Biochemical characterization of signal peptide processing enzymes: *Staphylococcus aureus* signal peptidase I and *Escherichia coli* signal peptide peptidase A2

by

Daniel Chiang

B.Sc., Simon Fraser University, 2008

Thesis Submitted In Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Molecular Biology and Biochemistry
Faculty of Science

SIMON FRASER UNIVERSITY

Spring 2015

Approval

Name: **Daniel Chiang**

Degree: **Doctor of Philosophy**

Title of Thesis: ***Biochemical characterization of signal peptide processing enzymes: Staphylococcus aureus signal peptidase I and Escherichia coli signal peptide peptidase A2***

Examining Committee: **Chair:** Dr. Nicholas Harden
Professor

Dr. Mark Paetzel
Senior Supervisor
Professor

Dr. Rosemary Cornell
Supervisor
Professor

Dr. David Vocadlo
Supervisor
Professor

Dr. Margo Moore
Internal Examiner
Professor
Department of Biological Sciences

Dr. Harry Brumer
External Examiner
Professor, Department of Chemistry
University of British Columbia

Date Defended/Approved: April 13th, 2015

Partial Copyright Licence

Abstract

Type I signal peptidase (SPase I) is an essential bacterial enzyme participating in the process of protein secretion. SPase I catalyzes the conversion of pre-proteins to mature proteins by cleaving off the amino-terminal signal peptides from the pre-proteins during protein secretion. The removal of these remnant signal peptides, required for the continuation of the secretion process, is not a well understood process in bacteria. In Escherichia coli, signal peptide peptidase A (SppA), together with other enzymes, is responsible for the removal of these remnant signal peptides. This thesis work focuses on characterizing and comparing *Staphylococcus aureus* SPase I (SpsB) with other bacterial SPase I as well as characterizing a previously unexamined SppA related enzyme, *E. coli* signal peptide peptidase A2 (SppA2).

A fluorescent lipidated peptide substrate with a Gram-positive signal peptide sequence was used to characterize the Michaelis-Menten kinetic constants for *S. aureus* SpsB along with SPase I from *Bacillus subtilis*, *Staphylococcus epidermidis*, and *E. coli*. *E. coli* SPase I has a significantly lower catalytic efficiency towards the substrate. A previously characterized *E. coli* pre-protein was mutated to match the sequence of the Gram-positive peptide sequence, leading to a significantly reduced maturation rate by *E. coli* SPase I, both *in vitro* and *in vivo*. These results have led to the discovery of a previously uncharacterized residue in the SPase I substrate binding groove (proline 88 in *E. coli*) which may contribute to the difference in catalytic efficiency observed between Gram-positive and Gram-negative SPase I enzymes.

Limited proteolysis has revealed that *E. coli* SppA2 has an N-terminal protease sensitive region (residues 39 to 91) and a C-terminal trypsin resistant (trSppA2) domain (residues 92 to 349). Light scattering results indicate that trSppA2 forms octamers in solution with a proposed dome-shape structure similar to that of *E. coli* and *B. subtilis* SppA. Activity and mutagenesis studies demonstrate that trSppA2 can digest both small peptides and folded proteins, with a preference for hydrophobic substrates, while the S178A and K230A mutants are inactive suggesting that SppA2 is a serine / lysine dyad enzyme. Lastly, the protease sensitive region is required for proper protein folding and may regulate substrate traffic into and out of the inner cavity of the SppA2 octamer.

Keywords: Signal peptidase I; Signal peptide peptidase A2; Ser/Lys proteases; Enzyme kinetics; Enzyme/substrate interactions; Protein translocation.

To my family

Acknowledgements

I would like to thank Dr. Mark Paetzel for taking me in as a Ph.D. student and for the opportunities he has provided me with to explore science. This has allowed me to learn many valuable lessons, techniques and skills which will be invaluable later in my life.

I would also like to thank Dr. Rosemary Cornell and Dr. David Vocadlo for their valuable advice over my Ph.D. career. Thank you to Dr. Margo Moore for being my internal examiner, Dr. Nicholas Harden for being my defense chair and Dr. Harry Brumer for being my external examiner.

I would like to express my gratitude to Dr. Apollos Kim for being my mentor, sparking my interest in science and ultimately inspiring me in pursue a Ph.D. degree. Many thanks to Deidre and Yun who have helped me through my Ph.D. career and provided me with insightful conversations on a wide variety of topics. Also many thanks to current and past Paetzel lab members and friends from the MBB department.

Lastly, I would like to thank my family for supporting me through my B.Sc. degree and now my Ph.D. degree.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AMC	Amino methylcoumarin
CD	Circular dichroism
CMC	Critical micelle concentration
CV	Column volume
DDM	n-Dodecyl β-D-maltoside
DMSO	Dimethyl sulfoxide
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
Em	Emission wavelength
Ex	Excitation wavelength
FRET	Förster resonance energy transfer
IB	Inclusion body
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal titration calorimetry
LepB	Leader peptidase I
MALS	Multiangle light scattering
MCA	Methylcoumaryl amide
mRFU	*Milli*-relative fluorescence unit
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PONA	Pro-OmpA nuclease A
PVDF	Polyvinylidene fluoride
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SipS	Signal peptidase IS
SipT	Signal peptidase IT
SP	Signal peptide
SPase I	Signal peptidase I
SpsB	Signal peptidase IB
SppA	Signal peptide peptidase A

SppA2	Signal peptide peptidase A2
WT	Wild type

Glossary

Ångström (Å)	A unit of length. 1Å = 1.0 × 10^{-10} m.
Amino methylcoumarin	See methylcoumaryl amide.
Critical micelle concentration	The concentration at which detergents spontaneously go into a micelle.
Column volume	The total volume of a chromatography column.
Competitive inhibitor	An enzyme inhibitor which directly competes for the enzyme's active site and prevents binding of the substrate.
Dabcyl	4-((4-(dimethylamino)phenyl)azo)benzoic acid is a molecule capable of absorbing light. It is commonly used as a fluorescence acceptor (quencher) in a FRET assays.
Dalton (Da)	A non-SI unit used to describe the molar mass of a protein molecule. 1 Da = 1g•mol^{-1}.
Dimethyl sulfoxide	A polar aprotic solvent commonly used to dissolve hydrophobic peptide substrates.
n-Dodecyl β-D-maltoside	A non-ionic detergent for solubilizing and stabilizing membrane proteins.
EDANS	5-((2-Aminoethyl)amino)naphthalene-1-sulfonic acid is a fluorescent reporter commonly used as a donor molecule in conjunction with Dabcyl as a quencher in FRET assays.
Hexahistidine tag	A protein affinity tag usually attached at the N-terminus and/or the C-terminus of the target protein. Used in conjunction with Ni^{2+}-NTA affinity chromatography.
Inclusion body	Aggregates of stainable substances inside a cell mostly made up of over expressed proteins which are misfolded with very little DNA, RNA and host proteins.
k_{cat}	A constant describing the rate of conversion from enzyme/substrate complex to enzyme and product which has a unit of per second (s^{-1}). Also known as turnover number or rate of catalysis.
k_{cat}/K_M	A measure of catalytic efficiency of an enzyme acting on a specific substrate which has a unit of per mole per second (M^{-1}•s^{-1}). Also know as specificity constant or kinetic efficiency.
k_d	A constant describing the formation of enzyme/substrate complexes from enzymes and substrate and *vice versa*.
K_M	A constant describing the amount of substrate required for an enzyme to reach half of its maximum reaction rate which has

	a unit of concentration in molar (M). Also known as Michaelis constant.
Methylcoumaryl amide	A fluorescent reporter attached to the C-terminus of a peptide substrate. Its intensity can be measured once it is liberated from the peptide substrate.
Micelle	An aggregation of detergent molecules where the micelle surface is formed by hydrophilic portion of the detergent while the hydrophobic portion forms the micelle center.
Ni^{2+}-NTA affinity chromatography	A protein purification method isolating proteins with hexahistidine tags.
Periplasm	A cellular compartment in between the outer membrane and the inner membrane found in Gram-negative bacteria.
Polymerase chain reaction	A technique used to amplify specific DNA fragments.
Size exclusion chromatography	A protein purification method separating proteins based on their Stokes radii.
Specific activity	The amount of a specific substrate converted to product per mg of a specific enzyme per unit time.
Stokes radius	The radius of a hard sphere which diffuses at the same rate as the protein molecule.
Tris	Tris(hydroxymethyl)aminomethane, a commonly used buffering component with a pKa value of 8.07 at 25 °C.
Triton X-100	A non-ionic detergent for solubilizing and stabilizing membrane proteins. Triton X-100 comes as a mixture with each Triton X-100 molecule varying in the length of polyethylene oxide chain.
Void volume	The volume of the mobile phase required to move a molecule which never enters the stationary phase through a chromatography column.

1. Introduction

1.1. Overview

In this chapter, the process of pre-protein secretion through the general secretory pathway in bacteria will be covered. The two enzymes of this thesis' focus, type I signal peptidase (SPase I) and signal peptide peptidase A2 (SppA2) will be discussed. The topics include what is known about these enzymes, how and where they fit into the secretory pathway, and the outstanding questions that this thesis will address.

1.2. Protein secretion in bacteria

Protein secretion is an essential cellular process found in all three domains of life. Secretory proteins include proteins that cross to the *trans* side of biological membranes and proteins that insert into biological membranes. Secretory proteins can differ greatly from each other in size, structure and function. All secreted proteins synthesized in the cytoplasm have to traverse at least one biological membrane before they reach their final destination.

Most bacteria can be categorized as either Gram-positive or Gram-negative based on their cell wall organization. Gram-positive bacteria have a single phospholipid membrane surrounded by a thick peptidoglycan layer (Silhavy, Kahne & Walker 2010). Gram-negative bacteria have two phospholipid membranes, an inner membrane and an outer membrane, with a thinner peptidoglycan layer sandwiched in between the two membranes (Tokuda 2009). The double membrane of Gram-negative bacteria creates an extra cellular compartment, called the periplasm, which is absent in Gram-positive bacteria (Figure 1.1). Biological membranes act as natural selective filters which allow relatively small and nonpolar molecules to freely diffuse across the membrane while

secretory proteins require specialized protein machinery to effectively translocate them across a biological membrane.

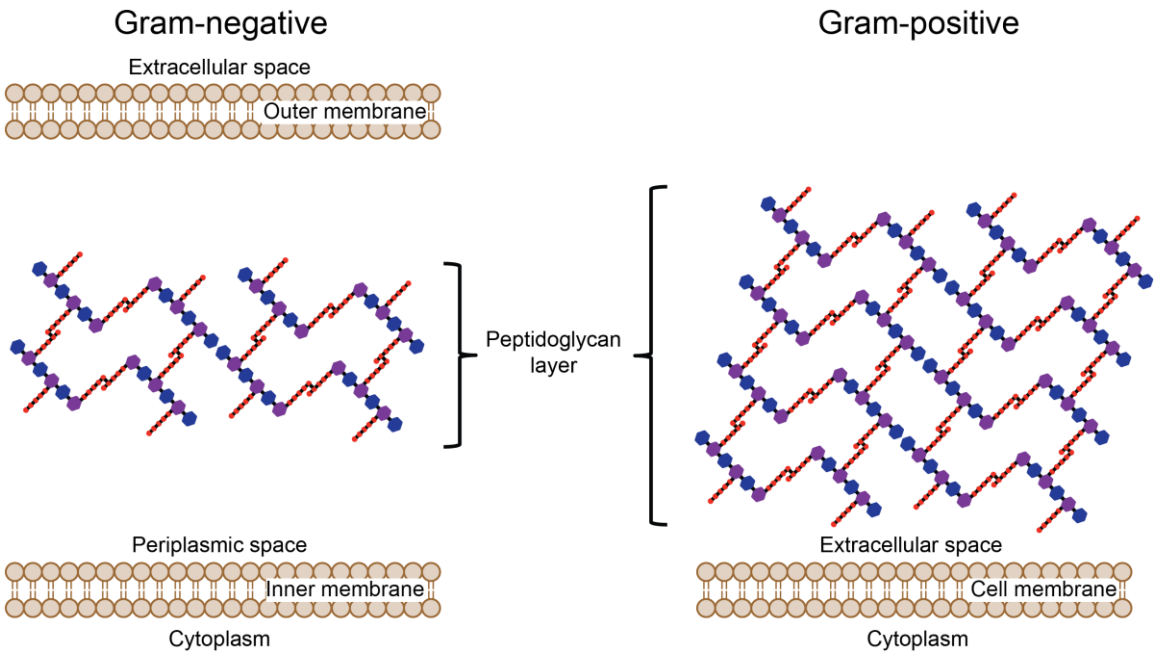


Figure 1.1. Gram-negative and Gram-positive bacterial cell wall.

A simplified depiction of Gram-negative and Gram-positive bacterial cell wall organization showing cell membranes and peptidoglycan layers.

To date, there are sixteen different bacterial secretion systems identified that handle various secretory proteins (Papanikou, Karamanou & Economou 2007). Of those sixteen, the seven major systems have been studied in detail (Table 1.1). Most bacteria have evolved to utilize at least two secretion systems to meet their needs. Of all the different translocation systems, the general secretory (Sec) translocation system is the one that is both ubiquitous and essential (Papanikou, Karamanou & Economou 2007)..The Sec export pathway can be broken down in to three distinct stages: pre-protein sorting and targeting, pre-protein translocation and pre-protein maturation.

Table 1.1. The seven major secretion systems in bacteria.

Secretion systems	Type of substrates secreted	Review
Sec	Majority of screted proteins	Beckwith 2013
Twin-Arginine (TAT)	REDOX proteins	Palmer & Berks 2012
Type I	S-layer proteins, polysaccharides, RTX toxin	Holland et al., 2005
Type II	Cell-wall degrading enzymes, intracellular toxins	Korotkov et al., 2012
Type III	Effector proteins, needle proteins	Cornelis 2006
Type IV	Effector proteins, DNA, DNA-protein complexes	Wallden et al., 2010
Type V	Effector proteins, autotransporters	Henderson et al., 2004

1.2.1. Signal peptides for pre-protein sorting and targeting

Most pre-proteins exported through the Sec translocation system are synthesized with a cleavable amino-terminal sorting sequence, a signal peptide, which is removed from the mature protein. A typical Sec signal peptide is made up of three distinct regions: a positively charged N-region, a hydrophobic H-region and a polar C-region with a SPase I cleavage site (Figure 1.2) (von Heijne 1985). Membrane-embedded proteins are synthesized with a non-cleavable signal peptide which remains with the mature protein as a transmembrane helix (Valent et al. 1998). Both membrane-embedded proteins and secreted pre-proteins must be kept in a non-native state to be exported through the Sec translocation machinery (Economou et al. 2006). This is achieved by two protein-mediated pathways: the signal recognition particle (SRP) mediated pathway and the SecB mediated pathway which are utilized by membrane-embedded proteins and secreted pre-proteins, respectively (Figure 1.3).

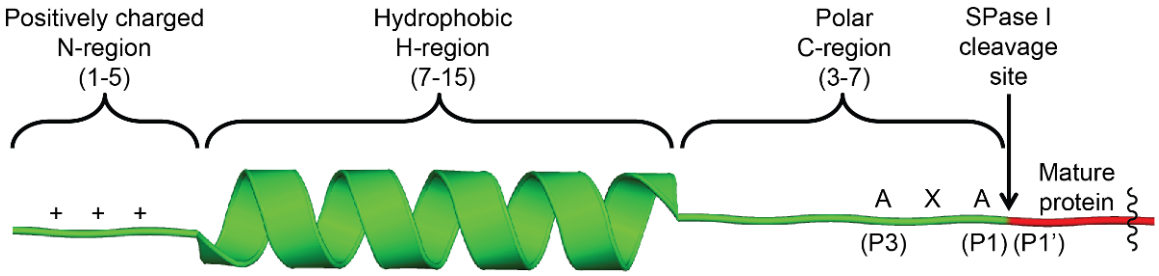


Figure 1.2. A typical Sec translocation signal peptide.

A cartoon depiction of the amino-terminal region of a pre-protein. The signal peptide is coloured in green and the mature protein is coloured in red. Numbers in brackets are the average number of residues found in each region. The SPase I recognition sequence is labeled as AXA where A stands for alanine and X can be any amino acid. P1 (Schechter and Berger nomenclature) (Schechter & Berger 1963) is defined as the residue right before the SPase I cleavage site and P1' is the residue immediately after the cleavage site.

The SRP pathway leads to co-translation translocation of membrane-embedded proteins, with SRP docking near the polypeptide exit site on a ribosome and binding to the nascent polypeptide chain of a membrane-embedded protein as it is being synthesized (Grudnik, Bange & Sinning 2009) . The SecB mediated pathway leads to post-translational translocation of pre-proteins where SecB, a tetrameric chaperone, binds to fully synthesized pre-proteins, keeping them in a non-native state (Randall et al. 1997).

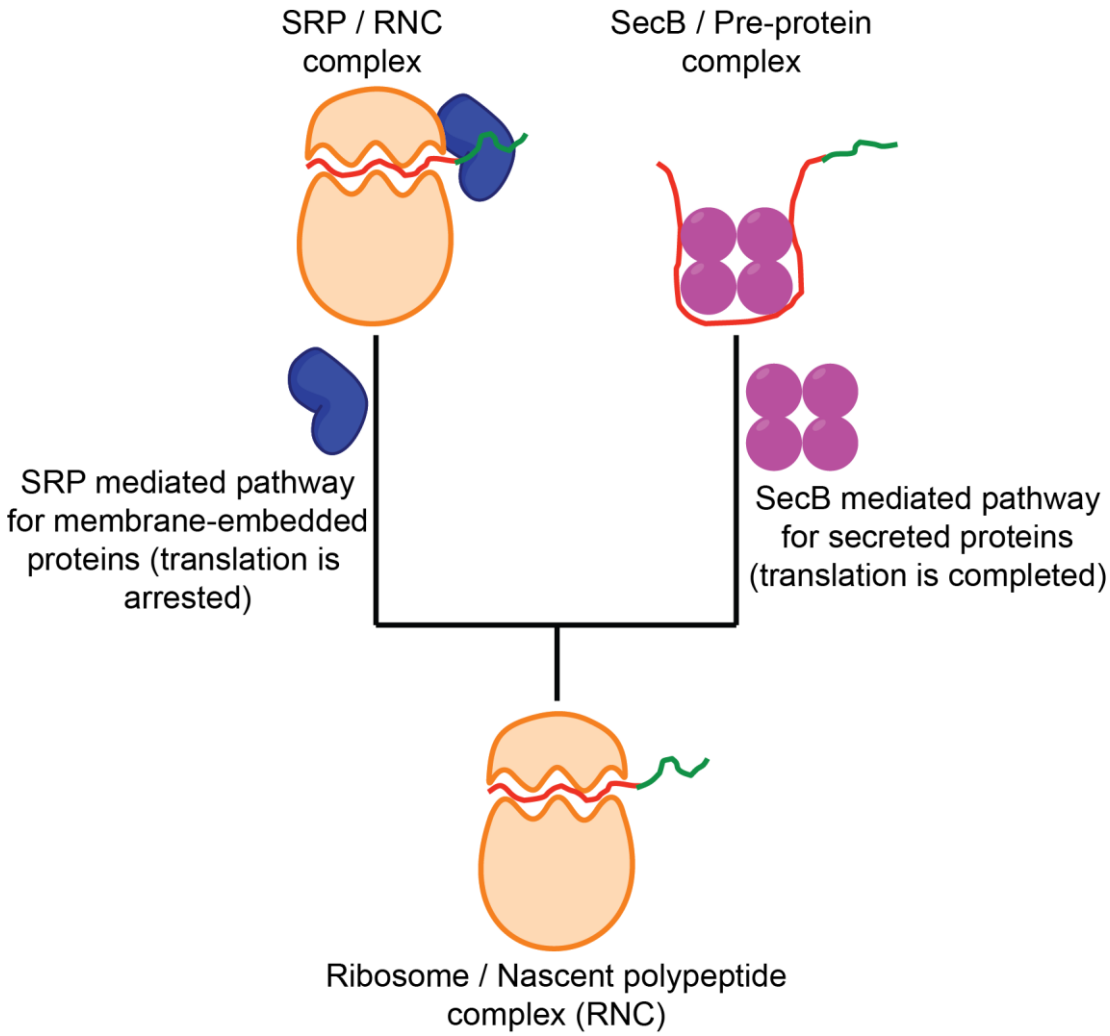


Figure 1.3. SRP and SecB mediated pre-protein translocation pathways.

