N100W, and N100Y, respectively. The lower values observed for the HSAN1 mimics N100W and N100Y (147- and 410-fold respectively) confirm that the mutations have had an impact on catalysis. We attribute the lower value for the N100C mutant (23-fold lower) to reactivity of the Cys^{100} thiol with the DTNB reagent (data not shown), since this mutant showed comparable KDS formation in the radioactive assay. The coupled assay was used to investigate the chain length specificity of wild-type SPT with various acyl-CoA substrates. We found palmitoyl-CoA (C16:0) to have the fastest turnover (k_{cat} = 1.150 s^{-1}) of all of the acyl-thioesters, with a K_m of 35.4 μM (Table 2). When comparing the k_{cat}/K_m values, stearoyl-CoA (C18:0) was the most efficient substrate (65,547 $M^{-1}\,s^{-1}$). The thioester with the longest chain, arachidoyl-CoA (C20:0) was a good substrate, with a K_m of <10 μM, but below this concentration, accurate rates could not be determined due to the limit of detection with the DTNB reagent. In contrast, the shortest substrate, decanoyl-CoA (C10:0), was the poorest substrate in terms of binding and turnover. In the first study of *S. paucimobilis* SPT carried out by Ikushiro *et al.* (39), they found a similar trend, with palmitoyl-CoA (C16) being the best substrate matching the human SPT. The recently discovered, unusual fused, viral SPT displayed highest activity with myristoyl-CoA (C14) (12, 49).

Similar analysis was also carried out on the SPT R378A and R378N mutants. These bound PLP in a manner similar to wild-type with absorbance maxima at 335 and 425 nm (data not shown). The values obtained were similar to that of the

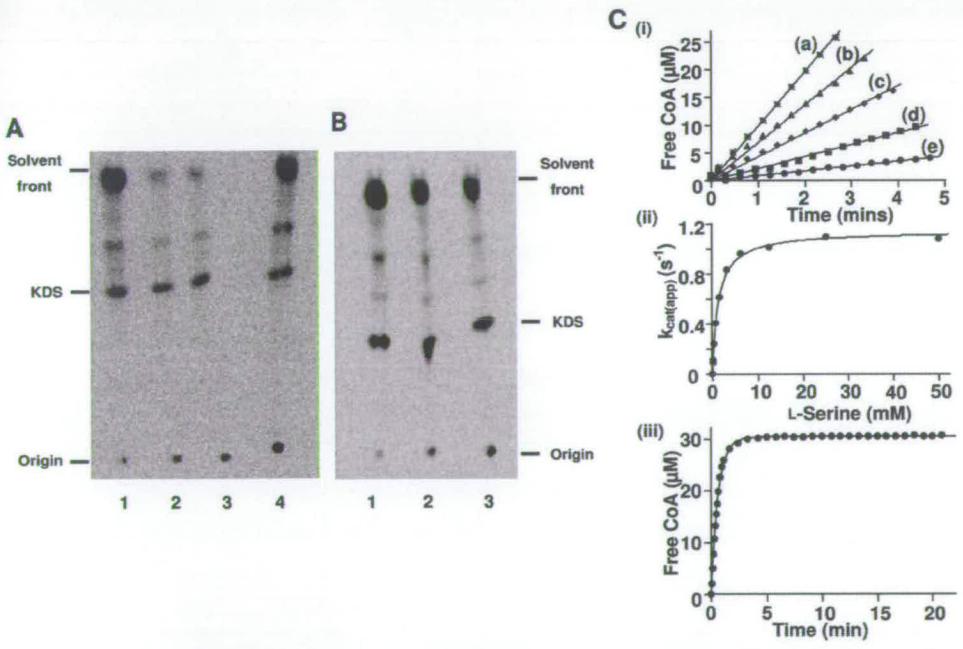


FIGURE 3. **Assay of SPT activity.** *A* and *B*, TLC of radiolabeled products obtained by SPT-catalyzed conversion of L-serine and palmitoyl-CoA to [^{14}C]KDS. Recombinant holo-SPT was incubated with U-^{14}C-labeled L-serine and palmitoyl-CoA for 20 min at 37 °C. The assay was stopped by the addition of NH_4OH and extracted with an equal volume of chloroform/methanol (2:1, v/v). Organic phases were removed, dried, separated by TLC, and analyzed by a PhosphorImager. The *arrow* marks the position of the standard KDS. *A*, products observed after reaction of the SPT N100C (*lane 1*), SPT N100W (*lane 2*), SPT N100Y (*lane 3*), and SPT WT (*lane 4*). *B*, products observed after reaction of the SPT WT (*lane 1*), SPT R378A (*lane 2*), and SPT R378N (*lane 3*). *C*, kinetic analysis of wild-type SPT using a spectrophotometric assay. *i*, initial rates of CoASH release measured spectrophotometrically at 412 nm and at several final concentrations of L-serine: 6.25 mM (*a*), 1.56 mM (*b*), 0.78 mM (*c*), 0.39 mM (*d*), and 0.12 mM (*e*). Each assay contained 0.16 μM enzyme, 250 μM palmitoyl-CoA, and 0.2 mM DTNB. *ii*, the initial rates were converted to k_{cat}^{app} values and plotted against L-serine concentration to determine the K_m and k_{cat} parameters. *iii*, complete condensation of 5 μM palmitoyl-CoA by 0.57 μM wild-type SPT in the presence of 100 mM L-serine. k_{cat}/K_m for palmitoyl-CoA was calculated by dividing k_{obs} by the enzyme concentration.

Downloaded from www.jbc.org at Edinburgh University Library, on September 24, 2009

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2009/06/17/M109.008680.DC1.html

wild-type SPT (K_d^{Ser} values of 1.1 and 3.1 mM for the R378A and R378N, respectively). Using the coupled spectrophotometric assay, we measured 60- and 40-fold decreases in k_{cat}/K_m^{Ser} for each mutant (R378A and R378N, respectively; Table 1).

SPT WT, N100Y, and R378A Quinonoid Formation—The ability of the wild-type enzyme and both the N100Y and R378A mutants to form a quinonoid intermediate was tested by the addition of a thioether analogue of palmitoyl-CoA reported by Ikushiro *et al.* (50). The addition of the analogue to the L-serine external aldimine form of the wild-type enzyme led to the appearance of a clear peak at 495 nm due to the quinonoid (increase in absorbance ~0.06), in agreement with that found previously (Fig. 4*A*). In contrast, the SPT N100Y mutant, under the same conditions, produced a small, broad shoulder 490–500 nm (increase in absorbance ~0.01 absorbance units; Fig. 4*B*). We could not observe the formation of a quinonoid species at all in the R378A mutant (Fig. 4*C*).

TABLE 2
Chain length specificity of wild-type SPT

Acyl-CoA	k_{cat}	K_m^{CoA}	k_{cat}/K_m^{CoA}
	s^{-1}	μM	$M^{-1}\,s^{-1}$
Decanoyl (10:0)	0.045 ± 0.002	2324.9 ± 175.1	19.4
Lauroyl (12:0)	0.262 ± 0.006	822.2 ± 32.6	318.7
Myristoyl (14:0)	0.601 ± 0.012	97.1 ± 5.1	6189.5
Palmitoyl (16:0)	1.150 ± 0.030	35.4 ± 2.0	32,485.9
Stearoyl (18:0)	0.898 ± 0.012	13.7 ± 0.8	65,547.4
Arachidoyl (20:0)	0.327 ± 0.014	<10	>32,700

The Structure of the Wild-type SPT·L-Ser External Aldimine Complex—The dimeric form of the L-Ser external aldimine SPT complex is shown in Fig. 5*A*. Of course, the major difference between the internal (holo-) and external aldimine forms is that in the internal adimine, the Schiff base is between Lys^{265} of the protein and the PLP co-factor, whereas in the L-Ser external aldimine, this C–N bond has broken. The PLP·L-ser external aldmine is clearly observed in the electron density (Fig. 5*B*). An overlay of the structures of the internal (41) and external aldimine shows an r.m.s. deviation of 0.3 Å and reveals that most of the interactions between the enzyme and the PLP cofactor are unchanged. In both structures, the PLP hydrogen-bonds to the side chains of Asp^{231}, His^{234}, Thr^{262}, and Ser^{264}, the main chain of Gly^{134} and Thr^{135}, and π-stacks with His^{159}. His^{159} is con-

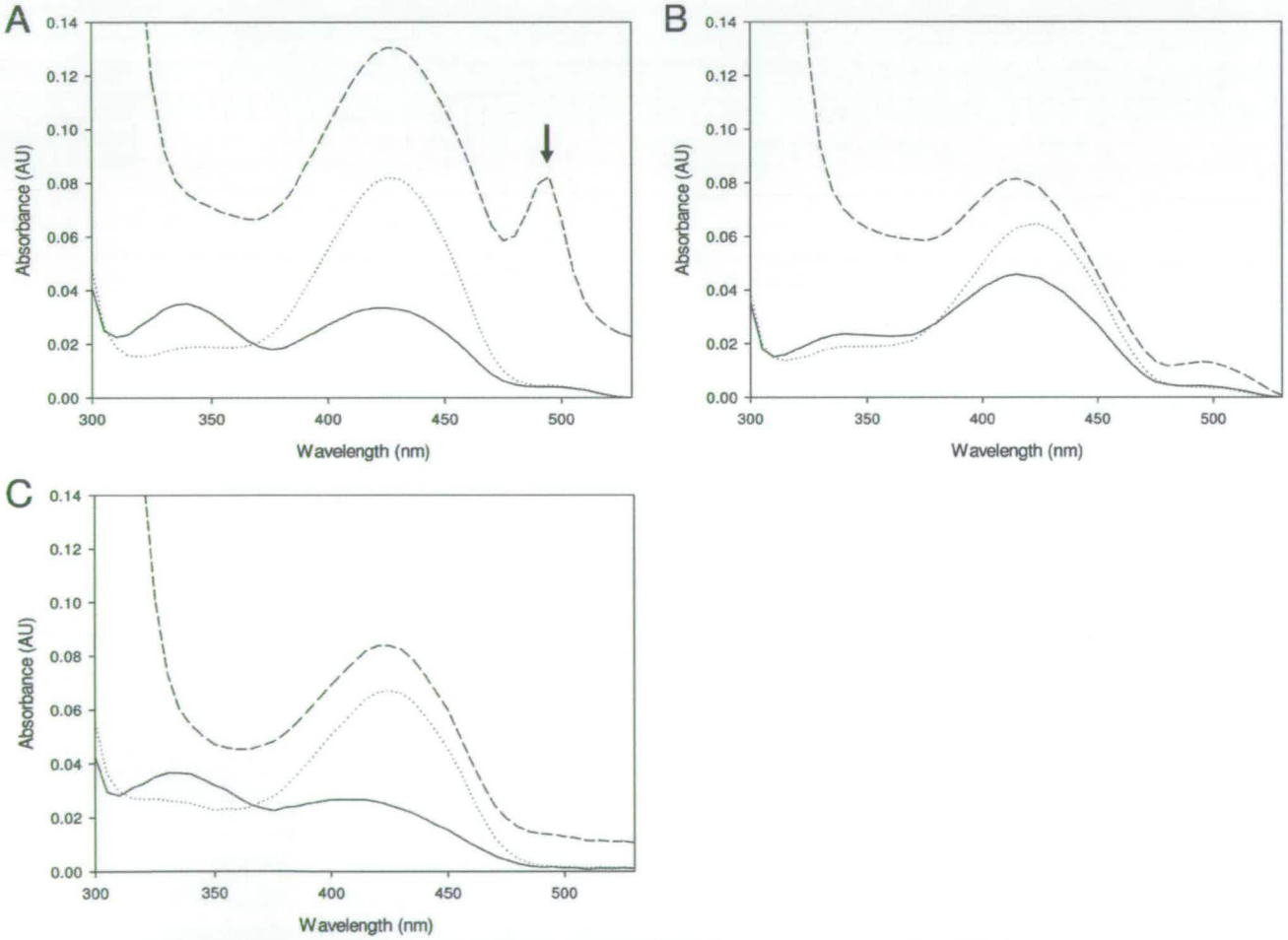


FIGURE 4. **Quinonoid formation in the wild-type SPT, N100Y, and R378A mutants.** UV-visible spectra of the SPT wild type (*A*), N100Y (*B*), and R378A (*C*) external aldimine forms after titration of the thioether analog (*S*-(2-oxoheptadecyl)-CoA) of palmitoyl-CoA. Incubations were carried out with 5 μM enzyme, 45 mM L-serine, 1.4 mM analog, 20 mM potassium phosphate buffer (pH 7.5), 150 mM NaCl for 15 min at 25 °C. The *solid line* in each spectrum represents the holo-form; the *dotted line* is the external aldimine form after the addition of L-ser; and the *dashed line* is after the addition of the analogue. An absorbance maximum at 495 nm (marked with an *arrow*), attributed to a quinonoid species (50), is clearly observed in the wild-type enzyme, but only a broad shoulder is produced in the N100Y mutant, and no peak is observed for the R378A mutant. *AU*, absorbance units.

Downloaded from www.jbc.org at Edinburgh University Library, on September 24, 2009

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2009/06/17/M109.008680.DC1.html

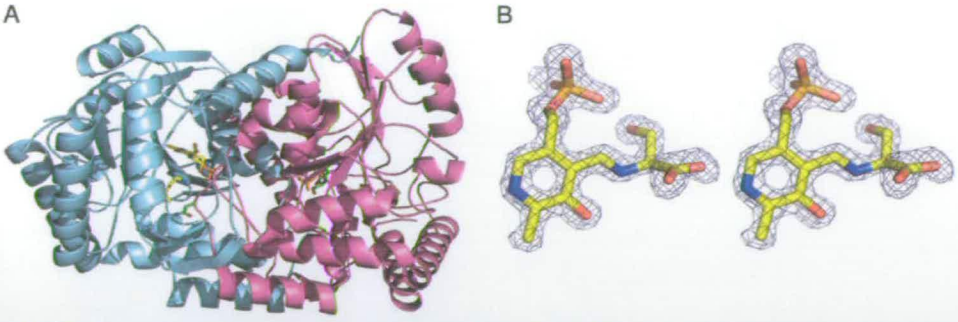


FIGURE 5. **Structure of the SPT·L-serine external aldimine complex.** *A*, overall structure of the SPT dimer showing one monomer in *light blue* and the other in *purple*. The residues Lys^{265} and Asn^{100} from the *light blue* monomer are shown in *stick form* and are *colored yellow*. The equivalent residues from the *purple monomer* are *colored green*. This clearly shows that the Asn^{100} from one monomer interacts with the active site of the other monomer. The PLP·L-ser external aldmine is also shown in *stick form*. *B*, a *stereo image* of an $F_o - F_c$ electron density map contoured at 3 σ around the PLP·L-ser external aldimine found in the native SPT. The phases were calculated from a model that had never included the PLP or aldimine.

served throughout all of the members of the α-oxoamine synthase family (41). As a result of external aldimine formation, the PLP has shifted, essentially rotating around the phosphate, with the Schiff base nitrogen moving over 1.5 Å. Lys^{265} adopts a different conformation as a result of breaking its link to PLP and now makes a hydrogen bond with the hydroxyl of the L-ser component of the external aldimine (Fig. 6*A*). Accompanying the change in Lys^{265} conformation, Tyr^{73} substantially alters its position. The carboxylate group of L-Ser hydrogen-bonds to the conserved His^{159} and makes a salt link with Arg^{378}, which has "swung" into the active site. In the native structure, this residue is at the end of a β-strand and hydrogen-bonded to the side chain of Gln^{357}. In order to make the salt contact, the β-strand has been disrupted, and the Cα of Arg^{378} has been moved over 5 Å. The loop following Arg^{378} (the "PPATP" loop) has significantly altered its conformation and position (*e.g.* the Cγ atom of Pro^{379} has moved over 10 Å). The reordering of the structure around Arg^{378} seems to be linked to other movements at Phe^{47} (1.2-Å Cα shift). These residues are contained within the one monomer. The phosphate of PLP binds to the side chain of Thr^{294} from one subunit and to the main chain of Ala^{294} from the other subunit. The hydroxyl of L-Ser makes a polar contact to the main chain of Ala^{294}. In the other monomer, the side chain Met^{104} adjusts its conformation, and there is a slight rigid body shift of Leu^{105} to Asn^{106}. The side chain of Asn^{100} from the other subunit has altered its position, but the hydrogen bond to Lys^{265} is maintained.

Structural Biology of the Holo-forms of the HSAN1 Mutant Mimics—The wild-type SPT and N100C structures superimpose with an r.m.s. deviation of 0.2 Å along the 396 Cα residues and are largely identical. Similarly, the N100W and N100Y mutant structures superimpose with 0.1 Å r.m.s. deviation for their Cα positions, mirroring their similar spectroscopic properties. However, both N100W and N100Y superimpose with the native structure (and N100C) with an r.m.s. deviation of 0.6 Å. For ease of discussion, only the differences between N100Y and the native enzyme are reported. The N100W structure is essentially identical to N100Y, whereas N100C is essentially identical to native. Reflecting the large deviation in Cα positions, there are a number of small shifts in secondary structure elements throughout the structure when comparing the N100Y mutant with the native SPT. These shifts are largely conserved, regardless of which structures one compares (N100W *versus* native, N100W *versus* N100C, N100Y *versus* N100C, and N100Y *versus* native), suggesting that they are real and not simply a crystallographic artifact. The most obviously visible change is the N-terminal helix (Asp^{23}–Gly^{42}), which has undergone an approximately 1.7-Å shift. This helix is an important part of the dimeric interface of SPT (41). At the site of mutation itself, the tyrosine side chain has "flipped" out the pocket occupied by Asn^{100} (Fig. 6*B*). At the Cα level, this is a displacement of 3.4 Å and is accomplished by an almost 180° rotation of the Ramachandran angle at Thr^{99}. As a result, the hydrogen bonds between the side chain of Asn^{100} with the key Lys^{265} from the other monomer have been broken. Instead, the Tyr^{100} now occupies a new pocket at the dimer interface formed by hydrophobic residues from both monomers (Phe^{109} from the same monomer; Met^{51}, Thr^{72}, and Ile^{69} from the other monomer) (Fig. 6*B*). These hydrophobic residues have adjusted their position to accommodate the bulky tyrosine side chain. The most pronounced of these changes appear to accompany the movement of Phe^{109}, which connects to an α-helix. Residues of the first turn of the helix, His^{110}, Asp^{111}, His^{112}, Met^{113}, and Glu^{114}, all show disturbed side chain and main chain positions relative to the native structure. The main chain at Leu^{105}, Asn^{106}, and Gly^{107} sits opposite and close to Asn^{100} in the native structure. The movement of the Cα atom at position 100 by 3.4 Å causes Leu^{105}–Gly^{107} to undergo large positional movements (over 4 Å for the Cα positions of Gly^{107}) and changes in their Ramachandran angles.

Structural Biology of Wild Type and N100Y L-Ser External Aldimines—The r.m.s. deviation between the two external aldimines, native SPT, and N100Y mutant at 0.4 Å is less than the value for the internal aldimine structures (0.6 Å). Thus, it appears that formation of the external aldimine drives N100Y toward a more "native" conformation. This can be seen most clearly in the loop between Tyr^{100} and Gly^{107}, which is more similar to the native structure in the external aldimine complex. The most striking difference between the two external aldimine structures is that in the N100Y structure, the "PPATP loop" remains in its internal aldimine conformation, and consequently Arg^{378} does not enter the active site and does not make salt links with the carboxylate of L-ser (Fig. 6*C*). Rather, in the external aldimine of N100Y, Arg^{378} is hydrogen-bonded to Gln^{357}, an interaction that is absent in the N100Y internal aldimine due to a small shift in structure.

DISCUSSION

Mechanistic Implications—The seminal hypothesis, forwarded by Dunathan (51) over 40 years ago, to explain how PLP-dependent enzymes can catalyze numerous chemical reactions is still valid. The key steps involve the deprotonation

Downloaded from www.jbc.org at Edinburgh University Library, on September 24, 2009

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2009/06/17/M109.008680.DC1.html

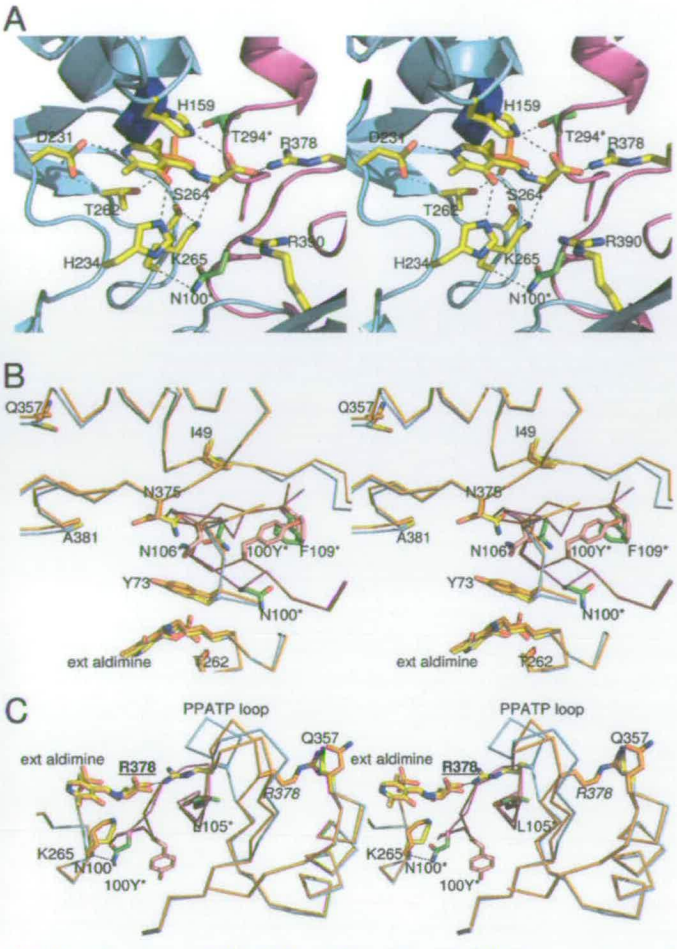


FIGURE 6. **Structure of the SPT·L-serine external aldimine complex and impact of the N100Y HSAN1 mutation.** *A*, a *stereo image* showing residues involved in binding the PLP·L-ser external aldimine complex are shown. His^{159}, Asp^{231}, His^{234}, Thr^{262}, Lys^{265}, and Arg^{378} are from monomer A (*light blue*) in *yellow stick format*. Residues Thr^{294} and Asn^{100} are from monomer B (*purple*), are *colored green*, and are annotated with an *asterisk*. Residues His^{159} and Arg^{378} interact with the carboxylate group of the L-serine; Lys^{265} (denoted by *dashed lines*) interacts with the hydroxyl group of the L-serine; and His^{234} interacts with the hydroxyl group of the PLP. Residue Asp^{231} binds to the nitrogen atom of the PLP cofactor, and the side chain of Asn^{100} (from monomer B) interacts with the backbone of Lys^{265} of the opposite monomer A. *B*, a *stereo image* of an overlay of the external aldimine forms of the wild-type and N100Y. The wild-type structure has been rendered with monomer A *colored* with a *blue* backbone and *yellow* side chains, and monomer B is shown with a *purple* backbone and *green* side chains. In the N100Y mutant, the monomer A backbone and side chains are *colored gold*, and monomer B is colored with a *red* backbone and *salmon pink* side chains. The N100 side chain from monomer B of the wild-type SPT (*green stick*, *asterisk*) interacts with residues from monomer A. However, in the SPT N100Y mutant, the Tyr^{100} side chain has flipped into a pocket formed by displacement of Phe^{109} of the same monomer *B*, residue Asn^{106} from monomer B has also undergone a large shift, and residue Asn^{375} from monomer A has also changed. The backbone containing residue Tyr^{100} has also undergone a large conformational change. *C*, a *stereo image* of an overlay of the external aldimine forms of the wild-type and N100Y showing the impact of mutation on Arg^{378} and the PPATP loop. The *coloring* is the same as in *B*. In wild-type SPT, residue Arg^{378} (side chain *yellow* on the *blue* backbone) is in the swung in position and interacts with the L-Ser carboxylic acid. In the SPT N100Y mutant, the "PPATP loop" and the Arg^{378} residue (backbone and side chain *colored gold*) are in the "swung out" conformation observed in the wild-type internal aldimine form. The interaction between Arg^{378} and Gln^{357} (both in *gold*) and the different positions of the Asn^{100} and Tyr^{100} residues are shown.

of the amino acid·PLP external aldmine to form the quinonoid species, which condenses with the incoming acyl-CoA thioester (Fig. 1). This is thought to generate a β-keto acid intermediate that then decarboxylates to give a product aldimine (23). The first intermediate in the catalytic cycle of SPT is the external aldimine formed between L-ser and PLP. Deprotonation of the Cα of this external aldimine intermediate is the crucial next step. In studies with ALAS, this deprotonation has been observed to be accelerated by over 250,000-fold by the binding of the incoming thioester substrate (14, 52). Studies on AONS, which uses L-alanine and pimeloyl-CoA, suggest that this enzyme undergoes conformational change during intermediate formation. Incubation of the enzyme with palmitoyl-CoA pre-equilibrated with L-alanine gave a 30-ms lag before quinonoid formation at 486 nm was observed with a rate of 45 s^{-1} (14). Recently, a thioether CoA substrate analogue (*S*-(2-oxoheptadecyl)-CoA) was shown to increase the rate of proton abstraction from the serine external aldimine of SPT by around 100-fold (50).

