THE UNIVERSITY
of EDINBURGH

Structural and mechanistic studies of the pyridoxal 5'-phosphate-dependent enzyme serine palmitoyltransferase

Bohdan Mykhaylyk

A Thesis Submitted for the Degree of

Doctor of Philosophy

The University of Edinburgh, 2018

Abstract

Sphingolipids (SLs) are complex lipid-derived structures that are essential components of cell membranes in eukaryotes and some bacteria. SLs and their complex derivatives ceramides are known to be involved in multiple processes such as the formation of lipid rafts, cell signalling and membrane trafficking.

The first step of SL biosynthesis is universal to all sphingolipid-producing organisms from bacteria to humans and is catalysed by the enzyme serine palmitoyltransferase (SPT). SPT is a member of the alpha-oxoamine synthase (AOS) family of pyridoxal-5'-phosphate-dependent enzymes. All AOS family enzymes retain a high degree of structural homology and catalyse the decarboxylative Claisen-like condensation of amino acids with thioester substrates. The SPT enzyme catalyses the formation of the universal SL precursor, 3-ketodihydrosphingosine (KDS), by condensation of L-serine and coenzyme A-derived palmitic acid. Being the key controller in SL biosynthesis, SPT plays a big role in regulating natural and pathological processes.

A lot of research interest has been recently generated by SLs isolated from bacterial members of the human microbiome and their roles in human health. Increasing evidence suggests that some of these SLs possess immunoregulatory effects and can have a direct impact on the immunity of the host. *Bacteroides fragilis* is a commensal gut-dwelling bacterium that belongs to a few human microbionts known to produce unique iso-branched sphingolipids (isoSLs); these have been shown to influence the human iNKT cell count. The production of SLs in *B.fragilis* is completely regulated by a gene product BF2461.

In this work, BF2461 was expressed and purified; using a combination of UV-vis spectrometry, enzymatic assays, mass spectrometry and protein X-ray crystallography, it has been confirmed to be an SPT. The substrate specificity of the *Bf*SPT has been assessed with a range of different chain-length substrates, including less common 15 and 17-carbon chain length coenzyme A substrates. The enzyme can produce different types of SL precursors with a preference for the 16-carbon chain substrate palmitoyl-CoA. However, at high levels of PCoA, a substrate inhibition is observed that might

point to a natural control mechanism employed by the bacterium in favour of producing iso-branched SLs (isoSLs).

The structure of *Bf*SPT has been elucidated in a complex with its amino acid substrate L-serine. Search and analysis of putative SPTs from other microbiome-associated bacteria that produce isoSLs show that they share high similarity with an average amino acid conservation of 74%, suggesting they might be adapted to a particular type of substrate. In this respect, *Bf*SPT might be the first isoSL-producing SPT to be structurally characterised, and the first one to have a direct impact on human health.

Further structural data were obtained on protein complexes with L-cycloserine and L-penicillamine, some common inhibitors of the PLP-dependent enzymes. The structure obtained in the presence of L-penicillamine provides the first direct structural evidence of the inhibitory mechanism by a thiazolidine complex formation in the active site of a PLP-dependent enzyme. These findings shed light on certain aspects of the reaction and inhibition mechanisms of *Bf*SPT as well as opening new prospects into researching this interesting target and its impact on the human microbiome.

Acknowledgements

First of all, I would like to express my biggest gratitude to my supervisor Prof. Dominic Campopiano who gave me the opportunity to join his research group and this project. His endless support, both academic and personal, great enthusiasm and passion have been guiding me through this long journey and always helping me to find the right direction.

I am endlessly grateful to the University of Edinburgh and Derek and Stewart Charitable Trust for funding my research.

I am thankful to Prof James Naismith for the opportunity I was given to work in his laboratory at the University of St Andrews. I thank Dr Lucile Moynie for her input in guiding me through the first steps of learning structural biology. A truly exceptional gratitude goes to Dr Magnus Alphey. His guidance and help in my structural studies were absolutely invaluable as he gets the credit for almost everything I have learned in the art of protein crystallography. Without your kind help and cooperation, most of this work would not have been accomplished, and I genuinely don't know how to express my appreciation fully.

I thank all my colleagues and fellow PhD students, members of lab 229 past and present, for all the help and good memories.

I thank all my friends, those who have been around with me in Edinburgh and those who have been far away. Great thanks to all you guys for keeping me going!

I finally thank all members of my family for their love, support and endless belief in me.

Declaration

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

Bohdan Mykhaylyk

The University of Edinburgh, 2018

Contents:

Abbreviations:

Å – Angström; 1 Å = 0.1 nm

AA – amino acid

Abs - absorbance

Ala (A) – alanine

ALAS - 5-aminolevulinate synthase

AOS - amino-oxononanoate

AONS - 8-amino-7-oxononanoate synthase

Arg (R) – arginine

Asn (N) - asparagine

Asp (D) – aspartic acids

A.U. – arbitrary unit

B-factor – isotropic measure of mobility for each atom within the electronic density expressed in $Å^2$

B.fragilis (*Bf*) – *Bacteroides fragilis*

BCA - β-chloroalanine

C.elegans - *Caenorhabditis elegans*

CoA/CoASH – coenzyme A

CHCA - α-cyano-4-hydroxycinnamic acid

CqsA - *Cholera* quorum sensing autoinducer

Cys (C) - Cysteine

Da – Dalton, 1Da = 1.66×10^{-27} kg

DCS – D-cycloserine

DHS – dihydrosphingosine

DNA – deoxyribonucleic acid

dNTP – deoxynucleotide phosphate

DTNB - 5,5'-dithio-bis-(2-nitrobenzoic acid)

D-pen – D-penicillamine

E.coli (*Ec*) – *Escherichia coli*

ESI MS – Electrospray Ionisation Mass Spectrometry

ETBF - enterotoxigenic strains of *B.fragilis*

F_o- F_c – Observed structure factor amplitudes – calculated structure factor amplitudes

αGC - α-galactosylceramide

GF – gel filtration (size exclusion chromatography)

Gly (G) – glycine

HEPES - 4-(2-hydroxyethyl)piperazine-1-ethanesulfonate

His (H) – histidine

HSAN I - Hereditary sensory autonomous neuropathy type I

iNKT - invariant natural killer T cells

IPTG - isopropyl β-D-1-thiogalactopyranoside

isoC15 – isopentadecanoyl

isoSL(s) – iso-branched sphingolipid(s)

k_{cat} - catalytic constant

K_d – dissociation (affinity) constant

K_i – inhibition constant

K_M - Michaelis-Menten constant

KBL - 2-Amino-3-ketobutyrate-CoA ligase

KDS – 3-ketodihydrosphingosine

KPhos - potassium phosphate buffer

LB – lysogeny broth medium

LCB – long chain bases

LCS – L-cycloserine

L-pen – L-penicillamine

LMW – low molecular weight

Lys (L) - lysine

MALDI - Matrix-assisted laser desorption/ionization mass spectrometry

MW – molecular weight

Ni-NTA - nickel-nitrilotriacetic acid

OD - optical density

PCoA – palmitoyl coenzyme A

PBS – phosphate buffered saline

Pdb – Protein Data Bank

PEG – polyethylene glycol

PCR - polymerase Chain Reaction

PLP – pyridoxal -5'-Phosphate

PMP – pyridoxal monophosphate

R.capsulatus (*Rc*) - *Rhodobacter capsulatus*

R factor – measure of relative discrepancy between experimentally measured factor amplitudes and calculated factor amplitudes from the model in crystallography defined as $\Sigma |F_{obs} - F_{calc}| / \Sigma F_{obs}$

RMSD – root mean square deviation, averaged measure of deviation in distances between atoms

Rpm – rotations per minute

SAM – S-adenosine methionine

S.cerevisiae – *Saccharomyces cerevisiae*

S.multivorum (Sm) – *Sphingobacterium multivorum*

S.paucimobilis (Sp) - *Sphingomonas paucimobilis*

S.wittichi (Sw) – *Sphingomonas wittichi*

S1P – sphingosine-1-phosphate

SB – sphingoid base

SDM - site-directed mutagenesis

SDS - sodium dodecyl sulphate

SDS PAGE - sodium dodecyl sulphate polyacrylamide gel electrophoresis

SEC – size-exclusion chromatography

Ser (S) – serine

SL - sphingolipid

SPT – serine palmitoyltransferase

*Bf*SPT – *Bacteroides fragilis* SPT

*Sp*SPT – *Sphingomonas paucimobilis* SPT

Sc *Sp*SPT – C-terminal His-tagged "Scottish" SPT

Am *Sp*SPT – N-terminal His-tageed "American" *Sp*SPT

ToF MS – Time of Flight Mass Spectrometry

Thr (T) – threonine

Tris - tris(hydroxymethyl)aminomethane

Tyr (Y) – tyrosine

UV-vis – Ultraviolet-visible range light spectrometry

V, V_{max} – observed and maximum enzyme reaction rates, respectively

Val (V) - valine

V.cholerae (*Vc*) – *Vibrio cholerae*

WT – wild type

Chapter 1: Introduction

1.1. Introduction to sphingolipids and their general features

Sphingolipids (SLs) are a diverse class of lipid molecules made of sphingoid or long-chain bases (LCB). Structurally LCBs consist of amino alcohols with a long hydrocarbon chain (a typical example being sphingosine, Fig.1.1). LCBs attach to a fatty acid via the amide linkage to produce ceramides, the simplest class of SLs that are further derivatised into more complex molecules. Acting as both structural components of cell membranes and signalling factors, SLs are responsible for a broad range of functions, not only enabling cell integrity but also being involved in cell signalling (Merrill, 2011; Fahy et al., 2005).

SLs are essential components of cells in all eukaryotes, however absent in most prokaryotes. While there is a high diversity of complex sphingolipids amongst various living organisms, they all share very common initial stages of biosynthesis with the first gateway step being universally catalysed by the serine palmitoyltransferase enzyme (SPT).

Sphingoid bases are defined as long-chain aliphatic amino-alcohols with their general structure usually represented by sphingosine (D-*erythro*-sphingosine, Fig.1.1). The molecule of sphingosine contains an 18-carbon chain with two alcohols (on C1 and C3, R- configuration), amine at C2, S- configuration, and E-double bond at C4. Another example of a major sphingoid base is sphinganine (Pruett et al., 2008).

There are numerous variations in structures of sphingoid bases that include the length of the chain, branching, presence and positions of certain head groups or non-saturated bonds. Appearance of specific sphingoid bases in the composition of SLs differs across taxonomic orders; one example being the higher proportion of phytosphingosine (4-hydroxysphinganine) in plants and fungi (Lynch et al., 2004). SL composition can also vary in different cells and tissues of one organism. The molecular differences between sphingoid bases govern certain types of interactions that have direct effects on the constituent state of biological membranes (Jaakishan et al., 2010; Kraft, 2017). It is generally acceptable to use the shorthand nomenclature for SLs stating the number of carbon atoms on their chains and number of unsaturated bonds; ceramide, for example,

is presented as d18:1/16:0 as seen on Fig.1.1.1; the unsaturated bond is assumed to be present at position 4 (Merrill, 2011; Fahy et al., 2005).

The most common mammalian SLs have their main chain lengths of 16 and 18 carbons. Odd linear chain lengths (17 and 19 carbons) containing branched alkyl chains have been detected in trace amounts in mammals (Carter & Hirschberg, 1968) while being more common in lower organisms. Most of these sphingolipids are thought to be the derivatives of fatty acids that are iso-branched, by having a methyl group on the penultimate carbon, or anteiso-branched, by having a methyl group on a carbon third from the end (Kaneda, 1991). Examples of species producing branched sphingolipids include the nematode worm *Caenorhabditis elegans* (Chitwood et al., 1995), bacteria from *Bacteroides* genus (Wieland Brown et al., 2013) and dental pathogen *Porphyromonas gingivalis* (Mun et al., 2007). Less frequent is branching in the middle of the LCB chain, one example being methylation occurring at C9 of glucosylceramide in *Pichia pastorris* (Ternes et al., 2006).

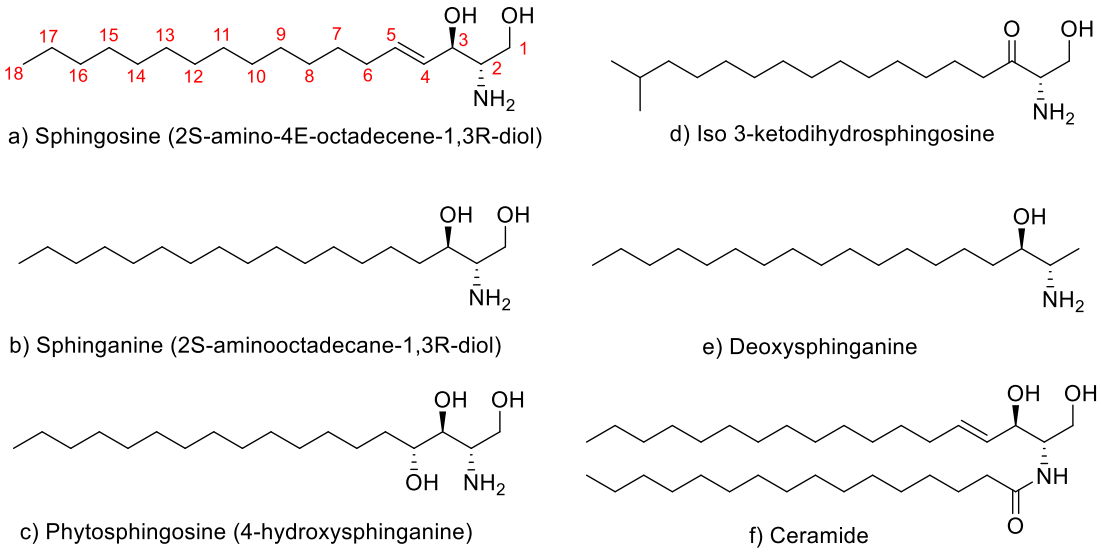


Fig. 1.1.1. The chemical structures of some sphingoid bases. Sphingosine (a) and sphinganine (b) are the two most common sphingoid bases in the majority of organisms except for plants and fungi that have the highest amounts of phytosphingosine (c). Iso 3-ketodihydrosphingosine (d, isoKDS) and deoxysphinganine (e, deoxyKDS) are two unusual products of SPT reactions. Several bacteria and nematode *C.elegans* produce iso-branched SLs, with isoKDS (d) being a proposed universal precursor. Deoxysphinganine (e) produced by a non-specific condensation of alanine and PCoA by SPT, is associated with several disorders such as HSAN type I and type 2 diabetes. The simplest type of sphingolipid is ceramide (f), made by condensation of a sphingoid base and a fatty acid base.

Branching of sphingolipids is reported to affect increasing membrane fluidity (Jaikishan et al., 2010); more importantly, branching modifications are thought to play a role in pathogen-host interactions, with the former protecting itself by bearing unique signature molecules not recognised by the latter (Buist, 2007).

Deoxy-sphingoid bases made due to a non-specific SPT-catalysed condensation of alanine instead of serine, are accumulated as toxic by-products and are associated with several disease conditions including rare genetical Hereditary sensory autonomous neuropathy type I (HSAN I), type 2 diabetes and Alzheimers (Duan & Merrill, 2015).

1.2. General mechanism of *de novo* sphingolipid biosynthesis

To appreciate how the vast abundance and diversity of sphingolipids arises it is necessary to look at the general mechanism of sphingolipid biosynthesis (Fig. 1.2.1).

The first universal step in sphingolipid biosynthesis involves condensation of L-serine with the coenzyme A-derived palmitic acid (palmitoyl-CoA) that is catalysed by the enzyme serine palmitoyltransferase (SPT). The product is the first sphingoid base 3-ketodihydrosphingosine (KDS); it is further processed via a cascade of modifications. Free sphingoid bases are usually present in small quantities as their accumulation is often cytotoxic; they are rapidly converted into ceramides - the simplest class of SLs consisting of a sphingoid base N-linked to a fatty acid (with the ceramide being a typical example) (Castro et al., 2014).

Ceramides themselves are an important and diverse class of SLs regulating membrane segregation and transport; they also have a profound effect on cell metabolism. Either an overproduction or deficiency of ceramides has been associated with numerous complex disorders such as obesity, diabetes, coronary heart diseases and cancer (Bikman & Summers, 2011).

Furthermore, ceramides are at the central point of the biosynthesis of sphingolipids as they are further converted into more complex SLs such as phosphosphingolipids involved in cell signalling, glycosphingolipids that act as cellular receptors and sphingomyelins that provide cover for the nerve cells (Fig. 1.2.1). Ceramides can also be converted back into simple sphingoid bases such as sphingosine; these, in turn, are

further used for a production of novel SLs or smaller signalling molecules such as sphingosine 1-phosphate (S1P), which is an important lipomediator.

The yeast biosynthetic pathway differs from the mammalian as certain points; most notably it uses phytosphingosine (4-hydroxysphinganine) as the main sphingoid base (Lynch et al., 2004). Yeast has served as a useful model organism for the study of SL biosynthesis and genetics (reviewed by Dickson, 2008). Several enzymes involved in the production of bacterial sphingolipids have been identified; however, the components of the full pathway so far have not been elucidated (Merrill, 2002; Merrill, 2011).

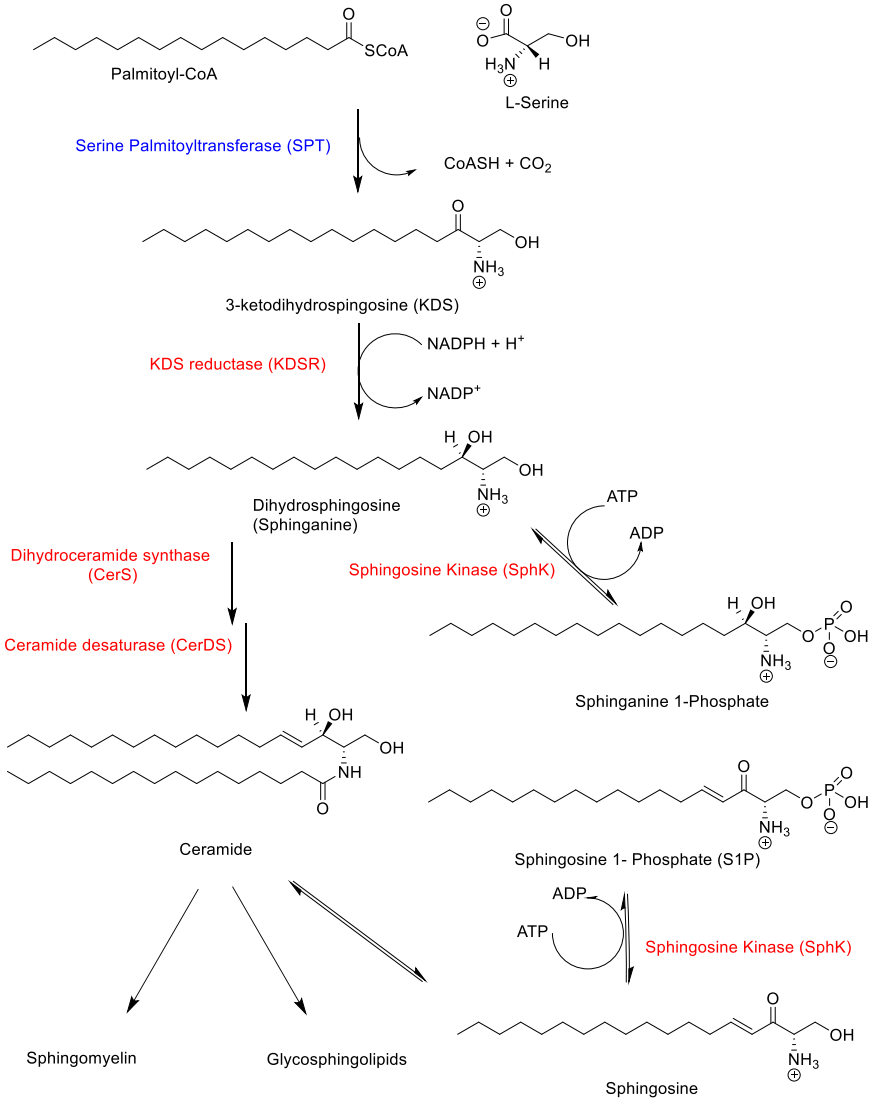


Fig. 1.2.1. A simplified general diagram of the sphingolipid biosynthesis in multicellular eukaryotes.

1.3. Chemistry of the pyridoxal 5'-phosphate (PLP)

Vitamin B6 is the general name given to a range of pyridine compounds and their phosphate derivatives (Fig. 1.3.1). These are precursors of biologically activated pyridoxal 5'-phosphate (PLP) - one of the most ubiquitous natural cofactors used by more than 160 enzymes (Percudani & Peracchi, 2003).

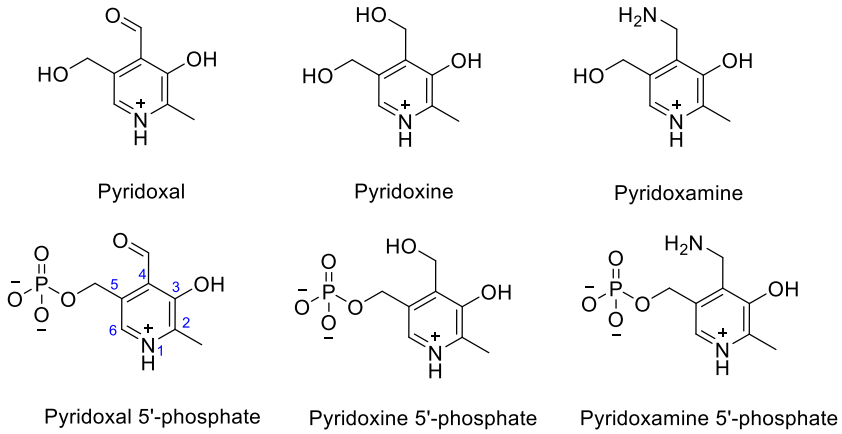


Fig. 1.3.1. Derivatives of vitamin B6 including pyridoxal 5'-phosphate

PLP-dependent enzymes are involved in a diverse range of amino acid reactions including transamination, decarboxylation, racemisation, substitution and elimination (reviewed by Metzler et al., 1954; Toney, 2005). Beyond this diversity lies a unifying principle first outlined by Dunathan, 1966.

In the active site of an enzyme, the PLP cofactor makes an internal aldimine bond via a Schiff base formation with a catalytic lysine residue. Upon binding of the amino acid substrate, a transaldimination reaction leads to a production of the PLP-substrate external aldimine. The universal stage in all PLP-catalysed reactions involves a group loss from the Cα of the bound amino acid. The reaction transition state assumes the so-called Dunathan intermediate (Fig 1.3.2), a geometric configuration where the Cα-R bond to-be broken is aligned perpendicularly to the PLP-aldimine plane. The overlap of the sigma bond with parallel pi orbitals causes hyperconjugation effect with the electronic density being withdrawn into the π-system. This weakens the sigma bond allowing it to be ultimately broken. Upon the departure of the leaving group R, the developing negative charge on the Cα is stabilised by the electron-withdrawing effects of PLP leading to a progression to the quinonoid transition state. The exact geometrical

configuration of the remaining groups around Cα in respect to the bond broken is what determines the type of reaction to occur (Dunathan 1966). This stage is a basis for the diversity of PLP catalysis as it determines the reaction type – either racemisation, substitution or elimination. The further reaction progression route is determined by the properties of an enzyme imposing certain specific 3D environment at the active site (Toney, 2011).

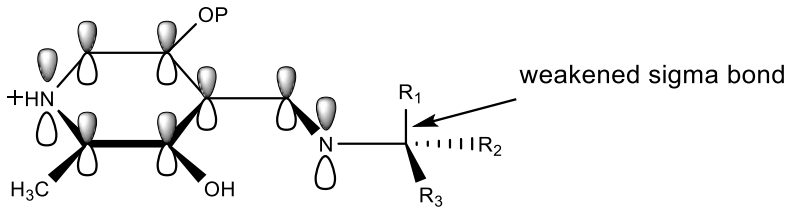


Fig. 1.3.2. The Dunathan intermediate formed by the PLP-amino acid external aldimine

The reaction specificity is therefore controlled by the enzyme which provides a spatial environment regulating the initial substrate binding and conformation of the groups around the Cα of the substrate. This is achieved by the specific residues around the active site. As proposed by Dunathan, the major factor in controlling the reaction specificity is the stabilisation of the substrate carboxylate group provided by the enzyme backbone.

In conclusion, enzyme architecture regulates the exact type of amino acid substrate that can bind to the PLP cofactor; it further determines the orientation of the substrate in the transition state, thus leading to a control of reaction specificity. (Eliot & Kirsch, 2004).

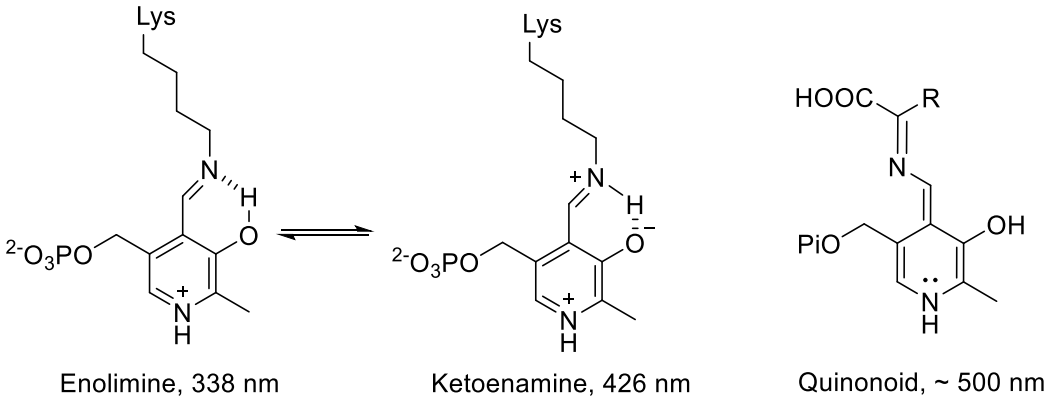


Fig. 1.3.3. The lysine-bound internal aldimine is present in two tautomeric forms, the enolimine (338 nm) and ketoenamine (426 nm). The quinonoid intermediate has the absorbance maxima at 500 nm.

The ability of the PLP cofactor to absorb in the near UV and visible spectrum is a significant advantage in studying of the PLP-dependent enzymes. Depending on a localisation of the proton, the PLP is present in two isoforms – enolimine and ketoenamine, absorbing at 338 and 426 nm respectively (Fig. 1.3.3). The equilibrium between these tautomeric forms changes upon the substrate binding, allowing spectral changes to be quantitatively monitored. The activated enzyme quinonoid intermediate species generally absorb around 500 nm (although this might vary), in some cases this enables detection of the transient intermediate during the reaction by means of UV-vis spectrophotometry.

1.4. Serine palmitoyltransferase (SPT)

For all known SL-containing organisms the first step in *de novo* sphingolipid biosynthesis is a Claisen-like condensation between the long-chain thioester (palmitoyl-CoA) and the amino acid L-serine yielding a sphingoid base 3-ketodihydrosphingosine (KDS) as the product (Fig. 1.4.1). This reaction is catalysed by the enzyme serine palmitoyltransferase (SPT).

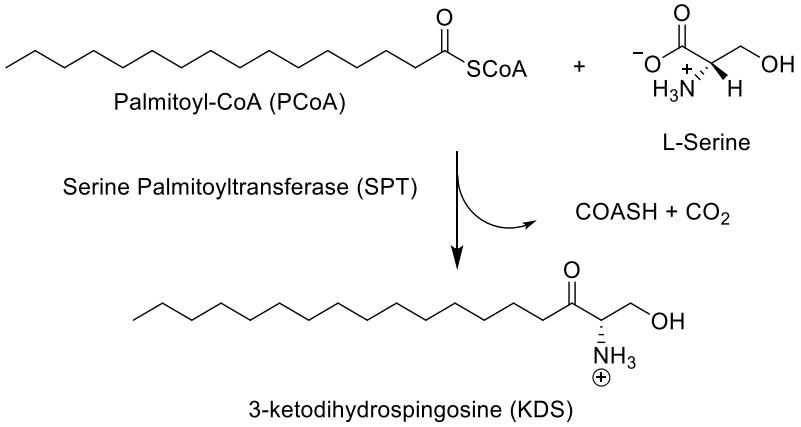


Fig. 1.4.1. Production of KDS from L-serine and palmitoyl-CoA by SPT

The above definition applies to most SPT-catalysed reactions including mammalian, yeast and *Sphingomonas paucimobilis* bacterial enzymes. However, the actual substrates and reaction products might vary either due to a natural preference for other substrate molecules in some organisms or due to particular mutations in SPT (such as the ones resulting in HSAN1 phenotype discussed later). As one of the purposes of this work is to gather and analyse the variety of substrates accepted by SPT, a broader definition of the SPT-catalysed reaction must be given to cover most of the cases. It is, therefore, can be described as decarboxylative Claisen-like condensation of acyl-derived long chain thioester and an amino acid that results in the formation of a sphingoid base as a product.

1.5. Eukaryotic and bacterial SPTs

Eukaryotic SPT is a complex membrane-associated enzyme composed of two major subunits LCB1 and LCB2 in yeastб with homological SPTLC1 and SPTLC2/3 in human (Weiss and Stoffel, 1997, reviewed by Lowther et al., 2012b); for clarity human subunits are referred to as hLCB1/hLCB2. The eukaryotic enzyme is a heterodimer with subunit hLCB2 bearing the catalytic Lys, while hLCB1 provides other residues involved in catalysis (Hanada et al., 2000b). Since its initial characterisation, smaller subunits of human enzyme ssSPTa and ssSPTb were discovered, presence of which is required for an increased activity of the main hLCB1/hLCB2 (Gable et al., 2000; Han et al., 2009) as well as a different type of the main subunit SPTLC3 (Hornemann et al., 2006), showing that eukaryotic SPT is a higher order heterodimeric complex. Expression and purification of eukaryotic SPT is, therefore, associated with numerous issues of solubility, and no structural information has been obtained so far. The notable exception is the SPT of a protozoan parasite *Toxoplasma gondii*, as up to date it has been the only eukaryotic SPT known to consist in a homodimeric form that is thought to be obtained via lateral gene transfer from bacteria (Mina et al., 2017).

In contrast, SPTs from sphingolipid-producing bacteria have become effective study models. SPT from the bacterium *Sphingomonas paucimobilis* was first identified and isolated by Ikushiro et al. in 2001. Unlike its eukaryotic counterpart, it is a readily soluble homodimer without membrane-associated domains. Due to the abundance of information, both structural and mechanistic, it remains an attractive target for studying the SPT mechanism.

The structures of two bacterial SPTs were determined by the Campopiano group in collaboration with the Jim Naismith group (St Andrews). SPT from *S.paucimobilis* was the first ever serine palmitoyltransferase to be structurally characterised in the PLP-bound holo- (Yard et al., 2007; pdb 2JG2) and the external aldimine form (Raman et al., 2009; pdb 2W8J), followed by the *Sphingomonas wittichii* SPT in the internal aldimine (Raman et al., 2010; pdb 2X8U). The structure of another SPT from *Sphingobacterium multivorum* was determined in the external aldimine form by Ikushiro et al. in 2009 (pdb 3A2B). Several other bacterial SPTs have been found and

characterised, including the membrane-associated SPT from *Bdellobacterium stolpii* (Ikushiro et al., 2007).

In *B.fragilis*, the gene BF2461 (reclassified into BF9343_2380 thereafter, for simplicity referred by the old name in the text) has been annotated as a putative SPT (57.4% seq. identity with *S.multovorum* SPT and 35.3% with *S.paucimobillis* SPT) and has been successfully targeted to generate SL-deficient organisms (Wieland Brown et al. 2013; An et al., 2014). Identical effects have been achieved in *P.gingivalis* upon deletion of homologous gene PG1780, resulting in SL deficiency and impaired survival of the strain (Moye et al., 2016). Furthermore, *B.fragilis* cells grown in the presence of myriocin, a known SPT inhibitor (Wadsworth et al., 2013) have also been SL-deficient (An et al., 2011).

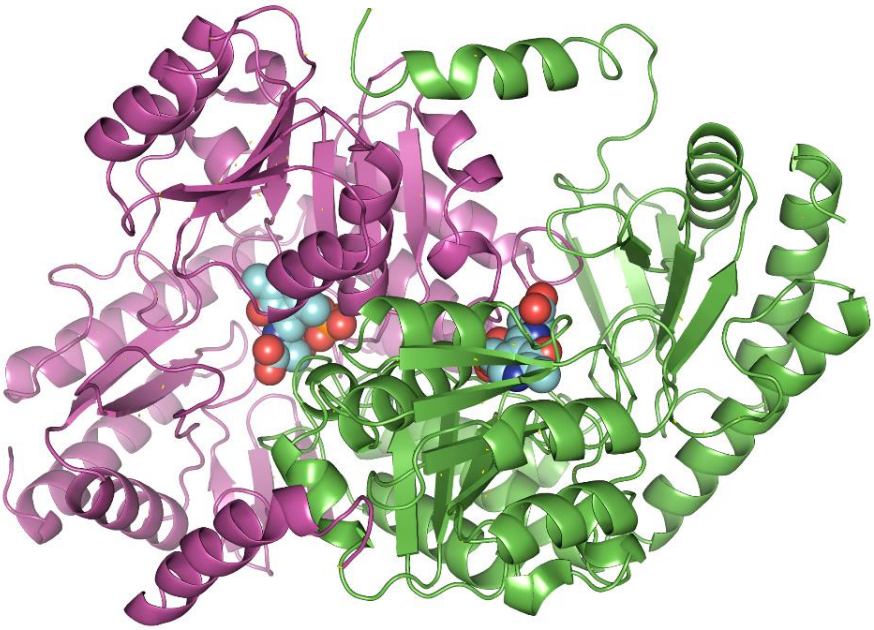
The BF2461 gene product from *B.fragilis* has been previously cloned, expressed and purified in the Campopiano group and confirmed to have an SPT activity (E.Bower, MRes thesis); one of the main aims of this work was processing and characterising this *Bf*SPT further.

1.6. The structure and the catalytic mechanism of *Sp*SPT

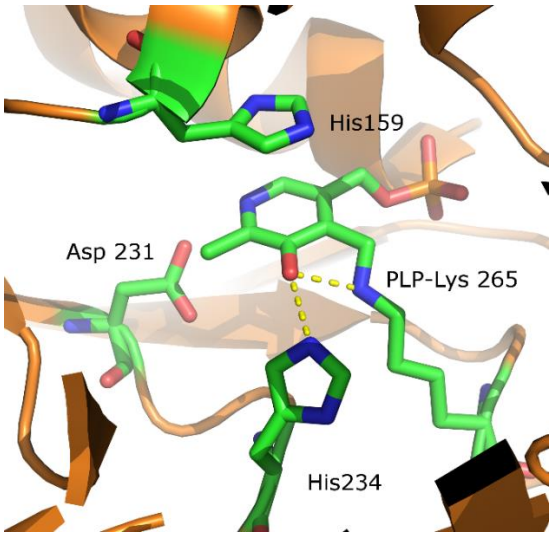
S.paucimobilis SPT (*Sp*SPT, pdb 2JG2) presents a homodimer with two active sites located at the domain interfaces (Fig. 1.6.1, A). It consists of three domains, the N-terminal, the central catalytic and the C-terminal (Yard et al., 2007). The PLP cofactor molecule in the active site is bond via the internal aldimine linkage to the catalytic Lys265 residue. Conserved residues His159, His234 and Asp239 (Fig 1.6.1, B) are involved in the stabilisation of the pyridine ring of PLP and are crucial for the activity of the enzyme (Fig 1.6.1, bottom). As stabilisation of the phosphate group of PLP has been shown to have a significant impact on the activity of SPT, residues involved in a formation of so-called "phosphate cup" – the hydrogen bonding network around it are also important (Beattie et al., 2013a). Arg378 and Arg390 (Fig 1.6.1, C) are two essential residues in the catalysis of the *Sp*SPT as both assist stabilisation of the carboxylate of L-serine at different stages of the reaction mechanism (Lowther et al., 2011). Arg 378 performs a conformational switch upon binding of L-serine and stabilises the carboxylate group of the external aldimine form. Arg 390 is involved in stabilisation of the reaction transition state and assisting decarboxylation.

His159 located in parallel to the PLP pyridine and is crucial for positioning of the cofactor. It also assists stabilisation of the carboxylate group of the external aldimine substrate (Fig. 1.6.1, C). The essential role of His159 is highlighted by the fact that replacement of the residue with alanine increases the K_d of L-serine by 55-fold and K_M by 10-20-fold; substitution of His159 with aromatic amino acids yields functionally inactive enzyme (Shiraiwa et al., 2009).

A



B



C

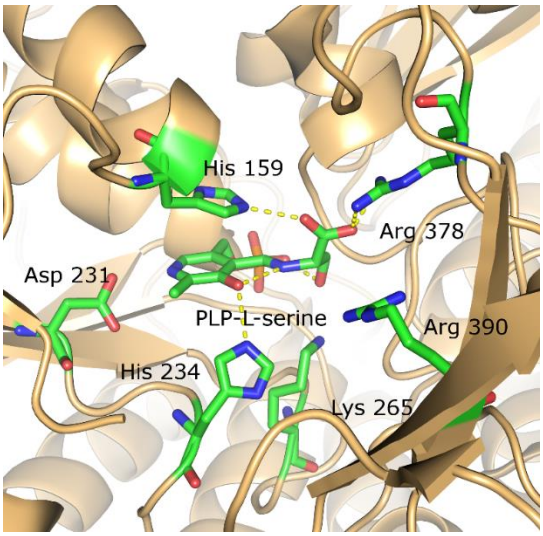


Fig.1.6.1. **A**: Structure of the *S.paucimobilis* SPT in PLP bound internal aldimine holo-form (pdb 2JG2). Bound PLP is shown as spheres. **B**: The *Sp*SPT in the PLP-bound internal aldimine form (pdb 2JG2). **C:** The external aldimine form of *Sp*SPT with the PLP bound to L-serine (pdb 2W8J).

The proposed mechanism of SPT reaction proceeds via the general Claisen-like condensation of L-serine and PCoA (Fig. 1.6.2) to yield KDS. The holo-enzyme has PLP bound to the catalytic Lys via an internal aldimine linkage (i). The amino acid substrate L-serine replaces Lys265 via transaldimination reaction. Upon formation of the external aldimine (ii), His159 forms a hydrogen bond with the carboxyl group of L-serine (Raman et al., 2009).

In the presence of palmitoyl-CoA, L-serine performs a rotational change into the Dunathan conformation (iii) where the Cα-H bond is located perpendicular to the imine bond and the pyridine ring of PLP. This is followed by the deprotonation of the Cα proton by Lys265. The deprotonation has been shown to occur in the absence of the second substrate as well but is 100-fold slower (Ikushiro et al., 2008).

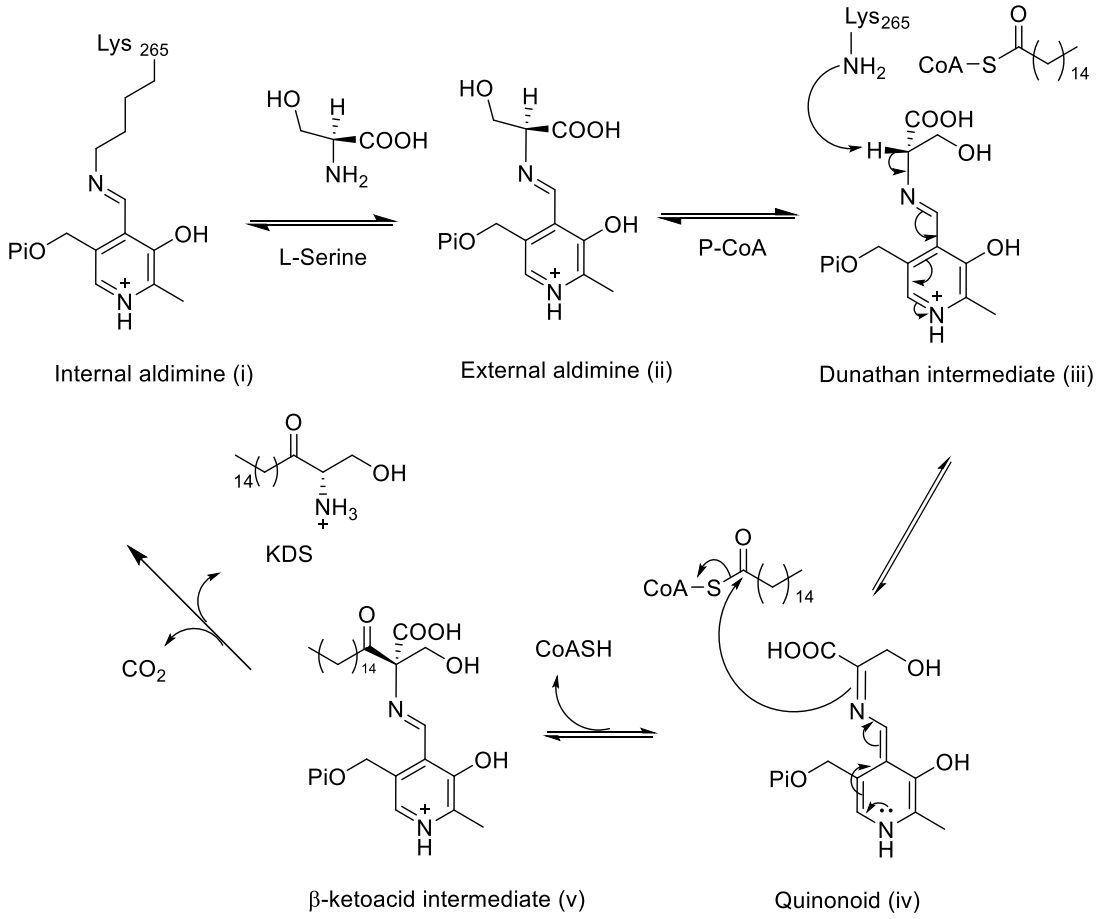


Fig. 1.6.2. The proposed reaction mechanism of SPT

Upon deprotonation of the Cα, the reaction proceeds via generation of reactive quinonoid intermediate (iv) that attacks the electrophilic carbonyl of the PCoA,

causing the formation of β-ketoacid intermediate (v). The following decarboxylation is assisted by interaction with R390 that is thought to stabilise the leaving carboxyl group. Upon the KDS-release step, the enzyme is recovered to the internal aldimine structure (i).

S.paucimobilis SPT accepts palmitoyl-CoA as its primary acyl-CoA substrate, although it can also utilise acyl-CoAs with different lengths of carbon chains such as myristoyl- (C14) and stearoyl-CoA (C18) (Reviewed by Hanada, 2003). Wild-type SPT strictly accepts L-serine as its first substrate, although mutated forms of human SPT are known to use L-alanine as a substrate leading to accumulation of toxic by-products (Penno et al., 2010).

1.7. Inhibition of SPT

Several natural inhibitors of SPT have been studied up to date, each revealing features of the enzyme reaction mechanism. The structures of some inhibitors are presented in the Figure 1.7.1.

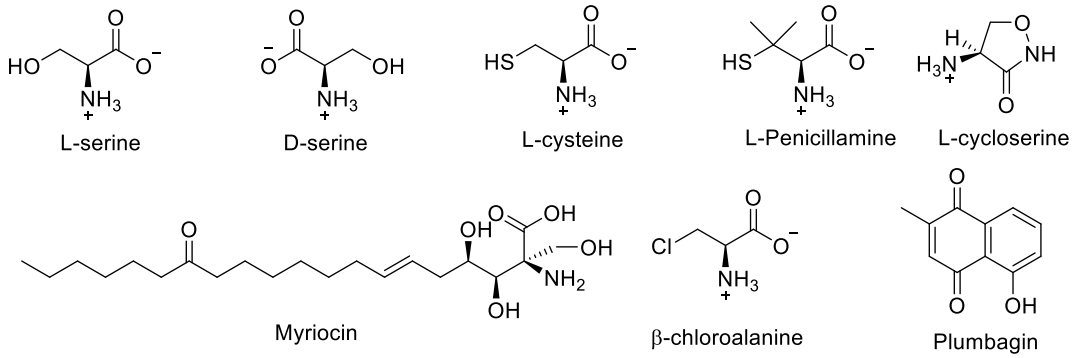


Fig. 1.7.1. Structures of the substrate L-serine and various known inhibitors of SPT

SPT is sensitive to the stereometrical configuration of the substrate. D-serine has been reported as a weak competitive inhibitor, being able to form an external aldimine with PLP. However, it cannot be utilised by SPT as substrate as a subsequent addition of PCoA to the reaction does not lead to production of KDS. It is likely that the position of hydrogen on Cα is crucial for deprotonation (Hanada et al., 2000a; Ikushiro et al., 2003).

L-cysteine and L-penicillamine, its non-proteinogenic dimethyl amino acid derivative, resemble serine with the hydroxyl group replaced by thiol. Both inhibit SPT via formation of covalent adducts with PLP (Hanada et al., 2000c); this is reflected by changes in the spectral profile of enzyme. The inhibition is reversible,[as SPT activity can be restored by dialysis with fresh PLP, indicating that no covalent modification of the protein is present. (Lowther et al., 2012a).

L-cycloserine, a natural product that resembles a cyclic version of serine, is also known to reversely inhibit PLP-dependent enzymes by a formation of isoxazole adducts (Peisach et al., 1998; Azam & Jayram, 2016). Surprisingly, inhibition of SPT with both D- and L- enantiomers of cycloserine reveals the formation of pyridoxal monophosphate (PMP) and an aldehyde product. This indicates a specific inhibitory mechanism for SPT involving hydrolytic ring opening on cycloserine with a subsequent decarboxylation and the amide bond hydrolysis (Lowther et al., 2010).

β-chloroalanine (BCA) is known as a suicide inhibitor forming an irreversible covalent adduct with the PLP-dependent enzymes (Medlock & Merrill, 1988). The proposed mechanism involves elimination of chloride and formation of aminoacrylate complex (Fig. 1.7.2). Upon release of the product, the reaction can proceed via two routes: the aminoacrylate is either hydrolysed to pyruvate and ammonia or inactivates the internal aldimine complex by a nucleophilic attack. The partition ratio between both reactions is 300:1; in the first case the catalytic cycle continues, while in the second, inactive complex is formed with a characteristic absorption at 330 nm (Ikushiro et al., 2004).

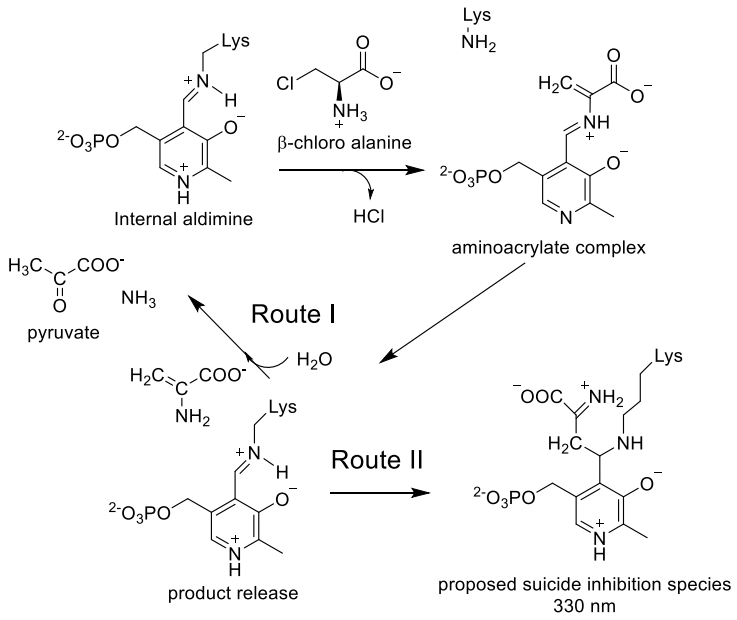


Fig. 1.7.2. The proposed inhibitory mechanism of PLP-dependent enzymes by β-chloroalanine.

Natural fungal product myriocin is the most commonly used SPT inhibitor in the studies of SL biosynthesis in all organisms (Chen et al., 1999). Its mechanism of inhibition was revealed by Wadsworth et al., 2013. Both mass spectrometry and structural data had shown that myriocin acts as a dual-mode inhibitor, firstly, by making a stable PLP adduct that can be reversely removed, and secondly, by slow degradation into C18- aldehyde containing species that condense with the active site lysine leading to an irreversible "suicide" inhibitory complex.

One of the most recent molecules reported to have inhibitory effects on AOS members is a natural plant compound plumbagin (Choi et al., 2012). It was successfully used to stop the synthesis of a natural product curvulamine in *Curvularia sp*, likely by inhibiting a novel, unknown enzyme from the AOS family (Han et al., 2014). Up to date, there were no studies proposing the inhibitory mechanism, although its structure shows a brief similarity with PLP. It would be interesting to investigate any possible effects of plumbagin and its inhibitory mechanism using bacterial SPT as a model.

Characterisation and study of SPT and other AOS family enzyme inhibitors can help to shed more light on the reaction mechanism and specificity of SPT. In this work, structural studies of SPT with D-Serine, L-penicillamine and L-cycloserine were attempted. Despite lack of the progress to obtain crystal data the with D-serine bound,

co-crystallisation studies of *Sp*SPT have revealed some interesting observations with regards to resolution of the substrate, thus pointing to what could be a previously overlooked property of the enzyme. Furthermore, structural data obtained on *Bf*SPT in the complex with L-penicillamine reveals evidence of the inhibitory mechanism as well as points towards the residues involved in the enzyme catalysis.

1.8. Sphingolipid metabolism in health and disease

Sphingolipids are essential components of the cell membrane and their production is essential in most eukaryotes. Studies targeting the sphingolipid biosynthesis pathways in different organisms show that while simple organisms, such as yeast and *C.elegans*, are viable without production of SLs, in higher eukaryotes they are essential for survival.

Studies of both *B.fragilis* (Wieland Brown et al., 2013; An et al., 2014) and *P.gingivalis* (Moye et al., 2016) demonstrated that bacteria can survive without making SLs, although their stress coping abilities and long-term survivability are impaired. Removal of SPT-encoding gene LCB1 in yeast *S.cerevisiae* results in the development of sphingolipid-deficient cells that are unable to control proton influx and thus are highly compromised at lower pH (Pinto et al., 1992).

Deletion of serine palmitoyltransferase gene results in lethal phenotypes in organisms such as *Drosophila* (Adachi-Yamada et al., 1999) and mice (Hojjati et al., 2005), unless sphingoid bases are supplemented externally. This evidently shows a dependence of higher organisms on *de novo* sphingolipid biosynthesis.

In mammals, sphingolipids are part of the diet; however, most SLs are degraded in the intestine, while only a fraction is used for biosynthesis (Schmelz et al., 1994). It is likely that dietary SLs are prevented from being incorporated into the synthetic pathway, as otherwise a complex mechanism would be necessary for selection of the appropriate bases present in given species (Vesper et al., 1999).

While mutations that completely disrupt the activity of SPT in humans are lethal, minor missense mutations can reduce the activity of SPT and change its specificity. The most common example of this is a rare genetic disorder known as hereditary sensory autonomous neuropathy type I (HSAN I) that leads to a progressive sensory loss as well as the development of ulcers and limb pathologies. The condition is caused by mutations in the hLCB1 subunit of human SPT (Dawkins et al., 2001). Along with reducing the overall activity of SPT, HSAN I-associated mutations allow the enzyme to process L-alanine as a substrate, ultimately leading to a build-up of toxic deoxysphingolipids (Penno et al., 2010). Several disease-related mutations have been

studied and characterised using bacterial SPT as a model of human SPT. The effects include general reduction of SPT activity, reduction of substrate affinity and production of non-soluble protein (Raman et al., 2009; Beattie et al., 2013b), although the mutant enzymes did not yield a deoxyKDS-producing phenotype. Studies used homology modelling based on the 3D structure of bacterial SPT have led to the production of mutants of the mammalian SPT that can generate deoxy-products (Gable et al., 2010; Bode et al., 2016).

More importantly, deoxysphingolipids associate not only with a rare HSAN I condition but also with more common disorders such as type II diabetes and metabolic dysfunction syndrome (Kowluru, 2014). Deoxysphingolipids are now used as specific biomarkers for these disorders (Othman et al., 2015). The accumulation of deoxySLs causes cytotoxic effects and impacts the insulin-producing cells. Targeting the build-up of these toxic metabolites could potentially form a basis for a novel therapy against type II diabetes; therefore, studies of SPT reaction specificity could provide a key milestone towards it.