A schematic showing the two separate pathways pre-proteins can take, depending on their final destination. The signal peptide of a pre-protein emerging from the ribosome is coloured in green while the mature protein is coloured in red. For membrane-embedded proteins, the signal recognition particle (SRP) binds to the newly emerged signal peptide and arrests further translation. For secreted proteins, the tetrameric SecB chaperones bind to fully synthesized pre-proteins and keeps them in unfolded states for translocation.

SRP/pre-protein and SecB/pre-protein complexes are targeted and docked to their respective receptors in complex with the Sec translocation machinery on the membrane. The SRP/pre-protein complex is targeted to a membrane receptor FtsY (Grudnik, Bange & Sinning 2009). Upon GTP hydrolysis by FtsY, the SRP/FtsY complex dissociates and the ribosome nascent chain is transferred to the Sec translocation machinery (Luirink et al. 1994). The SecB/pre-protein complex is targeted to SecA which

transfers the unfolded polypeptide chain to the Sec translocation machinery through SecA (Randall et al. 2005).

1.2.2. Pre-protein translocation through the Sec translocation machinery

The Sec translocation machinery consists of three functionally distinct units: the SecYEG protein conducting channel complex, the SecA ATPase motor and the SecDFYajC accessory protein complex (Figure 1.4).

SecY is an hourglass shaped transmembrane protein with a central pore which opens on both sides of the membrane. The hourglass shape provides the SecY channel with a constriction in the middle of the channel preventing ion leakage and allowing SecY to utilize an α-helix as a plug to close the pore in the absence of a substrate (Van den Berg et al. 2004). SecE is a membrane anchored protein which stabilizes the SecY channel (Kihara, Akiyama & Ito 1995) while SecG is a non-essential membrane anchored protein which stimulates pre-protein translocation by facilitating binding between the SecY channel and the SecA ATPase motor (Nishiyama, Suzuki & Tokuda 1996).

The SecA ATPase motor belongs to a helicase protein family which utilizes ATP hydrolysis to drive conformational change (Vrontou & Economou 2004). SecA contains an NBD (nucleotide binding domain) and an IRA2 (intramolecular regulator of ATP hydrolysis 2) domain to drive conformation change upon ATP hydrolysis (Sianidis et al. 2001). SecA also has two additional domains specific for pre-protein translocation: a PBD (pre-protein binding domain) within the NBD and a C-terminal domain fused to the IRA2. The PBD is responsible for binding signal peptides and the C-terminal domain interacts with the NBD and the IRA2 to prevent ATP hydrolysis in the absence of a pre-protein substrate (Karamanou et al. 2007). Conformational changes in both the PBD and the C-terminal domain upon ATP hydrolysis push the pre-protein through the SecYEG channel (Economou & Wickner 1994).

The SecD, SecF and YajC heterotrimer is not essential for life and not required for protein translocation, however, it can associate with the SecYEG channel (Duong & Wickner 1997a, Duong & Wickner 1997b). In some bacteria, the SecDF dimer regulates

SecA's association with the membrane, preventing backsliding of the pre-protein during translocation and facilitating the release of the pre-protein upon completion of translocation (Matsuyama, Fujita & Mizushima 1993).

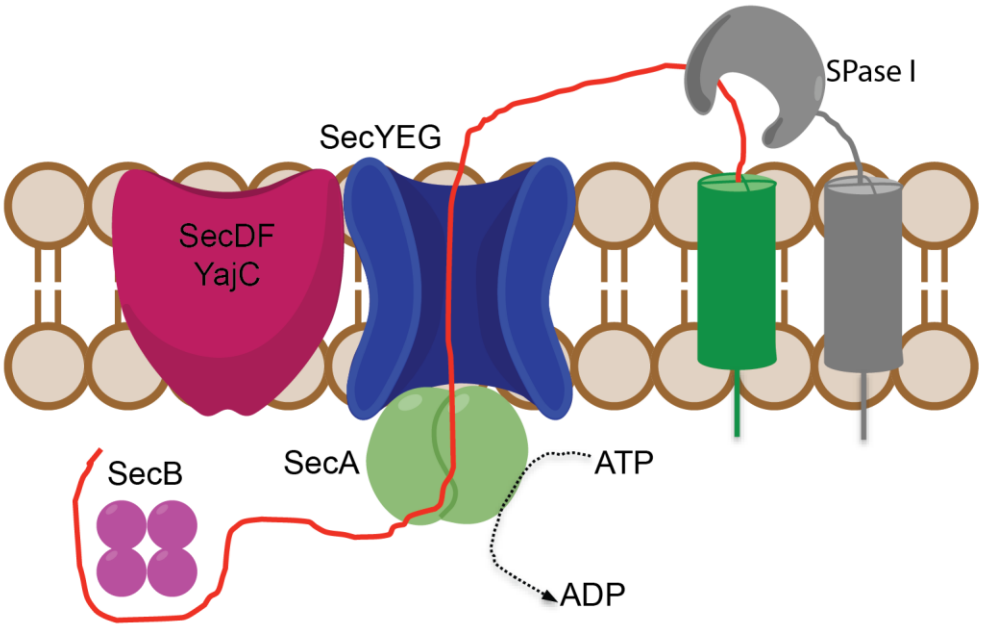


Figure 1.4. Pre-protein translocation through the SecYEG channel

Schematic representation of proteins involved in the Sec translocation machinery at the membrane surface. The pre-protein is pushed through the SecYEG protein-conducting channel by conformational changes in the SecA ATPase through cycles of ATP hydrolysis. Part of the pre-protein yet to enter the channel is kept unfolded and stabilized by SecB. The non-essential SecDFYajC heterotrimer stabilizes the SecYEG channel. SPase I removes the signal peptide from the pre-protein during or immediately after translocation.

1.2.3. Pre-protein maturation

Pre-protein maturation is a process where the signal peptide of a pre-protein is removed by a signal peptidase (SPase) during or shortly after translocation (Auclair, Bhanu & Kendall 2012) (Figure 1.4). In bacteria, there are three different types of SPase, each specific to a subset of pre-proteins. SPase II cleaves lipoprotein signal peptides and SPase IV cleaves prepilin signal peptides (Tettelin et al. 2002, Tjalsma et al. 1997) while SPase I is responsible for processing the majority of pre-proteins not processed by the other two SPases.

1.2.4. Remnant signal peptide processing

After a successful pre-protein secretion event, the mature protein is released on the *trans* side of the membrane while the cleaved signal peptide remains embedded within the membrane. Removal of these remnant signal peptides is an important task for the cell but little is known about the bacterial signal peptide degradation process (Weihofen & Martoglio 2003). Excess signal peptides embedded in the cell membrane can lead to inhibition of further pre-protein secretion, or in extreme cases, rupture the cell membrane (Chen et al. 1987, van der Wal et al. 1992). Due to the charged properties of signal peptides (a positively charged N-region and a polar C-region), they are unlikely to readily diffuse out of the cell membrane, suggesting enzymes may be involved to catalyze the removal of remnant signal peptides.

To date, proteases from the S49 protease family and site-2 protease family have been found to be responsible for cleaving bacterial signal peptides (Figure 1.5). *E. coli* signal peptide peptidase A (SppA) from the S49 protease family is a tetrameric compartmentalized protease anchored to the inner membrane in the periplasm (Kim, Oliver & Paetzel 2008, Wang et al. 2008). Inhibition of *E. coli* SppA using Antipain (an oligopeptide and known protease inhibitor) leads to signal peptide accumulation in the membrane (Ichihara, Beppu & Mizushima 1984). The deletion of the *sppA* gene in *B. subtilis* results in a reduced pre-protein maturation rate, likely due to the lack of availability of cell membrane space for insertion of newly synthesized pre-proteins. *E. coli* RseP and *B. subtilis* RasP from the site-2 protease family cleave in the H-region of membrane embedded signal peptides only after the liberation of the mature domain by SPase I (Saito et al. 2011). However, although enzymes responsible for remnant signal peptide processing have been discovered, there is no evidence suggesting they work independent of each other or in concert, nor is there evidence to suggest the order in which they work in the remnant signal peptide degradation pathway.

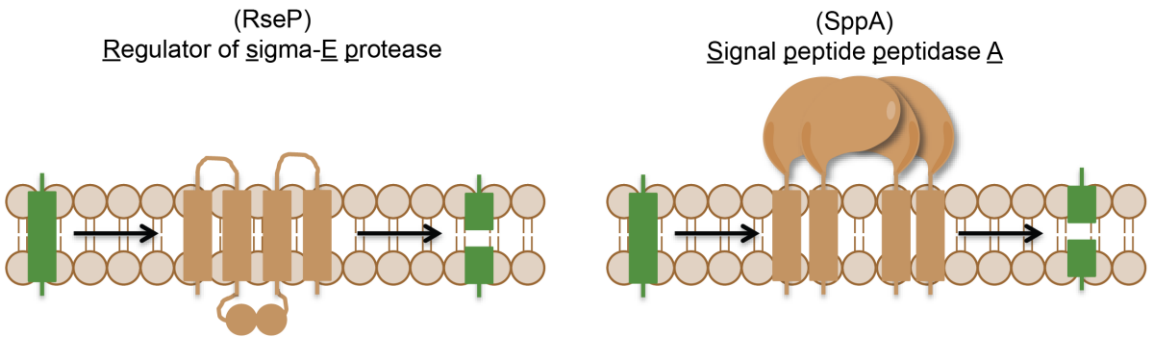


Figure 1.5. E. coli remnant signal peptide processing.

Regulator of sigma-E protease (RseP) has four transmembrane helices with two cytoplasmic PDZ domains between helices 2 and 3. Signal peptide peptidase A (SppA) has a signal transmembrane helix with a major periplasmic catalytic domain. Four SppA promoters come together to form a functionally active SppA tetramer.

1.3. Type I signal peptidase

Pre-protein processing by SPases was first hypothesized by César Milstein's group when they noticed that immunoglobulin light chains were being produced with a higher than expected molecular mass and then processed to a smaller form upon addition of the endoplasmic reticulum membrane fraction (Milstein et al. 1972). This observation sparked interest in other scientists to search for similar systems in different organisms. In bacteria, SPase I is essential for cell viability, making SPase I a potential antibiotic target (Date 1983). To prove SPase I is necessary for cell viability, Date inactivated the *E. coli* genomic SPase I gene through integration of a deletion plasmid containing a partial SPase I gene fragment. No colonies formed when *E. coli* was transformed with this deletion plasmid while transformation using control plasmids with the complete SPase I gene produced colonies. Becanuse PSase I has been show to be an essential enzyme, much effort has been spent in identifying its biological role, catalytic mechanism and three-dimensional structure. To date, *E. coli* SPase I is the most studied and well characterized bacterial SPase I, serving as the bacterial model for this enzyme. *E. coli* SPase I was first cloned, over-expressed and purified in the 1980s (Dalbey & Wickner 1985, Date & Wickner 1981, Zwizinski & Wickner 1980). The first kinetic characterization of *E. coli* SPase I was carried out using an SDS-PAGE gel method with a pre-protein substrate, pro-OmpA nuclease A (PONA) (Tschantz et al. 1995). PONA is a chimeric protein, having an *E. coli* OmpA signal peptide attached to

the N-terminus of nuclease A from *Staphylococcus aureus*. The first *E. coli* SPase I X-ray crystal structure was solved for the catalytically active soluble domain, without the two N-terminal transmembrane membrane segments (Paetzel, Dalbey & Strynadka 1998). The structure provided further evidence that *E. coli* SPase I uses a non-classical serine / lysine (Ser/Lys) dyad catalytic mechanism, which was previously suggested by earlier DNA mutagenesis and chemical modification studies (Black 1993, Paetzel et al. 1997). In addition to *E. coli*, SPase I from *B. subtilis*, *Streptomyces lividans*, *Streptococcus pneumonia*, *S. aureus*, *Mycobacterium tuberculosis* and *Pseudomonas aeruginosa* have since been biochemically characterized (Rao et al. 2009, van Roosmalen et al. 2004, Waite et al. 2012).

1.3.1. S26A peptidase family

Both Gram-positive and Gram-negative SPase I are categorized under the S26A peptidase family in MEROPS – the Peptidase Database (Rawlings et al. 2013). Peptidases from this family are characterized by their use of the Ser/Lys catalytic dyad. Despite being in the same peptidase family, Gram-positive and Gram-negative SPase I are different in many ways. Gram-positive SPase I typically has a single N-terminal transmembrane segment and a C-terminal soluble domain while Gram-negative SPase I has more than one N-terminal transmembrane segments and a larger C-terminal soluble domain (Figure 1.6). Gram-positive bacteria usually have multiple SPase I genes while Gram-negative bacteria usually carry one essential SPase I gene, with the exception of *P. aeruginosa* which has two (Waite et al. 2012).

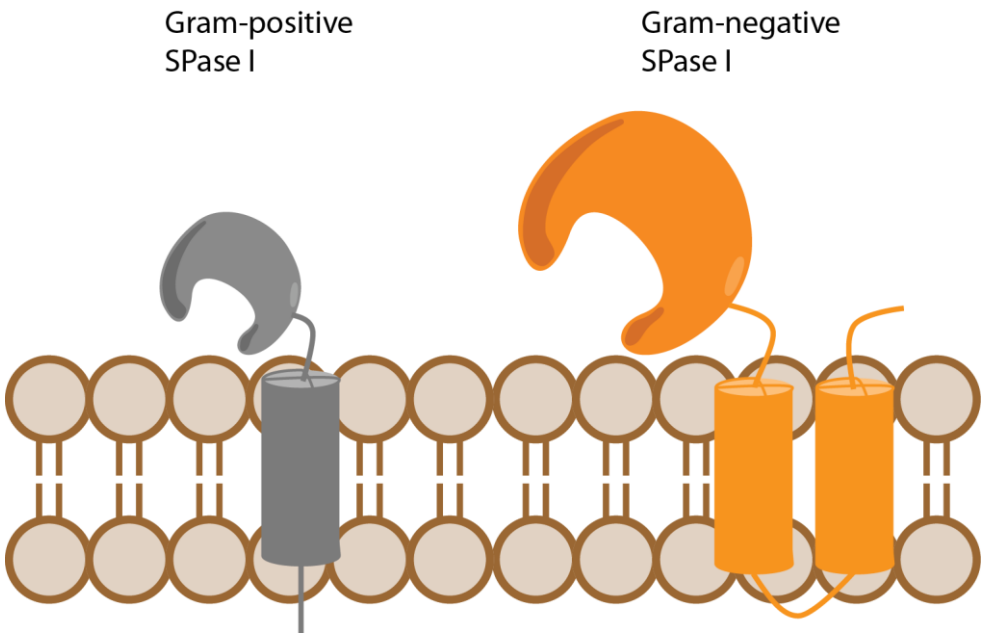


Figure 1.6. Comparison of typical Gram-positive and Gram-negative SPase I.

Schematic of Gram-positive SPase I (grey) and Gram-negative SPase I (orange) anchored to a cell membrane through their N-terminal transmembrane segments.

Although Gram-positive and Gram-negative SPase I are quite different, for example, *E. coli* SPase I and one of the major *B. subtilis* SPase I (SipS) have only 13% sequence identity, but there are five conserved region known as boxes A to E across various species (Figure 1.7) (Dalbey et al. 1997). Box A contains the hydrophobic residues forming the transmembrane segments in SPase I. Box B contains the catalytic serine nucleophile while box C has the residues forming the S3 substrate binding pocket. Box D contains the catalytic lysine general base and the residues making up the S1 substrate binding pocket and box E has a conserved serine thought to align the lysine general base for a better interaction with the serine nucleophile.

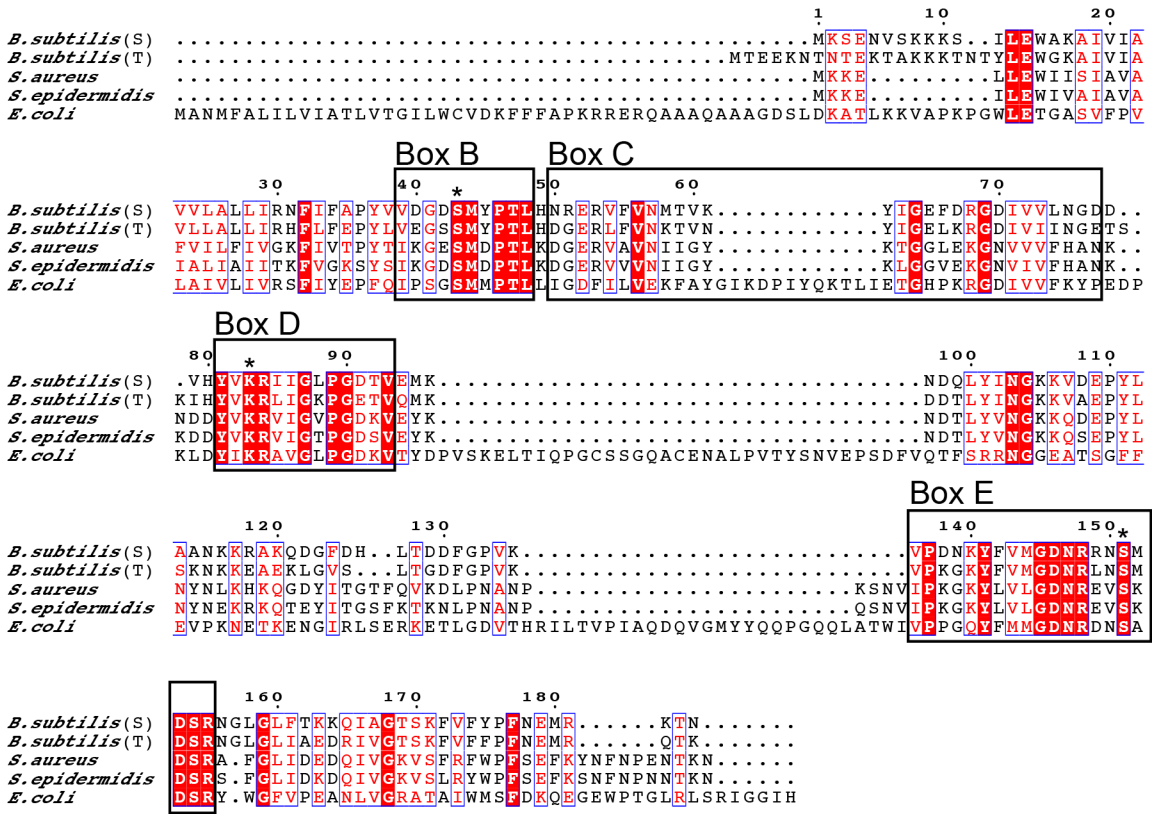


Figure 1.7. Multiple sequence alignments of Gram-positive and Gram-negative SPase I.