Our new structure provides insight into the residues that control the L-Ser external aldimine conformation (Fig. 7, *A–C*). How was this intermediate trapped in the crystal? The high resolution data confirm that the L-serine·PLP external aldimine is sp^3-hybridized at the Cα (the angles are 111.5, 106.4, and 109.6°, which are close to the theoretical values for a tetrahedral carbon). In the external aldimine structure, the Nz atom of Lys^{265} contacts (3.4 Å) the Cα proton of the L-Ser. This double role for the Schiff base forming lysine was suggested by studies of members of the α-oxoamine synthase family (14, 17, 18, 20), and SPT has now been shown to use the same feature. The angles between each functional group and the N-imine (which is co-planar with the PLP ring) were measured as follows: H-Cα-N, 107.7°; CO_2-Cα-N, 108.4°; CH_2-Cα-N, 113.3°. None of these have the desired perpendicular orientation to the plane, indicating that in the crystal structure, the external aldimine is not in the optimal "Dunathan conformation" suitable for deprotonation.

If, as seems likely, palmitoyl-CoA binding leads to rapid deprotonation, we suggest that it does so by triggering the rotation of the external aldimine about the imine N–Cα bond. A small turn of ~15–20° would bring the Cα hydrogen into the required perpendicular geometry (Fig. 7*B*). Of course, this would cause the simultaneous movement of the -CO_2^- group and the -CH_2OH side chain. Ikushiro proposed a role for the conserved residue Arg^{390} in the SPT mechanism (50). The structural data show that Arg^{390} is over 6.5 Å distant from the carboxylate; rotating around the Cα–N bond to reach the likely quinonoid conformation shortens this distance to 4.6 Å (Fig. 7, *C* and *D*). An additional rearrangement, promoted by palmitoyl-CoA binding, could allow Arg^{390} to interact with the carboxylate, as seen in the KBL external aldimine. We note that Arg^{390} plays an important role in stabilizing the conformation of Tyr^{73}, which forms the active site. The rotation to form the quinonoid would disrupt the salt link to Arg^{378} in the absence of any additional structural adjustment. The interplay of Arg^{390} and Arg^{378} in interacting with the carboxylate group may be an important component in controlling the formation and deprotonation of the external aldimine. Together, they could act to pick up, ferry, then anchor the carboxylic acid group. This functional group is ultimately lost from the β-keto acid as CO_2, and future work will be required to understand how this important

Downloaded from www.jbc.org at Edinburgh University Library, on September 24, 2009

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2009/06/17/M109.008680.DC1.html

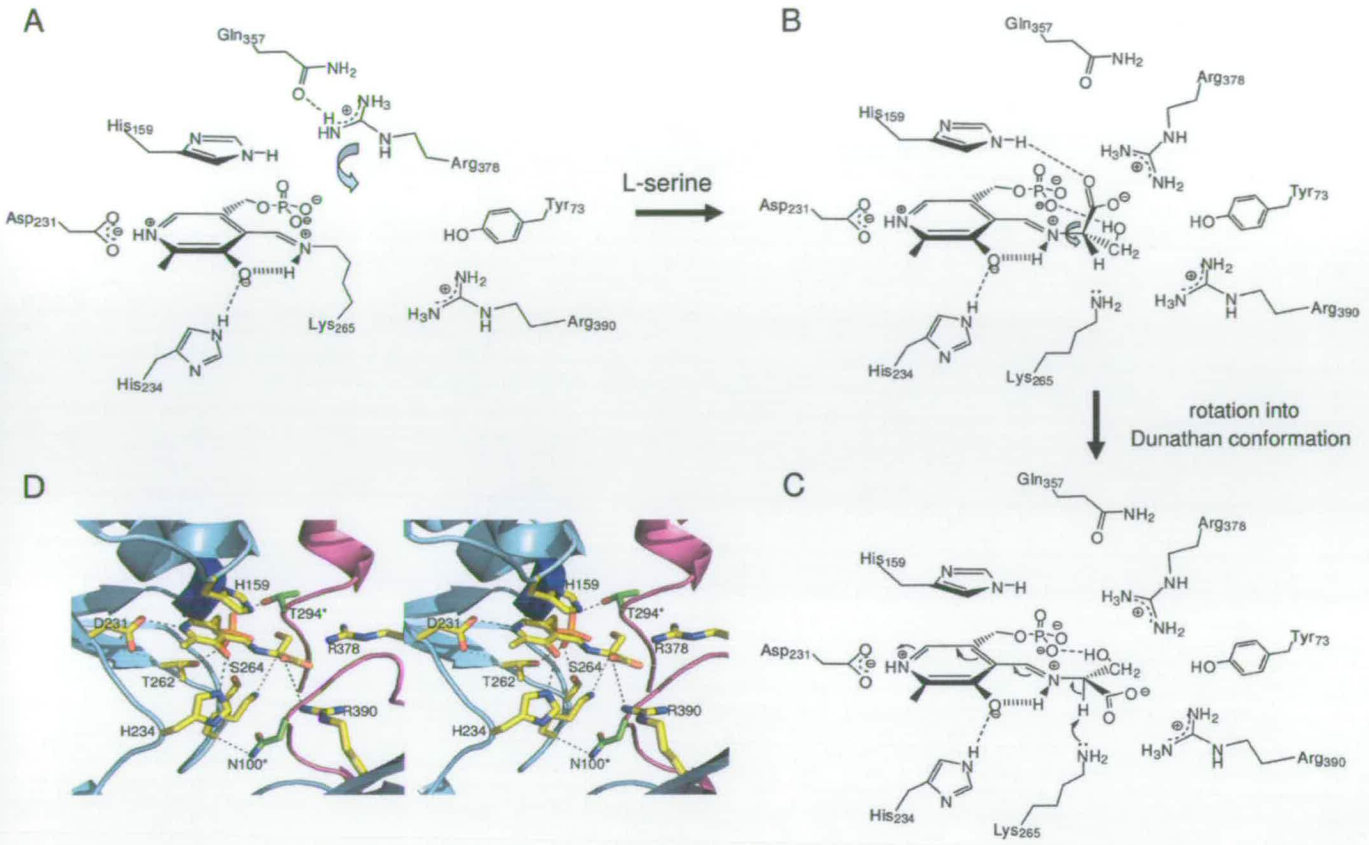


FIGURE 7. **Schematic representation of the residues involved in SPT catalysis, conformations of the external aldimines, and roles of active site arginine residues in controlling.** *A*, key residues involved in forming the active site are highlighted. The Arg^{378} side chain is in the swung out conformation and interacts with Gln^{357}. *B*, binding of L-Ser leads to PLP-L-Ser external aldimine formation. The residues observed in the wild-type crystal structure are shown. The carboxylate of L-Ser interacts with the side chains of His^{159} and Arg^{378}, which is now in the swung in conformation. Rotation around the Cα–Nα bond (*arrow*) brings the intermediate into the desired Dunathan conformation shown in *C*. The carboxylate is shown to possibly interact with Arg^{390}, and the Cα-H is now perpendicular to the PLP ring. This is then deprotonated by Lys^{265} to form the quinonoid. *D*, *stereo image* of a model of the external aldimine Dunathan conformation using the external aldimine structure.

step is controlled. Ikushiro *et al.* (50) also proposed a role for His^{159} acting as a hydrogen donor to accommodate the developing negative charge on the C=O of the thioester. The structure shows that the Nϵ atom of His^{159} hydrogen-bonds to the L-serine CO_2^- group as well as maintaining a π-stacking interaction with the PLP ring. This is consistent with the reported SPT H159F mutant, which bound L-serine weakly (K_d = 11.1 mM) and was incapable of KDS formation. Whether the residue also plays a role in the chemistry of the thioester activation remains an open question.

That Arg^{378} binds to the external aldimine was unsuspected, since it required a very large conformational change from the holoenzyme structure. Mutation of this residue (R378A and R378N) reduces enzyme efficiency by over a factor of 10 (Table 1). Thus, we suggest that this residue plays an important, but not essential, role possibly influencing the stabilization of an intermediate after both substrates have bound. Weight is given to this hypothesis by the fact that R378A cannot form a quinonoid in the presence of the thioether analogue (Fig. 4*C*). The nonessential nature of the residue is consistent with its lack of sequence conservation in all SPT isoforms, and an equivalent residue in the human SPT1/SPT2 heterodimer awaits identification. That there may be such a residue comes from analysis of enzymes in the α-oxoamine synthase family. In the two glycine-specific members, KBL and ALAS, residues Asn^{50} and Asn^{54}, respectively, make direct contact with the -CO_2^- group (17, 18) of the amino acid. In AONS, Webster *et al.* (14) used the AONS-AON product external aldimine structure to predict that the L-alanine carboxylate would coordinate with residue Asn^{47}. Sequence homology lines up residue Asn^{74} from *S. paucimobilis* SPT with the three Asn residues from AONS, ALAS, and KBL; however, analysis of the structure reveals that this residue in not involved in the active site. Thus, there is reason to suspect that some residue is required to coordinate to the carboxylic group of the external aldimine; however, given that the function is not essential, the conservation of the residue is not absolute.

The organization of the active sites of AONS, KBL, and ALAS is different compared with the bacterial SPT; the constellation of residues around the PLP cofactor is conserved, but there are subtle differences. SPT favors L-Ser as the substrate, and in the SPT external aldimine structure, the side chain hydroxyl makes a number of hydrogen bonds, which could explain this preference. Interestingly, although SPT can form an external aldimine with L-Thr (K_d = 3.8 mM) (40), it cannot be deprotonated to the quinonoid in our hands (data not shown). Looking at the structure of the aldimine, the additional methyl group of L-Thr would appear to clash with Ser^{102} and Arg^{378} in the proposed

Downloaded from www.jbc.org at Edinburgh University Library, on September 24, 2009

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2009/06/17/M109.008680.DC1.html

quinonoid model. In the ALAS structure, Thr^{83} would seem to select against any substituent at the Cα position, explaining the preference for Gly of this enzyme.

Mimics of the HSAN1 Disease Causing Mutations—The two most common mutations (C133W and C133Y) associated with the human disease HSAN1 have intrigued a number of researchers as to how they impact on sphingolipid metabolism and regulation. The fact that these mutations are in the inactive subunit, LCB1 (SPT1), has added to their scientific interest. Tissue from HSAN1 patients contains relatively normal levels of total sphingolipids (10, 31, 32). However, there are differing reports as to the intrinsic SPT activity of cells from HSAN1 individuals, and questions remain as to the precise consequence of the HSAN1 mutants on SPT turnover, stability, and regulation. Gable *et al.* (30) produced mutants of the yeast *lcb1* gene (C180Y and C180W) that mimic the HSAN1 substitutions (C133W and C133Y). Interestingly, these mutations dominantly inactivated the SPT activity of the yeast LCB1·LCB2 heterodimeric complex. This study proposed that the LCB1 HSAN1 mutations perturbed the active site of the LCB1·LCB2 dimer. McCampbell *et al.* (31) expressed the human LCB1 C133W mutant in a transgenic mouse and found that the SPT activity was decreased in various tissues and mice developed a number HSAN1 symptoms. However, the levels of total ceramide in these mice were unchanged, as was also observed in human HSAN1 individuals. This indicates that to some extent, SPT activity is preserved, and the mutants have a more subtle effect than simple loss of activity. It would also indicate that there is some regulation of sphingolipid biosynthesis after KDS formation by SPT.

We had previously mapped the human Cys^{133} residue to Asn^{100} on the bacterial SPT (41). We first engineered an N100C mutation as a better mimic for the human enzyme. This mutant enzyme behaves within the sensitivity of our UV-visible spectroscopic, kinetic, and high resolution structural analyses as native. This gave us confidence that the other mutations at Asn^{100} in the bacterial enzyme should provide a meaningful insight into human SPT.

N100W and N100Y mutants displayed marked spectroscopic, enzymatic, and structural differences when compared with wild-type SPT. Although both mutants are severely compromised in their activity, importantly, they do retain the ability to make the desired KDS product. It was obvious from the distinctive UV-visible properties of PLP-containing holo-forms of the enzymes that the mutants had altered the cofactor's chemical environment (Fig. 2, *A–D*). This change in the chemical environment apparently disturbs catalysis. After formation of an external aldimine, deprotonation gives a quinonoid intermediate (Fig. 1), which has been detected for both AONS and ALAS enzymes (14, 52). In the native SPT enzyme, the quinonoid form can be detected by adding the palmitoyl-CoA thioether analog. However, in the N100Y mutant, the signal for the quinonoid is essentially absent, indicating that the stabilization of this intermediate is significantly decreased (Fig. 4, *A* and *B*). In the wild-type holo-SPT (and N100C), the side chain of one monomer makes a direct contact with the amide backbone of the conserved PLP Schiff base Lys^{265} of the other monomer. This hydrogen bond is lost by the disruption introduced by the mutation. We suggest that this would increase the mobility of the internal aldimine. The most striking observation is that the N100Y and N100W mutations cause, because of the key location of Asn^{100} at the dimer interface, complex structural changes to ripple across the dimer interface (Fig. 6*B*). A striking illustration of this can be seen in the external aldimine of N100Y, particularly the role of Arg^{378}. In the native enzyme, this residue plays an important role in binding to the carboxylate of the external aldimine, but in the N100Y mutant, Arg^{378} remains uninvolved with the intermediate (Fig. 6*C*). Although it is difficult to identify a clear cut interaction that disfavors Arg^{378} from entering the active site, this residue is in contact with Leu^{105} from the other subunit. Leu^{105} is one of the residues directly affected by the mutation at Asn^{100}. The change in Leu^{105} position introduced by the mutation would increase the van der Waals contact between Arg^{378} (in the "swung in" state) and Leu^{105} in the mutant. It is this ability for structural changes to "reach across the interface" that we propose underlies the C133W and C133Y mutation effects in HSAN1. The perturbations in the neighboring structure give rise to the changed chemical environment of both the internal and external aldimine, altering the enzyme function.

Combined with previous studies on AONS, ALAS, and KBL, the SPT data presented here shed more light on the catalytic mechanism employed by members of the α-oxoamine synthase family. The eukaryotic SPT heterodimeric isozymes appear to be more complex than their bacterial homologs, since a third, tissue-specific SPT subunit has been identified recently (34). Furthermore, an abnormal, "dead end" sphingolipid product, 1-deoxysphinganine (also known as the natural product "spisulosine"), formed by SPT using L-Ala as a substrate, has been identified in mammalian cells (53, 54). Whether the SPT HSAN1 mutations impact directly on substrate and product specificity requires further detailed investigation. In the absence of large amounts of homogenous eukaryotic complexes, the bacterial enzymes provide some insight into how the wild-type and mutant forms behave and aid in the design of isolation strategies for the more complex SPTs (55). Studies aimed at the capture of complexes of various SPTs with a range of substrates, products, and inhibitors are under way.

Acknowledgments—We thank Dr. Thorsten Hornemann (Zürich) for help regarding the CoASH assay. We also extend thanks to Dr. Hiroko Ikushiro (Osaka) for the palmitoyl-CoA thioether analogue. The SSPF is supported by the Scottish Funding Council (grant reference SULSA).

REFERENCES

1. Merrill, A. H., Jr. (2002) *J. Biol. Chem.* **277,** 25843–25846
2. Futerman, A. H., and Hannun, Y. A. (2004) *EMBO Rep.* **5,** 777–782
3. Kobayashi, T., Takahashi, M., Nagatsuka, Y., and Hirabayashi, Y. (2006) *Biol. Pharm. Bull.* **29,** 1526–1531
4. Hirabayashi, Y., Igarashi, Y., and Merrill, A. H. J. (2006) *Sphingolipid Biology,* 1st Ed., Springer-Verlag, Tokyo
5. Auer-Grumbach, M. (2004) *Drugs Today* **40,** 385–394
6. Dyck, P. (1993) *Neuronal Atrophy and Degeneration Predominantly Affecting Peripheral Sensory and Autonomic Neurons,* W.B. Saunders Co., Philadelphia
7. Houlden, H., Blake, J., and Reilly, M. M. (2004) *Curr. Opin. Neurol.* **17,** 569–577

Downloaded from www.jbc.org at Edinburgh University Library, on September 24, 2009

Supplemental Material can be found at:
http://www.jbc.org/content/suppl/2009/06/17/M109.008680.DC1.html

8. Houlden, H., King, R., Blake, J., Groves, M., Love, S., Woodward, C., Hammans, S., Nicoll, J., Lennox, G., O'Donovan, D. G., Gabriel, C., Thomas, P. K., and Reilly, M. M. (2006) *Brain* **129,** 411–425
9. Bejaoui, K., Wu, C., Scheffler, M. D., Haan, G., Ashby, P., Wu, L., de Jong, P., and Brown, R. H., Jr. (2001) *Nat. Genet.* **27,** 261–262
10. Dawkins, J. L., Hulme, D. J., Brahmbhatt, S. B., Auer-Grumbach, M., and Nicholson, G. A. (2001) *Nat. Genet.* **27,** 309–312
11. Nicholson, G. A., Dawkins, J. L., Blair, I. P., Kennerson, M. L., Gordon, M. J., Cherryson, A. K., Nash, J., and Bananis, T. (1996) *Nat. Genet.* **13,** 101–104
12. Hanada, K. (2003) *Biochim. Biophys. Acta* **1632,** 16–30
13. Alexeev, D., Alexeeva, M., Baxter, R. L., Campopiano, D. J., Webster, S. P., and Sawyer, L. (1998) *J. Mol. Biol.* **284,** 401–419
14. Webster, S. P., Alexeev, D., Campopiano, D. J., Watt, R. M., Alexeeva, M., Sawyer, L., and Baxter, R. L. (2000) *Biochemistry* **39,** 516–528
15. Jordan, P. M., and Shemin, D. (1972) in *The Enzymes* (Boyer, P. D., ed) 3rd Ed., Academic Press, Inc., London
16. Ferreira, G. C., and Gong, J. (1995) *J. Bioenerg. Biomembr.* **27,** 151–159
17. Astner, I., Schulze, J. O., van den Heuvel, J., Jahn, D., Schubert, W. D., and Heinz, D. W. (2005) *EMBO J.* **24,** 3166–3177
18. Schmidt, A., Sivaraman, J., Li, Y., Larocque, R., Barbosa, J. A., Smith, C., Matte, A., Schrag, J. D., and Cygler, M. (2001) *Biochemistry* **40,** 5151–5160
19. Eliot, A. C., and Kirsch, J. F. (2004) *Annu. Rev. Biochem.* **73,** 383–415
20. Hunter, G. A., and Ferreira, G. C. (1999) *Biochemistry* **38,** 3711–3718
21. Zaman, Z., Jordan, P. M., and Akhtar, M. (1973) *Biochem. J.* **135,** 257–263
22. Alexeev, D., Baxter, R. L., Campopiano, D. J., Kerbarh, O., Sawyer, L., Tomczyk, N., Watt, R., and Webster, S. P. (2006) *Org. Biomol. Chem.* **4,** 1209–1212
23. Kerbarh, O., Campopiano, D. J., and Baxter, R. L. (2006) *Chem. Commun.* 60–62
24. Hanada, K., Hara, T., Fukasawa, M., Yamaji, A., Umeda, M., and Nishijima, M. (1998) *J. Biol. Chem.* **273,** 33787–33794
25. Verhoeven, K., Coen, K., De Vriendt, E., Jacobs, A., Van Gerwen, V., Smouts, I., Pou-Serradell, A., Martin, J. J., Timmerman, V., and De Jonghe, P. (2004) *Neurology* **62,** 1001–1002
26. Buede, R., Rinker-Schaffer, C., Pinto, W. J., Lester, R. L., and Dickson, R. C. (1991) *J. Bacteriol.* **173,** 4325–4332
27. Hanada, K., Hara, T., Nishijima, M., Kuge, O., Dickson, R. C., and Nagiec, M. M. (1997) *J. Biol. Chem.* **272,** 32108–32114
28. Nagiec, M. M., Baltisberger, J. A., Wells, G. B., Lester, R. L., and Dickson, R. C. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 7899–7902
29. Nagiec, M. M., Lester, R. L., and Dickson, R. C. (1996) *Gene* **177,** 237–241
30. Gable, K., Han, G., Monaghan, E., Bacikova, D., Natarajan, M., Williams, R., and Dunn, T. M. (2002) *J. Biol. Chem.* **277,** 10194–10200
31. McCampbell, A., Truong, D., Broom, D. C., Allchorne, A., Gable, K., Cutler, R. G., Mattson, M. P., Woolf, C. J., Frosch, M. P., Harmon, J. M., Dunn, T. M., and Brown, R. H., Jr. (2005) *Hum. Mol. Genet.* **14,** 3507–3521
32. Bejaoui, K., Uchida, Y., Yasuda, S., Ho, M., Nishijima, M., Brown, R. H., Jr., Holleran, W. M., and Hanada, K. (2002) *J. Clin. Invest.* **110,** 1301–1308
33. Gable, K., Slife, H., Bacikova, D., Monaghan, E., and Dunn, T. M. (2000) *J. Biol. Chem.* **275,** 7597–7603
34. Hornemann, T., Richard, S., Rütti, M. F., Wei, Y., and von Eckardstein, A. (2006) *J. Biol. Chem.* **281,** 37275–37281
35. Hornemann, T., Wei, Y., and von Eckardstein, A. (2007) *Biochem. J.* **405,** 157–164
36. Han, G., Gable, K., Yan, L., Natarajan, M., Krishnamurthy, J., Gupta, S. D., Borovitskaya, A., Harmon, J. M., and Dunn, T. M. (2004) *J. Biol. Chem.* **279,** 53707–53716
37. Hanada, K., Hara, T., and Nishijima, M. (2000) *J. Biol. Chem.* **275,** 8409–8415
38. Chen, J. K., Lane, W. S., and Schreiber, S. L. (1999) *Chem. Biol.* **6,** 221–235
39. Ikushiro, H., Hayashi, H., and Kagamiyama, H. (2001) *J. Biol. Chem.* **276,** 18249–18256
40. Ikushiro, H., Hayashi, H., and Kagamiyama, H. (2004) *Biochemistry* **43,** 1082–1092
41. Yard, B. A., Carter, L. G., Johnson, K. A., Overton, I. M., Dorward, M., Liu, H., McMahon, S. A., Oke, M., Puech, D., Barton, G. J., Naismith, J. H., and Campopiano, D. J. (2007) *J. Mol. Biol.* **370,** 870–886
42. Ellman, G. L. (1959) *Arch. Biochem. Biophys.* **82,** 70–77
43. Riddles, P. W., Blakeley, R. L., and Zerner, B. (1979) *Anal. Biochem.* **94,** 75–81
44. Collaborative Computational Project 4 (1994) *Acta Crystallogr. D* **50,** 760–763
45. Kabsch, W. (1993) *J. Appl. Crystallogr.* **26,** 795–800
46. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. D* **53,** 240–255
47. Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. D* **60,** 2126–2132
48. Williams, R. D., Wang, E., and Merrill, A. H., Jr. (1984) *Arch. Biochem. Biophys.* **228,** 282–291
49. Han, G., Gable, K., Yan, L., Allen, M. J., Wilson, W. H., Moitra, P., Harmon, J. M., and Dunn, T. M. (2006) *J. Biol. Chem.* **281,** 39935–39942
50. Ikushiro, H., Fujii, S., Shiraiwa, Y., and Hayashi, H. (2008) *J. Biol. Chem.* **283,** 7542–7553
51. Dunathan, H. C. (1966) *Proc. Natl. Acad. Sci. U. S. A.* **55,** 712–716
52. Zhang, J., and Ferreira, G. C. (2002) *J. Biol. Chem.* **277,** 44660–44669
53. Hornemann, T., Penno, A., and von Eckardstein, A. (2008) *Chem. Phys. Lipids* **154S,** S63, PO111
54. Zitomer, N. C., Mitchell, T., Voss, K. A., Bondy, G. S., Pruett, S. T., Garnier-Amblard, E. C., Liebeskind, L. S., Park, H., Wang, E., Sullards, M. C., Merrill, A. H., Jr., and Riley, R. T. (2009) *J. Biol. Chem.* **284,** 4786–4795
55. Ikushiro, H., Islam, M. M., Tojo, H., and Hayashi, H. (2007) *J. Bacteriol.* **189,** 5749–5761