1.9. *Bacteroides fragilis*

The human body is colonised by numerous bacteria species that, being either commensal, symbiotic or pathogenic, all together make up the human microbiome. Bacteria play a considerable part in human metabolism, immunity and protection against disease (reviewed by Reid, 2004). Organisms from *Bacteroides* genus are the most populous group of the human gut commensals making up to total 25% of the gut microflora (reviewed by Wexler, 2007). Despite not being the most abundant, *B.fragilis* is one of the species amongst the genus that draws most of the attention. Being part of the healthy human microflora, this gram-negative obligate anaerobe at certain conditions is often reported as a case of an opportunistic infection. Enterotoxigenic strains of *B.fragilis* (ETBF) have been linked with diarrhoea, inflammatory bowel disease and colon cancer (reviewed by Sears, 2009). The pathogenic effect of *B.fragilis* is contributed by two virulence factors: the bacterial capsule that is linked to abscess formation, and a specific protein toxin fragilysin. The latter one is a broad-spectrum metalloprotease that, in its active form, degrades E-cadherin receptors on the human gut epithelium, thus contributing to bacterial permeability into tissues (Goulas et al., 2011).

An even more striking feature of *B.fragilis* is its advanced role in modulation of the host's immune system. It produces a range of specific polysaccharides that initiate conversion of CD4 cells and subsequent production of anti-inflammatory cytokines in humans. Although this response is also related to bacterial pathogenicity, under normal conditions it is beneficial to the host, as it prevents gastrointestinal inflammation. Furthermore, the transplantation of *B.fragilis* to a deficient animal host was shown to prevent and even eliminate the symptoms of induced colitis (Mazmanian et al., 2008). As a commensal organism, *B.fragilis* is plays a role in regulating favourable conditions in the human intestine that lead to maintenance of a healthy bacterial microflora (reviewed by Surana and Kasper, 2012).

The most interesting aspect in the context of this review is that *B.fragilis* is one of the bacteria that produce sphingolipids in the human body. These SLs bring a unique molecular signature by being iso-branched on the end of their long carbon chain. Furthermore, compelling evidence suggests that functions of these bacterial isoSLs are

not limited to serving as a membrane building blocks; they also act as signalling factors with an impact on the host's immune system. The significance of these iso-branched sphingolipids is discussed in the next chapter.

1.10. Bacterial iso-branched sphingolipids in the human microbiome

While most of the fatty acids in bacteria are present in a straight-chain form, occurrence of less common branched-chain fatty acids is known for about 10% of bacterial species, amongst a few of which are the *Bacillus, Sphingobacterium spp*. (reviewed by Kaneda, 1990), and Bacteroides (Mayberry, 1980).

The branched fatty acids are produced from α-keto acids that are derivatives of valine, leucine and isoleucine. These fatty acids serve as precursors of phospholipids, having an impact on the bacterial cell membrane fluidity (Kaneda, 1977); they also appear to be the precursors in the biosynthesis of the branched sphingolipids in bacteria. The presence of the iso-branched SLs (containing a methyl group at the pre-terminal carbon of the LCB) has been detected in the soil-dwelling *Sphingomonas spp.* (Naka et al., 2003). Despite this, however, the biological roles of isoSLs in *Sphingomonas spp.* are not fully understood.

On the other hand, much of the recent interest is caused by the bacteria from the human microbiome that are associated with the production of isoSLs due to increasing evidence of their impact on human health (Heaver et al., 2018).

The iso-branched sphingolipids are not known to be made by humans or mammals; their presence in human tissue isolates can be traced to five bacterial genera of *Bacteroides*, *Parabacteroides*, *Prevotella*, *Tannerella* and *Porphyromonas (*Kato et al., 1995; Nichols et al., 2011); the former four being present in the human gut while the last one being a dental pathogen. While the exact functions of these bacterial sphingolipids have not been fully elucidated yet, the compelling evidence shows their involvement in intracellular signalling and mediating host immune cell response.

Porphynomonas gingivalis, a dental pathogen responsible for periodontal disease, produces a range of ceramides that are found in affected dental tissues and thought to

stimulate inflammatory responses (Nichols et al., 1998). Specific sphingolipid fractions isolated from diseased tissues have been shown to induce autoimmune reactions in a mouse model (Nichols et al., 2009). Mass spectrum analysis has identified that sphingolipids produced by *P.gingivalis* are composed of iso-branched sphingoid bases (Nichols et al., 2004; Nichols et al. 2006, Mun et al., 2007). Deletion of the gene PG1780 that encodes a putative SPT of *P.gingivalis* was shown to produce viable organisms, while their survival and stress response was impaired (Moye et al., 2016).

Bacteria colonising the gut are the largest group of SL-producing organisms in the human body. Due to its relevance in human health, *Bacteroides fragilis* also remains one of the most studied SL-producing organism in the human gut. While the presence of sphingolipids in *B.fragilis* was detected in the late 1970s, (Miyagawa et al., 1979), their functional importance has been highly underexplored until several recent studies. In 2011, An et al. produce sphingolipid-deficient cells of *B.fragilis* by chemical targeting of serine palmitoyltransferase, showing that the resulting cultures are viable, although they have lower resistance against the experimentally induced stress. The authors propose that SLs protect the cells by directly being the key components of membrane domains and also composing the signalling factors.

Further, in their major work, Wieland Brown et al., 2013, made an extensive structural mass analysis of *B.fragilis* sphingolipids, all of which are composed of iso-branched LCBs. Most importantly, amongst the bacterial SLs authors detected an analogue of α-galactosylceramide (α-GalCer, shown in figure 1.10.1) - a potent inducer of the invariant natural killer T cells (iNKT), initially known to be produced by a marine sponge (Bendelac & Savage, 2007). It has been proposed before that α-GalCer can be present in the human body in some form based on its high specificity towards the human immune cell receptors (Brossay et al., 1998). While earlier α-GalCer was found in the soil-dwelling *Sphingomonas* (Kinjo et al., 2005) where it has demonstrated the ability to activate iNKT cells, its presence in *B.fragilis*, a human commensal bacterium, is more significant due to a potential direct health impact.

Using immunology assays, Wieland Brown et al., 2013, showed that α-GalCer from *B.fragilis* activates both mouse and human iNKT cells, thus establishing a link towards the human immune regulation.

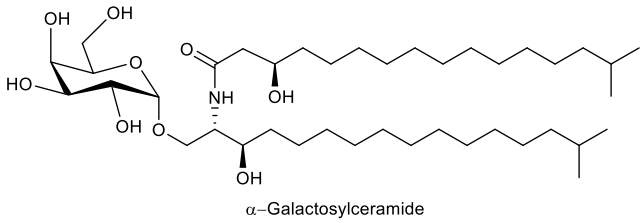


Fig. 1.10.1. The structure of α-galactosyl ceramide isolated from *B.fragilis* contains iso-branched fatty acid and sphingoid bases.

An et al., 2014, found that the total extract of *B.fragilis* SLs provides overall antagonistic effects on the gut receptors, thus increasing proliferation of iNKT cells; SL extracts of *B.fragilis* have been shown to provide protection against experimentally induced colitis in germ-free mice.

In several of the above studies, the authors have been targeting bacterial SPTs to investigate the impact on the SL-biosynthesis. Deletion of BF2641 in *B.fragilis* produced SL-deficient organisms (Weiland Brown et al., 2013; An et al., 2014) with similar results reported upon targeting the analogous gene PG1780 in *P.gingivalis*. Furthermore, *B.fragilis* cells grown in the presence of myriocin are also SL-deficient (An et al., 2011).

The study of sphingolipids in human microbiome remains a new and stimulating area of research. It is likely that SLs are one of the means of establishing evolutionary long microbiome-host interactions. While the specific roles of the iso-branched SLs produced by human microbiome bacteria are yet to be fully understood, their overall complexity and impact on human homeostasis are evident (Heaver et al., 2018).

However, what remains likely from the above evidence, is that the production of isoSLs in the bacteria from human microbiome can be totally controlled at the first rate-determining step catalysed by serine-palmitoyl transferase (SPT). This highlights the importance of *Bf*SPT as a study model for the enzymes controlling the isoSL-production in the human microbiome.

1.11. Iso-branched sphingolipids in Eukaryotes

Apart from the bacteria, *C.elegans*, a free-living nematode that is a common research model in biology, produces iso-branched fatty acids and SLs (Chitwood et al., 1995). SLs in *C.elegans* are critical for postembryonic development (Zhu et al., 2013); they provide an influence on ageing. Studies of the SL-deficient organisms show that they have slower developmental rates, but their lifespan increases (Cutler et al., 2014).

A recent study by Hannich et al., 2017, outlined the biosynthetic route in the production of the iso-branched fatty acids in *C.elegans* (Fig.1.11.1). By using heavy carbon amino-acid labelling, authors showed that the amino acid leucine is incorporated into branched-chain acyl-CoA via an α-ketoacid route, it then enters several cycles of fatty acid chain elongation. Authors detected two types of sphingoid bases by MS: C17-iso-sphinganine and C17-iso-deoxysphinganine (iso-branched analogues of the straight-chain sphinganine and deoxysphinganine). This suggests that the acyl-CoA substrate of *C.elegans* SPT is isopentadecanoyl-CoA, while the observed products are being made by its condensation with serine and alanine respectively. Similar mechanisms of iso-branched fatty acid production are thought to be present in bacteria (Kaneda, 1977). It is likely that the isoSL-producing bacteria in the human microbiome, including *B.fragilis*, have a similar route and make the isopentadecanoyl CoA (isoC15-CoA), that is used as a precursor in the biosynthesis of SLs.

In analogy to deoxysphingolipids, iso-sphingolipids are not metabolically processed by normal SL-producing organisms. Complementation of SL-deficient *C.elegans* with a straight chain SLs is not compatible with the metabolism of the worm; furthermore, an addition of isoSLs to an SL-deficient yeast *S.cerevisiae* does not support the growth of the cells and is toxic (Hannich et al., 2017). Incompatibility with the straight-chain SLs was recently used as a protection mechanism against plant-parasitic nematodes by Gao et al., 2016, demonstrating that non-branched sphingoid bases can induce nematocidal effects. These findings further highlight the importance of chemical differences between the straight-chain and iso- SLs. Iso-branched SLs have been discovered in parasitic nematodes, namely *Ascaris suum* living in a pig intestine (Lochnit et al., 1997). As parasitic nematodes share the same environment as the isoSL-producing bacteria in the mammal body, it might be more than a coincidence

that these species have adopted the use of isoSLs and could indicate a co-evolutionary strategy being adopted by the parasites in both utilising available food sources and protecting themselves from the host's immune system.

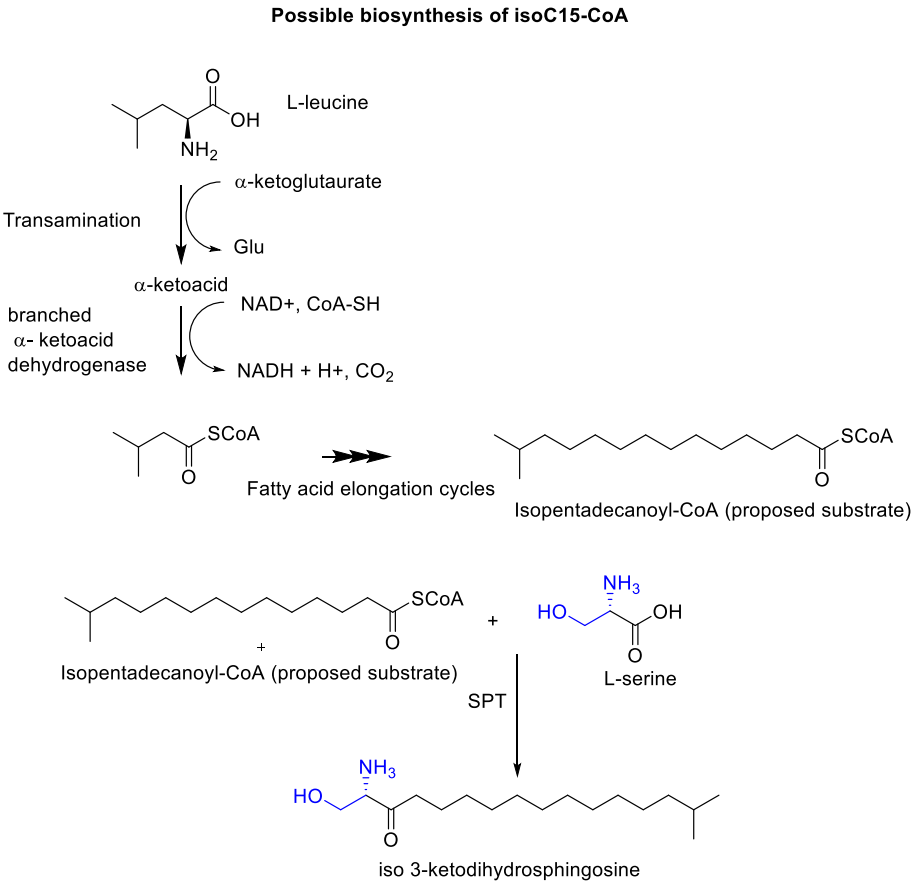


Fig. 1.11.1. A mechanism of iso-branched fatty acid formation as elucidated by Hannich et al., 2017 in *C.elegans*. It is possible that a similar mechanism occurs in *B.fragilis* to yield isopentadecanoyl-CoA (isoC15) as the primary substrate for SPT. The route to αGC production imposes that the substrate of *Bf*SPT is isopentadecanoyl-CoA (isoC15-CoA) (based on Hannich et al., 2017 and Wieland Brown et al., 2013).

1.12. The AOS family of enzymes

SPT belongs to the α-oxoamine synthase (AOS) family of enzymes within a wider fold type-I family of PLP-dependent enzymes. AOS enzymes catalyse the condensation of amino acids with acyl thioesters via general Claisen-like mechanism (Fig. 1.12.1). This ever-expanding family up to date counts 13 proteins that utilise various substrates and are involved in diverse metabolic pathways, 5 of which have been structurally characterised (table 1.12.1).

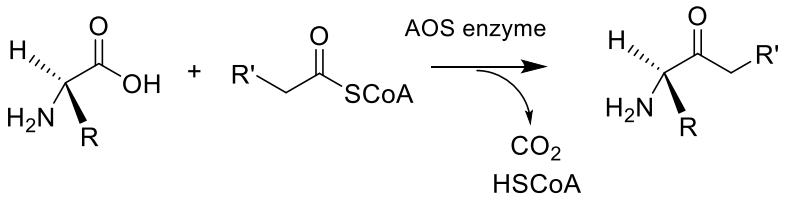


Fig. 1.12.1. Simplified general reaction mechanism of the AOS enzymes

Apart from the SPT, other structurally characterised members of the family are 5-aminolevulinate synthase (ALAS; Astner et al., 2005), 8-amino 7-oxononanoate synthase (AONS; Alexeev et al., 1998), 2-amino 3-ketobutyrate CoA ligase (KBL; Schmidt et al., 2001) and *Vibrio cholerae* CAI-1 synthase (CqsA; Jahan et al., 2009). The information on the enzymes with their pdb codes can be found in Appendix 5. All proteins have a 20-33% sequence identity (Fig. 1.12.3).

The enzyme 8-amino 7-oxononanoate synthase (AONS or BioF) catalyses the decarboxylative condensation of L-alanine with pimeloyl-CoA, which is the first step in the biosynthesis of the natural cofactor biotin in plants and bacteria (Alexeev et al., 1998; Webster et al., 2000). ALAS synthesises 5-aminolevulinate acid (ALA) from glycine and succinyl CoA in animals, fungi and some bacteria; ALA is an essential precursor in heme biosynthesis (Astner et al., 2005). The KBL is involved in threonine utilisation pathway as it converts its degradation product 2-amino 3-ketobutyrate into glycine and acetyl-CoA. Notably, it is the only known AOS family member that does not have a decarboxylation step (Schmidt et al., 2001). CqsA is responsible for the synthesis of the main quorum sensing factor in *V.cholerae* which acts as a mean of signalling between bacteria (Jahan et al., 2009). The substrate specificity of CqsA was

a matter of debate until it was reported to accept both decanoyl-CoA and S-adenosine methionine (SAM) as a substrate *in vivo* (Kelly et al., 2009; Ng et al., 2011).

Enzyme	Pathway	AA substrate	Thioester substrate	Product	Pdb ID
Serine Palmitoyltransferase (SPT)	Sphingolipid biosynthesis	L-serine	Palmitoyl-CoA	3-ketodihydro-sphingosine (KDS)	*Sp* holo: **2JG2** *Sp* PLP-L-serine: **2W8J**
8-amino 7-oxononanoate synthase (AONS)	Biotin	L-alanine	Pimeloyl-CoA	5-aminolevulinate	*Ec* Apo: **1BSO** *Ec* Holo: **1DJE** *Ec* AON-PLP: **1DJ9**
5-aminolevulinate synthase (ALAS)	Tetrapyrrole	Glycine	Succinyl-CoA	8-amino 7-oxononanoate	*Rc* holo: **2BWN** *Rc* glycine-PLP: **2BWP** *Rc* succinyl-CoA-PLP:**2BWO**
2-amino 3-ketobutyrate CoA ligase (KBL)	Threonine degradation	Glycine	Acetyl-CoA	2-amino 3-ketobutyrate	*Ec*: **1FC4**
***Cholerae* CAI-1 synthase (CqsA)**	Quorum sensing	S-adenosine-methione (SAM)	Decanoyl-CoA	Cholera autoinducer (CAI-1)	*Vc* apo: **2WK7** *Vc* holo: **2WK8**

Table 1.12.1. Summary of the AOS-family enzymes that have been structurally characterised. Abbreviations: *Sp - Sphingomonas paucimobilis, Ec - Escherichia coli, Rc - Rhodobacter capsulatus, Vc – Vibrio cholerae.*

Other AOS-family proteins have been characterised to a lesser extent; these involve bacterial enzymes involved in production of various metabolites including other autoinducer factors (Spirig et al., 2008, Hornung et al., 2013), pigments (Williamson et al., 2005; Garneau-Tsodikova et al., 2006) and natural products (Burke et al., 2007; Gerber et al., 2009). Recently it was reported that another AOS family member might be involved in the synthesis of an antibacterial alkaloid curvulamin in marine fish symbiont *Curvularia sp* (Han et al., 2014). One of the initial steps involves condensation of alanine with 4-carbon acyl thioester. No curvulamin was produced by bacteria in the presence of cycloserine and plumbagin, known inhibitors of SPT and AONS, suggesting that this unknown enzyme might be a member of the AOS family.

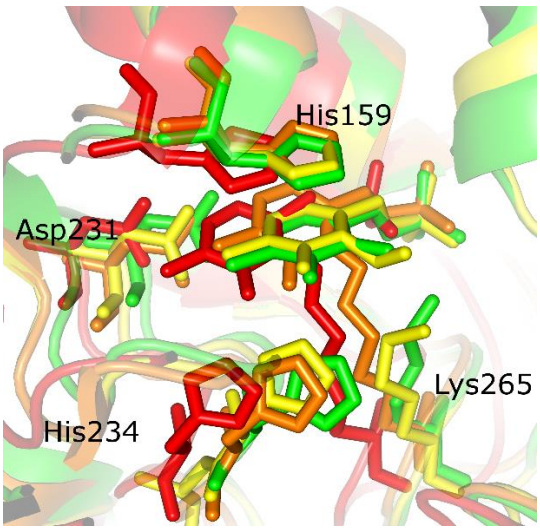


Fig. 1.12.2. Superimposed structures of 4 AOS family enzymes with the conserved residues involved in positioning of PLP shown as sticks. The proteins are KBL (1FC4, green), ALAS (2BW0, yellow), AONS, (1DJE, red), SPT (2JG2, orange). The *Sp*SPT sequence numbering is used.

Structural studies of the AOS-family enzymes reveal a range of similar features in the general fold symmetry and accommodation of the PLP molecule (Fig. 1.12.2). All proteins are present in the form of homodimers, with both catalytic sites located on the dimer interfaces. The position of PLP is superimposable in all structures, with the key stabilising residues (discussed as numbered in *Sp*SPT) being lysine 265, histidine 159, aspartic acid 231 and histidine 234. Conservation of the above residues across the five structurally characterised AOS enzymes is highlighted in the figure 1.12.3. Histidine 159 lays in parallel to the PLP ring; the residue is essential for the activity in the AOS enzymes as it regulates the position of the cofactor (Shiraiwa et al., 2009) and assists stabilisation of the amino acid carboxylate in the external aldimine form (Raman et al., 2009). Aspartic acid 231 is involved in PLP stabilisation by forming a salt bridge with the protonated nitrogen on the pyridine ring, while histidine 234 stabilises the hydroxyl group of the cofactor (Fig.1.12.2).

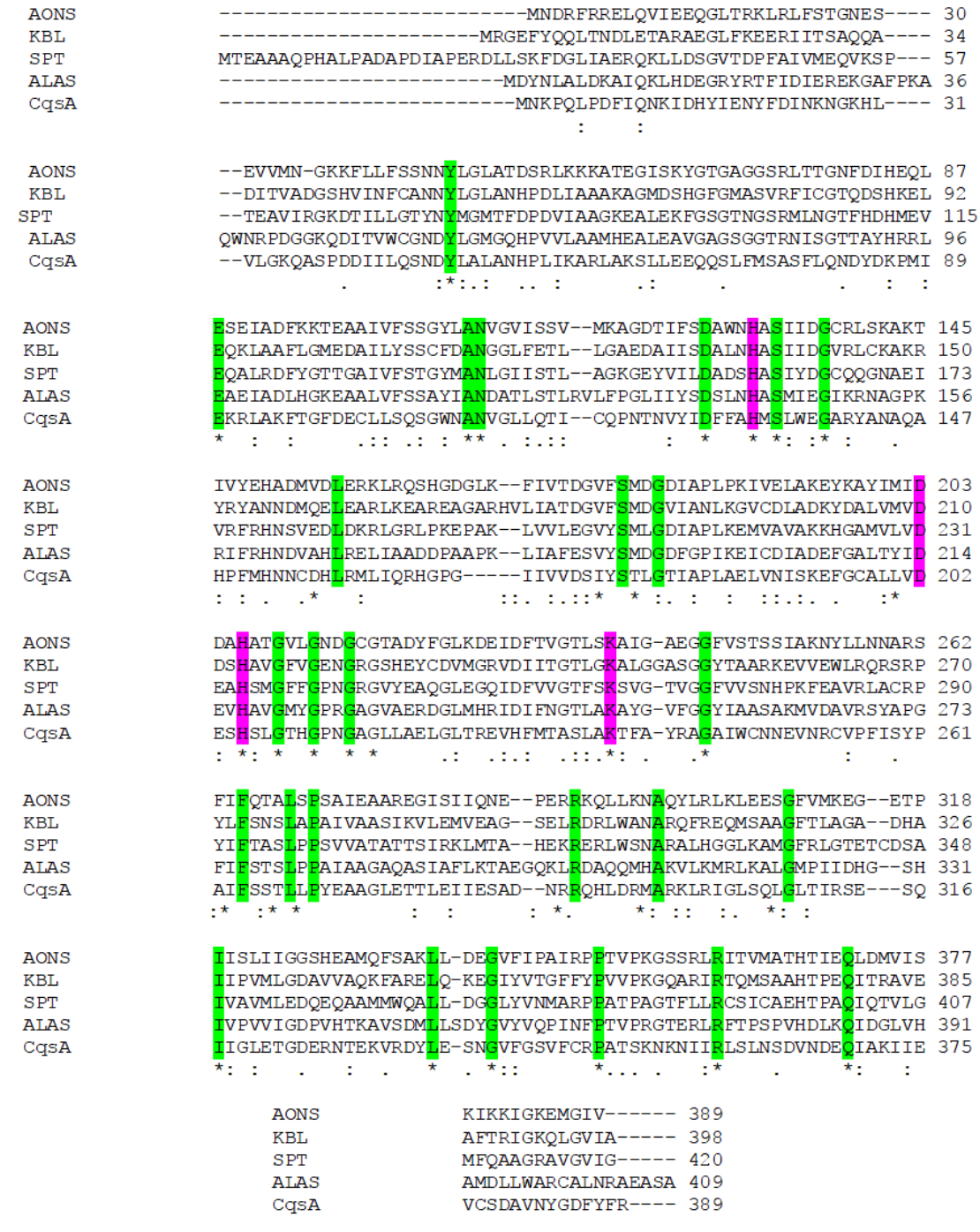


Fig. 1.12.3. Sequence alignment of five AOS family members whose X-ray structures have been determined. Conserved residues are highlighted in green, residues directly involved in PLP binding and stabilisation highlighted in pink. The alignment was generated in ClustalW2. The above UniProt genes were used: *E.coli* AONS: (P12998), *E.coli* KBL: (P0AB77), *S.paucimobilis* SPT: (Q93UV0), *Rhodobacter capsulatus* ALAS : (P18079), *V.cholerae* CqsA: (Q9KM65).

1.13. Towards expanding the catalytic potential of AOS family of enzymes

Enzyme catalysis is becoming a desirable alternative to chemical synthesis (reviewed by Jäckel et al., 2008). Enzymes work at moderate conditions without requiring expensive or toxic catalysts; their natural specificity allows to simplify production of high-value chiral compounds and reduce the costs of their subsequent purification. Enzymes are now finding more ways to be used as industrial catalysts in the production of biofuels and pharmaceuticals. Rational enzyme design and directed evolution strategies allow expanding catalytic potentials of existing enzymes and even creating novel ones (reviewed by Turner, 2009).

The enzymes from the AOS family are good candidates for biocatalytic engineering. By having high structural homology, conservation of key residues and utilisation of similar catalytic mechanism, the AOS enzymes can process a variety of different substrates and yield a broad range of products. Furthermore, certain AOS family members have been reported with dual catalytic functions. One example is ORF34 from *Streptomyces aizunenesis* with bifunctional ALAS and ALA-CoA cyclase activities (Zhang et al., 2010).

Another intriguing enzyme, AONS from *Thermus thermophilus* has been characterised to show activities at high temperatures (50-90°C) as well as ability to catalyse several other AOS-related reactions (Kubota et al., 2007). Apart from the AONS activity, the protein accepts glycine and serine as substrates and can catalyse KBL and SPT-like reactions. Such inherent catalytic promiscuity suggests that this thermophile bacterium might have adopted a single protein to be used in several biosynthetic pathways (Kubota et al., 2012).

Defining the exact features that determine the substrate specificities of the AOS enzymes and manipulating them can provide vast tools for broadening the enzyme catalysis.

1.14. Aims

This work aimed to study the structural and mechanistic properties of SPT enzymes from *S.paucimobilis* and *B.fragilis.* In particular, the aims were:

- To assess the possible mechanisms of amino acid substrate selection in *Sp*SPT using site-directed mutagenesis combined with the amino acid binding and activity studies.
- To express, purify and characterise the SPT from *B.fragilis* and compare it to other previously characterised SPTs. To see whether this microbiome-related SPT possesses unique mechanistic features relating to the production of iso-branched SLs.
- To perform a structural characterisation of *Bf*SPT using X-ray crystallography and compare the structural features of the enzyme with other reported SPTs.
- To assess the substrate selectivity of *Bf*SPT utilising a range of different chain length acyl-coenzyme substrates and to identify the reaction products.
- To study some general mechanistic features of SPT using inhibitors such as D-serine, penicillamine and L-cycloserine to reveal aspects of binding and inhibition modes of these compounds.

Chapter 2: Materials and methods

2.1. Materials

2.1.1. Constructs and plasmids

Genes encoding *Sp*SPT (C-terminal, C-terminal A295T and N-terminal His-tag versions) and *Bf*SPT (C-terminal His-tagged) were previously cloned into the pET28a vector and stored as plasmid stocks at -20°C.

Most of the previous work on *Sp*SPT was done using a C-terminal His-tagged construct, informally referred to as "Scottish SPT" in our lab, yet some complications with crystallisation arose. The crystallography work was done using the new "American SPT" construct (N-terminal His-tag version with additional 13 aa linker region) provided by our collaborator Teresa Dunn from Bethesda, USA. In this report, the constructs are referred to as Sc *Sp*SPT and Am *Sp*SPT respectfully. Sequences of both constructs are provided in Appendix 1.

2.1.2. Reagents

All the reagents were purchased from Sigma-Aldrich, Biorad or Ficher Scientific unless otherwise specified. Palmitoyl-CoA, pentadecenoyl-CoA and heptadecanoyl-CoA were purchased from Avanti Lipids. Other acyl-CoAs (myristoyl-CoA and stearoyl-CoA) were purchased from Sigma Aldrich.

2.1.3. Bacterial strains used

E.coli DH5α (Invitrogen^{TM}) was used as high plasmid copy number strain during molecular cloning applications. *E.coli* C2987 (New England Biolabs) was used for transformation and vector ligation of PCR products during the site-directed mutagenesis.

Protein expression was performed in *E.coli* BL21 DE3 (Novagen^{TM}) and *E.coli* BL21 Gold (Agilent^{TM}) strains.

2.1.4. List of general buffer solutions used and their abbreviations throughout the text

Buffer name/abbreviation	Mixture
Lysis/Nickel purification buffer (**LysB**)	20 mM KPhos pH 7.5 150 mM NaCl 10 mM imidazole 250 µM PLP*
Elution buffer (**EB**)	20 mM KPhos pH 7.5 150 mM NaCl 300 mM imidazole
Crystallography buffer (**X-trial**)	10 mM Tris pH 7.5 150 mM NaCl 250 µM PLP*
Gel filtration/dialysis buffer (**GF**)	20 mM KPhos pH 7.5 150 mM NaCl 250 µM PLP*

Table 2.1. List of the standard buffer solutions and their abbreviations used later in the text. *Freshly made aqueous PLP solution was added to the buffer directly before use. A composition of other buffer solutions is specified later in the text.

2.1.5. Growth and agar media

Bacterial cultures were grown in Lysogeny Broth (LB) media containing 10 grams per litre tryptone, 5 grams per litre yeast extract and 10 grams per litre NaCl.

The media were prepared using deionised water; the pH of solutions was brought to 7.5. The media were sterilised by autoclaving.

LB agar was purchased from Sigma-Aldrich; upon dissolving in water it was sterilised by autoclaving. During preparation of the agar antibiotic solutions were added when the agar cooled down to approximately 50°C; the agar was spread on Petri dishes that have been stored at 4°C and used within one week.

2.1.6. List of the primers used in site-directed mutagenesis reactions

The primers were designed according to the method by Liu & Naismith, 2008. To increase the annealing efficiency of *Bf*SPT C357del mutation, extended primers were designed. The primer oligonucleotides were ordered from Sigma Aldrich and stored as 10µM stocks at -20 °C.

***Sp*SPT A295T forward**

5'- CACCACCTCGCTGCCGCCCTC- 3'

***Sp*SPT A295T reverse**

5'- GCGAGGTGGTGAAGATGTACGGAC- 3'

***Bf*SPT C357del forward**

5'- CATCAACGTATCGTTCGGAGA[]TGCGGGAGGCACAAC- 3'

***Bf*SPT C357del reverse**

5'- ATCAACGTATCGTTCGGAGA[]TGCGGGAGGCACAAC- 3'

2.2. Methods

2.2.1 DNA manipulation

2.2.1.1 Site-directed mutagenesis

The site-directed mutagenesis (SDM) was based on the method reported by Liu & Naismith, 2008. The following mixture was used: 5 µl x10 polymerase enzyme buffer, 2.5 µl forward and 2.5 µl reverse primer (to the final concentration of 0.5 µM), 2.0 µl DNA template, 1 µl dNTP mix, 1 µl (2.5 U) polymerase; the mixture made up to total 50 µl by adding pure H_2O. Due to lack of initial success, three different DNA polymerases were used: Pfu Turbo Hotstart, Phu Phusion and Hi fidelity polymerase.

The general PCR reaction cycle was set up as:

1. Initial denaturation 95°C – 30 sec
2. (Denaturation 95°C – 30 sec; annealing 62°C – 30 sec; extension 72°C – 10 min) x 30 times
3. Final extension 72°C – 10 min

To increase the efficiency of SDM, parameters such as annealing temperature or extension time have been modified.

Upon completion of the reaction, a digest of methylated parental DNA strand was performed by adding 1 µl DpnI (10 U) enzyme and 6 µl of CutSmart buffer solution. The sample digestion time was increased from 1 to 5 hours at 37°C to increase the digestion quality. The resulting DNA products were visualised on the analytical scale using 1% agarose DNA gel; upon identification of PCR products, the multiplication band samples were used to transform *E.coli* CH5α cells.

2.2.1.2. Bacterial cell transformation

2 µl of plasmid DNA was added to a single aliquot of *E.coli* DH5α (cloning stages) or BL21 cells (for protein expression). The mixtures were incubated for 25 min at 4°C,

followed by a 40 sec heat shock at 42 °C. The cultures were returned to 4°C for another 5 min before addition of 100 µl SOC media. After incubation at 37 °C with shaking for 1 hour, cultures were used to seed the lysogeny broth (LB) agar plates containing an appropriate antibiotic (30 mg/ml kanamycin). The plates were incubated at 37°C for 18 hours.

2.2.1.3. Purification of plasmid DNA using Qiagen™ DNA miniprep kit

Upon successful transformation resulting in the growth of bacterial cell cultures, DNA miniprep procedure was performed to isolate the desired plasmid from the positive colonies. Several bacterial colonies were picked up and used to inoculate 5 ml of LB samples overnight in the presence of the antibiotic (30 mg/ml kanamycin). The overnight cultures were centrifugated for 10 min at 3000 rpm. The insoluble pellet was treated according to the Qiagen™ protocol, first by resuspending it in the presence of RNAse, this was followed by the alkaline lysis of the cells and neutralisation step. Cell debris was removed by centrifugation for 10 min at 13K rpm, the DNA containing soluble fraction was applied to the miniprep spin column. Upon several wash steps and removal of the residual buffer, the plasmid DNA was eluted using into a sterile Eppendorf tube. The purified plasmid DNA was used for sequencing to confirm the presence of the desired gene, a transformation of new bacterial cells or long-term storage at -20 °C.

2.2.1.4. Sequencing reaction

DNA sequencing reaction was performed using 5 µl of template DNA, 1 µl of either forward or reverse sequencing pET28a primer (10 µM stock), 2 µl of sequencing terminator buffer x5 and 2 µl of BigDye™ mixture. The sequencing reaction was programmed as 24 cycles: 94 ˚C for 30 seconds, 50 ˚C for 20 seconds and 60 ˚C for 4 minutes. Samples were analysed by GenePool, Edinburgh Genomics.

2.2.2. Protein expression and purification methods

2.2.2.1. Protein expression and cell harvesting

Upon successful transformation, single plasmid-containing BL21 DE3 colonies were used to set up large-scale protein expression. The individual colonies were used to inoculate 250 ml of LB medium in the presence of 30 mg/ml kanamycin; the cultures were incubated overnight at 37 ˚C shaking. These cultures were used as the stocks to inoculate larger 500 ml LB media containing the antibiotic. The usual growing volume per protein preparation was 3 l (6 flasks of 500 ml). Upon reaching the exponential growth phase at 37°C (OD_{600} = 0.6), the protein expression was induced by adding the isopropyl β-D-1-thiogalactopyranoside (IPTG) to the final concentration of 100 μM (0.1 mM). The protein expression was performed for 4 hours at 30 ˚C shaking. After this time the bacterial cell pellets were precipitated by centrifugation at 3000 rpm for 20 min at 4˚C. The supernatant was discarded while the pellets were resuspended in a small amount of phosphate buffered saline (PBS) buffer, followed by further centrifugation at 3000 rpm for 10 min at 4˚C. Upon discarding the supernatant, dry bacterial pellets containing the protein of interest were stored at -20˚C for further protein purification.

2.2.2.2. Solubilisation of protein content

Single bacterial cell pellets (~2g) were defrosted and resuspended in ~20 ml of LysB buffer at 4˚C with 1 protease inhibitor tablet (Roche) and 250 μl of PLP added. Upon resuspension, the cells were lysed by ultrasonication using a Soniprep 150 while being kept on ice; cell lysis was performed by sonication at moderate intensity for 30 sec followed by 30 sec of a cooling down period, this was repeated 15 times. Upon cell lysis, the soluble protein was separated from the cell debris by ultracentrifugation at 20k rpm, 4˚C for 30 min using Sorval^{TM} Ultracentrifuge. Separate samples from all the stages were taken for analysis on SDS PAGE. Soluble protein was taken for the further purification step by Ni-affinity chromatography.

2.2.2.3. Purification of protein by Ni-affinity chromatography

Soluble protein fraction was incubated with ~2 g of Qiagen™ Superflow Ni-NTA resin (Catalog number 30410) for 1 hour on a rotatory mixer. The resin was pre-equilibrated in LysB buffer containing 10 mM imidazole to reduce non-specific binding. Upon the incubation time, the sample was loaded onto a Qiagen™ disposable plastic column; the histidine-tagged protein was binding to the resin while unbound fractions were flown through the column. After another wash step with the LysB buffer, the protein was eluted using 5 ml of the EB buffer. The protein concentration was checked by measuring the absorbance at 280 nm by UV-vis scan. Samples from all stages were taken to be analysed by the SDS-PAGE. Eluted protein was taken to a further purification step by size exclusion chromatography.

2.2.2.4. Protein purification by size exclusion chromatography

The elution sample containing protein of interest from the previous Ni-affinity purification stage was passed through Millipore™ 0.45 µm to remove possible aggregates. The sample was loaded on the 120 ml GE Health Care Superdex S200 column that was pre-equilibrated with GF buffer containing 25 µM PLP. The protein was passed through the column, the protein-containing fractions were collected and analysed on SDS-PAGE; the protein concentration was checked by measuring the absorbance. Fractions containing the purified protein were frozen at -80°C with the addition of 20% glycerol.

When protein was purified for the purpose of setting crystal trials, the column was equilibrated with X-trial buffer instead; the protein elution fractions were stored on ice before being transported to the crystallography facility.

2.2.3. Protein characterisation methods

2.2.3.1. Protein analysis using polyacrylamide gel electrophoresis (SDS PAGE)

Protein samples were analysed under the denaturing conditions using SDS polyacrylamide gel electrophoresis. The protein sample was mixed with X2 loading buffer (3 mL of 0,5 M Tris, pH 6.8, 5 % w/v glycerol, 10 % w/v SDS, 0.4 mL β-mercaptoethanol and 0.05 % w/v bromophenol blue) and boiled for 5 min. It was then loaded onto 15% polyacrylamide gel and run for 50 min at 200V. The gel was visualised by staining in either CoomasieBlue stain (H_2O, 0.1 % w/v Coomassie brilliant blue R250, 40 % w/v methanol and 10 % w/v acetic acid) or Coomasie Quickstain™; images of gels were taken. The protein molecular mass was estimated by comparing the position of visible bands with the low molecular weight (LMW) markers (GE Healthcare).

2.2.3.2. Enzyme UV-vis spectroscopy

Before all spectroscopic studies, protein aliquots were dialysed in the GF buffer with 250 µM fresh PLP for about 1 hour to convert SPT into the holo form. The excess of PLP was then removed using a desalting PD-10 column (GE Healthcare).

Spectroscopy measurements were performed using Varian™ Cary UV-Vis spectrometer and 1ml quartz cuvette with 1cm path length. The protein absorbance spectra were normally recorded between 200 and 800 nm. The obtained data were processed using Microsoft Excel; the results were analysed and presented using Origin 9.0 software.

2.2.3.3. Determination of the enzyme concentration

The enzyme concentration was determined by measuring the protein absorbance peak at 280 nm and using the Beer-Lambert law:

$$c = \frac{A}{\varepsilon l}$$ (Equation 1),

where ***A*** is the absorbance reading at 280 nm, ***l*** is the light path length equal to the diameter of the cuvette (1 cm) and ***ε*** is the molar extinction coefficient. The molar extinction coefficient was determined for each protein using the ExPASy ProtParam tool (Gasteiger et al., 2005). It was estimated as 25900 $M^{-1}cm^{-1}$ for the WT *Sp*SPT and 23380 M^{-1} cm^{-1} for the WT *Bf*SPT.

2.2.3.4. Measuring enzyme dissociation constants for amino acids

*Sp*SPT and *Bf*SPT enzymes had dissociation constants (K_d) determined for several amino acids (L-serine, D-serine, L-alanine, glycine). A typical experiment for determination of a K_d of amino-acid binding was performed in the GF buffer (20 mM KPhos, pH 7.5, 150 mM NaCl) and had the initial concentration of SPT at 10 μM, the total volume of the mixture at the start was 1 ml. The instrument was blanked with the GF buffer and the initial protein spectrum was collected across the range of 260-500 nm. Small amounts of amino acid stock were gradually titrated into the mixture yielding the final amino acid concentration range of 0.1-100 mM (up to 160 mM for some AAs). Upon each addition, the SPT-AA mixture was left to equilibrate for 5 min and the protein spectrum was recorded afterwards. The titrations were made at 25°C.

Changes in the absorbance peak at 425 nm corresponding to formation of the ketoenamine peak were plotted over the increasing concentrations of each amino acid using Origin Software. The graphs were normalised based on their absorbance at 280 nm to account for a gradual dilution of the enzyme. The data were fitted by the hyperbolic saturation curve, K_d was obtained by the Origin software using the equation 2:

$$\Delta A_{obs} = \frac{\Delta A_{max}[\mathrm{AA}]}{K_d + [\mathrm{AA}]}, \text{ (Equation 2)}$$

where ΔA_{obs} and ΔA_{max} are the observed and maximum changes in the absorbance peak respectively and [AA] is the molar concentration of the amino acid. The experiments were done in triplicate unless otherwise stated; the error bars show the standard deviation within the samples.

2.2.3.5. Measuring enzyme activity using the DTNB assay

Kinetic activities of SPTs were observed using 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB) assay (developed by Ellman, 1959; the protocol from Raman et al., 2009). DTNB specifically reacts with free thiols by forming a covalent complex via a disulphide bond. This causes the release of free TNB^{2-} ions that have a maximum absorbance peak at 412 nm with $\varepsilon = 14140\ M^{-1}\ cm^{-1}$ (Riddles et al., 1983) (Fig.2.2.3.5).

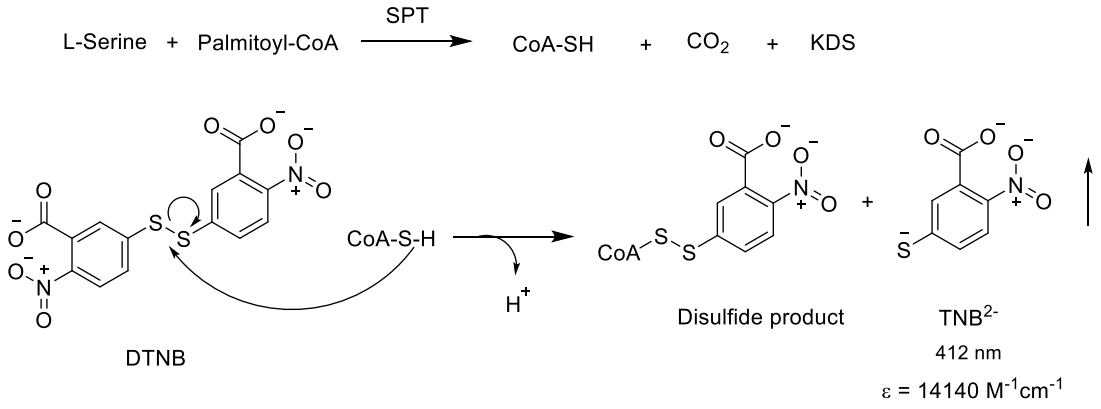


Fig. 2.2.3.5. Measuring the activity of SPT with DTNB assay. DTNB reacts with the thiol group on free CoASH, forming a disulphide product, this causes the release of coloured TNB^{2-} ion that is monitored spectrophotometrically.

The DTNB assay does not provide a direct measurement of the enzyme activity since the increase in absorbance at 412 nm is caused by the release of TNB^{2-} ion and not the enzymatic product itself. This was accounted for by taking the appropriate controls without the presence of an enzyme, CoA substrate and both for each experiment. The highest rate of absorbance increase obtained within these control readings was subtracted as a background when calculating the enzymatic rates; this value usually remained low and never exceeded 5% of the maximum observed reaction rate, remaining within the boundaries of experimental error. This allows assuming that the amount of TNB^{2-} released remains very close to the actual amount of the KDS (or KDS-like) product released by the SPT and can be used to establish reliable kinetic parameters for the reaction. For clarity, however, all reaction rates in the text and figures are presented in terms of the concentration of TNB^{2-} released per second.

2.2.3.6. Comparing relative activities of SpSPT WT and A295T mutant enzyme

The DTNB assay was used to compare the relative activities of the WT *Sp*SPT with the A295T mutant. The reactions were made up in 100 mM HEPES buffer of pH 8.0, 0.2 mM DTNB, 100 nM enzyme (WT or A295T) and 250 µM PCoA. Two sets of reactions were made with L-serine and L-alanine, using the excess concentrations of 40 mM for both amino acids. To account for the assay background, control reactions were set up for both WT and A295T mutant with no amino acids present. The reactions were made to a total volume of 150 µl with the deionised water. The reactions were monitored using Biotek Synergy HT plate reader on 96-well plates at 30°C.

The relative activities were presented in terms of comparing the initial rates that were determined from the absorbance increase at 412 nm for the first 5 min of reaction; the highest initial rate obtained for the reaction of WT *Sp*SPT with the L-serine was displayed as 100%, with the initial rates of A295T mutant presented as a % fraction of it. The relative activities of both WT and A295T enzymes with two amino acids were compared and presented as a bar chart using the Origin^{TM} software.

2.2.3.7. Determining the activity of SPT using DTNB assay

The activity of each *Bf*SPT reaction was monitored by DTNB assay on Biotek Synergy HT plate reader using 96-well plates, with the total volume for each reaction being 150 µl. As *B.fragilis* is a human commensal organism, all reactions were monitored at 37°C, which is close to the expected temperature conditions for *Bf*SPT *in vivo*. Experiments for determining the activities of *Bf*SPT with different substrates were performed individually for each substrate; each condition was set in triplicate; individual experiments were repeated up to three times with different enzyme samples to ensure consistency.

Each experimental condition for *Bf*SPT reaction was made in 100 mM KPhos buffer of pH 7.0 and contained 0.2 mM DTNB. The total reaction volume was brought to the final 150 µl using deionised water.

Experiments to determine the K_M of L-serine were made in the presence of 50 µM palmitoyl-CoA and 200 nM enzyme; the concentration range of L-serine tested was 0-40 mM. For determining the activities with acyl-CoA substrates experiments had an excess amount of 20 mM L-serine and 200 nM of the enzyme. The concentrations of acyl-CoAs were in the range of 0-250 µM. Due to the low observed reaction rates vs background, the activity of *Bf*SPT with myristoyl-CoA (C14:0) was performed at the enzyme concentrations of 200 nM and 400 nM; for valid comparison with other enzymes the data obtained at the enzyme concentration of 200 nM is presented in the text; however, presented kinetic parameters were calculated from the second set with at higher enzyme concentration of 400 nM.

The measured increase in absorbance at 412 nm was converted into the concentration of TNB^{2-} ion released using Beer-Lambert law (equation 1).

The mean initial reaction rates were plotted over the variable substrate concentrations [S]. The obtained saturation curves were fitted into the appropriate equations 2 or 3 from which the kinetic parameters were calculated using GraphPad™ software.

Kinetic parameters of K_M and V_{max} for *Bf*SPT reactions with L-serine, myristoyl- (C14:0) and stearoyl- (C18:0) CoAs were determined from the Michaelis Menten equation (3):

$$V = \frac{V_{max}\,[S]}{K_M + [S]} \quad \text{(Equation 3)},$$

where V is the observed initial rate expressed in nMs^{-1}, [S] is the variable substrate concentration.

*Bf*SPT reactions with pentadecenoyl- (C15:0), palmitoyl- (C16:0) and heptadecanoyl- (C17:0) CoAs showed a various presence of substrate inhibition. For these reactions, the modified equation for the substrate inhibition (4) was used to determine the K_M, V_{max} and K_i values:

$$V = \frac{V_{max}\,[S]}{K_M + [S]\left(1 + \frac{[S]}{K_i}\right)} \quad \text{(Equation 4)}$$

The enzyme turnover number k_{cat} was determined as V_{max} divided by the concentration of enzyme (200 nM, except for the reaction with myristoyl-CoA with 400 nM enzyme

as stated above). The substrate specificity index was presented as k_{cat}/K_M for each substrate.

2.2.3.8. Measuring the activity of *Bf*SPT upon time-dependent incubation with L-cycloserine

The DTNB assay was used to observe the inhibitory effect of L-cycloserine on the activity of *Bf*SPT. The reactions were set up in 96 well plates with the total volume of each reaction of 150 µl; the reactions were monitored using Biotek Synergy HT plate reader at 37°C.

The initial reaction mixtures contained the final 0.2mM DTNB buffered by 100 mM KPhos, pH 7.0; 20 mM L-serine and a range of concentrations of LCS (0-2.5 mM). The incubation with inhibitor was started by adding the 200 nM enzyme into the wells containing L-serine and LCS. Incubations were timed for the periods of 1, 5 and 30 min. The reactions were started upon addition of the PCoA into the wells to the final concentration of 25 µM. Notably, an imminent time lag between the start of reactions and launch of the data collection on the plate reader could not be avoided; this caused an increase in deviation, particularly affecting the samples with lower incubation times and concentrations of the inhibitor. The reactions were done in triplicate, the maximum averaged activity of *Bf*SPT in a control sample (no LCS) was presented as 100% with other relative activities calculated from it.

Cycloserine acts as a tight-binding inhibitor for PLP (Lowther, 2010), imposing that binding of the drug to the active site occurs more rapidly than the formation of the final inhibitory complex. In this case, the traditional Michaelis-Menten relation will not be valid as the concentration of free enzyme at steady state will not be equal to the concentration of free inhibitor. The Morrison equation (5) was designed for the cases of tight inhibitor binding and allows establishing of the apparent K_i, the constant used when the true inhibition constant K_i cannot be established:

$$V = V_0 - V_0 \left(\frac{(K_i + [E] + [I]) - \sqrt{(K_i + [E] + [I])^2 - 4[E][I]}}{2[E]} \right) \text{ (Equation 5)}$$

V – observed reaction rate; V_0 – maximum reaction rate observed without the presence of inhibitor; K_i – inhibition constant (apparent); E – concentration of enzyme; I – concentration of the inhibitor.

Separate values or the apparent K_i have been established for each incubation time; these do not account for the time factor directly but provided a convenient reference point for the working concentrations of LCS for further structural studies.

2.2.4. Biophysical methods of protein characterisation

2.2.4.1. Mass spectrometry

Native mass spectrometry of *Bf*SPT-L-penicillamine complex

*Bf*SPT sample was treated with 5 mM L-penicillamine for 30 min. The sample along with the enzyme control (no L-pen treatment) was buffer exchanged into 100 mM ammonium acetate using a PD-10 column (GE Healthcare). The final concentration of protein was 20.9 µM or 10.5 µM per monomer.

Samples were ionised using nano-electrospray ionisation (nano-ESI) using TriVersa NanoMate (Advion) and analysed on Synapt G2 Q-ToF (Waters) at 80°C, the pressure at 5 mbar and the sampling cone voltage at 150V. The spectra were averaged over 4 min of acquisition data and presented without processing. The m/z annotations were extracted using MassLynx V4 software (Waters).

MALDI-time of flight spectrometry

Detection of *Bf*SPT KDS-like products of *Bf*SPT reactions with 5 acyl-CoAs (C14, C15, C16, C17, C18) was performed using positive ion MALDI ToF spectrometry.

The reactions were made in 100 mM HEPES, pH 7.0, 2 µM DTNB, 40 mM L-serine and 50µM of the corresponding acyl-CoA substrate. For each acyl-CoA substrate tested 2 reactions were prepared, one of which had 200 nM of enzyme added, while other served as the negative non-enzyme control. Before mixing with the matrix, both positive and control reactions were checked for the activity monitored by the DTNB assay at 37°C. Upon detection of enzymatic activity, the reactions were further left for 1 hour at 37°C to incubate.

The matrix was made of 20 mg/ml α-cyano-4 hydroxycinnamic acid (CHCA) solution prepared in 50% acetonitrile, 40% methanol and 0.1% trifluoroacetic acid (w/v). 1µl of each reaction sample was mixed with 1µl of matrix solution; samples were transferred on a MALDI sample plate and allowed to dry.

The mass spectra were obtained using Bruker Ultraflex III equipment. The sample acquisition was performed over m/z range of 300-500 with the ion count range set as

2000-100 000. The peak analysis was performed using the Bruker Flex Software; the peaks corresponding to the positive +1 ions of the expected KDS-like products were identified. The obtained mass spectra were plotted as intensity versus m/z and presented using the GraphPad™ software.