Multiple sequence alignments of SPase I enzymes including *B. subtilis* (SipS and SipT), *S. aureus*, *S. epidermidis* and *E. coli* SPase I. The conserved transmembrane region (box A) is not shown but is should be noted that most Gram-negative bacteria have to transmembrane segments while most Gram-positive bacteria have a single transmembrane segment. The nucleophilic serine in box B, the lysine general base in box D and the coordinating serine in box E are marked with an asterisk sign (*). *B. subtilis* SipS numbering is used.

1.3.2. Signal peptidase I X-ray crystal structures

Solved in 1998, the first bacterial SPase I X-ray crystal structure determined was *E. coli* SPase I in complex with a β-lactam inhibitor (Paetzel, Dalbey & Strynadka 1998). Since then, X-ray crystal structures have also been solved for *E. coli* SPase I in complex with other inhibitors (Liu et al. 2011, Luo et al. 2009). However, to date, the *E. coli* SPase I remains the only bacterial SPase I whose structure is available. *E. coli* SPase I is consists primarily of β-strands which make up two large anti-parallel β-sheet domains, a catalytic domain and a cap domain (Figure 1.8). The catalytic domain is highly conserved across species containing the serine nucleophile and the lysine general base. The cap domain is less conserved and varies in size across species.

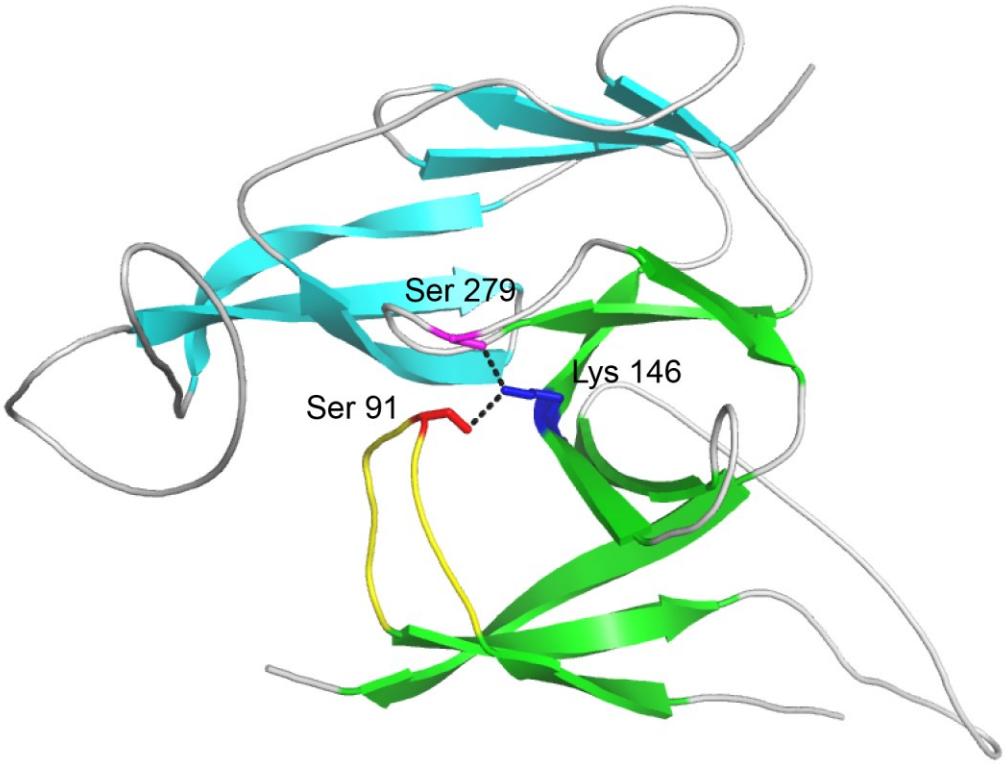


Figure 1.8. E. coli SPase I X-ray crystal structure.

E. coli SPase I in cartoon representation. For simplicity, small secondary structures are shown as loops. The catalytic domain is coloured in green, the loop containing the serine nucleophile in yellow, the cap domain in cyan and the rest of the molecule in white. Side chains of the serine nucleophile, the lysine general base and the coordinating serine are shown as sticks in red, blue and magenta, respectively. Hydrogen bonds are shown as black dotted lines.

The *E. coli* SPase I and arylomycin A_2, a peptide based inhibitor, complex structure has revealed detailed hydrogen bonding interactions between SPase I and the inhibitor (Figure 1.9) (Paetzel et al. 2004). Arylomycin A_2 forms hydrogen bonds with the two SPase I catalytic residues and additional hydrogen bond interactions with the SPase I main chain and side chains of the residues lining the substrate binding groove. In general, SPase I interacts with a peptide substrate or a peptide based inhibitor in an extended β conformation fashion.

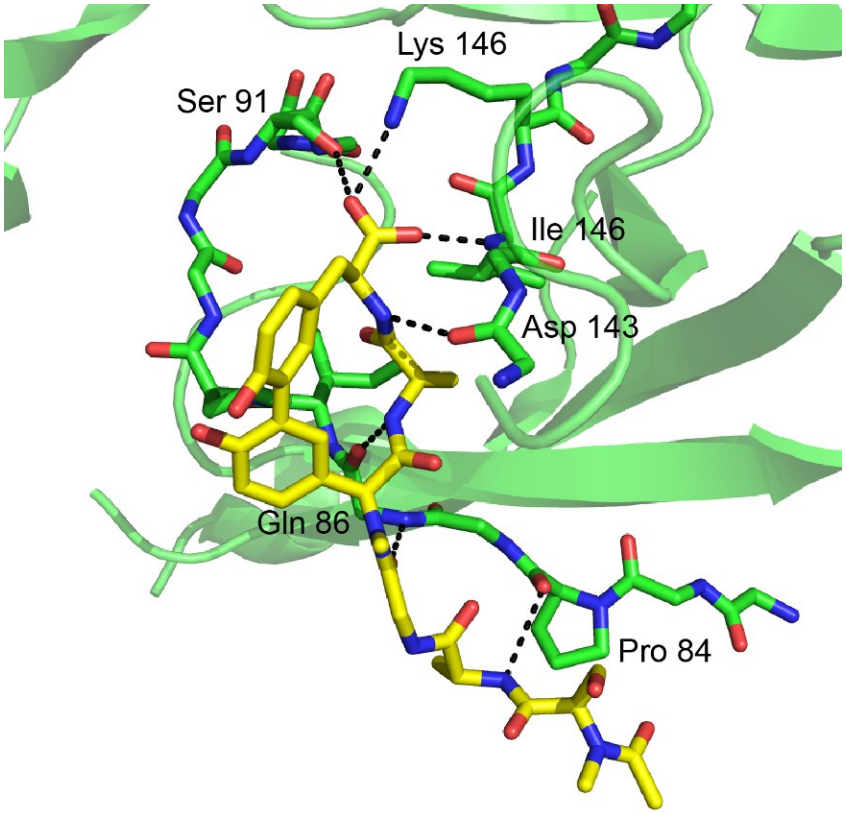


Figure 1.9. E. coli SPase I in complex with arylomycin A_2.

E coli SPase I is shown in cartoon and residues lining the active site are shown in sticks with nitrogen in blue, oxygen in red and carbon in green. For simplicity, side chains which are not hydrogen bonding with arylomycin A_2, are not shown. Arylomycin A_2 is shown in sticks with nitrogen in blue, oxygen in red and carbon in yellow. Hydrogen bonds are shown as dotted black lines.

1.3.3. Signal peptidase I substrates and inhibitors

Schechter and Berger nomenclature will be used in this section to describe the positions of the peptide substrate residues and their respective binding pocket positions within the enzyme (Figure 1.10) (Schechter & Berger 1967).

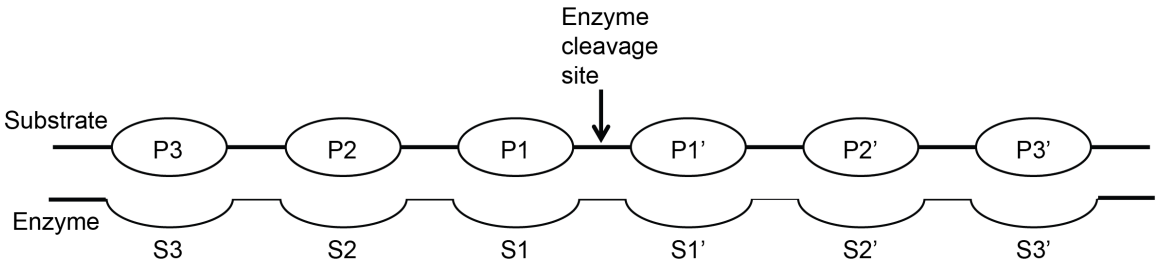


Figure 1.10. Schechter and Berger enzyme / substrate nomenclature.

Substrate residues are given a P designation and a number. The residue immediately preceding the enzyme cleavage site is labeled as P1 and the number increases going towards the N-terminus of the substrate. The residue immediately after the enzyme cleavage site is labeled with a prime sign as P1' and the number increases going towards the C-terminus of the substrate. The corresponding substrate binding pockets within the enzyme are labeled the same way except with an S designation.

Pre-proteins are the endogenous substrates of SPase I; more specifically, the signal peptide portion of the pre-protein directly interacts with SPase I. Most of the SPase I activity characterizations have been done with either pre-proteins or synthetic peptides with sequences based on signal peptides. Signal peptides vary from pre-protein to pre-protein but there are several key features which are shared (Figure 1.11). SPase I has substrate specificity for alanine at both the P1 (-1) and the P3 (-3) positions (Paetzel, Dalbey & Strynadka 2000), large basic residues, such as lysine or arginine, at P2 (-2) and a proline or glycine at P6 (Bruton et al. 2003, Shen et al. 1991). The P1' (+1) position can be any residue, except a proline which renders the signal peptide non-cleavable by SPase I (Barkocy-Gallagher & Bassford 1992, Nilsson & von Heijne 1992)

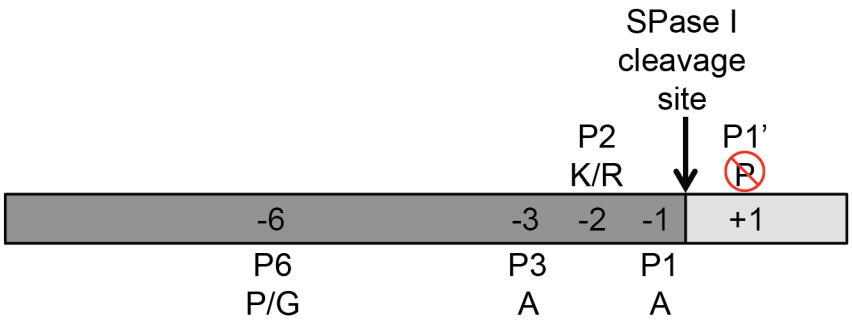


Figure 1.11. A typical synthetic signal peptide substrate for SPase I.

Alanine is the preferred residue at P1 and P3. Large basic residues, lysine and arginine, are preferred at P2 while a glycine or proline is usually found at P6. P1' can tolerate any residue except a proline.

Serine protease inhibitors can be classified into two general categories: protein-based inhibitors and small molecule inhibitors. Many protein-based inhibitors present themselves as substrates to their target proteases and inhibit by forming stable enzyme / inhibitor complexes with the proteases. For example, ecotin is a small, dimeric β-barrel, protein-based serine protease inhibitor secreted by *E. coli* (Maurizi 1992). The function of ecotin is thought to be to protect *E. coli* from digestive enzymes in the mammalian gut and from neutrophil elastase (Chung et al. 1983, Eggers et al. 2004). There are no known naturally occurring protein-based inhibitors for SPase I. However, a maltose binding protein (MBP) with a proline introduced at the P1' position of the signal peptide can become a competitive inhibitor for *E. coli* SPase I (Barkocy-Gallagher & Bassford 1992). Small molecule inhibitors have two sub-categories: chemical compound inhibitors and peptide-based inhibitors. Chemical compound inhibitors can inhibit through covalent modification of the protease's catalytic residues, disrupting the catalytic mechanism, or by acting as competitive inhibitors. For example, chloromethyl ketone inhibitors work by alkylating the histidine in the Ser/His/Asp catalytic triad, inhibiting enzymes such as chymotrypsin (Schoellmann & Shaw 1963). Benzamidine is an example of a competitive chemical compound, inhibiting trypsin (Krieger, Kay & Stroud 1974) . *E. coli* SPase I is not inhibited by many standard chemical compound inhibitors such as chloromethyl ketone, benzamidine, diisopropyl fluorophosphate (DFP) and phenylmethylsulfonyl fluoride (PMSF) but is inhibited by β-lactam containing inhibitors (Kim, Muramatsu & Takahashi 1995, Kuo et al. 1994). Peptide-based inhibitors function like chemical compound inhibitors, with an added short polypeptide for increased specificity and binding affinity for their target protease. Arylomycin A_2 and a synthetic MBP signal peptide with a proline at the P1' position are examples of non-covalent binding peptide-based SPase I inhibitors (Barkocy-Gallagher & Bassford 1992, Smith, Roberts & Romesberg 2010). Examples of covalent binding peptide-based inhibitors are signal peptide aldehydes (a signal peptide with an aldehyde group instead of a carboxylic acid group at the C-terminus of the P1 residue), which lock SPase I in its acyl-enzyme intermediate stage (Buzder-Lantos et al. 2009).

1.4. Signal peptide peptidase A

The idea that remnant signal peptides are removed by enzymes was first proposed by Hussain's group after conducting SPase II catalyzed pre-protein maturation experiments (Hussain et al. 1982, Hussain, Ichihara & Mizushima 1982) . Briefly, the *E. coli* membrane was isolated in the presence of globomycin, a SPase II inhibitor, which causes the accumulation of pre-lipoproteins in the membrane. Globomycin was then washed away to initiate pre-lipoprotein maturation and the process was monitored on SDS-PAGE gels. Surprisingly, only mature lipoproteins were seen but not their associated signal peptides. Hussain named the unknown remnant signal peptide degrading enzyme, signal peptide peptidase (SppA). Soon after Hussain's discovery of *E. coli* SppA, it was cloned, sequenced, expressed and purified (Ichihara et al. 1986). Based on results from cross-linking experiments and inhibition studies with known serine protease inhibitors, early reports suggested *E. coli* SppA was a tetrameric serine protease (Ichihara, Beppu & Mizushima 1984, Ichihara et al. 1986). These findings were later confirmed by DNA mutagenesis studies and the determining of *E. coli* SppA's structure using X-ray crystallography techniques (Kim, Oliver & Paetzel 2008, Wang et al. 2008). *E. coli* SppA is not an essential enzyme and *E. coli* cells lacking the *sppA* gene do not show accumulation of remnant signal peptides, suggesting that *E. coli* carries at least two enzymes capable of remnant signal peptide degradation (Suzuki et al. 1987).

1.4.1. S49A peptidase family

The S49A peptidase family includes SppA from bacteria, archaea, plants and SohB , the bacterial suppressor of *hrtA* B. All members of the S49A family utilize the Ser/Lys dyad catalytic mechanism (Rawlings et al. 2013). Bacterial SppA can be classified into two general types; the 67K type or the 36K type, based on their molecular mass (Figure 1.12). *E. coli* SppA is the 67K type while *B. subtilis* SppA is the 36K type, with only 25% sequence identity between the two.

Figure 1.12. Schematic of 67K type and 36K type SppA.

The polypeptide N-terminus and C-terminus are labelled N and C, respectively. The red cylinder represents the transmembrane segment (TM) while the blue rectangular box represents the SppA protease domain. Both 67K and 36K type SppA have one TM, with two protease domains in the 67K type and one protease domain in the 36K type.

Although different in size, *E. coli* and *B. subtilis* SppA both use the same strategy to assemble into active enzymes. In *E. coli* SppA, a complete active site is formed between two protease domains within one SppA protomer and four protomers form a final tetrameric complex with four active sites (Kim, Oliver & Paetzel 2008). In *B. subtilis* SppA, a complete active site is formed between two neighbouring SppA protomers and eight protomers forming a final octameric complex with eight active sites (Figure 1.13) (Nam, Kim & Paetzel 2012).

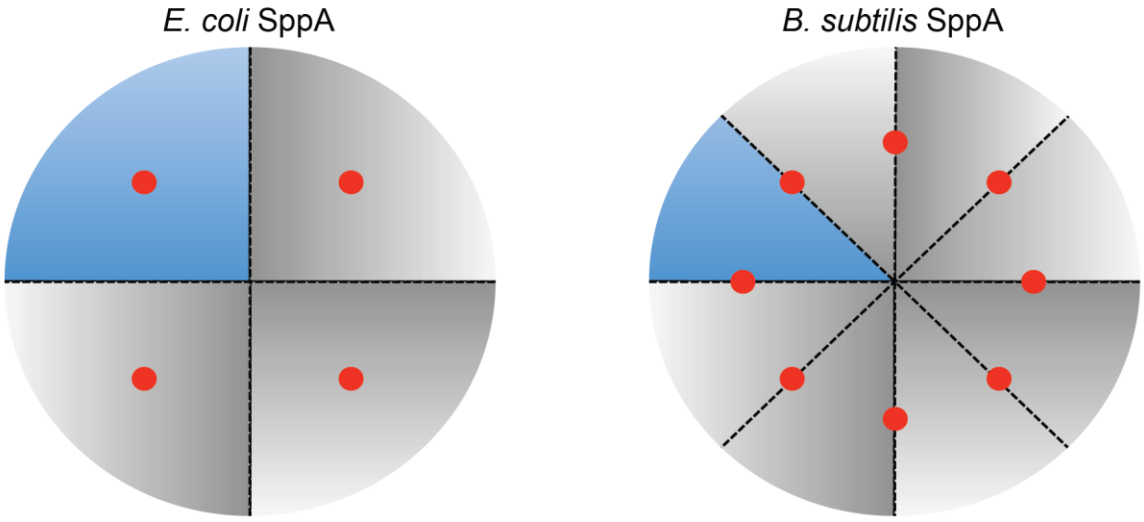


Figure 1.13. Schematic of the E. coli SppA tetramer and the B. subtilis SppA octamer.