Downloaded from www.jbc.org at Edinburgh University Library, on September 24, 2009

Inhibition of the PLP-dependent enzyme serine palmitoyltransferase by cycloserine: evidence for a novel decarboxylative mechanism of inactivation†‡

Jonathan Lowther,[a] Beverley A. Yard,[a] Kenneth A. Johnson,[b] Lester G. Carter,[b] Venugopal T. Bhat,[a] Marine C. C. Raman,[a] David J. Clarke,[a] Britta Ramakers,[a] Stephen A. McMahon,[b] James H. Naismith[b] and Dominic J. Campopiano*[a]

Received 1st March 2010, Accepted 13th April 2010
First published as an Advance Article on the web
DOI: 10.1039/c003743e

Cycloserine (CS, 4-amino-3-isoxazolidone) is a cyclic amino acid mimic that is known to inhibit many essential pyridoxal 5′-phosphate (PLP)-dependent enzymes. Two CS enantiomers are known; D-cycloserine (DCS, also known as Seromycin) is a natural product that is used to treat resistant *Mycobacterium tuberculosis* infections as well as neurological disorders since it is a potent NMDA receptor agonist, and L-cycloserine (LCS) is a synthetic enantiomer whose usefulness as a drug has been hampered by its inherent toxicity arising through inhibition of sphingolipid metabolism. Previous studies on various PLP-dependent enzymes revealed a common mechanism of inhibition by both enantiomers of CS; the PLP cofactor is disabled by forming a stable 3-hydroxyisoxazole/pyridoxamine 5′-phosphate (PMP) adduct at the active site where the cycloserine ring remains intact. Here we describe a novel mechanism of CS inactivation of the PLP-dependent enzyme serine palmitoyltransferase (SPT) from *Sphingomonas paucimobilis*. SPT catalyses the condensation of L-serine and palmitoyl-CoA, the first step in the *de novo* sphingolipid biosynthetic pathway. We have used a range of kinetic, spectroscopic and structural techniques to postulate that both LCS and DCS inactivate SPT by transamination to form a free pyridoxamine 5′-phosphate (PMP) and β-aminooxyacetaldehyde that remain bound at the active site. We suggest this occurs by ring opening of the cycloserine ring followed by decarboxylation. Enzyme kinetics show that inhibition is reversed by incubation with excess PLP and that LCS is a more effective SPT inhibitor than DCS. UV-visible spectroscopic data, combined with site-directed mutagenesis, suggest that a mobile Arg^{378} residue is involved in cycloserine inactivation of SPT.

Introduction

Sphingolipids are a large family of bioactive molecules that are found in all eukaryotic and some prokaryotic membranes. An important example is sphingomyelin, a constituent of the protective myelin sheath that surrounds nerve cells.[1] Sphingolipids can associate with cholesterol to form 'lipid rafts' or sphingolipid-based microdomains necessary for signal transduction and membrane trafficking.[2] Sphingolipid metabolites such as ceramides and sphingosine-1-phosphate play important roles in cell proliferation, differentiation and apoptosis[3–5] and it follows that pharmaceutical intervention that regulates the sphingolipid metabolic pathway could help to combat pathological processes such as carcinogenesis,[6] atherosclerosis[7] and Parkinson's disease.[8]

The *de novo* biosynthetic pathway for sphingolipids varies from one organism to another but the first and rate-limiting step is common to all: condensation of L-serine with palmitoyl-CoA to form 3-ketodihydrosphingosine.[9] This step is catalysed by the pyridoxal 5′-phosphate (PLP)-dependent enzyme, serine palmitoyltransferase (SPT), a member of the α-oxo-amine synthase (AOS) subfamily. Other members of this subfamily that have been well characterised include 8-amino-7-oxononanoate synthase (AONS),[10,11] 5-aminolevulinate synthase (ALAS)[12,13] and 2-amino-3-ketobutyrate-CoA ligase (KBL).[14] These enzymes catalyse reactions in heme biosynthesis, biotin biosynthesis and threonine degradation, respectively. Recent additions to the growing list of identified AOS enzymes are the bacterial quorum-sensing autoinducer synthases CqsA in *Vibrio cholerae*[15–17] and LqsA in *Legionella pneumophila*.[17]

The AOS enzymes utilise a PLP cofactor at the active site to catalyse the Claisen-like condensation between an amino acid and an acyl-CoA substrate. The first high resolution crystal structure of the holo-form of a bacterial, homodimeric SPT

[a] *EaStChem, School of Chemistry, The University of Edinburgh, West Mains Road, Edinburgh, EH9 3JJ, Scotland, UK*
[b] *Centre for Biomolecular Sciences, Scottish Structural Proteomics Facility, The University of St Andrews, Fife, KY16 9ST, Scotland, UK*

† The atomic coordinates and structure factors of the structure reported in this paper have been submitted to the Protein Data Bank and are awaiting acceptance.
‡ Electronic supplementary information (ESI) available: [DETAILS]. See DOI: 10.1039/c003743e

This journal is © The Royal Society of Chemistry 2010

from *Sphingomonas paucimobilis* clearly shows the cofactor covalently attached to the side-chain of a conserved Lys^{265} residue *via* a Schiff's base (also known as an internal aldimine) at the dimer interface.[18] Transaldimination occurs when the L-serine substrate binds at the active site to form an external aldimine that is stabilised by other conserved residues: a His^{159} that stacks above the PLP ring and His^{234} that hydrogen bonds with the PLP-bound intermediate.[19] The crystal structure of the enzyme-bound PLP:L-serine external aldimine complex has recently been resolved for SPTs from the sphingolipid-producing bacteria *S. paucimobilis*[20] and *Sphingobacterium multivorum*.[21] The proposed steps subsequent to the formation of the external aldimine are deprotonation at C_α of the external aldimine complex to form a quinonoid (carbanion equivalent) intermediate; a Claisen condensation with the acyl-CoA substrate and loss of free CoASH to form a β-ketoacid intermediate; decarboxylation to form a product quinonoid; protonation of this quinonoid to form the product external aldimine; release of the α-oxoamine product and regeneration of the enzyme PLP internal aldimine (Fig. 1).

Both enantiomers of cycloserine (Fig. 2A) can be thought of as cyclic analogues of serine and/or alanine and have been shown to be irreversible inhibitors of many PLP-dependent transaminases,[22] racemases[23] and decarboxylases.[24] L-Cycloserine (LCS) is prepared synthetically whereas D-cycloserine (DCS) is a natural product isolated from *Streptomyces* strains[25] and is a broad spectrum antibiotic. Due to its severe side effects DCS is most commonly used as a second-line drug in combination therapy to treat tuberculosis.[26] DCS is also used in neurological studies since it is a potent agonist of *N*-methyl-D-aspartic acid (NMDA) receptors that are involved in neurotransmission.[27] One of its main anti-bacterial targets is the PLP-dependent alanine racemase,[28,29] an essential enzyme that generates D-alanine for the formation of the D-alanyl–D-alanine dipeptide incorporated into the bacterial peptidoglycan layer. Unlike many irreversible inhibitors that inactivate their protein targets by covalent modification, CS renders its targets inactive by forming a stable adduct with the essential PLP cofactor. It is not surprising that CS has also been shown to be a potent inhibitor of bacterial SPT[30] as well as mammalian SPT in mouse brain microsomes.[31] Although LCS is commonly used as a regulator of lipid metabolism in biological research, its mechanism of inhibition is still unknown. In this paper we used a combination of X-ray crystallography, mass spectrometry, UV-Vis spectroscopy and enzyme kinetics to elucidate the mechanism of SPT inactivation by both enantiomers of CS. We highlight differences in the enantiospecific inhibition observed for both LCS and DCS and provide further insight into PLP-dependent chemistry.

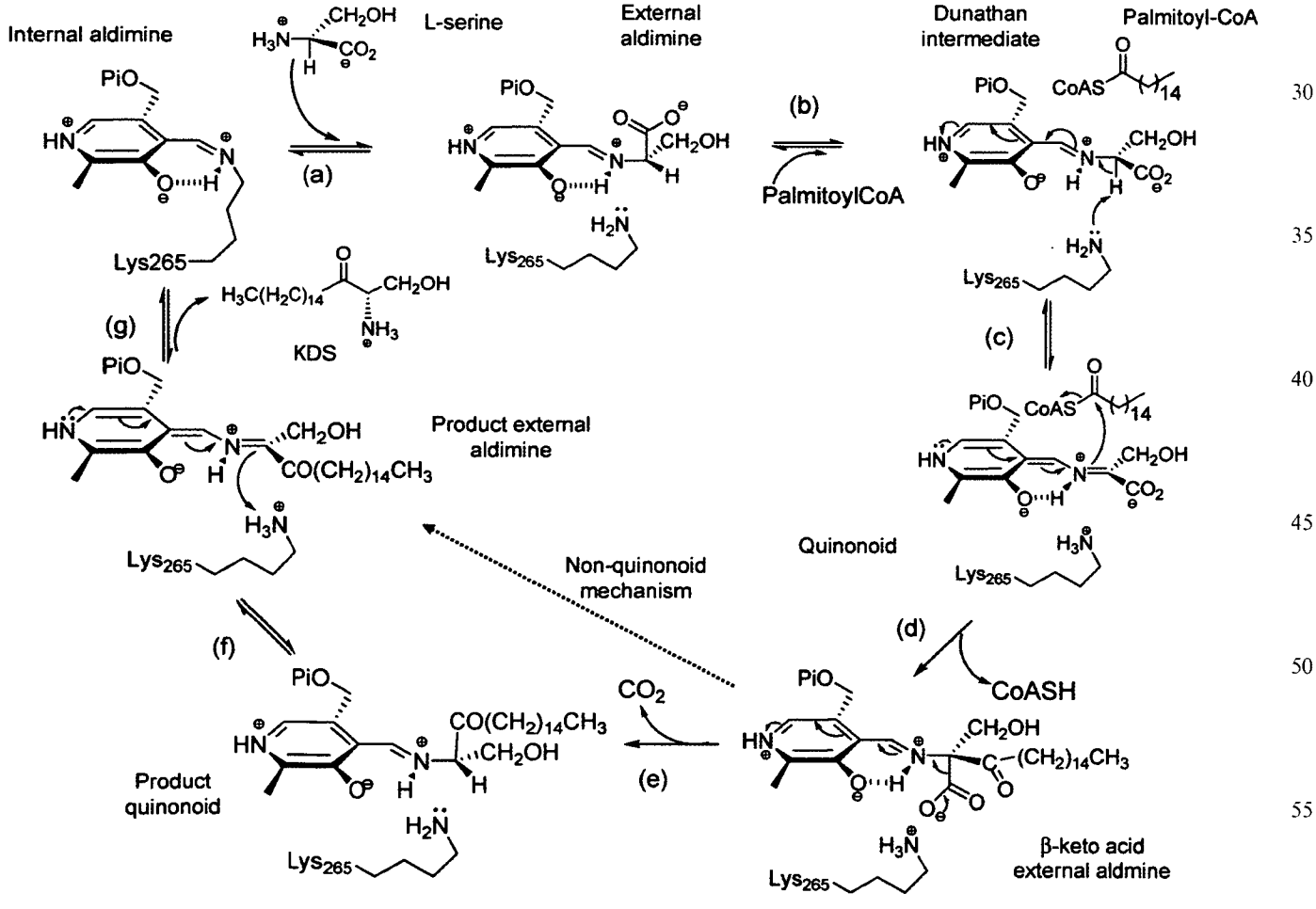


Fig. 1 Catalytic mechanism of SPT.

This journal is © The Royal Society of Chemistry 2010

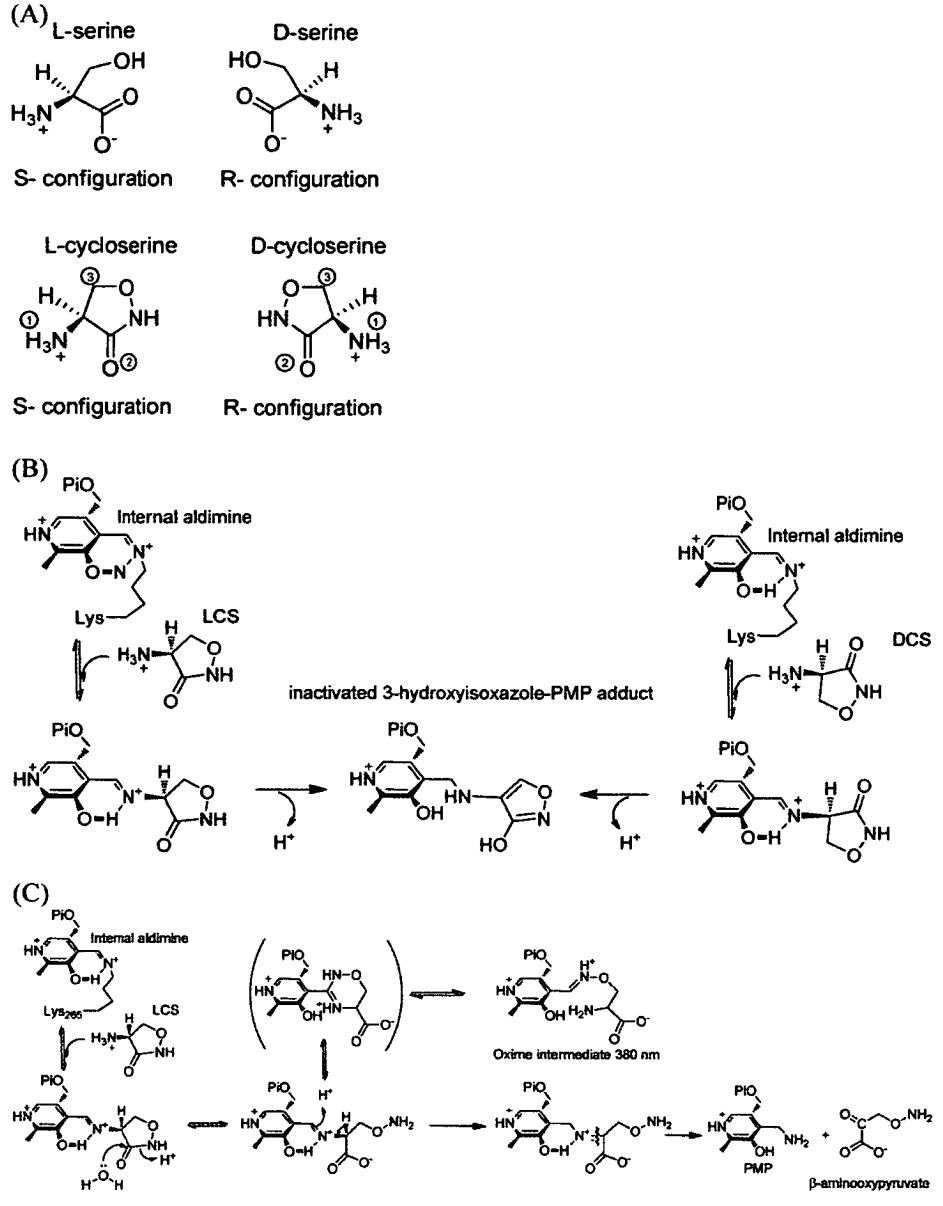


Fig. 2 (A) Structures of SPT substrates and inhibitors. (B) Mechanism of formation of the aromatic adduct observed in each of the crystal structures shown in Table 1. (C) Proposed mechanism of LCS inactivation of SPT by Ikushiro *et al.*[30]

Results

Both LCS and DCS are irreversible inhibitors of serine palmitoyltransferase

When either LCS or DCS was incubated with SPT, enzyme activity decreased over time. Log plots of % activity remaining *versus* time, according to the method of Kitz and Wilson,[32] were linear for each enantiomer indicating classical time-dependent inactivation of SPT activity (Fig. 3A and B). We noted that the Kitz and Wilson secondary plots (Fig. 3C) appeared to pass very close to the origin—there are many examples of this in the literature such as the inhibition of the thiamine-dependent enzyme benzaldehyde lyase (BAL) by methyl benzoylphosphonate[33] and the inhibition of the PLP-dependent γ-aminobutyric acid aminotransferase (GABA-AT) by (1*S*,3*S*)-3-amino-4-difluoromethylene-cyclopentanecarboxylic acid.[34] Kitz and Wilson replots of this nature suggest that both DCS and LCS form very weak complexes with SPT but inactivation is fast compared with the formation of the CS:SPT complexes. Moreover, these prevent the estimation of the k_{inact} and K_I values directly. In comparison with LCS, we found that approximately 15-fold higher concentrations of DCS were required to inactivate SPT activity to the same extent and over a similar period of time. From the secondary plot of $1/k_{app}$ *versus* 1/[cycloserine] the second-order rate constant of inactivation k_{inact}/K_I was calculated as 1/slope for each inhibitor (Fig. 3C). LCS ($k_{inact}/K_I = 0.83 \pm 0.5$ M^{-1} s^{-1}) was found to be ~14-fold more effective at inactivating SPT than DCS ($k_{inact}/K_I = 0.06 \pm 0.002$ M^{-1} s^{-1}).

Activity of SPT, measured by monitoring formation of the free thiol of CoASH using a continuous DTNB assay,[20] was reduced to 1% and 22% after 2 hours inhibition with 5 mM

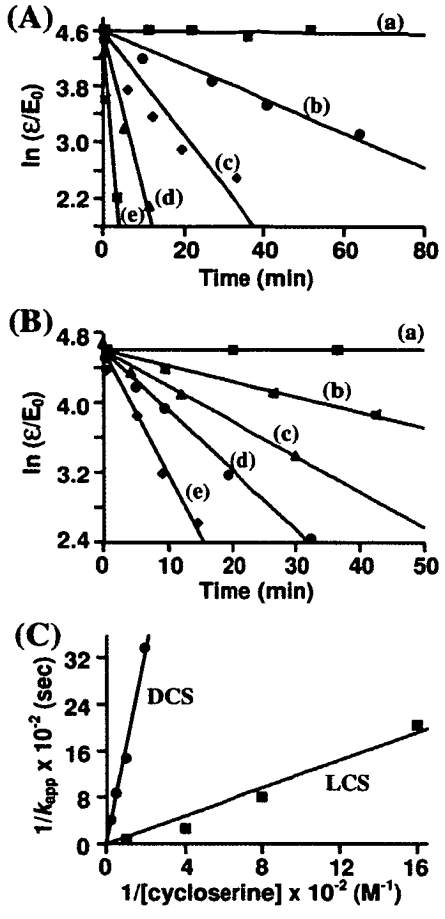


Fig. 3 (A) Inactivation of SPT by LCS at inhibitor concentrations of 0, 0.625, 1.25, 2.5 and 5 mM. (B) Inactivation of SPT by DCS at inhibitor concentrations of 0, 10, 20, 40 and 80 mM. (C) Secondary plot of $1/k_{app}$ *versus* 1/[inhibitor] for LCS and DCS.

LCS and DCS, respectively (Fig. 4). When these samples were dialysed against buffer in the absence of free PLP we noted that enzyme activity was not recovered. In contrast, activity returned to 83% (LCS) and 79% (DCS) of the original activity after dialysis against buffer containing 25 μM PLP. The fact that activity can be recovered upon dialysis against PLP shows that inactivation by both enantiomers does not occur through covalent modification of the protein but rather by disabling the PLP cofactor.

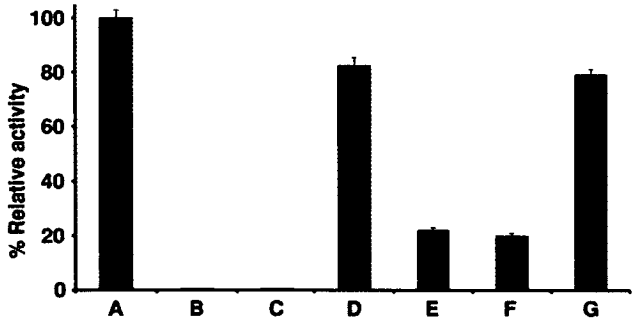


Fig. 4 Inhibition of SPT activity by LCS and DCS and regeneration of activity by dialysis against PLP buffer. SPT activity of the following samples was measured by DTNB assay: (A) 20 μM SPT, (B) 20 μM SPT and 5 mM LCS at 25 °C after 2 hour incubation, (C) sample B after dialysis against 20 mM potassium phosphate buffer, pH 7.5, (D) sample B after dialysis against 20 mM potassium phosphate, pH 7.5, and containing 25 μM PLP, (E) 20 μM SPT and 5 mM DCS at 25 °C after 2 hour incubation, (F) sample E after dialysis against 20 mM potassium phosphate buffer, pH 7.5, and (G) sample E after dialysis against 20 mM potassium phosphate, pH 7.5, and containing 25 μM PLP.

UV-visible spectroscopy analysis of LCS and DCS binding to holo-SPT

The UV-visible spectrum of holo-SPT displays absorbance maxima at 336 nm and 425 nm corresponding to the enolimine and ketoenamine forms of the PLP-bound enzyme, respectively (Fig. 5A and B, solid line). When LCS (5 mM) was added to the enzyme at pH 7.5 and 25 °C, notable changes in these peaks occurred suggesting that LCS interacts with the PLP cofactor (Fig. 5A, dashed lines). Over a period of ~30 minutes the ketoenamine peak (425 nm) was reduced to ~10% of its original value with concomitant growth of new peaks at 330 nm and 380 nm. No further changes in this UV-vis spectrum were observed when the sample was incubated for ~8 h. This suggests that the PLP cofactor bound to residue Lys^{265} as an internal aldimine is displaced and one or more new species are formed.