2.2.4.2. X-ray crystallography

2.2.4.2.1. *S.paucimobilis* N-terminal SPT trials

The N-terminal *Sp*SPT was expressed and purified using Qiagen™ Ni-NTA superflow resin following the standard protocol. As it was not suitable to have phosphates in the crystallisation trials due to a high amount of salt crystals, the protein was washed into a different X-trial buffer containing 20 mM TRIS pH 7.5, 150 mM NaCl and 250 µM PLP during the size-exclusion chromatography stage. Upon purification, the protein samples were concentrated to 25 mg/ml using VivaSpin™ concentrator tube with 45 kDa cut off.

Prior to setting up the trials, the sample was centrifuged at 14 krpm, 4°C for 15 min to remove possible aggregates. Crystal trials were set by the sitting drop method in 96 well plates using ARI Crystal Gryphon robot; the plates were stored at 20°C. The final concentration of the protein was 20 mg/ml, the protein to precipitant ratio being 1:1 and 2:1. For co-crystallisation with D-serine a 20 mM concentration of the amino acid was used. Several commercial screens were set up, namely PEG I, PEG II, Wizard, JCSG+. Crystals appeared from as little as two days, they were monitored using Rigaku Minstrel™ imager and tested within two weeks.

Crystals suitable for testing were transferred to cryoprotectant solution (same mother liquor with 20% glycerol) and frozen in liquid nitrogen. The best quality crystal was obtained at 0.1M Sodium HEPES pH 7.5, 25% w/v PEG 4000 in the presence of 20 mM D-serine and gave the diffraction to a maximum resolution of 1.55 Å. Full diffraction dataset on the crystal was collected at the Diamond Light Source by a collaborator Dr Lucile Moynie, University of St Andrews.

2.2.4.2.2. *Bf*SPT SPT crystal trials

The protein was expressed and purified in the same way as the N-terminal *Sp*SPT into the X-trial buffer. Initial trials were set up automatically using ARI Crystal Gryphon robot; the commercial screens set up were PEG I, PEG II, Wizard, JCSG+; the screenings were made in 96 well plates using the sitting drop method.

Initial screens were performed with the protein concentration of 15, 20 and 25 mg/ml; the protein to precipitant ratios were set as 1:1 and 2:1. The best results were obtained at the concentrations of 15 and 20 mg/ml and the protein to precipitant ratio of 2:1.

The most suitable initial crystallisation condition was identified as 100 mM Tris pH 8.0, 20% w/v PEG 8K. These conditions were used to set up optimisation screens, with the pH range of 7.0-9.0 in 0.2 increments and the range of 10-26% w/v of PEG 8 K (2% increments). The follow-up screens were manually set up using the sitting drop method in 96 well plates.

Upon several rows of optimisation, the best reproducible crystallisation conditions were found to be around 100 mM Tris pH 8.5, 17% w/v PEG 8K, 150 mM NaCl, 12% glycerol. The presence of glycerol showed a stabilising effect on the protein by reducing unwanted precipitation and allowed to use the original crystallisation solutions for cryoprotection (Teng, 1990). Several other conditions gave good quality crystals; the exact crystal growth conditions are specified for each of the structures obtained. The crystals were frozen and tested for initial diffraction, with the ones diffracting to the resolution of 3 Å and higher being processed for the collection of the full dataset.

The data collection was performed in collaboration with Dr Magnus Alphey. Datasets were collected either at the Diamond Light Source (L-serine and L-penicillamine) or using the in-house X-ray source at the University of St Andrews (L-cycloserine structure).

L-serine

Trials of the *Bf*SPT have been carried in the presence of 20 mM L-serine. The crystal grown in 100 mM Tris pH 7.7, 16% w/v PEG 8K, 150 mM NaCl, 250 µM PLP and

13% w/v glycerol has provided the best quality diffraction to the highest resolution of 2.45 Å. The full dataset was collected at the Diamond Light Source.

Penicillamine

*Bf*SPT co-crystallisation attempts with both 20 mM L-penicillamine and 20 mM D-penicillamine did not result in a formation of reasonable quality crystals. A soaking method was performed to obtain the *Bf*SPT structure in the complex with L-penicillamine. Protein crystals were placed in a small droplet of the original crystallisation solution with a single solid granule of L-penicillamine added. In the first 30 min of incubation time, the crystals have visibly changed colour from yellow to pale. The crystals were left overnight before being taken out of the well solution and frozen.

The exact soaking method was carried out with D-penicillamine. The soaking resulted in the visible changes of crystal colour, however, none of the resulting crystals did provide a diffraction to a resolution higher than ~3Å.

The highest quality crystal soaked with L-penicillamine has provided a diffraction to a maximum resolution of 2.39 Å. The crystal was grown in 100 mM Tris pH 7.7, 16% w/v PEG 8K, 150 mM NaCl, 250 µM PLP and 10% w/v glycerol. The full dataset was collected at the Diamond Light Source.

L-Cycloserine

The co-crystallisation attempts with 2.5 mM LCS did not yield formation of any crystals. The soaking method was performed in analogy to penicillamine studies using the *Bf*SPT crystals grown in 100 mM Tris pH 8.0, 18% w/v PEG 8K, 150 mM NaCl, 250 µM PLP and 12% w/v glycerol. A small solid granule of LCS was added to the crystallisation well with the crystals present. In about 2 hours the crystals have visibly changed colour from yellow to pale; the sample was left to incubate overnight before being taken out of the well solution and frozen. The most successful crystal provided diffraction up to a maximum resolution of 2.75 Å. The full dataset was collected using the in-house X-ray source at the University of St Andrews.

2.2.4.2.3. Structure solution and refinement

The obtained diffraction data were collected and processed automatically using iMOSFLM (Battye et al., 2011), with scaling performed by AIMLESS (Evans & Murshudov, 2013). The initial structures were solved by molecular replacement using PHASER (McCoy et al., 2007) and *S.paucimobilis* SPT (pdb code 2W8G) as a model for both *Sp* and *Bf* SPTs. The *Bf*SPT structures obtained by soaking with L-pen and LCS were solved after obtaining the structure of *Bf*SPT-L-serine external aldimine structure, using the later as the model for a molecular replacement. Ligand dictionaries were generated using PRODRG (Schüttelkopf & Van Aalten, 2004). Ligand fitting and manual refinements were performed in Coot (Emsley et al., 2010); automatic model refinements were done by REFMAC5 (Murshudov et al., 2011); the structure validation and model quality was checked using MOLPROBITY (Chen et al., 2010). The pdb validation reports were obtained using wwPDB server (Gore et al., 2017). Protein structures are displayed using PyMOL Molecular Graphics System, Version 1.8.2.2 Schrödinger, LLC.

2.2.4.2.4. Crystallographic data collection and refinement statistics

Protein	*B.fragilis* SPT	*B.fragilis* SPT	*B.fragilis* SPT	*S.paucimobilis* SPT
Ligand	PLP-L-serine	PLP-L-penicillamine	PLP-L-cycloserine adduct*	PLP-L-serine
Data collection				
Space group	P1 21 1	P1 21 1	P1 21 1	P1 21 1
Cell dimensions				
a, *b*, *c* (Å)	60.26, 148.12, 103.24	60.22, 146.66, 103,12	59.45, 145.35, 101.75	61.38, 103.89, 63.75
a, b, g (°)	90.00, 105.08, 90.00	90.0, 104.98, 90.00	90.00, 105.16, 90.00	90.00, 116.53, 90.00
Resolution (Å)[a]	82.70-2.45 (2.51-2.45)	82.37-2.39 (2.45-2.39)	28.20 – 2.75 (2.90-2.75)	54.92-1.55 (1.59-1.55)
R_{merge}	0.078 (0.547)	0.076 (0.518)	0.153 (0.689)	0.073 (0.754)
$I / \sigma(I)$	9.7 (2.3)	11.0 (2.7)	6.8 (1.3)	15.6 (2.5)
Completeness (%)	97.8 (92.9)	98.4(97.9)	96.5(95.2)	98.3(97.9)
Average redundancy	4.8 (4.5)	5.7 (4.9)	4.9 (4.8)	5.7(5.5)
Chains in A.U.	4	4	4	2
Refinement				
Resolution	82.70-2.45	82.37-2.39	28.20-2.75	54.92-1.55
No. of unique reflections	62701	67068	41751	103385
R_{work}/R_{free}	0.2112/0.2502	0.2075/0.2514	0.276/0.325*	0.1811/0.2098
No. atoms				
Protein	12235	12268	12260	6012
Ligand atoms	88	96	80*	44
Number of waters	171	189	4	285
B-factors (Å^2)				
Protein	59.25	62.88	49.66	20.42
Protein B-factor chain range	46.51(A)-74.65(D)	51.80 (A) -73.42(D)	37.36(A) - 61.07(D)	18.53(B)-22.30 (A)
Waters	43.98	48.88	16.55	21.21
Ligand	60.81	69.17	56.82*	20.4
Ligand B-factor range	54.36(A)-69.4(D)	58.2(A) - 78.8 (B)	50.0 (A) - 65.0 (D)*	19.70 (B) -21.10 (A)
R.M.S. deviations				
Bond lengths (Å)	0.0115	0.012	0.01*	0.018
Bond angles (°)	1.3701	1.434	1.39*	1.8556
Ramachandran plot summ. (%)[b]	97.2/2.8/0	98.3/1.7/0	96.6/4.4/0	98.0/2.0/0

* The presented refinement parameters are for the proposed PLP-decarboxyoxime complex. Discussion of other possible adducts of L-cycloserine are presented in chapter 3.3.6.

[a] Numbers in brackets represent the highest resolution shell

[b] % of residues in most favoured regions/allowed regions/unfavoured regions

2.2.5. Bioinformatic analysis of conservation amongst putative isoSPTs

Search for putative isoSPT sequences was performed using protein BLAST (Altschul et al., 1990). The amino acid sequence of *Bf*SPT was used as an input and searched independently against the sequence databases of *Parabacteroides, Prevotella, Tannerella* and *Porphyromonas* taxa, with the top hit results selected for each.

For the top hit results, the identities of the corresponding genes and proteins were obtained from GeneBank (Benson et al., 2005) and UniProt (Bateman et al., 2017) respectively. The sequence alignments and phylogenetic trees were made using Clustal Omega (Sievers et al., 2011). The aligned sequences were presented using ESPript 3.0 (Robert & Gouet, 2014).

The structural models displaying sequence conservation were made by Consurf (Ashkenazy et al., 2010) using the sequence and the pdb coordinates of the *Bf*SPT L-serine external aldimine structure as an input template. The models were displayed using PyMOL Molecular Graphics System, Version 1.8.2.2 Schrödinger, LLC.

Generation of the proposed PLP-isoKDS external aldimine model was made using PRODRG (Schüttelkopf & Van Aalten, 2004). Fitting of the PLP-isoKDS molecule into the conservation model was manually performed in Coot (Emsley et al., 2010) using the structure of the PLP-decarboxymyriocin complex (4BMK) as the initial template. The modelling of the 17-carbon chain of the ligand into the possible substrate-binding tunnel of *Bf*SPT was manually performed in Coot using the local geometric restraints provided by the protein backbone.

Chapter 3. Results and Discussion

3.1. *Sphingomonas paucimobilis* SPT

The initial approach of this project was to continue studies on SPT from *Sphingomonas paucimobilis*. The enzyme has been established as a reliable model in our research group and has been extensively characterised by the previous group members (Yard, Raman, Lowther, Wadsworth, Beattie). The aim was to expand the knowledge of certain mechanisms of enzyme inhibition further as well as to study the substrate specificity of the enzyme and the possible mechanisms leading to the formation of deoxysphingolipids.

The enzyme was previously studied in the form of two constructs derived from specific cloning sites on the pET28a vector. The original construct contained a poly-histidine tag after the C-terminal (appendix 1); most of the previous work has been carried on it. Due to some undetermined difficulties with crystallisation of the protein arising after 2012, the further structural work has been carried on another construct containing N-terminal His tag joined by 13 additional linker amino acids (appendix 1). This second construct, provided by Teresa Dunn, has been codon-optimised for expression in *Arabidopsis thaliana;* this also led to some implications for designing the primers for SDM. The protein construct has shown a slightly reduced binding affinity for L-serine and a kinetic activity lower for about two times, that was most likely caused by the presence of the flexible linker tag. Colloquially the first construct was named as the "Scottish" (Sc) SPT and the former as the "American" (Am) SPT.

Due to the differences in the activities of both constructs, all spectrophotometric studies were carried out with Sc SPT, while Am SPT was used for some structural studies. Despite some differences in mass (45975 Da for Sc and 47232 Am) and kinetic activity, both constructs have shown generally similar parameters and were treated using the same protocol.

As this project has changed its direction quickly and not much significant work has been carried on both enzymes, most of the data presented here were obtained using ScSPT, while some of the work carried on AmSPT has been left off the scope of this report. Therefore further here the SPT from *S.paucimobilis* is referred to as *Sp*SPT (indicating the C-terminal Sc construct unless stated otherwise).

3.1.1. Expression and purification of *Sp*SPT (C- and N-terminal)

Both C- and N-terminal SPTs have been expressed in pET28a vector and introduced into *E.coli* BL21 cells using kanamycin resistance gene for selection of the positive colonies. Upon the expression, the protein was readily purified using metal affinity chromatography, followed by purification with size-exclusion chromatography (Fig. 3.1.1.1).

The purified SPT appeared yellow in solution, indicating the presence of PLP as its cofactor. Characterisation of fractions collected over the course of protein purification by SDS-PAGE revealed the presence of a strong band roughly corresponding to 45.0 kDa (with the theoretical mass of *Sp*SPT being 45975.6 Da). Upon the second purification by the SEC, the protein has appeared as a broad single peak with the retention volume of 68.5ml on the Superdex^{TM} S200 column; this corresponds to an estimated MW of roughly 92.0 kDa (the calibration curve and the calculation is shown in Appendix 3) indicating that the protein is present in the form of homodimer.

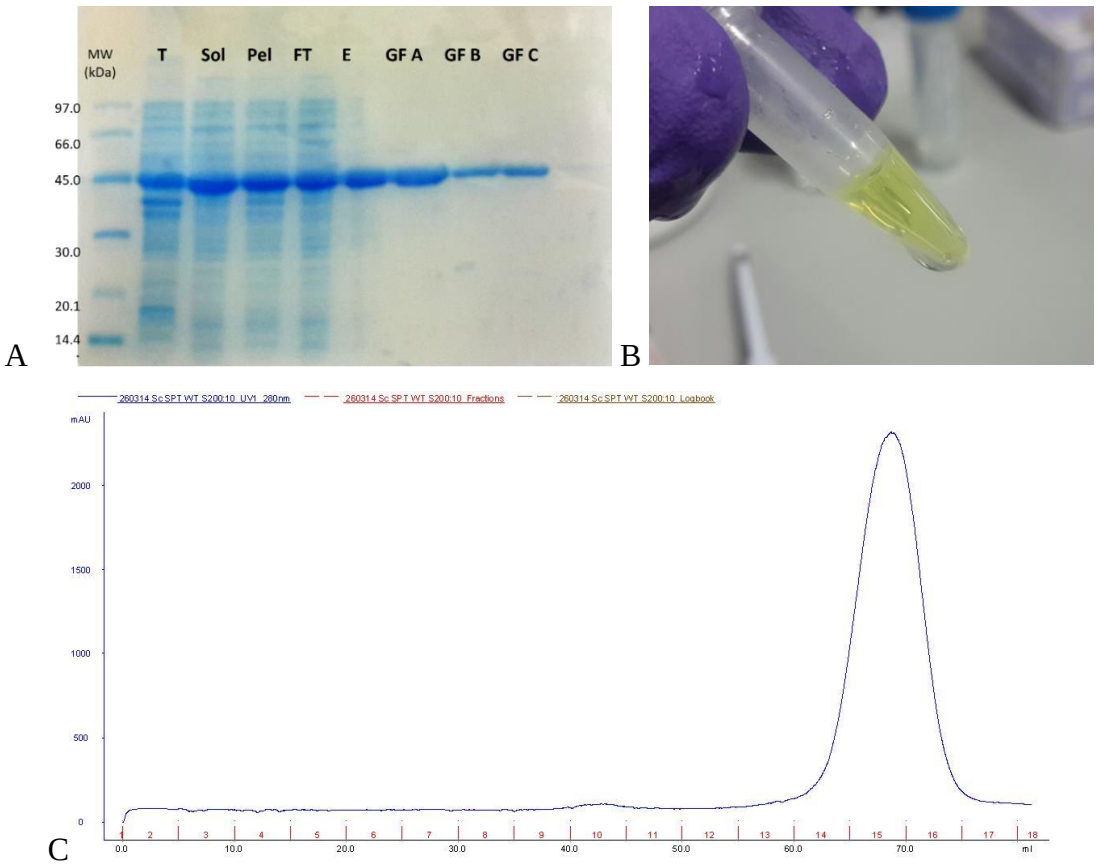


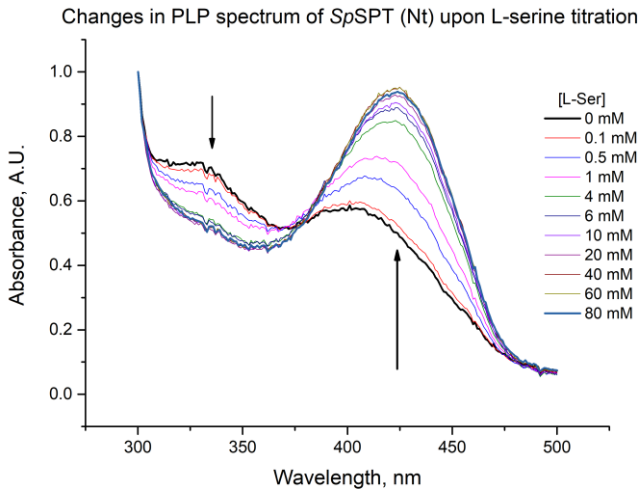
Fig. 3.1.1.1. A) The SDS-PAGE analysis of the fractions collected during the purification of *Sp*SPT. The fractions are marked as follows: T – total protein fraction obtained upon cell lysis; Sol – soluble fraction loaded on the Ni-affinity column; Pel – resuspended non-soluble fraction obtained from precipitated cell debris (discarded); FT – the flow through fraction during the Ni-affinity purification (discarded); E – elution fraction; GF A, B, C – peak fractions containing the purified protein collected after the size-exclusion chromatography. **B**) A photograph of the purified *Sp*SPT as it appears bright yellow in solution. **C**) The size-exclusion chromatogram of *Sp*SPT; the fractions 14, 15 and 16 contain the purified protein (GF A, B and C from the top gel). The column used is Superdex^{TM} S200 (GE Healthcare^{TM}).

3.1.2. Spectroscopic studies of *Sp*SPT

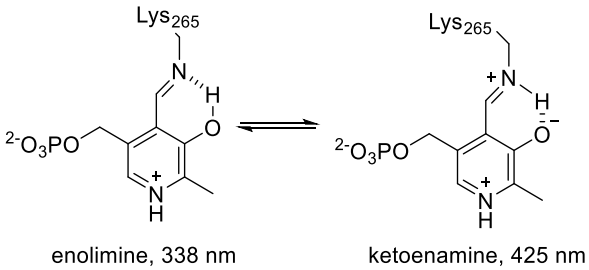
The SPT from *S.paucimobilis* has a characteristic UV-Vis profile of a PLP-dependent enzyme (Ikushiro et al., 2001; Raman et al., 2009). The PLP molecule is in equilibrium between the two resonance forms, enolimine and ketoenamine, having the absorbance maxima at 338 and 425 nm respectfully (Fig. 3.1.2, A). Upon formation of the external aldimine with L-serine a change in the equilibrium is observed as the enzyme predominantly forms the ketoenamine complex. It is possible to follow the formation of the external aldimine by monitoring the change in absorbance at 425 nm. This change is proportional to the increase in the substrate concentration and can be used for determining the enzyme-substrate dissociation constant K_d (Fig. 3.1.2, C).

The Sc SPT K_d (L-serine) was determined to be 1.1 mM; this value well agrees to the previously reported in the literature (Raman, 2009); the Am *Sp*SPT (N terminal) had its K_d (L-serine) determined as 1.34 mM, it is likely that the flexible 13-amino acid extension of the Am SPT might slow down the substrate binding and thus reduce the affinity.

A



B



C

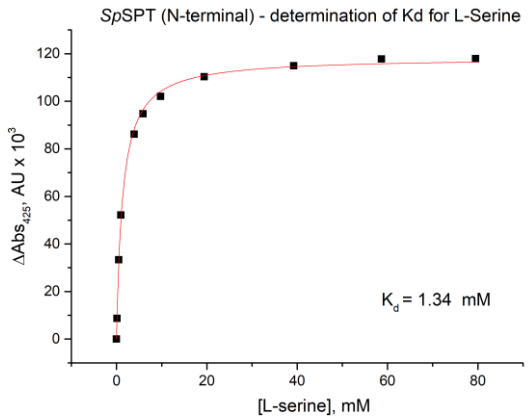


Fig 3.1.2. A). The PLP binding profile of *Sp*SPT (N-terminal)**.** The cofactor is present in an equilibrium between enolimine and ketoenamine forms (**B**). Addition of the substrate causes a shift in the equilibrium between these forms that can be monitored by the UV-vis spectrometry. **C**) Changes in absorbance of the ketoenamine peak at 425 nm are plotted against the increasing concentration of L-serine and used to determine the binding constant K_d

3.1.3. Exploring of the substrate specificity of the A295T mutant

Previous studies by A. Beattie have highlighted the role of the phosphate group interactions with SPT and its impact on the catalytic activity (Beattie et al., 2013a). In the external aldimine structure of *Sp*SPT (pdb 2W8J), the active site residues around the phosphate group of PLP form a network of hydrogen bonds (Fig. 3.1.3.1, A). Notably, a 2.5 Å length H-bond is formed with the hydroxyl group of the substrate L-serine. It was proposed that mutating the alanine 295 to threonine will introduce additional hydroxyl group from the protein backbone that would stabilise the phosphate and, at the same time, create a steric clash with the hydroxyl group of L-serine, thus reducing its binding affinity (Fig. 3.1.3.1, B). Both ALAS and AONS have a threonine residue present at the position corresponding to A295 in *Sp*SPT; thus, in theory, an introduction of the mutation might facilitate a conversion of L-alanine by SPT.

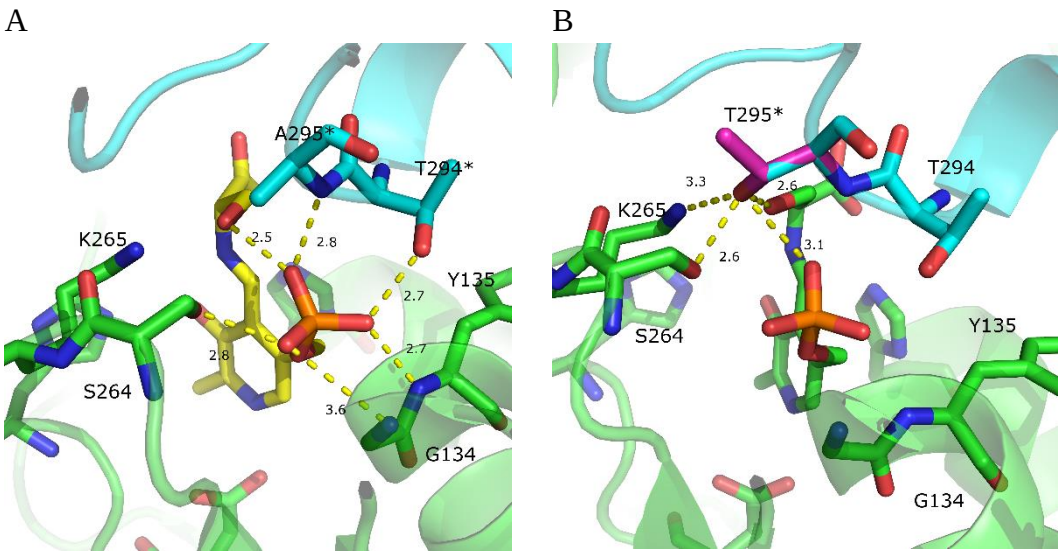


Fig. 3.1.3.1. Proposed effect of the A295T mutation on the substrate binding mode of ScSPT. **A)** The stabilisation of the phosphate group of the PLP-L-serine external aldimine by the backbone of *Sp*SPT (2W8J). The key contacts are presented with the numerical distances shown in Å; the residues from the opposite protein chain are marked with * **B)** A theoretical model of the previous structure showing a possible effect of the A295T mutation (marked in purple).

The C-terminal His-tagged A295T mutant was produced and spectroscopically characterised by A. Beattie (PhD thesis). This work has looked further at the properties

of the given mutant and its binding to amino acids glycine and alanine, as well as the comparison of reactivity with the WT enzyme.

The A295T mutant enzyme was expressed from pET28a plasmid in *E.coli* BL21 DE3 cells in the same manner as WT *Sp*SPT. The protein was purified with metal affinity and size exclusion chromatography; the purified protein has shown a single band of 45kDa on the SDS PAGE.

Comparison of amino acid binding in WT and A295T *Sp*SPT

To assess the effects of the mutation on the binding parameters of several amino acids, UV-vis binding assays were performed as described in methods. Binding of both WT *Sp*SPT and the A295T mutant were studied with L- and D-serine, glycine and L-alanine. The data are presented over the next pages in figures 3.1.3.2 (A-D); the obtained binding constants are summarised in table 3.1.3.1.

The A295T mutant shows different binding properties with L-serine and L-alanine as compared to the WT. It is possible to propose that introduction of the hydroxyl group from threonine 295 has weakened L-serine binding in about 13 times (K_d WT/mut = 1.1/14.7); it also resulted in a different spectral configuration upon L-alanine binding. Due to non-linear changes in the PLP spectrum, however, the determined K_d of WT enzyme lacks accuracy, while no K_d could be determined for the A295T mutant. The mutation, however, did not have any relevant effect on glycine binding as was expected (K_d WT/mut = 14.6/16.1), suggesting that other molecular interactions might be involved in regulation of the amino acid recognition. The mutation also had a minor effect on reducing D-serine binding affinity compared to the wild type (K_d WT/mut = 9.7/15.2), resulting in K_d's of both serine enantiomers being of similar values.

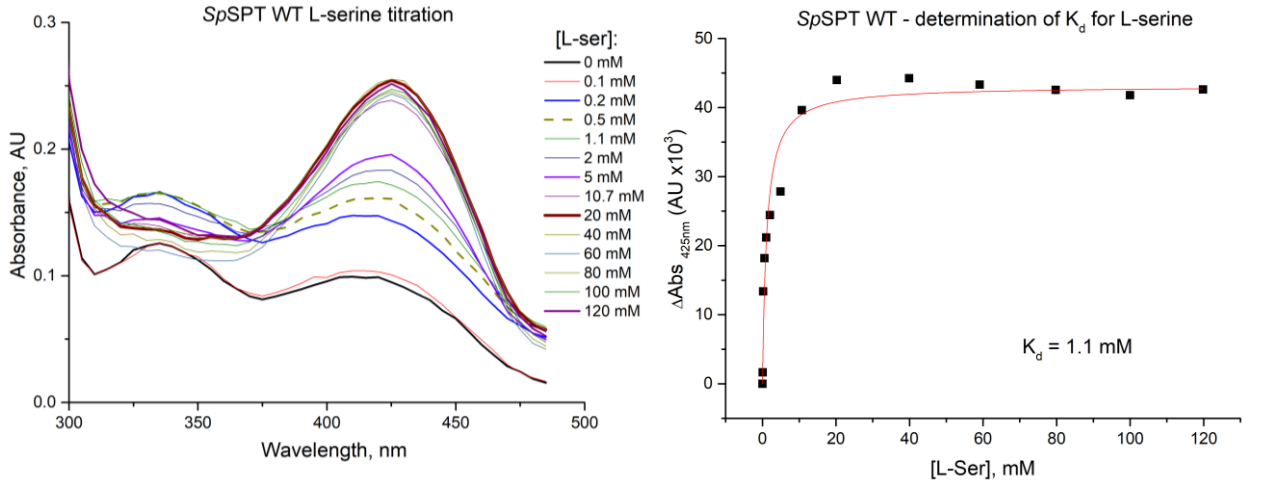
Enzyme	K_d (L-serine), mM	K_d (L-alanine), mM	K_d (glycine), mM	K_d (D-serine), mM
*Sp*SPT WT	1.1	12.3	14.6	9.7
*Sp*SPT A295T	14.7	-	16.1	15.2

Table 3.1.3.1. Comparison of the obtained amino acid binding constants of the WT and A295T mutant

Figures 3.1.3.2 (1-4). Comparison of amino acid binding in WT and A295T *Sp*SPT

L-serine

A



B

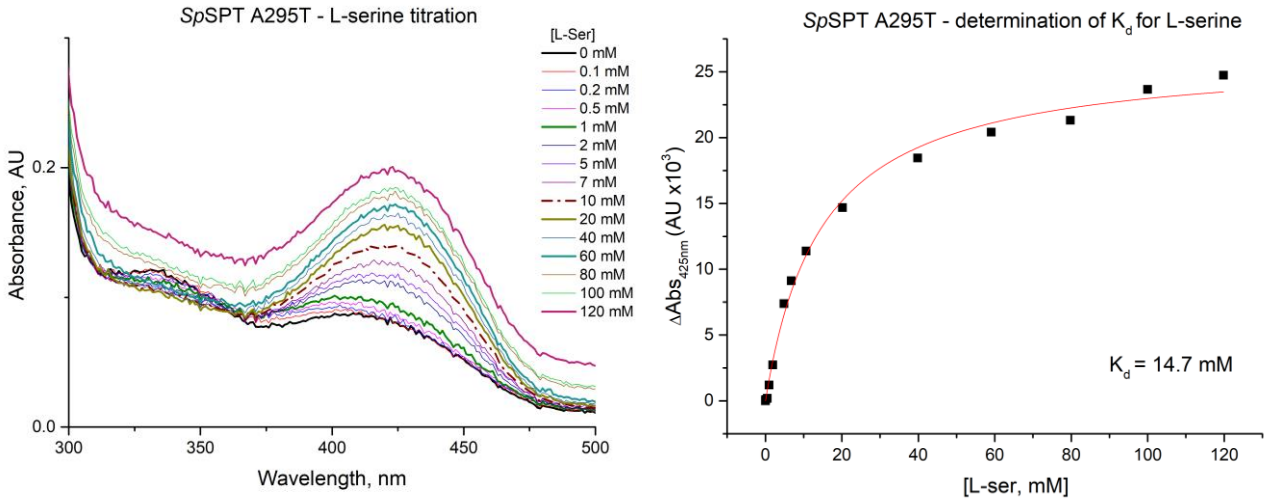
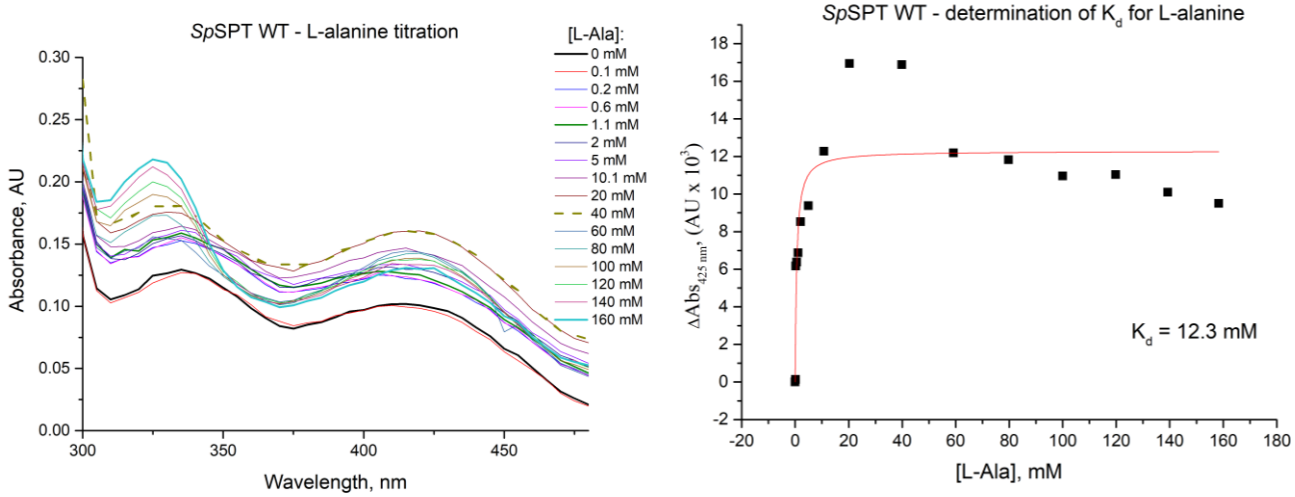


Figure 3.1.3.2 (A). Left: The UV-vis profile changes of *Sp*SPT WT (**A**) and the A295T mutant (**B**) upon binding of L-serine. **Right**: Changes in the absorbance at 425 nm peak plotted vs amino acid concentration used to determine the binding constants K_d for *Sp*SPT WT (**A**) and A295T mutant (**B**).

L-alanine

A



B

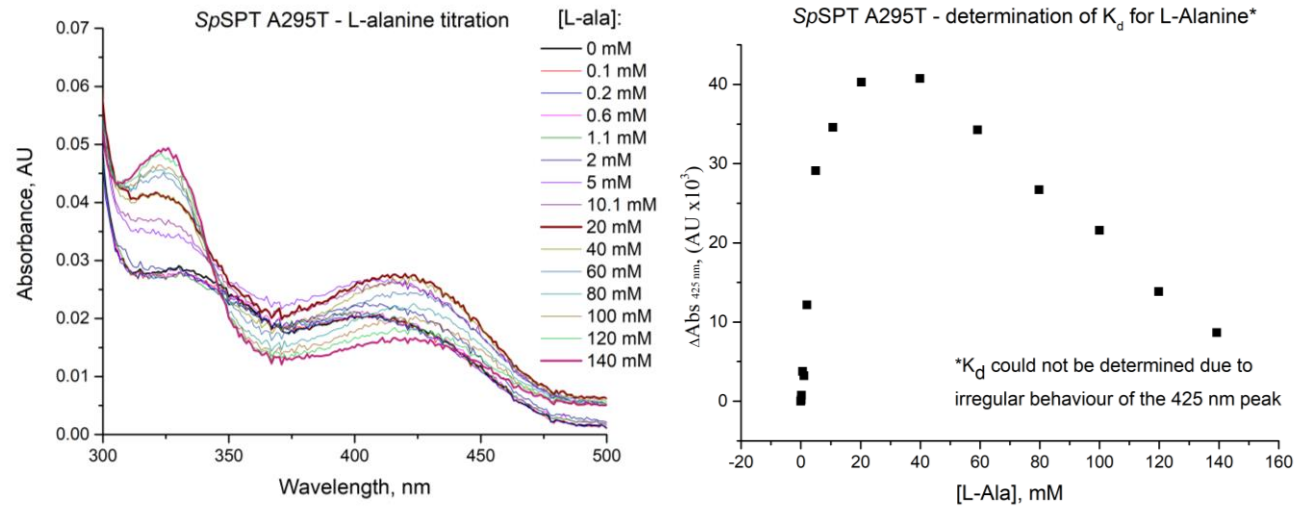
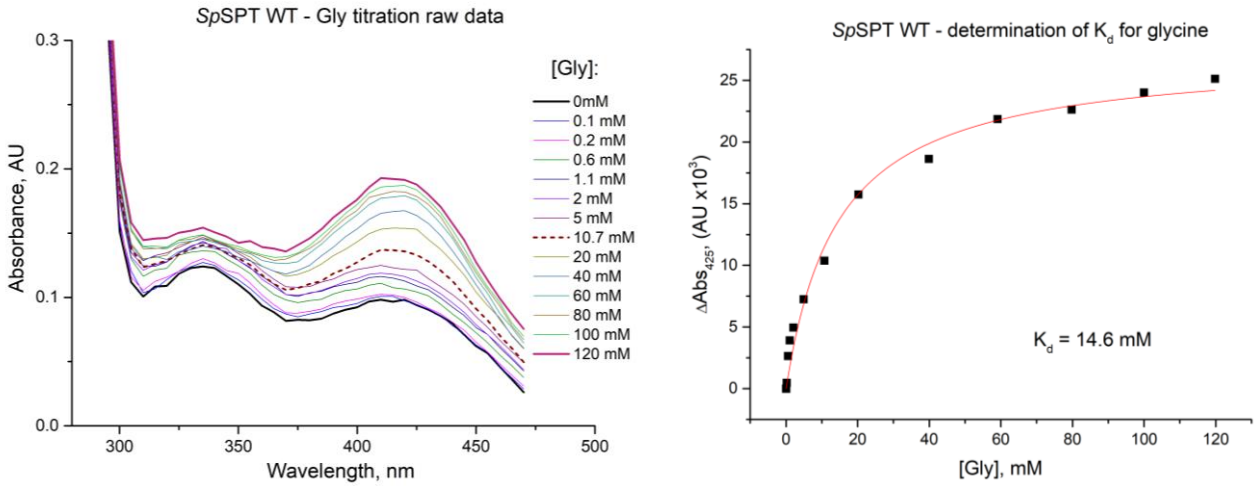


Figure 3.1.3.2 (B). Left: The UV-vis profile changes of *Sp*SPT WT (**A**) and the A295T mutant (**B**) upon binding of L-alanine. **Right**: Changes in the absorbance at 425 nm peak plotted vs amino acid concentration used to determine the binding constants K_d for *Sp*SPT WT (**A**) and A295T mutant (**B**). Due to non-linear change in absorbance of the A295T mutant it was not possible to determine the K_d.

Glycine

A



B

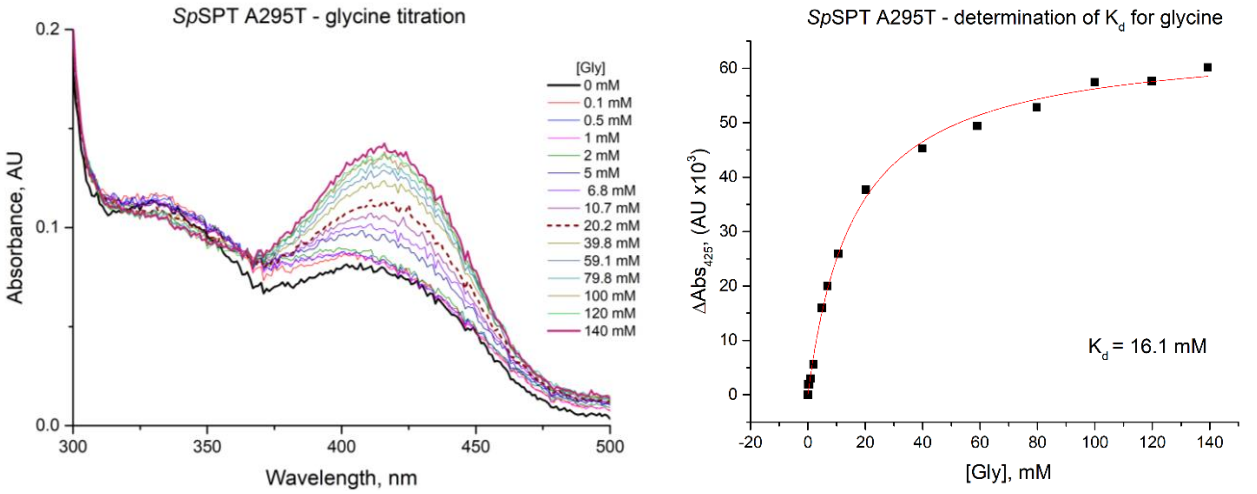
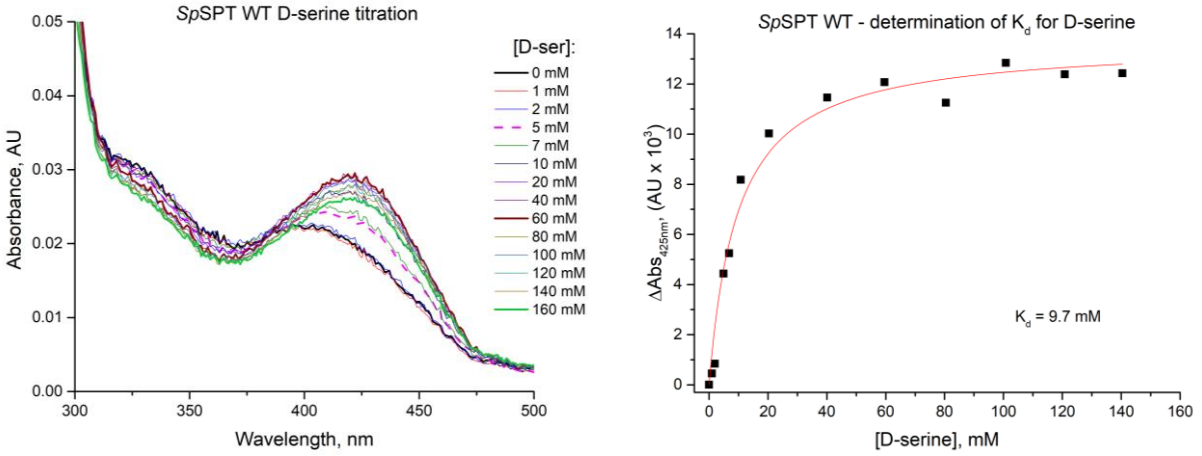


Figure 3.1.3.2 (C). Left: The UV-vis profile changes of *Sp*SPT WT (**A**) and the A295T mutant (**B**) upon binding of glycine. **Right**: Changes in the absorbance at 425 nm peak plotted vs amino acid concentration used to determine the binding constants K_d for *Sp*SPT WT (**A**) and A295T mutant (**B**).

D-serine

A



B

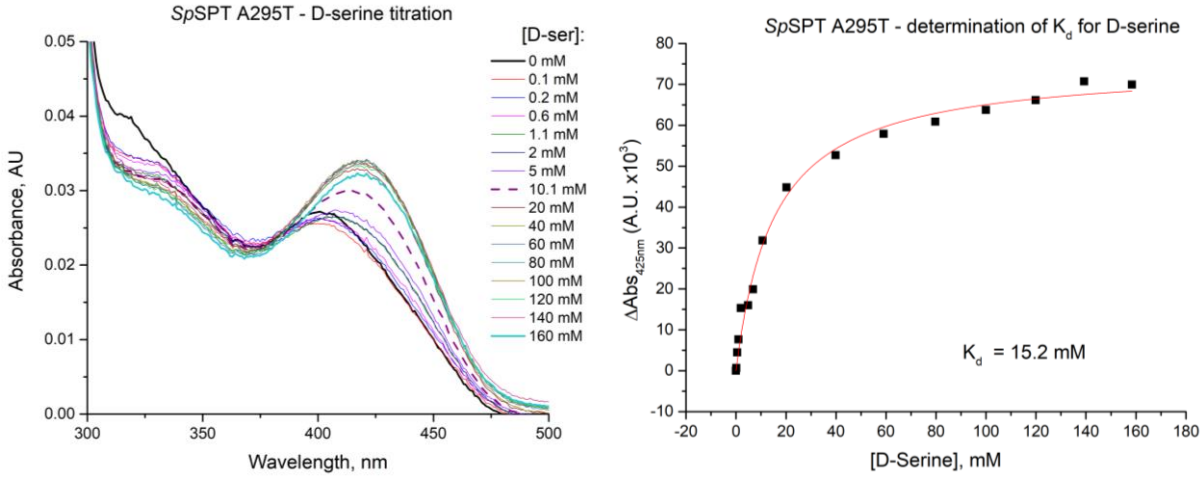


Figure 3.1.3.2 (D). Left: The UV-vis profile changes of *Sp*SPT WT (**A**) and the A295T mutant (**B**) upon binding of D-serine. **Right**: Changes in the absorbance at 425 nm peak plotted vs amino acid concentration used to determine the binding constants K_d for *Sp*SPT WT (**A**) and A295T mutant (**B**).

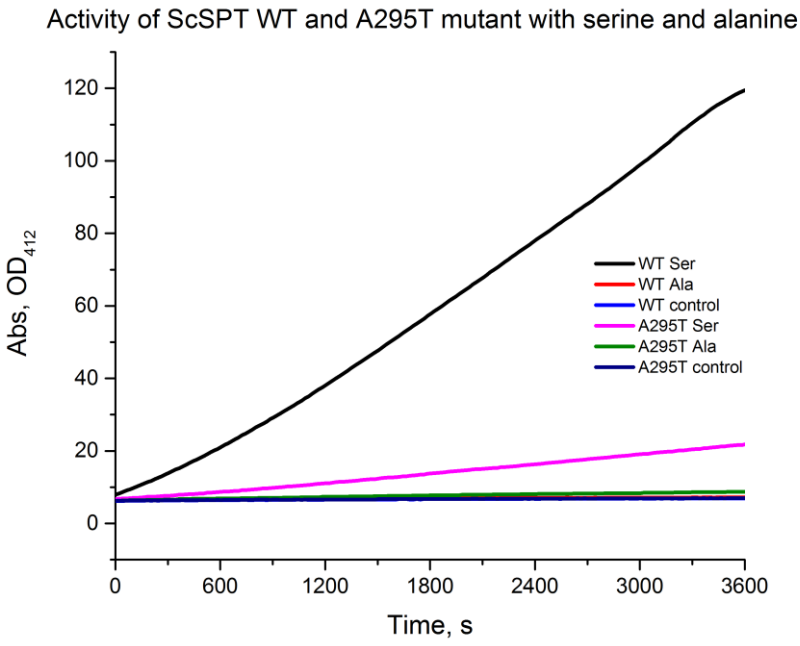
Observing relative activities of WT and A295T *Sp*SPT using DTNB assay

Due to a very low activity of the A295T mutant, it was not possible to obtain the kinetic parameters of the enzyme variant. Therefore, a simple comparative experiment was carried out to compare the relative observed activities of both WT and A295T enzymes in terms of the observed initial reaction rate using the DTNB assay. The observed initial reaction rate of the WT enzyme was presented as 100%, with the rates of the A295T mutant presented as percentage fractions of it. To see whether the mutation has led to any changes in the substrate specificity, separate reactions have been carried with the excess concentrations of 40 mM L-serine and 40 mM L-alanine for both enzyme variants. The enzyme concentration was kept at 100 nM, other reaction conditions were kept as described in the materials and methods section.

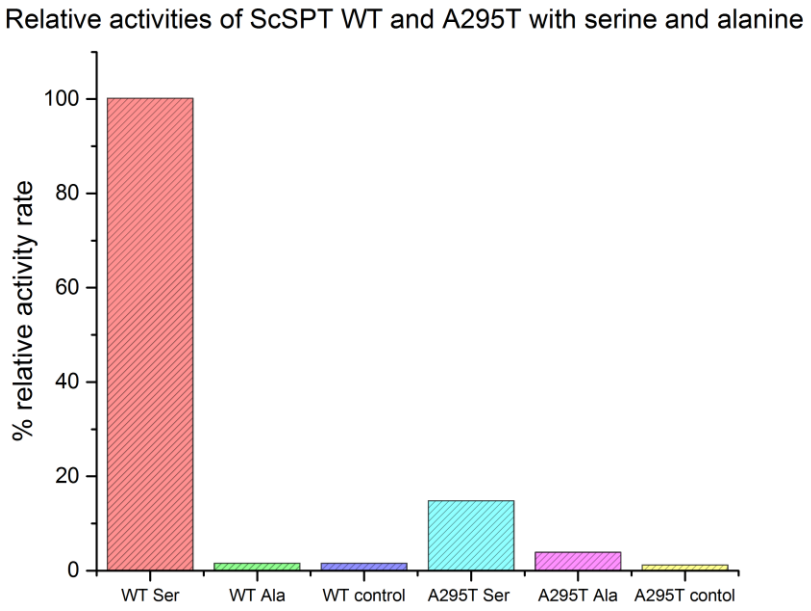
Relative activities of both WT and A295T mutant enzymes are shown in figure 3.1.3.3. The reaction rate of the mutant with L-serine shows only 13% of the WT activity. This correlates well with the predictions and binding studies as it is expected that threonine 295 introduces a steric clash in the active site, making binding of L-serine less favourable.

In contrast, a very slow activity was observed for the mutant with L-alanine. The WT *Sp*SPT is known as not able to process L-alanine as a substrate (Hanada et al., 2000c), Higher increase in absorbance at 412 nm has been observed for the A295T mutant in reaction with alanine compared to the background, giving a total of 0.48 nM of TNB^{2-} released per second.

The observed activity of A295T with alanine represents only 1.56% of the WT enzyme activity with serine; therefore, the specificity of the enzyme has not been subverted yet. While the mutant form still prefers serine as its substrate, the serine/alanine preference rate has decreased from > 600 in the WT to 8.3 in A295T.



A



B

Fig. 3.1.3.3. Activities of WT and A295T mutant SPT with 40 mM L-serine and 40 mM L-alanine measured by DTNB assay. **A**) Total enzyme activity observed in terms of the TNB^{2-} ion released over time. **B**) Relative activity presented in terms of the initial reaction rates with the initial reaction rate of WT SPT with L-serine presented as 100%. The control samples had no amino acid present.

3.1.4. Structural studies of *Sp*SPT (N-terminal): Co-crystallisation attempt with D-serine

Crystal trials of N-terminal *Sp*SPT were set up in the presence of 20 mM D-serine. The crystals grown in 0.1M Sodium HEPES pH 7.5, 25% w/v PEG 4000 gave the best diffraction data with the resolution of 1.55 Å.

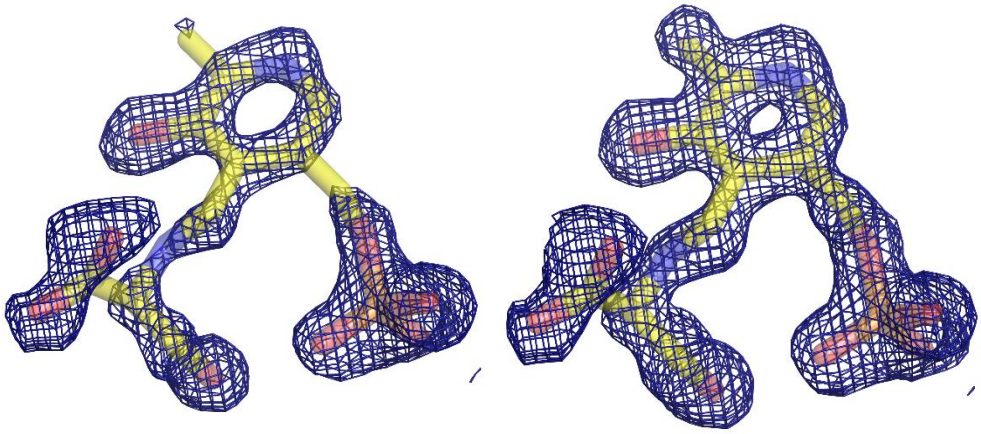
Resolving of the structure was carried according to the materials and methods; the structure of *Sp*SPT in complex with L-serine (pdb 2W8J) was used as the model for molecular replacement.

The obtained structure is generally comparable in its quality to the already published *Sp*SPT L-serine external aldimine structure (Raman et al., 2009; pdb code 2W8J) with the similar resolution (1.55 Å and 1.5 Å of the published structure), despite having slightly higher R_{work}/R_{free} values of 0.181/0.210 compared to 0.157/0.187 published.

However, unexpectedly upon solving the initial structure without the ligand, the unbiased 3.0 σ F_o-F_c map presented electron density that was corresponding to the L-serine external aldimine. No residual density that could indicate the shape of D-serine was observed in the active site (Fig. 3.1.4.1, A). To account for the possibility of a co-crystallisation error, the trials were repeated for the second time with 20 mM D-serine; the results were reproducible indicating the presence of L-serine in the active site.

D-serine powder provided by Sigma-Aldrich™ is marketed as 99% pure by the manufacturer, therefore it can contain up to 1% L-serine. As D-serine binding to SPT is ~9 times weaker compared to L-serine (K_d 9.8 mM and 1.1 mM respectively), selective binding of the natural substrate would be preferred. However, this could not entirely contribute to the presence of L-serine in the active site as its maximum concentration would be 0.25 mM. This is 4 times lower than the K_d for L-serine and would leave most of the enzyme unoccupied and available for D-serine; that would be reflected by the proportional split in the electron density but would not result in all of the bound substrate being L-serine. It is therefore likely that D-serine was converted into L-serine in the active site of SPT during the two-week crystallisation period; thus, SPT might have a weak serine racemase activity.