Each SppA complex has one protomer coloured in blue with the rest of the protomers coloured in grey and each protomer is separated by dash lines. Red spheres represents complete Ser / Lys dyad active sites formed between protease domains within a protomer (*E. coli*) or between neighbouring protomers (*B. subtilis*).

1.4.2. Signal peptide peptidase A X-ray crystal structures

To date, only structures for *E. coli* SppA and *B. subtilis* SppA have been solved (Kim, Oliver & Paetzel 2008, Nam, Kim & Paetzel 2012). *E. coli* SppA has an α/β fold with a globular domain and a extension arm (Figure 1.14 A) while domains 1 and 2 are structural duplicates of each other, despite sharing only 20% sequence identity (Figure 1.14 B). Upon oligomerization, the *E. coli* SppA extension arms form the roof of adome while the globular domains form the body of this dome (Figure 1.14 C).

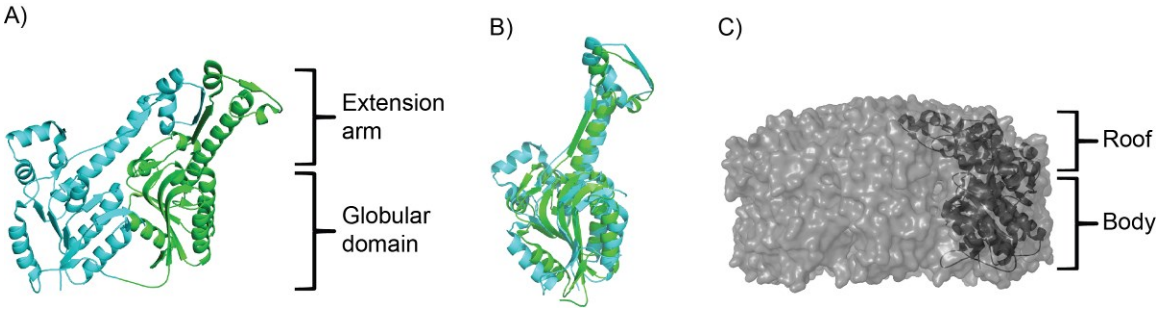


Figure 1.14. E. coli SppA X-ray crystal structure.

Only the *E. coli* SppA catalytic domains from residues 56 to 549 are shown in this figure. A) The *E. coli* SppA monomer in cartoon. Domain 1 is cyan and domain 2 is green. B) Superposition of *E. coli* SppA domains 1 and 2 colored the same way as in panel A. C) A side view of the *E. coli* SppA protomer with a semitransparent surface overlay representing the SppA tetrameric dome complex.

B. subtilis SppA is very similar to *E. coli* SppA, in both its protomer structure and the dimension and shape of its final octameric complex. The major difference between the two SppA enzymes lies in their substrate binding pockets. *B. subtilis* SppA has a deep S1 binding pocket and a narrow S3 binding pocket whereas *E. coli* SppA has a shallow S1 binding pocket and a wide S3 binding pocket which results in notably different substrate preferences for the two enzymes (Nam, Kim & Paetzel 2012).

1.4.3. Signal peptide peptidase A substrates and inhibitors

E. coli SppA enzymatic activity was first demonstrated using αS1 casein and two small synthetic substrates, Boc-Ala-ONp and Cbz-Leu-ONp (Pacaud 1982, Regnier 1981). Later studies using pre-lipoprotein signal peptides have suggested that *E. coli* SppA preferentially cleaves in the hydrophobic (H) region of the signal peptide with a preference for hydrophobic residues in the P1 and P1' positions (Novak & Dev 1988).

Currently there has been no in-depth kinetic characterization of *E. coli* SppA, although recently, kinetic analysis was carried out on *B. subtilis* SppA using a synthetic fluorogenic peptide (Nam & Paetzel 2013). Due to the lack of intensive kinetic characterization of bacterial SppA, only a few inhibitors have been characterized so far. *E. coli* SppA is inhibited by peptide-based inhibitors including antipain, leupeptin and chymostatin (Ichihara et al. 1986). A short peptide from the C-terminus of *B. subtilis* SppA has been proposed to have a self-regulatory role in a recent study (Nam & Paetzel 2013).

1.4.4. Signal peptide peptidase A2

E. coli SohB (suppressor of *hrtA* B), a member of the S49A family, was first identified through a rescue experiment (Baird et al. 1991). Briefly, the *E. coli hrtA* null strain, which cannot survive at heat shock temperatures (> 42°C), was transformed with high-copy number plasmids bearing the *sohB* gene. It was observed that overexpression of the SohB protein provided partial resistance to heat shock. *E. coli* SohB was later named SppA2 by the MEROPS database due to its sequence similarities to other bacterial SppA enzymes (Rawlings et al. 2013). To date, there is no experimental evidence suggesting that *E. coli* SppA2 is a Ser/Lys dyad protease which can degrade remnant signal peptides.

1.5. Proposed serine / lysine dyad catalytic mechanism

Proteases can be categorized into one of four mechanistic classes of serine, cysteine, aspartic or metallo-proteases based on the residues used to carry out the catalytic activity (Rawlings et al. 2013). The Ser/His/Asp catalytic triad is the most characterized and studied catalytic mechanism in the serine protease class. The three residues in the catalytic triad form a charge relay system with the serine acting as a nucleophile to carry out a nucleophilic attack on the carbonyl carbon of the scissile bond of an incoming substrate. The histidine general base activates the nucleophilic serine by extracting a proton from the side chain hydroxyl group of serine while aspartate aligns the histidine general base and neutralizes the charge on the histidine through out catalysis (Ekici, Paetzel & Dalbey 2008). Ser/Lys dyad enzymes go through a two-step

catalysis similar to that of the classic Ser/His/Asp triad with formation of a stable acyl-enzyme intermediate (Figure 1.15).

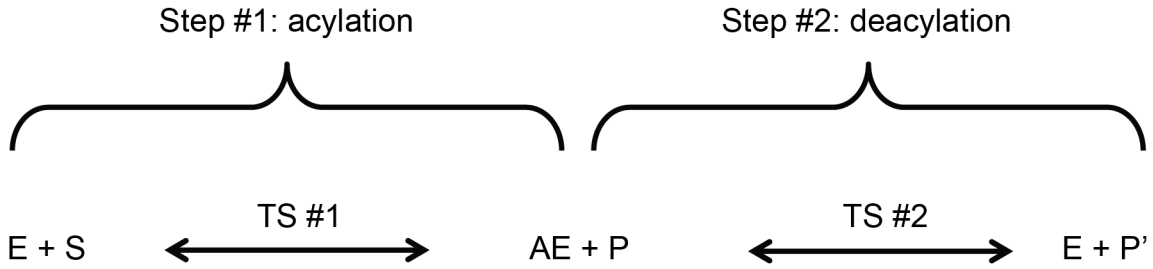


Figure 1.15. A typical two stage Ser/Lys dyad mechanism reaction.

The acylation step consists of the enzyme (E) binding to its substrate (S) to initiate catalysis. The complex goes through the first high energy tetrahedral transition state (TS #1) to form the acyl-enzyme (AE) and the first product (P) is released. The deacylation step consists of the acyl-enzyme going through the second high energy tetrahedral transition state (TS #2) to release the second product (P') and regenerate the enzyme.

The Ser/Lys catalytic dyad is similar to the Ser/His/Asp triad in many ways with the serine acting as a nucleophile, the lysine acting as a general base and a third hydroxyl containing residue, serine or threonine, acting as a coordinating residue (Figure 1.16) (Paetzel 2013). Study of the *E. coli* SPase I structure has revealed several key features of the Ser/Lys catalytic dyad. The nucleophilic serine attacks on the *si*-face of its substrate rather than the *re*-face commonly attacked by Ser/His/Asp triad enzymes (Paetzel & Strynadka 1999). In order to activate the nucleophilic serine, the lysine general base with a pKa of 10.5 has to be in a deprotonated form at physiological pH. It has been proposed that the hydrophobic environment surrounding the lysine general base helps to suppress its pKa (Paetzel et al. 1997).

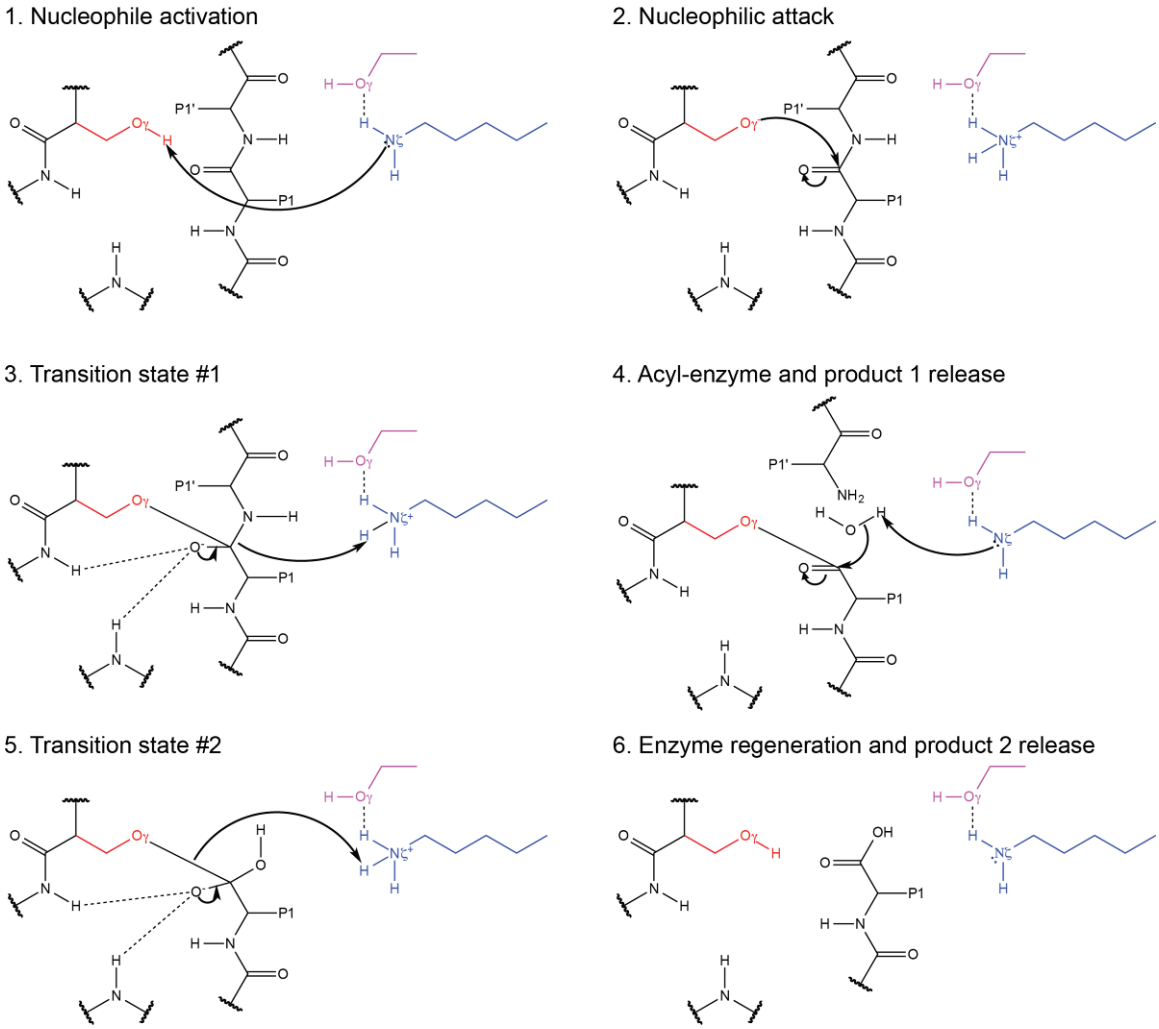


Figure 1.16. The proposed general Ser/Lys dyad mechanism.

The nucleophile serine, general base lysine and coordinating serine side chains are coloured in red, blue and purple, respectively. Solid lines represent covalent bonds, dotted lines represent hydrogen bonds and black arrows denote electron movements.

1.6. Research objectives

The first goal of the *S. aureus* SpsB project was to solve, for the first time, the structure of a type I signal epptidase from a Gram-positive bacterium, specifically, SpsB from *S. aureus*. The aims were as follows:

- To design and prepare a truncated SpsB more amenable to crystallization and evaluate the construct based on its enzymatic activity.

- To design a substrate for use in kinetic characterization of *S. aureus* SpsB and other bacterial SPase I and for use as a protein crystallization additive.
- To crystallize truncated *S. aureus* SpsB with or without a bound substrate/inhibitor.
- To investigate why there is a significant difference in catalytic efficiency towards the same substrate observed between Gram-positive and Gram-negative SPase I.

The second goal was to characterize signal peptide peptidase A2, SppA2 from *E. coli*. Currently, the only known function of *E. coli* SppA2 is to provide partial resistance to heat shock in *E. coli* cells lacking the HtrA protein which is a chaperone/protease. The following questions were addressed in this thesis.

- Is *E. coli* SppA2 a protease?
- What types of substrates can *E. coli* SppA2 degrade?
- What are the substrate preferences for *E. coli* SppA2?
- Does *E. coli* SppA2 assemble into an oligomer?
- What is the structure of *E. coli* SppA2?

2. Biochemical characterization of *Staphylococcus aureus* signal peptidase I

2.1. Overview

The primary goal of this part of the thesis project was to design a truncated *Staphylococcus aureus* SPase I (SpsB) to be more amenable to crystallization than the full length enzyme. Enzymatic activity was chosen as the main criterion to evaluate how well the truncated SpsB represents the full length SpsB. The peptide substrate sequence used (NGEVAKA, P7 to P1) came from the signal peptide C-region of a putative glycoside hydrolase from *Bacillus licheniformis*. This sequence was chosen because this signal peptide gave the highest subtilisin secretion rate in *B. subtilis* when this signal peptide was fused at the N-terminus of subtilisin among a total of 393 signal peptides examined from *B. subtilis* and *B. licheniformis* (Degering et al. 2010). Our peptide design strategy is quite different from a previous substrate discovery experiment using phage display library technique (Sharkov & Cai 2002). Sharkov & Cai have discover several highly reactive phage clones with *S. aureus* SpsB but synthetic peptide substrates containing peptide sequences from these clones were poor substrates for S. aureus SpsB. A second peptide substrate using the same peptide sequence but with the addition of a dodecanoyl hydrocarbon chain added to the N-terminus of the peptide substrate to mimic the hydrophobic region was also synthesized. It was hypothesized that the hydrocarbon chain may help in substrate presentation when used in a micelle containing system. This design stragegy was used by Stein, Barbosa & Bruckner who have make a signal peptide substrate containing five lysines and leucines leucines N-terminal to a previously characterized peptide substrate has drastically increased the catalytic efficiency of E. coli SPase I toward the same substrate minus the N-terminal additions (Stein, Barbosa & Bruckner, 2000). The secondary goal was to kinetically characterize and to crystallize other Gram-positive SPase I.

Methicillin-resistant *S. aureus* (MRSA) SpsB was chosen as the main target SPase I for this project since MRSA infection is one of the leading opportunistic infections in Canadian hospitals (Simor et al. 2010). The full length *spsB* gene was cloned from *S. aureus* (MRSA252) genomic DNA and inserted into a pET28a+ vector incorporating a hexahistidine tag at the N-terminus of SpsB for future use in conjunction with Ni^{2+}-NTA affinity purification. A truncated construct, *S. aureus* SpsB Δ2-20 missing part of the predicted transmembrane segment, was constructed to serve as a potential crystallization candidate (Figure 2.1). Removing a transmembrane helix may facilitate protein expression, purification and crystallization in the absence of detergents. Unfortunately, *S. aureus* SpsB Δ2-20 expressed mostly as inclusion bodies, with a small amount of natively folded population in the supernatant of the cell lysate. However, *S. aureus* SpsB Δ2-20 retained approximately 80% of its catalytic efficiency compared to the full length *S. aureus* SpsB, making it a good crystallization candidate.

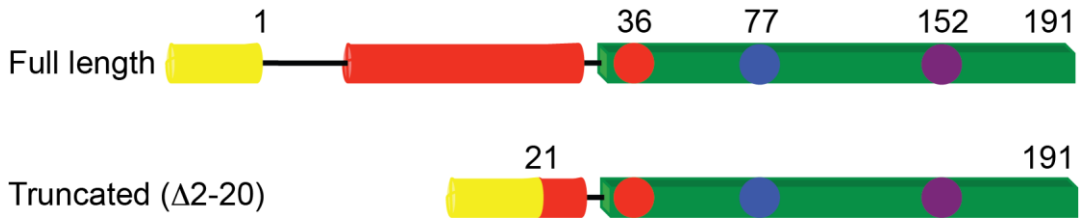
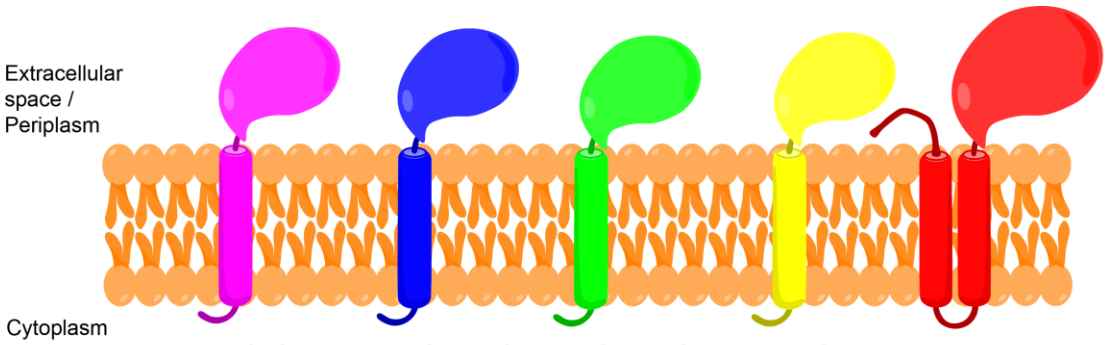


Figure 2.1. Schematics of full length and truncated S. aureus SpsB constructs.