A number of mechanisms have been proposed for cycloserine inhibition of a range of PLP-dependent enzymes and have been collated by Olson *et al.*[35] In one mechanism originally proposed for the inhibition of alanine racemase, the authors speculate that the cycloserine ring of the inhibitor–PLP adduct is opened after attack by either an active site

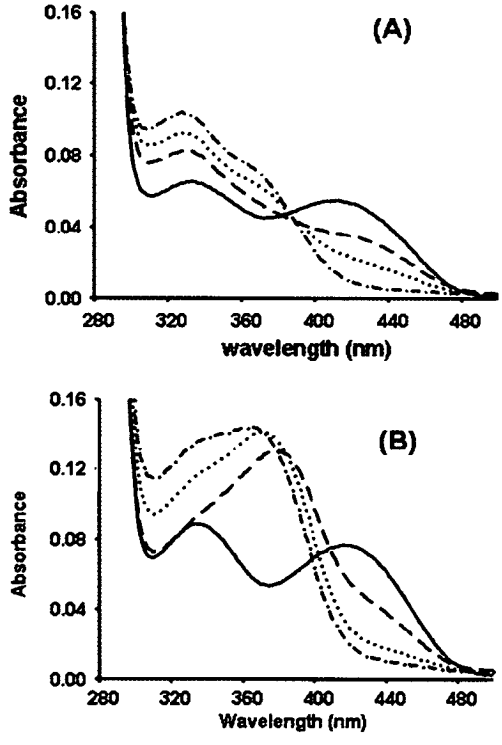


Fig. 5 (A) SPT and 5 mM LCS at time 0 (solid line), 30 seconds (long dash), 1 minute (dotted line) and 30 minutes (dash dot). (B) SPT and 5 mM DCS at time 0 (solid line), 30 seconds (long dash), 2 hours (dotted line) and 10 hours (dash dot).

This journal is © The Royal Society of Chemistry 2010

nucleophile on the enzyme or hydroxide ion.[36,37] In the case of *S. paucimobilis* SPT Ikushiro *et al.*[30] proposed that hydrolysis of the ring-open adduct leads to the formation of free pyridoxamine 5′-phosphate (PMP) and β-aminooxypyruvate (Fig. 2C). They speculated that the 380 nm peak that they observed upon incubation of SPT with LCS could be due to formation of a transient oxime intermediate while the 330 nm peak is due to the ring-opened adduct prior to hydrolysis (Fig. 2C). Of interest, a similar UV-vis spectrum was obtained in a study of the interaction of LCS with ArnB, the PLP-dependent aminotransferase involved in amino arabinose biosynthesis.[38] X-Ray structure analysis of that inactivated enzyme revealed a bound hydroxyisoxazole–PLP adduct at the ArnB active site. Therefore it appears that LCS can interact with its target enzymes in different ways with the cycloserine ring remaining intact or undergoing ring cleavage.

Dramatic changes in the PLP absorbance spectrum of holo-SPT were also observed within 30 seconds after addition of DCS (Fig. 5B, dashed line) but changes thereafter were much slower compared to LCS, occurring over a period of ~10 hours (Fig. 5B, dotted and dot-dash lines). These spectroscopic changes correlate with the observation that DCS inactivates SPT activity at a slower rate than LCS (see Fig. 3). Loss of the 425 nm peak was accompanied by appearance of a new peak at 380 nm. Over time (~2–10 h) this peak shifted to 365 nm along with the slow appearance of a pronounced, broad shoulder with λ_{max} at 330 nm. In contrast to the changes that occurred in the presence of LCS the 380 nm peak was the most dominant and not the 330 nm peak. Since DCS is a much slower inhibitor than LCS a build-up of intermediates along with slower formation of products might be expected and tends to agree with Ikushiro's proposition[30] that the LCS-derived 380 nm peak could be that of an oxime intermediate, while the slow build-up of products is represented by the slowly-formed 330 nm shoulder.

The UV-vis spectra of holo-SPT in the presence of LCS and DCS (Fig. 5A and B) are very different to each other. In contrast, the spectra of free PLP in the presence of both CS enantiomers are identical and display maxima at 360 nm (Fig. S1A and B, ESI‡). Therefore it is clear that the enzyme does not simply accommodate a PLP–cycloserine aldimine that is formed between the cofactor and inhibitor in solution. The new 330 nm peak observed in the spectrum of the enzyme in the presence of both CS enantiomers is similar to the λ_{max} (320 nm) for free PMP (Fig. S1C, ESI‡) and could be due to bound PMP.

Identification of SPT–cycloserine reaction products by LC-MS

The reaction products after inactivation of SPT by both CS enantiomers were analysed by LC ESI-MS (Fig. S2, ESI‡). After 30 min inactivation, low molecular weight products were separated from the protein by centrifugation through a filter with a 3 kDa cut-off. Analysis of this sample by LC ESI-MS showed a peak in the single ion monitoring (SIM) at m/z = 248 with a total ion count (TIC) above background at 2.8 min that corresponds to PMP (Fig. S2B, ESI‡). We were unable to detect an ion above background corresponding to the aminooxyaldehyde so the filtrate was derivatised with 2,4-dinitrophenylhydrazine (2,4-DNP) to form the corresponding hydrazone (chemical formula $C_8H_9N_5O_5$ = 255.06, Fig. S2A, ESI‡). When the 2,4-DNP derivatised sample was analysed by LC-MS a clear peak eluted from the C18 column at 3.1 minutes that ionized very well above background (Fig. S2C, ESI‡). Electrospray MS analysis of this peak revealed it to have m/z = 256 (Fig. S2C, ESI‡) which corresponds to the $[M + H]^+$ ion of the 2,4-DNP derivative of β-aminooxyacetaldehyde derived from LCS. No peaks were observed for free PLP and a 2,4-DNP derivative of the PLP aldehyde. These peaks were not observed for control SPT samples treated in the same way but in the absence of cycloserine (data not shown). We carried out a similar analysis of SPT after incubation with DCS. We observed the same peaks as seen in the LCS analysis but the total ion counts were much lower overall and the UV trace was more complicated, unsurprising given the slower rate of reaction with this enantiomer (data not shown). We conclude that LCS inactivation of SPT results in the formation of PMP and β-aminooxyacetaldehyde from the cycloserine ring.

Structure of *S. paucimobilis* SPT after inactivation by L-cycloserine

Crystals of SPT:LCS-derived complex were prepared by co-crystallisation in the presence of 1 mM LCS. Colourless crystals were obtained that diffracted to a resolution of 1.4 Å. We used our previous 1.3 Å resolution structure of the holo-form of *S. paucimobilis* SPT with the PLP and water molecules to solve the structure of this new complex.[18] It was clear that in the SPT:LCS complex the PLP cofactor is no longer covalently attached to the sidechain of Lys^{265} (contoured at 3σ and 0.2 $Å^{-3}$) and that the lysine side-chain adopts a conformation similar to that observed in the SPT:L-ser external aldimine complex[20] (Fig. 6A). The experimental electron density cannot definitively assign the co-factor as either PLP (an aldehyde) or PMP (amine), as the density is very weak for the terminal atom. However, in light of the mass spectrometry results we have modeled the co-factor as PMP. There is no additional density that could be modeled by either free cycloserine or an isoxazole ring, which have been observed in other cycloserine-inactivated PLP dependent enzymes (Table 1). Nor could such additional groups fit into the structure given the position of the side chains and a clearly identified water molecule (Fig. 6B). We do note that unusually the side chain of Tyr^{73} is disordered in the structure. The aromatic ring of the pyridoxal co-factor essentially rotates ~23° about an axis which links the C2 and C5 atoms. The phosphate group has not moved and the interactions with the protein are largely conserved, although there are some changes in water structure. In this new structure the side chains of Arg^{378} and Gln^{357}, which were noted to make a salt interaction in the holo-form,[18] are disordered confirming this is a flexible region of the protein. In contrast, in the PLP:L-ser complex the loop containing Arg^{378} moves into the active site to allow the Arg^{378} side-chain to make a salt contact with L-ser carboxylate

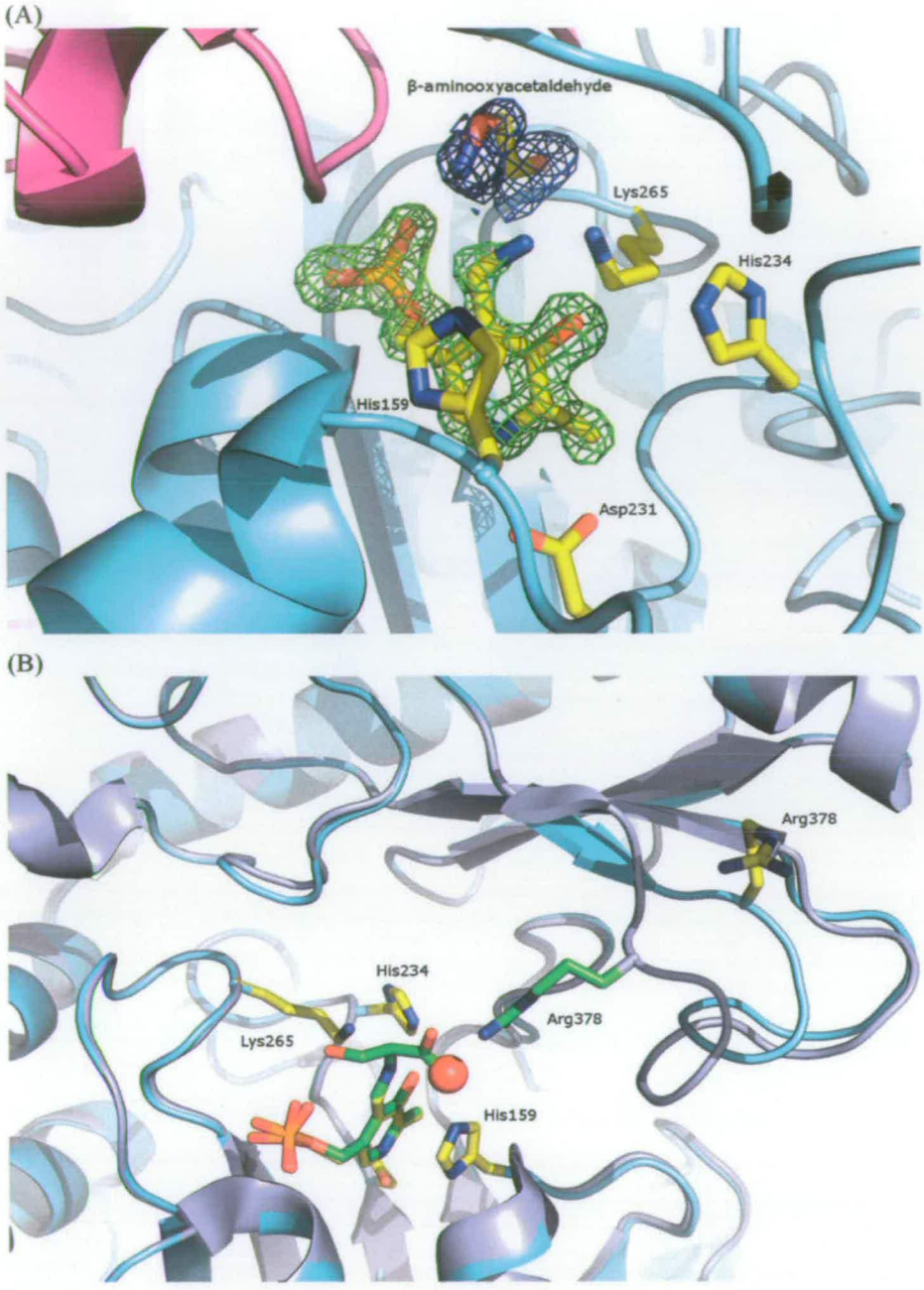


Fig. 6 (A) Fo − Fc electron density for the PMP molecule. This was calculated from molecular replacement model which was refined and had omitted the co-factor. The map is contoured in green at 3σ (0.2 $Å^{-3}$). Also shown are the side chains of Lys265, His159, Asp231 and His234. Monomer A is shown in ribbon in cyan and monomer B in magenta. The additional Fo − Fc electron density "blob" is shown in blue, contoured at 2.7σ (0.2 $Å^{-3}$). A molecule of the β-aminooxyacetalaldehyde identified by mass spectrometry is placed in the density. Carbons are colored yellow, nitrogen blue, oxygen red and phosphorous orange. (B) Overlay of SPT:L-ser (2bwj) and the LCS inactivated form. The loop containing R378 adopts the "swung-out" conformation in the LCS form, in contrast to the "swung in" conformation in SPT:L-ser. The color scheme for the LCS inhibited form is as (A). For the SPT:L-ser structure, monomer A is colored light blue. Carbons are colored green, and other atoms are colored the same as in (A). The main chain of R378 adopts a very different conformation from the SPT:L-ser structure because the salt contact with L-ser is missing. A well ordered molecule (red sphere) in the LCS structure is found in the same location as the L-ser carboxylate. The side chain of L-ser points towards the unfitted blob at the active site. This figure was prepared using PyMol (http://www.pymol.org/).

(Fig. 6B). Interestingly there is an additional 'blob' of difference electron density lying adjacent to the PMP cofactor and parallel to Lys^{265}. This blob sits at the interface of the two subunits of the functional SPT dimer. If the density is contoured at 0.16 e $Å^{-3}$ (2.7σ) it has a continuous nature (Fig. 6B). We have not fitted this density in the final model but in light of the mass spectrometry data have rendered it as the β-aminooxyacetaldehyde. We have been unable to grow X-ray diffraction quality crystals of DCS-inactivated SPT and this is the focus of current research.

This journal is © The Royal Society of Chemistry 2010

Table 1 PDB entries for PLP-dependent enzymes inactivated by CS

Enzyme	Organism	Enantiomer	PDB	References
Alanine racemase	*Enterococcus faecalis*	DCS	3E6E	60
Alanine racemase	*Escherichia coli*	DCS	2RJH	61
Alanine racemase	*Streptomyces lavendulae*	LCS	1VFT	62
Alanine racemase	*Streptomyces lavendulae*	DCS	1VFS	62
Alanine racemase	*Bacillus stearothermophilus*	LCS	1NIU	23
Alanine racemase	*Bacillus stearothermophilus*	DCS	1EPV	23
Aminodeoxychorismate lyase	*Escherichia coli*	DCS	1I2L	63
ArnB aminotransferase	*Salmonella typhimurium*	LCS	1MDZ	38
Dialkylglycine decarboxylase	*Pseudomonas cepacia*	DCS & LCS	1D7S	24
D-Amino acid transferase	*Bacillus species*	DCS	2DAA	46

Inactivation of a SPT R378N mutant by LCS and DCS

We recently determined the crystal structure of the SPT:L-serine external aldimine complex that allowed us to identify any changes that occur during the formation of this key intermediate from the internal aldimine holo-SPT form. The major difference between the two forms was the movement of Arg^{378} from a "swung-out" position (bound to the side-chain of Gln^{357}) to a "swung in" position allowing it to form a salt bridge with the $-CO_2^-$ group of the L-ser substrate[20] (Fig. 6B). This Arg^{378} residue appears to be highly mobile in *S. paucimobilis* SPT and we had probed its role in catalysis and found that a SPT R378N mutant was still active albeit at ~40× fold reduced rate when compared with the wild-type enzyme. We inhibited the SPT R378N mutant with LCS and DCS to investigate the mechanism of cycloserine inhibition. The UV-visible spectrum of the SPT R378N holoenzyme is similar to that of the wild-type SPT showing absorbance maxima at 336 nm and 415 nm corresponding to the enolimine and ketoenamine tautomers of the PLP-bound enzyme, respectively (Fig. 7A and B, solid line). When LCS (5 mM) was added to the mutant enzyme, a decrease in the 415 nm peak was observed over 30 min along with the appearance of a new peak at 330 nm and a small shoulder at 380 nm (Fig. 7A). Upon completion of the LCS inactivation of the SPT R378N mutant we noted that the 380 nm peak (Fig. 7A) is notably smaller than that observed in the wild-type enzyme (Fig. 5A, dashed line) suggesting that in the absence of Arg^{378} the mutant enzyme is unable to stabilize the putative oxime intermediate.

In contrast to the LCS enantiomer, when DCS was added to the SPT R378N mutant, dramatic changes in the UV-visible spectrum occurred within the mixing time (Fig. 7B, dashed line). Most noticeable was the appearance of a new peak at 495 nm which was transient and disappeared after 30 seconds. This could be due to the formation of a quinonoid species—these have been observed with SPT and other members of the AOS family only when incubated with substrates and substrate analogues and products, and have characteristic absorbance maxima >480 nm.[11,15,16,39,40] In the case of SPT, Ikushiro *et al.*[39] used a palmitoyl-CoA thioether analogue to initiate deprotonation of the PLP:L-ser external aldimine and the resulting quinonoid was observed as a new peak in the UV-visible spectrum with λ_{max} = 493 nm. In addition to the 495 nm peak, DCS binding to SPT R378N induced formation of a broad peak with a plateau from

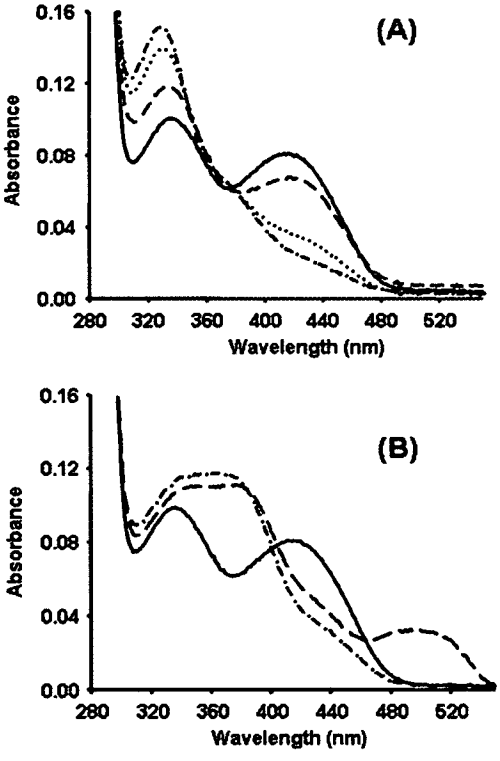


Fig. 7 (A) R378N and 5 mM LCS at time 0 (solid line), 30 seconds (long dash), 4 minutes (dotted line) and 30 minutes (dot dash). (B) R378N and 5 mM DCS at time 0 (solid line), mixing time (long dash), 30 min (dot dash).

340 nm to 380 nm along with a small shoulder at 425 nm (Fig. 7B). We suggest that the broad peak is likely to be a combination of overlapping bands due to a number of PLP-derived species. The peak shifted only slightly over the following 30 minutes and was an indication that maybe a DCS-derived intermediate becomes 'stalled' in the SPT R378N mutant. LC ESI-MS analyses of the low molecular weight fraction from both the CS enantiomer-inactivated R378N incubations were inconclusive *e.g.* no ions corresponding to the PMP and β-aminooxyacetaldehyde products could be detected (data not shown). This is unsurprising since the R378N mutant has such a slow turnover compared with the wild-type enzyme. Taken together, the analyses of the SPT R378N mutant with both LCS and DCS provide further evidence that the SPT enzyme is controlling the reaction pathway of each inhibitor in a stereospecific manner that

requires residues not simply involved in binding the PLP cofactor directly. Current efforts are focused on capturing CS-derived intermediates in the mutant enzyme by X-ray crystallography.

Discussion

The antibiotic DCS was first isolated from the soil bacterium Streptomyces and its structure and activity determined over 50 years ago.[41–45] Cycloserine inactivates many PLP-dependent enzymes by disabling the essential PLP cofactor at the active site and a list of the X-ray structures from the Protein Data Bank with cycloserine adducts bound in the active site is presented in Table 1. In each case inhibitor binding results in transaldimination of the PLP bound, holo-enzyme and leads to formation of a PLP:cycloserine external aldimine. Deprotonation of this species gives rise to a stable 3-hydroxy-isoxazole–PMP adduct whereby the cycloserine ring remains intact and covalently linked to the PLP cofactor—the so-called "aromatisation mechanism"[35] (Fig. 2B). Olson *et al.* studied LCS inactivation of γ-aminobutyric acid aminotransferase and Peisach *et al.* investigated D-amino acid aminotransferase inhibition by DCS.[46] The aromatized CS adduct was observed in both cases where the active site lysine residue was inferred as the base which removes the CS α-proton. It is interesting that inhibition by both enantiomers of CS leads to formation of the same isoxazole–PMP adduct in the alanine racemase that was studied by Fenn *et al.*[23] This could be due to the alternating acid/base nature of the residues (Lys^{39}, that is also involved in formation of the initial internal aldimine with PLP, and the side-chain of Tyr^{265}) from an enzyme that catalyses racemisation of both enantiomers of alanine. However, this study revealed that the kinetics of inactivation of the two CS enantiomers was different; DCS inactivated alanine racemase much faster than the LCS. This explains why DCS (seromycin) is the better antibiotic. We also found different kinetics for each CS enantiomer with SPT; the apparent half-life of inactivation was approximately 15 minutes for 1.25 mM LCS and 20 mM DCS. Kitz and Wilson analysis also revealed that both enantiomers inhibit *via* a bimolecular mechanism *i.e.* they are weak, but fast inhibitors, and suggests that carrying out incubations at lower temperatures would provide further insight into the inhibition mechanism.[32]

Since SPT only catalyses condensation between palmitoyl-CoA and L-serine (although it does bind D-serine and forms an external aldimine) we expected to observe differences in SPT inactivation by each CS enantiomer. These were manifested in differences in both the rates of inactivation, with LCS being 14 times faster than DCS, and time-dependent changes in the UV-vis spectra. We could not detect an aromatised CS derivative after inactivation, but we did identify PMP and the derivatised form of a novel aldehyde by mass spectrometry. This CS-derived product was formed from both enantiomers but to a greater extent (based on LC ESI-MS profiles) from LCS inactivation. Moreover, we also identified a PMP form of the cofactor in crystals of the LCS-inactivated SPT and a low molecular weight species non-covalently bound at the active site which clearly no longer contains a cycloserine ring and whose size and shape are consistent with the proposed β-aminoacetaldehyde product.