A



3.0 σ F_o-F_c map 1.2 σ $2F_o$-F_c map

B

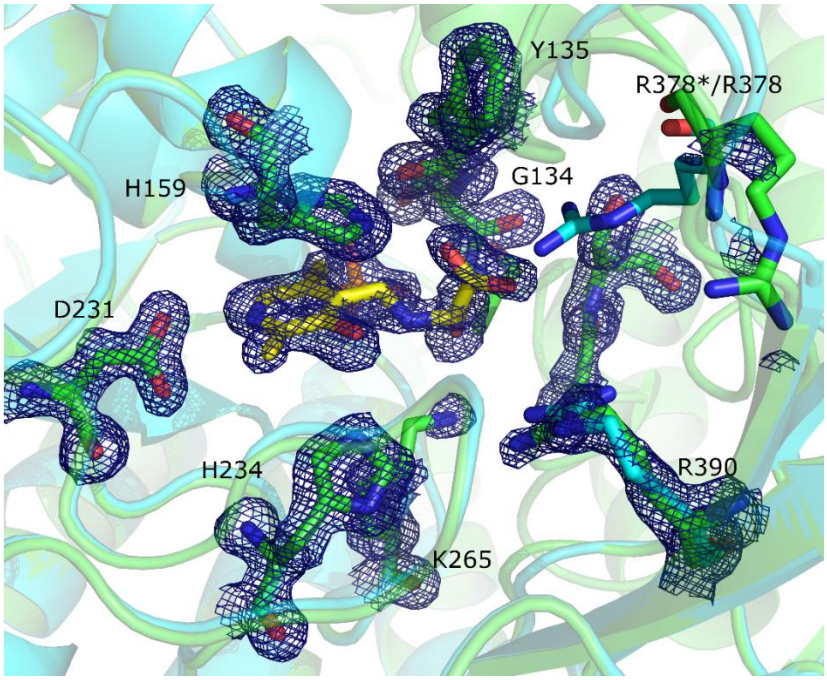


Fig. 3.1.4.1. A) The electron density maps (initial 3.0 σ F_o-F_c map and the final refined 1.2 σ $2F_o$-F_c map) obtained from the co-crystallisation dataset of *Sp*SPT with D-serine, with the PLP-L-serine external aldimine modelled inside. The electron density contributes to the presence of L-serine in the active site. **B)** An overlay of the *Sp*SPT structure obtained via co-crystallisation with D-serine (green) and the *Sp*SPT external aldimine structure (2W8J, coloured in cyan). The $2F_o$-F_c electron density map from the refined D-serine co-crystallisation structure is contoured around 1.2 σ. While the density around R378 is weak suggesting flexibility of the residue, no contact with the external aldimine is observed as opposed to the published L-serine structure in cyan (the same residue marked with *). RMSD variation of the obtained structure with 2W8J – 0.558 Å.

Comparisons of the obtained structure of *Sp*SPT with the already published L-serine external aldimine (2W8J) and the PLP internal aldimine (Yard et al., 2007; pdb code 2JG2) structures reveal that it has more similarities with the latter one, despite having the same PLP-ligand complex as 2W8J. The core RMSD between the D-serine co-crystallisation product structure and 2W8J is 0.558 Å across 398 aligned residues; the RMSD between the structure and 2JG2 is 0.501 Å across 398 aligned residues (calculated in Coot). The differences are mainly contributed by the flexibility of the "PPAP" substrate-selection loop, yet more importantly by the position of the arginine 378 (Fig. 3.1.4.1, B).

In the internal aldimine structure of *Sp*SPT (2JG2), the residue points outwards from the active site of an enzyme and towards the protein surface. Upon binding of L-serine, the R378 performs a distinctive conformational switch and twists towards the carboxylate group of L-serine, providing a stabilising interaction for the enzyme-substrate complex (Raman et al., 2009, pdb code 2W8J). A similar conformational twist is expected to occur in *S.wittichii* SPT with the corresponding R358; while the residue is located near the surface of the protein in the published internal aldimine structure (Raman et al., 2010, pdb id 2X8U), it is likely to be involved in stabilisation of the external aldimine.

In the structure obtained by the co-crystallisation with D-serine, the density around the "PPAP" loop remains weaker than for the rest of the protein, suggesting that the loop is flexible in the crystal complex. While the R378 is not completely visible due to flexibility, the residual electron density shows that its guanidinium group is located on the surface of the protein and not near the active site where the L-serine is bound in the form of external aldimine (Figure 3.1.4.1, B).

Although indirectly, this observation points to some important features of the SPT binding mechanism. While D-serine can bind the PLP in the active site, it causes no conformational change of the R378 observed during the L-serine binding, suggesting that the enzyme has a distinctive and rapid mechanism for identifying its correct substrate. The absence of the conformational change of R378 indicates that initially D-serine has been bound in an external aldimine form and was slowly converted into

L-serine afterwards. No further conformational changes of R378 could occur as the enzyme has already retained its shape in the crystal lattice.

Remarkably, previous 2008 study by Ikushiro et al. provided evidence that similar deprotonation of the external aldimine form occurs in *Sp*SPT. The authors monitored the hydrogen-deuterium exchange on the Cα of serine bound to the SPT using NMR spectroscopy. Upon binding of the second substrate, reduction of the Cα hydrogen peak occurs for L-serine, but not for D-serine. However, in the absence of second substrate binding, very slow deprotonation is observed for both enantiomers with relatively close reaction rates of 0.19×10^{-5} s^{-1} for L-serine and 0.12×10^{-5} s^{-1} for D-serine, and half-lives of 103 h and 150 h respectively.

Interconversion of L- and D-serine is catalysed by the enzyme serine racemase belonging to β-family of PLP-dependent enzymes (Smith et al., 2010, pdb id 3L6B). In the reaction mechanism, the substrate (either L- or D- serine) replaces the catalytic Lys56 to form an external aldimine with the PLP in a similar manner to SPT; subsequent deprotonation and reprotonation of the Cα proceeds via acid/base catalysis enabled by Ser84 and Lys56.

Another amino acid racemisation mechanism is occurring in N-acylamino acid racemase (NAAAR, Thoden et al., 2004; pdb 1SJA). While NAAAR is not a PLP-dependent enzyme, it involves two basic residues, Lys163 and Lys 263, located opposite to each other across the Cα carbon of the substrate; depending on the amino acid enantiomer bound, one of the lysines acts as a base that deprotonates the Cα; this is followed by stereoinversion and reprotonation. Lack of the PLP cofactor that would act as an electron-withdrawing sink during the transition state limits the reaction to N-acylated amino acids only. A similar racemisation mechanism involving two basic residues located in one plane across the Cα-H bond, each specifically deprotonating either L- or D- enantiomer, might as well occur in SPT.

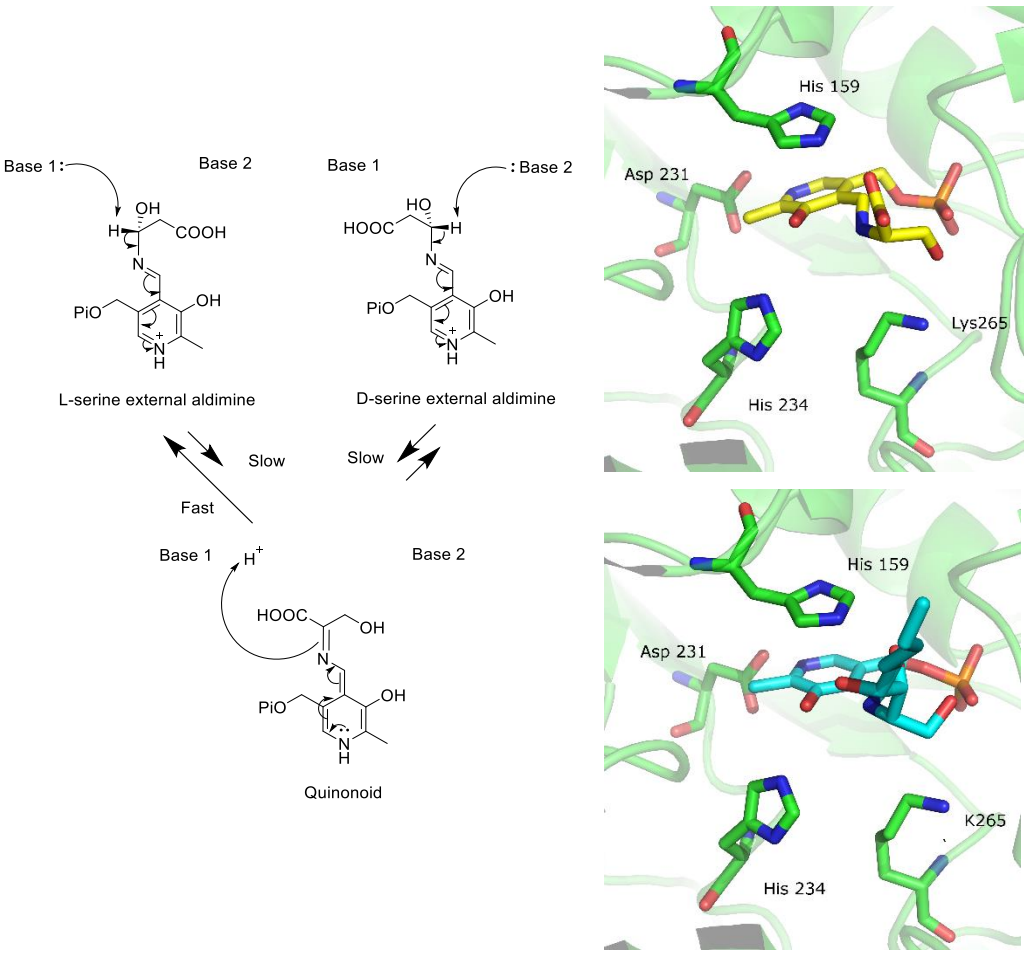


Fig. 3.1.4.2. Left: The proposed mechanism of serine resolution in the active site of SPT. Slow deprotonation of the Cα on both L-and D-serine leads to a formation of the planar quinonoid species with the subsequent resolution to L-serine. **Right**: Positioning of the L-serine in the obtained structure (top) and the decarboxymyriocin (bottom, pdb 4BMK) in the active site of SPT with the residues possibly involved in deprotonation of the Cα of the substrate.

During the SPT reaction mechanism, Lys265 acts as a base and rapidly deprotonates the Cα on the external aldimine of L-serine upon binding of the second substrate. However, as slow deprotonation occurs for both L- and D-serine in the absence of second substrate binding, it could be mediated by two basic residues that each can deprotonate a specific enantiomer. This will result in the formation of a planar quinonoid intermediate that will be rapidly reprotonated back, either by the catalytic lysine or the solvent. The restrictions provided by the active site of the enzyme would almost exclusively result in the formation of L-serine, thus leading to stereospecific resolution (Fig. 3.1.4.2, left). The catalytic Lysine265 residue would most likely act as one of the bases; possible second base could be the His159 located across the plane to

L265 and in parallel to the pyridine ring of PLP. His234, while being found at the opposite side of the Cα, is unlikely to play a role of the base due to a large distance between the N1 of the imidazole ring and the Cα of the bound substrate (5.8 Å).

His159 and Lys265 are located on either side of the substrate α-carbon (Fig.3.1.4.2, top right), thus allowing both residues to act as a deprotonating base. Upon binding of D-serine via external aldimine, the Cα-H bond might be oriented towards His159 and therefore be susceptible to slow deprotonation. After formation of the quinonoid intermediate planar to the PLP ring, a stereoselective reprotonation would occur yielding the resolution to L-serine external aldimine.

Myriocin has a D-configuration at its C2 and therefore is expected to have the same binding mode as D-serine. In the structure of SPT K265A with decarboxymyriocin (4BMK) the bound ligand retains D-configuration, while the Cα-H bond formed due to decarboxylation is directed towards His 234 (Fig. 3.1.4.2, bottom right). It remains likely that the D-configuration is retained due to the absence of the catalytic Lys265 that would act as a deprotonating base otherwise.

The observations derived from the co-crystallisation study are very intriguing as they suggest that SPT might have an inherent racemase activity. This activity could be tested using a variety of different methods such as chiral resolution HPLC and coupled enzyme assays. Although supplement of D-serine does not lead to the production of KDS by SPT (Hanada et al., 2000a), its possible conversion to L-serine might gradually result in the formation of KDS in the long term. It would not be possible to monitor this conversion using SPT only due to rapid degradation of PCoA, yet it might be possible to develop a coupled enzyme assay using an L-serine specific enzyme such as dehydratase. Studies could be further aided by using radioactively labelled serine.

It is important to see the effect of active site lysine on conversion of D-serine. Valuable information could be obtained by exploring the properties of the K265A mutant, particularly testing its ability to form quinonoid species and its interaction with D-serine. To test the validity of this assumption, a co-crystallisation study of the SPT K265A mutant could be performed with D-serine.

Finally, the differences the flexibility of the “PPAP” loop and the R378 between the obtained structure and the published L-serine external aldimine indicate on specific mechanisms employed by *Sp*SPT in substrate selection. This is in contrast to the SPT from *B.fragilis*, where different mechanisms are shown to occur upon the substrate binding, as discussed later in the text. The practical impact of this observation could explain the difficulties in obtaining crystal complexes of *Sp*SPT with L-penicillamine that have been accomplished using the *Bf*SPT as a model.

3.2. *Bacteroides fragilis* SPT

3.2.1. Expression and purification of *Bf*SPT

The putative *B.fragilis* SPT from the BF2461 gene (the nucleotide and protein sequences are presented in appendix 2) was successfully expressed in *E.coli* BL21 strain and purified according to the materials and methods. Visualisation of the protein fractions, taken at different expression and purification stages, by SDS PAGE shows the appearance of a band slightly below the 45 kDa (Fig 3.2.1.1, A). During purification by size-exclusion chromatography, the protein eluted in a peak with the retention time corresponding to an approximate MW of 90 kDa (Fig.3.2.1.1, B; the calibration curve is shown in appendix 3). This suggested that the protein is present in a form of dimer in solution as expected.

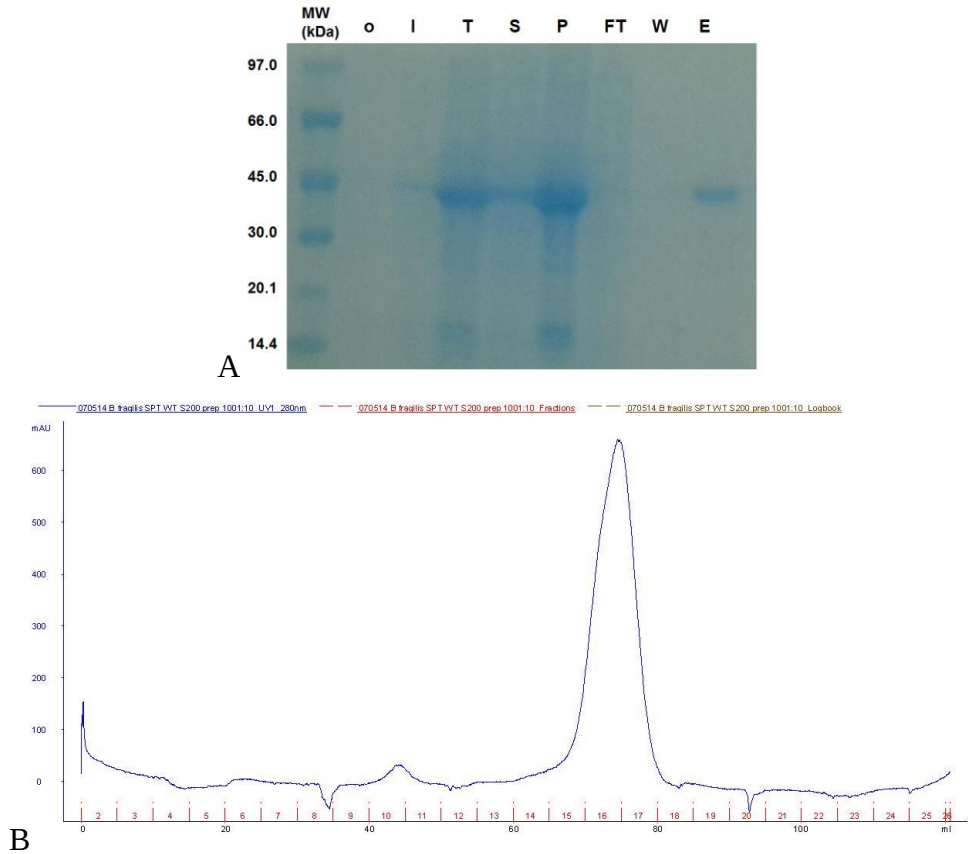


Fig 3.2.1.1. **A**) The SDS PAGE analysis of fractions collected during the expression and purification of *Bf*SPT. Labels are: o – pre-induction, I – induced sample, T – total cell fraction, S – soluble fraction, P – non-soluble pellet fraction (discarded); Ni-affinity purification flow through (FT), wash (W) and elution (E). **B)** Size exclusion chromatogram of *Bf*SPT with the retention peak appearing at fractions 16-17.

The analysis of the purified protein by ESI LC-MS has yielded a single species with the mass of 44613.8 Da (Fig.3.2.1.2); this corresponds to the mass of the His-tagged *Bf*SPT construct without the initial methionine (theoretical MW 44614.3 Da, ExPASy ProtParam).

Treatment of the sample with 10 mM $NaBH_4$ resulted in the chemical reduction of the internal aldimine bond between the PLP and the catalytic lysine (Fig. 3.2.1.2, bottom). The resulting PLP-enzyme adduct is stable enough to be observed by LC-MS. The spectrum of *Bf*SPT under these reducing conditions shows the appearance of the species with MW of 44845.6 Da. The increase of 232 Da well corresponds to the theoretical mass of the PLP (233 Da) being covalently bound to the protein in the reduced form.

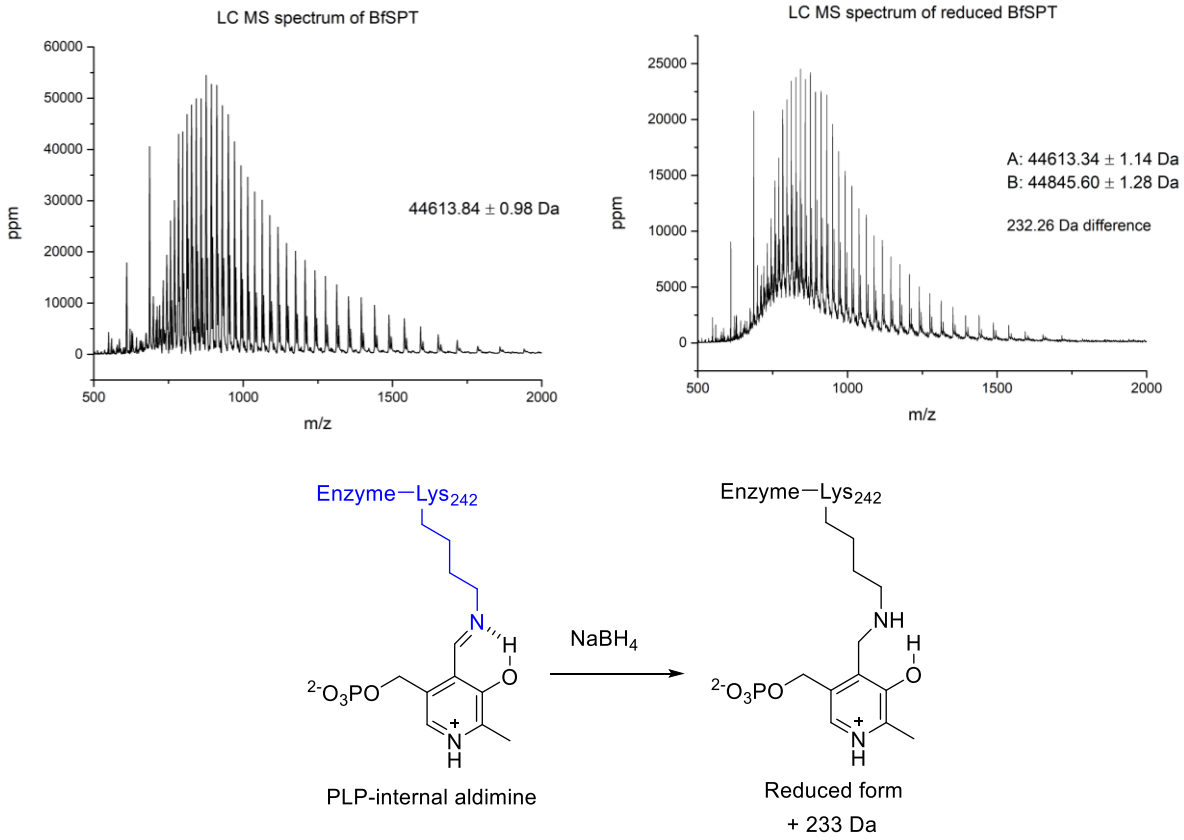


Figure 3.2.1.2. The LC-MS spectrum of *Bf*SPT in both as-purified and reduced forms. Theoretical MW of *Bf*SPT is 44614.30 Da (ExPASy ProtParam). The spectrum on the right has been obtained after the protein treatment with 10 mM $NaBH_4$. The aldimine bond is reduced and the PLP forms a stable covalent adduct with the enzyme; this was captured by LC-MS with the mass of 44845.6 Da (in agreement with the expected 233 Da increase).

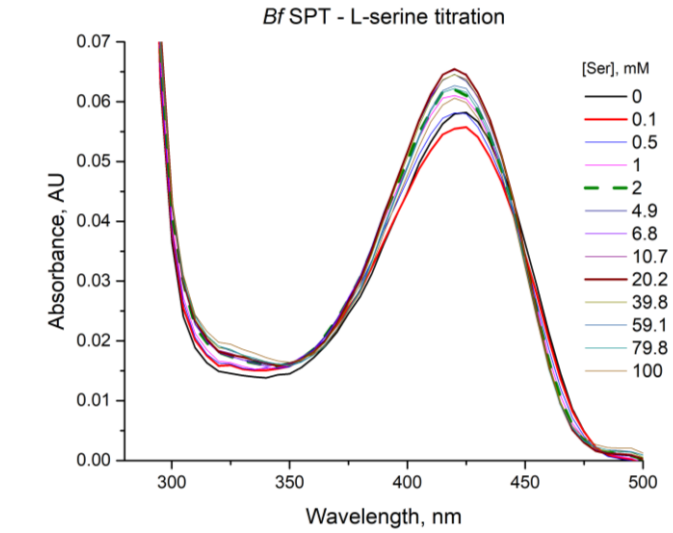
While having generally acceptable levels of expression and protein stability to be used in further studies, *Bf*SPT did show slightly weaker long-term stability compared to the *Sp*SPT. The protein stability has also varied broadly between different preparations. To address this issue, extra control was contributed to the protein induction, expression and purification stages.

3.2.2. Spectrophotometric properties of *Bf*SPT

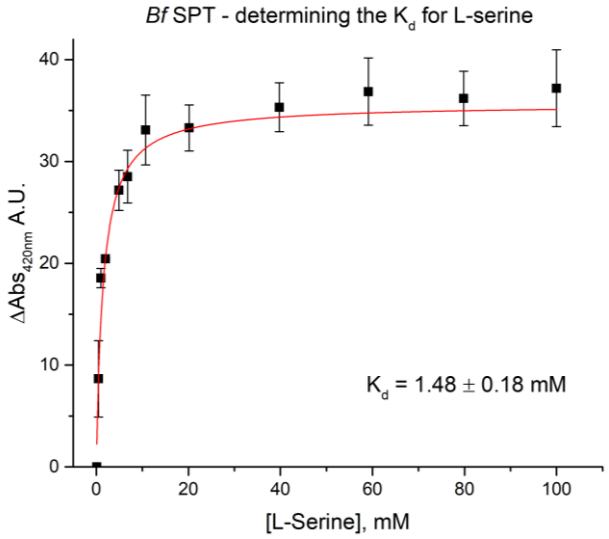
The purified *Bf*SPT gave a characteristic UV-vis spectrum of a PLP-dependent enzyme with the appearance of peaks at 330 and 420 nm corresponding to the enolimine and ketoenamine isoforms of PLP. Notably, in *Bf*SPT the maximum absorbance peak is shifted to 420 nm compared to 425 nm in the *Sp*SPT.

In contrast to the *Sp*SPT, the PLP of *Bf*SPT is present almost entirely in the ketoenamine form (Fig.3.2.2.1, A). This agrees well with the previous observations (E. Bower, MRes thesis), and has also been reported in the literature for the bacterial SPT from *Bdellovibrio stolpii* (Ikushiro et al., 2007) that has a 48% sequence identity with the *Bf*SPT (protein BLAST, AB259216 gene).

Binding of L-serine causes some visible changes in the equilibrium between both isoforms, however, these are less prominent than those of the *S.paucimobilis* SPT but look similar to a closer-related *S.multivorum* SPT (Ikushiro et al., 2007). Based on the spectral changes at 420 nm, the binding constant K_d of *Bf*SPT to L-serine has been determined as 1.48 mM (Fig. 3.2.2.1, B); this is relatively higher compared to some other SPTs yet remains in a very close range to them (the comparison of the K_d values of *Bf*SPT and other SPTs is presented in Table 3.2.3.3.2 further in a text).



A



B

Fig 3.2.2.1. A) Changes in the PLP spectrum of *Bf*SPT upon addition of increasing concentrations of L-serine. The saturated spectra observed at concentrations of L-serine > 20 mM have been omitted for simplicity **B)** The mean increase in the absorbance at 420 nm plotted against the concentration of L-serine to determine the binding constant K_d. The experiment was done in triplicate with the error bars presenting the standard deviation between absorbance changes in samples.

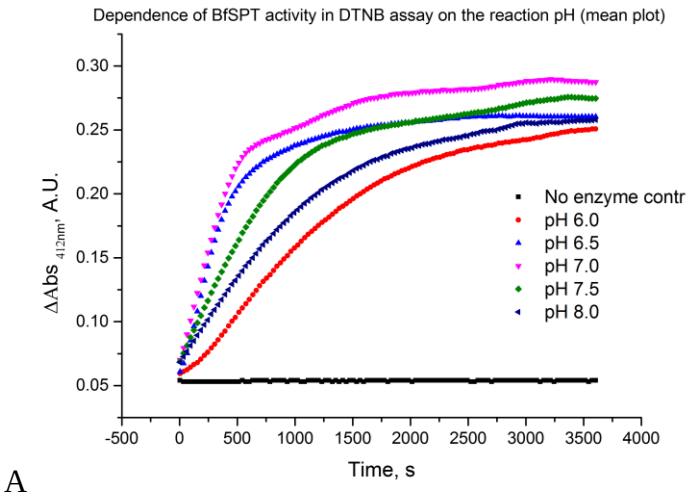
3.2.3. Kinetic characterisation of *Bf*SPT

3.2.3.1. The effect of pH on the activity of *Bf*SPT

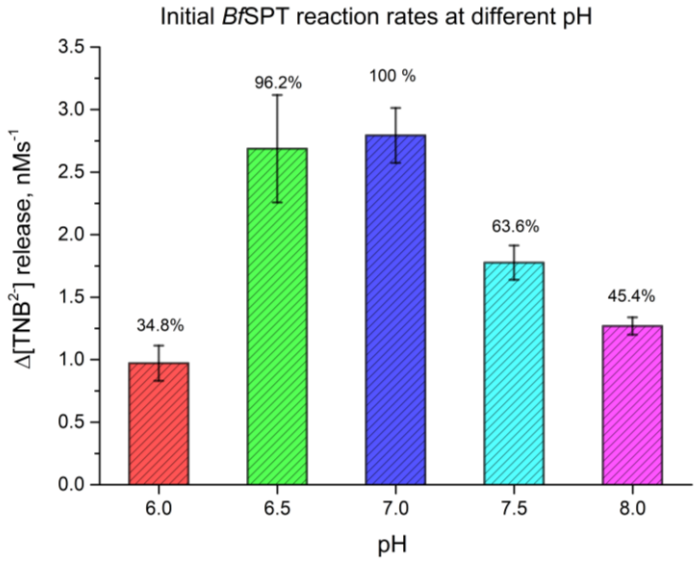
The DTNB assay (Ellman, 1959) has been widely utilised in measuring enzyme activities in reactions that involve thiol release (Riddles et al., 1983), and SPT in particular (Raman et al., 2009). Although the inability to monitor the release of the actual products is the biggest drawback, DTNB assay remains a convenient and routine way of monitoring the activity of SPT in the real time.

The activity of *Sp*SPT has been conventionally measured at pH 7.5 – 8.0. However, as the pH level of a human gut is typically ranged at 6.4-7.5 (Pye et al., 1990), pH optimisation for the *Bf*SPT activity was performed. The activity was tested using 5 different 100 mM KPhos buffers with the pH range from 6 to 8 (with 0.5 increments). All the reaction conditions apart from the pH were kept constant with the enzyme concentration of 200 nM, L-serine and PCoA concentrations of 20 mM and 25 µM respectively.

The activity of *Bf*SPT varies considerably across the pH range tested (Fig.3.2.3.1.1), with the highest reaction rate being observed at pH 7.0; at pH 6.5 the enzyme retains 96.2% of its highest activity, while at other pH values tested the activity drops to 34.8% (pH 6), 63.6% (pH 7.5) and 45.4% (pH 8.0). It is therefore evident that a careful pH control is needed for the optimal activity of the enzyme; hence all further kinetic experiments have been performed using buffered KPhos solutions of pH 7.0.



A



B

Fig. 3.2.3.1.1. The dependence of *Bf*SPT activity on the pH measured by the DTNB assay. **A**) The mean absorbance readings over the time of the reaction calculated from the raw data. **B**) Comparison of the relative activity of *Bf*SPT at different pH levels presented as a relative percentage rate (calculated from the maximum activity observed). The reactions were made in triplicate with the error bars representing the standard deviation between individual reactions.

3.2.3.2. Activity of *Bf*SPT with different amino acids

Wild-type *Bf*SPT was tested for inherent catalytic promiscuity towards other amino acid substrates using the optimised DTNB assay. Alongside L-serine, activities with the excess of glycine, L-alanine and L-threonine were tested (40 mM each). However, apart from L-serine, there was no TNB^{2-} release with the absorbance at 412 nm staying at the baseline level, suggesting that wild-type *Bf*SPT cannot utilise other amino acids tested as substrates (Fig.3.2.3.2.1).

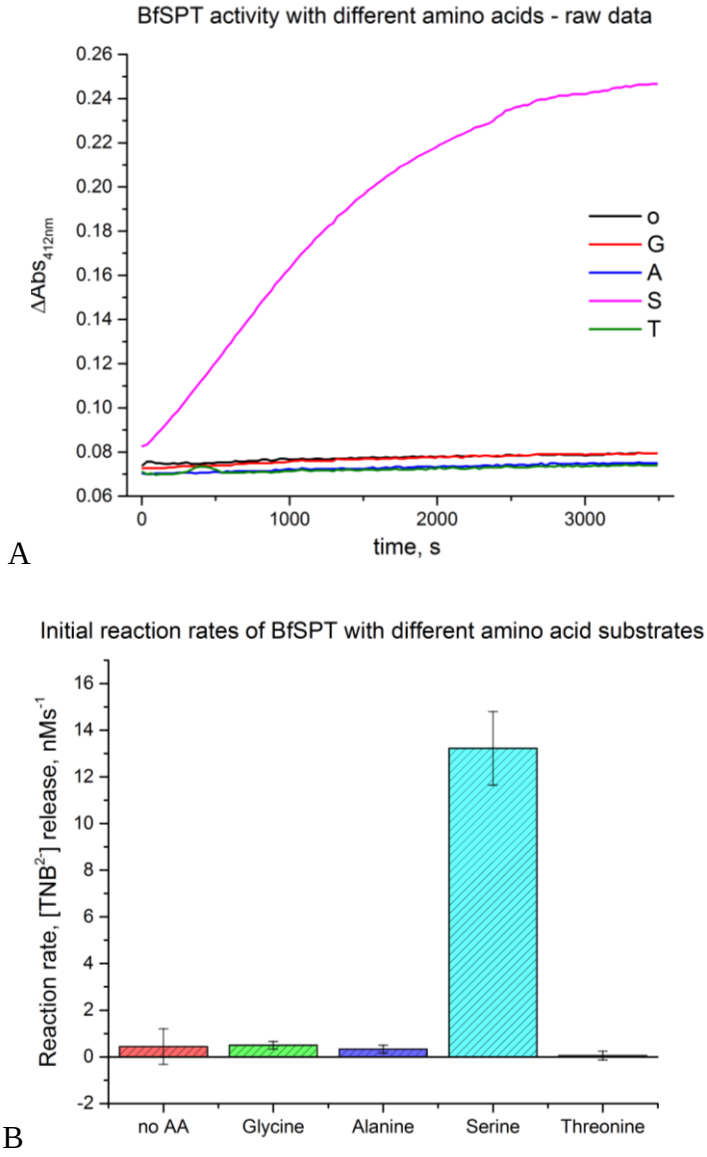


Figure 3.2.3.2.1. The activity of *Bf*SPT with different amino acids as measured by DTNB assay using PCoA as the second substrate. **A**) Averaged mean output of the raw data. Samples are labelled as: no amino acid control (o), glycine (G), L-alanine (A), L-serine (S) and L-threonine (T). **B**) Comparison of the mean reaction rates. Reactions were made in triplicate; the error bars present the standard deviation.

Just as most of the AOS family of enzymes, and its homolog *S.paucimobilis* SPT in particular, *Bf*SPT was able to bind to several amino acids. Apart from the expected substrate L-seine, the spectral changes in the protein indicated binding of similar amino acids glycine, L-alanine and L-threonine. However, when monitored by DTNB assay, *Bf*SPT was only able to utilise L-serine as its first substrate (Fig 3.2.3.2.1) while the activity with the three other amino acids did not increase above the negative control. In this respect, the enzyme reacts in a similar manner to all already reported SPTs.

To obtain the kinetic parameters for L-serine, palmitoyl-CoA was used as the best performing acyl-CoA substrate. The measurements were made at the constant concentration of PCoA of 50 µM, giving the optimal reaction rate (Fig. 3.2.3.2.2). The enzyme K_M for L-serine was determined as 0.64 ± 0.06 mM; while being lower than the already reported values for other bacterial SPTs, it is of a similar magnitude (compared in table 3.2.3.3.2). The determined enzyme turnover rate k_{cat} is 67.9 x $10^{-3}s^{-1}$. This is considered to be at the lower end of SPT activity, being very close to *S.witichii* (68.7 x $10^{-3}s^{-1}$) but about 15 times less of the *S.paucimobilis* SPT. However, the value was derived at conditions where the enzyme rates followed the Michaelis-Menten relations and did not account for any possible impacts of the second substrate, which are to be discussed later.

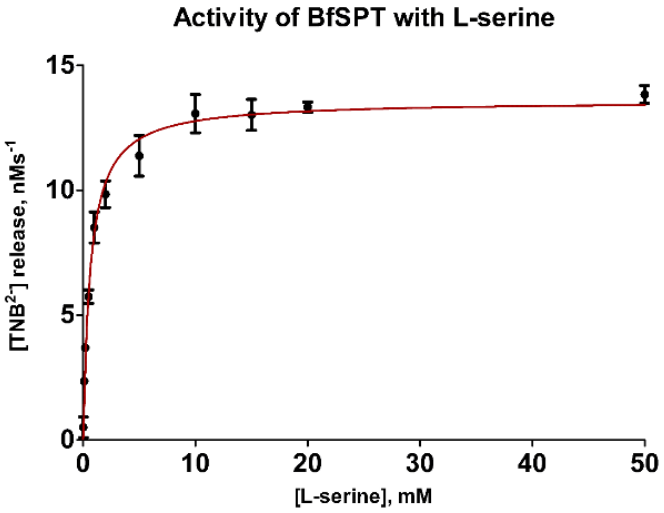


Fig. 3.2.3.2.2. The activity of *Bf*SPT in reaction with L-serine presented in terms of the initial rates of TNB^{2-} released per second (measured by the DTNB assay). The reaction rates are plotted against the increasing concentrations of L-serine; the saturation curve was fitted into the equation 3 to obtain the K_M. Each reaction condition was measured in triplicate; the error bars show the standard deviation

3.2.3.3. Studies of *Bf*SPT specificity towards acyl-CoA substrates

To test the enzyme specificity towards the second CoA-derived substrate, activities were measured by DTNB assay with L-serine and each of the following acyl-CoA substrates of different chain lengths (increasing by an extra CH_2 unit): myristoyl-CoA (C14), pentadecenoyl-CoA (C15), palmitoyl-CoA (C16), heptadecanoyl-CoA (C17) and stearoyl-CoA (C18).

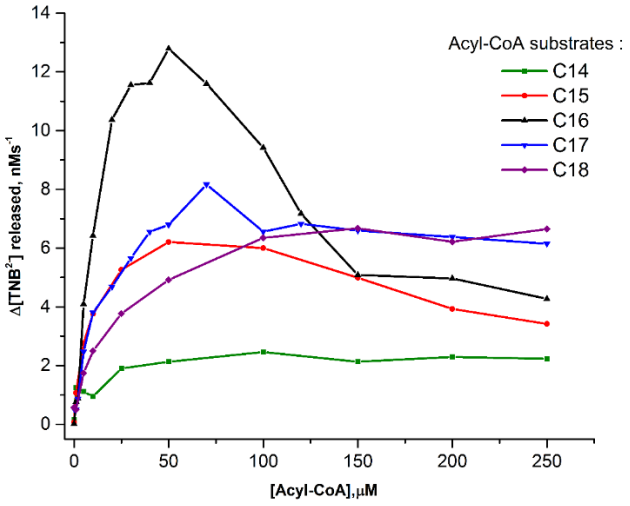


Figure 3.2.3.3.1. A summarised comparison of the relative activities of *Bf*SPT with different-length acyl-CoA substrates measured by the DTNB assay (only the same scale data points are included). The acyl-CoA substrates are labelled as indicated in the text above.

Individual reaction rates of *Bf*SPT plotted over the acyl-CoA concentrations are presented further in the Fig. 3.2.3.3.2. The obtained kinetic parameters are summarised in the Table 3.2.3.3.1.

The *Bf*SPT has shown activities with all acyl-CoA substrates, while changes in the activities with different-length substrates varied considerably (Fig. 3.2.3.3.1). The highest enzyme activities were observed for pentadecenoyl- (C15), palmitoyl- (C16)- and heptadecanoyl-CoAs (C17), yet these substrates also caused clear enzyme inhibition at high concentrations. The highest reaction rates were observed for palmitoyl-CoA, yet the substrate also caused the biggest drop in the reaction rates at higher concentrations. Activities with myristoyl- (C14) and stearoyl- (C18) CoAs were lower, while these substrates did not appear to inhibit the enzyme.

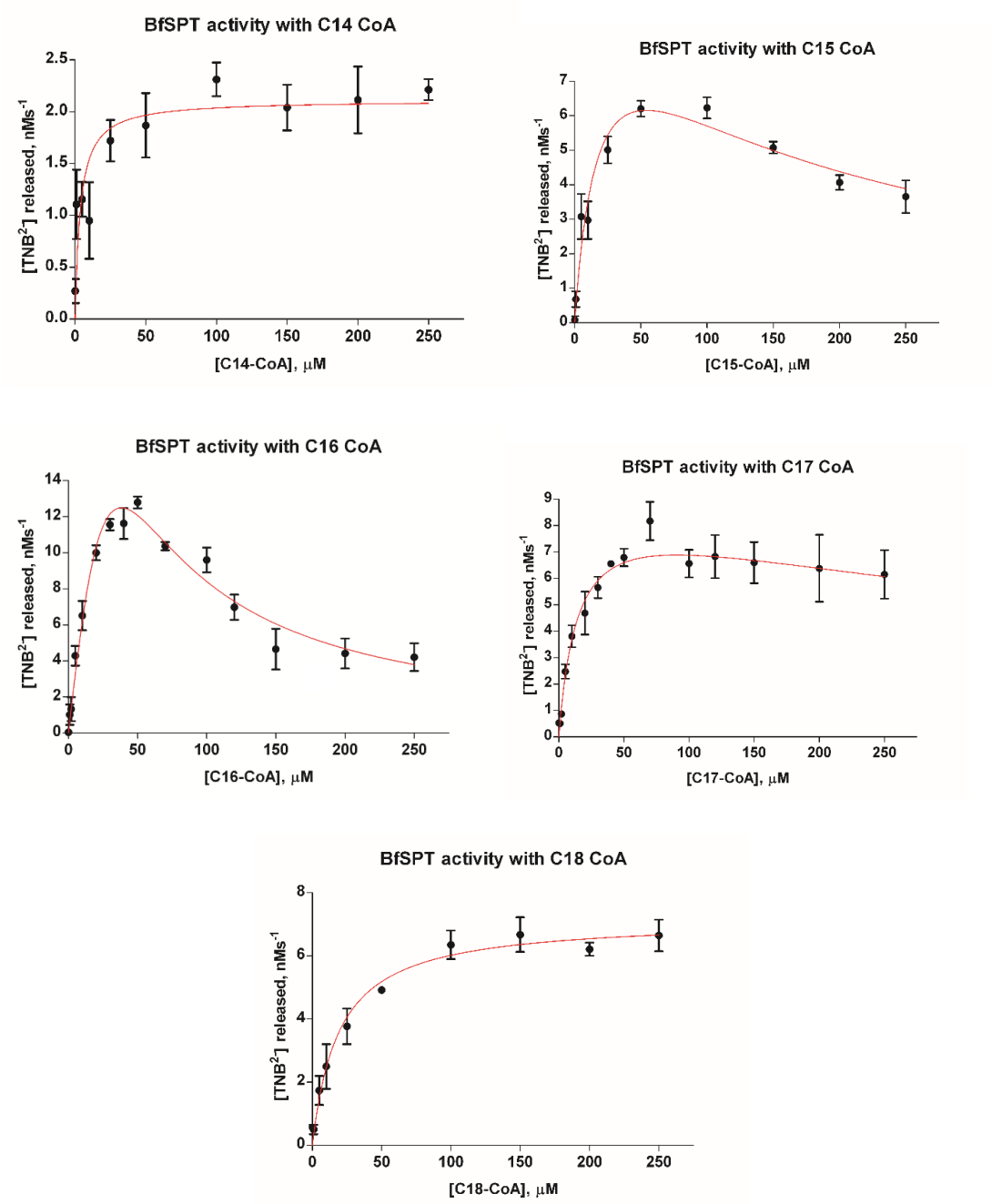


Fig. 3.2.3.3.2. Determining the activity of *Bf*SPT with various acyl-CoA substrates using the DTNB assay. The initial reaction rates are presented in terms of the amounts of TNB^{2-} ion released per second and plotted across the increasing concentrations of acyl-CoA substrates. The reaction rates were fitted into the appropriate equations 3 or 4 as discussed in section 2.2.3.7 and used to determine the kinetic parameters summarised in Table 3.2.3.3.1. Each reaction condition was performed several times (at least 5), with the error bars presenting the standard deviation between the individual readings.

Substrate	**K_M**	**k_{cat} x 10^3**	**k_{cat}/K_M**	**K_i**
	µM	s^{-1}	M^{-1} s^{-1}	*µM*
L-serine	636 ± 59	67.9 ± 1.3*	107	-
Myristoyl (14:0)	6.7 ± 2.8	6.2 ± 0.4	922	-
Pentadecanoyl (15:0)	19.2 ± 6.2	52.6 ± 8.8	2733	153.4 ± 50.8
Palmitoyl (16:0)	91.6 ± 39.7	361.0 ± 138**	3940	17.8 ± 8.3
Heptadecanoyl (17:0)	14.5 ± 4.7	45.4 ± 6.1	3135	569.1 ± 326.1
Stearoyl (18:0)	19.2 ± 3.7	35.9 ± 1.6	1871	-

Table 3.2.3.3.1. Summary of *Bf*SPT kinetic data collected on different substrates

Organism	**k_{cat} x 10^3**	**K_M (L-serine)**	**K_M(PCoA)**	**K_d (L-serine)**	**K_i (PCoA)**	**Pdb code**
	s^{-1}	*mM*	*µM*	*mM*	*µM*	
S.paucimobilis	1150 ± 30.0	1.4 ± 0.10	35.4 ± 2.0	1.1 ± 0.1	-	2W8J
S.wittichi	68.7 ± 1.5	0.78 ± 0.10	23.4 ± 4.5	0.80 ± 0.1	-	2X8U
S.multivorum	120 ± 10.0	4.8 ± 0.06	100 ± 10.0	0.47 ± 0.1	-	3A2B
B.fragilis	67.9 ± 1.3* 361.0 ± 138**	0.64 ± 0.06	91.6 ± 39.7	1.48 ± 0.18	17.8 ± 8.3	tbc

Table 3.2.3.3.2. Comparison of *Bf*SPT kinetic parameters with those of other published SPTs.

* - derived from the obtained V_{max} at fixed concentration of PCoA where enzyme shows the highest observed activity

** - derived from the estimated value of V_{max} obtained by analysing the impact of substrate inhibition on the reaction rate

*Bf*SPT retains a degree of catalytic promiscuity towards its second substrate. In this respect, the enzyme is similar to other characterised bacterial SPTs (Ikushiro et al., 2004, Raman et al., 2009). As monitored by the DTNB assay, the enzyme shows activity with all five acyl-CoA substrates tested, while the extent of this activity varies depending on the substrate.

Most strikingly, while the enzyme had shown the highest activity towards palmitoyl-CoA (C16) compared to any other substrates at concentrations ≤50 μM, it had its reaction rates steeply reduced at higher concentrations, producing a characteristic inhibition pattern. To check whether this inhibition could be caused by the substrate and not by a formation of the product instead, several reaction conditions were repeated in the presence of 20-300 μM KDS. However, it was found that the reaction rates were very similar to the observed before, and not affected by the addition of various concentrations of the KDS product.

To determine the kinetic parameters of the enzyme towards the PCoA, the Michaelis-Menten equation modified for substrate inhibition was used. Since the true values of K_M, V_{max} and K_i are unknown and appear to have a high interference between them, the presented apparent values are subject to a large statistical error.

The observed substrate inhibition pattern indicates that while the enzyme processes PCoA as a substrate with the K_M of 91.6 μM, it produces an inhibitory complex at much lower concentrations (K_i of 17.8 μM), giving approximately 5-fold preference towards the latter. As $K_i >>> K_M$ it is not possible to obtain an accurate value of V_{max} and therefore k_{cat} with high certainty. The determined k_{cat} of 361.0 x $10^{-3}s^{-1}$ is being several magnitudes higher than the previously reported value of 67.9 x $10^{-3}s^{-1}$ determined at the steady-state conditions with the optimal concentration of PCoA at 50 μM. The enzyme has the highest specificity ratio of k_{cat}/K_M of 3940 $M^{-1}s^{-1}$ for PCoA, however, due to the observed high inhibition, it is not clear whether it utilises the substrate *in vivo*.

Two of the second best-performing substrates, pentadecenoyl- (C15) and heptadecanoyl – (C17) CoAs also showed some degree of substrate inhibition although in both cases it was not as severe as with the palmitoyl-CoA.

The enzyme K_M for pentadecenoyl-CoA (C15) at 19.2 µM is considerably lower than its K_i of inhibitory complex formation at 153.4 µM. This appears similar to the K_M of 14.5 µM for heptadecanoyl-CoA (C17), while the substrate has much higher K_i of 569.1 µM. Both substrates have lower enzyme turnover rates k_{cat} (52.6 x 10^{-3} and 45.4 x 10^{-3} s^{-1} respectively), yet due to the lower K_M as compared to PCoA, the enzyme specificity indexes remain high for both substrates (2733 $M^{-1}s^{-1}$ for C15-CoA and 3135 $M^{-1}s^{-1}$ for C17-CoA).

Notably, the C17 substrate has produced the biggest variation in the enzyme activity at concentrations ≥50 µM in comparison with other tested acyl-CoAs. During multiple readings performed at the same conditions, it was possible to observe reductions of the initial rate at high concentrations, while in other instances no inhibition was present. This unexpected variation could not be due to the experimental limitations only as all the simultaneous readings were taken across the whole range of acyl-CoA substrates at constant conditions. It was, however, possible to observe that the enzyme had seen bigger reductions in reaction rates developing into a substrate inhibition pattern in repeating readings as it was kept over time (while the rate would not be affected in the same manner for similar substrates C16 and C18 CoA).

It might be possible to speculate about the presence of specific substrate selection mechanism in *Bf*SPT. The enzyme could be specifically inhibited by palmitoyl-CoA as it is the main acyl-CoA precursor of sphingoid bases in humans, and thus the biggest competitor to the proposed substrate isopentadecanoyl-CoA. The SPT remains less specific towards the one carbon-longer heptadecanoyl-CoA. Yet it was not possible to pinpoint the very exact nature of this phenomenon due to the experimental limitations.

No substrate inhibition was observed at high concentrations of both stearoyl- and myristoyl-CoAs. K_M for stearoyl-CoA was 19.2 µM with the turnover rate k_{cat} of 35.9 x 10^{-3} s^{-1}, giving the specificity index of 1871 $M^{-1}s^{-1}$.

While the K_M for myristoyl-CoA (C14:0) of 6.7 µM is the lowest amongst all determined, due to the slowest turnover (k_{cat} = 6.7 x $10^{-3}s^{-1}$) it was the weakest substrate for *Bf*SPT, giving specificity index of 922 $M^{-1}s^{-1}$.

In summary, the kinetic results show that *Bf*SPT has the highest specificity for the 16-carbon long PCoA, with slightly reduced specificity levels for the 15- and 17- carbon length substrates. Yet it presents a tight inhibitory mechanism towards the PCoA that is less evident towards the one carbon shorter C15-CoA and one carbon longer C17-CoA. This indicates a specific selection mechanism that is likely to play role in a competition of PCoA with the proposed natural substrate isopentadecenoyl-CoA in *Bf*SPT. The biggest setback so far was an inability to use the proposed natural substrate of the enzyme. Up to our knowledge both the iso-branched CoA or its precursor isopentadecanoic acid are not available commercially.

While the findings highlight the ways *Bf*SPT can process some of the acyl-CoA substrates *in vitro*, the enzyme activity *in vivo* might be subject to a very different array of factors. The most significant question remaining is how the findings relate to the sphingolipid metabolism of *B.fragilis in vivo* and the particular impact of *Bf*SPT on the relation of the commensal and its host.

3.2.3.4. Characterisation of *Bf*SPT products by MALDI MS

To confirm that the monitored *Bf*SPT reactions with acyl-CoAs resulted in the formation of the corresponding KDS-like products, reaction spectra were obtained using positive ion MALDI ToF.

Spectra for *Bf*SPT reactions with the C14, C15, C16, C17 and C18 acyl-CoA substrates were analysed; to provide a suitable control negative samples for each reaction condition were set up without the presence of the enzyme.

The obtained reaction spectra are presented in the Figure 3.2.3.4.1 (parts 1-3) over the next few pages. The negative control samples are shown on top for each substrate.

Positive ion peaks were detected at m/z of 272.2 (C14), 286.2 (C15), 300.2 (C16), 314.2 (C17), 328.3 (C18), each corresponding to the monoisotopic mass of the positive ion (+1) of the expected KDS product. The peaks did not appear in the negative control reactions. The obtained spectra show the presence of the KDS-like products and therefore confirm that *Bf*SPT catalyses the condensation of L-serine with all the acyl-CoA substrates tested.

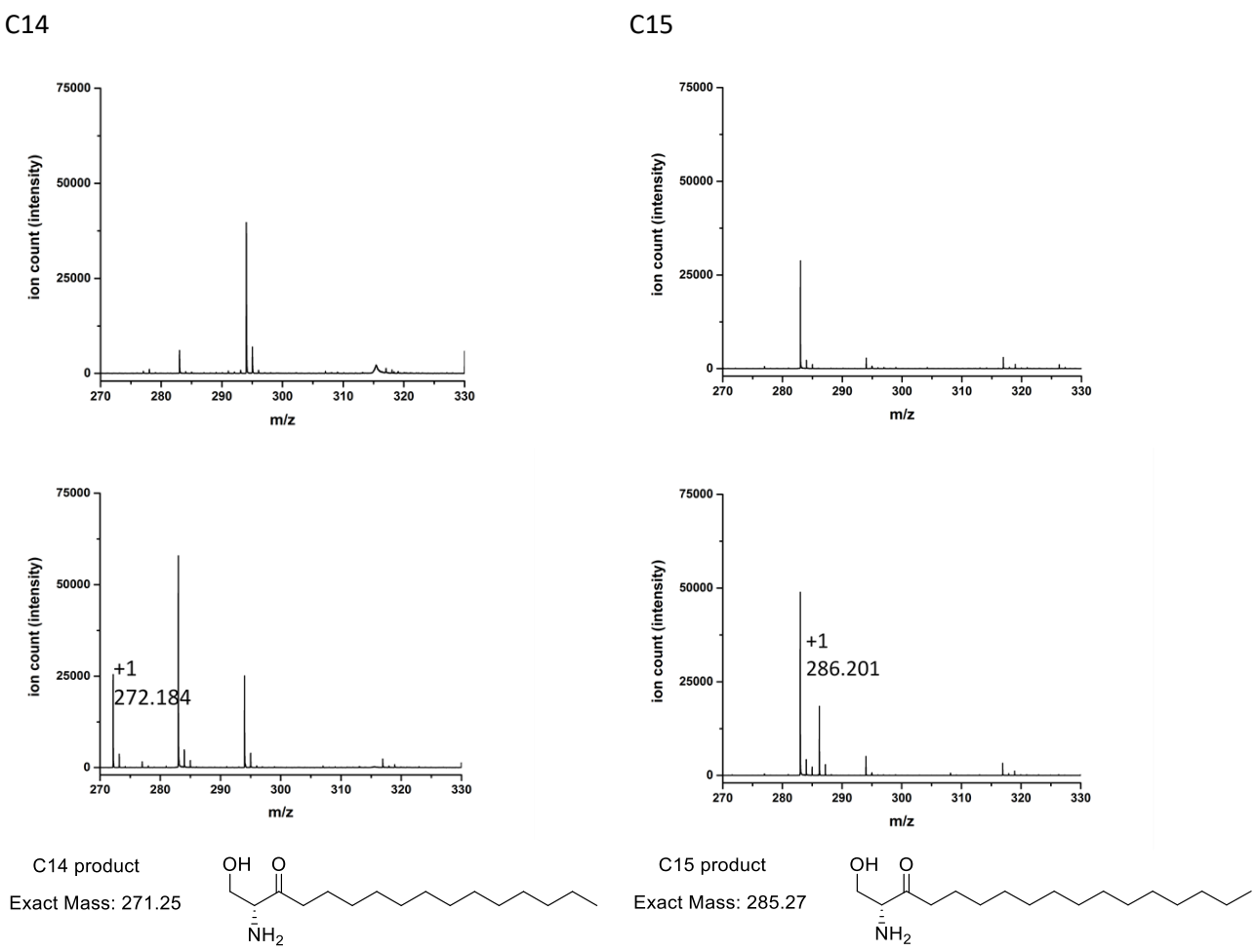


Fig. 3.2.3.4.1 (part 1). MALDI-ToF spectra of the *Bf*SPT reaction products with C14 and C15. The top spectra correspond to negative control reactions containing both L-serine and corresponding acyl-CoAs but no enzyme. The bottom spectra of the enzymatic reactions each present a peak corresponding to the positive ion of the expected KDS-like product.