Full length *S. aureus* SpsB is 191 residues in length, made up of a single predicted N-terminal transmembrane segment (red cylinder) and a C-terminal protease domain (green box). The three catalytic residues-- serine nucleophile, lysine general base and the coordinating serine-- are shown as circles in red, blue and purple, respectively. The hexahistidine tag added by the pET28a+ vector is shown as a yellow cylinder at the N-terminus of each construct.

In order to prepare a suitable quantity for crystallization (>10mg/mL) of *S. aureus* SpsB Δ2-20, the inclusion bodies of *S. aureus* SpsB Δ2-20 were denatured, refolded and the quality of the refolded enzyme was assessed by enzymatic activity assays. Refolded *S. aureus* SpsB Δ2-20 showed no discernable difference in its catalytic efficiency compared to the natively folded *S. aureus* SpsB Δ2-20.

Other SPase I enzymes including; SpsB from the Gram-positive opportunistic pathogen *S. epidermidis*, SipS and SipT from the Gram-positive model organism *B. subtilis* and LepB from the Gram-negative model organism *E. coli*, were also cloned, expressed, purified and kinetically characterized (Figure 2.2).



	B. subtilis SipS	*B. subtilis* SipT	*S. aureus* SpsB	*S. epidermidis* SpsB	*E. coli* LepB
MW (kDa)	21	22	22	22	36
pI	9.4	9.5	9.0	9.5	6.9
Accession number	P28628	P71013	Q6GIC3	Q5HJQ6	P00803

Sequence Identity (%)

	B. subtilis SipS	*B. subtilis* SipT	*S. aureus* SpsB	*S. epidermidis* SpsB	*E. coli* LepB
B. subtilis SipS	-----	66	40	39	13
B. subtilis SipT	66	-----	37	37	14
S. aureus SpsB	40	37	-----	80	14
S. epidermidis SpsB	39	37	80	-----	14
E. coli LepB	13	14	14	14	-----

Figure 2.2. B. subtilis, S. aureus, S. epidermidis and E. coli SPase I properties.

Schematic of *B. subtilis* primary SPase I (SipS and SipT), *S. aureus* SPase I (SpsB), *S. epidermidis* SPase I (SpsB) and *E. coli* SPase I (LepB). The molecular weight (MW) in kDa, isoelectric point (pI), Uniprot accession number and a sequence identity matrix for each are provided.

2.2. Materials and methods

2.2.1. Cloning and mutagenesis

B. subtilis SipS and SipT, *S. aureus* SpsB and SpsB Δ2-20 and *S. epidermidis* SpsB were each amplified using PCR with gene specific primer pairs containing NdeI and XhoI restriction sites in the forward and reverse primers, respectively (Table 2.1). *E. coli* LepB was previously cloned (Klenotic et al. 2000). Each SPase I gene was digested, then ligated into the expression vector pET28a+ (Novagen) using the NdeI and XhoI sites. Using *S. aureus* SpsB as the template, four active-site mutants were constructed for each SPase I where: the serine nucleophile was mutated to alanine (S36A), the

lysine general base to alanine (K77A) and the coordinating serine to either alanine, threonine or cysteine (S152A, S152T, S152C). Each mutation was introduced with mutation specific primer pairs (Table 2.2) using the QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent Technologies) following the manufacturer's protocol. All DNA sequences were confirmed by DNA sequencing (Genewiz) and the translated protein sequences matched the protein sequences deposited in the UniProt database (UniProt Consortium 2014) for *B. subtilis* SipS (P28628), *B. subtilis* SipT (P71013), *S. aureus* SpsB (Q6GIC3) and *S. epidermidis* SpsB (Q5HJQ6). The expressed proteins from these constructs each carry an N-terminal hexahistidine tag, followed by a thrombin cleavage site (MGSSHHHHHHSSGLVPRGSHM), in addition to their native protein sequences.

Table 2.1. Bacterial source and primer pairs used for each SPase I.

Organism (strain)	SPase I gene	Primer pairs (forward / reverse)
Staphylococcus aureus (MRSA252)	*spsB*	5'-GCCATATGAAAAAAGAATTATTGGAATGGATTATTTCAATTG-3' 5'-GCCTCGAGTTAATTTTTAGTATTTTCAGGATTGAAATTATATTTAAAT-3'
Staphylococcus epidermidis (RP62A)	*spsB*	5'-GCCATATGATGAAAAAAGAAATTTTAGAGTGGATTGTTG-3' 5'-GCCTCGAGTTAATTTTTAGTGTTATTTGGATTAAAGTTAGATTTAAAT-3'
Bacillus subtilis (168)	*sipS*	5'-GCCATATGAAATCAGAAAATGTTTCGAAGAAAAAGTCA-3' 5'-GCCTCGAGCTAATTTGTTTTGCGCATTTCGTTAAA-3'
Bacillus subtilis (168)	*sipT*	5'-GCCATATGACCGAGGAAAAAAATACGAATACTGAGAA-3' 5'-GCCTCGAGTTATTTTGTTTGACGCATTTCGTTAAAC-3'

Restriction sites NdeI and XhoI are underlined.

Table 2.2. Site-directed mutagenesis primers pairs for S. aureus SpsB active site mutants.

Mutation	Primer Pairs (forward / reverse)
S36A	5'-ATTAAAGGTGGAGCAATGGATCCAACT-3' 5'-AGTTGGATCCATTGCTTCACCTTTAAT-3'
K77A	5'-GATGACTATGTTGCACGTGTCATCGGT-3' 5'-ACCGATGACACGTGCAACATAGTCATC-3'
S152A	5'-AATCGTGAAGTAGCAAAAGATAGCCGT-3' 5'-ACGGCTATCTTTTGCTACTTCACGATT-3'

S152T	5'-AATCGTGAAGTAACTAAAGATAGCCGT-3'
	5'-ACGGCTATCTTTAGTTACTTCACGATT-3'
S152C	5'-AATCGTGAAGTATGTAAAGATAGCCGT-3'
	5'-ACGGCTATCTTTACATACTTCACGATT-3'

Codons changed for each mutant are underlined.

2.2.2. Full length SPase I over-expression and purification

Plasmids bearing each full length SPase I gene were individually transformed into *E. coli* BL21(DE3). For each construct, one liter of LB media supplemented with kanamycin (0.05mg/mL) was inoculated with 50mL of overnight culture. The cells were grown at 37°C with shaking (250rpm) until the $O.D._{600}$ reached 0.6, whereupon 1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to the culture to induce protein expression. The cells grew for four hours following induction. At the end of the four hours, the cells were pelleted by centrifugation (6000xg, 5min, 4°C) and the cell pellet was then stored at -80°C until needed. Later, following thawing, 5g of the cell pellet was resuspended in 25mL of buffer A (20mM Tris pH 8.0, 100mM NaCl). The resuspended cells were then sonicated using a Sonic Dismembrator Model 500 (Fisher Scientific), with a 15sec pulse and 30sec rest cycle three times at 30% amplitude, on ice and lysed by passing through an EmulsiFlex-C3 cell disruptor (Avestin) for 5min. The resulting cell lysate was clarified by centrifugation (35,000xg, 30min, 4°C) and the supernatant was discarded. The cell lysate pellet was resuspended in 25mL of buffer A with 0.1% Triton X-100 and dounced 50 times followed by centrifugation (35,000xg, 30min, 4°C). The clarified supernatant was then subjected to Ni^{2+}-NTA affinity chromatography purification. The clarified supernatant, containing the His-tagged SPase I protein, was poured onto a column containing 3mL of Ni^{2+}-charged resin (Qiagen) pre-equilibrated with kinetic buffer (buffer A with 0.01% n-dodecyl β-D-maltoside (DDM)). The column was washed twice with 30mL of kinetic buffer containing first 10mM, and then 20mM imidazole. Next, the His-tagged SPase I protein was eluted using a step gradient, created by 6mL of kinetic buffer containing imidazole in increasing concentrations of 100 to 500mM imidazole. Fractions containing the purified SPase I protein were concentrated to 1mL using an Amicon ultracentrifugal filter (Millipore) with a 3-kDa cutoff

and dialysed against 1L of kinetic buffer at 4°C overnight. Protein concentration was determined by BCA assay (Thermo Fisher) following the manufacturer's protocol.

2.2.3. S. aureus SpsB Δ2-20 overexpression and purification

Truncated *S. aureus* SpsB (Δ2-20) was prepared following the same method as the full length SPase I preparation described above, except the kinetic buffer was used for both the cell pellet's resuspension and lysis steps. The clarified supernatant from this cell lysate was used directly in the subsequent Ni^{2+}-NTA purification steps. Elution fractions containing *S. aureus* SpsB Δ2-20 protein were further purified with cation exchange chromatography using SP-sepharose beads (GE Healthcare Life Sciences). *S. aureus* SpsB Δ2-20 protein was passed through a column containing 3mL of SP-sepharose beads pre-equilibrated with kinetic buffer. Next, the column was washed twice with 10mL of kinetic buffer containing first 100mM NaCl, and then 150mM NaCl. Following the washes, *S. aureus* SpsB Δ2-20 protein was eluted in a step gradient, using 6mL of kinetic buffer each time containing an increasing concentration of NaCl (200, 400, 600 and 800mM).

2.2.4. S. aureus SpsB Δ2-20 inclusion bodies preparation and refolding

Cell pellets containing *S. aureus* SpsB Δ2-20 protein in inclusion bodies were resuspended in buffer A and lysed as described above for the full length SPase I preparation. The inclusion bodies were collected by centrifugation (5,000xg, 20min, 4°C) following which the supernatant was discarded while the pelleted inclusion bodies were collected. The inclusion bodies were washed by resuspending in buffer A with 0.1% Triton X-100 and pelleting by centrifugation (5,000xg, 10min, 4°C). The supernatant was discarded. The wash step was repeated five times. After washing, 200mg (wet weight) of inclusion bodies was dissolved in 100mL of 4M guanidinium-HCl with 0.1% Triton X-100 and stirred at room temperature for one hour. The dissolved inclusion body solution was centrifuged (35,000xg, 30min, 4°C) to remove insoluble materials. The resuspended *S. aureus* SpsB Δ2-20 protein was refolded by dialyzing against 8L of refolding buffer (20mM Tris pH8.0, 1M NaCl, 0.1% Triton-X 100) at room temperature overnight. After dialysis, the solution was clarified by centrifugation (35,000xg, 30min, 4°C) to remove

any misfolded proteins. Refolded *S. aureus* SpsB Δ2-20 protein for kinetic analysis was further dialyzed against 1L of kinetic buffer at 4°C overnight while *S. aureus* SpsB Δ2-20 protein for crystallization experiments was dialyzed against crystallization buffer (10mM HEPES pH7.5, 150mM NaCl, 0.05% Tween-20).

2.2.5. Kinetic characterization of S. aureus SpsB and SpsB Δ2-20 using a FRET peptide substrate

A FRET peptide substrate, Dabcyl-NGEVAKAAE(EDANS)T-NH_2, was custom synthesized (CanPeptide) and dissolved in dimethyl sulfoxide (DMSO) to make a 10mM stock solution. Dabcyl (4-((4-(dimethylamino)phenyl)azo)benzoic) is the quencher and EDANS (5-((2Aminoethyl)amino)naphthalene-1-sulfonic acid) is the donor of the FRET pair. Measuring the fluorescence emitted by EDANS when the peptide substrate is cleaved allows monitoring of the kinetic reaction's progress over time. 100µL reaction mixtures containing 100nM of either full length *S. aureus* SpsB or SpsB Δ2-20 in kinetic buffer were pipetted into 96 well microplates (Greiner Bio-One, Cat-No. 675096). The reactions were then initiated by adding 1µL of various concentrations of the peptide substrate in DMSO (final DMSO concentration is 1%), to the enzyme solutions in the microplates, to give a final concentration from 5µM to 100µM. The reaction progress was monitored on a SpectraMax M5 (Molecular Devices) with an excitation wavelength (Ex) of 340nm and an emission wavelength (Em) of 530nm for 5 minutes at 23°C. Kinetic constants, k_{cat} and K_M, were determined by fitting data points, collected in triplicates to the Michaelis−Menten equation using the Prism5 software (GraphPad). Fluorescent intensity was converted into molar units using a standard curve generated with EDANS and Dabcyl at various concentrations (Appendix B).

2.2.6. Kinetic characterization of S. aureus SpsB using an MCA lipidated peptide substrate

S. aureus SpsB, *S. epidermidis* SpsB, *B. subtilis* SipS and SipT, and *E. coli* LepB were screened using an MCA lipidated peptide substrate, dodecanoyl-NGEVAKA-MCA, which was custom synthesized (CanPeptide) and dissolved in DMSO to make a 10mM stock solution. MCA (7-Amino-4-methylcoumarin) is a fluorophore whose signal can be detected once it is liberated from the peptide substrate. 2µM of the MCA lipidated

peptide substrate was added to each well of 100µL reaction mixture containing 200nM of one of the SPase I enzymes. The reaction progress was monitored as described above except Em380 / Ex460 were used. Kinetic characterization of *S. aureus* SpsB was carried out using 50nM of the enzyme while the substrate varied in concentration from 1µM to 40µM. Kinetic constants were determined as described above. Fluorescent intensity was converted into molar units using a standard curve generated with MCA at various concentrations (Appendix B).

2.2.7. Kinetic characterization of full length SPase I using a FRET lipidated peptide substrate

A FRET lipidated peptide substrate, dodecanoyl-K(Dabcyl)NGEVAKAAE(EDAN-S)T-NH_2, was custom synthesized (CanPeptide) for this experiment and dissolved in DMSO to make a 10mM stock solution. Reaction mixtures were set up and reaction progress was monitored as described for the kinetic characterization using the FRET peptide substrate (See Section 2.2.5). For *S. epidermidis* SpsB, *B. subtilis* SipS and *B. subtilis* SipT, 10nM of the enzyme was used with the peptide substrate ranging in concentration from 2.5µM to 60µM. For *E. coli* LepB, 500nM of the enzyme was used with substrate concentrations covering 10µM to 300µM. Kinetic constants were determined as described above.

2.2.8. Activity characterization of S. aureus SpsB, and the mutants S36A, K77A, S152A, S152C and S152T, using a FRET lipidated peptide substrate

To compare the activity of *S. aureus* SpsB to the *S. aureus* SpsB S36A, K77A, S152A, S152T and S152C mutants, 2µM of the FRET lipidated peptide substrate was incubated with either 200nM of wild type SpsB or 1µM of mutant SpsB. Reaction mixtures were set up and reaction progress was monitored as described above for the FRET peptide substrate (See Section 2.2.5).

2.2.9. *Effect of pH, detergents, ions, and small chemical compounds on S. aureus SpsB Δ2-20*

100µL reaction mixtures containing 100nM *S. aureus* SpsB Δ2-20 and 20µM of Dabcyl-NGEVAKAAE(EDANDS)AET-NH_2 FRET peptide substrate, were pipetted into 96 well microplates. *S. aureus* SpsB Δ2-20 was incubated with various buffers (see buffers listed below) for 30 minutes at room temperature before the reaction was initiated with the addition of 1µL of the FRET peptide substrate. The reaction progress was monitored as described above for the FRET peptide substrate (See Section 2.2.5). To determine the effect of pH, *S. aureus* SpsB Δ2-20 was incubated in 100mM NaCl and 0.01% DDM, with one of the following buffers in each reaction solution: 20mM MES at pH5.0 or pH6.0, HEPES at pH7.0, Tris at pH8.0, CHES at pH9.0 or CAPS at pH10.0 or pH11.0. To determine the effect of detergents, *S. aureus* SpsB Δ2-20 was incubated in 20mM Tris pH8.0 and 100mM NaCl buffer containing either DDM (0, 0.0025, 0.005 or 0.01% w/v) or Triton X-100 (0, 0.025, 0.05 or 0.1% v/v). To measure the effect of divalent cations, *S. aureus* SpsB Δ2-20 was incubated with kinetic buffer containing 1mM of either $CaCl_2$, $MgCl_2$, $ZnCl_2$, $CoCl_2$ or $NiCl_2$. To identify the effect of ionic strength, *S. aureus* SpsB Δ2-20 was incubated in kinetic buffer with either 0, 100, 250, 500, 750, 1000 or 2000mM NaCl. Finally, to determine the effect of small chemical compounds, *S. aureus* SpsB Δ2-20 was incubated with kinetic buffer containing either DMSO (1, 2.5, 5, 7.5, 10, 12.5 or 15% v/v) or DTT (0, 1, 5mM) or EDTA (0, 1, 5mM) or glycerol (0, 10, 25, 50% v/v).

2.2.10. *S. aureus SpsB Δ2-20 S36A crystallization*

S. aureus SpsB Δ2-20 S36A (9mg/mL) was incubated with arylomycin A_2 inhibitor (1:1 molar ratio) on ice for two hours prior to crystallization. Crystallization was carried out using the sitting drop vapour diffusion method. 1µL of *S. aureus* SpsB Δ2-20 S36A enzyme/inhibitor mixture was mixed with 1µL of reservoir solution (16.5% w/v PEG3350, 0.18M tri-sodium citrate dihydrate) and sealed in a well with 1mL of reservoir at room temperature. Crystals were observed after two weeks. Other crystallization screening solutions used included Crystal Screens I, II and PEG Ion Screen (Hampton Research).

2.2.11. S. aureus SpsB homology model construction

The full length *S. aureus* SpsB protein sequence was submitted to the Swiss-Model server for homology model construction (Guex et al., 2009) with *E. coli* LepB (PDB: 1B12) as the template model (Paetzel et al., 1998). The constructed model was then energy minimized using the YASARA server (Kreiger et al., 2009). The final energy minimized *S. aureus* SpsB model was used for analysis without any further modifications.

2.3. Results

2.3.1. Full length SPase I purification

Full length SPase I from *S. aureus*, *S. epidermidis*, *B. subtilis* and *E. coli* exhibit low expression levels, likely due to the limited inner membrane space where the over-expressed SPase I resides, while being expressed in the *E. coli* BL21(DE3) expression host. A typical yield for each of the SPase I enzymes from 2L (~5g of cell pellet) of cell culture is approximately 500µg. Although this amount is not suitable for crystallization experiments, it is sufficient for kinetic characterization. *S. aureus* SpsB, with an N-terminal hexahistidine tag and a theoretical molecular mass of 24kDa, was expressed as an insoluble component in *E. coli* BL21(DE3) (Figure 2.3 lane 3). It was partially extracted by Triton X-100 (Figure 2.3 lane 4) and then purified to homogeneity using Ni^{2+}-NTA affinity chromatography (Figure 2.3 lanes 8 to 10).