These combined data lead us to propose a novel decarboxylative, ring-opening mechanism for inactivation of SPT by cycloserine (Fig. 8). The first step in the normal catalytic cycle of SPT is a transaldimination step when the substrate L-ser replaces the PLP-Lys^{265} internal aldimine to form a PLP:L-ser external aldimine (Fig. 1). Deprotonation of this species does not occur until the second substrate palmitoyl-CoA binds. It is interesting to note that SPT acts upon both CS enantiomers in the absence of palmitoyl-CoA suggesting that the CS binds in the SPT active site and forms a PLP:CS external aldimine complex which then reacts readily. Previously, we captured the PLP:L-ser external aldimine and the structure revealed the Lys^{265} poised below this intermediate at a position where it would act as the base that removes the Cα proton (to form a quinonoid) once binding of the thioester causes rotation around the Cα–N bond.[20] Interestingly, D-ser can form an external aldimine with SPT but is a poor substrate presumably because the α-H is not in the correct position when palmitoyl-CoA binds.[47] To rationalise why we did not observe aromatization of the LCS inhibitor (which has the same *S*-configuration as L-ser) we suggest that the Cα proton in the PLP:LCS external aldimine cannot be in the optimal orientation to be removed by the lysine (Fig. 1 and 8A). Instead we propose ring-opening of the LCS ring by cleavage of the amide bond. Amide bond hydrolysis of peptides is catalysed by various proteases which use a powerful nucleophile (*e.g.* serine or cysteine in the serine or cysteine proteases) to attack this bond and generate an acylated intermediate which is subsequently hydrolysed. We cannot identify such a residue in the active site of SPT at present but conformational changes upon ligand binding may bring such a residue into play. An acylated ring-opened adduct was recently captured by Macheboeuf *et al.* who studied the inhibition of penicillin-binding protein (PBP) by the natural product lactivicin (LTV).[48] This interesting molecule contains separate cycloserine and γ-lactone rings and is the only known natural PBP inhibitor that does not contain a β-lactam. Crystallographic analysis of the PBP revealed that LTV inhibition involves opening of both the cycloserine and γ-lactone rings upon attack by the active site serine nucleophile. We have accommodated this possible route (Fig. 8, path (a)) into the mechanism and the acylated intermediate formed could be hydrolysed to give the carboxylated PLP intermediate.

We also suggest an alternative mechanism involving direct attack of water on the CS ring (Fig. 8, path (b)). This is justified when one considers the properties of cycloserine. During the initial characterization of the DCS natural product Kuehl *et al.* and Hidy *et al.* found the CS ring opened upon treatment with methanol and HCl to afford the β-amino-D-alanine methyl ester.[43,44] Treatment of this ester with alkali converted it back to DCS. Further studies revealed unusual chemical properties of DCS; it was found to dimerise to give a six-membered *cis*-3-6-bis(aminooxymethyl)-2,5-piperazinedione adduct in acid.[44,49] The proposed mechanism of adduct formation involves nucleophilic attack of one CS amino group at the carbonyl position of another CS

This journal is © The Royal Society of Chemistry 2010

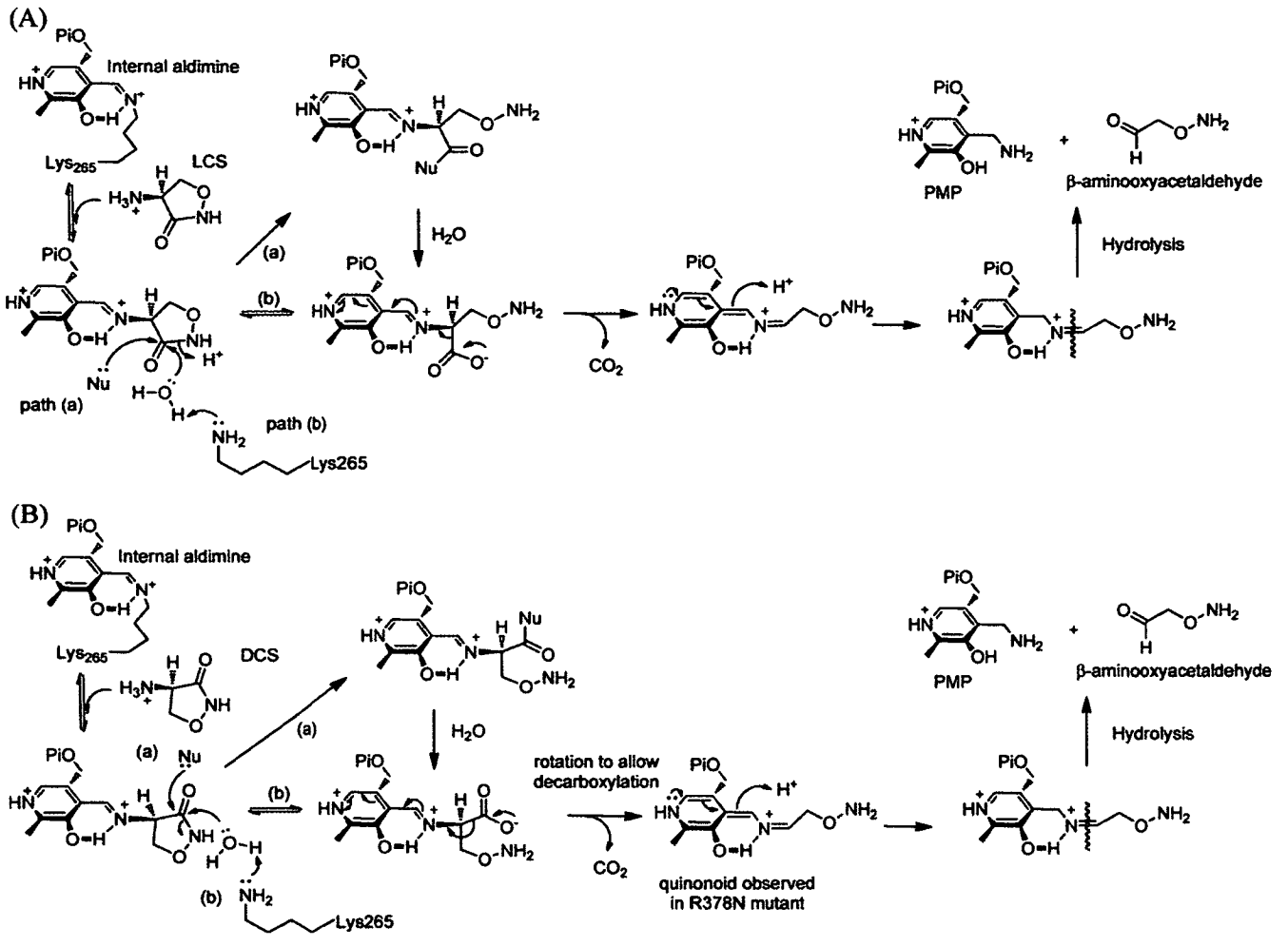


Fig. 8 (A) Novel ring-opening, decarboxylative mechanism for inactivation of SPT by LCS. Path (a) denotes an enzymatic, nucleophile(Nu)-mediated mechanism with an acylated intermediate. Path (b) is the direct hydrolytic mechanism. (B) Inactivation of SPT by DCS highlighting a proposed quinonoid intermediate observed in the SPT R378N mutant. The paths are the same as in (A).

molecule, which then ring opens at the amide bond to generate the aminooxy form. This reaction occurs again in an intramolecular fashion to generate the piperazine ring. So, in the context of a SPT catalysed mechanism ring-opening could also proceed through acid catalysis and a number of potential proton donors are found in the active site (*e.g.* His159).

We predict two possible routes for the ring-opened adduct. Firstly, it can form an oxime intermediate, observed as a 380 nm peak in the UV-visible spectrum (Fig. 2C and 5A). In the second, novel route, the ring-open adduct decarboxylates, which is then followed by imine hydrolysis to form the products, PMP and β-aminooxyacetaldehyde (Fig. 8). This fits well with the SPT mechanism since the enzyme catalyses decarboxylation of the putative β-keto acid Claisen condensation intermediate to give a product external aldimine which is then hydrolysed to release KDS (Fig. 1). Since LCS effectively acts as an amino donor to generate PMP, the enzyme has to be regenerated by addition of excess PLP (Fig. 4). In contrast to LCS, DCS is a much slower inactivator of SPT and this suggests that it forms a PLP:DCS external aldimine which can react *via* a similar decarboxylative mechanism, possibly requiring some conformational change of the ring opened intermediate to promote this step (Fig. 8B). To probe the formation of intermediates during SPT inactivation by both enantiomers we utilised a SPT R378N mutant since this residue forms a salt bridge with the carboxylate of the external aldimine with L-ser. Again, we saw clear enantiospecific differences in the behaviour of this mutant with LCS and DCS by UV-vis spectroscopy (Fig. 7). Incubation of SPT R378N with LCS gave rise to a similar spectrum to the wild-type enzyme but lacked a shoulder at 380 nm (Fig. 7A). In contrast, we were excited to observe transient formation of a quinonoid species at 510 nm during DCS inactivation of the R378N mutant (Fig. 7B) and we suggest that this is due to the quinonoid formed by decarboxylation (Fig. 8B). Of interest in a previous study Fenn *et al.* used a Y265F mutant of alanine racemase to probe the different LCS and DCS inactivation mechanisms of this enzyme.[37] The Y265F mutant produced only the aromatised isoxazole PLP adduct with DCS, whereas LCS inactivation led to formation of a range of products including PMP and acetate. They concluded that this active site residue "steers" the pathway of inactivation. In our work, mutation of Arg^{378} has revealed the presence of

transient species and may be useful to capture intermediates in future.

LCS is commonly used in biological research to inhibit *de novo* biosynthesis of sphingolipids in mammalian cells but it remains debatable whether it specifically targets SPT. LCS was shown to be a potent inhibitor of SPT activity derived from mouse brain microsomes[50] and inhibited mouse brain SPT activity *in vivo* after intraperitoneal injection.[51] On the other hand cycloserine toxicity in Chinese hamster ovary cells was not attributed to SPT inhibition since exogenous sphingosine could not rescue cell death.[52] Further research is required to examine whether LCS targets SPT alone or can inhibit other PLP-dependent enzymes on the sphingolipid biosynthetic pathway such as sphingosine-1-phosphate lyase. With knowledge gained from our study of the bacterial, homodimeric SPT, it will be interesting to characterise the products of LCS and DCS inhibition of the heterodimeric human enzyme. Whether the β-aminooxyacetaldehyde derivative is produced by the more complex mammalian isoforms remains to be seen. The identification of two small subunits (ssSPTa and ssSPTb) that stimulate human SPT activity > 100 fold should facilitate the analysis of cycloserine inactivation products.[53]

Interestingly DCS is a much more potent inhibitor of alanine racemase compared to the synthetic LCS.[23] When the antibacterial effect of DCS and LCS was tested on a large number of bacterial strains, most had lower tolerance towards DCS.[54] In contrast, the sphingolipid-producing bacterium *Bacteroides levii* was much more susceptible to LCS than DCS which correlated to inhibition of SPT.[50] Future investigations are necessary to determine whether LCS could target sphingolipid-producing bacterial pathogens.

Materials

Plasmids and *Escherichia coli* competent cells were purchased from Novagen, and all chromatography columns were from GE Healthcare. All buffers and reagents including LCS and DCS were from Sigma. Palmitoyl-CoA was from Avanti Lipids.

Methods

Cloning and expression of *S. paucimobilis* SPT wild-type and R378N mutant

The SPT wild-type gene and R378N mutant were cloned in pET28a expression vector (Novagen) as previously described.[20] The plasmids were transformed into *E. coli* BL21 (DE3) competent cells and selection was carried out on LB agar containing 30 μg ml^{-1} kanamycin. A single colony was used to inoculate an overnight culture grown at 37 °C in 500 ml 2YT broth (16 g l^{-1} Bacto-tryptone, 10 g l^{-1} Bacto-yeast extract, 5 g l^{-1} sodium chloride (pH 7.5)). This culture was added to 4 litres of 2YT supplemented with kanamycin and grown to OD_{600} of 0.6 before addition of 0.1 mM isopropyl 1-thio-β-D-galactopyranoside to induce protein expression. Growth was continued for 5 hours at 30 °C. Cells were harvested (Sorvall RC5B centrifuge) by centrifugation at 3500 rpm for 20 minutes at 4 °C. The enzyme was purified using IMAC on nickel resin (Invitrogen) as well as size exclusion chromatography (Sephadex S200HR, GE Healthcare). Before UV-vis spectroscopy and inactivation assays, the enzyme was freshly converted to the PLP-bound form by dialysis against buffers containing 25 μM PLP.

UV-visible spectroscopy of SPT inhibition by cycloserine

All UV-visible spectra were recorded on a Cary 50 UV-visible spectrophotometer (Varian) and analysed using Cary WinUV software (Varian). Enzyme was dialysed against 20 mM potassium phosphate (pH 7.5) containing 150 mM NaCl and 25 μM PLP for 4 hours at 4 °C. Excess PLP was removed on a PD-10 (Sephadex G-25M) desalting column (GE Healthcare). For UV-visible assays, the concentration of recombinant protein was 20 μM. The spectrophotometer was blanked with 20 mM potassium phosphate (pH 7.5) containing 150 mM NaCl and spectra were collected from 800 nm to 200 nm. Quartz cuvettes from NSG Precision Cells, Type 18-BM (Material quartz, Lightpath 10 mm), had a sample volume of 500 μl. By setting the instrument to cycle mode and collecting spectra at time intervals, changes in the UV-visible spectrum were monitored after addition of cycloserine.

Rates of SPT inactivation by cycloserine using DTNB assay

SPT activity was measured using the DTNB assay as previously described.[20] Assays contained enzyme, substrates and DTNB with final concentrations as follows: 1 μM enzyme, 25 mM L-serine, 250 μM palmitoyl-CoA, 0.2 mM DTNB in 20 mM HEPES buffer, pH 8.0. SPT was incubated with LCS (0 mM, 0.63 mM, 1.25 mM, 2.5 mM and 5 mM) and DCS (0 mM, 10 mM, 20 mM, 40 mM and 80 mM). At time intervals an aliquot of the enzyme plus inhibitor solution was diluted into the assay buffer to record the activity. The remaining activity was calculated as a percentage of the total activity in the absence of inhibitor *i.e.* at time zero. Kitz and Wilson plots of ln % remaining activity against time were linear.[32] Graphs were plotted using WinCurveFit software, Kevin Raner, Australia.

Isolation and derivatisation of aldehyde followed by detection by LC-MS

SPT (100 μM) was inactivated with 5 mM LCS for 2 hours followed by centrifugation at 13 000 rpm for 10 minutes in a Vivaspin 500 concentrators (VWR) to separate the low and high molecular weight material. The filtrate (0.1 ml) was added to a 2,4-DNP solution (0.5 ml) prepared in acetonitrile and containing 0.5% HCl. The solution was incubated at 37 °C for 30 minutes to allow derivatisation of the aldehyde. The hydrazone product was detected by LC-MS on a 6130 quadrupole mass spectrometer with a 1200 Series quaternary LC system (Agilent Technologies). Sample volumes were typically 10 μl under automated injection into a LC system equipped with a C18 column followed by separation in a methanol : water : formic acid (80 : 19 : 1) mobile phase. The mass spectrometer was set up in selected ion monitoring (SIM) mode to search for ions with masses corresponding

This journal is © The Royal Society of Chemistry 2010

to the β-aminoacetaldehyde and its 2,4-DNP hydrazone derivative (Fig. S2, ESI‡).

Structural biology

Crystals of LCS-inactivated SPT were prepared by co-crystallisation. A 10 mM stock solution of LCS was prepared and added to freshly prepared holo-SPT (20 μM, in 10 mM Tris 7.5, 150 mM NaCl and 25 μM PLP) to give a final LCS concentration of 1 mM. After overnight incubation at 4 °C, the protein was concentrated to 20 mg ml^{-1} and this solution was used to set up crystal trials. Colourless SPT:LCS crystals grew from a 2 μl drop consisting of 1 μl of the SPT/LCS solution and 1 μl of well solution (100 mM HEPES, pH 6.5, 110 mM $MgCl_2$, 21.5% PEG 3350). A single crystal was mounted in a cryo-loop and cryo-protected in 22% PEG 3350, 120 mM $MgCl_2$, 100 mM HEPES, pH 6.5, 20% PEG 400. The crystal was frozen and data obtained that diffracted to a resolution of 1.40 Å on beamline ID14-1 (wavelength = 0.934 Å) at 100 K. Data were collected on and processed to 1.40 Å using XDS.[55] The structure was solved by molecular replacement using the 1.3 Å holo-form as the model (with co-factor and water molecules removed). Due to an oversight a new Free-R set was devised for the structure, more rigorous would have to be use the same as the isomorphous search model. To overcome the bias in the Free-R set, all the ligands and waters were removed, the B-factors reset to a common value and then the structure was refined. TLS groups were defined using the TLS server[56] and once again the B-factors reset to a single value. The asymmetric unit contains a monomer of protein, but as with the native contains a functional dimer. The structure was refined using REFMAC5,[57] TLS groups were used and in the final steps of refinement anisotropic thermal factors were refined for all atoms. Inclusion of anisotropic B-factors decreases R-free by 0.6% and was therefore judged appropriate. PMP was modelled into the difference electron density immediately after molecular replacement, the density was unambiguous. The choice of PMP not PLP was made based on mass spectrometric data not the X-ray data themselves. Water molecules were added manually and checked using COOT.[58] A dictionary of the aldehyde was made using the PRODRG server.[59] The aldehyde was not included in the final model. Its inclusion did not perturb either the R-factor or the stereochemical quality of the model. The data collection and refinement statistics are given in Supplementary Table.‡

Abbreviations

2,4-DNP	2,4-dinitrophenylhydrazine
DCS	D-cycloserine
DTNB	dithionitrobenzoic acid
LC ESI-MS	liquid chromatography electrospray ionisation mass spectrometry
LCS	L-cycloserine
PLP	pyridoxal 5′-phosphate
PMP	pyridoxamine 5′-phosphate
SPT	serine palmitoyltransferase

Acknowledgements

This work was supported by Biotechnology and Biological Sciences Research Council (BBSRC) grants BB/F009739/1 (supporting JL) and BBS/B/14434 to DJC and JHN. VTB and MCCR were supported by University of Edinburgh, EastChem studentships. BAY was supported by a University of Edinburgh/Syngenta PhD studentship. We thank Dr Mike Greaney for access to mass spectrometry instrumentation.

References

1 J. S. O'Brien and E. L. Sampson, *J. Lipid Res.*, 1965, **6**, 537–544.
2 K. Simons and E. Ikonen, *Nature*, 1997, **387**, 569–572.
3 A. H. Futerman and Y. A. Hannun, *EMBO Rep.*, 2004, **5**, 777–782.
4 Y. A. Hannun and L. M. Obeid, *Nat. Rev. Mol. Cell Biol.*, 2008, **9**, 139–150.
5 S. T. Pruett, A. Bushnev, K. Hagedorn, M. Adiga, C. A. Haynes, M. C. Sullards, D. C. Liotta and A. H. J. Merrill, *J. Lipid Res.*, 2008, **49**, 1621–1639.
6 H. Fyrst, B. Oskouian, P. Bandhuvula, Y. Gong, H. S. Byun, R. Bittman, A. R. Lee and J. D. Saba, *Cancer Res.*, 2009, **69**, 9457–9464.
7 T. S. Park, W. Rosebury, E. K. Kindt, M. C. Kowala and R. L. Panek, *Pharmacol. Res.*, 2008, **58**, 45–51.
8 J. Bras, A. Singleton, M. R. Cookson and J. Hardy, *FEBS J.*, 2008, **275**, 5767–5773.
9 K. Hanada, *Biochim. Biophys. Acta*, 2003, **1632**, 16–30.
10 D. Alexeev, M. Alexeeva, R. L. Baxter, D. J. Campopiano, S. P. Webster and L. Sawyer, *J. Mol. Biol.*, 1998, **284**, 401–419.
11 S. P. Webster, D. Alexeev, D. J. Campopiano, R. M. Watt, M. Alexeeva, L. Sawyer and R. L. Baxter, *Biochemistry*, 2000, **39**, 516–528.
12 P. M. Jordan, *Biosynthesis of Tetrapyrroles*, Elsevier, Amsterdam, 1991.
13 G. C. Ferreira and J. Gong, *J. Bioenerg. Biomembr.*, 1995, **27**, 151–159.
14 A. Schmidt, J. Sivaraman, Y. Li, R. Larocque, J. A. Barbosa, C. Smith, A. Matte, J. D. Schrag and M. Cygler, *Biochemistry*, 2001, **40**, 5151–5160.
15 N. Jahan, J. A. Potter, M. A. Sheikh, C. H. Botting, S. L. Shirran, N. J. Westwood and G. L. Taylor, *J. Mol. Biol.*, 2009, **392**, 763–773.
16 R. C. Kelly, M. E. Bolitho, D. A. Higgins, W. Lu, W. L. Ng, P. D. Jeffrey, J. D. Rabinowitz, M. F. Semmelhack, F. M. Hughson and B. L. Bassler, *Nat. Chem. Biol.*, 2009, **5**, 891–895.
17 T. Spirig, A. Tiaden, P. Kiefer, C. Buchrieser, J. A. Vorholt and H. Hilbi, *J. Biol. Chem.*, 2008, **283**, 18113–18123.
18 B. A. Yard, L. G. Carter, K. A. Johnson, I. M. Overton, M. Dorward, H. Liu, S. A. McMahon, M. Oke, D. Puech, G. J. Barton, J. H. Naismith and D. J. Campopiano, *J. Mol. Biol.*, 2007, **370**, 870–886.
19 A. C. Eliot and J. F. Kirsch, *Annu. Rev. Biochem.*, 2004, **73**, 383–415.
20 M. C. Raman, K. A. Johnson, B. A. Yard, J. Lowther, L. G. Carter, J. H. Naismith and D. J. Campopiano, *J. Biol. Chem.*, 2009, **284**, 17328–17339.
21 H. Ikushiro, M. M. Islam, A. Okamoto, J. Hoseki, T. Murakawa, S. Fujii, I. Miyahara and H. Hayashi, *J. Biochem. (Tokyo)*, 2009, **146**, 549–562.
22 T. S. Soper and J. M. Manning, *J. Biol. Chem.*, 1981, **256**, 4263–4268.
23 T. D. Fenn, G. F. Stamper, A. A. Morollo and D. Ringe, *Biochemistry*, 2003, **42**, 5775–5783.
24 V. N. Malashkevich, P. Strop, J. W. Keller, J. N. Jansonius and M. D. Toney, *J. Mol. Biol.*, 1999, **294**, 193–200.
25 M. L. Svensson and S. Gatenbeck, *Arch. Microbiol.*, 1982, **131**, 129–131.
26 G. Di Perri and S. Bonora, *J. Antimicrob. Chemother.*, 2004, **54**, 593–602.