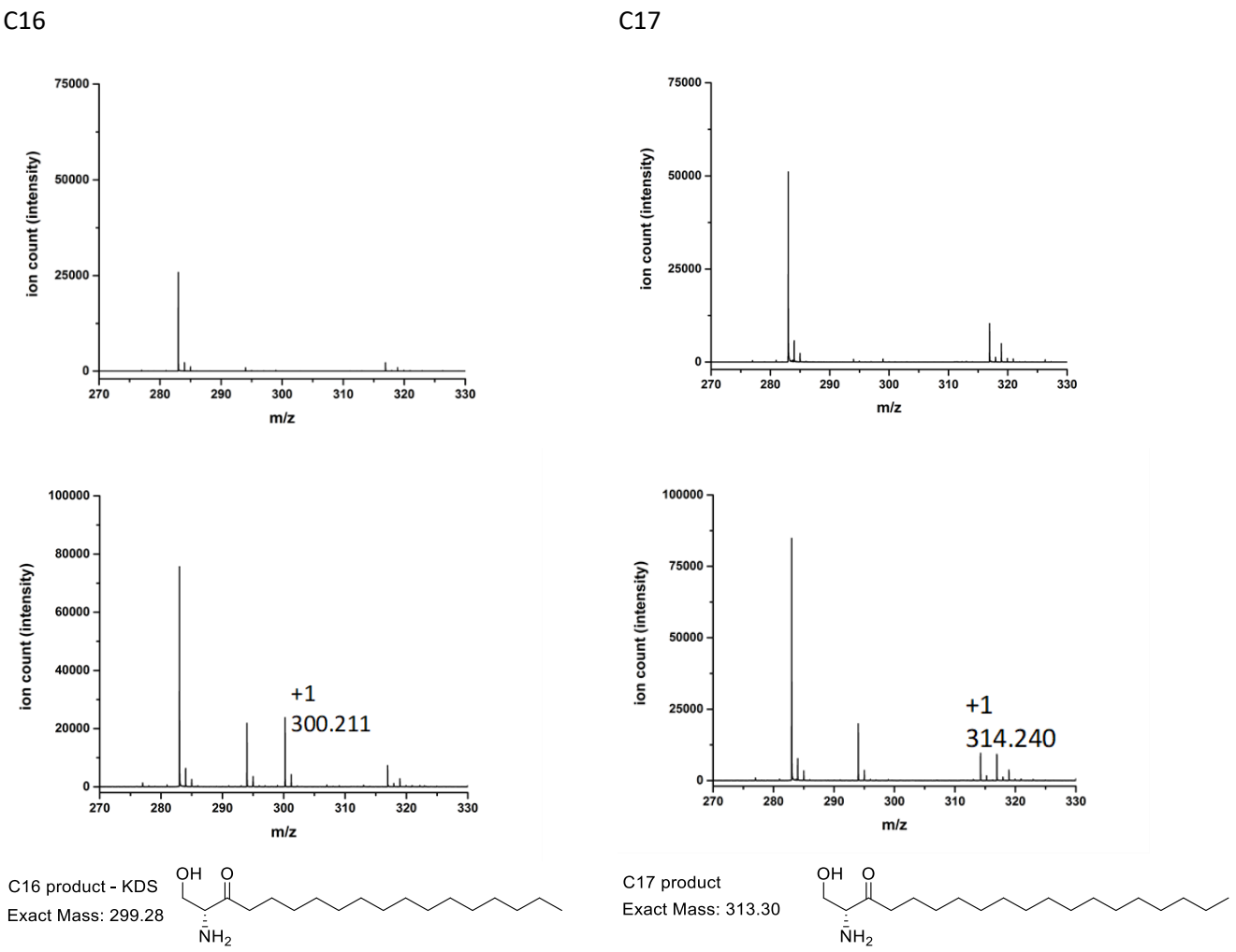


Fig. 3.2.3.4.1 (part 2). MALDI-ToF spectra of the *Bf*SPT reaction products with C16 and C17. The top spectra correspond to negative control reactions containing both L-serine and corresponding acyl-CoAs but no enzyme. The bottom spectra of the enzymatic reactions each present a peak corresponding to the positive ion of the expected KDS-like product.

C18

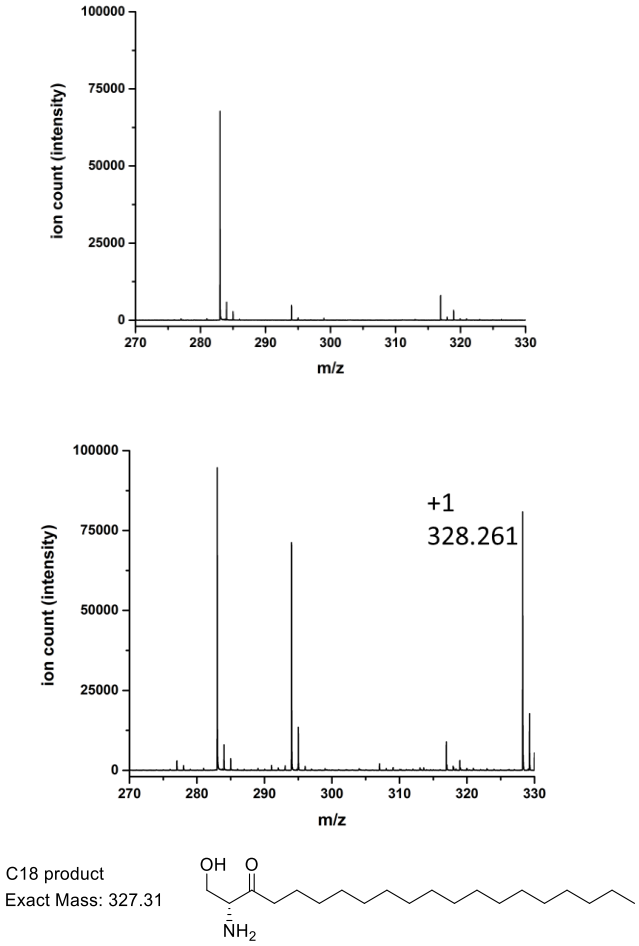


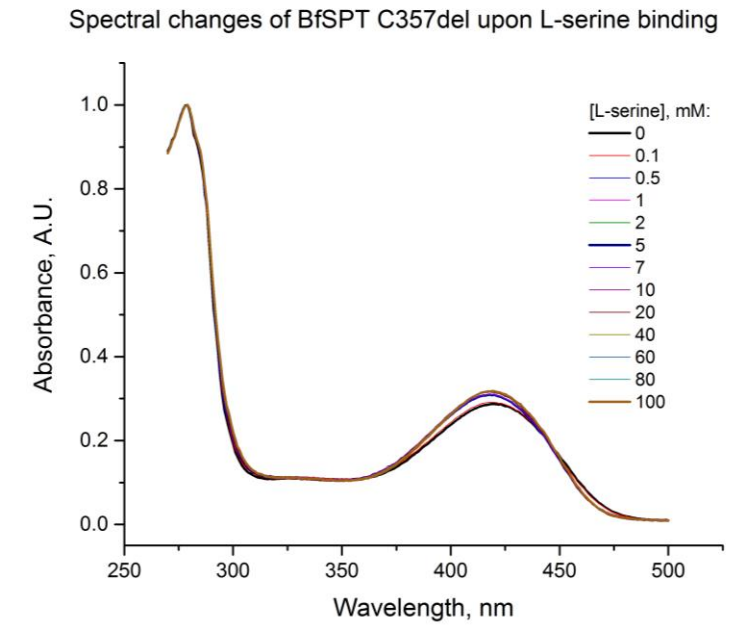
Fig. 3.2.3.4.1 (part 3). MALDI-ToF spectra of the *Bf*SPT reaction products with C18. The top spectrum corresponds to negative control reactions containing both L-serine and C18-CoA but no enzyme. The bottom spectrum of the enzymatic reaction presents a peak corresponding to the positive ion of the expected KDS-like product.

3.2.3.5. Assessing the activity of *Bf*SPT Cys357del mutant

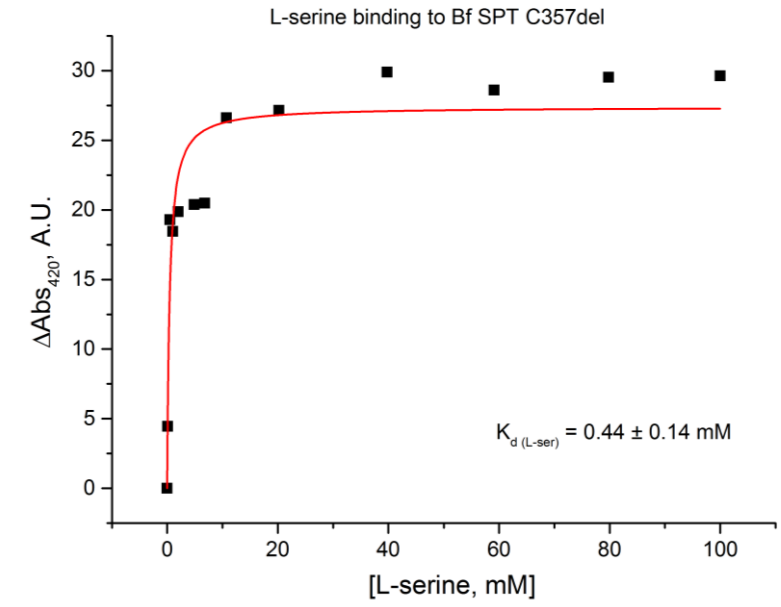
The flexible "PPAP" loop is present in other SPTs and AONS enzymes and thought to be involved in a selection of the second substrate. The loop contains generally distinctive motive of conserved proline and alanine followed by another one or two amino acids and another proline; the exact conservation of residues tends to vary a lot even within the proposed group of more related isoSPTs (discussed later in chapter 3.3.7). Within the structure of the loop, *Bf*SPT contains the unusual cysteine residue followed by serine (PPACSP) similarly. The significance of this cysteine residue has long been questioned, with possible theories including involvement in the specific selection of the second substrate and controlling the second substrate inhibition. To test the significance of cysteine 357 in the PPACSP loop of the *Bf*SPT, a deletion mutant lacking the amino acid has been made. The protein was expressed and purified using the general protocol.

The *Bf*SPT Cys357del mutant showed a generally similar spectral profile to the WT protein with low enolimine and high ketoenamine peaks (Fig 3.2.3.5.1, A). Upon titration with L-serine, only minor spectral changes occurred with a slight increase of the 420 nm peak; yet the binding saturation occurred very rapidly with the derived binding constant of L-serine K_d equal to 0.44 mM ((Fig 3.2.3.5.1, B). This is approximately 3 times lower than of the WT enzyme (1.48 mM), suggesting that removal of the cysteine in the flexible substrate selection loop increases the binding affinity of L-serine. This remains consistent with the theoretical expectations; it is likely that the cysteine thiol group would provide certain repulsion to the serine molecule before it enters the active site of the enzyme.

The kinetic activity of the C357del mutant was tested using the DTNB assay under the same conditions as the WT. The main aim was to observe whether the deleted cysteine residue would have any profound effect on the second substrate inhibition observed in the WT enzyme. The enzyme reaction was set up with increasing concentrations of PCoA over the range of 0-250 µM; the concentration of L-serine was kept in excess at 40 mM; concentration of enzyme was set at 200 nM.



A



B

Fig 3.2.3.5.1. A) Changes in the PLP spectrum of *Bf*SPT upon addition of L-serine. **B**) Changes in the absorbance peak at 420 nm plotted against the concentration of substrate and used to determine K_d of L-serine binding to *Bf*SPT C357del

The activity rates of the C357del mutant were approximately 4-5 times lower than of the wild type (Fig 3.2.3.5.2); yet the enzymatic rates at higher concentrations of PCoA were still reduced, showing a similar substrate inhibition pattern to the wild-type enzyme. Just as in the WT, determination of the exact kinetic parameters of the mutant remained ambitious; despite observing reduced rates of enzymatic reactions V_{obs}, it was not possible to determine exactly which of the kinetic parameters (K_M, K_i or V_{max}) is affected by the mutation. Nonetheless, it remains evident that removal of cysteine 357 from the substrate-selection PPACSP loop does not eliminate the substrate inhibition by PCoA, and therefore it is controlled by other factors.

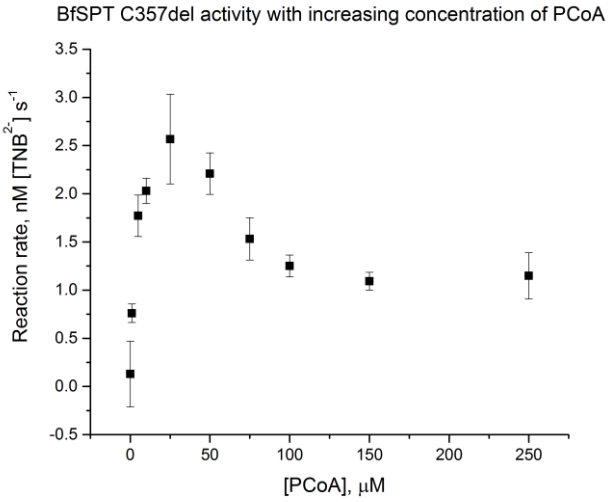


Fig 3.2.3.5.2. Initial reaction rates of *Bf*SPT C357del mutant plotted across the different PCoA concentrations ranging between 0-250 μM, measured by DTNB assay. The reaction was done in triplicate with the average mean rates plotted; the error bars are showing the standard deviation between the samples

The residue S358 also remains characteristic for *Bf*SPT only and is not present in other SPTs that have been experimentally characterised. It could be reasonable to assume that both residues C357 and S358 are involved in the specific substrate binding mechanisms of *Bf*SPT. A generation of the S358A mutant might be used for further exploration of the PACSP loop and its effect on the enzyme mechanism, with a later potential of generating a double C375del S358A mutant.

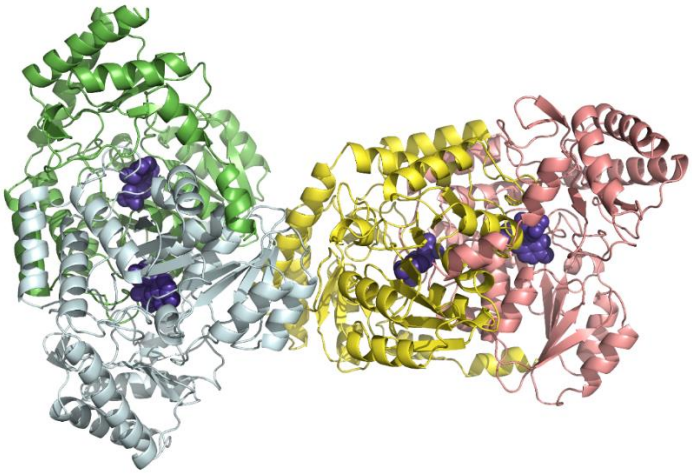
3.3. Structural and inhibitor-binding studies of *Bf*SPT

3.3.1. The crystal structure of *Bf*SPT with L-serine

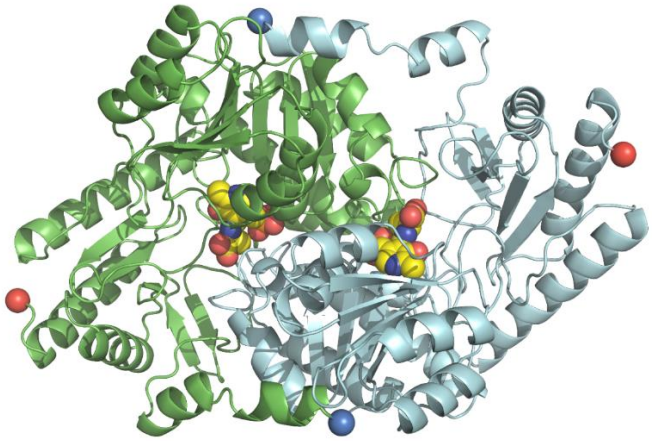
The crystal structure of *Bf*SPT complexed with L-serine was solved by molecular replacement using the homologous structure of *Sp*SPT (2W8J) as a template. The obtained highest resolution was 2.45 Å with the final refinement values of R_{work}/R_{free} being 0.21/0.25.

One asymmetric unit consists of four monomeric *Bf*SPT chains composing two homodimers AB and CD oriented perpendicularly to each other (Fig. 3.3.1.1, A). This type of assembly appears to be crystallisation related as the purification and native mass spectrometry data suggests that *Bf*SPT is present as a homodimer of two chains in the biological form (Fig. 3.3.1.1, B).

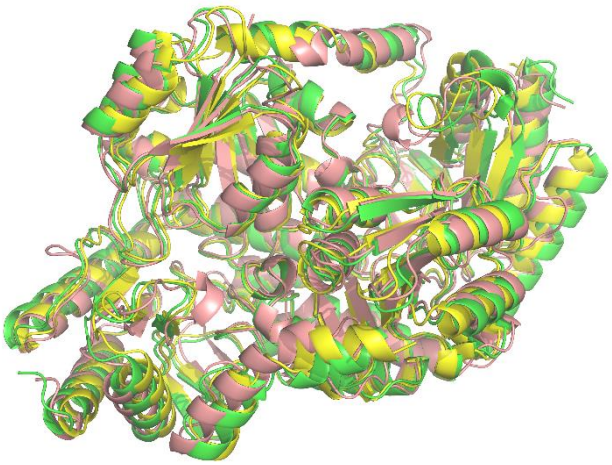
*Bf*SPT retains a general structure of other AOS enzymes resembling high homology with other the known SPTs. The Cα RMSD (calculated using Coot, Emsley et al., 2010) of the *Bf*SPT relative to the other SPTs (aligned with chain A of *Bf*SPT) are: 1.39 Å over 384 residues of the *S.paucimobilis* enzyme (2W8J), 1.36 Å over 381 residues of *S.witichii* SPT (2X8U, internal aldimine) and 0.93 Å over 381 residues of the *S.multivorum* SPT (3A2B). The average RMSD difference between the individual chains of the *Bf*SPT is 0.34 Å (the smallest of 0.18 Å being within chains A and C; the biggest of 0.44 Å between C and D). Least-square alignment of *Bf*SPT with the external aldimine structures of *Sp* and *Sm* SPT is shown in Fig. 3.3.1.1, C.



A



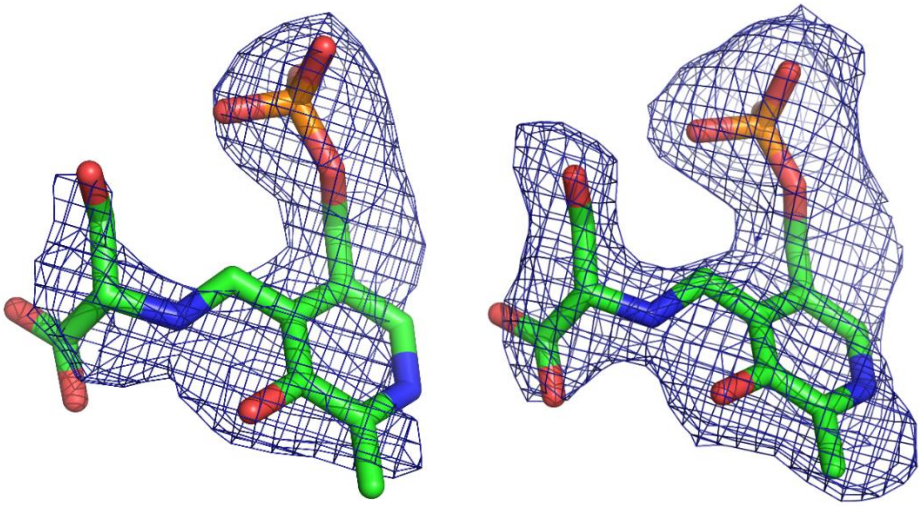
B



C

Fig. 3.3.1.1. The 2.45 Å structure of *Bf*SPT in a complex with the L-serine external aldimine. **A**) The asymmetric unit is made of 4 protein chains forming two dimers; dimer formed with chains A (green) and B (cyan) is perpendicular to the second dimer made of chain C (yellow) and D (magenta). The position of the PLP-L-serine external aldimine complex is marked by spheres in dark blue. **B**) The biological unit formed by two monomeric subunits A and B. The protein chain termini are marked by blue (N) and red (C) spheres. The PLP-L-serine external aldimine complex is presented in spheres with carbon atoms coloured yellow. **C**) The least-square superposition of *Bf*SPT with *S.paucimobils* SPT (2W8J, yellow) and *S.multivorum* SPT (3A2B, magenta) external aldimine structures. *Bf*SPT has RMSD of 1.39 Å with *Sp*SPT (35.4% seq. identity) and RMSD of 0.93 Å with *Sm*SPT (57.4% seq. identity).

The initial unbiased F_o-F_c omit map contoured at 3.0 σ level included positive density in the active site of the enzyme well corresponding to the PLP-L-serine external aldimine (Fig. 3.3.1.2); the structure was refined with PLP-L-serine external aldimine with the final F_o-F_c density map well corresponding to the shape of the ligand.



3.0 σ F_o-F_c map 2.6 σ F_o-F_c map

Fig. 3.3.1.2. The PLP-L-serine external aldimine density models. **Left**: The initial F_o-F_c density maps (contoured at 3.0 σ, cut off at 2.0 Å) calculated from the structure that did not have the PLP-L-serine complex modelled in. **Right**: Final F_o-F_c map obtained after refinement with the PLP-L-serine external aldimine complex modelled in, shown at 2.6 σ, cut off 2.1Å.

The PLP cofactor is located on the dimer interface (Fig. 3.3.1.3**)**; its pyridine ring is held in parallel to the essentially conserved H136 residue and stabilised by polar contacts with D208 (2.9 Å) and H211 (3.2 Å). The phosphate group is stabilised by a 2.6 Å polar contact with S241 of the same subunit and the contacts with S271 and A272 of the opposite subunit (2.8 and 2.7 Å respectively). The L-serine is bound to PLP via aldimine linkage (external aldimine); the catalytic lysine242 is not bound to the PLP. The carboxyl group of L-serine is pointing into the opposite direction from K242 and makes a 3.1 Å polar contact with H136; the hydroxyl group of serine makes a 3.1 Å polar contact with the phosphate group of PLP. The hydroxyl group of L-serine is oriented towards the conserved N50 yet there is no polar contact due to a large distance of 4.4 Å.

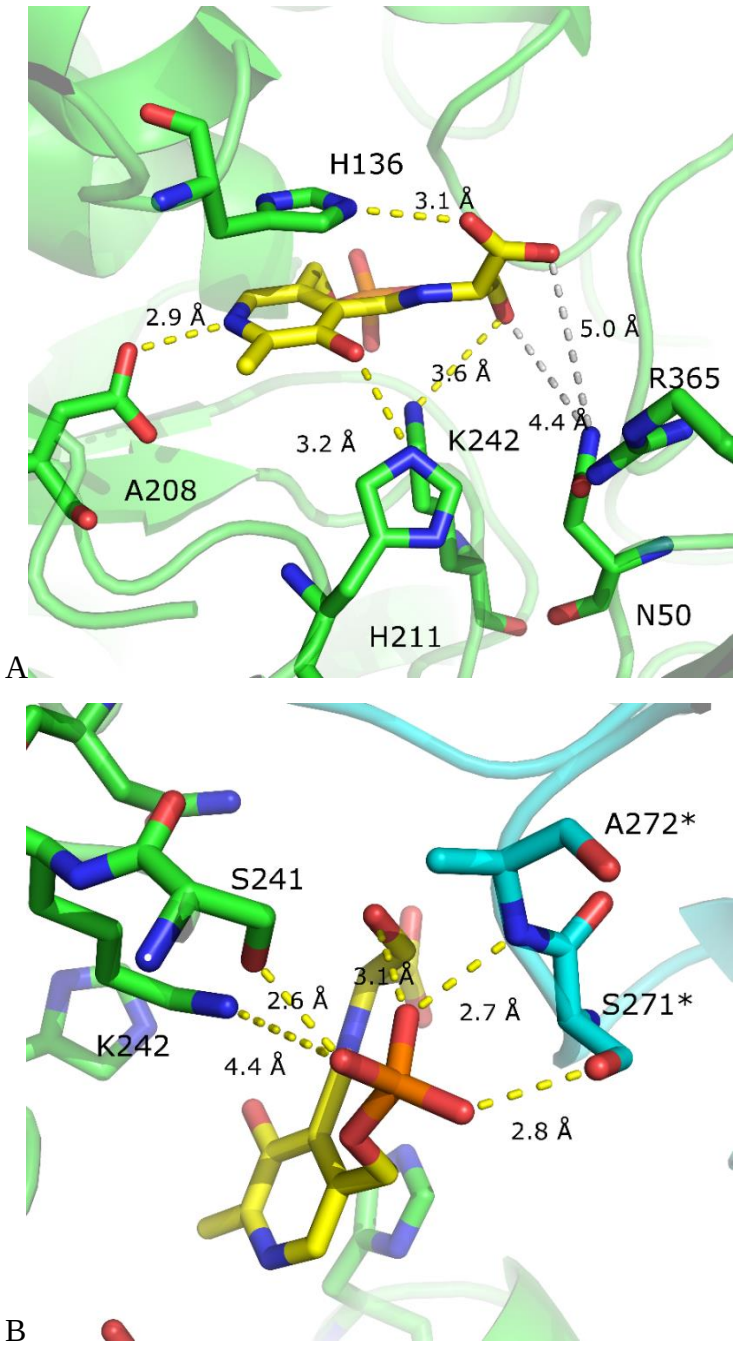


Fig. 3.3.1.3. The position of L-serine external aldimine in the active site of the *Bf*SPT structure (Chain A). **A**) The key residues involved in stabilising the PLP-L-serine external aldimine. **B**) Key interatomic distances between the protein backbone and the phosphate group of PLP. The residues marked with * belong to the opposite protein subunit.

While the general geometry of the active site is similar to *Sp*SPT, the noticeable differences between the external aldimine forms of two enzymes arise in the stabilisation of bound L-serine by the additional residues. In the *Sp*SPT (2W8J, Fig.

3.3.1.4, A) the carboxylate of L-serine is stabilised by a contact with the flexible R378 residue located right next to the flexible PPATP loop. The analogical residue R358 is present in *S.wittichi* SPT near the protein surface in the internal aldimine structure (Raman et al., 2010, pdb id 2X8U) and expected to perform a conformational twist upon the substrate binding, while no external aldimine structure was published. In a contrast, *Bf*SPT (Fig. 3.3.1.4, B) is missing a similar interaction with the carboxylate of L-serine; the analogical residue of *Bf*SPT is V353 that would not be able to interact with the L-serine. In this respect, *Bf*SPT is reminiscent of *S.multivorum* SPT (Ikushiro et al., 2009, pdb id 3A2B) that has a similar substrate binding interactions (Fig. 3.3.1.4. C).

Both *Bf*SPT and *Sm*SPT have N50 and N52, respectively, located at the opposite plane in the active site. In the structure of *Bf*SPT the interatomic distances between the carboxamide group of N50 and the hydroxylate and carboxylate groups of L-serine are 4.4 Å and 5.0 Å respectfully. While the distances are too large to involve a polar contact in the crystal structure, it is highly likely that the given residue is involved in the reaction mechanism and stabilisation of the transition form.

The similar catalytic mechanism has been proposed to occur in *S.multivorum* SPT by Ikushiro et al., 2009 where the corresponding N52 is involved in coordination of the departing carboxylate of L-serine. Further stabilisation is provided by a distant R367 (corresponding to R365 in *Bf*SPT). The experimental evidence obtained by Lowther et al., 2011 in *S.wittichii* SPT shows that replacement of the corresponding arginine 370 by lysine reduces enzymatic activity down to 3% while, change to alanine yields a completely inactive enzyme.

Upon deprotonation of Cα from L-serine N50 plays part in stabilising the planar quinonoid intermediate by interacting with the carboxylate of the deprotonated serine. The given asparagine residue is one of the key essentially conserved residues in all AONS enzymes where its alteration or removal reportedly results in complete loss of enzyme activity. Notably, the residue N74 is also conserved in *Sp*SPT yet it occupies a position further away from the active site where it is less likely to be directly involved it the reaction mechanism; the position analogical to N50 of *Bf*SPT is occupied by Y73 in *Sp*SPT.

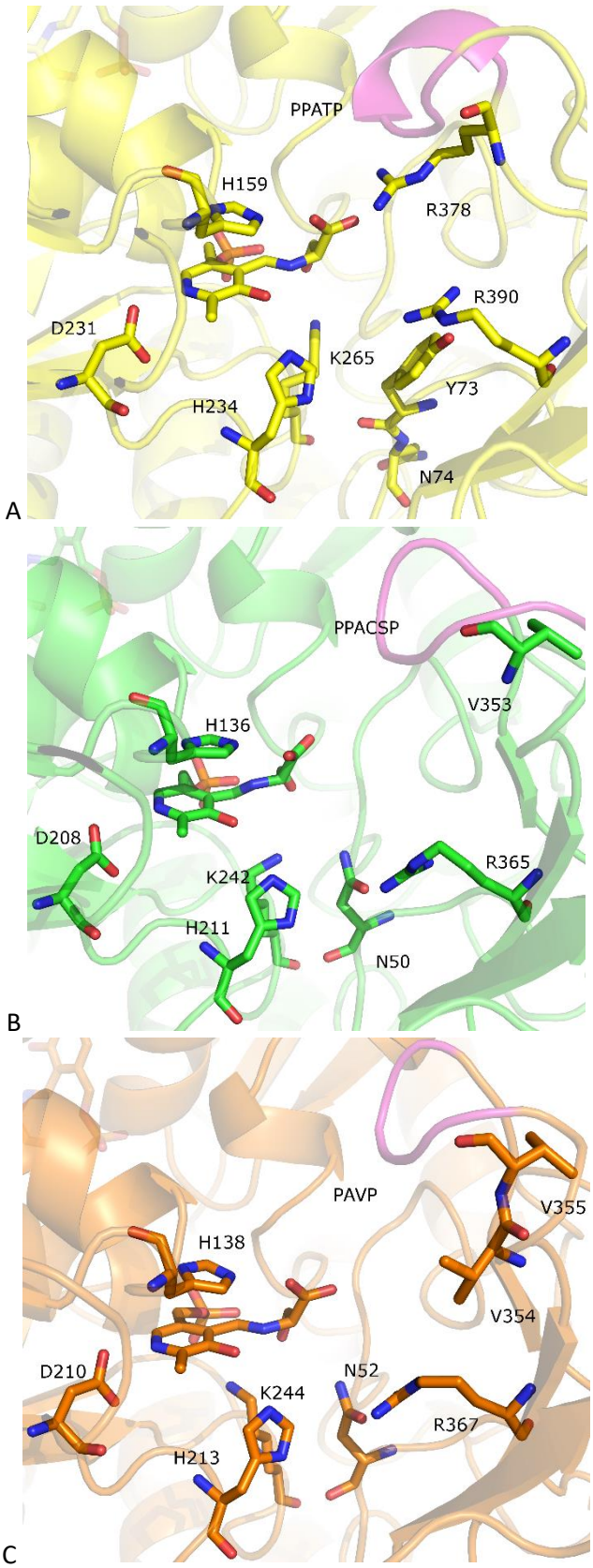


Fig. 3.3.1.4. Comparison of the active sites of L-serine external aldimine forms of *S.paucimobilis* SPT (A, yellow, pdb code 2W8J), *B.fragilis* SPT (B, green) and *S.multivorum* SPT (C, orange, pdb code 3A2B). The variants of "PPAP"-like loop are marked as the purple cartoon on all structures.

It is evident that the L-serine external aldimine form of *Bf*SPT is stabilised by the same mode as in *S.multivorum*; the N50 of *Bf*SPT is oriented towards L-serine yet it does not form a contact with the residue. This is in striking contrast with *Sp*SPT, where R378 performs a conformational shift upon binding of L-serine and provides a stabilising interaction for the carboxylate group of the substrate. No homologous interaction of the same type is present in *Bf* and *Sm*SPTs; analogical residues V353 and V355 respectively are located far from the active site. This might allow additional flexibility for the "PPAP" loops in the enzymes.

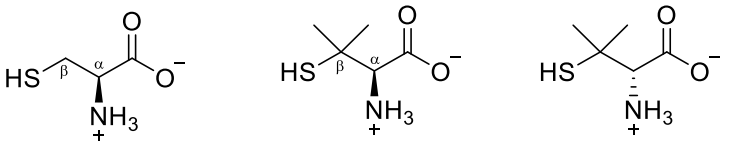
Presence of the N50 in *Bf*SPT can also suggest that the enzyme might be a better model for studying the structural binding of inhibitors compared to *Sp*SPT, as no further conformational changes occur in the structure upon formation of external aldimine (as opposed to the R378 switch in *Sp* enzyme).

The key differences between the active sites of *Bf*SPT and *Sp*SPT highlight the specific strategies that the two enzymes have adopted. The lack of stabilising arginine residue in *Bf*SPT results in a weaker stabilisation of carboxylate in the reaction transition state and therefore directly explains the observed reaction rate differences between *Bf* and *Sp*SPT. At the same time *Bf*SPT might have evolved to keep the substrate selection loop (PPACSP) more flexible; while this reduces the enzyme reaction rate, it might be one of the mechanisms adopted by *B.fragilis* to get a tighter control over the reaction flux (and therefore a controlling step in production of all its sphingolipids). Although the actual evidence with regards to *Bf*SPT is limited, independent experiments performed in our group with similar *P.gingivalis* SPT suggest that the structure of PPACSP loop might indeed have an impact on controlling the second acyl substrate inhibition.

3.3.2. Studies of *Bf*SPT-penicillamine complex formation using UV-vis spectroscopy

Penicillamine (Pen) or β,β-dymethylcysteine is a non-proteinaceous sulphur-containing amino acid. It is present in L and D forms, with D-penicillamine occurring naturally as a degradation product of penicillins (Abraham et al., 1943). Both compounds are general inhibitors of the PLP-dependent enzymes by a formation of thiazolidine adducts with the cofactor (Buell & Hansen, 1960).

D-penicillamine is a weaker PLP inhibitor and therefore approved in use as a drug; the *in vivo* inhibitory effects are mild and can be subverted by administering pyridoxal (Jaffe et al., 1964). It is used as a copper chelator in the treatment of Wilson disease (Das & Ray, 2006) and as a treatment for rheumatoid arthritis (Camp, 1977). L-penicillamine, on the other hand, is a stronger PLP inhibitor and presents toxic effects *in vivo* (Wilson & Vigneaud, 1950) by reducing the brain levels of the PLP-dependent glutamate decarboxylase and thus affecting neurotransmitter levels (Abe & Masuda, 1979).



L-cysteine L-Penicillamine D-Penicillamine

The structure of penicillamine is resembling a structure of cysteine with an additional two methyl groups on the C_β. While cysteine can form a cyclic thiazolidine adduct with PLP, penicillamine forms it more rapidly due to gem-dimethyl effect, and thus acts as a strong inhibitor of the PLP-dependent enzymes.

The inhibitory effects of penicillamine on *Sp*SPT have been studied by Lowther et al., 2012a. Using spectrophotometric UV-vis and end-point inhibition data, it was shown that both enantiomers form complexes with SPT. L-penicillamine had a stronger effect in reducing the activity of the enzyme to 5% at a 5 mM concentration; under the same conditions, D-penicillamine reduced the activity to 34%. The formation of thiazolidine adduct was also reported by observing the expected mass increases using native mass spectrometry.

In this work L- and D-penicillamine studies were performed to explore the binding mechanisms in *Bf*SPT. Studying and comparing binding modes of the two enantiomers allowed to observe certain stereospecific features of *Bf*SPT which are useful for a deeper understanding of the enzyme substrate-binding mechanism. Formation of the penicillamine-derived adducts was monitored by the UV-vis spectrometry; the protein complexes were analysed using native mass spectrometry. X-ray crystallography has also revealed structural information of the PLP-penicillamine binding mode.

The formation of the *Bf*SPT-penicillamine adduct could be directly observed upon monitoring the spectral changes of the enzyme UV-vis profile. The excess amount of L-penicillamine (5 mM) was added to a 20 μM aliquot of the enzyme; changes in the absorbance have been monitored over 30 min with readings taken every 5 min.

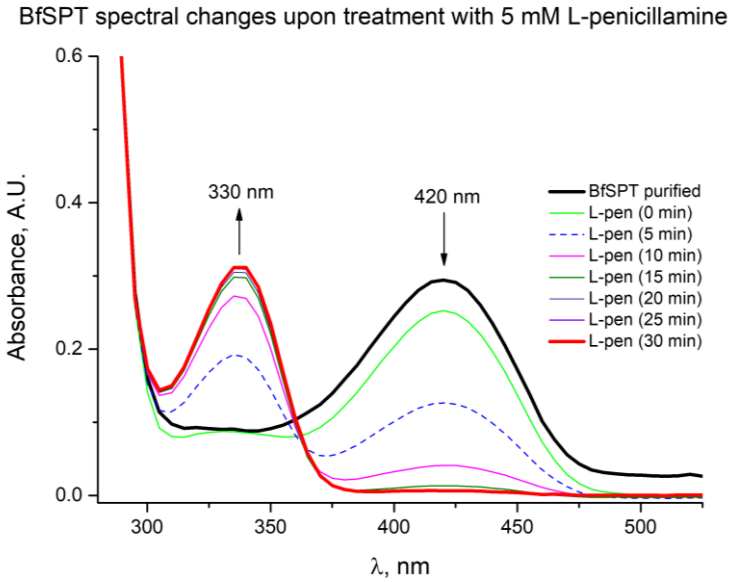
Upon the addition of L-pen, the enzyme profile starts changing rapidly with the increase in the 330 nm peak and the decrease in the 420 nm peak (Fig.3.3.2.1, A).

The change in both peaks is observable from the start at t=0 min; the increase in 330 nm and decrease on 420 nm rapidly passes the middle point after 5 min; further incubation leads to a complete elimination of the 420 nm peak and maximum increase in the 330 nm peak after 30 min. It is important to note that the proposed thiazolidine peak has the absorbance at 333nm, therefore it overlaps with the already existing PLP enolimine peak (330 nm). Therefore, the decrease in 420 nm ketoenamine peak is more informative when considering the enzyme-inhibitor complex formation.

Treatment of *Bf*SPT with 5 mM D-penicillamine under the same condition resulted in similar spectral changes (Fig.3.3.2.1, B), yet these occurred over the much slower time; at the end-point at 35 min, the 330 nm peak was still increasing and the 420 nm peak was decreasing. Upon further time-dependent incubation with D-penicillamine it was shown that the total equilibrium point occurs after a 2 hour incubation time. In this respect, D-penicillamine forms a complex with the enzyme at a slower rate. To determine the approximate binding constants of both enantiomers, equal aliquots of the enzyme at 20 μM were treated with a range of concentrations of L-and D-penicillamine and incubated for 30 mins and 2 hours respectively to reach the end saturation point. The spectra of the enzyme were collected after the incubation time

and plotted together. The changes in absorbance in both peaks were plotted together with the K_d calculated as described in methods.

A



B

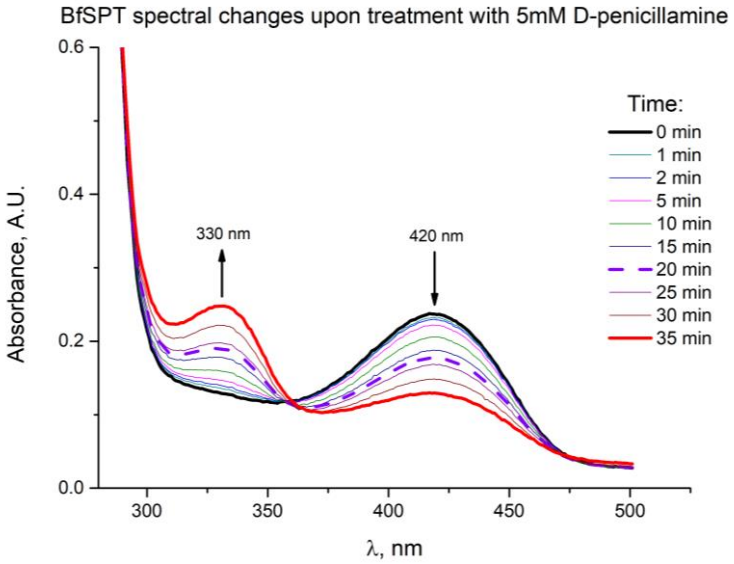
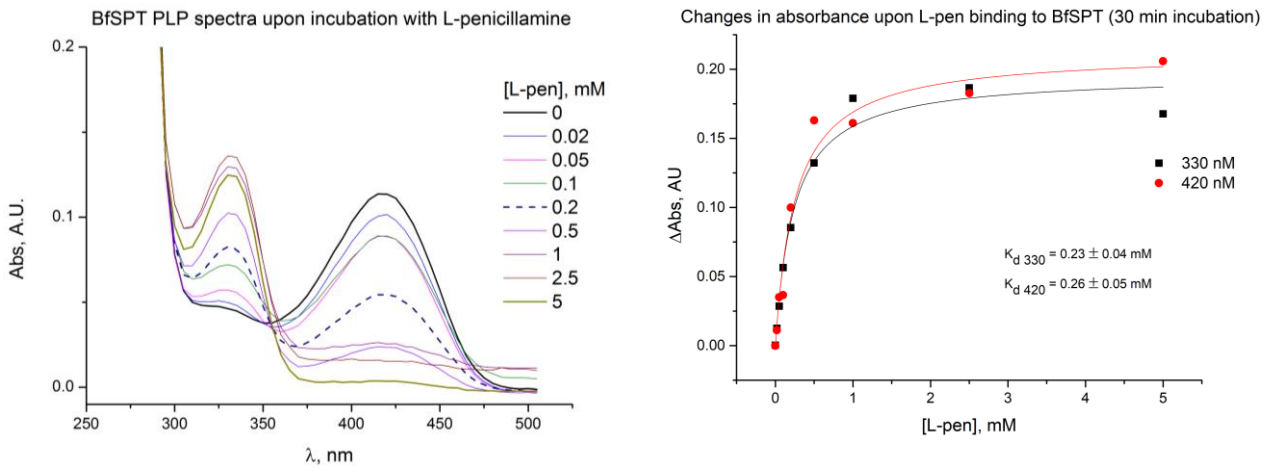


Figure 3.3.2.1. Spectral changes of *Bf*SPT upon time-dependent incubation with L-penicillamine (**A**) and D-penicillamine (**B**)

By observing the changes in absorbance of *Bf*SPT upon treatment with different concentrations of L- and D-penicillamine it was possible to approximate the binding constants for both enantiomers (Fig.3.3.2.2). The K_d values for L-pen were 0.23 mM

as determined by 330 nm peak and 0.26 mM as determined by 420 nm peak. D-penicillamine showed a weaker binding with K_d of 0.31 mM (330 nm) and 0.48 mM (420 nm).

A



B

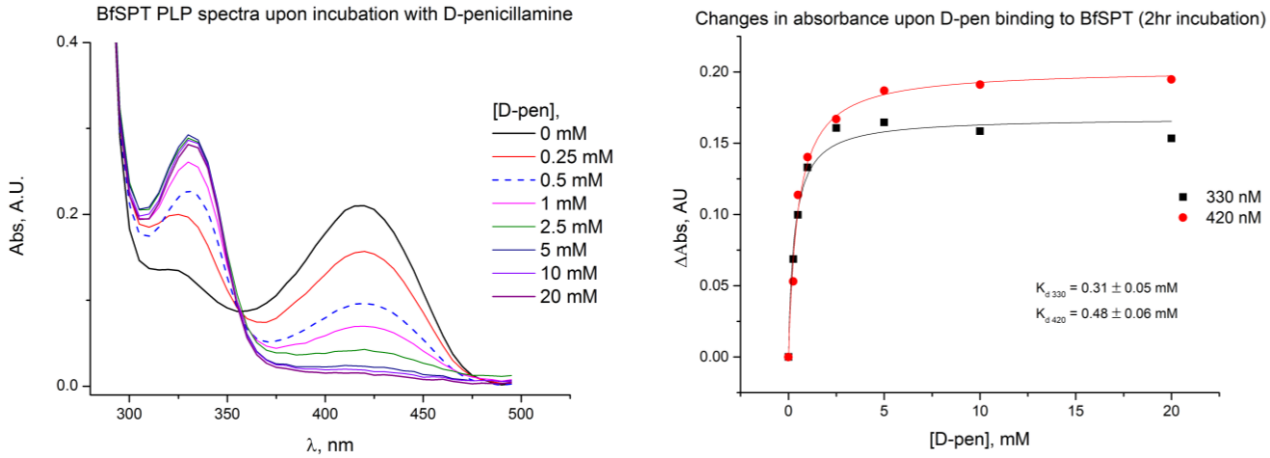


Figure 3.3.2.2 Left: The spectral changes of *Bf*SPT upon the concentration-dependent incubation with L- penicillamine (**A**) and D-penicillamine (**B**). **Right**: Changes in the absorbances at 330 nm and 420 nm peaks of both enantiomers plotted versus the concentrations of each enantiomer.

The discrepancy between the readings at two peaks brings in some important observations that should be addressed with regards to the proposed PLP-penicillamine binding mechanism (Figure 3.3.2.3). Due to an overlap of the thiazolidine peak with

the PLP in enolimine form, it is not possible to distinguish between the two forms and therefore reliably determine the K_d by viewing the 330 nm peak only. Therefore, the K_d values obtained from the readings at 420 nm appear to be more accurate as the reduction in the peak strength corresponds to the conversion of the PLP in the ketoenamine form into the PLP-pen adduct. Therefore, the K_d values of 0.26 mM for L-pen and 0.48 mM for D-pen are considered to be more reliable. The difference in K_d constants shows that the binding of L-penicillamine occurs approximately 1.8 times stronger than of its enantiomer.

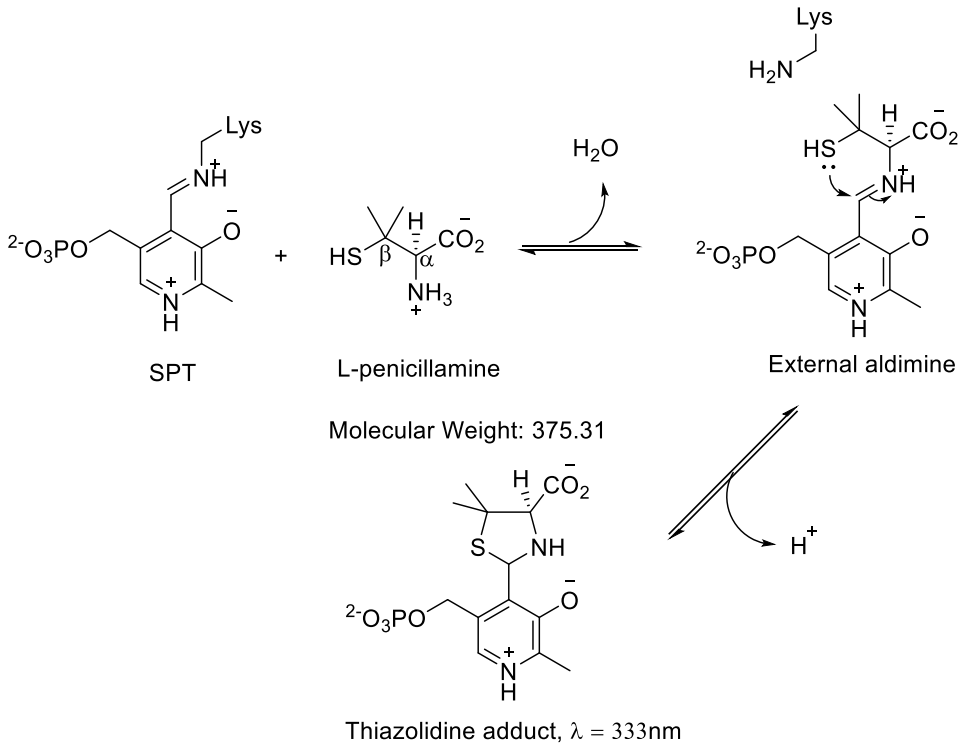


Figure 3.3.2.3. The proposed mechanism of PLP-L-penicillamine binding with the subsequent thiazolidine adduct formation.

The formation of the PLP-thiazolidine complex occurs in two stages (Fig. 3.3.2.3). First, upon arrival into the active site of the enzyme, L-penicillamine replaces the catalytic lysine to form the PLP-external aldimine. During the second stage, the thiol of L-pen processes with the nucleophilic attack on the aldimine carbon, forming a cyclic thiazolidine adduct. The formation of the adduct is further facilitated due to the gem-dimethyl or Thorpe-Ingold effect, where the presence of the two methyl groups bound to Cβ of L-pen causes an increase in a compression angle amongst them and

therefore makes the cyclisation reaction more favourable. The final product is not covalently attached to the enzyme.

The binding mode of D-penicillamine is expected to be very similar to its enantiomer, however, as the observed data show, it binds with a weaker affinity and at a slower rate, making it a less potent inhibitor of *Bf*SPT.

It was not possible to observe the inhibition of *Bf*SPT activity using the DTNB assay due to the interference of the thiol group present in penicillamine with the DTNB reagent. The previous methods by Lowther et al., 2012a, applied to *Sp*SPT involved measuring of the activity using the assay after treatment with both L-and D-penicillamine with subsequent dialysis into a buffer containing no PLP. During the process, the non-covalent adduct was removed from the active site of the enzyme; this allowed monitoring the remaining SPT activity. This approach did not work successfully with *Bf*SPT since the enzyme was prone to precipitation upon dialysis in no-PLP containing buffer after being treated with both L- and D- penicillamine, therefore inhibition data could not be obtained.

3.3.3. Observing the formation of *Bf*SPT-penicillamine complexes by native mass spectrometry

The penicillamine adduct that forms in the active site of *Bf*SPT is not bound to the enzyme covalently, therefore it was not possible to capture it using mass spectrometry under denaturing conditions. Yet some evidence of the complex formation was obtained using native ESI ToF MS. L-enantiomer of penicillamine was chosen as it is a stronger binder to SPT.

*Bf*SPT has proven to be not as robust as *Sp*SPT under the conditions of ESI-MS for native studies. Several attempts have been performed using fresh protein preparation every time; however, the native protein spectrum has been appearing smeared, most likely due to a lower stability of the protein that has been discussed before. This itself provides an implication for the accurate mass determination; however, it still allows observing spectral mass changes of the native protein complexes upon treatment with L-penicillamine.