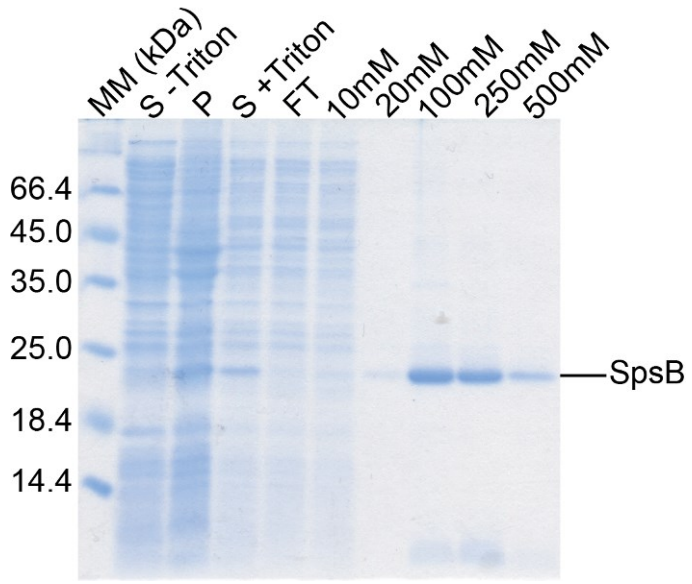


Figure 2.3. S. aureus SpsB Ni^{2+}-NTA affinity chromatography.

Protein samples from each step of the Ni^{2+}-NTA affinity chromatography process were run on a 15% SDS-PAGE gel stained with PageBlue protein staining solution. The molecular mass ladder (MM) in kDa is shown in the left lane. The supernatant without Triton X-100 (S –Triton) from the clarified cell lysate did not contain a significant amount of SpsB and was discarded. SpsB was extracted from the insoluble pellet (P) using buffer containing 0.1% Triton X-100 (S +Triton). No significant amount of SpsB was found in the flow through (FT) fraction. The column was washed twice with buffer containing first10mM and then, 20mM imidazole. SpsB was eluted stepwise with buffer containing 100mM, 250mM and 500mM imidazole.

2.3.2. S. aureus Δ2-20 SpsB refolding and purification

Due to *S. aureus* SpsB's low level expression and low efficiency of extraction from the cell membrane using 0.1% Triton X-100, a truncated *S. aureus* SpsB (Δ2-20), lacking part of the predicted transmembrane segment, was constructed in hopes of increasing protein expression level and solubility. *S. aureus* SpsB Δ2-20 was purified using Ni^{2+}-NTA affinity chromatography and cation exchange chromatography (Figure 2.4). Although *S. aureus* SpsB Δ2-20 has a higher expression level compared to full length SpsB, the majority of the *S. aureus* SpsB Δ2-20 protein expressed as inclusion bodies (Figure 2.4A lane P).

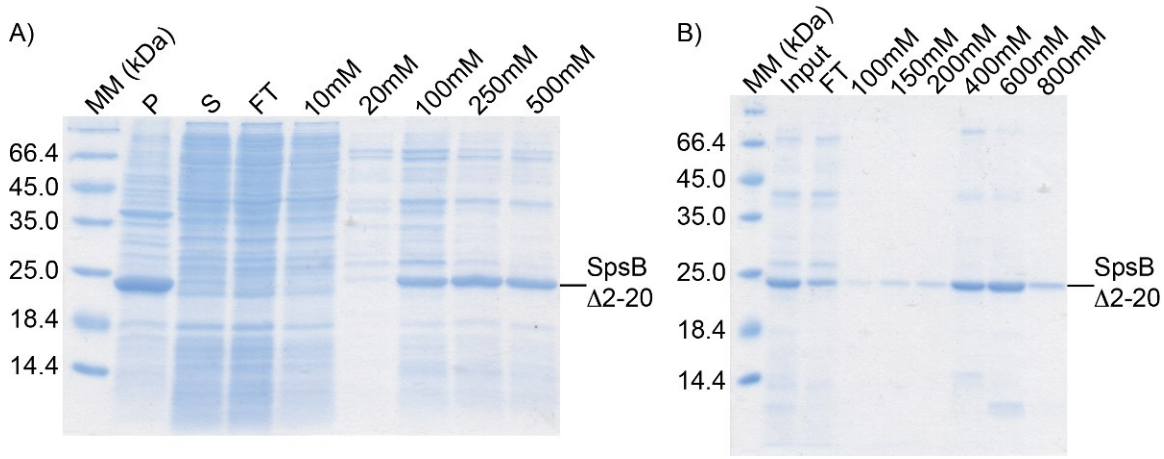


Figure 2.4. S. aureus SpsB Δ2-20 Ni^{2+}-NTA affinity chromatography and cation exchange chromatography.

A) Pellet (P), supernatant (S) and column flow through (FT) samples were analyzed. The column was washed twice with buffer containing first10 and then, 20mM imidazole. *S. aureus* SpsB Δ2-20 was eluted stepwise with buffer containing 100, 250 and finally, 500mM imidazole. B) Protein samples from cation exchange chromatography were run on a 15% SDS-PAGE gel. Elution fractions from the Ni^{2+}-NTA process were used as the input fraction. A significant quantity of *S. aureus* SpsB Δ2-20 was found in the column flow through (FT). The column was washed twice with buffer containing first 100mM and then, 150mM NaCl. *S. aureus* SpsB Δ2-20 was eluted stepwise with buffer containing 200, 400, 600 and finally, 800mM NaCl. Protein samples from each step in the Ni^{2+}-NTA affinity chromatography process were run on a 15% SDS-PAGE gel stained with PageBlue protein staining solution. The molecular mass ladder (MM) in kDa is shown in the left lane.

S. aureus SpsB Δ2-20 was successfully refolded (as described in Section 2.2.4) and purified by $Ni2^{+}$-NTA affinity chromatography (Figure 2.5).

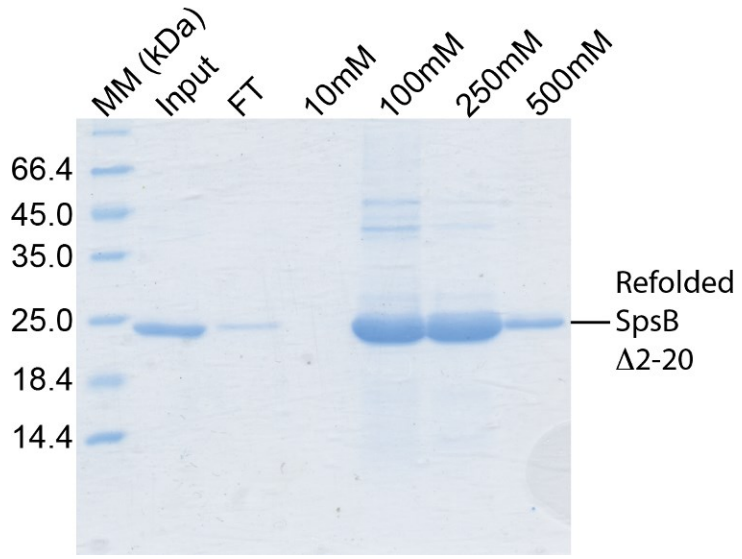


Figure 2.5. Refolded S. aureus SpsB Δ2-20 Ni^{2+}-NTA affinity chromatography.

Protein samples from each step of the $Ni2^{+}$-NTA affinity chromatography process were run on a 15% SDS-PAGE gel stained with PageBlue protein staining solution. The molecular mass ladder (MM) in kDa is shown in the left lane. Refolded *S. aureus* SpsB Δ2-20 was used as the input fraction and the input and the column flow through (FT) were analyzed. The column was washed once with buffering containing 10mM imidazole. *S. aureus* SpsB Δ2-20 was eluted stepwise with buffer containing 100mM, 250mM and finally, 500mM imidazole.

2.3.3. Natively folded and refolded S. aureus SpsB Δ2-20 proteins have similar catalytic efficiency

Enzymatic assay was chosen as the primary method to assess the quality of the refolded *S. aureus* SpsB Δ2-20 enzyme. A fluorogenic peptide, Dabcyl-NGEVAKAAE(EDANS)T-NH_2 containing the C-region (P7 to P1) of a signal peptide and part of the mature protein (P1' to P3'), was synthesized. Fluorescence produced by the product, $^{-}$OOC-AE(EDANS)T-NH_2, was measured to monitor the progress of the reaction. Refolded *S. aureus* SpsB Δ2-20 shows no discernable k_{cat} and K_M differences compared to natively folded *S. aureus* SpsB Δ2-20 (Table 2.3). The *S. aureus* SpsB truncation design and refolding protocol were successful in producing crystallization quantities of refolded enzyme that retains 80% of its catalytic efficiency compared to the natively folded full length SpsB.

Table 2.3. ***S. aureus SpsB, SpsB Δ2-20 and refolded SpsB Δ2-20 kinetic constants determination using the Dabcyl-NGEVAKAAE(EDANS)T-NH_2 substrate.***

	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($s^{-1} \bullet M^{-1}$)
SpsB	$7.3 \times 10^{-2} \pm 0.7 \times 10^{-2}$	25 ± 6	2.9×10^3
SpsB Δ2-20	$5.0 \times 10^{-2} \pm 0.4 \times 10^{-2}$	29 ± 6	1.7×10^3
SpsB Δ2-20 (refolded)	$5.3 \times 10^{-2} \pm 0.4 \times 10^{-2}$	23 ± 5	2.3×10^3

Error shown is two standard deviations determined from triplicate experiments.
See Appendix C (B, C, D) for initial velocity vs substrate concentration curves.

2.3.4. Full length SPase I kinetic characterization using a FRET lipidated signal peptide substrate

Kinetic characterization was further expanded to include Gram-positive SPase I, *S. epidermidis* SpsB and *B. subtilis* SipS and SipT, to screen for potential crystallization candidates. *E. coli* LepB was included as a reference since it is the most highly studied bacterial SPase I. k_{cat} and K_M were determined using an optimized peptide substrate with a dodecanoyl hydrocarbon chain added to the N-terminus of the FRET peptide substrate which mimics the H-region of a signal peptide, dodecanoyl-K(Dabcyl)NGEVAKAAE (EDANS)T-NH_2 (Figure 2.6).

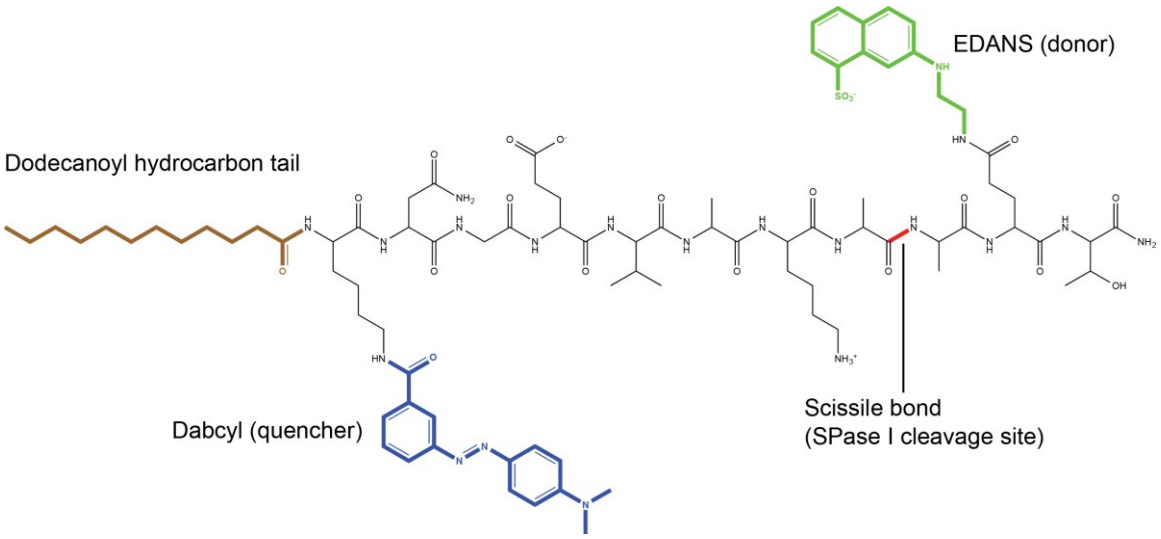


Figure 2.6. Schematic of the dodecanoyl-K(Dabcyl)NGEVAKAAE(EDANS)T-NH_2 FRET lipopeptides substrate.

The lipopeptide substrate contains 11 residues (P8 to P3') with a dodecanoyl hydrocarbon chain coloured in brown attached at the N-terminus. The quencher, Dabcyl, coloured in blue is attached to the side chain of lysine at the P8 position. The donor, EDANS, coloured in green is attached to the side chain of glutamic acid at the P2' position. The scissile bond, which will be cleaved by SPase I, is highlighted in red.

Among the five SPase I enzymes tested, *S. aureus* SpsB has the highest catalytic efficiency towards the lipopeptide substrate, followed by *S. epidermidis* SpsB, and then *B. subtilis* SipS and SipT which share similar catalytic efficiency levels. *E. coli* LepB has a significantly lower catalytic efficiency due to both a lower k_{cat} and a higher K_M compared to Gram-positive SPase I (Table 2.4).

Table 2.4. Gram-positive and Gram-negative SPase I kinetic constants determination using the dodecanoyl-K(Dabcyl)NGEVAKAAE (EDANS)T-NH_2 substrate.

SPase I	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($s^{-1} \bullet M^{-1}$)
S. aureus SpsB	$1.3 \times 10^{-1} \pm 0.1 \times 10^{-1}$	0.9 ± 0.2	1.5×10^{5}
S. epidermidis SpsB	$1.2 \times 10^{-1} \pm 0.1 \times 10^{-1}$	2.9 ± 0.5	4.2×10^{4}
B. subtilis SipS	$6.2 \times 10^{-2} \pm 0.3 \times 10^{-2}$	9.3 ± 1.4	6.7×10^{3}
B. subtilis SipT	$7.2 \times 10^{-2} \pm 0.4 \times 10^{-2}$	8.5 ± 1.3	8.7×10^{3}
E. coli LepB	$1.5 \times 10^{-3} \pm 0.1 \times 10^{-3}$	32.0 ± 4.5	4.7×10

Error shown is two standard deviations determined from triplicate experiments.
See Appendix C (A, G, H, I, J) for initial velocity vs substrate concentration curves.

2.3.5. *S. aureus SpsB kinetic characterization using an MCA lipidated signal peptide substrate*

A peptide substrate with a methylcoumaryl amide (MCA) fluorogenic reporter immediately following the SPase I cleavage site, dodecanoyl-NGEVAKA-MCA, was made to probe whether or not the residues in the P' region (residues after the SPase I cleavage site) of the peptide substrate affects SPase I catalysis. Fluorescence from the free MCA released after SPase I cleavage was measured to monitor reaction progress. SPase I enzymes from *S. aureus*, *S. epidermidis*, *B. subtilis*, and *E. coli* were screened to find a suitable SPase I for kinetic characterization. Among the SPase I enzymes screened, *S. aureus* SpsB has the highest activity followed by *S. epidermidis* SpsB, and then *B. subtilis* SipS and SipT. *E. coli* LepB has no detectable activity (Figure 2.7). Using the MCA substrate, *S. aureus* SpsB k_{cat} and K_M were determined to be $7.5 \times 10^{-3} \pm 0.7 \times 10^{-3}$ s^{-1} and 8.6 ± 2.2 μM, respectively (see Appendix C for the initial velocity vs substrate concentration curve E). By substituting P' residues with MCA, *S. aureus* SpsB showed an approximately 18 times reduction in k_{cat}, a 9 times increase in K_M and an overall 168 times reduction in catalytic efficiency compared to using dodecanoyl-K(Dabcyl)NGEVAKAAE(EDANS)T-NH_2 as the substrate.

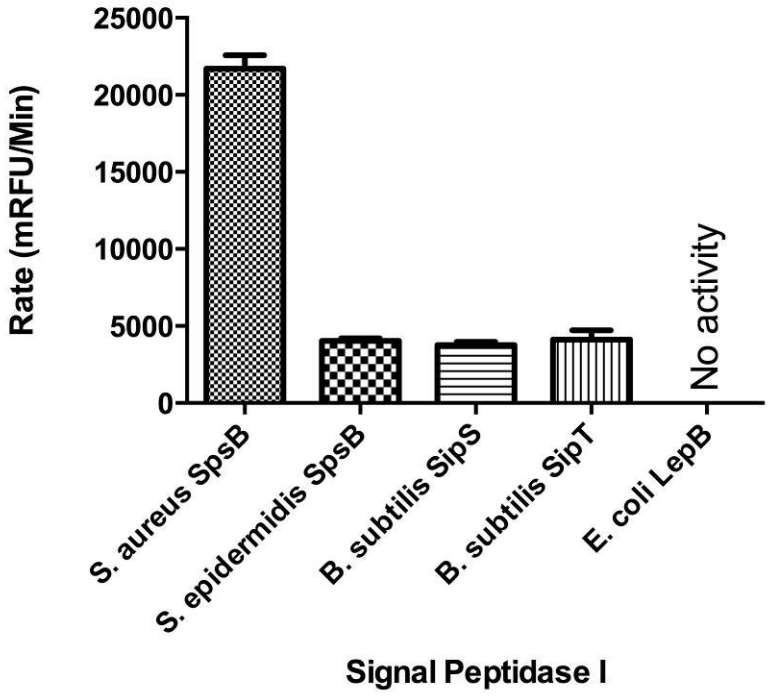
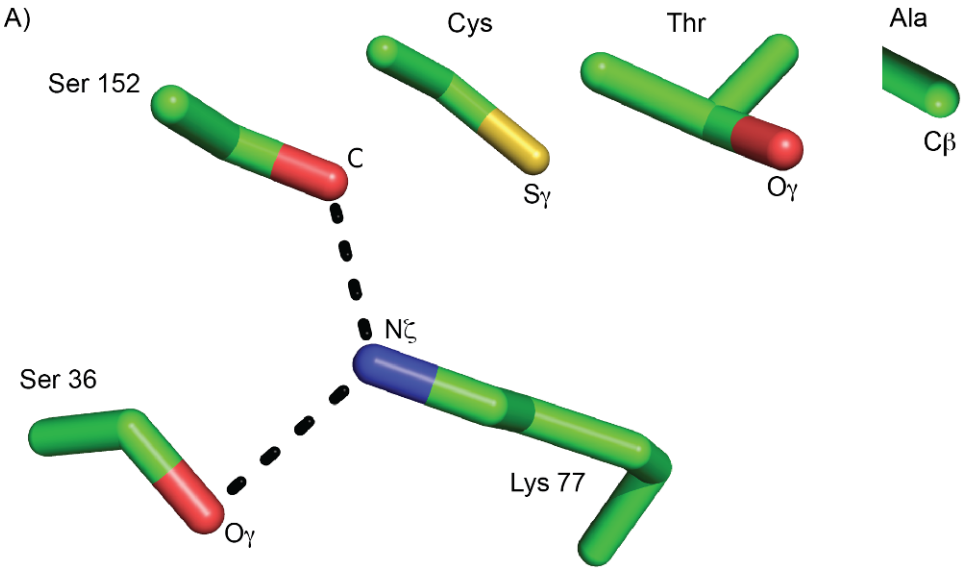


Figure 2.7. Gram-positive and Gram-negative SPase I activity comparison using the dodecanoyl-NGEVAKA-MCA substrate.