27 A. Sheinin, S. Shavit and M. Benveniste, *Neuropharmacology*, 2001, **41**, 151–158.
28 J. L. Strominger, E. Ito and R. H. Threnn, *J. Am. Chem. Soc.*, 1960, **82**, 998–999.
29 F. C. Neuhaus and J. L. Lynch, *Biochemistry*, 1964, **3**, 471–480.
30 H. Ikushiro, H. Hayashi and H. Kagamiyama, *Biochemistry*, 2004, **43**, 1082–1092.
31 K. S. Sundaram and M. Lev, *J. Neurochem.*, 1984, **42**, 577–581.
32 R. Kitz and I. B. Wilson, *J. Biol. Chem.*, 1962, **237**, 3245–3249.
33 G. S. Brandt, N. Nemeria, S. Chakraborty, M. J. McLeish, A. Yep, G. L. Kenyon, G. A. Petsko, F. Jordan and D. Ringe, *Biochemistry*, 2008, **47**, 7734–7743.
34 H. Yuan and R. B. Silverman, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 1651–1654.
35 G. T. Olson, M. Fu, S. Lau, K. L. Rinehart and R. B. Silverman, *J. Am. Chem. Soc.*, 1998, **120**, 2256–2267.
36 R. R. Rando, *Biochem. Pharmacol.*, 1975, **24**, 1153–1160.
37 T. D. Fenn, T. Holyoak, G. F. Stamper and D. Ringe, *Biochemistry*, 2005, **44**, 5317–5327.
38 B. W. Noland, J. M. Newman, J. Hendle, J. Badger, J. A. Christopher, J. Tresser, M. D. Buchanan, T. A. Wright, M. E. Rutter, W. E. Sanderson, H. J. Muller-Dieckmann, K. S. Gajiwala and S. G. Buchanan, *Structure*, 2002, **10**, 1569–1580.
39 H. Ikushiro, S. Fujii, Y. Shiraiwa and H. Hayashi, *J. Biol. Chem.*, 2008, **283**, 7542–7553.
40 J. Zhang and G. C. Ferreira, *J. Biol. Chem.*, 2002, **277**, 44660–44669.
41 R. L. Harned, P. H. Hidy and E. K. La Baw, *Antibiot. Chemother.*, 1955, **5**, 204–205.
42 D. A. Harris, M. Ruger, M. A. Reagan, F. J. Wolf, R. L. Peck, H. Walick and H. B. Woodroff, *Antibiot. Chemother.*, 1955, **5**, 183–190.
43 F. A. Kuehl, Jr., F. J. Wolf, N. R. Trenner, R. L. Peck, R. P. Buhs, E. Howe, I. Putter, B. D. Hunnewell, R. Ormond, G. Downing, J. E. Lyons, E. Newstead, L. Chaiet and K. Folkers, *J. Am. Chem. Soc.*, 1955, **77**, 2344–2345.
44 P. H. Hidy, E. B. Hodge, V. V. Young, R. L. Harned, G. A. Brewer, W. F. Phillips, W. F. Runge, H. E. Stavely, A. Pohland, H. Boaz and H. R. Sullivan, *J. Am. Chem. Soc.*, 1955, **77**, 2345–2346.
45 C. H. Stammer, A. N. Wilson, F. W. Holly and K. Folkers, *J. Am. Chem. Soc.*, 1955, **77**.
46 D. Peisach, D. M. Chipman, P. Van Ophem, J. M. Manning and D. Ringe, *J. Am. Chem. Soc.*, 1998, **120**, 2268–2274.
47 K. Hanada, T. Hara and M. Nishijima, *FEBS Lett.*, 2000, **474**, 63–65.
48 P. Macheboeuf, D. S. Fischer, T. Brown, Jr., A. Zervosen, A. Luxen, B. Joris, A. Dessen and C. J. Schofield, *Nat. Chem. Biol.*, 2007, **3**, 565–569.
49 F. O. Lassen and C. H. Stammer, *J. Org. Chem.*, 1971, **36**, 2631–2634.
50 K. S. Sundaram and M. Lev, *Antimicrob. Agents Chemother.*, 1984, **26**, 211–213.
51 K. S. Sundaram and M. Lev, *J. Lipid Res.*, 1985, **26**, 473–477.
52 K. Hanada, M. Nishijima, T. Fujita and S. Kobayashi, *Biochem. Pharmacol.*, 2000, **59**, 1211–1216.
53 G. Han, S. D. Gupta, K. Gable, S. Niranjanakumari, P. Moitra, F. Eichler, R. H. Brown, Jr., J. M. Harmon and T. M. Dunn, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 8186–8191.
54 F. C. Neuhaus, *D-cycloserine and O-carbamyl-D-serine*, Springer-Verlag, New York, 1967.
55 W. Kabsch, *J. Appl. Crystallogr.*, 1993, **26**, 795–800.
56 J. Painter and E. A. Merritt, *Acta. Crystallogr., Sect. D: Biol. Crystallogr.*, 2006, **62**, 439–450.
57 G. N. Murshudov, A. A. Vagin and E. J. Dodson, *Acta. Crystallogr., Sect. D: Biol. Crystallogr.*, 1997, **53**, 240–255.
58 P. Emsley and K. Cowtan, *Acta. Crystallogr., Sect. D: Biol. Crystallogr.*, 2004, **60**, 2126–2132.
59 A. W. Schuettelkopf and D. M. F. van Aalten, *Acta Crystallogr.*, 2004, **D60**, 1355–1363.
60 A. Priyadarshi, E. H. Lee, M. W. Sung, K. H. Nam, W. H. Lee, E. E. Kim and K. Y. Hwang, *Biochim. Biophys. Acta*, 2009, **1794**, 1030–1040.
61 D. Wu, T. Hu, L. Zhang, J. Chen, J. Du, J. Ding, H. Jiang and X. Shen, *Protein Sci.*, 2008, **17**, 1066–1076.
62 M. Noda, Y. Matoba, T. Kumagai and M. Sugiyama, *J. Biol. Chem.*, 2004, **279**, 46153–46161.
63 P. Y. Jensen, J. F. Parsons, K. E. Fisher, A. S. Pachikara, M. Tordova, A. J. Howard, E. Eisenstein and J. E. Ladner, 2003, DOI: 10.2210/pdb1i2l/pdb. Q7

This journal is © The Royal Society of Chemistry 2010

The Serine Palmitoyltransferase from *Sphingomonas wittichii* RW1: An Interesting Link to an Unusual Acyl Carrier Protein

Marine C. C. Raman,[1] Kenneth A. Johnson,[2] David J. Clarke,[1] James H. Naismith,[2] Dominic J. Campopiano[1]

[1] *School of Chemistry, EaStCHEM, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ*

[2] *Centre for Biomolecular Sciences, Scottish Structural Proteomics Facility, The University of St Andrews, Fife, KY16 9ST, Scotland, UK*

Received 16 March 2010; revised 22 April 2010; accepted 22 April 2010
Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/bip.21482

ABSTRACT:

Serine palmitoyltransferase (SPT) catalyses the first step in the de novo biosynthesis of sphingolipids (SLs). It uses a decarboxylative Claisen-like condensation reaction to couple L-serine with palmitoyl-CoA to generate a long-chain base product, 3-ketodihydrosphingosine. SLs are produced by mammals, plants, yeast, and some bacteria, and we have exploited the complete genome sequence of Sphingomonas wittichii to begin a complete analysis of bacterial sphingolipid biosynthesis. Here, we describe the enzymatic characterization of the SPT from this organism and present its high-resolution x-ray structure. Moreover, we identified an open reading frame with high sequence homology to acyl carrier proteins (ACPs) that are common to fatty acid biosynthetic pathways. This small protein was co-expressed with the SPT and we isolated and characterised the apo- and holo-forms of the ACP. Our studies suggest a link between fatty acid and sphingolipid metabolism. © 2010 Wiley Periodicals, Inc. Biopolymers 93: 811–822, 2010.

Additional Supporting Information may be found in the online version of this article.

Correspondence to: Dominic J. Campopiano; e-mail: dominic.campopiano@ed.ac.uk

Contract grant sponsor: BBSRC
Contract grant number: BBS/B/14434
Contract grant sponsors: EPSRC
Contract grant number: EP/C543289/1
Contract grant sponsors: EaStChem PhD studentship
Contract grant number: BBS/B/14434, BBF0085031 and BB/S/B14450

©2010 Wiley Periodicals, Inc.

Keywords: sphingolipids; Sphingomonas; serine palmitoyltransferase; pyridoxal phosphate; acyl carrier protein

INTRODUCTION

Sphingolipids (SLs) are essential structural components of eukaryotic membranes and more recently they have been found to play important roles in cell signaling, inflammation and apoptosis.[1–4] They are also found in various microbial species such as yeast and fungi, as well as bacteria such as Sphingomonas. The early steps in SL biosynthesis appear to be conserved in all species studied to date; the first reaction in the pathway is catalysed by the pyridoxal 5′-phosphate (PLP)-dependent enzyme serine palmitoyltransferase (SPT). It produces 3-ketodihydrosphingosine (KDS) by a Claisen-like condensation of L-serine and palmitoyl-CoA (see Figure 1). SPT belongs to the α-oxoamine synthase (AOS)[5] family which contains three other well-characterized members: 8-amino-7-oxononanoate synthase (AONS),[6,7] 5-aminolevulinate synthase,[8–10] and 2-amino-3-ketobutyrate-CoA ligase.[11,12] Together they catalyse similar chemical reactions but have specificity for their respective amino acid and fatty acid CoA thioester substrates.

Genes encoding SPT have been identified in numerous organisms but the species-specific traits of each enzyme have only emerged in recent years.[13] In humans, SPT is a heterodimeric, membrane-bound enzyme encoded by two homologous genes making its isolation in a catalytically active form extremely challenging.[14] Recent evidence suggests that the human enzyme is even more complex with the identification of two small subunits that activate in vitro SPT activity ~10

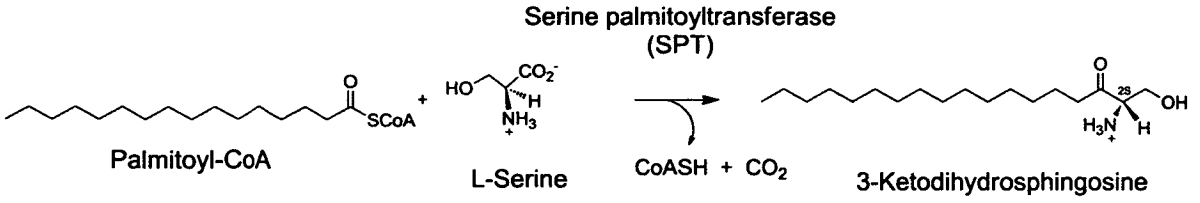


FIGURE 1 Reaction catalysed by SPT.

fold.[15] Insight into the structure and mechanism of SPT has been obtained from studies of a cytoplasmic, water-soluble, homodimer that was isolated from *Sphingomonas paucimobilis*.[16] The interactions between this SPT and various substrates and inhibitors have also been explored.[17–19] Our group solved the first SPT structure, that of the PLP-bound, holo-form of *S. paucimobilis* that brought to light important catalytic residues and substrate specificity.[20] This also allowed a model of the human heterodimeric enzyme to be built to probe residues involved in human neurological disorders such as hereditary autonomic neuropathy type I (HSAN1). More recently, we captured the PLP:L-ser external aldimine complex which provided insight into how this key intermediate is stabilised.[21] Recently, Ikushiro et al.[22] reported the structure of the holo- and external aldimine forms of SPT from the bacterium *Sphingobacterium multivorum* showing some differences with the *S. paucimobilis* SPT.

As well as detailed studies on the individual enzymes involved in SL and lipid synthesis and metabolism, increasing effort has also been placed on expanding our knowledge of the structures and functions of all lipids within an organism, the field of "lipidomics."[23] Reports of the lipid content of mammals, plants and microbes have begun to appear and it is important to investigate interesting model organisms. Here, we describe our work on *Sphingomonas wittichii* RW1, a strictly aerobic Gram-negative bacterium that produces both glucuronosyl and galacturonosyl ceramide types of SLs.[24] This is also an environmentally useful organism as it metabolises toxic pollutants such as dioxin.[25] The genome sequence of *S. wittichii* RW1 was recently completed and this stimulated our interest in using this organism as a good model for bacterial SL biosynthesis. We identified the gene encoding *S. wittichii* SPT using the amino acid sequence of the SPT from *S. paucimobilis*. Curiously, upstream from SPT, we found a small gene product that showed high sequence homology to fatty acid synthase (FAS) type II acyl carrier proteins (ACPs; Supporting Information Figure 1). ACPs are small (~10 kDa), acidic proteins that play essential roles in the biosynthesis of fatty acids and other important natural products including polyketides.[26,27] ACPs are expressed in an apo-form and then post-translationally modified to their 4′-phosphopantetheine (4′-PP), holo-form using CoASH by an ACP synthase (ACPS also known as a phosphopantetheinyl transferase, PPTase).[28] The fatty acid is built up step-by-step by enzymes within the multi-subunit FAS complex whilst attached to the ACP as a thioester.[26] ACPs have also been shown to be key components of polyketide synthases and their peptidyl carrier protein (PCP) homologs also carry out essential roles in nonribosomal peptide synthases.[29] The carrier protein must interact with each of the enzymes within each complex and the exact details of these protein–protein interactions are the subject of current research.[30]

The discovery that the *S. wittichi* ACP and SPT genes are encoded by consecutive open reading frames (ORFs) on the genome suggest a possible direct link between fatty acid and sphingolipid biosynthesis. In this study, we report the structural and functional characterization of *S. wittichii* SPT. We have also isolated and purified the linked apo-ACP and have successfully converted it into the holo-form. Sequence analysis, coupled with chemical and molecular modeling, suggest that it is an unusual apo-ACP. Our work provides the platform for future studies of bacterial sphingolipid biosynthesis.

MATERIALS AND METHODS

Strains, DNA, and Expression Plasmids

The *S. wittichii* RW1 strain was a kind gift from Dr. Rolf Halden, John Hopkins University, USA. The complete genome sequence is available (http://img.jgi.doe.gov). The plasmids encoding the *Streptomyces coelicolor* ACPS (pET15b-ACPS) and *Bacillus subtilis* Sfp (pET28a/Sfp) were generous gifts from Dr. Matt Crump (University of Bristol) and Prof. Jason Micklefield (University of Manchester), respectively. The synthetic DNA construct (encoding *Streptomyces verticillus* SVP) was purchased from Mr. Gene, Regensburg, Germany.

Molecular Biology Reagents, Chemicals, and Chromatography Media

Plasmids (pET) and *Escherichia coli* chemically competent cells BL21 (DE3), HMS174 (DE3) were purchased from Novagen. PCR products were cloned using the pGEM-T Easy cloning kit from Progema. Taq Polymerase Beads were from GE Healthcare. Plasmid and chromosomal DNA isolation kits were from Qiagen and used as per manufacturers' instructions. All buffers and chemical reagents were

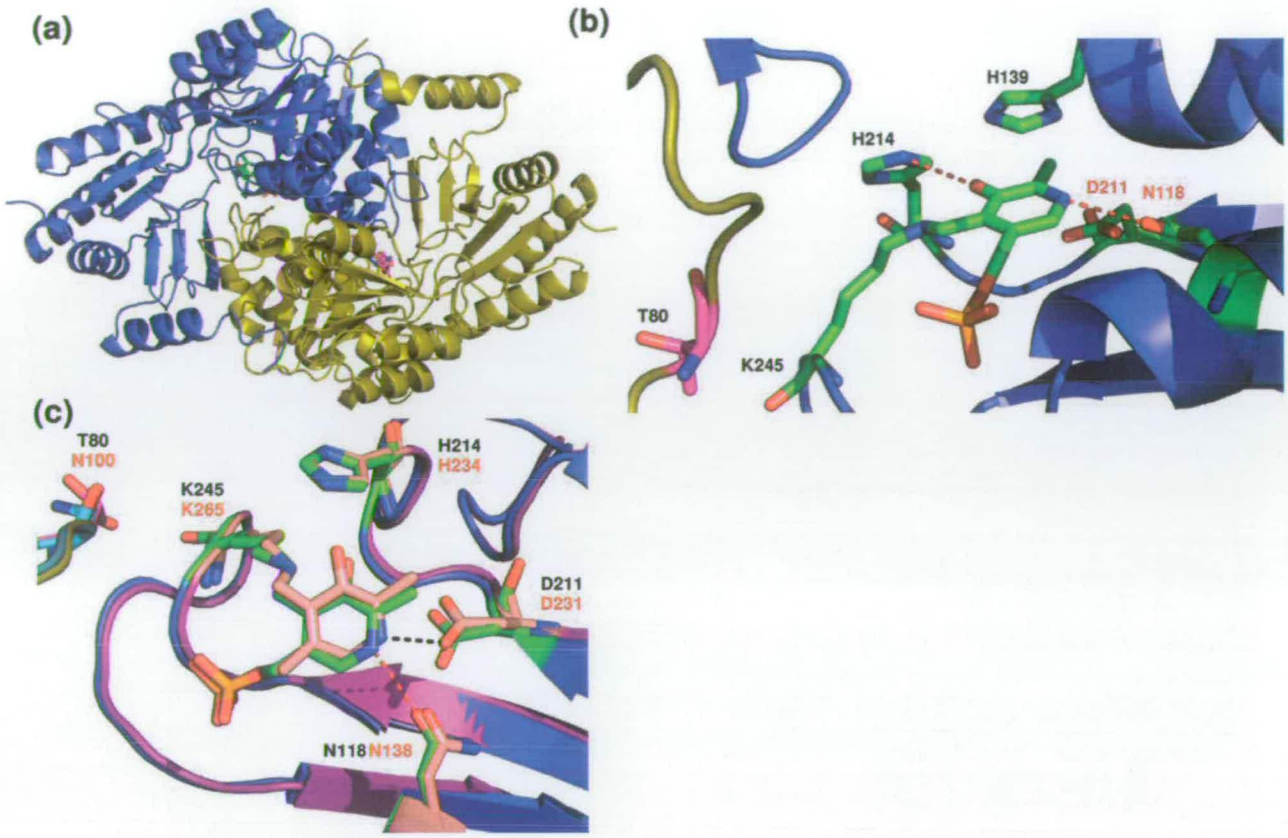


FIGURE 2 Structure of *Sphingomonas wittichii* SPT. (a) Overall structure of the SPT dimer showing one monomer in blue and the other in olive. The cofactor PLP bound to the residue Lys245 in the blue monomer is shown in stick form and is colored green. The equivalent residue from the olive monomer is colored magenta. (b) Catalytic site view showing the cofactor PLP of monomer A and the residues His139, His214, and Asn118 involved in its stabilization. Monomer A is drawn in blue and monomer B is drawn in olive. (c) Overlay of SPT SW and SPT SP showing similar residues involved at the active sites. SPT SW is colored blue and the PLP plus active sites residues are represented in green. SPT SP is colored magenta and the PLP cofactor and the active site residues are pink.

from Sigma. Palmitoyl-CoA was from Avanti Polar Lipids and [U-^{14}C]L-serine (specific activity 1.85 MBq/mL) was from GE Healthcare. The decanoyl-CoA (10:0), lauroyl-CoA (12:0), myristoyl-CoA (14:0), stearoyl-CoA (18:0), and arachidoyl-CoA (20:0) were from Sigma. Oligonucleotide primers were purchased from SigmaGenosys. PEG 3350 was from Fluka. All chromatography columns (Ni-sepharose, gel filtration) were from GE Healthcare.

Gene Cloning

Using the Sphingomonas gene sequence (Supporting Information Figure 2) four PCR products that amplified the genes encoding ACP-SPT together, ACP alone, SPT alone and ACPS were amplified using purified genomic DNA from *S. wittichii* strain RW1 as a template and the following primers:

(ACP-SPT and ACP forward BspHI)
5′-CATGTCATGAGCAGCAGAGAAGACATCTTCACC-3′
(ACP-SPT reverse HindIII)
5′-GGAAAAGCTTGGGGATTACTCCCGTGGCGCGCC-3′
(ACP reverse HindIII)
5′-AAGCTTGGAGAGGAGGTCGGCCACC-3′
(SPT forward NcoI)
5′-GGCTGACCCCATGGCCGACCTCCT-3′
(SPT reverse HindIII)
5′-TGCGGCCGCAAGCTTGGGGATTAC-3′
(ACPS forward NdeI)
5′-AGCGGGGAACCCATATGATCATCGGCATC-3′
(ACPS reverse XhoI)
5′-ATCCCCCTCGAGCAGCGGGCGCGCATACA-3′

These primers introduce appropriate restriction enzyme sites (underlined) for subsequent cloning in the *E. coli* expression plasmids; pET-28a (for ACP-SPT, ACP, and SPT), pET-22b (for ACPS) designed to include a His_6 tag at the C-terminus of the recombinant SPT and ACPS proteins.

The predicted sequences of the *S. wittichii* gene products (or ORFs) are annotated in the whole genome sequence (http://genome.jgi-psf.org/sphwi/sphwi.annotation.html). This suggested that the "Swit_3900" ORF encoded an "AONS" of 400 amino acids with high sequence homology to SPT from *S. paucimobilis.*

Sequence analysis also suggests that this ORF was translated from a GTG start codon, 10 amino acids downstream from a possible alternative Met start codon. Consequently, the SPT expressing clone was designed in order to produce the "isolated" SPT sequence and allow determination of the sequence of the translated protein.

The gene encoding the promiscuous PPTase (also known as ACPS) from *S. verticillus* ATCC15003 "Svp" was generated by a commercial company (Mr. Gene, Germany) using the sequence described in Sanchez et al.[31] The synthetic gene contained the restriction site *Nco*I and *Xho*I to clone the Svp gene in pET-22b and obtain a His_6 Tag at the C-terminus to enable purification by immobilised metal affinity chromatography (IMAC) on nickel resin.

Protein Expression and Purification

Except where specified protein expression was carried out in *E. coli* BL21 (DE3) host cells transformed with the appropriate pET plasmid and grown in the appropriate selective antibiotic (ampicillin at 100 μg/mL or kanamycin at 30 μg/mL). To co-express *S. wittichii* ACP and SPT, cells transformed with the appropriate plasmid (pET-28a/ACP_SPT SW) were cultured in 2YT media [16 g/L Bacto-tryptone, 10 g/L Bacto-yeast extract, 5 g/L sodium chloride (pH 7.5)] including the appropriate antibiotic at 30°C for 3–5 h in the presence of 0.1 m*M* isopropyl β-D-1-thiogalactopyranoside (IPTG) to induce protein expression. To express the *S. verticillus* Svp protein, the expressing plasmid was used to transform HMS 174 (DE3) cells which were grown at 30°C for no longer than 3 h to avoid any misfolding and inactivation of the enzyme. It is important to note that it was not possible to express the *S. wittichii* ACP alone, thus we co-expressed the ACP with SPT from the T7 promoter on the pET plasmid. Induced cells were harvested (Sorvall RC5B centrifuge) by centrifugation at 3500 rpm for 20 min at 4°C.

For the isolation of all recombinant proteins, the cell pellet was resuspended in 20 m*M* potassium phosphate buffer (pH 7.5) containing one EDTA-free protease tablet (Roche) before being disrupted by sonication (Soniprep 150) for 15 cycles (30 s on, 30 s off) on ice. The cell debris was removed by centrifugation at 16,000 rpm for 30 min at 4°C and the supernatant cell-free extract was filtered (0.45 μm syringe filter), before being loaded onto the appropriate chromatography column. All chromatography steps were carried out at 4°C using an ATKA (GE Healthcare) to control flow rates and salt gradients.

Each of the poly-histidine tagged recombinant proteins, *S. wittichii* SPT, *B. subtilis* Sfp, and *S. coelicolor* ACPS were purified according to the previously described methods.[21,32,33] Cell free extracts containing recombinant *S. wittichii* ACP protein in 20 m*M* potassium phosphate buffer (pH 7.5) were loaded onto a HiLoad Q-sepharose anion exchange column (50 mL, GE Healthcare) pre-equilibrated with 20 m*M* potassium phosphate buffer (pH 7.5). The ACP protein was eluted with a linear salt gradient (0–1.0*M* KCl) over 20 column volumes (1 L). Fractions containing ACP were identified by SDS–PAGE analysis, pooled and loaded onto a calibrated HiLoad 16/60 Superdex 75 prep grade (120 mL) size exclusion column. Elution was carried out in 20 m*M* potassium phosphate buffer (pH 7.5), 150 m*M* NaCl.

The his-tagged *S. wittichii* ACPS in 20 m*M* potassium phosphate buffer (pH 7.5), 150 m*M* NaCl, 10% glycerol (v/v), and 20 m*M* imidazole was loaded onto a HiTrap Chelating HP column pre-equilibrated with the same buffer. The protein was eluted with a linear gradient (20–500 m*M* imidazole) over 10 column volumes. Fractions containing ACPS were pooled and loaded onto a previously calibrated HiLoad™ 16/60 Superdex™ 75 prep grade (120 mL) size exclusion column. The pure protein was eluted at a flow rate of 1.2 mL/min in 20 m*M* potassium phosphate buffer (pH 7.5), 150 m*M* NaCl, 10% Glycerol. The *S. verticillus* Svp protein was purified following the same protocol that was used for *S. wittichii* ACPS isolation but the purification buffer did not contain any glycerol.