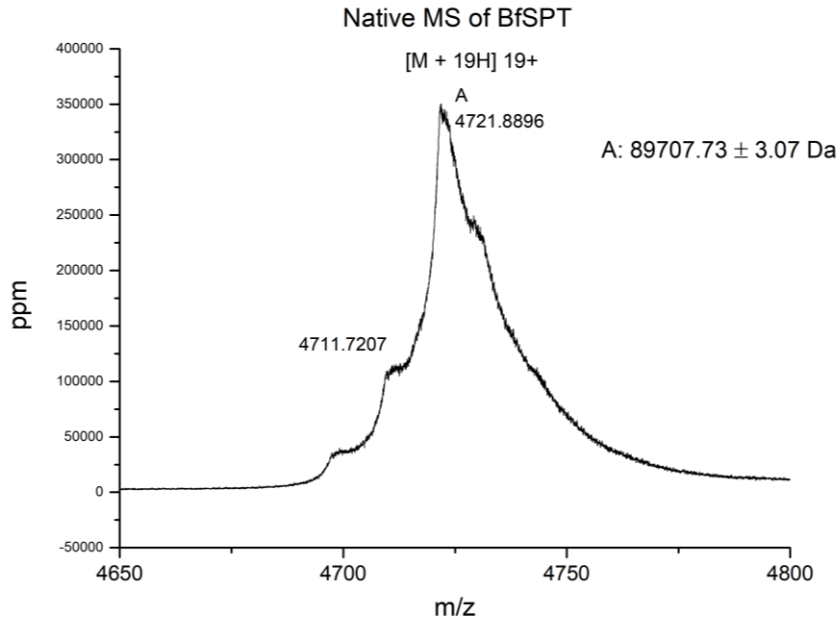
*Bf*SPT treated with 5 mM L-penicillamine for 30 min was analysed using native MS. A non-treated aliquot of the enzyme was used as a control. The obtained spectra are presented in figure 3.3.3.1.

The spectrum of the native *Bf*SPT (Fig.3.3.3.1, A) shows a dominant peak (deconvoluted from m/z charge ion +19) of 89707.73 Da; this corresponds to the monomeric mass of 44853.87 Da within 10 Da difference to the expected mass of *Bf*SPT with the PLP (44843 Da). Two smaller peaks appear that, despite not being resolved, fit within the theoretical mass ranges of *Bf*SPT dimer with no PLP and *Bf*SPT with the cofactor on one monomer only.

The native spectrum of *Bf*SPT treated with L-penicillamine shows a different profile with the several additional peaks being present (Fig.3.3.3.1, B). The major peak A of 4697.7 at +19 m/z corresponds to 89229.2 Da. This agrees well with the expected mass of dimeric *Bf*SPT that has lost PLP in both subunits (89226 Da); the reduction in mass compared to the control sample (89707.7Da) is 478.7 Da or 239.4 Da per monomer. The expected thiazolidine adduct is not bound to the enzyme covalently therefore

under the ionising conditions of the experiment a large fraction of the PLP-thiazolidine adduct is lost from the active site of the enzyme.

A



B

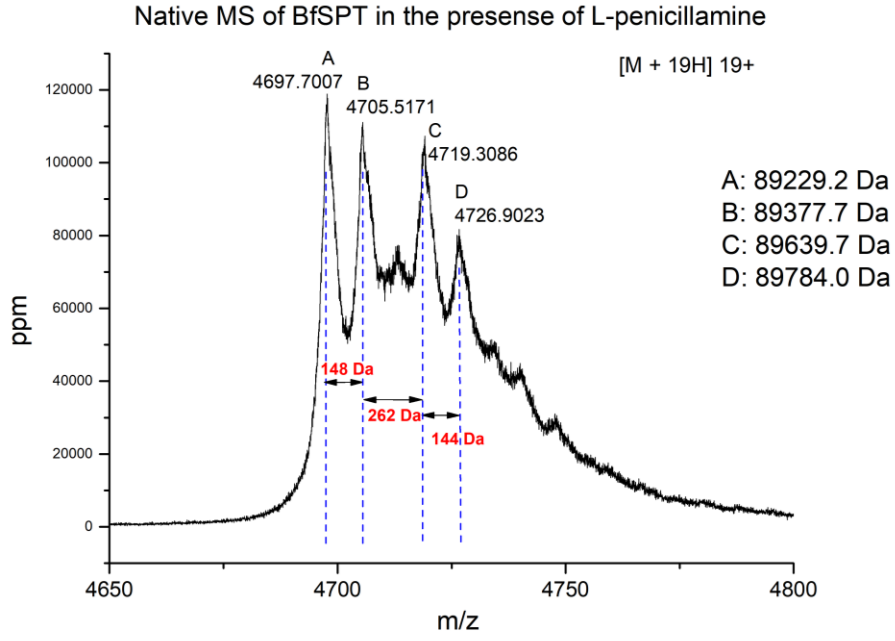


Fig. 3.3.3.1 Native mass spectra of *Bf*SPT (m/z peak [M +19H] 19+). **A**) the control enzyme spectra. **B**) Native mass spectra of the enzyme treated with 5 mM L-pen.

The next peak B at 4705.517 corresponds to 89377.7 Da; the increase in mass is 148 Da, this corresponds to a mass of L-penicillamine, however as the previous peak indicates the enzyme with no PLP, it is unclear how this peak arises; most likely it is yielded due to non-specific binding of single L-pen molecule to a *Bf*SPT dimer.

The next resolved peak C corresponds to 89639.7 Da, giving the increase in 262 Da compared with the peak B and 410 Da with the peak A (enzyme-no PLP). The exact mass of the peak does not correspond to any of the expected species. The most likely adduct to be present within the close mass range would be an enzyme dimer with 2 PLP units on it (89686 Da); yet as the difference between the theoretical and actual peaks is 46.3 Da, it is not possible to resolve the peak with high certainty.

A minor peak appearing in between of the previous two larger peaks could not be accurately resolved and might correspond to a mixture in protein states, including dimer with one PLP and dimer with one PLP unit + L-pen.

The last resolved peak D corresponds to 89784.0 Da which is an increase in 144 Da compared to the peak C; this increase in mass corresponds well to an enzyme dimer with PLP unit forming a thiazolidine adduct with L-penicillamine (theoretical mass increase 144 Da). However, as there are some discrepancies in between the mass of the previous peak C and the control sample it is not possible to deduce that the peak responds to 1 PLP and 1 PLP-thiazolidine unit with high certainty. The remaining smaller peaks at higher mass could not be resolved but they might be indicative of 2PLP-2 thiazolidine complexes and some further non-specific adducts.

While there are some implications in determining the accurate masses of the *Bf*SPT-L-pen complexes, it is possible to observe changes in the mass spectrum of the enzyme upon the treatment with the inhibitor and formation of several intermediate forms; while the spectrum of the native enzyme has 2 PLP units under the ionisation conditions, treatment with L-penicillamine causes major loss of the non-covalent adduct with majority of the enzyme appearing in apo-form; several further forms are present; these most likely include 1PLP, 2PLP, 1PLP-TA; 2PLP-TA; 2PLP-2TA complexes; not all of these could be resolved from the obtained spectra. Additional peak increases of 148 Da suggest that L-penicillamine might also bind non-specifically to the enzyme under the experimental conditions.

3.3.4. Structure of *Bf*SPT in a complex with L-penicillamine

As opposed to the L-cycloserine adduct discussed further in the Chapter 3.3.6, the L-penicillamine-derived thiazolidine complex is stable and therefore, could be easier captured using the X-ray crystallography. The 2.39 Å structure of *Bf*SPT with the PLP-thiazolidine adduct has been obtained by directly soaking *Bf*SPT crystals with L-penicillamine granules for 30 minutes, during which the yellow crystals have turned colourless. Analogical soaking with D-penicillamine was performed, none of the resulting crystals provided a diffraction to a resolution higher than ~3 Å. This is most likely because D-penicillamine is a slower and weaker binder to *Bf*SPT and forms a complex less readily at the same conditions; with some optimisation it might be possible to obtain diffraction-quality crystals with D-penicillamine in future.

The structure of *Bf*SPT complexed with L-penicillamine was solved using the previously obtained structure in the complex with L-serine as the model. The obtained highest resolution was 2.39 Å, with the final refinement values of R_{work}/R_{free} being 0.21/0.25.

The crystallised complex is formed by 4 *Bf*SPT monomers making two homodimers in the asymmetric unit (Fig.3.3.4.1.). The core RMSD between the L-pen and L-ser structures is 0.228 Å across 1572 aligned residues (Coot). The obtained protein structure complex displays the positive electron density in the initial F_o-F_c omit map, well corresponding to the theoretical PLP-thiazolidine adduct (Fig 3.3.4.2); the PLP-L-pen ligand was fitted in and used for a refinement. The refined density model is shown in Fig. 3.3.4.2, middle. The density for the complex is clearly present in all subunits, although the subunit B does lack the density around the carboxylate of L-penicillamine. For display and discussion, the subunit chain C was chosen. The shape of the ligand clearly shows the presence of the ring-closed thiazolidine complex with the two isomethyl groups at Cβ and the chiral carboxylate group of L-penicillamine well defined.

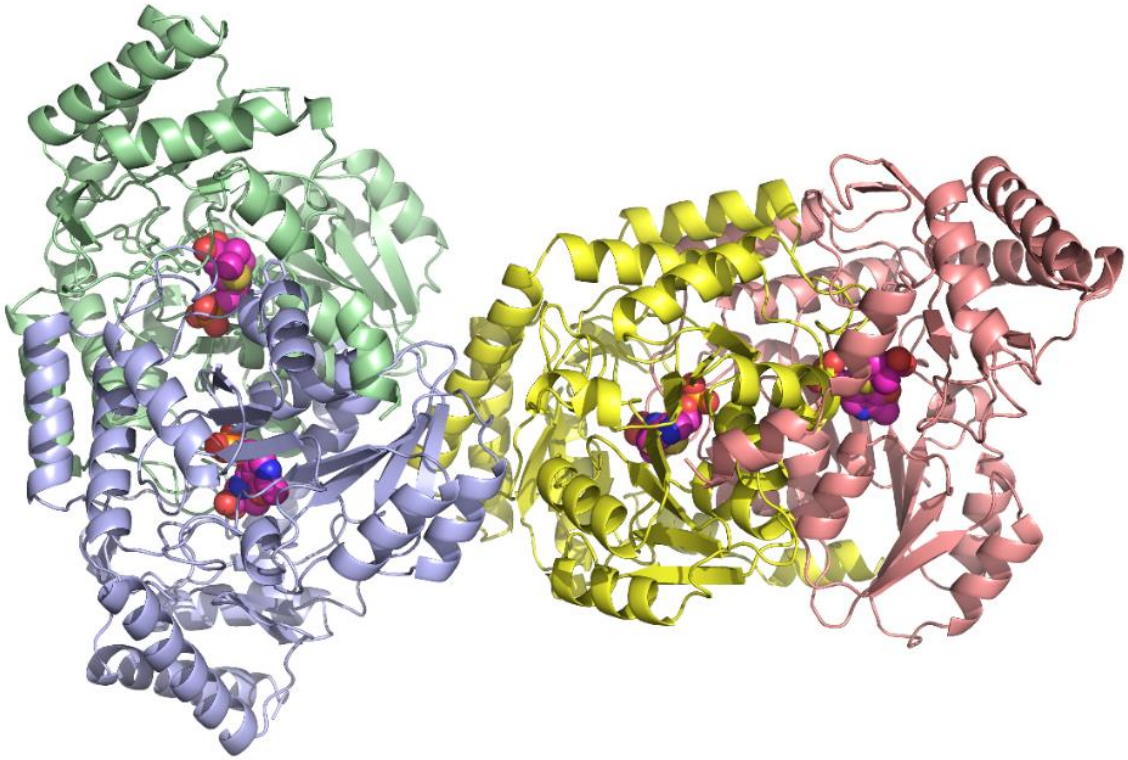
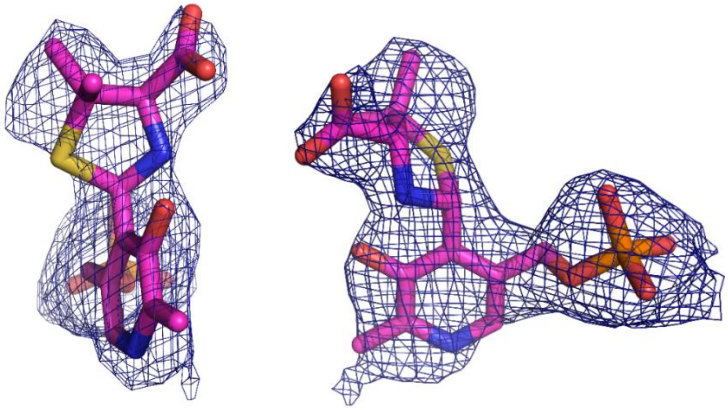
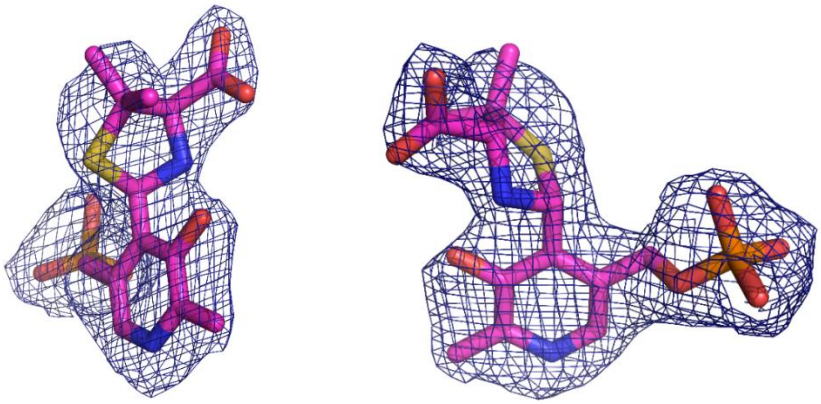


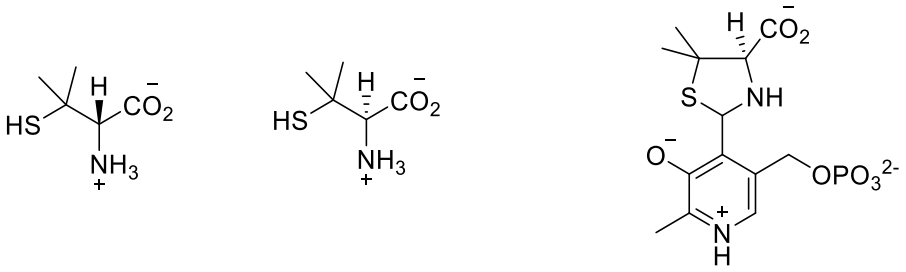
Fig. 3.3.4.1. The 2.39 Å *Bf*SPT structure obtained by soaking the protein crystal with solid L-penicillamine. The positions of thiazolidine adducts are shown in purple spheres. The asymmetric unit consists of 4 chains forming two homodimers AB and CD. Chains are coloured as follows: A – turquoise, B – slate blue, C – yellow, D – pink.



3.0 σ F_o-F_c omit map



1.2 σ $2F_o$-F_c refinement map



D-penicillamine L-penicillamine PLP-L-pen thiazolidine adduct

Figure 3.3.4.2. Ligand density maps for 2.39 Å *Bf*SPT structure obtained by soaking with L-penicillamine, shown in two different orientations (top and bottom); protein chain C is shown. **Top**: The original F_o-F_c map (contoured at 3.0 σ, 1.7 Å cut off) that had no ligand modelled in, with the structure of the PLP-thiazolidine complex rendered on top. **Middle**: The final $2F_o$-F_c map obtained after refinement with the PLP-thiazolidine adduct (contoured at 1.2 σ, 1.6 Å cut off). **Bottom:** Chemical structures of D- and L- penicillamine and the PLP-thiazolidine adduct forming upon L-penicillamine binding.

While there is no direct contact of L-serine with the N50 in the external aldimine structure of *Bf*SPT (fig 3.3.4.3, B), the L-penicillamine-derived thiazolidine adduct forms a 2.96 Å hydrogen bond between the carboxylate oxygen of L-penicillamine and the carboxamide nitrogen of N50 (Fig.3.3.4.3, A). Furthermore, there is a 3.57 Å polar contact between the second carboxylate oxygen of the thiazolidine adduct and the imidazole nitrogen of H211 (Fig. 3.3.4.3, A) that further stabilises the complex. Both these contacts are not present in the L-serine external aldimine structure (Fig 3.3.4.3, B).

Both N50 and H211 residues are essentially conserved between all the AOS-family enzymes; the presence of these contacts in the inhibitor complex further highlights their potential involvement in a stabilisation of the transition complex during the course of the reaction. The structural evidence itself suggests that in the Dunathan conformation, L-serine performs a swing in the active site upon binding of the second substrate, where the carboxylate group would be positioned in a similar manner as seen in the thiazolidine complex.

While we have no structural data for the D-penicillamine binding mode, it is viable to propose that the carboxylate group of the latter would be placed in the opposite orientation than of L-penicillamine, preventing the contact with the carboxamide of the N50 and the imidazole of H211. This would result in a weaker binding of the adduct in the active site. This is the most likely reason for the slower adduct formation, observed during the UV-vis studies, as well as generally weaker levels of inhibition reported for the D-enantiomer with the PLP-dependent enzymes.

The observed PLP-thiazolidine complex serves as a structural model for other PLP-dependent enzymes present in the human body reportedly inhibited by L-penicillamine, and the human SPT in particular. In this respect, the complex provides a structurally-based evidence of the inhibitory effect of L-penicillamine, and despite the lack of D-penicillamine complex, allows discussing the variety in the binding modes of the two enantiomers and hence explains the observable difference in their biological activities.

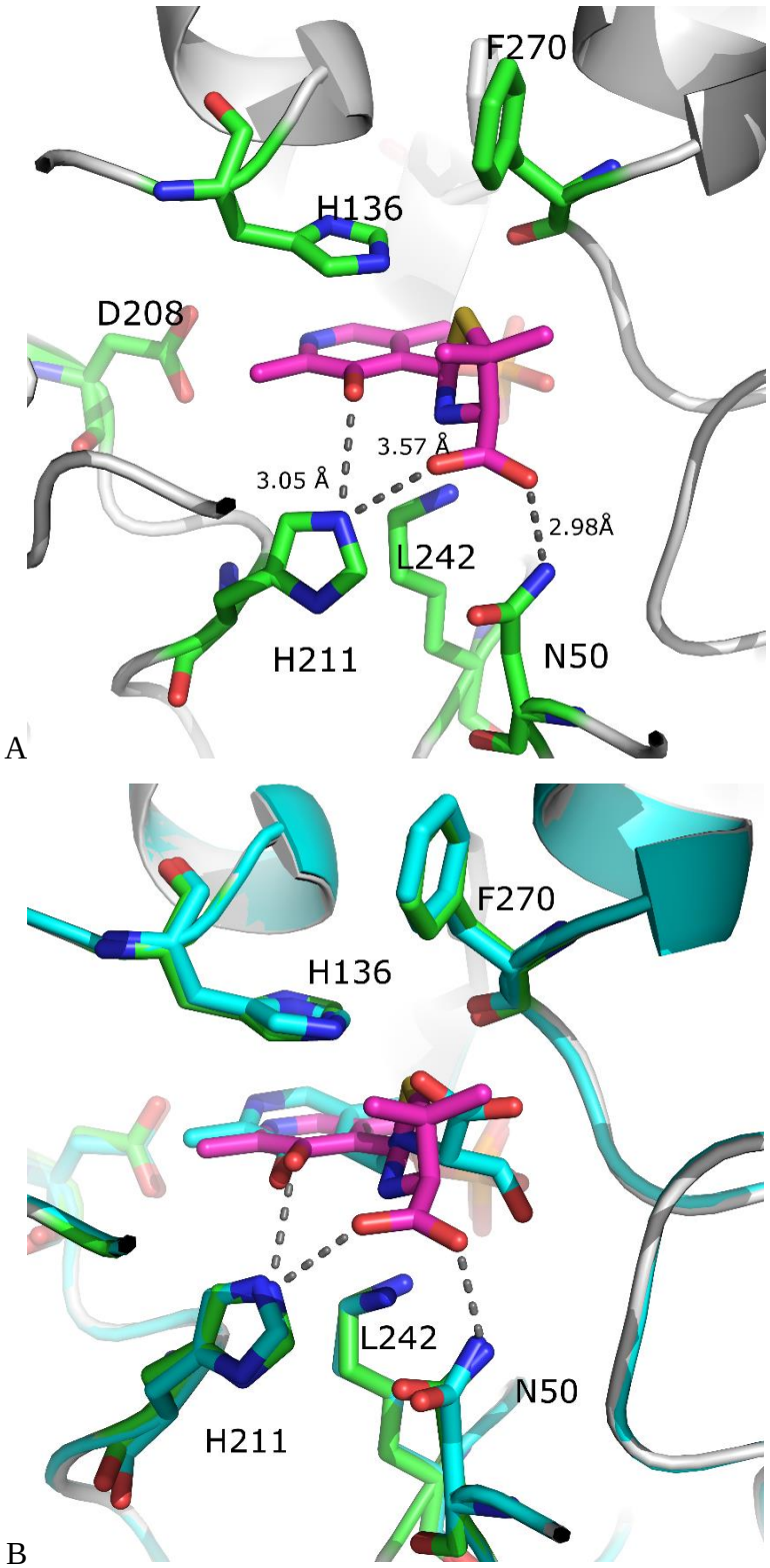


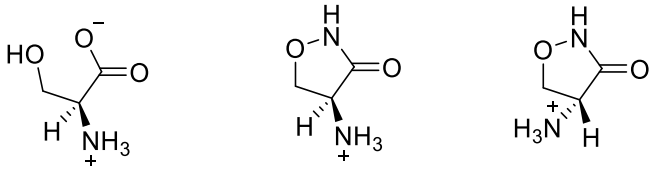
Figure 3.3.4.3.: **A**) The binding mode of L-penicillamine derived thiazolidine adduct in the active site of *Bf*SPT. The PLP-thiazolidine ring complex is stabilised by the carboxylate group contacts to the N50 and H211. **B)** An overlay of the active sites of the PLP-L-serine external aldimine structures (residues and ligand in cyan) and the L-penicillamine derived thiazolidine complex (ligand in purple, residues in green). The thiazolidine carboxylate is positioned differently to the L-serine carboxylate and makes contacts with N50 and H211.

3.3.5. Inhibition of *Bf*SPT with L-cycloserine

D-cycloserine (trade name Seromycin^{TM}) is a natural amino acid derivative used in the second line of treatment against the drug-resistant *Mycobacterium tuberculosis* (Perri & Bonora, 2004). By mimicking the shape of D-alanine and D-glutamine it acts as an inhibitor targeting D-aminotransferase enzymes involved in peptidoglycan biosynthesis (Peisach et al., 1998), thus effectively disrupting the synthesis of bacterial cell wall components. It also acts as a partial agonist of N-methyl D-aspartate receptor (NMDA) in nerve cells and has potential as a treatment in neuropsychological disorders (Schade & Paulus, 2016).

L-cycloserine (LCS) is a synthetic enantiomer that shows similar inhibitory effects on aminotransferases; however, unlike DCS, the potential to use LCS for medical treatment is limited due to a non-specific inhibition of other pyridoxal 5'-phosphate-dependent enzymes, in particular, this affects mamal serine palmitoyltransferase (Williams et al., 1987). Nevertheless, due to its high binding affinity to PLP, LCS is a very useful molecule for studying mechanisms of the PLP-dependent enzymes.

The earlier study by Lowther et al., 2010, has described effects of prolonged deactivation of serine palmitoyltransferase from *Sphingomonas paucimobillis* by L-cycloserine, leading to the formation of a product identified as β-aminooxyacetalaldehyde and a reduced pyridoxamine 5'-phosphate (PMP). To further supplement the earlier data and elucidate the early stages of deactivation mechanism, spectral studies of *Bf*SPT deactivation with LCS have been performed. These were further complemented by obtaining the crystal structure of the enzyme in the presence of L-cycloserine.



L-serine L-cycloserine D-cycloserine

The UV spectrum of *Bf*SPT spectrum has been observed upon incubation with 2.5 mM LCS over the period of 30 mins. The spectral shift over the time indicated changes in the equilibrium between the PLP isoforms as binding of LCS occurred (Fig 3.3.5.1).

The absorbance peak at 420 nm (ketoenamine) decreased over the 30 mins being reduced almost to baseline; the enolimine peak (330 nm) permanently increases; the peak at 380 nm starts appearing upon further incubation. The peak appearing at 380 nm (evident from 15-30 min) is thought to be the ring-opened oxime intermediate (Ikushiro et al., 2004; Lowther et al., 2010) – a reversible complex that has a faster formation rate than the final inhibitory complex, causing its build up.

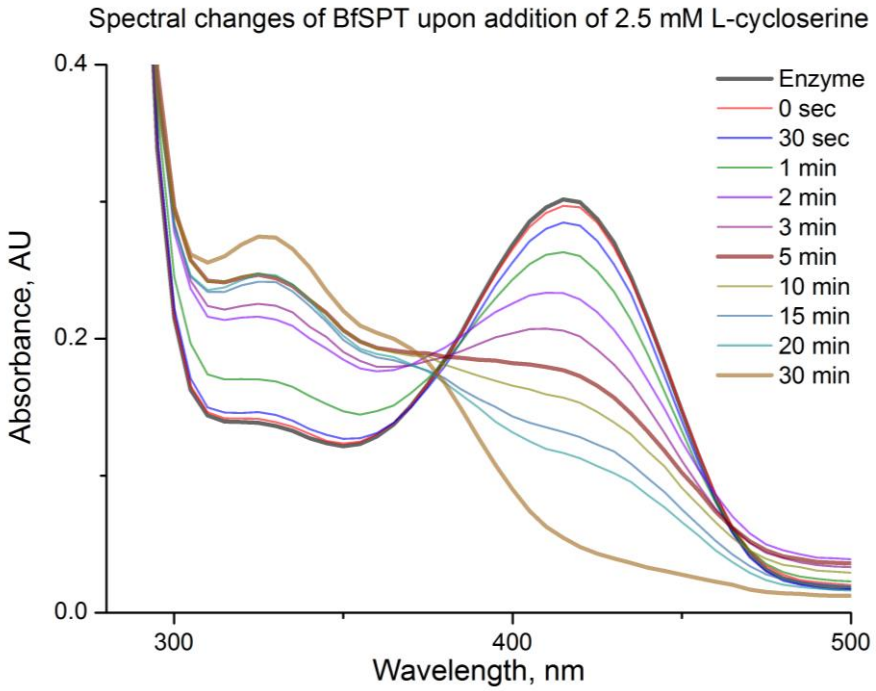


Fig. 3.3.5.1. Changes in the spectral profile of *Bf*SPT over time upon addition of 2.5mM LCS.

To assess the inhibitory effects of L-cycloserine in *Bf*SPT, the activity of the enzyme was monitored by the DTNB assay after a time-dependent incubation of the enzyme with various concentrations of LCS. The reactions were done in triplicate, the maximum averaged activity of *Bf*SPT in the control samples (no LCS) was presented as 100%, with other relative activities presented as the % fraction of it (Fig 3.3.5.2).

The LCS has shown a strong inhibitory effect on the activity of *Bf*SPT even at small concentrations and over a low incubation period, with apparent K_i values being at the micromolar level: 9.0 ± 2.5 μM for 1 min incubation, 8.4 μM ± 2.1 μM for 5 min incubation and 1.7 ± 0.4 μM for 30 min incubation. The results show that LCS is a

very potent inhibitor of *Bf*SPT, with the excess concentrations starting from 0.1 mM and higher even after a very short incubation time.

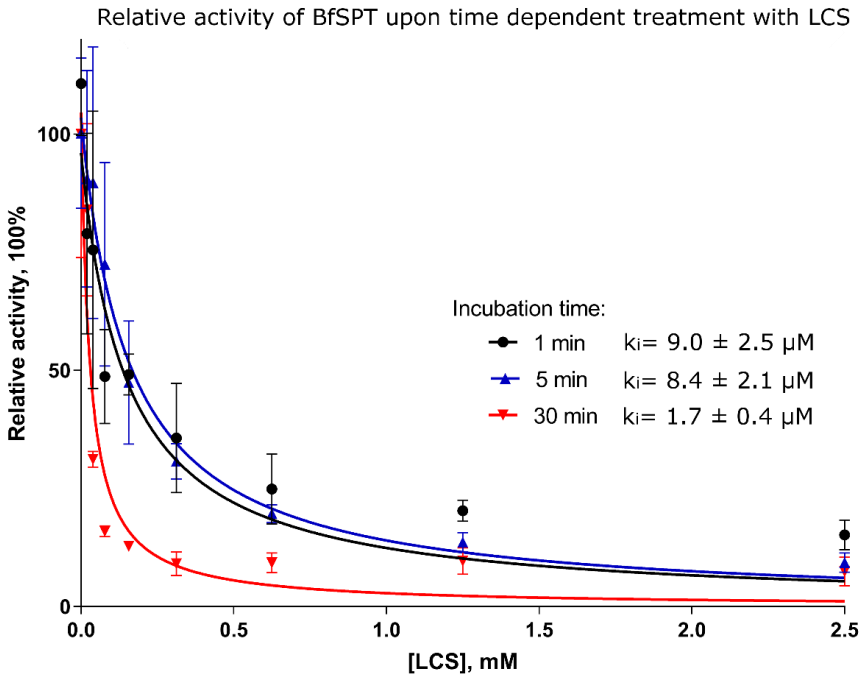


Fig. 3.3.5.2. Relative activity of *Bf*SPT (determined from measured initial reaction rates by the DTNB assay), that has been prior incubated with various amounts of LCS, over time; the black, blue and red lines correspond to the samples incubated for 1, 5 and 30 mins respectively. Reactions were done in triplicate with the error bars presenting the standard deviation between the samples.

It is important to note that the experiment was designed to provide an estimation of the inhibitory concentrations of LCS to be used with *Bf*SPT and not to determine the true parameters of the enzyme inhibition. In the last case, the true kinetic parameters could not be established under the available conditions since the reaction is processed using the PCoA and not the expected isopentadecanoyl-CoA. Furthermore, the concentration of PCoA is set at 25 µM at which the optimal rate of the reaction is reached, while the impact of the substrate's own inhibition is also unaccounted for.

3.3.6. Structural studies of the L-cycloserine binding mode to *Bf*SPT

Upon monitoring of the binding and inhibition, structural studies of *Bf*SPT and LCS have been performed. The most successful crystal obtained by a soaking method provided diffraction up to a maximum resolution of 2.75 Å.

The structure was solved using the *Bf*SPT L-serine structure with no PLP as a template for a molecular replacement. As in the previous *Bf*SPT structures, the complex consists of 4 enzyme chains in an asymmetric unit forming two separate homodimers. The core RMSD between the structure and the *Bf*SPT L-serine complex is 0.544 Å across 1572 aligned residues in 4 chains (Coot).

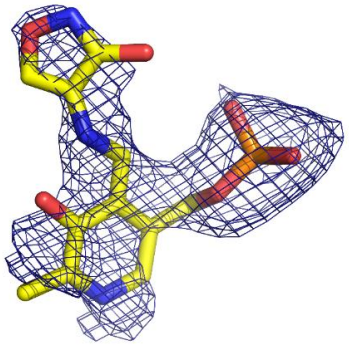
After initial refinement without the PLP cofactor modelled, the F_o-F_c map had shown weak density signals for the ligand in the active sites of most of the chains, with only the phosphate ion and minor parts of the PLP chain visible. The density in the chain C, however, corresponded well to shape the PLP aromatic ring, phosphate group and the likely cycloserine-derived adduct visible at 3 σ - contour level of the unbiased F_o-F_c map.

The initial density observed around the ligand binding site was not conclusive enough to determine the exact nature of the LCS adduct and the mechanism of its formation. Resolution of several possible adducts has been attempted based on the proposed mechanisms of LCS deactivation of the PLP-dependent enzymes and SPT in particular (Fig 3.3.6.1).

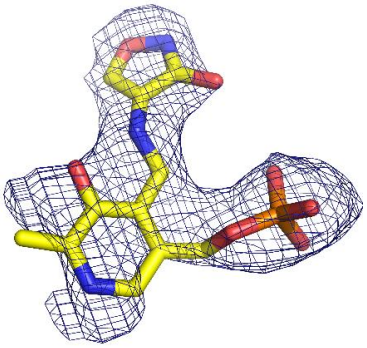
3 possible adducts were fitted into the active site (Fig 3.3.6.1): the cyclic 3-hydroxyisoxazole-PMP (A), PLP-oxime (an open ring form, B) and PLP-decarboxylated oxime (C), a decarboxylated form of the oxime. The structures of the ligands fitted into the active site of *Bf*SPT are shown in figure Fig 3.3.6.5 later in the discussion.

The analysis of the obtained R values and temperature B-factors indicates that out of the three complexes, the PLP-decarboxylated oxime complex model (C) is the one that most closely corresponds to the experimental data, while the hydroxyisoxazole-PMP adduct (A) is the least likely. However, relatively low resolution limits the possibility of making a definite conclusion about the exact nature of the ligand complex observed.

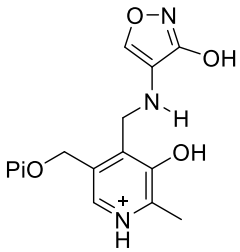
Structures of possible PLP-cycloserine derivatives



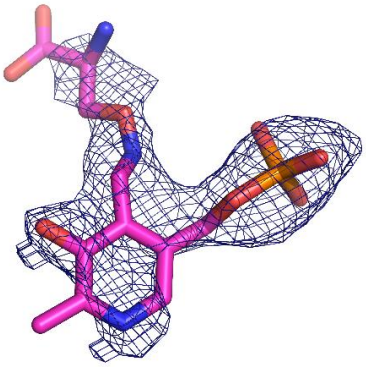
2.5σ (F_o-F_c)



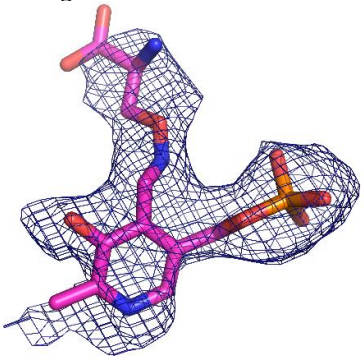
2.1 σ (F_o-F_c)
R_{work}/R_{free} = 0.282/0.329
Av. ligand B factor = 60.6



3-hydroxyisoxazole-PMP (A)

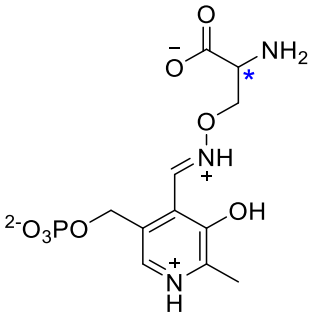


2.5σ (F_o-F_c)

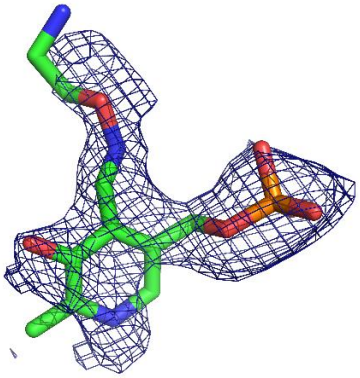


2.1 σ (F_o-F_c)

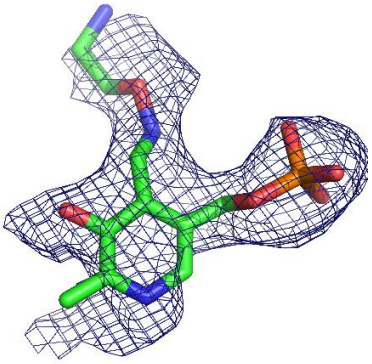
R_{work}/R_{free} = 0.277/0.323
Av. ligand B factor = 62.6



Oxime complex (B)

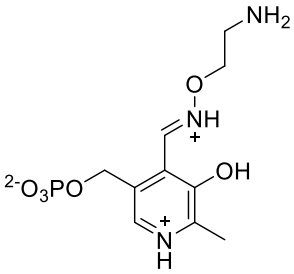


2.5σ (F_o-F_c)



2.1 σ (F_o-F_c)

R_{work}/R_{free} = 0.276/0.325
Av. ligand B factor = 56.8



Decarboxyoxime complex (C)

Fig. 3.3.6.1. The F_c-F_o electron density maps around the active site of *Bf*SPT-LCS soaked complex. The figures to the left present the same initial unbiased F_o-F_c electron density map contoured at 2.5 σ level. The figures to the right show biased F_o-F_c electron density maps contoured at 2.1 σ level, obtained after independent refinements with each of the three possible complexes. The obtained values of R_{work}/R_{free} and average ligand B-factors are shown for each of the proposed ligands. The possible adducts include hydroxyisoxazole-PMP (yellow, A), PLP-oxime complex (purple, B) and PLP-decarboxylated oxime complex (green, C).

Mechanisms of LCS deactivation of PLP-enzymes

There are numerous examples of the PLP-cycloserine complexes in the Protein Data Bank (Appendix 4), these include alanine racemases (Fenn et al., 2003; Noda et al., 2004; Asojo et al., 2012), decarboxylases (Malashkevich et al., 1999), lyases (Kuznetsov et al., 2015), D-amino acid aminotransferases (Peisach et al., 1998) and D-cysteine desulfhydrase (Bharath et al., 2012). Apart from the serine palmitoyltransferase structure presented by our group (Lowther et al., 2010; pdb id 2XBN), all the structures report a closed-ring oxazole adduct bound to the cofactor in the active site; it is present in either PLP-3-isoxazolidinone or PMP 3-hydroxyisoxazole forms.

Upon formation of the PLP-cycloserine external aldimine, a transamination reaction occurs, resulting in the formation of 3-hydroxyisoxazole PMP adduct (Fig.3.3.6.2); during this, the hydrogen is lost from the chiral Cα leading to a configurational change from sp^3 to sp^2 and adoption of the planar geometry by the ligand. The resulting adduct has a reported absorbance peak at 320 nm (Fenn et al., 2003). No mechanism involving cycloserine ring-opening has been reported for any of the aminotransferase enzymes.

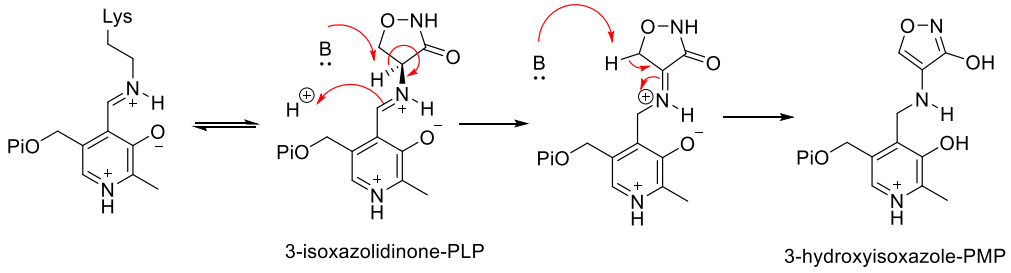


Fig. 3.3.6.2. The mechanism of cycloserine binding to PLP-dependent enzymes; the cyclic adducts appear in most of the published protein structures (summarised in Appendix 4).

Malashkevich et al., 1999, present structures of dialkyglicine decarboxylase from *Burkholderia cepacia* with both LCS (pdb:1D7U) and DCS (pdb:1D7S) bound. Both structures solved at high resolution (1.95 and 2.05 Å respectively) present the PLP-oxazole ring cyclic complexes; yet the LCS-derived complex adopts a planar ring configuration indicating deprotonation at the Cα, while the DCS-derived complex retains the sp^3 configuration at the Cα. Fenn et al., 2003, report that in the alanine

racemase from *Geobacillus stearothermophilus* a transamination reaction occurs in both enantiomers leading to a formation of a planar product (hydroxyisoxazole-PMP).

Contrary to most of the other classes of PLP-dependent enzymes that have been shown to form cyclic adducts with cycloserine, SPT-LCS complex is proposed to form via different, ring opening mechanism (Fig 3.3.6.3); the adduct is either rapidly and reversibly converted to a sterically preferable oxime intermediate, with the characteristic absorption at 380 nm, or slowly hydrolysed to yield the PMP and a product, β-aminooxopyruvate (Ikushiro et al., 2004).

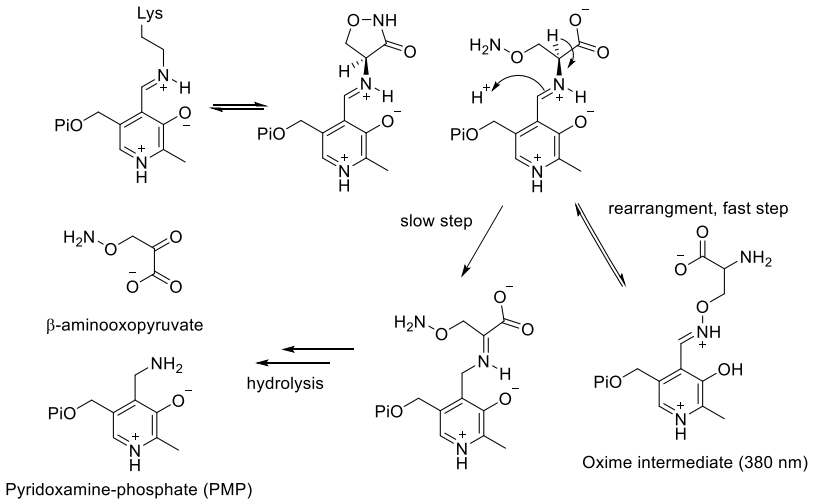


Fig. 3.3.6.3. The proposed mechanism of cycloserine inhibition of SPT (Ikushiro et al., 2004)

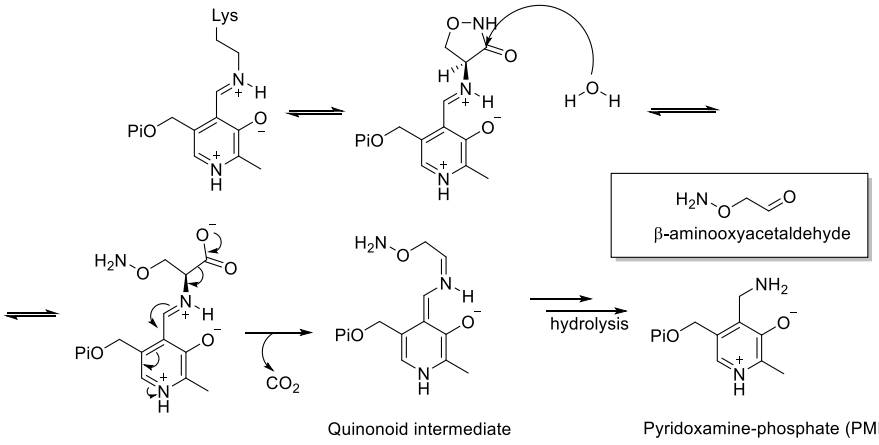


Fig. 3.3.6.4. The proposed mechanism of LCS inhibition of *Sp*SPT observed by Lowther et al., 2010. As opposed to the earlier mechanism by Ikushiro et al., 2004, the reaction involves decarboxylation step after cycloserine ring opening; this results in the formation of β-aminooxyacetaldehyde and PMP.

The structural evidence obtained in *Sp*SPT by Lowther et al., 2010, suggests that in *S.paucimobilis* SPT a ring-opening mechanism occurs; however, the obtained crystallisation product corresponded to β-aminooxyacetaldehide, showing that loss of the product carboxylate group had occurred (Fig.3.3.6.4). This allowed proposing a novel inhibitory mechanism present in *Sp*SPT that involves a decarboxylation step before the product release.

Provided that LCS similarly deactivates *Bf*SPT, formation of the adduct would occur via the ring opening mechanism; this would lead to a very slow product release and accumulation of the oxime-PLP intermediate. As the UV-vis spectrum of *Bf*SPT shows the appearance of a characteristic 380 nm peak upon incubation with LCS, it is likely that the enzyme deactivation occurs via the ring-opening mechanism; this might also involve a decarboxylation step as in *Sp*SPT.

The nature of the observed LCS-derived adduct

As the mechanism of PLP inactivation by LCS in *Bf*SPT may take several routes, different enzyme-inhibitor complexes could potentially be present in the active site. Resolving of the structure was attempted with 3 adducts: the closed-ring 3-hydroxyisoxazole-PMP (A) which has been reported for most of the PLP-dependent enzymes in the pdb (Appendix 4), the PLP-oxime complex (B) as based on the mechanism of SPT inhibition outlined by Ikushiro et al., 2004, and the PLP-decarboxylated oxime (C), based on the structural data presented by Lowther et al., 2010.

The active sites of the refined model complexes are presented in Figure 3.3.6.5.

The closed-ring cycloserine adduct has been fitted as 3-hydroxyisoxazole-PMP (Fig.3.3.6.5, A). Upon formation of the PLP-3-isoxazolidinone adduct, a transamination reaction can occur yielding a change in Cα of LCS from sp^3 to sp^2 hybridisation state; this would give a formation of the planar 3-hydroxyisoxazole – PMP adduct and loss of chirality of the molecule. Due to a low resolution, it is not possible to distinguish between the 3-hydroxyisoxazole and the 3-isoxazolidinone forms in the structure. The shape of the density, however, suggests that the PLP-

nitrogen bond does not lay in the plane of the pyridoxal ring, therefore a presence of the transaminated planar product is more likely. The model least agrees with the experimental data with the R_{work}/R_{free} values of 0.282/0.329 and the average ligand temperature B-factor of 60.6. Furthermore, the appearance of the observed 380 nm peak upon treatment of *Bf*SPT with LCS may suggest that the ring-opening mechanism is present in the enzyme, thus making the cyclic complex one of the least likely to be present.

The proposed PLP-oxime complex (Fig 3.3.6.5, B) fits into the general shape, yet the subsequent refinement does not cover the carboxylate group completely; the orientation of the group also appears less favourable as under the density restraints it does not make a stabilising contact with H136. Both L and D enantiomers of the oxime complex were fitted, yet neither of those was entirely covered by the electron density. This model better agrees with the experimental data with the R_{work}/R_{free} = 0.277/0.323 and the average ligand temperature B-factor of 62.6. However, it would remain highly likely that the carboxylate group of the intermediate B would make a polar contact either with H136 or N50 of *Bf*SPT; this would result in a stronger electronic density around the carboxylate ion. As these theoretical estimations are not observed in the model, it remains less likely that the oxime complex is present in the crystal structure.

The apparent density, however, corresponds better to the shape of the theoretical decarboxylated oxime complex (Fig. 3.3.6.5, C) that is likely to form in *Bf*SPT in analogy with *Sp*SPT as observed by Lowther et al., 2010. Based on the values of R_{work}/R_{free} = 0.276/0.325 and the average ligand temperature B-factor of 56.8, the complex model is the one that best agrees with the experimental data. The limited resolution, however, does not allow making a definite conclusion about the actual ligand present in the active site and hence discussing the exact mechanism of the *Bf*SPT inactivation by L-cycloserine.

The results, however, indicate that the inhibitor soaking approach works well for *Bf*SPT and can be potentially used to obtain a better resolution data. Small molecule mass spectrometry can be used in combination with the structural studies to elucidate the exact mechanism of LCS inhibition in *Bf*SPT.

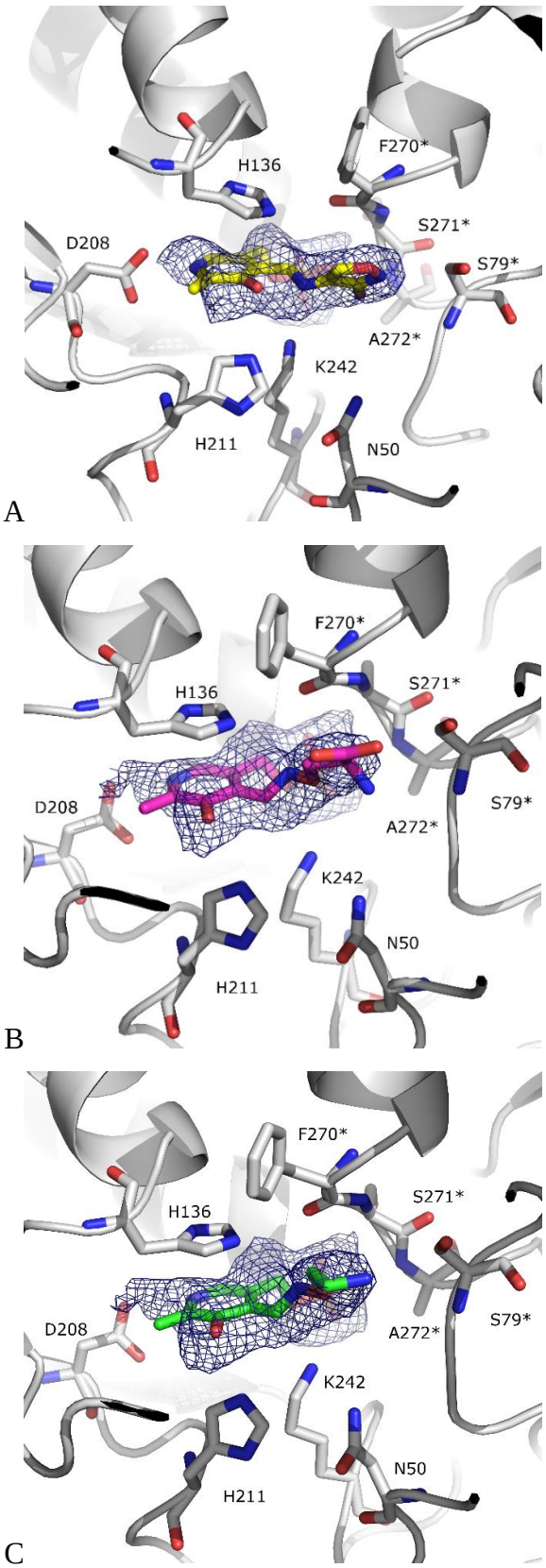


Fig 3.3.6.5. Three possible LCS-derived complexes rendered on top of the corresponding refined 2.1 σ (F_o-F_c) electron density maps in the active sites of the obtained 2.75Å structure of *Bf*SPT. Yellow: hydroxyisoxazole-PMP (A); Purple: oxime (B); Green: decarboxylated oxime (C).

3.3.7. Structural conservation amongst microbiome-related iso-branched SPTs

B.fragilis is a typical example of *Bacteroides*, one of the five genera of microbiome-related bacteria that are associated with a production of iso-branched SLs, the other four being *Parabacteroides*, *Prevotella*, *Tannerella* and *Porphyromonas* (Kato et al., 1995; Nichols et al., 2011). By producing unique iso-branched SLs, the bacteria from these genera can impact the host's immune system and therefore are directly relevant to human health. Being SL-producing organisms, these species are expected to have an SL-producing pathway starting with SPT. Further significance of these isoSL-producing organisms in the human microbiome was recently reviewed by Heaver et al., 2018.

In *B.fragilis*, SL-deficient bacteria have been produced by growing cells in the presence of myriocin (An et al., 2011) or knocking out the gene BF2461 (Wieland Brown, 2013; An et al., 2014) that has been confirmed to encode for *Bf*SPT by this study. Identical effects have been achieved in *P.gingivalis* upon deletion of the homologous PG1780 gene, resulting in SL deficiency and impaired survival of the strain (Moye et al., 2016). It is likely that SL metabolism could also be controlled at the SPT level in the rest of the isoSL-producing microbiome-related bacteria.

*Bf*SPT is the first microbiome-related SPT associated with a production of iso-branched SLs to be structurally characterised. To identify other potential isoSPT candidates and assess their similarities a protein-sequence bioinformatics approach was used. A protein BLAST (Altschul et al., 1990) search was performed against the other four bacterial taxa using the WT sequence of *Bf*SPT (Uniprot ID Q5LCK4) as an input; this allowed to identify four proteins for each of the taxa with a very high sequence identity (74.3% on average; summary of hits shown in table 3.3.7.1; the protein sequences are presented in Appendix 5). Notably, the highest matching identity scores were picked for each respective genus; variation of putative SPTs between species within each taxon was not assessed.

The average sequence identity in various AOS-family proteins amongst different species remains within 30-35%; the average sequence identity amongst the four structurally characterised bacterial SPTs (*Sp*, *Sm*, *Sw* and *Bf*) is 42.4%.

The high sequence conservation amongst the five proteins identified by the BLAST search (average 74.3%) and association of the query species with the production of iso-branched SLs allows consideration of the hit results as putative isoSPTs.

Species	GeneBank ID[1]	UniProtID[2]	% similarity (*Bf*SPT)
B.fragilis	CAH08161.1	Q5LCK4	-
Porphyromonas gingivalis	WP_021679668.1	W1R7E5	76%
Parabacteroides goldsteinii	WP_010802415.1	S0GTM6	77%
Prevotella sp.	CDE86132.1	R7LFZ5	73%
Tannerella sp.	CCY38459.1	R5IDR1	76%

[1] Benson et al., 2005, [2] Bateman et al., 2017.