2μM of dodecanoyl-NGEVAKA-MCA was digested by 200nM of either *S. aureus* SpsB, *S. epidermidis* SpsB, *B. subtilis* SipS, *B. subtilis* SipT or *E. coli* LepB in a 100μL reaction. Reaction rate in *milli*-fluorescence unit per minute (mRFU/Min) was measured at Ex 380nm and Em 460nm. The error bar shown are the standard errors calculated from triplicate experiments.

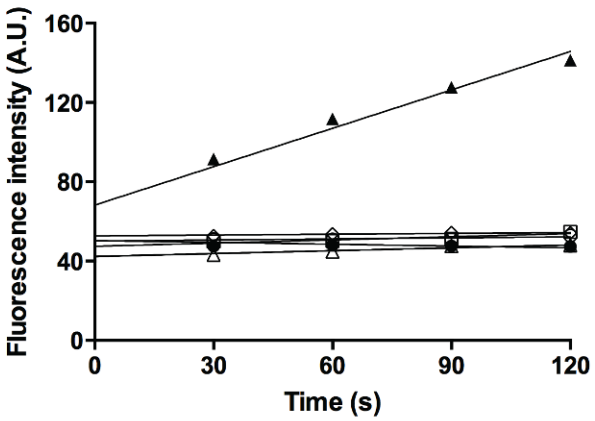
2.3.6. Serine 152 plays a critical role in the Ser / Lys dyad mechanism of S. aureus SpsB

Site-directed mutagenesis and enzymatic assays were carried out to determine whether or not *S. aureus* SpsB utilizes a coordinating serine in its Ser / Lys dyad mechanism. The proposed coordinating serine was predicted to be at position 152 based on sequence alignment against other Gram-positive SPase I enzymes in this study. The coordinating serine was mutated either to an alanine, a cysteine or a threonine (Figure 2.8A) and the effects of these mutations were examined using enzymatic assays. S152A, S152T and S152C mutants significantly reduced SpsB's specific activity, almost as much as when the two catalytic residues, serine 36 and lysine 77, were mutated to alanines (Figure 2.8B).



B)

Specific activity (μmol min^{-1} mg^{-1})



WT (200nM)	4.74×10^{-3}
S36A (1μM)	N/A
K77A (1μM)	N/A
S152T (1μM)	N/A
S152C (1μM)	N/A
S152A (1μM)	N/A

C)

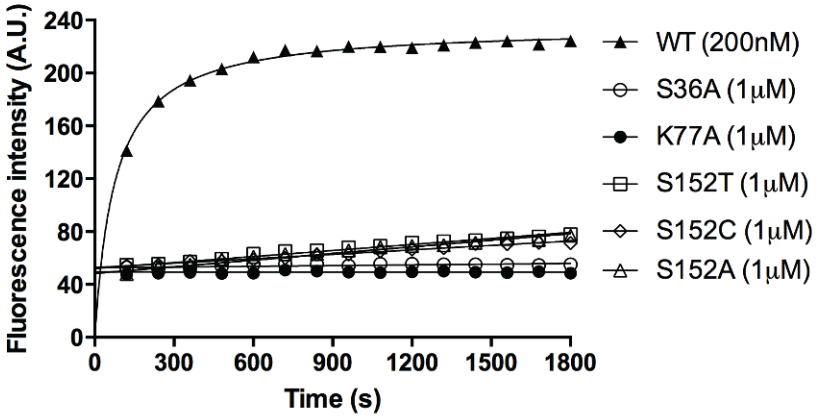


Figure 2.8. S. aureus SpsB WT, S36A, K77A, S152A, S152T and S152C mutants activity comparison.

A) The serine nucleophile, the lysine general base and the coordinating serine side chains are shown in sticks with carbon in green, nitrogen in blue, oxygen in red and sulphur in yellow. Threonine and cysteine mutant side chains are shown beside the coordinating serine side chain. Side chain functional group atoms are labeled. Hydrogen bonds are shown as dotted lines. B) 2μM of dodecanoyl-K(Dabcyl)NGEVAKAAE(EDA NS)-NH_2 was digested with either 200nM SpsB or 1μM mutant SpsB in a 100μL reaction mixture. Fluorescence intensities of the cleavage products were measured using Em 340nm and Ex 530nm and plotted over two minutes and the specific activity of the wild type enzyme calculated. Fluorescence intensity was converted to molar unit using the standard curve from Appendix B. C) *S. aureus* SpsB activity plotted over 180 minutes.

2.3.7. Effect of salt concentration, pH and other small chemical compounds on S. aureus Δ2-20 SpsB activity

Refolded *S. aureus* SpsB Δ2-20 activity was examined using buffers that varied in salt concentration, pH and small chemical compound additives to identify buffer components which may interact with SpsB Δ2-20 and potentially be used as additives to increase the likelihood of crystallizing the enzyme. All activity assays in this section were carried out using the Dabcyl-NGEVAKAAE(EDANS)T-NH_2 FRET peptide as the substrate. *S. aureus* SpsB Δ2-20 has a detergent requirement for optimal enzymatic activity, although a significant reduction in activity was observed at detergent concentrations below the critical micelle concentration (CMC) for the two different detergents tested, DDM and Triton X-100 (Figure 2.9). *S. aureus* SpsB Δ2-20 has a higher activity when DDM was used as the detergent, showing a two-fold increase in activity compared to Triton X-100, when the detergent concentration is at the CMC. Formation of micelles in the kinetic buffer may help increase enzymatic activity by i) providing a cell membrane like surface where SPase I enzymes can dock on a proposed membrane interacting surface; ii) by allowing lipidated peptide substrates to insert themselves into the micelle, helping in substrate presentation, and better mimicing *in vivo* conditions; and iii) by recruiting both SPase I and the substrate to the micelle, thus reducing the three dimensional search for the substrate by SPase I to a two dimensional search.

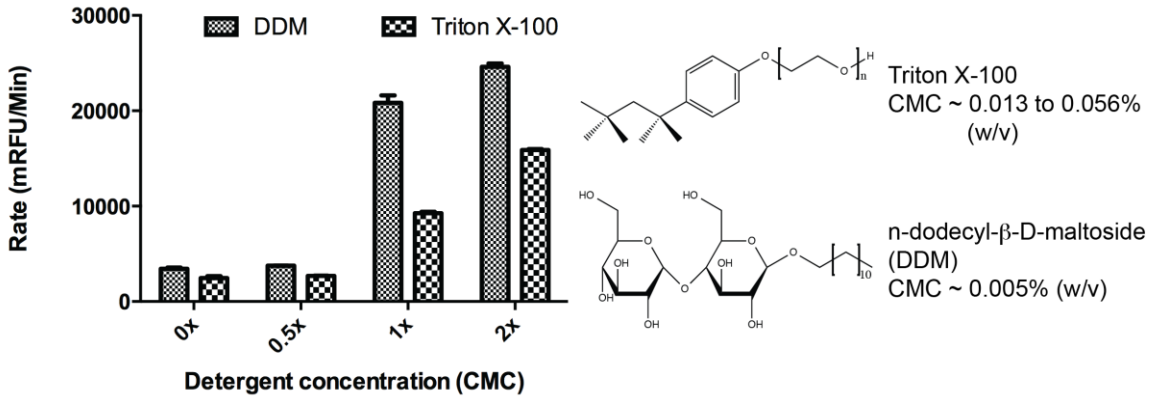


Figure 2.9. Effect of two detergents, Triton X-100 and n-dodecyl-β-D-maltoside (DDM), on S. aureus SpsB Δ2-20 activity.

20μM of Dabcyl-NGEVAKAAE(EDANS)T-NH_2 was incubated with 100nM *S. aureus* SpsB Δ2-20. Intensity of the cleavage products was measured at Ex 340nm and Em 530nm and the reaction rate in *milli*-relative fluorescence unit per minute (mRFU/Min) was calculated. The bar graph shows SpsB Δ2-20's reaction rate in 20mM Tris pH8.0, 100mM NaCl buffer containing 0, 0.0025, 0.005 or 0.01% n-dodecyl-β-D-maltoside (DDM) or 0, 0.025, 0.05 or 0.1% Triton X-100 which corresponds to 0, 0.5, 1 or 2 times the critical micelle concentration (CMC). Error bars shown are the standard errors calculated from triplicate experiments. The chemical structures of Triton X-100 and DDM are shown with their respective CMC. Triton X-100 is a mixture of detergents where each individual Triton X-100 molecule may have a different number of ethylene oxide repeats. The average number of ethylene oxide repeats (n) is 9.

The effect of pH on *S. aureus* SpsB Δ2-20 activity was considered, using buffers with pH ranging from 5 to 11. SpsB Δ2-20 has the highest activity at pH8; with higher activity levels seen at basic pH rather than at acidic pH. *S. aureus* SpsB Δ2-20 activity was abolished at pH 5 and 6 (Figure 2.10).

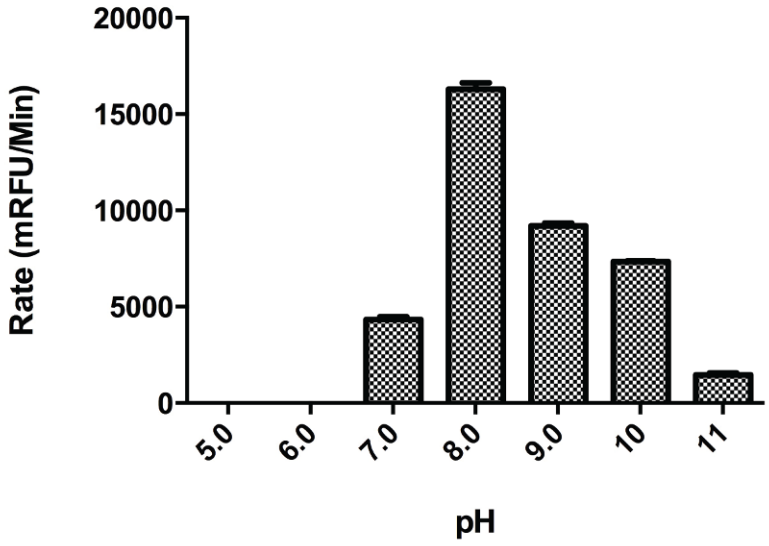


Figure 2.10. S. aureus SpsB Δ2-20 pH profile.

20μM of Dabcyl-NGEVAKAAE(EDANS)T-NH_2 was incubated with 100nM of *S. aureus* SpsB Δ2-20 in buffers with pH ranging from 5 to 11. Intensity of the cleavage product was measured at Ex 340nm and Em 530nm and the reaction rate in *milli*-relative fluorescence unit per minute (mRFU/Min) was calculated. The bar graph shows the reaction rate at each pH point with error bars showing the standard errorr calculated from triplicate experiments.

Lastly, small chemical compounds were tested to examine their effect on *S. aureus* SpsB Δ2-20 activity. SpsB Δ2-20 showed an increase in activity with an increase in NaCl concentration. The activity was increased 3-fold when going from an NaCl free buffer to a buffer containing 100mM NaCl. The activity continued to increase to a maximum at 750mM NaCl, then decreased as the NaCl concentration continued to increase (Figure 2.11A). Various divalent cations, Ca^{2+}, Mg^{2+}, Zn^{2+}, Co^{2+} and Ni^{2+}, at 1mM were tested. Co^{2+} and Zn^{2+} abolished *S. aureus* SpsB Δ2-20 activity while the other cations did not affect activity of SpsB Δ2-20 (Figure 2.11B). The effect of DMSO on SpsB Δ2-20 activity was studied because all the substrates tested were prepared in DMSO and all reaction mixtures contained at least 1% of DMSO. A significant decrease in activity was observed at DMSO concentrations over 5%, although *S. aureus* SpsB Δ2-20 still retained most of its activity at 15% DMSO (Figure 2.11C). SpsB Δ2-20 activity was affected by the presence of a reducing agent, dithiothreitol (DTT), being reduced by half in a buffer containing 5mM DTT (Figure 2.11D). This is interesting given that *S. aureus* SpsB Δ2-20 does not have any disulfide bonds. Glycerol and EDTA were seen to

have a negative impact on SpsB Δ2-20 activity; as glycerol or EDTA concentration increases, SpsB Δ2-20 activity decreases (Figure 2.11E and F).

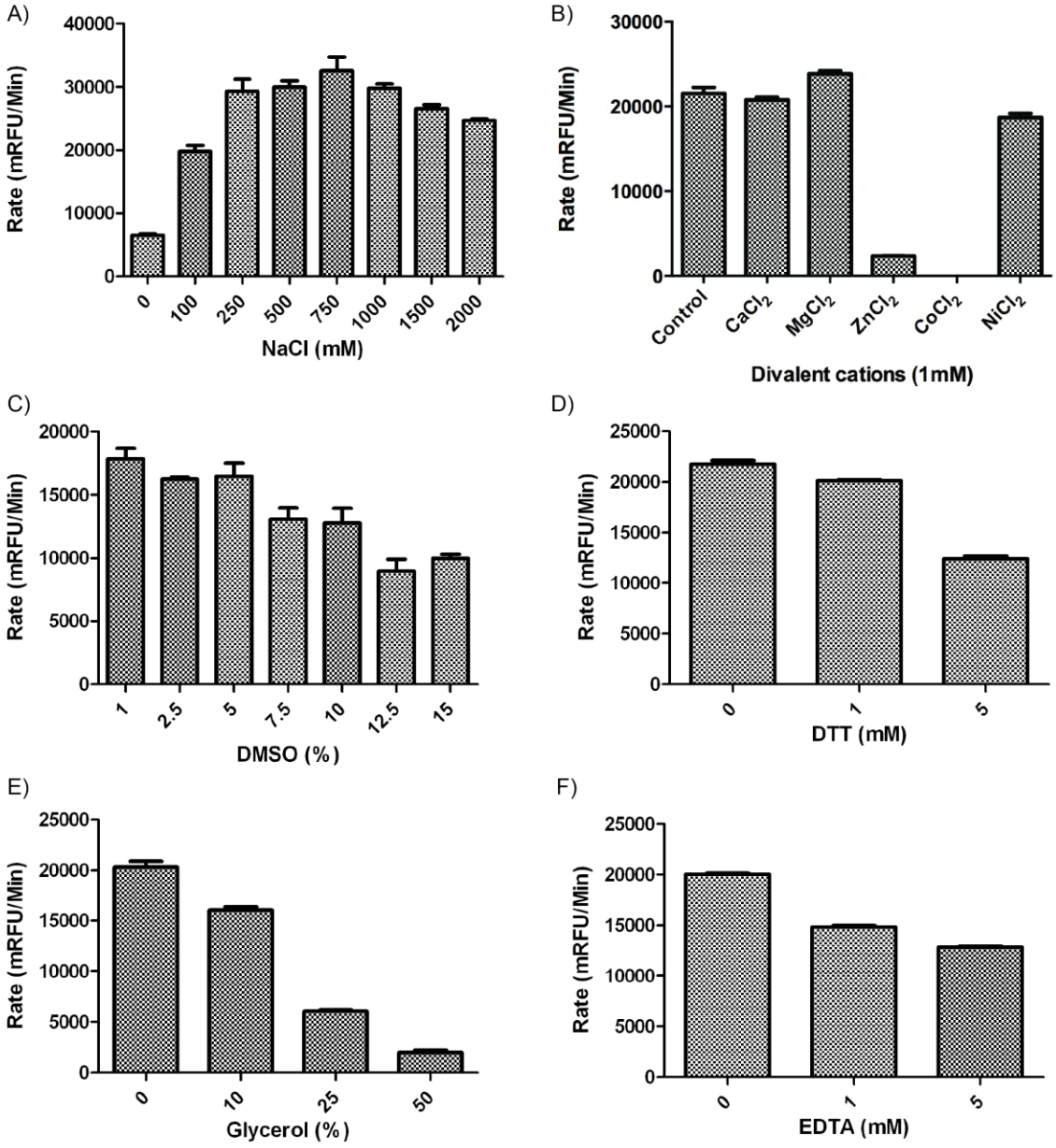


Figure 2.11. Effect of small chemical compounds on S. aureus SpsB Δ2-20's activity.

20μM of Dabcyl-NGEVAKAAE(EDANS)T-NH_2 was incubated with 100nM of *S. aureus* SpsB Δ2-20 in kinetic buffer (20mM Tris pH8.0, 100mM NaCl, 0.01% DDM) containing varying concentrations of small chemical compounds. Intensity of the cleavage product was measured at Ex 340nm and Em 530nm and the reaction rate was calculated in *milli*-relative fluorescence unit per minute (mRFU/min). The error bars shown are the standard errors calculated from triplicate experiments. A) SpsB Δ2-20 activity in kinetic buffer containing 0, 100, 250, 500, 1000, 1500 and 2000mM of NaCl. B) SpsB Δ2-20 activity in kinetic buffer containing 1mM of either $CaCl_2$, $MgCl_2$, $ZnCl_2$, $CoCl_2$ or $NiCl_2$. The control is the activity of SpsB Δ2-20 activity in kinetic buffer with no divalent cations added. C) SpsB Δ2-20 activity in kinetic buffer containing 1, 2.5, 5, 7.5, 10, 12.5 and 15% (v/v) DMSO. D) SpsB Δ2-20 activity in kinetic buffer containing 0, 1 and 5mM DTT. E) SpsB Δ2-20 activity in kinetic buffer containing 0, 10, 25 and 50% (v/v) glycerol. F) SpsB Δ2-20 activity in kinetic buffer containing 0, 1 and 5mM EDTA.

2.3.8. S. aureus SpsB Δ2-20 S36A crystallization

S. aureus SpsB Δ2-20 S36A was first crystallized as needle clusters in 20% w/v PEG3350, 0.2M sodium citrate (PEG Ion Screen condition #46, Hampton Research) during initial screening for protein crystallization conditions (Figure 2.12A). The initial hit was optimized by co-crystallizing *S. aureus* SpsB Δ2-20 S36A with arylomycin A_2 inhibitor and the resulting complex formed single rectangular crystals in the same condition (PEG Ion #46) (Figure 2.12B). Changing the crystallization reservoir to 16.5% w/v PEG 3350, 0.18M tri-sodium citrate dihydrate further optimized this hit, increasing crystal growth (Figure 2.12C). X-ray diffraction experiments were carried out but none of the crystals diffracted beyond 4.5Å.