For storage, the enzyme solutions were brought to a final concentration of 20% glycerol (v/v) (*S. wittichii* SPT, *S. wittichii* ACP, *S. coelicolor* ACPS, and *B. subtilis* Svp) or 40% glycerol (*S. verticillus* Svp and *S. wittichii* ACPS) and stored at −80°C until use.

Protein identity and integrity were confirmed by routine high pressure liquid chromatography electrospray mass spectrometry (HPLC-ESI MS) on a MicroMass Platform II quadrupole mass spectrometer equipped with an electrospray ion source. The experimentally determined masses of each recombinant protein were within 0.1% of the predicted masses based on their amino acid sequences.

Determination of the "Isolated" *S. wittichii* SPT N-Terminus

The ORF annotated as "Swit_3900" in the *S. wittichii* genome sequence is predicted to encode an "AONS". The translated protein is predicted to contain 400 amino acids and a theoretical mass of 43,095 Da. To determine the actual size of the ORF we cloned ORFs "Swit_3899" and "Swit_3900" together in a pET plasmid with a single T7 promoter and a six histidine tag at the C-terminal end of ORF Swit_3900. Expression of both genes was induced with IPTG and the Swit_3900-encoded protein was purified using standard IMAC procedure. After liquid chromatography (LC)-ESI-MS analysis, the mass of the purified recombinant protein was 44,515 Da which is in good agreement with a sequence containing 412 amino acids that has lost the N-terminal methioine residue and includes the C-terminal KLAAALEHHHHHH affinity fusion tag. Therefore, the purified Swit_3900 protein, hereafter named SPT SW, resembles a N-terminally truncated SPT SP (lacking 20 amino acids).

Spectroscopic Measurements and Determination of Dissociation Constants

All UV-visible spectra were recorded on a Cary 50 UV-visible spectrophotometer (Varian) analyzed using Cary WinUV software (Varian). We used the protocol previously described in Raman et al.[21], to characterise the spectroscopic properties of the recombinant *S. wittichii* SPT. Briefly, the SPT was converted to its internal aldimine (holo-) form by dialysis against freshly prepared buffer containing 20 m*M* potassium phosphate buffer (pH 7.5), 150 m*M* NaCl, 25 μ*M* PLP for 1 h at 4°C. We removed excess PLP by passage through a PD-10 column (GE Healthcare) eluting with the same buffer without PLP. The protein was concentrated to 10–20 mg/mL using a VivaSpin 30 kDa cut-off ultrafiltration spin filter. The dissociation constant for L-serine (K_d) was determined spectroscopically by monitoring the changes (at 425 nm) upon incubation of the protein with increasing concentrations of L-ser (0–80 m*M*). Data were fitted to a hyperbolic saturation curve using Sigma Plot software as described in Raman et al.[21]

Table I Kinetic Parameters for the Purified SPT *S. wittichii* and *S. paucimobilis*

Enzyme	$k_{cat} \times 10^3$ (s^{-1})	K_m^{Ser} (mM)	K_m^{PCoA} (μM)	k_{cat}/K_m^{Ser} ($M^{-1}s^{-1}$)	k_{cat}/K_m^{PCoA} ($M^{-1}s^{-1}$)	K_d^{Ser} (mM)
SPT *S. wittichii*	68.7 ± 1.5	0.78 ± 0.10	23.4 ± 4.5	88.1	2,936	0.80 ± 0.1
SPT *S. paucimobillis*	1150.0 ± 30.0	1.40 ± 0.10	35.4 ± 2.0	821.4	32,486	1.1 ± 0.1

Assay of Recombinant SPT Activity

We used methods previously described in Raman et al.,[21] to determine the catalytic activity of *S. wittichii* SPT with a range of substrates. Product formation was monitored by following production of [^{14}C] KDS from [^{14}C] L-serine using thin layer chromatography and autoradiography. SPT activity was also monitored continuously by measuring the release of CoASH by reaction with 5,5′-dithiobis-2-nitrobenzoic acid (DTNB) at 412 nm.

Structural Biology

The SPT protein was screened for suitable crystallization conditions at the Scottish Structural Proteomics Facility. The proteins were dialyzed in the presence of excess PLP to ensure complete reloading prior to crystallization. The screens were built on a Hamilton Microstar liquid-handling robot controlled by Rhombix system software (Thermo).

The *S. wittichii* SPT was crystallized at 20 mg/mL in 10 m*M* Tris (pH 7.5), 150 m*M* NaCl, 25 μ*M* PLP using 30% PEG monomethylether (MME) 2000, 0.1*M* KSCN, (NEXTAL JCSG+ screen condition 81). Crystals used for data collection were soaked for a few minutes in a solution of 27% PEG MME 2000, 0.22*M* KSCN, 18% PEG400 plus 1–2 m*M* PLP.

The crystal was mounted in a cryo-loop (Molecular Dimensions) and frozen by plunging it into liquid nitrogen and carried in a dry cryogenic Dewar to the European Synchrotron Radiation Facility (Grenoble, France) for data collection. Multiple crystals were screened at beam line BM14.The data sets were collected at 100 K to 2.1 Å using beam line 14 (Supporting Information Table I). The data for *S. wittichii* SPT were processed with Mosflm and scaled with Scala from the CCP4 suite of programs.[34,35] The density revealed that the PLP was covalently bound to Lys245 confirming that we had obtained the internal aldimine, holo-form of the protein. The SPT model was refined using Refmac5[34] and manually adjusted, including the addition of water molecules with WinCoot.[36] Data and structure files have been deposited in the Protein Data Bank (PDB ID code 2×BU).

Conversion of *S. wittichii* apo-ACP to holo-ACP and acyl-ACP

The protocol used to modify the *S. wititchii* apo-ACP with the different PPTases was modified from previously described methods.[31–33]

For brevity, the protocol is described using the generic name, "PPTase" that refers to each of the recombinant ACPS enzymes used; *S. coelicolor* ACPS, *B. subtilis* Sfp, *S. verticillus* Svp, and *S. wittichii* ACPS.

A final enzyme concentration of 60 μ*M* apo-ACP in reaction buffer (see below) was incubated with 1 μ*M* PPTase, 1 m*M* $MgCl_2$ and a range of concentrations of CoASH/acyl-CoA (250 μ*M* to 1 m*M*) in a final volume of 250 μL. The reaction was incubated at 37°C overnight and then analyzed by LC-ESI-MS to determine the extent of modification of the apo-ACP.

For each PPTase a different reaction buffer was used. For *S. coelicolor* ACPS and *S. verticillus* Svp, 50 m*M* Tris buffer, pH 7.8 was used. For *B. subtilis* Sfp we carried out the reaction in 75 m*M* MES buffer at pH 6.0 and for *S. wittichii* ACPS this was performed in 20 m*M* potassium phosphate buffer (pH 7.5); 5 m*M* DTT was added to the reaction mixture when using free CoASH as a substrate.

It is important to note that $MgCl_2$ concentrations of >1 mM could not be used for reactions containing acyl-CoAs as this caused precipitation of these substrates.

High-resolution LC-FT-ICR Mass Spectrometry Analysis

For high-resolution mass spectrometry (LC-FT-ICR MS), an Ultimate 3000 HPLC system (Dionex, Sunnyvale, CA), equipped with a monolithic PS-DVB (500 μ*M* × 5 mm) analytical column (Dionex), was used. MS data was acquired on a Bruker 12 T Apex Qe FT-ICR (Bruker Daltonics, Billerica, MA) equipped with an electrospray ionization source. Desolvated ions were transmitted to a 6 cm Infinity cell® penning trap. Trapped ions were excited (frequency chirp 48–500 kHZ at 100 steps of 25 μs) and detected between *m/z* 600 and 2000 for 0.5 s to yield a broadband 512 Kword time-domain data. Fast Fourier Transforms and subsequent analyses were performed using DataAnalysis (Bruker Daltonics) software. Multiple charge states could be observed in this way for each of the major species.

Isotopic Fitting. Isotope distributions of specific charge states were predicted using IsotopePattern software (Bruker Daltonics) from theoretical empirical formulae. These were overlaid upon the recorded experimental data as scatter plots, with the theoretical apex of each isotope peak designated by a circle.

Top-Down Fragmentation. Top-down fragmentation was performed on the 12 T Qh-FT-ICR. Two specific ions, *m/z* 1521 and 1304, were sequentially isolated using the mass resolving quadrupole, and MS/MS was performed using collision induced dissociation (CID). For CID, the collision voltage was typically set between 20 and 35 V. Fragmentation data was the sum of 200 acquisitions and data analyses were performed using DataAnalysis (Bruker Daltonics). The SNAP 2.0 algorithm was used for automated peak picking. The resulting top-down fragment mass lists were combined and searched against the primary sequence of ACP using Prosight-PTM software.[37] Mass error tolerances were set for all searches at 10 ppm.

RESULTS AND DISCUSSION

Purification and Characterization of *S. wittichii* SPT

Using BLAST analysis we searched the annotated genome of the sphingolipid-producing organism *S. wittichii* RW1 for the gene encoding a homologue of *S. paucimobilis* SPT. We identified a hit ORF with 70% sequence identity that was annotated as "Swit_3900". It had been labeled as an AONS catalysing the first step in biotin biosynthesis (condensation of L-alanine and pimeloyl-CoA to generate 8-amino-7-oxononanoate, AON[6]) as there is highly conserved sequence homology between members of the AOS family. We also noted that just upstream of "Swit_3900" another small gene, "Swit_3899" had been annotated with "unknown function." Upon sequence analysis of this translated 80 aa protein we found a conserved phosphopantetheinylation sequence motif (DSLT) characteristic of ACPs (Supporting Information Figures 1 and 2).[29,38] The conserved serine residue is post-translationally modified using CoASH as a substrate to convert the apo-ACP into the holo-ACP form containing a 4′-PP prosthetic arm.[28] Further sequence analysis of the two genes suggested the presence of a typical bacterial promoter upstream of the predicted ACP gene and ribosome binding sites (RBSs) before the putative ACP gene (Supporting Information Figure 2). A consensus RBS could not be detected for the SPT gene.

To explore the structure and function of this pair of proteins we cloned both genes into a pET expression plasmid to determine if both genes could be transcribed from the strong plasmid-encoded T7 promoter. We also fused a six-histidine tag at the C-terminal end of the putative SPT to enable facile purification by IMAC as this tagging strategy worked well with our previous *S. paucimobilis* SPT. We induced transcription from the T7 promoter of the expression plasmid pET-28a/ACP_SPT SW using IPTG and we analyzed the expressed protein products by SDS–PAGE (Supporting Information Figure 1b). Two clearly overexpressed bands were observed at ~10 kDa and ~45 kDa corresponding to the ACP and SPT gene products, respectively. This result confirmed that both genes could be expressed from a single external promoter and suggested that the SPT product was translated from a RBS on the RNA transcript.

A large-scale culture was carried out with these expressing cells, a cell-free extract was prepared in the presence of PLP (25 μM) and then subjected to IMAC to isolate the His-tagged putative SPT SW. Upon elution with imidazole we obtained fractions highly enriched in a single 42 kDa protein which was further purified using size exclusion chromatography. The protein eluted as a dimer and its UV-vis spectrum showed characteristic absorbance maxima at 340 and 425 nm which correspond to the enolimine and ketoenamine forms of the PLP-bound enzyme (Supporting Information Figure 3). By measuring the change in absorbance at the ketoenamine-specific wavelength at 425 nm, the dissociation constant for L-serine (K_d L-Ser) was found to be 0.8 mM which is in the same range as the value (1.1 mM) obtained for the *S. paucimobilis* SPT (Table I).[18,21]

We noted that during isolation of the His-tagged SPT SW by IMAC, the putative ACP protein did not bind to the column, so this unbound fraction was used for the isolation of the ACP by ion-exchange chromatography (see below).

SPT Activity and Kinetics

We measured KDS production catalysed by the SPT SW with [^{14}C] L-serine and palmitoyl-CoA using the method described previously (Supporting Information Figure 3b).[39] We also used a continuous, spectroscopic DTNB-linked assay that detects formation of CoASH to determine kinetic parameters of the enzyme. The SPT SW bound L-serine and palmitoyl-CoA with K_m values of 0.78 mM and 23.4 μM, respectively (Table I). The enzyme turned over with a k_{cat} of 0.069 s^{-1} and a catalytic efficiency (k_{cat}/K_m^{Ser}) of 88.1 M^{-1} s^{-1} for L-serine. A comparison of the kinetics of the SPTs from *S. wittichii* and *S. paucimobilis* revealed that both enzymes display similar binding affinities for both substrates but that the SPT SW had a much lower turn over and efficiency (~17 fold for k_{cat}; and 9 and 11 fold for k_{cat}/K_m for L-ser and palmitoyl-CoA, respectively) compared with the SPT SP.

The acyl-CoA chain length specificity of the SPT SW was also investigated. We found stearoyl-CoA (C18:0) to have the fastest turnover (k_{cat} = 0.083 s^{-1}) of all of the acyl-thioesters tested and a K_m of 12.9 μM (Supporting Information Table II). When comparing the k_{cat}/K_m values, stearoyl-CoA (C18:0) was also the most efficient substrate (6416 M^{-1} s^{-1}). In contrast, the decanoyl-CoA (C10:0) and lauroyl-CoA (C12:0), were the poorest substrates and accurate kinetic constants could not be determined.

These combined results suggested that although the SPTs from these different strains displayed high sequence identity and similar substrate binding specificities there could be subtle differences that would lead to differences in turnover rates. It appeared that the SPT from both Sphingomonas strains bound L-serine with the same affinity (similar K_d values) and similar K_m values for CoASH ester substrates but there were large differences in catalytic rates. The presence of a putative ACP that co-expressed with the SPT suggested that instead of using palmitoyl-CoA as a thioester substrate, the *S. wittichii* SPT could possibly use an acyl-ACP thioester substrate.

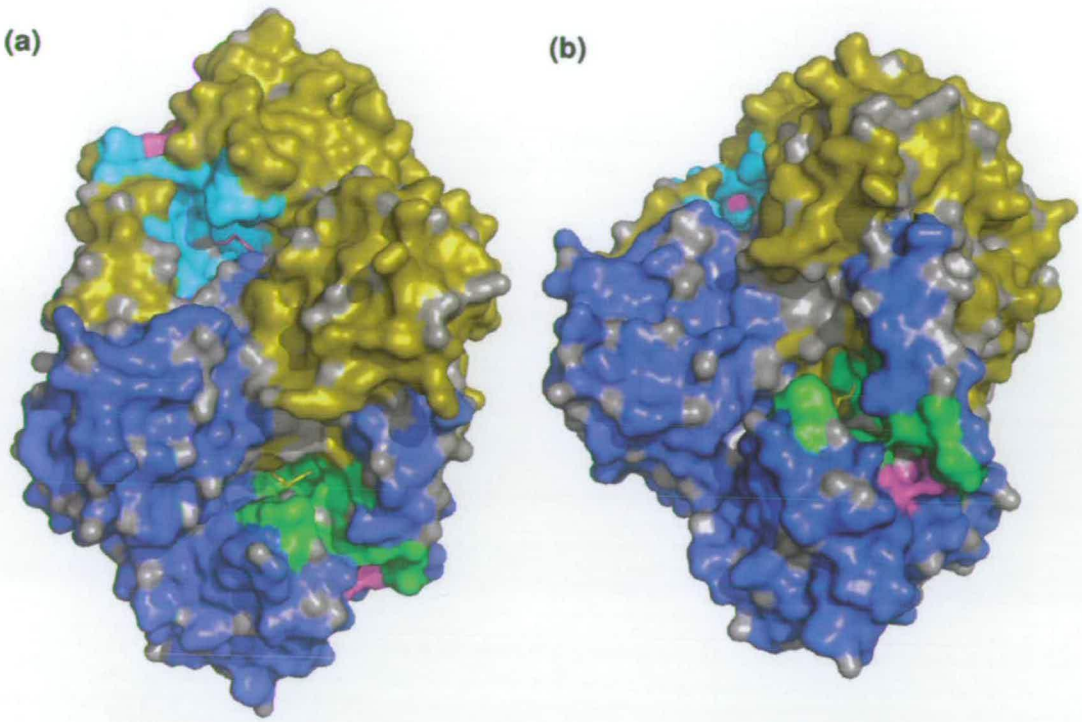


FIGURE 3 Surface structure of SPT showing potential palmitoyl-CoA binding cavity. (a) View from the top. (b) Top/side view. Monomer A is show in blue and monomer B in olive. The two hydrophobic cavities are colored cyan and green. The end of each channel is delimited by two valine residues Val23 and Val341 shown in magenta.

High-resolution Crystal Structure of *S. wittichii* SPT

S. wittichii SPT crystallised as a dimer in the asymmetric unit, in contrast to the *S. paucimobilis* SPT where the dimer is generated by crystal symmetry (Figure 2a).[20–22] However, both the dimer and monomer structures are similar (with a few differences) between the Sphingomonas strains, root mean square deviation for the dimer is 0.65 Å (796 C alpha backbone) and 0.61 for the monomer (398 C alpha backbone). In our previous SPT SP structure, it was not possible to model the N-terminal 22 amino acids and in both SPTs the C-terminal hexa-histidine affinity tags were not visible. The SPT SW is shorter at the N-terminus compared to the SPT SP and therefore the SPT SW model contains 398 of the 412 amino acids from Ala2 to Pro400 inclusive. The PLP cofactor is bound in both monomers in a similar way. All SPT:PLP interactions observed with the cofactor in one active site are also present in the second PLP-binding site within the dimer. The PLP cofactor forms hydrogen bonds to the side chains of Asn118, Thr242, Thr274, the main chains of Gly114, Tyr115, Ala275, and π-stacks with His139 (Figure 2b). Overall the SPT SW and SPT SP active sites bear very close resemblance to each other (Figure 2c).

Two important interactions present in *S. paucimobilis* SPT and also found in other members of the AOS family appear to be absent in the SPT SW structure. Both Asp211 and His214 appear to be too far away to form interactions to the PLP pyridine nitrogen and PLP oxygen atoms, respectively that have been observed in other structures. Both these residues are absolutely conserved among all SPTs and AOS family members and are known to stabilise and anchor the PLP in the active site. Even if these residues are too far away to interact with PLP in this crystal structure, they may be involved during catalysis in solution possibly by stabilization of other intermediates (e.g., formation of the PLP:L-Ser external aldimine). We do note that a possible compensatory interaction involves the side chain of Asn118 which is closer to the pyridine nitrogen of the PLP ring than the equivalent residue (Asn138) in the SPT SP structure. In the *S. paucimobilis* SPT structure we found that an arginine residue, Arg378, played a role in anchoring the carboxylate residue of L-Ser in the PLP:L-Ser external aldimine form of the enzyme. In SPT SP the side chain of Arg378 undergoes a large conformational change from the holo-"swung-out" form to the external aldimine "swung-in" form.[21] We note that in SPT

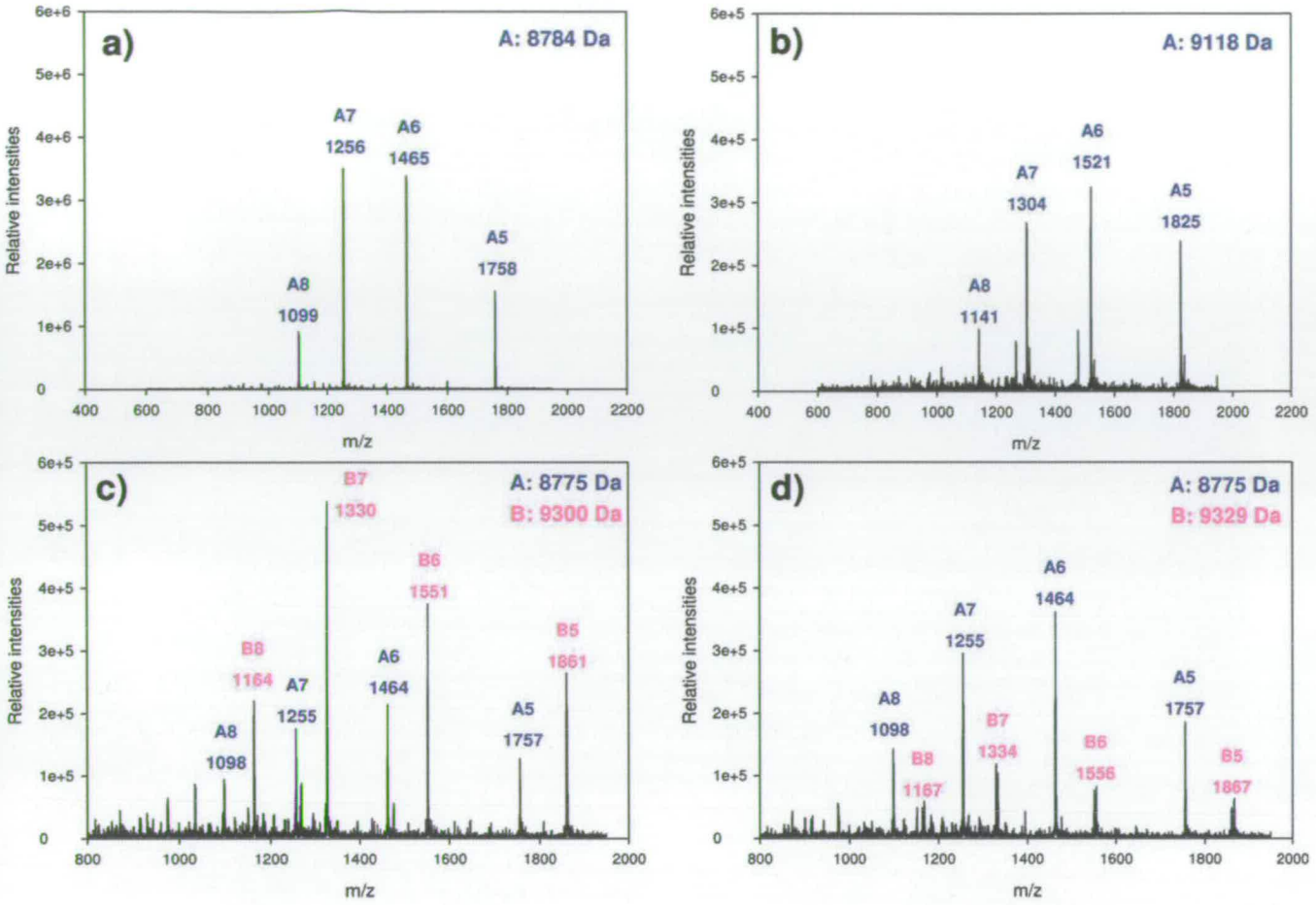


FIGURE 4 Mass spectrometry of ACP. Ion envelopes of (a) apo-ACP, (b) holo-ACP, (c) lauroyl-ACP, and (d) myristoyl-ACP. Apo-forms of ACP are shown in blue. Holo- and acyl-forms are shown in red. The deconvoluted experimental masses based on the ion envelope are shown.

SW residue Arg358 is suggested to play this role and is in the "swung-out" position characteristic of the holo-form of the enzyme.