Table 3.3.7.1. Hit results for the search of provisional isoSPTs in microbiome-related bacteria identified by BLAST

To assess the similarities amongst the putative isoSPTs and to identify possible conservation motifs, sequence alignments were made for five isoSPTs as well as other three SPTs that have been structurally characterised. It is important to note that such comparison is not representative to the general population of bacterial SPTs and is not intended to be used as one; in contrast, it is used to highlight the similarities between *Bf*SPT and other SPTs from bacteria that have been known to be a part of the human microbiome and generate iso-branched SLs.

	S.paucimobilis	*S.wittichi*	*S.Multivorum*	*Prevotella*	*B.fragilis*	*Porphyromonas*	*Parabacteroides*	*Tannerella*
S.paucimobilis		74.69	35.35	35.44	35.28	33.93	34.43	34.43
S.wittichi	74.69		34.35	36.83	37.69	35.13	38.11	36.83
S.Multivorum	35.35	34.35		53.16	57.36	59.18	57.72	58.99
Prevotella	35.44	36.83	53.16		72.84	66.07	66.84	66.84
B.fragilis	35.28	37.69	57.36	72.84		75.77	76.65	76.14
Porphyromonas	33.93	35.13	59.18	66.07	75.77		80.61	80.1
Parabacteroides	34.43	38.11	57.72	66.84	76.65	80.61		81.01
Tannerella	34.43	36.83	58.99	66.84	76.14	80.1	81.01	

Table 3.3.7.2. Sequence identity matrix generated by Clustal Omega; higher percentage of sequence identities amongst the isoSPTs is marked in red.

The alignment of 5 iso SPTs is shown in figure 3.3.7.2; the alignment of five isoSPTs with *Sp*, *Sm* and *Sw* SPTs is presented in appendix 6. Additionally, the eight SPTs were aligned with some other AOS-family enzymes to reveal the conserved structural core within the family; the alignment is shown in appendix 6. The numbers of totally conserved and partially conserved residues in each of the alignments as well as percentage sequence similarities and identities are given in table 3.3.7.3.

Comparison of all three alignments allows assessing the approximate structural cores shared between AOS enzymes, SPTs and isoSPTs in particular. A phylogenetic tree of all 14 proteins was constructed using Clustal Omega (Sievers et al., 2011) to show approximate evolutionary distance amongst the enzymes (Fig. 3.3.7.1).

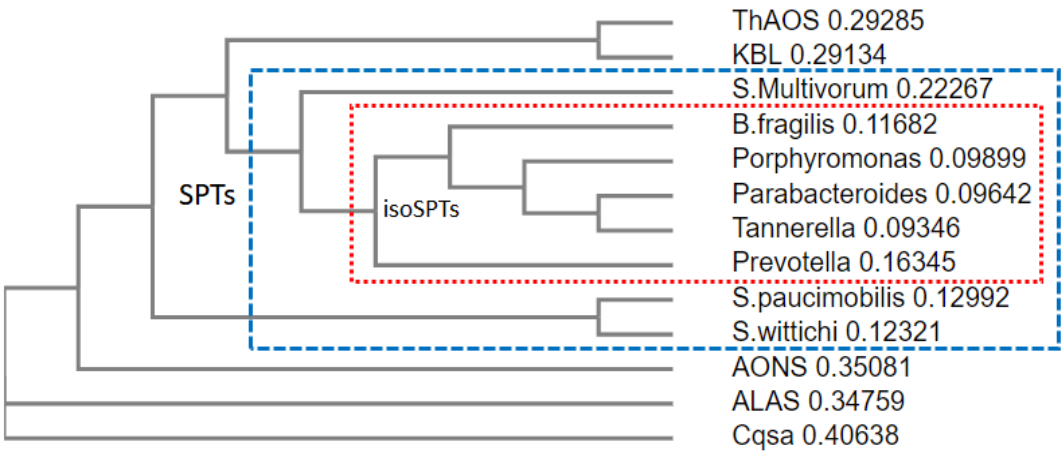


Fig. 3.3.7.1. Phylogenetic tree showing evolutionary relations between different bacterial SPTs and within some common AOS-family enzymes. The numbers represent a fraction of amino acid substitutions per sequence length. SPTs are named after a particular species they belong to; other proteins are from: *Th*AOS (*Thermus thermophilus*), KBL (*E.coli);* AONS *(E.coli);* ALAS *(Rhodobacter capsulatus),* Cqsa – *Vibrio cholerae.* The sequence alignment and the phylogenetic tree was obtained using Clustal Omega software from EMBL (Sievers et al., 2011). All putative isoSPTs sequences are taken from closely related organisms belonging to the same taxonomic order *Bacteroidales*.

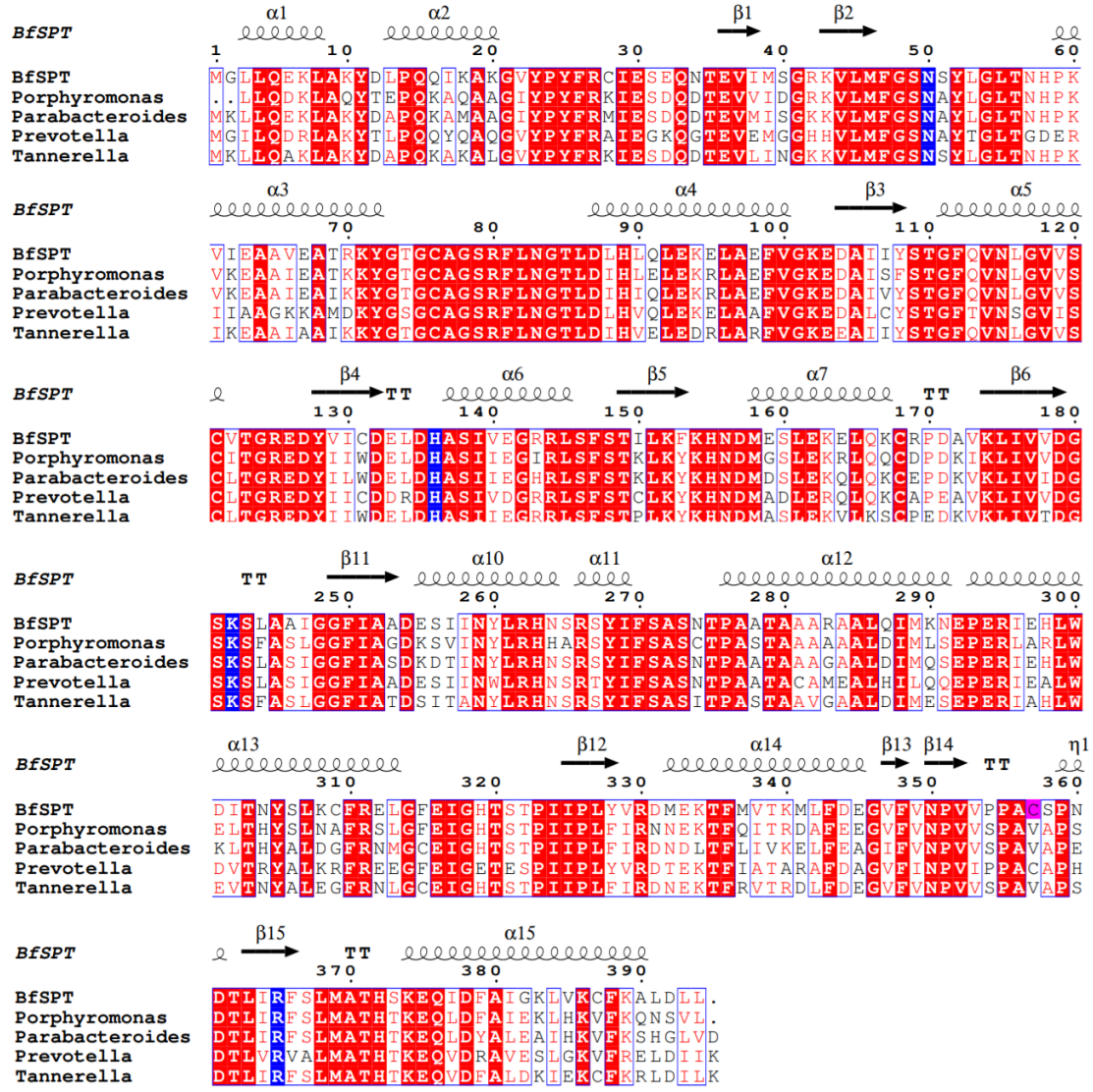


Figure 3.3.7.2. Sequence alignment of 5 isoSPTs (protein IDs presented in table 3.3.7.1). The alignment was performed using Clustal Omega and visualised using ESPript 3.0 (Robert & Gouet, 2014). The residue numbering and the secondary structure information is obtained from the structure of *Bf*SPT. The key catalytic residues involved in PLP or substrate binding are highlighted in blue; the C357 of *Bf*SPT located within the PACSP loop is highlighted in purple.

While the proteins from the AOS family all retain a conserved structural core, the sequence conservation amongst them varies differently with only 16% of residues being fully or partially conserved and 5.7% totally conserved in the alignment sample of 14 AOS enzymes (presented in Appendix 6).

Alignment	Reference	No. fully conserved	No.partially conserved	Length N_{total}	Similarity,%	Identity, %
14 AOS	App 6	22	40	384	16.1	5.7
8 SPTs	App 6	91	90	395	45.8	23.0
5 isoSPT	3.3.7.2	225	72	395	75.2	57.0

Table 3.3.7.3. Summary of the alignments of the protein sequences, presenting the numbers of fully and partially concerned residues within each of the alignments, total lengths of the alignments; sequence similarity $(N_{cons} + N_{partially\ cons})/N_{total}$ and sequence identity N_{cons}/N_{total} amongst all of the compared proteins.

The structures of five putative isoSPTs share a core of 225 conserved and 72 partially conserved residues, giving the sequence similarity and identity scores of 75.2% and 57.0% respectively. The protein sequences in this group were compared including other bacterial SPTs that have been structurally characterised (*Sp*, *Sm* and *Sw*); this gave a sequence similarity score of 45.8%. and a sequence identity score of 23.0%.

It is evident that all five isoSPTs share a common structural core way more considerable than with other SPTs compared (75% including partially conserved residues, as opposed to 45.8% among all the SPTs tested). In this respect, high homology amongst the five putative SPTs from bacteria living in the human body is remarkably striking, considering these organisms are known to produce iso-branched sphingolipids. It is very likely that this conservation is functionally related and is preserved in order to adapt for the specific substrates of these enzymes.

To visualise the protein conservation on a 3D structure the conserved and partially conserved residues were mapped by Consurf (Ashkenazy et al., 2010) with *Bf*SPT structure used as a template. The obtained visualisations for 8 SPTs and 5 isoSPTs have been presented in a surface mode and compared in figure 3.3.7.3.

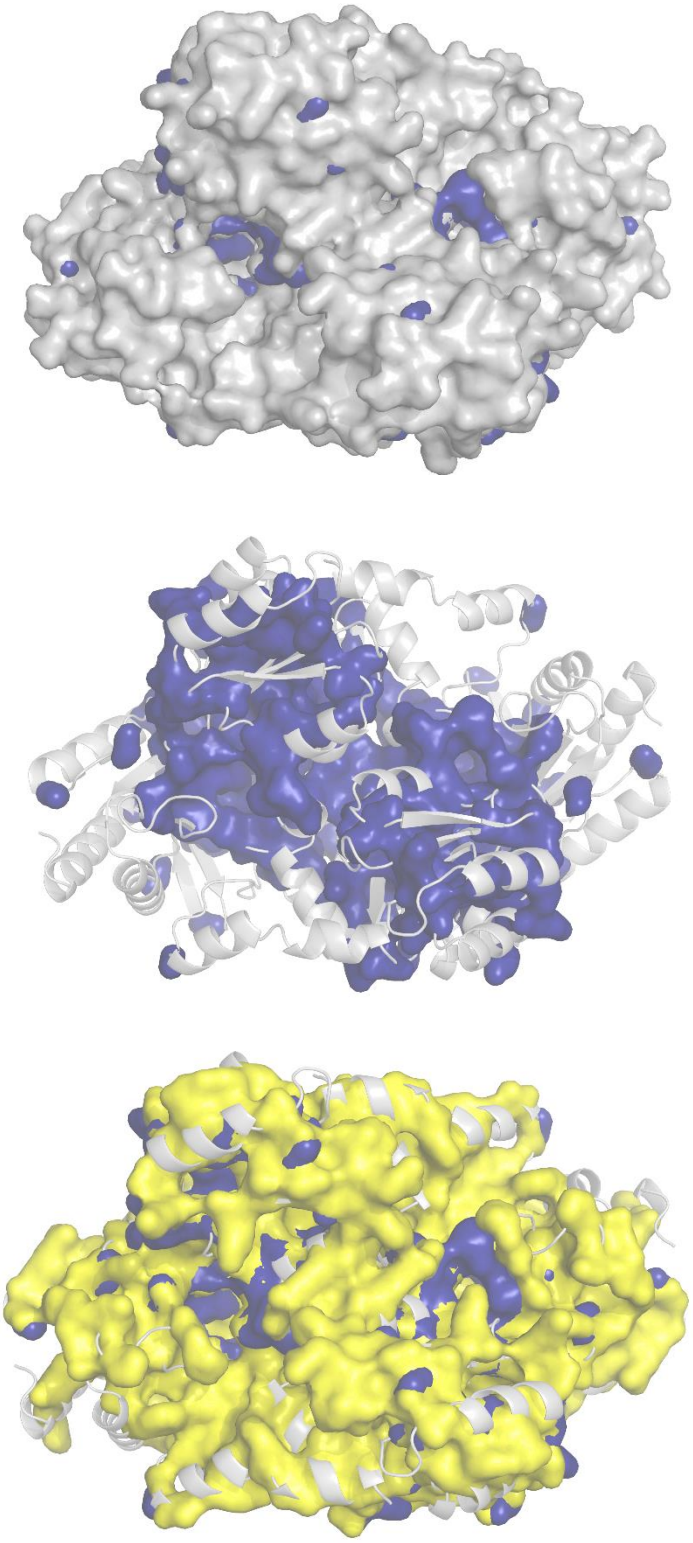


Fig. 3.3.7.3. Structural representation of conservation in the eight SPTs and five isoSPTs only (model: *Bf*SPT). Light grey: all *Bf*SPT residues presented as surface; Dark Blue: residues fully or partially conserved among 8 STPs (45.8%); Yellow: residues fully or partially conserved in five isoSPTs (75%).

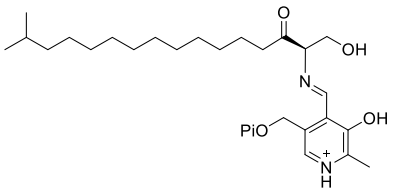
All 8 SPTs have the key PLP-binding residues conserved including the catalytic lysine loop (G238-S243, *Bf*SPT numbering used). 26 out of the 91 residues fully conserved between 8SPTs are glycines; these are generally involved in preserving the shape of the linker regions by adopting flexible geometry.

It is notable that within the isoSPTs the highest sequence variety is observed within the α-helices of proteins. While some β-sheets show a degree of variation, most of them also remain conserved, forming central core domains.

When the conservational alignment is presented as a surface model (Fig. 3.3.7.3), a visible cleft can be observed with the entrance near the "PPAP" loop; the tunnel proceeds to the active site. While there is no direct structural evidence obtained, the cleft is thought to be the entry of the second substrate. Furthermore, the structure of *Sp*SPT in a complex of decarboxymyriocin (4BMK, Wadsworth et al., 2013) shows that the visible 9-carbon chain of the substrate-mimicking inhibitor extends towards the tunnel.

All eight SPTs show a high degree of conservation in the loops forming the proposed substrate-entry tunnel, however, while some variation is present in the SPTs from *Sp, Sm* and *Sw*, the same loops remain almost exclusively identical in all isoSPTs. Notably, the protein surface near the proposed substrate entry site and within the tunnel remains very highly conserved in all isoSPTs; this suggests that all the enzymes share a very similar mode of substrate recognition and might be evolutionary strictly conserved to accept an iso-branched substrate. The high conservation indicates that the proteins are highly specific in their binding mode of the second substrate which is most likely the 15-carbon isopentadecanoyl-CoA.

In order to visualise a possible second substrate/product binding mode, a theoretical PLP-isoKDS external aldimine has been modelled into the active site of *Bf*SPT (Fig. 3.3.7.4)



product-external aldimine

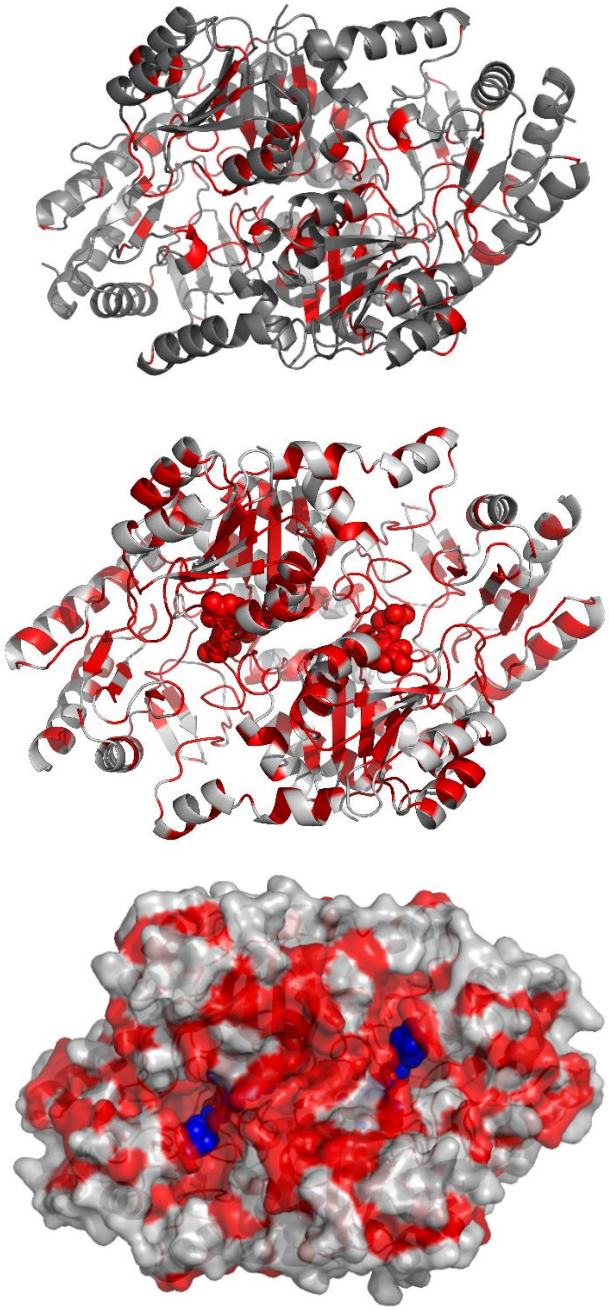


Fig. 3.3.7.4. Mapping of structurally conserved residues (coloured in red) amongst SPTs. **Top**: cartoon representation of residue conservation in 8 SPTs (*Sp*, *Sm*, *Sw* and 5 isoSPTs) mapped on the structure of *Sp*SPT external aldimine (2W8J). Only fully conserved residues are mapped. **Middle**: representation of residues remaining totally conserved among the isoSPTs mapped on the structure of *Bf*SPT-L-serine external aldimine; the ligand is presented as red spheres. Bottom: The structure of *Bf*SPT presented in a surface mode; residues fully conserved between the isoSPTs are mapped. A theoretical PLP-isoKDS product external aldimine has been modelled into the proposed substrate-entry tunnel to represent a possible binding mode; the positioning of decarboxymyriocin from 4BMK was used for initial reference.

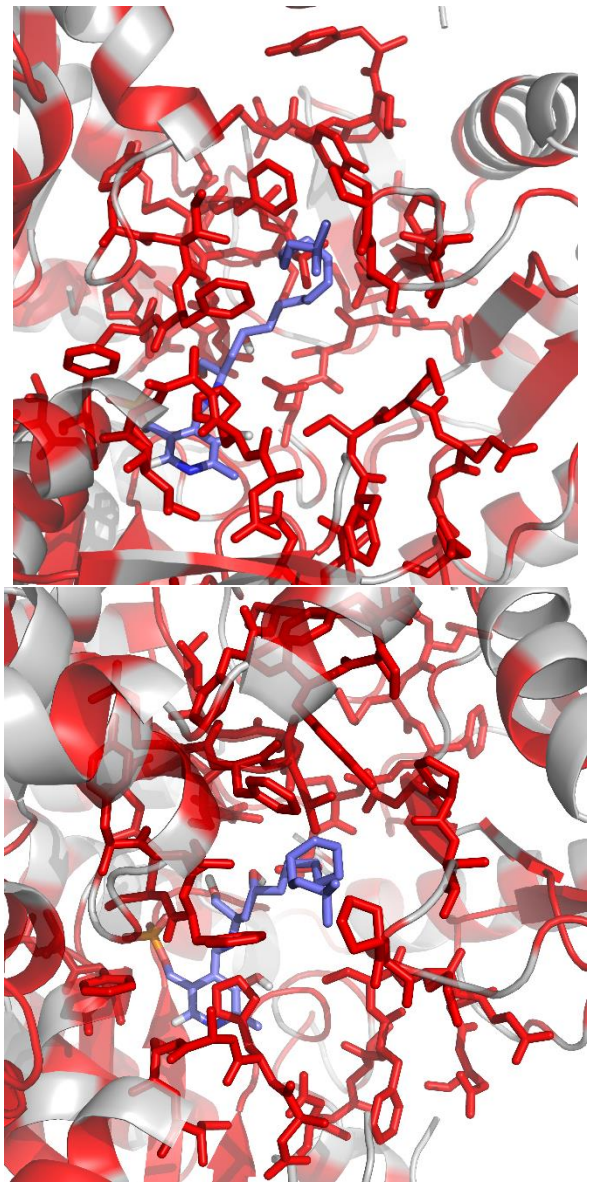


Fig. 3.3.7.5. Conserved residues in five isoSPTs around the proposed substrate-entry tunnel mapped on a structure of *Bf*SPT shown in two orientations. The modelled isoKDS-external aldimine shown in blue; the conserved residues forming the proposed second-substrate entry pocket are shown as red sticks.

From observing both the sequence alignments and the 3D models, it is evident that isoSPTs share unusually high homology amongst them. Conserved loops that form the substrate binding cleft (*Bf*SPT numbering) are G23-27, V44-N50, G75-D87; K174-D187 (Fig.3.3.7.5). These contain a higher number of large aliphatic and aromatic residues as compared to *Sp*SPT that are more likely to form interactions with a proposed iso-branched substrate.

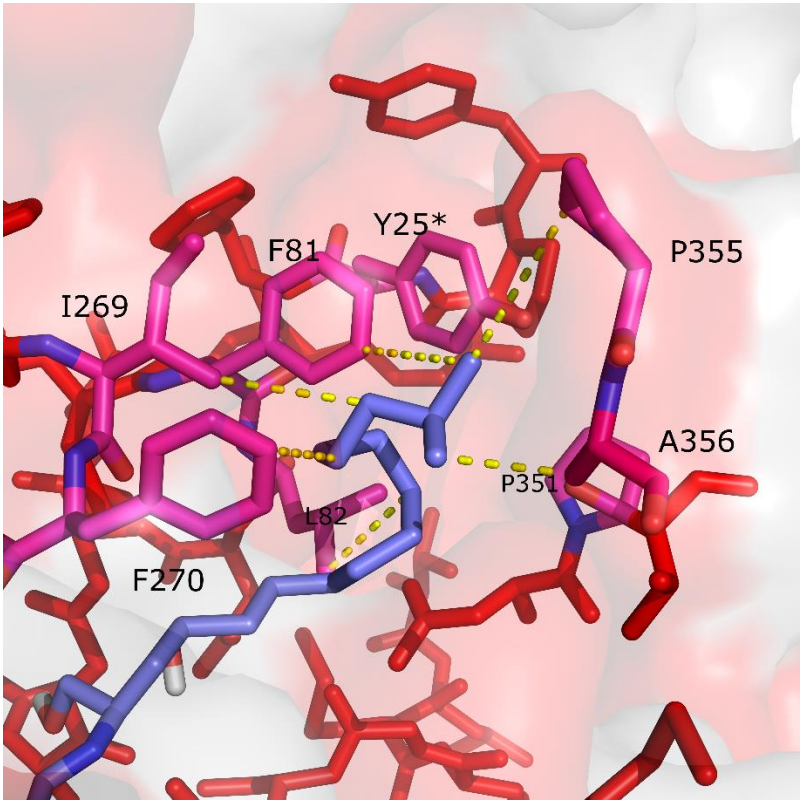


Fig. 3.3.7.6. A model showing a possible mode of substrate binding in isoSPTs. The residues coloured in purple are conserved in all 5 isoSPTs and could provide stabilisation of the bound iso-substrate (cyan). * - residue belonging to the opposite protein chain.

It is also possible to speculate that the conserved residues F81, L82, I269, F270, residue Y25 of the opposite chain and residues P351, P355, A356 belonging to the "PAPP" loop are responsible for the selection and stabilisation of the iso-branched substrate binding by providing a defined hydrophobic pocket environment around the iso-methyl group at the end of the substrate chain (Fig. 3.3.7.6). The very conserved arrangement of these residues might be a key to the observed substrate-inhibition pattern towards a longer non-branched palmitoyl-CoA. The importance of these residues could be experimentally tested by site-directed mutagenesis in *Bf*SPT. Modification of these residues might alter the inhibition patterns observed in reactions of *Bf*SPT with palmitoyl-CoA and reveal novel aspects that control the second-substrate specificity of the enzyme.

It can not be excluded that the high sequence conservation amongst the five isoSPTs is not directly related to the enzyme specificity and is contributed by the close

phylogenetic distance of their species only. However, the 3D-visualisation of the conserved residues overall and around the proposed substrate binding-pocket suggests that some of the structure conservation motifs might be functionally important in binding of the iso-substrate and could be associated with the production of isoSL bases.

The lack of real experimental data does not allow to confirm the specific features of the acyl-CoA substrate binding or determine the nature of substrate inhibition observed towards PCoA in *Bf*SPT. Yet the structural model analysis of the enzyme in comparison with other putative isoSPTs present in human microbiome strongly highlights their similarities as compared to other SPTs. It might be possible that we are looking at the new sub-class of SPT enzymes, of which *Bf*SPT is a first characterised example, that are highly conserved, associated with production of unique isoSL bases and have a direct relevance to the human health.

3.4. Conclusions and further work

There is a vast array of sphingolipids in the human body, adopting a variety of forms and shapes. Furthermore, as humans, we host numerous range of other organisms that compose our microbiome. Some of these organisms produce their own, unique SLs. While the total extent of the microbiome-related SLs and their functions remain largely unknown, it is evident that they play a more significant role in human health than was previously understood. Nevertheless, despite the enormous diversity of sphingolipids in nature, up to our knowledge, the first gateway step in the production of every SL molecule is controlled by the enzyme serine palmitoltransferase.

The work aimed to study the mechanistic and structural features of SPTs from *S.paucimobilis* and *B.fragilis*. Studies of *Sp*SPT highlighted some interesting observations with respect to the amino acid binding. While the X-ray structure of the *Sp*SPT co-crystallised with D-serine yielded the other enantiomer L-serine in the active site, observation of changes in the enzyme backbone allowed to make some deductions about possible racemase activity of SPT. Furthermore, generation of A295T mutant has shown that *Sp*SPT might have a potential to be used for altering substrate specificity. The primary focus of this work, however, remained on the *Bf*SPT.

Bacteroides fragilis is an interesting member of the human microbiome and presents a fascinating target for SL research. It produces some unique iso-branched SLs, the significance of which drives a lot of attention due to the growing evidence of their impact on the human immune system. This work has shed light on some of the structural and mechanistic features of *B.fragilis* SPT. The enzyme has been successfully purified and characterised by a combination of spectrophotometric assays, mass spectrometry and X-ray crystallography. It was capable of utilising a number of substrates *in vitro*, namely myristoyl-, pentadecenoyl-, palmitoyl-, heptadecanoyl- and stearoyl-CoAs, with all reaction products being confirmed by MALDI-ToF mass spectrometry. The high levels of substrate inhibition for palmitoyl-CoA, the substrate expected to be naturally competitive with the proposed isopentadecanoyl-CoA, suggest a presence of a control mechanism regulating the selection of a second substrate.

X-ray crystallography was used to obtain the structure of *Bf*SPT, making it the first SPT with a direct link to human health to be structurally characterised. Several structural complexes of *Bf*SPT have been obtained in the presence of the substrate L-serine as well as with L-cycloserine and L-penicillamine complexes. The structure of *Bf*SPT obtained in the presence of L-penicillamine showed the formation of the PLP-thiazolidine adduct, providing the first structural evidence for the inhibitory mechanism of the drug. While the structural information obtained from the *Bf*SPT-L-cycloserine complex remains non-conclusive, it shows that the taken approach can be used further to get structural information about the drug-binding mode.

While the main aims of this work were successfully achieved, there is an extensively broad area for further studies. The main drawback of this study was the inability to test the activity of *Bf*SPT with its proposed natural substrate isopentadecanoyl-CoA. Obtaining and characterising the iso-branched substrate would be a top priority for future work as this would allow making definite conclusions about the reaction specificity of *Bf*SPT and its role in a production of isoSLs.

Structural and bioinformatic analyses of *Bf*SPT revealed some essential residues within the proposed substrate-binding tunnel that could play a role in the substrate recognition. Testing the significance of these residues using site-directed mutagenesis could provide insights into the substrate recognition and inhibition modes of the enzyme.

As the *Bf* enzyme has proven itself to be a good model for structural studies, much work can be carried in this direction. In particular, further attempts can be made to obtain the structure of the enzyme in complex with D-penicillamine.

Arguably some of the most significant findings of this study were made by analysing the similarity of *Bf*SPT and other putative SPTs from the human microbiome that have been associated with the production of iso-branched sphingolipids. This approach, in combination with the structural data already obtained, has allowed identification of some high structural conservation amongst the enzymes. Further exploration of this area might yield a key to understanding how the bacteria in the human microbiome produce their unique sphingolipids. In this respect, *Bf*SPT can serve as an established model for future studies.

References

Abe, M., & Matsuda, M. (1979). Effect of antivitamin B6 on regional GABA metabolism in mouse brain and its relation to convulsions. *Journal of Nutritional Science and Vitaminology*, *25*(6), 459–68. https://doi.org/10.3177/jnsv.25.459

Abraham, E. P., Chain, E., Baker, W., & Robinson, R. (1943). Penicillamine, a characteristic degradation product of penicillin. *Nature*, *151*, 107. https://doi.org/10.1038/151107a0

Adachi-Yamada, T., Gotoh, T., Sugimura, I., Tateno, M., Nishida, Y., Onuki, T., & Date, H. (1999). De novo synthesis of sphingolipids is required for cell survival by down-regulating c-Jun N-terminal kinase in Drosophila imaginal discs. *Molecular and Cellular Biology*, *19*(10), 7276–86. https://doi.org/10.1128/MCB.19.10.7276

Alexeev, D., Alexeeva, M., Baxter, R. L., Campopiano, D. J., Webster, S. P., & Sawyer, L. (1998). The crystal structure of 8-amino-7-oxononanoate synthase: a bacterial PLP-dependent, acyl-CoA-condensing enzyme. *J. Mol. Biol.*, *284*(2), 401–419. https://doi.org/10.1006/jmbi.1998.2086

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–10. https://doi.org/10.1016/S0022-2836(05)80360-2

Amorim Franco, T. M., Favrot, L., Vergnolle, O., & Blanchard, J. S. (2017). Mechanism-Based Inhibition of the Mycobacterium tuberculosis Branched-Chain Aminotransferase by d - And l - Cycloserine. *ACS Chemical Biology*, *12*(5), 1235–1244. https://doi.org/10.1021/acschembio.7b00142

An, D., Na, C., Bielawski, J., Hannun, Y. A., & Kasper, D. L. (2011). Membrane sphingolipids as essential molecular signals for Bacteroides survival in the intestine. *Proceedings of the National Academy of Sciences*, *108*(Supplement_1), 4666–4671. https://doi.org/10.1073/pnas.1001501107

An, D., Oh, S. F., Olszak, T., Neves, J. F., Avci, F. Y., Erturk-Hasdemir, D., … Kasper, D. L. (2014). Sphingolipids from a symbiotic microbe regulate homeostasis of host intestinal natural killer T cells. *Cell*, 156(1–2), 123–133. https://doi.org/10.1016/j.cell.2013.11.042

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., & Ben-Tal, N. (2010). ConSurf 2010: Calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Research*, *38*(SUPPL. 2), 529–533. https://doi.org/10.1093/nar/gkq399

Asojo, O. A., Nelson, S. K., Mootien, S., Lee, Y., Rezende, W. C., Hyman, D. A., … Anthony, K. G. (2014). Structural and biochemical analyses of alanine racemase from the multidrug-resistant Clostridium difficile strain 630. *Acta Crystallographica Section D: Biological Crystallography*, *70*(7), 1922–1933. https://doi.org/10.1107/S1399004714009419

Astner, I., Schulze, J. O., van den Heuvel, J., Jahn, D., Schubert, W.-D., & Heinz, D. W. (2005). Crystal structure of 5-amweinolevulinate synthase, the first enzyme of heme biosynthesis, and its link to XLSA in humans. *The EMBO Journal*, *24*(18), 3166–3177. http://doi.org/10.1038/sj.emboj.7600792

Azam, M. A., & Jayaram, U. (2016). Inhibitors of alanine racemase enzyme: A review. *Journal of Enzyme Inhibition and Medicinal Chemistry*, *31*(4), 517–526. https://doi.org/10.3109/14756366.2015.1050010

Bateman, A., Martin, M. J., O'Donovan, C., Magrane, M., Alpi, E., Antunes, R., … Zhang, J. (2017). UniProt: The universal protein knowledgebase. *Nucleic Acids Research*, *45*(D1), D158–D169. https://doi.org/10.1093/nar/gkw1099

Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., & Leslie, A. G. W. (2011). iMOSFLM: A new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallographica Section D: Biological Crystallography*, *67*(4), 271–281. https://doi.org/10.1107/S0907444910048675

Beattie, A. E., Clarke, D. J., Wadsworth, J. M., Lowther, J., Sin, H., & Campopiano, D. J. (2013a). Reconstitution of the pyridoxal 5′-phosphate (PLP) dependent enzyme serine palmitoyltransferase (SPT) with pyridoxal reveals a crucial role for the phosphate during catalysis. *Chemical Communications*, *49*(63), 7058. https://doi.org/10.1039/c3cc43001d

Beattie, A. E., Gupta, S. D., Frankova, L., Kazlauskaite, A., Harmon, J. M., Dunn, T. M., & Campopiano, D. J. (2013b). The pyridoxal 5'-phosphate (PLP)-dependent enzyme serine palmitoyltransferase (SPT): effects of the small subunits and insights from bacterial mimics of human hLCB2a HSAN1 mutations. *BioMed Research International*, *2013*, 194371. https://doi.org/10.1155/2013/194371

Bendelac, A., Savage, P. B., & Teyton, L. (2007). The Biology of NKT Cells. *Annual Review of Immunology*, *25*(1), 297–336. https://doi.org/10.1146/annurev.immunol.25.022106.141711

Benson, D. A., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Wheeler, D. L. (2005). GenBank. *Nucleic Acids Research*, *33*(DATABASE ISS.), 34–38. https://doi.org/10.1093/nar/gki063

Bharath, S. R., Bisht, S., Harijan, R. K., Savithri, H. S., & Murthy, M. R. N. (2012). Structural and mutational studies on substrate specificity and catalysis of Salmonella typhimurium D-cysteine desulfhydrase. *PLoS ONE*, *7*(5). https://doi.org/10.1371/journal.pone.0036267

Bikman, B. T. & Summers, S. A. Ceramides as modulators of cellular and whole-body metabolism. *J Clin Invest.* 2011;121(11):4222–4230. doi:10.1172/JCI57144

Bode, H., Bourquin, F., Suriyanarayanan, S., Wei, Y., Alecu, I., Othman, A., … Hornemann, T. (2016). HSAN1 mutations in serine palmitoyltransferase reveal a close structure-function-phenotype relationship. Human Molecular Genetics, 25(5), 853–865. https://doi.org/10.1093/hmg/ddv611

Brossay, L., Chioda, M., Burdin, N., Koezuka, Y., Casorati, G., Dellabona, P., & Kronenberg, M. (1998). CD1d-mediated Recognition of an α-Galactosylceramide by Natural Killer T Cells Is Highly Conserved through Mammalian Evolution. *The Journal of Experimental Medicine*, *188*(8), 1521–1528. https://doi.org/10.1084/jem.188.8.1521

Buell, M. V., & Hansen, R. E. (1960). Reaction of Pyridoxal-5-phosphate with Aminothiols. *Journal of the American Chemical Society*, *82*(23), 6042–6049. https://doi.org/10.1021/ja01508a018

Buist, P. H. (2007). Exotic biomodification of fatty acids. *Natural Product Reports*, *24*(5), 1110–1127. https://doi.org/10.1039/b508584p

Burke, C., Thomas, T., Egan, S., & Kjelleberg, S. (2007). The use of functional genomics for the identification of a gene cluster encoding for the biosynthesis of an antifungal tambjamine in the marine bacterium Pseudoalteromonas tunicata: Brief report. *Environmental Microbiology*, *9*(3), 814–818. https://doi.org/10.1111/j.1462-2920.2006.01177.x

Camp, a V. (1977). Penicillamine in the treatment of rheumatoid arthritis. *Proceedings of the Royal Society of Medicine*, *70*(2), 67–69. https://doi.org/10.1177/003591577707000201

Carter, H. E., & Hirschberg, C. B. (1968). Phytosphingosines and Branched Sphingosines in Kidney. *Biochemistry*, *7*(6), 2296–2300. https://doi.org/10.1021/bi00846a036

Castro, B. M., Prieto, M., & Silva, L. C. (2014). Ceramide: A simple sphingolipid with unique biophysical properties. *Progress in Lipid Research*, *54*(1), 53–67. https://doi.org/10.1016/j.plipres.2014.01.004

Chen, J. K., Lane, W. S., & Schreiber, S. L. (1999). The identification of myriocin-binding proteins. *Chemistry and Biology*, *6*(4), 221–235. https://doi.org/10.1016/S1074-5521(99)80038-6

Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., … Richardson, D. C. (2010). MolProbity: All-atom structure validation for macromolecular crystallography. *Acta Crystallographica Section D: Biological Crystallography*, *66*(1), 12–21. https://doi.org/10.1107/S0907444909042073

Chitwood, D. J., Lusby, W. R., Thompson, M. J., Kochansky, J. P., and Howarth, O. W (1995). The glycosylceramides of the nematode Caenorhabditis elegans contain an unusual, branched-chain sphingoid base. *Lipids* 1995 Jun;30(6):567-73.

Choi, J. S., Park, N. J., Lim, H. K., Ko, Y. K., Kim, Y. S., Ryu, S. Y., & Hwang, I. T. (2012). Plumbagin as a new natural herbicide candidate for Sicyon angulatus control agent with the target 8-amino-7-oxononanoate synthase. *Pesticide Biochemistry and Physiology*, *103*(3), 166–172. https://doi.org/10.1016/j.pestbp.2012.04.007

Cutler, R. G., Thompson, K. W., Camandola, S., Mack, K. T., & Mattson, M. P. (2014). Sphingolipid metabolism regulates development and lifespan in Caenorhabditis elegans. *Mechanisms of Ageing and Development*, *143–144*, 9–18. https://doi.org/10.1016/j.mad.2014.11.002

Das, S. K., & Ray, K. (2006). Wilson's disease: An update. *Nature Clinical Practice Neurology*, *2*(9), 482–493. https://doi.org/10.1038/ncpneuro0291

Dawkins, J. L., Hulme, D. J., Brahmbhatt, S. B., Auer-Grumbach, M., & Nicholson, G. A. (2001). Mutations in SPTLC1, encoding serine palmitoyltransferase, long chain base subunit-1, cause hereditary sensory neuropathy type I. *Nature Genetics*, *27*(3), 309–312. https://doi.org/10.1038/85879

Dickson, R. C. (2008). *Thematic Review Series: Sphingolipids.* New insights into sphingolipid metabolism and function in budding yeast. *Journal of Lipid Research*, *49*(5), 909–921. https://doi.org/10.1194/jlr.R800003-JLR200

Duan, J., & Merrill, A. H. (2015). 1-deoxysphingolipids encountered exogenously and made de novo: Dangerous mysteries inside an enigma. *Journal of Biological Chemistry*, *290*(25), 15380–15389. https://doi.org/10.1074/jbc.R115.658823

Dunathan, H. C. (1996). Confirmation and reaction specificity in pyridoxal phosphate enzymes. *Proc Natl Acad Sci U S A*. 1966 Apr;55(4):712-6.

Eliot, A. C., & Kirsch, J. F. (2004). Pyridoxal Phosphate Enzymes: Mechanistic, Structural, and Evolutionary Considerations. *Annual Review of Biochemistry*, *73*(1), 383–415. https://doi.org/10.1146/annurev.biochem.73.011303.074021

Ellman, G. L. (1959). Tissue sulfhydryl groups. *Archives of Biochemistry and Biophysics*, *82*(1), 70–77. https://doi.org/10.1016/0003-9861(59)90090-6

Emsley, P., Lohkamp, B., Scott, W. G., & Cowtan, K. (2010). Features and development of Coot. *Acta Crystallographica Section D: Biological Crystallography*, *66*(4), 486–501. https://doi.org/10.1107/S0907444910007493

Evans, P. R., & Murshudov, G. N. (2013). How good are my data and what is the resolution? *Acta Crystallographica Section D: Biological Crystallography*, *69*(7), 1204–1214. https://doi.org/10.1107/S0907444913000061

Fahy, E., Subramaniam, S., Brown, H. A., Glass, C. K., Merrill, A. H., Murphy, R. C., … Dennis, E. A. (2005). A comprehensive classification system for lipids 1, *46*. http://doi.org/10.1194/jlr.E400004-JLR200

Fenn, T. D., Stamper, G. F., Morollo, A. A., & Ringe, D. (2003). A Side Reaction of Alanine Racemase : Transamination of Cycloserine †,‡, *5775*–5783.

Gable, K., Slife, H., Bacikova, D., Monaghan, E., & Dunn, T. M. (2000). Tsc3p is an 80-amino acid protein associated with serine palmitoyltransferase and required for optimal enzyme activity. *Journal of Biological Chemistry*, *275*(11), 7597–7603. https://doi.org/10.1074/jbc.275.11.7597

Gable, K., Gupta, S. D., Han, G., Niranjanakumari, S., Harmon, J. M., & Dunn, T. M. (2010). A disease-causing mutation in the active site of serine palmitoyltransferase causes catalytic promiscuity. *Journal of Biological Chemistry*, *285*(30), 22846–22852. https://doi.org/10.1074/jbc.M110.122259

Gao, H., Qi, G., Yin, R., Zhang, H., Li, C., & Zhao, X. (2016). Bacillus cereus strain S2 shows high nematicidal activity against Meloidogyne incognita by producing sphingosine. *Scientific Reports*, *6*(1), 28756. https://doi.org/10.1038/srep28756

Garneau-Tsodikova, S., Dorrestein, P. C., Kelleher, N. L., & Walsh, C. T. (2006). Protein assembly line components in prodigiosin biosynthesis: Characterization of PigA,G,H,I,J. *Journal of the American Chemical Society*, *128*(39), 12600–12601. https://doi.org/10.1021/ja063611l

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., & Bairoch, A. (2005). The Proteomics Protocols Handbook. *The Proteomics Protocols Handbook*, 571–607. https://doi.org/10.1385/1592598900

Gerber, R., Lou, L., & Du, L. (2009). A PLP-dependent polyketide chain releasing mechanism in the biosynthesis of mycotoxin fumonisins in fusarium verticillioides. *Journal of the American Chemical Society*, *131*(9), 3148–3149. https://doi.org/10.1021/ja8091054

Gore, S., Sanz García, E., Hendrickx, P. M. S., Gutmanas, A., Westbrook, J. D., Yang, H., … Kleywegt, G. J. (2017). Validation of Structures in the Protein Data Bank. *Structure*, *25*(12), 1916–1927. https://doi.org/10.1016/j.str.2017.10.009

Goulas, T., Arolas, J. L., & Gomis-Rüth, F. X. (2011). Structure, function and latency regulation of a bacterial enterotoxin potentially derived from a mammalian adamalysin/ADAM xenolog. *Proceedings of the National Academy of Sciences*, *108*(5), 1856–1861. https://doi.org/10.1073/pnas.1012173108

Han, G., Gupta, S. D., Gable, K., Niranjanakumari, S., Moitra, P., Eichler, F., … Dunn, T. M. (2009). Identification of small subunits of mammalian serine palmitoyltransferase that confer distinct acyl-CoA substrate specificities. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(20), 8186–91. https://doi.org/10.1073/pnas.0811269106

Han, W. B., Lu, Y. H., Zhang, A. H., Zhang, G. F., Mei, Y. N., Jiang, N., … Tan, R. X. (2014). Curvulamine, a new antibacterial alkaloid incorporating two undescribed units from a Curvularia species. *Organic Letters*, *16*(20), 5366–5369. https://doi.org/10.1021/ol502572g

Hanada, K., Hara, T., & Nishijima, M. (2000a). D-Serine inhibits serine palmitoyltransferase, the enzyme catalyzing the initial step of sphingolipid biosynthesis. *FEBS Letters*, *474*(1), 63–65. https://doi.org/10.1016/S0014-5793(00)01579-9

Hanada, K., Hara, T., & Nishijima, M. (2000b). Purification of the serine palmitoyltransferase complex responsible for sphingoid base synthesis by using affinity peptide chromatography techniques. *Journal of Biological Chemistry*, *275*(12), 8409–8415. https://doi.org/10.1074/jbc.275.12.8409

Hanada, K., Nishijima, M., Fujita, T., & Kobayashi, S. (2000c). Specificity of inhibitors of serine palmitoyltransferase (SPT), a key enzyme in sphingolipid biosynthesis, in intact cells. A novel evaluation system using an SPT-defective mammalian cell mutant. *Biochemical Pharmacology*, *59*(10), 1211–1216. https://doi.org/10.1016/S0006-2952(00)00251-3

Hanada, K. (2003). Serine palmitoyltransferase, a key enzyme of sphingolipid metabolism. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1632*(1–3), 16–30. https://doi.org/10.1016/S1388-1981(03)00059-3

Hannich, J.T., Mellal, D., Feng, S., Zumbuehl, A., & Riezman, H. (2017). Structure and conserved function of iso-branched sphingoid bases from the nematode Caenorhabditis elegans. *Chemical Science*, *8*(5), 3676–3686. https://doi.org/10.1039/c6sc04831e

Heaver, S. L., Johnson, E. L., & Ley, R. E. (2018). Sphingolipids in host–microbial interactions. *Current Opinion in Microbiology*, *43*, 92–99. https://doi.org/10.1016/j.mib.2017.12.011

Hojjati, M. R., Li, Z., & Jiang, X. C. (2005). Serine palmitoyl-CoA transferase (SPT) deficiency and sphingolipid levels in mice. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, *1737*(1), 44–51. https://doi.org/10.1016/j.bbalip.2005.08.006

Hornemann, T., Richard, S., Rütti, M. F., Wei, Y., & Von Eckardstein, A. (2006). Cloning and initial characterization of a new subunit for mammalian serine-palmitoyltransferase. *Journal of Biological Chemistry*, *281*(49), 37275–37281. https://doi.org/10.1074/jbc.M608066200

Hornung, C., Poehlein, A., Haack, F. S., Schmidt, M., Dierking, K., Pohlen, A., … Streit, W. R. (2013). The Janthinobacterium sp. HH01 Genome Encodes a Homologue of the V. cholerae CqsA and L. pneumophila LqsA Autoinducer Synthases. *PLoS ONE*, *8*(2). https://doi.org/10.1371/journal.pone.0055045

Ikushiro, H., Hayashi, H., & Kagamiyama, H. (2001). A water-soluble homodimeric serine palmitoyltransferase from Sphingomonas paucimobilis EY2395T strain: Purification, characterization, cloning, and overproduction. *Journal of Biological Chemistry*, *276*(21), 18249–18256. https://doi.org/10.1074/jbc.M101550200

Ikushiro, H., Hayashi, H., & Kagamiyama, H. (2003). Bacterial serine palmitoyltransferase: A water-soluble homodimeric prototype of the eukaryotic enzyme. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1647*(1–2), 116–120. https://doi.org/10.1016/S1570-9639(03)00074-8

Ikushiro, H., Hayashi, H., & Kagamiyama, H. (2004). Reactions of serine palmitoyltransferase with serine and molecular mechanisms of the actions of serine derivatives as inhibitors. *Biochemistry*, *43*(4), 1082–1092. https://doi.org/10.1021/bi035706v

Ikushiro, H., Islam, M. M., Tojo, H., & Hayashi, H. (2007). Molecular characterization of membrane-associated soluble serine palmitoyltransferases from *Sphingobacterium multivorum* and *Bdellovibrio stolpii*. *Journal of Bacteriology*, *189*(15), 5749–5761. https://doi.org/10.1128/JB.00194-07

Ikushiro, H., Fujii, S., Shiraiwa, Y., & 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*(12), 7542–7553. https://doi.org/10.1074/jbc.M706874200

Ikushiro, H., Islam, M. M., Okamoto, A., Hoseki, J., Murakawa, T., Fujii, S., … Hayashi, H. (2009). Structural insights into the enzymatic mechanism of serine palmitoyltransferase from *Sphingobacterium multivorum*. *Journal of Biochemistry*, *146*(4), 549–562. https://doi.org/10.1093/jb/mvp100

Jahan, N., Potter, J. A., Sheikh, M. A., Botting, C. H., Shirran, S. L., Westwood, N. J., & Taylor, G. L. (2009). Insights into the Biosynthesis of the Vibrio cholerae Major Autoinducer CAI-1 from the Crystal Structure of the PLP-Dependent Enzyme CqsA. *Journal of Molecular Biology*, *392*(3), 763–773. https://doi.org/10.1016/j.jmb.2009.07.042

Jaikishan, S., Björkbom, A., & Slotte, J. P. (2010). Sphingomyelin analogs with branched N-acyl chains: The position of branching dramatically affects acyl chain order and sterol interactions in bilayer membranes. *Biochimica et Biophysica Acta - Biomembranes*, *1798*(10), 1987–1994. https://doi.org/10.1016/j.bbamem.2010.07.006

Jäckel, C., Kast, P., & Hilvert, D. (2008). Protein Design by Directed Evolution. *Annual Review of Biophysics*, *37*(1), 153–173. https://doi.org/10.1146/annurev.biophys.37.032807.125832

Jaffe, I. A., Altman, K., & Merryman, P. (1964). The Antipyridoxine Effect of Penicillamine in Man. *The Journal of Clinical Investigation*, *43*(10), 1869–1873. https://doi.org/10.1172/JCI105060

Kaneda, T. (1977). Fatty acids of the genus Bacillus: an example of branched-chain preference. *Bacteriological Reviews*, *41*(2), 391–418.

Kaneda, T. (1991). Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance. *Microbiological Reviews*, *55*(2), 288–302. https://doi.org/10.1038/nrm1745

Kato, M., Watanabe, K., Tanaka-Bandoh, K., Watanabe, K., & Ueno, K. (1995). Sphingolipid composition in Bacteroides species. *Anaerobe*. 1995 Apr;1(2):135-9. https://doi.org/10.1006/anae.1995.1009

Kelly, R. C., Bolitho, M. E., Higgins, D. A., Lu, W., Ng, W.-L., Jeffrey, P. D., … Bassler, B. L. (2009). The Vibrio cholerae quorum-sensing autoinducer CAI-1: analysis of the biosynthetic enzyme CqsA. *Nature Chemical Biology*, *5*(12), 891–895. https://doi.org/10.1038/nchembio.237

Kinjo, Y., Wu, D., Kim, G., Xing, G. W., Poles, M. A., Ho, D. D., … Kronenberg, M. (2005). Recognition of bacterial glycosphingolipids by natural killer T cells. *Nature*, *434*(7032), 520–525. https://doi.org/10.1038/nature03407

Kowluru, A. (2014). Deoxysphingolipids: β-cell, beware of these new kids on the block. *Diabetes*, *63*(4), 1191–3. https://doi.org/10.2337/db14-0022

Kraft, M. L. (2017). Sphingolipid Organization in the Plasma Membrane and the Mechanisms That Influence It. *Frontiers in Cell and Developmental Biology*, *4*(January), 1–19. https://doi.org/10.3389/fcell.2016.00154

Kubota, T., Shimono, J., Kanameda, C., & Izumi, Y. (2007). The first thermophilic alpha-oxoamine synthase family enzyme that has activities of 2-amino-3-ketobutyrate CoA ligase and 7-keto-8-aminopelargonic acid synthase: cloning and overexpression of the gene from an extreme thermophile, Thermus thermophilus, and. *Bioscience, Biotechnology, and Biochemistry*, *71*(12), 3033–40. https://doi.org/10.1271/bbb.70438

Kubota, T., & Izumi, Y. (2012). Detection and characterization of a thermophilic biotin biosynthetic enzyme, 7-keto-8-aminopelargonic acid synthase, from various thermophiles. *Bioscience, Biotechnology, and Biochemistry*, *76*(4), 685–90. https://doi.org/10.1271/bbb.110807

Kuznetsov, N. A., Faleev, N. G., Kuznetsova, A. A., Morozova, E. A., Revtovich, S. V., Anufrieva, N. V., … Demidkina, T. V. (2015). Pre-steady-state kinetic and structural analysis of interaction of methionine γ-lyase from Citrobacter freundii with inhibitors. *Journal of Biological Chemistry*, *290*(1), 671–681. https://doi.org/10.1074/jbc.M114.586511

Liu, H., & Naismith, J. H. (2008). An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnology*, *8*, 91. https://doi.org/10.1186/1472-6750-8-91

Lochnit, G., Dennis, R. D., Zähringer, U., & Geyer, R. (1997). Structural analysis of neutral glycosphingolipids from Ascaris suum adults (Nematoda: Ascaridida). *Glycoconjugate Journal*, *14*(3), 389–399. https://doi.org/10.1023/A:1018530914067

Lowther, J., Yard, B. A., Johnson, K. A., Carter, L. G., Bhat, V. T., Raman, M. C. C., … Campopiano, D. J. (2010). Inhibition of the PLP-dependent enzyme serine palmitoyltransferase by cycloserine: evidence for a novel decarboxylative mechanism of inactivation. *Molecular Biosystems*, *6*(9), 1682–1693. https://doi.org/10.1039/c003743e

Lowther, J., Charmier, G., Raman, M. C., Ikushiro, H., Hayashi, H., & Campopiano, D. J. (2011). Role of a conserved arginine residue during catalysis in serine palmitoyltransferase. *FEBS Letters*, *585*(12), 1729–1734. https://doi.org/10.1016/j.febslet.2011.04.013

Lowther, J., Beattie, A. E., Langridge-Smith, P. R. R., Clarke, D. J., & Campopiano, D. J. (2012a). l-Penicillamine is a mechanism-based inhibitor of serine palmitoyltransferase by forming a pyridoxal-5′-phosphate-thiazolidine adduct. *MedChemComm*, *3*(8), 1003. https://doi.org/10.1039/c2md20020a

Lowther, J., Naismith, J. H., Dunn, T. M., & Campopiano, D. J. (2012b). Structural, mechanistic and regulatory studies of serine palmitoyltransferase. *Biochemical Society Transactions*, *40*(3), 547–554. https://doi.org/10.1042/BST20110769

Lynch, D. D. V, & Dunn, T. M. T. (2004). An introduction to plant sphingolipids and a review of recent advances in understanding their metabolism and function. *New Phytologist*, *161*(3), 677–702. https://doi.org/10.1111/j.1469-8137.2003.00992.x

McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C., & Read, R. J. (2007). Phaser crystallographic software. *Journal of Applied Crystallography*, *40*(4), 658–674. https://doi.org/10.1107/S0021889807021206

Malashkevich, V. N., Strop, P., Keller, J. W., Jansonius, J. N., & Toney, M. D. (1999). Crystal structures of dialkylglycine decarboxylase inhibitor complexes. *Journal of Molecular Biology*, *294*(1), 193–200. https://doi.org/10.1006/jmbi.1999.3254

Mayberry, W. R. (1980). Hydrooxy fatty acids in Bacteroides species: d-(-)-3 hydroxy-15-methylhexadecanoate and its homologs. *J. Gen. Bact.*, *143*(2), 582–587. http://doi.org/0021-9193/80/08-0582/06$02.00/0

Mazmanian, S. K., Round, J. L., & Kasper, D. L. (2008). A microbial symbiosis factor prevents intestinal inflammatory disease. *Nature*, *453*(7195), 620–625. https://doi.org/10.1038/nature07008

Medlock, K. A., & Merrill, A. H. (1988). Inhibition of Serine Palmitoyltransferase In Vitro And Long-chain Base Biosynthesis in Intact Chinese Hamster Ovary Cells By β-chloroalanine. *Biochemistry*, *27*(18), 7079–7084. https://doi.org/10.1021/bi00418a061

Merrill, A. H. (2002). De novo sphingolipid biosynthesis: A necessary, but dangerous, pathway. *Journal of Biological Chemistry*, *277*(29), 25843–25846. https://doi.org/10.1074/jbc.R200009200

Merrill, A. H. (2011). Sphingolipid and glycosphingolipid metabolic pathways in the era of sphingolipidomics. *Chemical Reviews*, *111*(10), 6387–6422. https://doi.org/10.1021/cr2002917

Metzler, E., Ikawa, M. & Snell, E. A General Mechanism for Vitamin B6 -catalyzed Reactions. *J. Am. Chem. Soc*. 13, (1954).

Mina, J. G., Thye, J. K., Alqaisi, A. Q. I., Bird, L. E., Dods, R. H., Grøftehauge, M. K., … Denny, P. W. (2017). Functional and phylogenetic evidence of a bacterial origin for the first enzyme in sphingolipid biosynthesis in a phylum of eukaryotic protozoan parasites. *Journal of Biological Chemistry*, *292*(29), 12208–12219. https://doi.org/10.1074/jbc.M117.792374

Miyagawa, E., Azuma, R., Suto, T., & Yano, I. (1979). Occurrence of free ceramides in Bacteroides fragilis NCTC 9343. *Journal of Biochemistry*, *86*(2), 311–320. https://doi.org/10.1093/oxfordjournals.jbchem.a132528

Moye, Z. D., Valiuskyte, K., Dewhirst, F. E., Nichols, F. C., Davey, M. E., Lamont, R., & Davey, M. E. (2016). Synthesis of Sphingolipids Impacts Survival of Porphyromonas gingivalis and the Presentation of Surface Polysaccharides, *7*(November), 1–13. http://doi.org/10.3389/fmicb.2016.01919

Mun, J., Onorato, A., Nichols, F. C., Morton, M. D., Saleh, A. I., & Smith, M. B. (2007). Structural confirmation of the dihydrosphinganine and fatty acid constituents of the dental pathogen Porphyromonas gingivalis †, 3826–3833. http://doi.org/10.1039/b712707c

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., … Vagin, A. A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallographica Section D: Biological Crystallography*, *67*(4), 355–367. https://doi.org/10.1107/S0907444911001314

Naka, T., Fujiwara, N., Yano, I., Maeda, S., Doe, M., Minamino, M., … Kumazawa, Y. (2003). Structural analysis of sphingophospholipids derived from Sphingobacterium spiritivorum , the type species of genus Sphingobacterium, *1635*, 83–92. http://doi.org/10.1016/j.bbalip.2003.10.010

Ng, W.-L., Perez, L. J., Wei, Y., Kraml, C., Semmelhack, M. F., & Bassler, B. L. (2011). Signal production and detection specificity in Vibrio CqsA/CqsS quorum-sensing systems. *Molecular Microbiology*, *79*(6), 1407–1417. https://doi.org/10.1111/j.1365-2958.2011.07548.x

Nichols, F. C. (1998). Novel ceramides recovered from Porphyromonas gingivalis: relationship to adult periodontitis. *Journal of Lipid Research*, *39*(12), 2360–72. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9831624

Nichols, F. C., Riep, B., Mun, J., Morton, M. D., Bojarski, M. T., Dewhirst, F. E., & Smith, M. B. (2004). Structures and biological activity of phosphorylated dihydroceramides of Porphyromonas gingivalis. *Journal of Lipid Research*, *45*(12), 2317–2330. http://doi.org/10.1194/jlr.M400278-JLR200

Nichols, F. C., Riep, B., Mun, J., Morton, M. D., Kawai, T., Dewhirst, F. E., & Smith, M. B. (2006). Structures and biological activities of novel phosphatidylethanolamine lipids of Porphyromonas gingivalis. *Journal of Lipid Research*, *47*(4), 844–853. https://doi.org/10.1194/jlr.M500542-JLR200

Nichols, F. C., Housley, W. J., Conor, C. A. O., Manning, T., Wu, S., & Clark, R. B. (2009). Unique Lipids from a Common Human Bacterium Represent a New Class of Toll-Like Receptor 2 Ligands Capable of Enhancing Autoimmunity. *The American Journal of Pathology*, *175*(6), 2430–2438. http://doi.org/10.2353/ajpath.2009.090544

Nichols, F. C., Yao, X., Bajrami, B., Downes, J., Finegold, S. M., Knee, E., … Clark, R. B. (2011). Phosphorylated dihydroceramides from common human bacteria are recovered in human tissues. *PLoS ONE*, *6*(2). http://doi.org/10.1371/journal.pone.0016771

Noda, M., Matoba, Y., Kumagai, T., & Sugiyama, M. (2004). Structural evidence that alanine racemase from a D-cycloserine-producing microorganism exhibits resistance to its own product. *Journal of Biological Chemistry*, *279*(44), 46153–46161. https://doi.org/10.1074/jbc.M404605200

Noland, B. W., Newman, J. M., Hendle, J., Badger, J., Christopher, J. A., Tresser, J., … Buchanan, S. G. (2002). Structural studies of Salmonella typhimurium ArnB (PmrH) aminotransferase: A 4-amino-4-deoxy-L-arabinose lipopolysaccharide-modifying enzyme. *Structure*, *10*(11), 1569–1580. https://doi.org/10.1016/S0969-2126(02)00879-1

Othman, A., Saely, C. H., Muendlein, A., Vonbank, A., Drexel, H., von Eckardstein, A., & Hornemann, T. (2015). Plasma 1-deoxysphingolipids are predictive biomarkers for type 2 diabetes mellitus. *BMJ Open Diabetes Research & Care*, *3*(1), e000073. https://doi.org/10.1136/bmjdrc-2014-000073

Peisach, D., Chipman, D. M., Van Ophem, P. W., Manning, J. M., & Ringe, D. (1998). D-Cycloserine inactivation of D-amino acid aminotransferase leads to a stable noncovalent protein complex with an aromatic cycloserine-PLP derivative. *Journal of the American Chemical Society*, *120*(10), 2268–2274. http://doi.org/10.1021/ja973353f

Penno, A., Reilly, M. M., Houlden, H., Laurá, M., Rentsch, K., Niederkofler, V., … Hornemann, T. (2010). Hereditary sensory neuropathy type 1 is caused by the accumulation of two neurotoxic sphingolipids. *The Journal of Biological Chemistry*, *285*(15), 11178–87. https://doi.org/10.1074/jbc.M109.092973

Percudani, R., & Peracchi, A. (2003). A genomic overview of pyridoxal-phosphate-dependent enzymes. *EMBO Reports*, *4*(9), 850–854. https://doi.org/10.1038/sj.embor.embor914

Perri, G., & Bonora, S. (2004). Which agents should we use for the treatment of multidrug-resistant Mycobacterium tuberculosis? *Journal of Antimicrobial Chemotherapy*. https://doi.org/10.1093/jac/dkh377

Pinto, W. J., Wells, G. W., & Lester, R. L. (1992). 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. *Journal of Bacteriology*, *174*(8), 2575–2581. https://doi.org/10.1128/jb.174.8.2575-2581.1992

Priyadarshi, A., Lee, E. H., Sung, M. W., Nam, K. H., Lee, W. H., Kim, E. E., & Hwang, K. Y. (2009). Structural insights into the alanine racemase from Enterococcus faecalis. *Biochimica et Biophysica Acta - Proteins and Proteomics*, *1794*(7), 1030–1040. https://doi.org/10.1016/j.bbapap.2009.03.006

Pruett, S. T., Bushnev, A., Hagedorn, K., Adiga, M., Haynes, C. A., Sullards, M. C., … Merrill, A. H. (2008). *Thematic Review Series: Sphingolipids.* Biodiversity of sphingoid bases ("sphingosines") and related amino alcohols. *Journal of Lipid Research*, *49*(8), 1621–1639. https://doi.org/10.1194/jlr.R800012-JLR200

Pye, G., Evans, D. F., Ledingham, S., & Hardcastle, J. D. (1990). Gastrointestinal intraluminal pH in normal subjects and those with colorectal adenoma or carcinoma. *Gut*, *31*(12), 1355–1357. https://doi.org/10.1136/gut.31.12.1355

Raman, M. C. C., Johnson, K. A., Yard, B. A., Lowther, J., Carter, L. G., Naismith, J. H., & Campopiano, D. J. (2009). The external aldimine form of serine palmitoyltransferase: Structural, kinetic and spectroscopic analysis of the wild-type enzyme and HSAN1 mutant mimics. *Journal of Biological Chemistry*, *284*(25), 17328–17339. https://doi.org/10.1074/jbc.M109.008680

Raman, M. C. C., Johnson, K. A., Clarke, D. J., Naismith, J. H., & Campopiano, D. J. (2010). The serine palmitoyltransferase from sphingomonas wittichii RW1: An interesting link to an unusual acyl carrier protein. *Biopolymers*, *93*(9), 811–822. https://doi.org/10.1002/bip.21482

Reid, G. (2004). When Microbe Meets Human. *Clinical Infectious Diseases*, *39*(6), 827–830. https://doi.org/10.1086/423387

Riddles, P. W., Blakeley, R. L., & Zerner, B. (1983). Reassessment of Ellman's Reagent. *Methods in Enzymology*, *91*(C), 49–60. https://doi.org/10.1016/S0076-6879(83)91010-8

Robert, X., & Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, *42*(W1), 320–324. https://doi.org/10.1093/nar/gku316

Schade, S., & Paulus, W. (2016). D-Cycloserine in Neuropsychiatric Diseases: A Systematic Review. *International Journal of Neuropsychopharmacology*. https://doi.org/10.1093/ijnp/pyv102

Schmelz, E. M., Crall, K. J., Larocque, R., Dillehay, D. L., & Merrill, A. H. (1994). Uptake and Metabolism of Sphingolipids in Isolated Intestinal Loops of Mice. *The Journal of Nutrition*, *124*(December 1993), 702–712.

Schmidt, a, Sivaraman, J., Li, Y., Larocque, R., Barbosa, J. a, Smith, C., … Cygler, M. (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*(17), 5151–5160. http://doi.org/10.1021/bi002204y

Schüttelkopf, A. W., & Van Aalten, D. M. F. (2004). PRODRG: A tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallographica Section D: Biological Crystallography*, *60*(8), 1355–1363. https://doi.org/10.1107/S0907444904011679

Sears, C. L. (2009). Enterotoxigenic Bacteroides fragilis: A rogue among symbiotes. *Clinical Microbiology Reviews*, *22*(2), 349–369. https://doi.org/10.1128/CMR.00053-08

Shiraiwa, Y., Ikushiro, H., & Hayashi, H. (2009). Multifunctional role of His159 in the catalytic reaction of Serine palmitoyltransferase. *Journal of Biological Chemistry*, *284*(23), 15487–15495. https://doi.org/10.1074/jbc.M808916200

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., … Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, *7*(539). https://doi.org/10.1038/msb.2011.75

Smith, M. A., Mack, V., Ebneth, A., Moraes, I., Felicetti, B., Wood, M., … Barker, J. (2010). The structure of mammalian serine racemase: Evidence for conformational changes upon inhibitor binding. *Journal of Biological Chemistry*, *285*(17), 12873–12881. https://doi.org/10.1074/jbc.M109.050062

Spirig, T., Tiaden, A., Kiefer, P., Buchrieser, C., Vorholt, J. A., & Hilbi, H. (2008). The Legionella autoinducer synthase LqsA produces an α-hydroxyketone signaling molecule. *Journal of Biological Chemistry*, *283*(26), 18113–18123. https://doi.org/10.1074/jbc.M801929200

Surana, N. K., & Kasper, D. L. (2012). The yin yang of bacterial polysaccharides: Lessons learned from B. fragilis PSA. *Immunological Reviews*, *245*(1), 13–26. https://doi.org/10.1111/j.1600-065X.2011.01075.x

Teng, T.-Y. (1990). Mounting of crystals for macromolecular crystallography in a free-standing thin film. *Journal of Applied Crystallography*, *23*(5), 387–391. https://doi.org/10.1107/S0021889890005568

Ternes, P., Sperling, P., Albrecht, S., Franke, S., Cregg, J. M., Warnecke, D., & Heinz, E. (2006). Identification of fungal sphingolipid C9-methyltransferases by phylogenetic profiling. *Journal of Biological Chemistry*, *281*(9), 5582–5592. https://doi.org/10.1074/jbc.M512864200

Thoden, J. B., Taylor Ringia, E. A., Garrett, J. B., Gerlt, J. A., Holden, H. M., & Rayment, I. (2004). Evolution of Enzymatic Activity in the Enolase Superfamily: Structural Studies of the Promiscuous o-Succinylbenzoate Synthase from Amycolatopsis. *Biochemistry*, *43*(19), 5716–5727. https://doi.org/10.1021/bi0497897

Toney, M. D. (2005). Reaction specificity in pyridoxal phosphate enzymes. *Archives of Biochemistry and Biophysics*, *433*(1), 279–287. https://doi.org/10.1016/j.abb.2004.09.037

Toney, M. D. (2011). Controlling reaction specificity in pyridoxal phosphate enzymes. *Biochim. Biophys. Acta.* doi.org/10.1016/j.bbapap.2011.05.019

Turner, N. J. (2009). Directed evolution drives the next generation of biocatalysts. *Nature Chemical Biology*, *5*(8), 567–573. https://doi.org/10.1038/nchembio.203

Vesper, H., Schmelz, E.-M., Nikolova-Karakashian, M. N., Dillehay, D. L., Lynch, D. V., & Merrill, A. H. (1999). Sphingolipids in Food and the Emerging Importance of Sphingolipids to Nutrition. *The Journal of Nutrition*, *129*(7), 1239–1250. https://doi.org/10.1093/jn/129.7.1239

Wadsworth, J. M., Clarke, D. J., Mcmahon, S. A., Lowther, J. P., Beattie, A. E., Langridge-Smith, P. R. R., … Campopiano, D. J. (2013). The chemical basis of serine palmitoyltransferase inhibition by myriocin. *Jourxnal of the American Chemical Society*, *135*(38), 14276–14285. https://doi.org/10.1021/ja4059876

Weiss, B., & Stoffel, W. (1997). Human and murine serine-palmitoyl-CoA transferase--cloning, expression and characterization of the key enzyme in sphingolipid synthesis. *European Journal of Biochemistry / FEBS*, *249*(1), 239–247. https://doi.org/10.1111/j.1432-1033.1997.00239.x

Wexler, H. M. (2007). Bacteroides: The good, the bad, and the nitty-gritty. *Clinical Microbiology Reviews*, *20*(4), 593–621. https://doi.org/10.1128/CMR.00008-07

Wieland Brown, L. C., Penaranda, C., Kashyap, P. C., Williams, B. B., Clardy, J., Kronenberg, M., … Fischbach, M. A. (2013). Production of α-Galactosylceramide by a Prominent Member of the Human Gut Microbiota. *PLoS Biology*, *11*(7). https://doi.org/10.1371/journal.pbio.1001610

Williams, R. D., Sgoutas, D. S., Zaatari, G. S., & Santoianni, R. A. (1987). Inhibition of serine palmitoyltransferase activity in rabbit aorta by L-cycloserine. *J Lipid Res*, *28*(12), 1478–1481.

Williamson, N. R., Simonsen, H. T., Ahmed, R. A. A., Goldet, G., Slater, H., Woodley, L., … Salmond, G. P. C. (2005). Biosynthesis of the red antibiotic, prodigiosin, in Serratia: Identification of a novel 2-methyl-3-n-amyl-pyrroie (MAP) assembly pathway, definition of the terminal condensing enzyme, and implications for undecylprodigiosin biosynthesis in Streptomyces. *Molecular Microbiology*, *56*(4), 971–989. https://doi.org/10.1111/j.1365-2958.2005.04602.x

Wilson, J.E. & Vigneaud, V. Du (1950). Inhibition of the growth of the rat by L-penicillamine and its prevention by aminoethanol and related compounds. *J. Biol. Chem*., 1950, 184 (1), 63-70.

Wu, D., Hu, T., Zhang, L., Chen, J., Du, J., Ding, J., … Shen, X. U. (2008). Residues Asp164 and Glu165 at the substrate entryway function potently in substrate orientation of alanine racemase from *E. coli* : Enzymatic characterization with crystal structure analysis, 1066–1076. http://doi.org/10.1110/ps.083495908.1066

Yard, B. A., Carter, L. G., Johnson, K. A., Overton, I. M., Dorward, M., Liu, H., … Campopiano, D. J. (2007). The Structure of Serine Palmitoyltransferase ; Gateway to Sphingolipid Biosynthesis, 870–886. http://doi.org/10.1016/j.jmb.2007.04.086

Zhang, W., Bolla, M. L., Kahne, D., & Walsh, C. T. (2010). A three enzyme pathway for 2-amino-3-hydroxycyclopent-2-enone formation and incorporation in natural product biosynthesis. *Journal of the American Chemical Society*, *132*(18), 6402–11. https://doi.org/10.1021/ja1002845

Zhu, H., Shen, H., Sewell, A. K., Kniazeva, M., & Han, M. (2013). A novel sphingolipid-TORC1 pathway critically promotes postembryonic development in Caenorhabditis elegans. *eLife*, *2013*(2), e00429.

Appendix 1. Nucleotide and protein sequences of *Sp*SPT constructs

N-terminal (American) *Sp*SPT nucleotide sequence:

```
   1 GGCAGCAGCC ATCATCATCA TCATCACAGC AGCGGCCTGG TGCCGCGCGG
  51 CAGCCATATG GCTAGCACAG AGGCTGCCGC TCAGCCGCAC GCCCTGCCTG
 101 CAGACGCGCC TGACATTGCT CCTGAACGTG ATTTGTTAAG TAAATTTGAT
 151 GGCCTAATCG CCGAAAGACA AAAACTACTC GATTCAGGTG TAACAGATCC
 201 TTTTGCCATC GTTATGGAGC AAGTTAAGTC TCCAACTGAA GCTGTGATAC
 251 GAGGAAAAGA TACAATCCTG CTCGGAACTT ATAATTACAT GGGAATGACC
 301 TTTGATCCGG ACGTAATTGC AGCAGGTAAG GAAGCATTGG AGAAATTTGG
 351 TAGCGGTACT AATGGTAGCA GAATGCTGAA CGGGACCTTT CATGATCATA
 401 TGGAAGTTGA ACAAGCCTTG CGTGATTTCT ACGGTACGAC GGGAGCTATT
 451 GTCTTTTCAA CTGGTTACAT GGCAAACCTT GGAATCATAT CAACACTTGC
 501 GGGTAAAGGT GAGTATGTTA TTTTGGATGC AGACTCTCAT GCTTCTATCT
 551 ACGATGGCTG TCAACAAGGA AACGCGGAAA TCGTAAGATT CCGGCACAAC
 601 TCAGTGGAGG ATCTTGATAA GAGACTTGGT AGGCTTCCAA AGGAACCAGC
 651 TAAACTCGTT GTTTTGGAAG GCGTTTATTC GATGCTTGGT GATATTGCTC
 701 CTCTCAAGGA GATGGTTGCT GTGGCCAAGA AGCATGGAGC CATGGTGCTT
 751 GTCGACGAGG CACATTCTAT GGGCTTCTTC GGACCAAATG GTAGAGGAGT
 801 GTACGAAGCT CAGGGACTCG AAGGACAGAT AGACTTCGTC GTCGGCACTT
 851 TCTCTAAGTC TGTTGGAACT GTTGGCGGGT TTGTTGTGAG CAATCATCCA
 901 AAATTCGAGG CGGTGAGGTT AGCTTGTAGG CCATATATAT TCACCGCTAG
 951 TTTGCCCCCG AGTGTCGTAG CTACAGCTAC CACATCTATA CGAAAACTTA
1001 TGACAGCGCA CGAGAAACGG GAGAGATTAT GGTCTAATGC AAGAGCACTG
1051 CATGGAGGGC TTAAGGCTAT GGGGTTTAGG TTAGGAACGG AGACTTGCGA
1101 CTCCGCTATT GTAGCTGTGA TGCTTGAGGA TCAGGAACAA GCTGCTATGA
1151 TGTGGCAGGC CTTGTTAGAT GGTGGACTAT ATGTTAACAT GGCAAGACCT
1201 CCTGCAACCC CTGCTGGTAC GTTTCTCCTA CGTTGTTCCA TTTGCGCTGA
1251 ACACACTCCC GCACAAATCC AGACTGTTTT GGGGATGTTT CAAGCTGCGG
1301 GACGCGCAGT CGGGGTTATT GGTTGAGCGA ATTCGAGCTC CGTCGACAGG
1351 CTTGCGGCCG CA
```

1362 bp

N-terminal *Sp*SPT amino acid sequence ("American") in pET28a plasmid

MGSSHHHHHHSSGLVPRGSHMASTEAAAQPHALPADAPDIAPERDLLSKFDGLIAERQKLLDS
GVTDPFAIVMEQVKSPTEAVIRGKDTILLGTYNYMGMTFDPDVIAAGKEALEKFGSGTNGSRM
LNGTFHDHMEVEQALRDFYGTTGAIVFSTGYMANLGIISTLAGKGEYVILDADSHASIYDGCQQ
GNAEIVRFRHNSVEDLDKRLGRLPKEPAKLVVLEGVYSMLGDIAPLKEMVAVAKKHGAMVLV
DEAHSMGFFGPNGRGVYEAQGLEGQIDFVVGTFSKSVGTVGGFVVSNHPKFEAVRLACRPYIF
TASLPPSVVATATTSIRKLMTAHEKRERLWSNARALHGGLKAMGFRLGTETCDSAIVAVMLED
QEQAAMMWQALLDGGLYVNMARPPATPAGTFLLRCSICAEHTPAQIQTVLGMFQAAGRAV
GVIG

MW = 47232.0 Da (ExPASy ProtParam)

Molar extinction coefficient = 25900 M^{-1} cm^{-1}

C-terminal (Scottish) *Sp*SPT nucleotide sequence:

```
   1 ATGACCGAAG CCGCCGCTCA GCCCCACGCC CTCCCCGCCG ACGCGCCCGA
  51 CATCGCGCCG GAACGCGACC TGCTCTCCAA GTTCGACGGC CTGATCGCCG
 101 AGCGGCAGAA GCTGCTCGAC TCCGGCGTCA CCGATCCCTT CGCGATCGTG
 151 ATGGAACAGG TGAAGTCGCC GACCGAGGCC GTGATCCGTG GCAAGGACAC
 201 GATCCTGCTC GGCACGTACA ACTATATGGG CATGACCTTC GATCCGGACG
 251 TGATCGCAGC GGGCAAGGAA GCGCTGGAGA AGTTCGGGTC GGGCACCAAT
 301 GGCAGCCGGA TGCTCAACGG CACCTTCCAC GACCATATGG AAGTCGAACA
 351 GGCGCTGCGC GATTTCTACG GCACGACCGG CGCGATCGTC TTTTCGACCG
 401 GTTACATGGC CAATCTCGGC ATCATCTCAA CGCTGGCGGG CAAGGGTGAG
 451 TATGTCATCC TCGACGCCGA CAGCCATGCG TCGATCTATG ACGGCTGCCA
 501 GCAGGGCAAT GCCGAGATCG TCCGCTTCCG CCACAATTCG GTCGAGGATC
 551 TCGACAAGCG GCTGGGCCGT CTGCCCAAGG AACCTGCCAA GCTGGTCGTG
 601 CTGGAGGGCG TCTATTCGAT GCTCGGCGAC ATCGCTCCGC TGAAGGAGAT
 651 GGTCGCGGTC GCCAAGAAGC ATGGCGCAAT GGTCTTGGTC GACGAAGCGC
 701 ATTCGATGGG CTTTTTCGGC CCCAACGGGC GCGGCGTGTA CGAGGCGCAA
 751 GGGTTGGAAG GCCAGATCGA TTTCGTCGTC GGCACCTTCT CCAAATCGGT
 801 CGGCACAGTC GGCGGCTTCG TCGTGTCCAA TCATCCGAAG TTCGAGGCGG
 851 TCCGCCTCGC CTGCCGTCCG TACATCTTCA CCGCCTCGCT GCCGCCCTCG
 901 GTGGTAGCGA CCGCGACGAC GTCGATCCGC AAGCTGATGA CCGCGCATGA
 951 AAAGCGTGAG CGGCTGTGGT CGAATGCCCG CGCGTTGCAT GGCGGGCTGA
1001 AGGCGATGGG CTTCAGGCTC GGCACCGAGA CCTGCGACAG CGCGATCGTC
1051 GCGGTCATGC TGGAGGATCA GGAACAGGCC GCGATGATGT GGCAGGCGCT
1101 GCTCGACGGC GGGCTCTACG TCAACATGGC GCGCCCGCCC GCGACCCCGG
1151 CCGGCACCTT CCTGCTGCGC TGCTCCATCT GTGCCGAGCA CACGCCGGCG
1201 CAGATCCAGA CCGTGCTGGG CATGTTCCAG GCCGCGGGCC GCGCGGTCGG
1251 CGTCATC GGC CTCGAGCACC ACCACCACCA CCAC
```

1284 bp

C-terminal *Sp*SPT amino acid sequence ("Scottish") in pET28a plasmid

MTEAAAQPHALPADAPDIAPERDLLSKFDGLIAERQKLLDSGVTDPFAIVMEQVKSPTEAVIRG
KDTILLGTYNYMGMTFDPDVIAAGKEALEKFGSGTNGSRMLNGTFHDHMEVEQALRDFYGTT
GAIVFSTGYMANLGIISTLAGKGEYVILDADSHASIYDGCQQGNAEIVRFRHNSVEDLDKRLGRL
PKEPAKLVVLEGVYSMLGDIAPLKEMVAVAKKHGAMVLVDEAHSMGFFGPNGRGVYEAQGL
EGQIDFVVGTFSKSVGTVGGFVVSNHPKFEAVRLACRPYIFTASLPPSVVATATTSIRKLMTAHE
KRERLWSNARALHGGLKAMGFRLGTETCDSAIVAVMLEDQEQAAMMWQALLDGGLYVNM
ARPPATPAGTFLLRCSICAEHTPAQIQTVLGMFQAAGRAVGVIGLEHHHHHH

MW (without initial methionine) = 45975.6 Da (ExPASy ProtParam)

Molar extinction coefficient = 25900 M^{-1} cm^{-1}

Appendix 2. Nucleotide and amino acid sequence of *B.fragilis* SPT (BF9343_2380, old locus tag BF2461)

Bf2461 nucleotide sequence:

```
   1 ATGGGATTAT TACAAGAGAA GTTAGCTAAA TACGACCTCC CTCAGCAGAT
  51 AAAGGCTAAA GGCGTATATC CATACTTTCG TTGTATCGAA AGTGAACAGA
 101 ACACAGAGGT GATAATGAGT GGCAGAAAGG TGTTAATGTT TGGCTCAAAC
 151 TCATACTTAG GCCTGACTAA TCATCCGAAA GTAATTGAAG CTGCTGTTGA
 201 AGCTACCCGC AAATATGGTA CAGGTTGCGC CGGATCGCGT TTTCTGAACG
 251 GTACACTCGA CCTCCATCTT CAATTGGAGA AAGAATTGGC CGAATTTGTT
 301 GGTAAAGAAG ATGCTATCAT TTATTCTACC GGATTTCAGG TAAATCTGGG
 351 TGTGGTTTCG TGTGTGACAG GTCGTGAAGA TTATGTGATC TGTGATGAAC
 401 TTGACCACGC TTCTATTGTT GAAGGACGCC GCCTTTCTTT TTCTACCATT
 451 CTTAAGTTCA AGCATAACGA TATGGAATCT CTTGAGAAAG AGTTGCAGAA
 501 ATGTCGTCCT GATGCAGTGA AACTGATTGT AGTAGATGGA GTATTCAGTA
 551 TGGAGGGTGA TATTGCCAAT TTGCCTGAGA TCGTCCGTTT GTCTAAAAAA
 601 TATGATGCTA ATATCATGGT AGATGAAGCG CATGGACTGG GAGTTTTGGG
 651 TAATCACGGA CGCGGTACTT GTGATCATTT CGGATTGACT AAAGAGGTGG
 701 ATCTTATTAT GGGTACATTC AGTAAGTCAT TGGCCGCTAT CGGTGGCTTT
 751 ATTGCAGCAG ACGAGTCCAT CATTAATTAT TTGCGTCACA ATTCACGTTC
 801 ATATATCTTT AGTGCAAGTA ATACGCCTGC TGCTACAGCT GCCGCTCGTG
 851 CTGCACTTCA GATTATGAAA AACGAACCGG AACGTATTGA GCATTTGTGG
 901 GATATAACCA ATTACTCTTT AAAGTGTTTC CGTGAACTTG GTTTTGAGAT
 951 CGGACATACC TCCACTCCTA TCATTCCTCT ATATGTACGT GATATGGAGA
1001 AGACATTTAT GGTAACTAAG ATGTTATTTG ACGAAGGTGT GTTTGTAAAT
1051 CCAGTTGTGC CTCCCGCATG TTCTCCGAAC GATACGTTGA TTCGTTTCTC
1101 GTTGATGGCT ACACACTCTA AAGAACAGAT TGATTTTGCT ATCGGTAAGT
1151 TAGTGAAATG TTTCAAGGCA CTTGATCTTT TATAA
```

***Bf*SPT amino acid sequence in pET28a plasmid:**

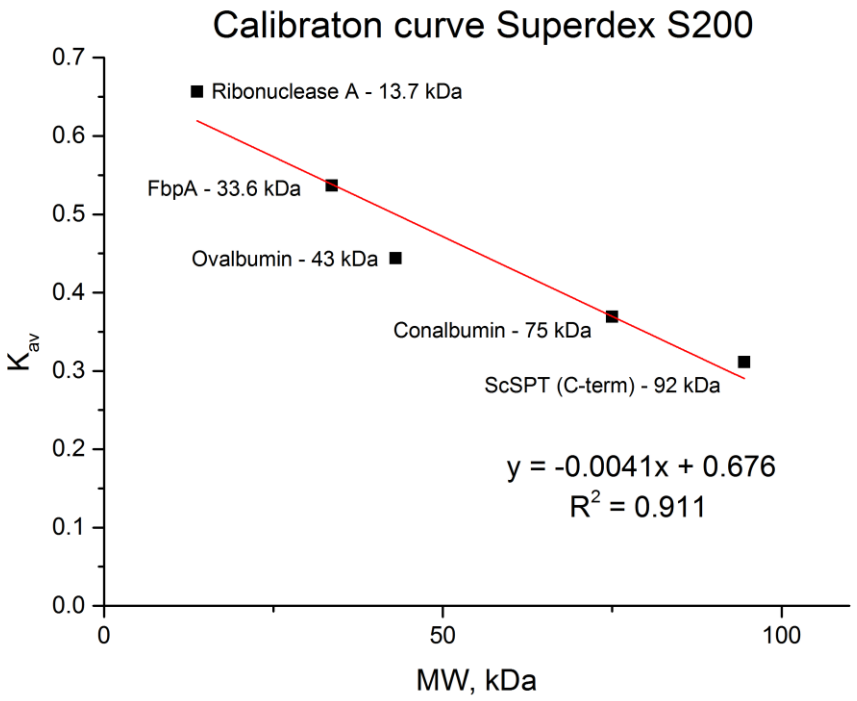
MGLLQEKLAKYDLPQQIKAKGVYPYFRCIESEQNTEVIMSGRKVLMFGSNSYLGLTNHPKVIEAAVEATR
KYGTGCAGSRFLNGTLDLHLQLEKELAEFVGKEDAIIYSTGFQVNLGVVSCVTGREDYVICDELDHASIVE
GRRLSFSTILKFKHNDMESLEKELQKCRPDAVKLIVVDGVFSMEGDIANLPEIVRLSKKYDANIMVDEAH
GLGVLGNHGRGTCDHFGLTKEVDLIMGTFSKSLAAIGGFIAADESIINYLRHNSRSYIFSASNTPAATAAA
RAALQIMKNEPERIEHLWDITNYSLKCFRELGFEIGHTSTPIIPLYVRDMEKTFMVTKMLFDEGVFVNPV
VPPACSPNDTLIRFSLMATHSKEQIDFAIGKLVKCFKALDLLEHHHHHH

Amino acid sequence of *B.fragilis* SPT in pET28a plasmid. Coloured are the catalytic residues (red), PLP-binding motif (green), PACSP loop (blue) and C-terminal histidine extension (purple).

MW (without the initial methionine) = 44614.30 Da (ExPASy ProtParam)

Molar extinction coefficient = 23380 M^{-1} cm^{-1}

Appendix 3. Superdex S200 Gel filtration column calibration curve



App. 3. Calibration curve for the HiLoad 16/600 Superdex 200 column (GE Healthcare) used to estimate the molecular weight of the protein based on its retention time.

The MW of the protein is estimated as follows:

$$\mathrm{MW} = \mathrm{e}\left(\frac{K_{av} - 0.676}{-0.0041}\right)$$

$$K_{av} = \frac{V_e - \mathrm{V}_o}{V_t - \mathrm{V}_o}$$

V_e = elution volume

V_o = Void volume (43.4 mL)

V_t = Total bed volume (120 mL)

Appendix 4. Cyclsoserine-derived structures in the pdb

Pdb code:	Molecule:	Organism:	Enzyme:	Ligand config:	Reference:
1D7S	DCS	*Burkholderia cepacia*	dialkylglycine decarboxylase	3-isoxazolidinone (sp^3)	Malashkevich, 1999
1D7U	LCS	*Burkholderia cepacia*	dialkylglycine decarboxylase	3-hydroxyisoxazole (sp^2)	Malashkevich, 1999
1EPV	DCS	*Geobacillus stearothermophilus*	alanine racemase	3-hydroxyisoxazole (sp^2)	Fenn et al., 2003
1NIU	LCS	*Geobacillus stearothermophilus*	alanine racemase	3-hydroxyisoxazole (sp^2)	Fenn et al., 2003
1I2L	DCS	*E.coli*	deoxychorismate lyase	3-hydroxyisoxazole (sp^2)	Jensen et al. (unpubl)
1VFS	DCS	*Streptomyces lavendulae*	alanine racemase	3-hydroxyisoxazole (sp^2)	Noda et al., 2004
1VFT	LCS	*Streptomyces lavendulae*	alanine racemase	3-hydroxyisoxazole (sp^2)	Noda et al., 2004
1MDZ	LCS	*Salmonella enterica*	ArnB aminotransferase	3-hydroxyisoxazole (sp^2)	Noland et al., 2004
2DAA	DCS	*Bacillus sp.*	D-aminoacid aminotransferase	3-isoxazolidinone (sp^3)	Peisach et al., 1998
2RJH	DCS	*E.coli*	alanine racemase	3-isoxazolidinone (sp^3)	Wu et al., 2008
3E6E	DCS	*Enterococcus faecalis*	alanine racemase	3-hydroxyisoxazole (sp^2)	Priyadarshi et al., 2009
4OMA	LCS	*Citrobacter freundii*	methionine gamma-lyase	3-hydroxyisoxazole (sp^2)	Kuznetsov et al., 2015
4LUT	*DCS*	*Clostridium difficile*	alanine racemase	3-isoxazolidinone (sp^3)	Asojo et al., 2014
4D9E	LCS	*Salmonella enterica*	D-cysteine desulfhydrase	3-isoxazolidinone (sp^3)	Bharath et al., 2012
2XBN	LCS	*Sphingomonas paucimobilis*	serine palmitoyltransferase	β-aminooxyacetaldehyde	Lowther et al., 2010
5U3F	DCS	*Mycobacterium Tuberculosis*	branched-chain aminotransferase	3-hydroxyisoxazole (sp^2)	Amorim Franco et al., 2017

PDB structural entries of various PLP-dependent enzymes containing cycloserine-derived adducts. 3-isoxazolidinone has a tetrahedral sp^3 configuration around C of , while 3-hydroxyisoxazole is a planar ring in sp^2 configuration.

Appendix 5. Sequences of putative microbiome-related isoSPTs identified by protein BLAST

*Bf*SPT sequence Genbank CAH08161.1 (Q5LCK4):

```
  1 MGLLQEKLAK YDLPQQIKAK GVYPYFRCIE SEQNTEVIMS GRKVLMFGSN SYLGLTNHPK
 61 VIEAAVEATR KYGTGCAGSR FLNGTLDLHL QLEKELAEFV GKEDAIIYST GFQVNLGVVS
121 CVTGREDYVI CDELDHASIV EGRRLSFSTI LKFKHNDMES LEKELQKCRP DAVKLIVVDG
181 VFSMEGDIAN LPEIVRLSKK YDANIMVDEA HGLGVLGNHG RGTCDHFGLT KEVDLIMGTF
241 SKSLAAIGGF IAADESIINY LRHNSRSYIF SASNTPAATA AARAALQIMK NEPERIEHLW
301 DITNYSLKCF RELGFEIGHT STPIIPLYVR DMEKTFMVTK MLFDEGVFVN PVVPPACSPN
361 DTLIRFSLMA THSKEQIDFA IGKLVKCFKA LDLL
```

Porphyromonas gingivalis WP_021679668.1 (W1R7E5);

76% identity with *Bf*SPT

```
  1 LLQDKLAQYT EPQKAQAAGI YPYFRKIESD QDTEVVIDGR KVLMFGSNAY LGLTNHPKVK
 61 EAAIEATKKY GTGCAGSRFL NGTLDIHLEL EKRLAEFVGK EDAISFSTGF QVNLGVVSCI
121 TGREDYIIWD ELDHASIIEG IRLSFSTKLK YKHNDMGSLE KRLQQCDPDK IKLIVVDGVF
181 SMEGDVCNLP EIVRLAKRYN ANVMVDEAHG IGVMGDHGRG VCNHFGLTDE VDLIMGTFSK
241 SFASLGGFIA GDKSVINYLR HHARSYIFSA SCTPASTAAA AAALDIMLSE PERLARLWEL
301 THYSLNAFRS LGFEIGHTST PIIPLFIRNN EKTFQITRDA FEEGVFVNPV VSPAVAPSDT
361 LIRFSLMATH TKEQLDFAIE KLHKVFKQNS VL
```

Parabacteroides goldsteinii (77% identity with *Bf*SPT) WP_010802415.1 (S0GTM6):

```
  1 MKLLQEKLAK YDAPQKAMAA GIYPYFRMIE SDQDTEVMIS GKKVLMFGSN AYLGLTNHPK
 61 VKEAAIEAIK KYGTGCAGSR FLNGTLDIHI QLEKRLAEFV GKEDAIVYST GFQVNLGVVS
121 CLTGREDYIL WDELDHASII EGHRLSFSTK LKYKHNDMDS LEKQLQKCEP DKVKLIVIDG
181 VFSMEGDIAK LPEIVALAKK YNASIMVDEA HGLGVLGDHG RGTCNHFGVT DDVDLIMGTF
241 SKSLASIGGF IASDKDTINY LRHNSRSYIF SASNTPAATA AAGAALDIMQ SEPERIEHLW
301 KLTHYALDGF RNMGCEIGHT STPIIPLFIR DNDLTFLIVK ELFEAGIFVN PVVSPAVAPE
361 DTLIRFSLMA THTKEQLDYA LEAIHKVFKS HGLVD
```

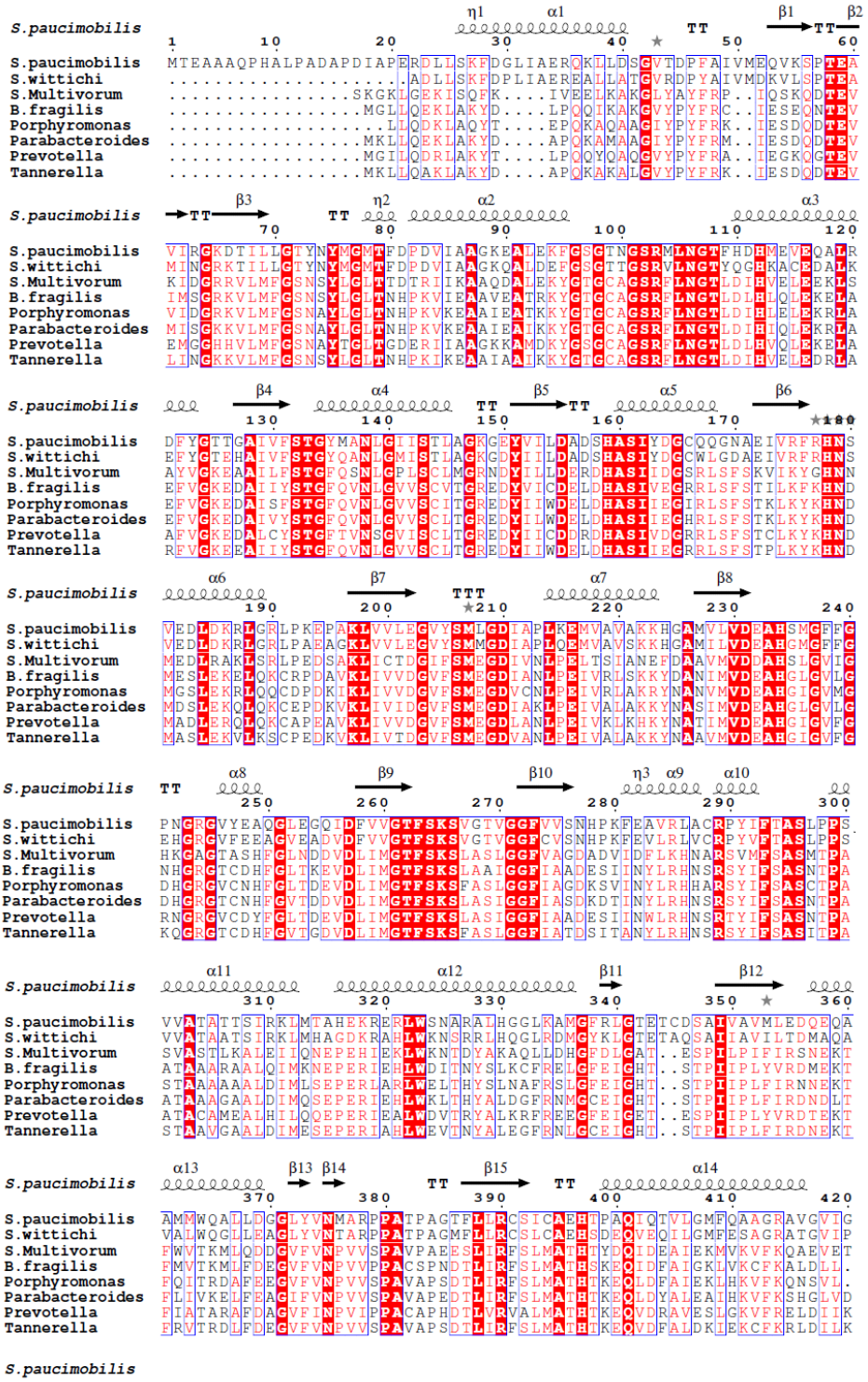
Prevotella sp 73% identity: GenBank ID: CDE86132.1 (R7LFZ5)

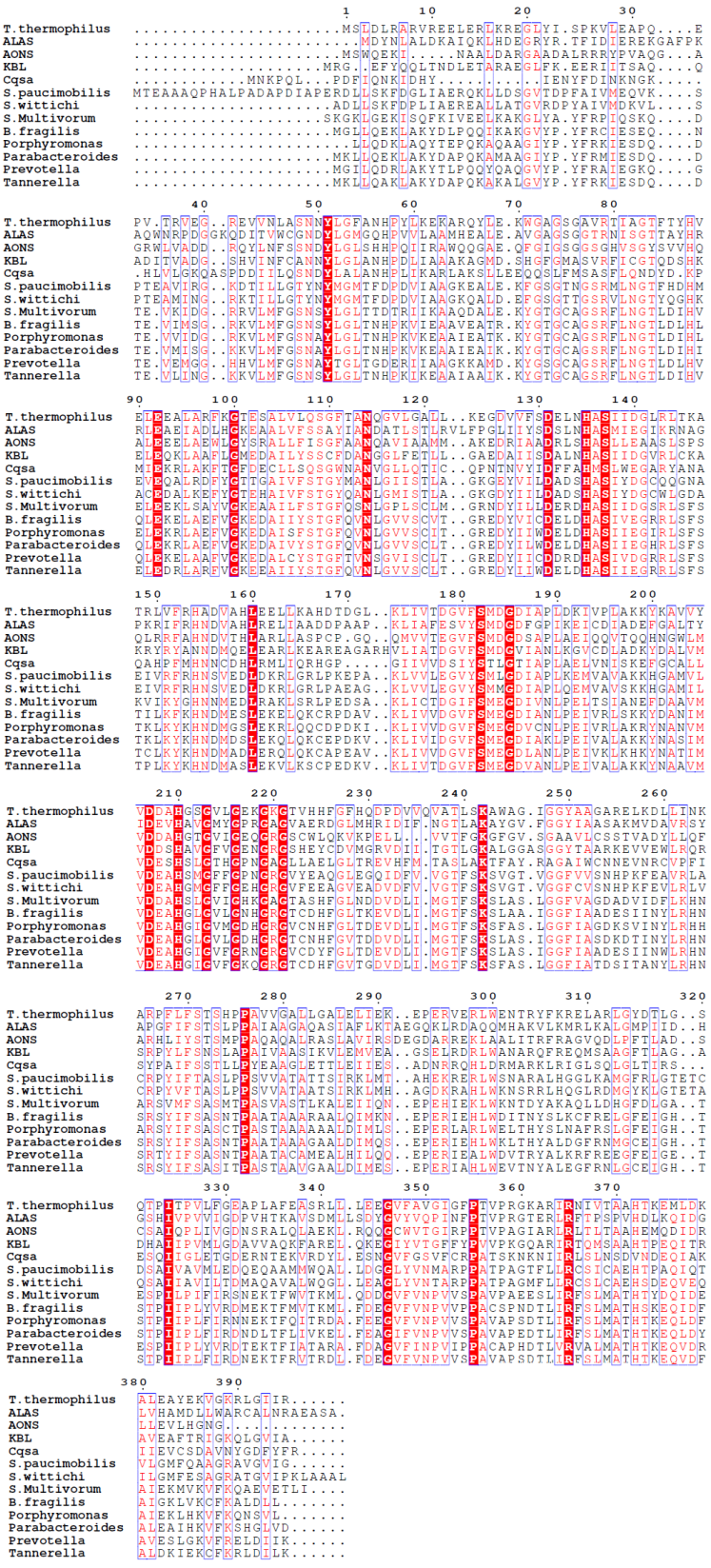
```
  1 MGILQDRLAK YTLPQQYQAQ GVYPYFRAIE GKQGTEVEMG GHHVLMFGSN AYTGLTGDER
 61 IIAAGKKAMD KYGSGCAGSR FLNGTLDLHV QLEKELAAFV GKEDALCYST GFTVNSGVIS
121 CLTGREDYII CDDRDHASIV DGRRLSFSTC LKYKHNDMAD LERQLQKCAP EAVKLIVVDG
181 VFSMEGDLAN LPEIVKLKHK YNATIMVDEA HGIGVFGRNG RGVCDYFGLT DEVDLIMGTF
241 SKSLASIGGF IAADESIINW LRHNSRTYIF SASNTPAATA CAMEALHILQ QEPERIEALW
301 DVTRYALKRF REEGFEIGET ESPIIPLYVR DTEKTFIATA RAFDAGVFIN PVIPPACAPH
361 DTLVRVALMA THTKEQVDRA VESLGKVFRE LDIIK
```

Tannerella sp. 76% (CCY38459.1), (R5IDR1)

```
  1 MKLLQAKLAK YDAPQKAKAL GVYPYFRKIE SDQDTEVLIN GKKVLMFGSN SYLGLTNHPK
 61 IKEAAIAAIK KYGTGCAGSR FLNGTLDIHV ELEDRLARFV GKEEAIIYST GFQVNLGVVS
121 CLTGREDYII WDELDHASII EGRRLSFSTP LKYKHNDMAS LEKVLKSCPE DKVKLIVTDG
181 VFSMEGDVAN LPEIVALAKK YNAAVMVDEA HGIGVFGKQG RGTCDHFGVT GDVDLIMGTF
241 SKSFASLGGF IATDSITANY LRHNSRSYIF SASITPASTA AVGAALDIME SEPERIAHLW
301 EVTNYALEGF RNLGCEIGHT STPIIPLFIR DNEKTFRVTR DLFDEGVFVN PVVSPAVAPS
361 DTLIRFSLMA THTKEQVDFA LDKIEKCFKR LDILK
```

Appendix 6. Sequence alignment of 8 SPTs and 14 AOS enzymes




T.thermophilus
ALAS
AONS
KBL
Cqsa
S.paucimobilis
S.wittichi
S.Multivorum
B.fragilis
Porphyromonas
Parabacteroides
Prevotella
Tannerella