Other small differences can be observed between the two bacterial SPTs. In our previous study of *S. paucimobilis* SPT we used the bacterial, homodimeric enzyme to generate a model of the heterodimeric (SPT1/2) human SPT.[20] We used this model to predict the site of mutation in the human SPT1 subunit (Cys133, which is mutated to Tyr and Trp) that causes the sphingolipid metabolic disease HSAN1. Our model suggested *S. paucimobilis* SPT residue Asn100 mapped to human SPT1 residue Cys133 and this side chain plays a role in contacting the backbone amide of the PLP-binding residue Lys265. In a recent study we probed its structural and functional role by making Asn100Tyr and Asn100Trp mutant forms of SPT SP.[21] This revealed that mutations on one SPT monomer do indeed impact on the other. In *S. wittichii* SPT the Asn100 residue is replaced by Thr80, other sequence changes place it in a hydrogen bonding environment at the interface between the two subunits. In contrast to Asn100 of SPT SP, the side chain of Thr80 does not make a direct hydrogen bond contact with the other subunit but is immediately adjacent to PLP binding Lys245 of the neighboring subunit (Figure 2c).

We also observed other small differences between the SPT structures from both Sphingomonas strains. The entrance to the active site that is situated at the "top" of the SPT SW protein surface is wider than that observed in *S. paucimobilis* SPT, possibly indicating the binding of a larger long-chain fatty acid substrate (e.g., acyl-ACP) instead of an acyl-CoA thioester (Supporting Information Figure 4). This results from a rigid movement of the loop Pro360-Pro363 of about 2 Å. The residues within this loop are conserved between SPT SP and SPT SW, however Met366 is found in SPT SW and is replaced with Thr386 in SPT SP; whilst residue Thr356 in SPT SW is replaced by Met376 in SPT SP. As a result the methionine side chain sits in different hydrophobic pockets in both enzymes and this may be responsible for the

different loop positions. These comparisons suggest that there are subtle structural differences between the bacterial SPT isoforms that could explain observed differences in their catalytic activities.

To further probe the Nature of the acyl-CoA binding site we used the structure of the recently solved enzyme CqsA from *Vibrio cholerae*, a recent addition to PLP-dependent, AOS family, which resembles SPT and AONS.[40] CqsA is involved in the biosynthesis of cholerae autoinducer-1 (CAI-1) from (S)-2-aminobutyrate and decanoyl-CoA.[41] Of particular interest, the external aldimine of the final product from the condensation of L-threonine and decanoyl-CoA has been trapped in the active site (PDB ID code: 2WKA). Superposition of the SPT SW and *V. cholerae* CqsA structures places the acyl-CoA binding site in a long hydrophobic channel at the interface of the two monomers of SPT SW (Figures 3a and 3b). The channel is lined by residues Leu10, Arg15, Ala17-Leu19, Thr21-Tyr27, Val341, Arg358-Pro361 from monomer A and Val84-Leu85, Leu266, and Val272 from Monomer B. Two valine residues (Val23 and Val341 from monomer A) delimit the end of the channel. We note that SPT SW has a lower chain length limit with regard to its acyl-CoA substrates and we detected no activity with substrates of C12 and below. The exact details of chain length specificity within the AOS family are still unknown but this analysis will allow us to probe the role of the residues lining this putative channel.

Purification, Characterization, and Modification of Apo-ACP

S. wittichii ACP was co-expressed with SPT in *E. coli* and a rapid two-step purification procedure using anion exchange and size-exclusion chromatography led to the isolation of milligram quantities of the protein. We used LC-ESI-MS to determine the mass of the purified protein (8784 Da) and this revealed that the ACP SW had been expressed in the apo-form (predicted mass without Met residue 8782.5 Da, Figure 4a). This suggests that the enzymes involved in post-translational modification of the type II FAS machinery in *E. coli* do not recognise the *S. wittichii* ACP and cannot convert it to the 4′-PP holo-form.

Generally, ACPs are modified by a PPTase (also known as ACPS) by the addition of the 4′-PP group and thereafter the fatty acid chain is built up step-by-step by a FAS.[26] Individual members of the PPTase superfamily tend to display high substrate specificity and will only recognise their cognate ACP.[28] This has caused technical problems with regard to production of various natural products in recombinant hosts (and in vitro) where the ACP is not modified correctly. Some PPTases are able to catalyse the direct transfer of an acyl group to an apo-ACP using an acyl-CoA substrate, thus generating an acylated-ACP in a single step. Some of these PPTases are well known for their broad substrate recognition and inherent "promiscuity" and have been used to solve problems of production of natural products in alternative hosts.[27,31–33] As we isolated the ACP SW in an apo-form we wished to convert it to the holo- and various acylated forms. We used CoASH to install the 4′-PP arm, as well as acyl-CoAs of varying fatty acid chain length (decanoyl C10; lauroyl C12; myristoyl C14; palmitoyl C16 and steroyl C18) in combination with a number of PPTases.

The first attempt to modify ACP SW (with CoASH and acyl-CoAs) was made with the *S. coelicolor* ACPS since it is known to have broad specificity.[33,42] Unfortunately, this did not give any positive results (data not shown) and this suggested that *S. coelicolor* ACPS does not recognise ACP SW. We next tried the broad-range ACPS *B. subtilis* Sfp and found it did successfully convert the apo-ACP into the holo-form with an increase mass of 4′-PP (experimental mass 9118 Da, Figure 4b) as well as the C10 (data not shown), C12 (mass 9300 Da, Figure 4c) and C14 (mass 9329, Figure 4d) acyl-ACP derivatives. The % conversions of apo- to the C12- and C14-modified ACPs (based on MS peak height intensities) were ~50% and ~20%, respectively. We also used high-resolution MS coupled with fragmentation techniques to identify the site of 4′-PP attachment on the holo-ACP SW as the Ser39 residue within the conserved motif (Supporting Information Figure 5). Unfortunately, *B. subtilis* Sfp failed to transfer a palmitoyl (C16:0) or a stearoyl (C18:0) chain onto the apo-ACP SW. The transfer of the palmitoyl chain by *B. subtilis* Sfp was further explored by testing several reaction conditions (varying enzyme, substrate, and co-factor concentrations as well as temperature and time) without any success (data not shown). Previous studies reported that *B. subtilis* Sfp was successfully used to modify a large number of ACPs and PCPs with various substrates but never with long-chain fatty acids (≥C16).[32,43,44]

Another useful promiscuous PPTase, *S. verticillus* Svp, with similar characteristics to Sfp, has been shown to have a preference for longer carbon chains.[45] It was also shown to modify both type I and type II ACPs and PCPs from either *S. verticillus* or other Streptomyces species.[31] However, although we found that *S. verticillus* Svp converted apo-ACP SW to the holo-form it failed to transfer any acyl-thioester substrates (data not shown). Since the *S. verticillus* Svp and other PPTases are able to catalyse formation of holo-ACP SW but are not able to transfer an acyl-4′-PP group, we suggest that the problem with the long-chain transfer must come from the structure of the ACP SW itself. Finally, we used sequence homology to identify the PPTase (ACPS) in the *S. wittichii*

(a)

```
                1                                                   50
   ACP EC     (1) MSTIEERVKKIIGEQLGVKQE--EVTNNASFVEDLGADSLDTVELVMALE
   ACP SW     (1) MSSREDIFTRVAEQIEPFNKKGIDLAETTSFAGDLEWDSLTVMDFVAAVE
Consensus     (1) MSS ED   KI            DL    SF  DL  DSL  MD V ALE

                51                              80
   ACP EC    (49) EEFDTEIPDEEAEKITTVQAAIDYINGHQA
   ACP SW    (51) DEFDIIITMNMQAEIETVGQLVDAVAKLRG
Consensus    (51) DEFD  I       I TV   ID I    A
```

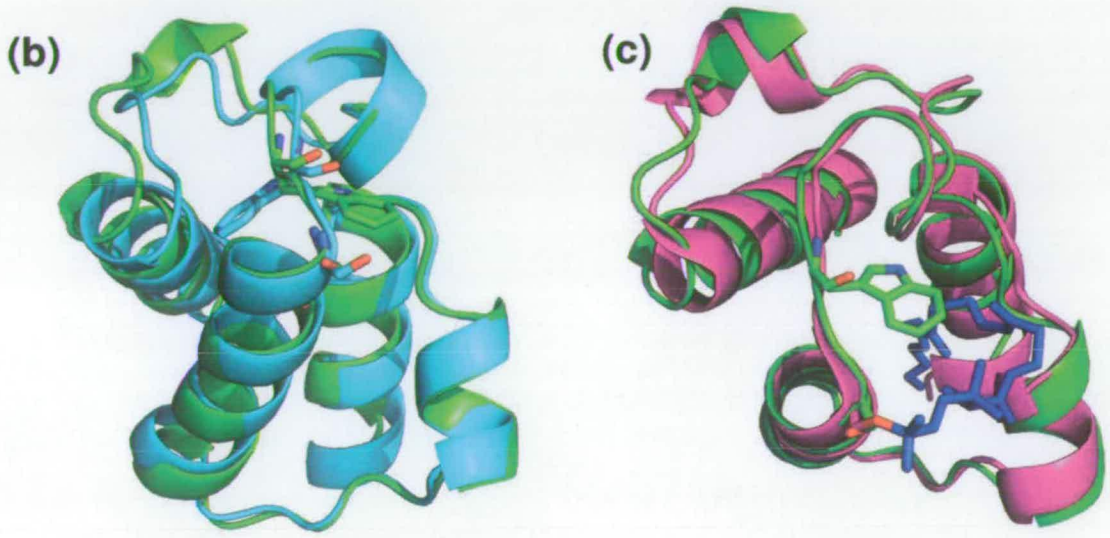


FIGURE 5 (a) Sequence alignments of *E. coli* and *S. wittichii* ACPs. (b) 3D models of ACP *S. wittichii* showing superposition of Swiss and JIGSAW ACP 3D models. The ACP Swiss-model is colored cyan and JIGSAW-model is colored green. In each model, residue Trp37 is shown in stick form. (c) Superposition of the JIGSAW *S. wittichii* ACP model and *E. coli* ACP structure (PDB ID: 1LOH). The ACP SW JIGSAW-model is colored green and *E. coli* structure in magenta with the C4 acyl group shown in stick form and colored in dark blue.

genome (ORF code Swit_3872) that enabled us to overexpress this PPTase in *E. coli*. We were able to purify the recombinant *S. wittichii* ACPS but found that in assays with apo-ACP SW this enzyme catalysed only the formation of the holo-form and provided no acyl-ACP. Similarly, co-expression of ACP SW and ACPS SW inside the same *E. coli* host provided only holo-ACP (data not shown).

In conclusion, we found that *B. subtilis* Sfp was the best PPTase to act upon the apo-ACP SW. Unfortunately, we were unable to generate sufficient quantities of long-chain acylated-ACP to carry out a study of its ability to act as an acyl-donor in the SPT reaction with L-serine. In a recent study on rat type I ACP the authors also noted the tendency of the C16-acylated form of this protein to precipitate over 48 h.[42] We also realised through our acyl-CoA chain length kinetic analysis that SPT SW failed to accept substrates with carbon chain lengths of <C14 (Supporting Information Table II). To generate sufficient quantities of long-chain acyl-ACP SW we may need to turn to a chemical, nonenzymatic procedure using N-acylimidazole derivatives and this is a target of future work.[46]

ACP Structural Model

Since we found the apo-ACP SW was so difficult to convert to an acylated form using promiscuous PPTases, we carried out sequence analysis to identify any novel characteristics that could explain its recalcitrant nature. Protein sequence alignment revealed that ACP SW is rather unusual and does not fit perfectly into an ACP sub-family described in the literature. Its closest homologs are proteins with no published experimental data associated with them. However, it does show 26% amino acid sequence identity to *E. coli* ACP which has been characterised extensively (Figure 5a).[47] In ACP SW one particular residue stands out—a highly uncommon tryptophan residue (Trp37) positioned just before the conserved

(WDSLT) 4′-PP sequence motif (Figure 5a). Numerous ACP structures are available in the literature in both apo-, holo-, and acylated forms.[29,38] Despite acting in diverse biosynthetic pathways, interacting with many different enzymes and having low overall sequence similarity, their secondary structures and overall folds are very similar.[27] For example the *E. coli* ACP is typical and displays a common fold of four α-helices. Thus, using the 3D-JIGSAW server[48–50] and Swiss-model[51–53] we were able to generate two 3D models of ACP SW to gain some insight into the role of this unusual tryptophan residue (Figures 5b and 5c). The ACP SW model was based on a best fit to that of the ACP from *E. coli*.[47] This ACP was shown to be able to accommodate a growing acyl chain within a hydrophobic cavity that could expand as fatty acid synthesis proceeded. The ACP SW models adopted the typical four α-helical bundle fold. Superposition of the two models generated by the different software revealed two possible positions for the Trp37 side chain both of which are close to the site of 4′-PP attachment (Ser39) (Figure 5b). In both models, the residue is positioned on a solvent-exposed flexible loop and so one could imagine that the tryptophan could adopt both conformations. The superposition of the 3D-JIGSAW model with the C4-ACP *E. coli* structure (Figure 5c) lends weight to the attractive idea that this unusually placed tryptophan could control transfer of 4′-PP and/or long-chain acyl groups. Indeed, we found that the holo-ACP SW could be generated from the apo-form by incubation with various PPTases but these structural models also provide a possible explanation for the lack of modification of the ACP SW with longer chain acyl-CoAs (>C16) in appreciable yields. It may be that the ACP SW is converted to a long-chain acyl-CoA form only within *S. wittichii* as part of fatty acid synthesis. Thus, this directly links fatty acid and sphingolipid biosynthesis. Further work will be required to find conditions appropriate for its formation and the role of the Trp37 in this conversion will be probed by site directed mutagenesis.

CONCLUSIONS

To begin a study of bacterial sphingolipid biosynthesis, we have isolated and purified recombinant proteins from the bacterium *S. wittichii*. Using various spectroscopic, kinetic, and structural techniques we characterised SPT SW, and showed that it binds its co-factor PLP and its substrate L-serine. We proved that SPT SW was enzymatically active with L-serine and a range of long-chain acyl-CoA substrates. Furthermore, we solved the high-resolution x-ray structure of the enzyme which allowed a structural comparison with the SPT from *S. paucimobilis*. We noted that despite high amino acid sequence identity the SPT isozymes displayed differences in kinetics and in structure. We found that the SPT gene resides alongside a putative small protein with sequence homology to bacterial type II ACPs and this led us to hypothesise that an ACP-thioester, rather than a CoA-thioester may be the natural substrate for SPT. We co-expressed the *S. wittichii* ACP-SPT gene pair in *E. coli*, isolated the recombinant ACP and used mass spectrometry to show it was recovered in an apo-form. We investigated various PPTases in an effort to convert the ACP into its holo-form and acylated forms but succeeded in isolating only the holo-, C10, C12, and C14 acyl-forms. Analysis of the ACP primary sequence identified an unusual tryptophan residue next to the conserved phosphopantetheinylation motif. Using *E. coli* ACP as a structural homolog, models of ACP SW were generated and suggest that this residue may block the attachment of longer acyl chains. Probing the role of this residue will be the focus of future structural and functional studies as will be the identification and characterization of other downstream enzymes in the Sphingomonas sphingolioid biosynthetic pathway.

We thank Dr. Jonathan Lowther (University of Edinburgh) and Dr. Alan Brown (University of Exeter) for helpful discussions.

REFERENCES

1. Merrill, A. H. J Biol Chem 2002, 277, 25843–25846.
2. Futerman, A. H.; Hannun, Y. A. EMBO Rep 2004, 5, 777–782.
3. Kobayashi, T.; Takahashi, M.; Nagatsuka, Y.; Hirabayashi, Y. Biol Pharm Bull 2006, 29, 1526–1531.
4. Hirabayashi, Y.; Igarashi, Y.; Merrill, A. H. J. Sphingolipid Biology; Springer-Verlag: Tokyo, 2006.
5. Eliot, A. C.; Kirsch, J. F. Annu Rev Biochem 2004, 73, 383–415.
6. Alexeev, D.; Alexeeva, M.; Baxter, R. L.; Campopiano, D. J.; Webster, S. P.; Sawyer, L. J Mol Biol 1998, 284, 401–419.
7. Webster, S. P.; Alexeev, D.; Campopiano, D. J.; Watt, R. M.; Alexeeva, M.; Sawyer, L.; Baxter, R. L. Biochemistry 2000, 39, 516–528.
8. Jordan, P. M.; Shemin, D. In The Enzymes; Boyer, P. D., Ed.; Academic Press: London, New York, 1972; p 339–356.
9. Ferreira, G. C.; Gong, J. J Bioenerg Biomembr 1995, 27, 151–159.
10. Astner, A.; Schulze, J. O.; van den Heuvel, J.; Jahn, D.; Schubert, W. D.; Heinz, D. W. Embo J 2005, 24, 3166–3177.
11. Schmidt, A.; Sivaraman, J.; Li, Y.; Larocque, R.; Barbosa, J. A. R. G.; Smith, C.; Matte, A.; Schrag, J. D.; Cygler, M. Biochemistry 2001, 40, 5151–5160.
12. Bashir, Q.; Rashid, N.; Akhtar, M. Chemical Commun 2006, 5065–5067.
13. Hanada, K. Biochim Biophys Acta 2003, 1632, 16–30.
14. Hanada, K.; Hara, T.; Fukasawa, M.; Yamaji, A.; Umeda, M.; Nishijima, M. J Biol Chem 1998, 273, 33787–33794.
15. Han, G.; Gupta, S. D.; Gable, K.; Niranjanakumari, S.; Moitra, P.; Eichler, F.; Brown, R. H., Jr.; Harmon, J. M.; Dunn, T. M. Proc Natl Acad Sci USA 2009, 106, 8186–8191.

16. Ikushiro, H.; Hayashi, H.; Kagamiyama, H. J Biol Chem 2001, 276, 18249–18256.
17. Ikushiro, H.; Hayashi, H.; Kagamiyama, H. Biochim Biophys Acta 2003, 1647, 116–120.
18. Ikushiro, H.; Hayashi, H.; Kagamiyama, H. Biochemistry 2004, 43, 1082–1092.
19. Ikushiro, H.; Islam, M. M.; Tojo, H.; Hayashi, H. J Bacteriol 2007, 189, 5749–5761.
20. Yard, B. A.; Carter, L. G.; Johnson, K. A.; Overton, I. M.; Dorward, M.; Liu, H.; McMahon, S. A.; Oke, M.; Puech, D.; Barton, G. J.; Naismith, J. H.; Campopiano, D. J. J Mol Biol 2007, 370, 870–886.
21. Raman, M. C.; Johnson, K. A.; Yard, B. A.; Lowther, J.; Carter, L. G.; Naismith, J. H.; Campopiano, D. J. J Biol Chem 2009, 284, 17328–17339.
22. Ikushiro, H.; Islam, M. M.; Okamoto, A.; Hoseki, J.; Murakawa, T.; Fujii, S.; Miyahara, I.; Hayashi, H. J Biochem 2009, 146, 549–562.
23. Shaner, R.; Allegood, J.; Park, H.; Wang, E.; Kelly, S.; Haynes, C.; Sullards, M.; Merrill, A. J. J Lipid Res 2009, 50, 1692–1707.
24. Wittich, R. M.; Wilkes, H.; Sinnwell, V.; Francke, W.; Fortnagel, P. Appl Environ Microbiol 1992, 58, 1005–1010.
25. Habe, H.; Ashikawa, Y.; Saiki, Y.; Yoshida, T.; Nojiri, H.; Omori, T. FEMS Microbiol Lett 2002, 211, 43–49.
26. White, S. W.; Zheng, J.; Zhang, Y. M.; Rock, C. O. Annu Rev Biochem 2005, 74, 791–831.
27. Lai, J. R.; Koglin, A.; Walsh, C. T. Biochemistry 2006, 45, 14869–14879.
28. Lambalot, R. H.; Gehring, A. M.; Flugel, R. S.; Zuber, P.; LaCelle, M.; Marahiel, M. A.; Reid, R.; Khosla, C.; Walsh, C. T. Chem Biol 1996, 3, 923–936.
29. Mercer, A. C.; Burkart, M. D. Nat Prod Rep 2007, 24, 750–773.
30. Weissman, K. J.; Muller, R. Chembiochem 2008, 9, 826–848.
31. Sanchez, C.; Du, L.; Edwards, D. J.; Toney, M. D.; Shen, B. Chem Biol 2001, 8, 725–738.
32. Quadri, L. E.; Weinreb, P. H.; Lei, M.; Nakano, M. M.; Zuber, P.; Walsh, C. T. Biochemistry 1998, 37, 1585–1595.
33. Cox, R. J.; Crosby, J.; Daltrop, O.; Glod, F.; Jarzabek, M. E.; Nicholson, T. P.; Reed, M.; Simpson, T. J.; Smith, L. H.; Soulas, F.; Szafranska, A. E.; Westcott, J. J Chem Soc, Perkin Trans 2002, 1, 1644–1649.
34. CCP4. Acta Crystallogr D 1994, 50, 760–763.
35. Kabsch, W. J Appl Cryst 1993, 26, 795–800.
36. Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. Acta Crystallogr D Biol Crystallogr 1997, 53, 240–255.
37. LeDuc, R. D.; Taylor, G. K.; Kim, Y. B.; Januszyk, T. E.; Bynum, L. H.; Sola, J. V.; Garavelli, J. S.; Kelleher, N. L. Nucleic Acids Res 2004, 32, W340–345.
38. Byers, D. M.; Gong, H. Biochem Cell Biol 2007, 85, 649–662.
39. Williams, R. D.; Wang, E.; Merrill, A. H., Jr. Arch Biochem Biophys 1984, 228, 282–291.
40. Jahan, N.; Potter, J. A.; Sheikh, M. A.; Botting, C. H.; Shirran, S. L.; Westwood, N. J.; Taylor, G. L. J Mol Biol 2009, 392, 763–773.
41. Kelly, R. C.; Bolitho, M. E.; Higgins, D. A.; Lu, W.; Ng, W. L.; Jeffrey, P. D.; Rabinowitz, J. D.; Semmelhack, M. F.; Hughson, F. M.; Bassler, B. L. Nat Chem Biol 2009, 5, 891–895.
42. Ploskon, E.; Arthur, C. J.; Evans, S. E.; Williams, C.; Crosby, J.; Simpson, T. J.; Crump, M. P. J Biol Chem 2008, 283, 518–528.
43. Garneau-Tsodikova, S.; Dorrestein, P. C.; Kelleher, N. L.; Walsh, C. T. J Am Chem Soc 2006, 128, 12600–12601.
44. Yin, J.; Straight, P. D.; McLoughlin, S. M.; Zhou, Z.; Lin, A. J.; Golan, D. E.; Kelleher, N. L.; Kolter, R.; Walsh, C. T. Proc Natl Acad Sci USA 2005, 102, 15815–15820.
45. Gerber, R.; Lou, L.; Du, L. J Am Chem Soc 2009, 131, 3148–3149.
46. Cronan, J. E.; Jr.; Klages, A. L. Proc Natl Acad Sci USA 1981, 78, 5440–5444.
47. Roujeinikova, A.; Simon, W. J.; Gilroy, J.; Rice, D. W.; Rafferty, J. B.; Slabas, A. R. J Mol Biol 2007, 365, 135–145.
48. Bates, P. A.; Sternberg, M. J. Proteins 1999,Suppl. 3, 47–54.
49. Bates, P. A.; Kelley, L. A.; MacCallum, R. M.; Sternberg, M. J. Proteins 2001,Suppl. 5, 39–46.
50. Contreras-Moreira, B.; Bates, P. A. Bioinformatics 2002, 18, 1141–1142.
51. Arnold, K.; Bordoli, L.; Kopp, J.; Schwede, T. Bioinformatics 2006, 22, 195–201.
52. Schwede, T.; Kopp, J.; Guex, N.; Peitsch, M. C. Nucleic Acids Res 2003, 31, 3381–3385.
53. Guex, N.; Peitsch, M. C. Electrophoresis 1997, 18, 2714–2723.

Reviewing Editor: Sylvie Garneau-Tsodikova