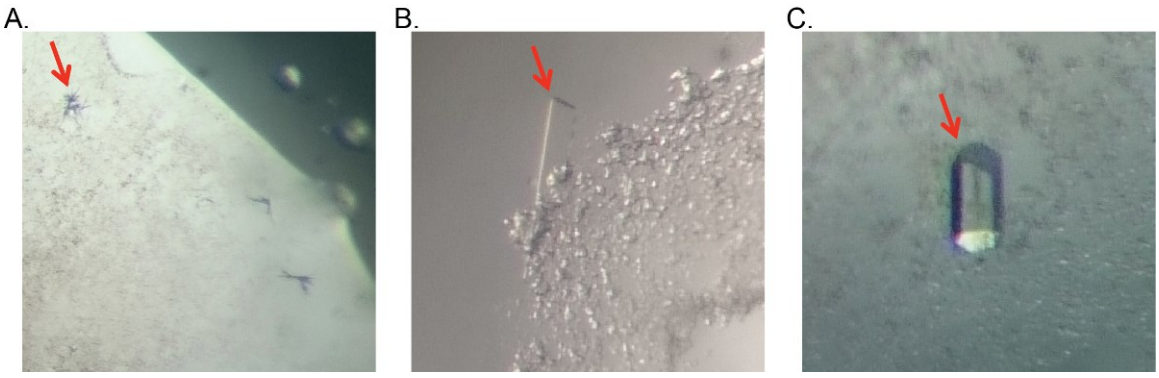


Figure 2.12. S. aureus SpsB Δ2-20 S36A protein crystals.

S. aureus SpsB Δ2-20 S36A protein crystals are marked by red arrows and photos were taken under 1,000x magnification. A) Apo-enzyme (*S. aureus* SpsB Δ2-20 S36A) crystallized in 20% v/w PEG 3350 and 0.2M tri-sodium citrate dihydrate. B) SpsB Δ2-20 S36A in complex with arylomycin A2 crystallized in 20% v/w PEG 3350 and 0.2M tri-sodium citrate dihydrate. C) Enzyme/inhibitor complex crystallized in 16.5% v/w PEG 3350 and 0.18M tri-sodium citrate dihydrate.

2.4. Discussion

In this chapter we utilized activity assays to assess our *S. aureus* SpsB constructs, peptide substrates and buffer components for use in crystallization experiments. A truncated *S. aureus* SpsB construct was designed, lacking the first twenty residues which still retained part of its predicted single N-terminal transmembrane (residues 7 to 29). Although this construct expressed as inclusion bodies, it was successfully refolded through simple dialysis with common laboratory reagents. Our construct design and refolding protocols were considered successful as crystallization amounts (>10 mg/mL) of a truncated *S. aureus* SpsB, which has enzymatic activities close to that of the full-length enzyme, were obtained.

Various catalytically impaired constructs were made, with S36A, K77A, S152A, S152T and S152C mutations, which may be useful in future crystallization trials as *S. aureus* SpsB is plagued by the self-degradation common among proteases. The K77A will be especially useful for co-crystallization with a substrate as it may slow down catalysis which may allow the capturing of the acyl-enzyme intermediate. The lysine, which is mutated to an alanine in this construct, is required to activate the deacylating water in the proposed Ser/Lys catalytic dyad mechanism. This strategy was previously used successfully to trap an acyl-enzyme intermediate of a mutant viral protease VP4 with the catalytic lysine mutated to an alanine in complex with the C-terminal internal cleavage site of another VP4 protease (Chung & Paetzel, 2013). We have tried co-crystallizing the K77A mutant with a lipidated FRET peptide substrate that has a sub-μM K_M towards the enzyme but no crystals have been observed to date. This peptide could also be chemically modified to become an inhibitor (eg. substituting the carboxylic acid group at the C-terminus to an aldehyde or a borate group) and used as a co-crystallization additive.

One interesting observation we have made is that high NaCl concentration in a buffer promotes protein folding, solubility and enzymatic activity. The refolding buffer contained 1M NaCl which, when reduced to 100mM will cause heavy precipitation of truncated *S. aureus* SpsB during dialysis for protein refolding. We saw the enzymatic activity of truncated *S. aureus* SpsB more than triple when the enzyme was moved from

a low salt buffer into a high salt buffer. This is the opposite of *E. coli* SPase I which is inhibited at NaCl concentrations above 160mM (Zwizinski, Date & Wickner 1981).

Through enzymatic characterization we have shown that truncated *S. aureus* SpsB has a detergent dependence for optimal enzymatic activity; more specifically, the detergent will only exert its effect at or above the CMC. The enzymatic activity of truncated *S. aureus* SpsB more than quadrupled in a buffer containing DDM at the CMC compared to a detergent free buffer while truncated *S. aureus* SpsB in a buffer containing DDM at half the CMC has a similar enzymatic activity compared to the detergent free buffer. This suggests that truncated *S. aureus* SpsB may have a hydrophobic surface which interacts with the micelle surface. The *S. aureus* SpsB homology model shows a large surface rich in hydrophobic residues and aromatic residues which are usually found at the protein / membrane interface (Figure 2.12) It is possible that through this micelle surface interaction, truncated *S. aureus* SpsB adopts a more enzymatically active conformation. This would explain why detergents only exert their effect on the enzymatic activity of *S. aureus* SpsB when the detergent concentration is at or above the CMC where detergent molecules can form micelles. Different detergents also have different degrees of effect on truncated *S. aureus* SpsB's enzymatic activity. Based on this observation, further studies should examine other detergents and their effects on enzymatic activity using activity assays to search for detergents may be used as additives in crystallization trials.

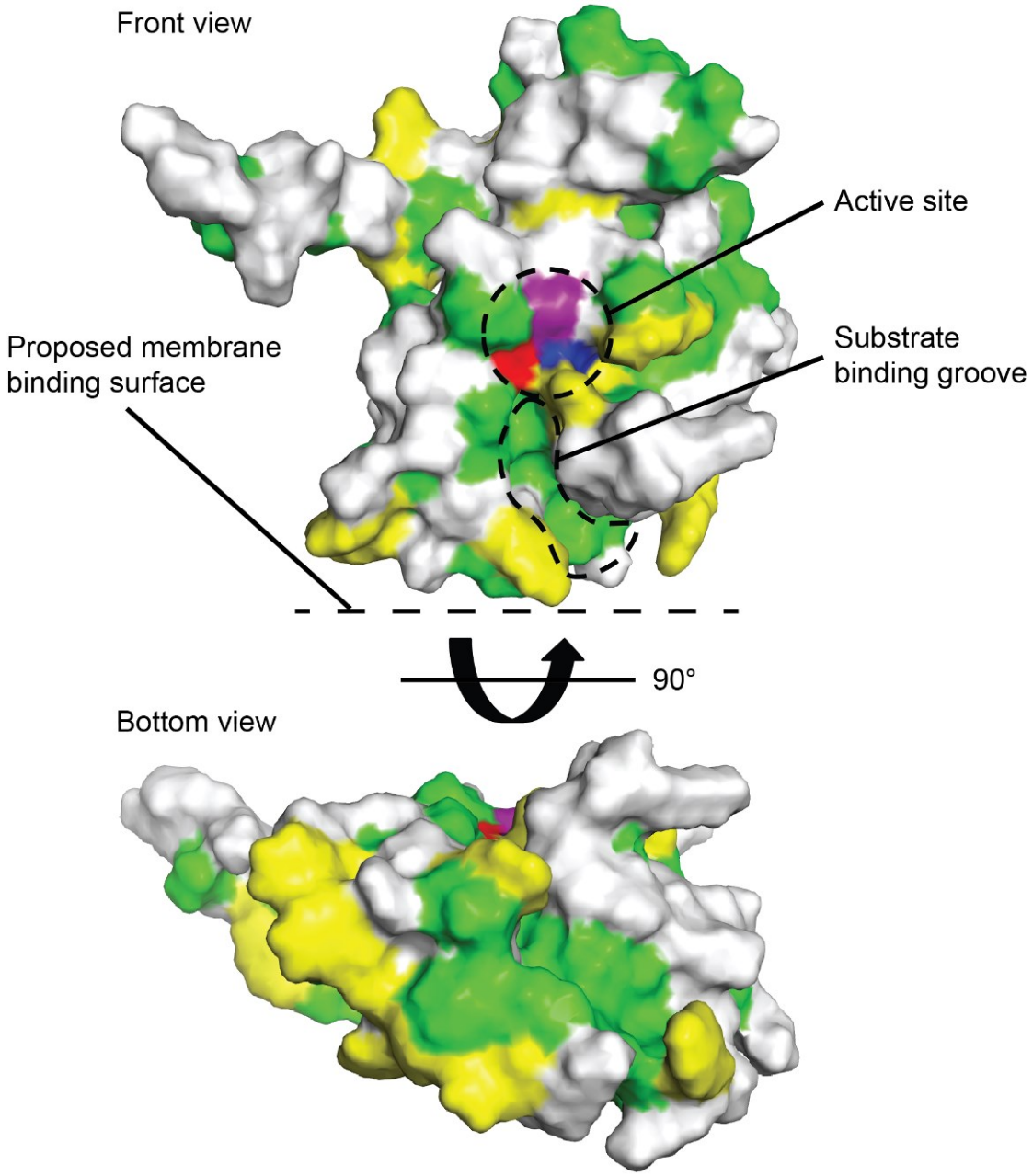


Figure 2.13. S. aureus SpsB homology model

The *S. aureus* SpsB homology model generated by SWISS-MODEL (Guex, Peitsch & Schwede 2009) using *E. coli* SPase I (PDB: 1KN9) as the template and energy minimized by YASARA Energy Minimization Server (Krieger et al. 2009). This model includes residues 24 to 178. Hydrophobic residues (alanine, valine, leucine, isoleucine, glycine and methionine) are green, aromatic residues (phenylalanine, tyrosine, tryptophan) are yellow, with the rest of the molecule in white. The catalytic residues, serine nucleophile, lysine general base and coordinating serine are coloured in red, blue and purple, respectively.

3. Investigation of SPase I box B residues

3.1. Overview

This chapter will examine two of the many highly conserved residues in box B of SPase I enzymes using Gram-negative *E. coli* LepB as the model SPase I. The two residues of interest are *E. coli* LepB proline 88 and serine 89 and their equivalent residues in Gram-positive *S. aureus* SpsB, lysine 33 and glycine 34, respectively (Figure 3.1). In Gram-negative SPase I enzymes, the proline and serine are highly conserved but in Gram-positive SPase I enzymes, only the glycine is highly conserved while other residues may be substituted at the lysine position but never a proline. *E. coli* LepB serine 89 has been proposed to play a role in stabilizing the tetrahedral oxyanion transition state intermediate through its side chain hydroxyl group; and when it was mutated to an alanine, a significant reduction in enzymatic activity was observed (Carlos et al. 2000). Until now, there have been no investigations into proline 88 and its equivalent residue in *S. aureus* SpsB. To study proline 88 and its role, mutagenesis was done to introduce Gram-positive conserved residues into *E. coli* LepB, making three Gram-positive like mutants (P88A, S89G, and P88A/S89G). The opposite procedure was carried out on *S. aureus* SpsB (K33A, K33P, G34S and K33P/G34S), generating Gram-negative like mutants. Double mutants (i.e. P88A/S89G and K33P/G34S) were made to investigate whether or not the two highly conserved residues have a synergistic effect on enzymatic activity. Activity assays were performed using a Gram-positive signal peptide substrate, dodecanoyl-K(Dabcyl) NGEVAKAAE (EDANS)T-NH_2 (described in Section 2.1), to examine the effect of these mutations. The *E. coli* LepB P88A mutant was found to have a higher catalytic efficiency compared to the wild type while *S. aureus* SpsB K33P mutant was discovered to be catalytically inactive. Surprisingly, *S. aureus* SpsB K33A catalytic efficiency is similar to that of the wild type enzyme, which suggests that this position can be any residue other than a proline, as seen in the Gram-positive sequence alignment (Figure 3.1). *E. coli* LepB and the P88A mutant were also examined using the pre-protein substrates, PONA and DANA. PONA (described in Section 1.3) is a pre-

protein carrying an *E. coli* OmpA signal peptide while DANA is a derivative of PONA where the C-region of the OmpA signal peptide (P7 to P1) is substituted with the sequence from the FRET peptide substrate (FATVAQA to NGEVAKA) derived from a *B. licheniformis* pre-protein. Results showed that *E. coli* LepB processes PONA faster than DANA and that the *E. coli* LepB P88A mutant has higher enzymatic activity than wild type LepB. This is in agreement with the results from the FRET peptide substrate assay. Lastly, *in vivo* processing of PONA and DANA was monitored by following the growth of BL21(DE3) cells (using $O.D._{600}$) overexpressing either PONA or DANA. The culture overexpressing DANA has a lower $O.D._{600}$ compared to the culture overexpressing PONA during stationary phase, which may be the result of differential processing by endogenous *E. coli* SPase I.

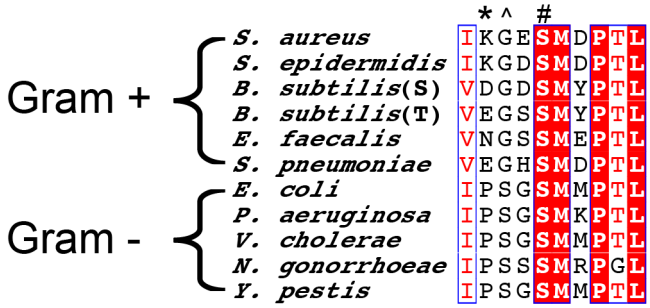


Figure 3.1. Gram-positive and Gram-negative SPase I enzymes box B sequence alignments.

Multiple sequence alignment of SPase I box B from Gram-positive bacteria: *Staphylococcus aureus*, *Staphylococcus epidermidis*, *Enterococcus faecalis*, and *Streptococcus pneumonia*e and Gram-negative bacteria: *Escherichia coli*, *Pseudomonas aeruginosa*, *Vibrio cholera*, *Neisseria gonorrhoeae*, and *Yersinia pestis*. The nucleophile serine (residue 91 in *E. coli*, 36 in *S. aureus*) is marked with an #. The proposed oxyanion hole stabilizing residue (serine 89 in *E. coli*, glycine 33 in *S. aureus*) is marked with an ^. The conserved proline of interest in Gram-negative SPase I (proline 88 in *E. coli*) and its equivalent residue in Gram-positive SPase I (lysine 33 in *S. aureus*) is marked with an *.

3.2. Materials and methods

3.2.1. S. aureus SpsB and E. coli LepB mutagenesis

S. aureus SpsB and *E. coli* LepB box B mutants and mutagenesis primer pairs used are listed in Table 3.1. Site-directed mutagenesis was done following procedures outlined in Section 2.2.1.

Table 3.1. Mutagenesis primer pairs used for each S. aureus SpsB and E. coli LepB box B mutant.

SPase I	Mutation	Primer pair (forward / reverse)
S. aureus SpsB	K33A	5'-CCATATACAATTGCAGGTGAATCAATG-3' 5'-CATTGATTCACCTGCAATTGTATATGG-3'
	K33P	5'-CCATATACAATTCCAGGTGAATCAATG-3' 5'-CATTGATTCACCTGGAATTGTATATGG-3'
	G34S	5'-TATACAATTAAAAGTGAATCAATGGAT-3' 5'-ATCCATTGATTCACTTTTAATTGTATA-3'
	K33P, G34S*	5'-TATACAATTCCAAGTGAATCAATGGAT-3' 5'-ATCCATTGATTCACTTGGAATTGTATA-3'
E. coli LepB	P88A	5'-CCGTTCCAGATCGCGTCAGGTTCGATG-3' 5'-CATCGAACCTGACGCGATCTGGAACGG-3'
	S89G	5'-TTCCAGATCCCGGGAGGTTCGATGATG-3' 5'-CATCATCGAACCTCCCGGGATCTGGAA-3'
	P88A, S89G#	5'-TTCCAGATCGCGGGAGGTTCGATGATG-3' 5'-CATCATCGAACCTCCCGCGATCTGGAA-3'

Mutated codons are underlined. * *S. aureus* SpsB K33P used as the template. # *E. coli* LepB P88A used as the template.

3.2.2. S. aureus SpsB and E. coli LepB mutant protein overexpression and purification

All mutant enzymes were purified as described in Section 2.2.2. Briefly, each mutant enzyme was overexpressed in *E. coli* BL21(DE3), an expression host with IPTG induction, purified using Ni^{2+}-NTA affinity chromatography, dialyzed against kinetic buffer (20mM Tris pH8.0, 100mM NaCl and 0.01% DDM) and kept at -80°C until needed.

3.2.3. Activity characterization of S. aureus SpsB and E. coli LepB enzymes and their mutants using a FRET lipidated peptide substrate

Activity assays were done as described in Section 2.2.7. Briefly, initial activity screening was done using 100nM *S. aureus* SpsB and the K33A, K33P, G34S and K33P/G34S mutant enzymes or 500nM *E. coli* LepB and the P88A, S89G and P88A/S89G mutant enzymes with 20μM dodecanoyl-K(Dabcyl)NGEVAKAAE(EDANS)

T-NH_2 FRET peptide substrate. k_{cat} and K_M were determined for the *S. aureus* SpSB K33A mutant (5nM) with the substrate concentration ranging from 0.25μM to 8μM and the *E. coli* LepB P88A mutant (100nM) with the substrate concentration ranging from 2.5μM to 40μM.

3.2.4. Activity characterization of E. coli LepB and the P88A mutant enzymse using pre-proteins, PONA and DANA

E. coli LepB and the P88A mutant enzymes were prepared in kinetic buffers as 20μM stocks and serially diluted (10, 20, 100, 200, 1000, 2000, and 10000x dilution). 20μL of reaction mixtures were set up containing 3μg of either PONA or DANA in kinetic buffer and the reaction was initiated by adding 1μL of enzyme solution. The reaction was allowed to proceed for an hour at 37°C, and then terminated by adding 5μL of 6X SDS-PAGE loading buffer followed by boiling at 95°C for 10 minutes. Digestion results were analyzed on a 15% SDS-PAGE gel stained with PageBlue protein staining solution.

3.2.5. E. coli BL21(DE3) growth curve assay

Competent cells of the E. coli BL21(DE3) expression host were transformed with the pET28a+ (Novagen) vector harbouring either PONA, DANA, PONA +1 Pro or DANA +1 Pro. Overnight cultures for each of the above constructs were prepared by inoculating 5mL of LB containing 50μg/mL kanamycin with freshly streaked single colonies and grown at 37°C overnight. The next day, 100mL of LB containing 50μg/mL kanamycin was inoculated with 1mL of the overnight culture and grown at 37°C with shaking (250rpm). $O.D._{600}$ was measured every 30 minutes using a quartz cuvette (path length = 10mm) on a SpectraMax M5 plate reader (Molecular Devices). Protein expression was induced using IPTG (1mM final concentration) at $O.D._{600}$ between 0.6 to 0.7.

3.3. Results

3.3.1. E. coli LepB P88A has higher catalytic efficiency compared to the wild type enzyme

All *E. coli* LepB and *S. aureus* SpsB mutants were subjected to a pre-screening for enzymatic activity to identify feasible enzymes on which to carry out full kinetic characterization. Among all the enzymes tested, only the *E. coli* LepB P88A and *S. aureus SpsB K33A* mutants were selected, based on their enzymatic activity relative to their respective wild type enzyme (Figure 3.2). *E. coli* LepB P88A has an increased activity while *S. aureus* SpsB K33A has a similar activity, compared to the wild type enzyme. All other mutant enzymes did not show detectable activity within the experimental parameters used.

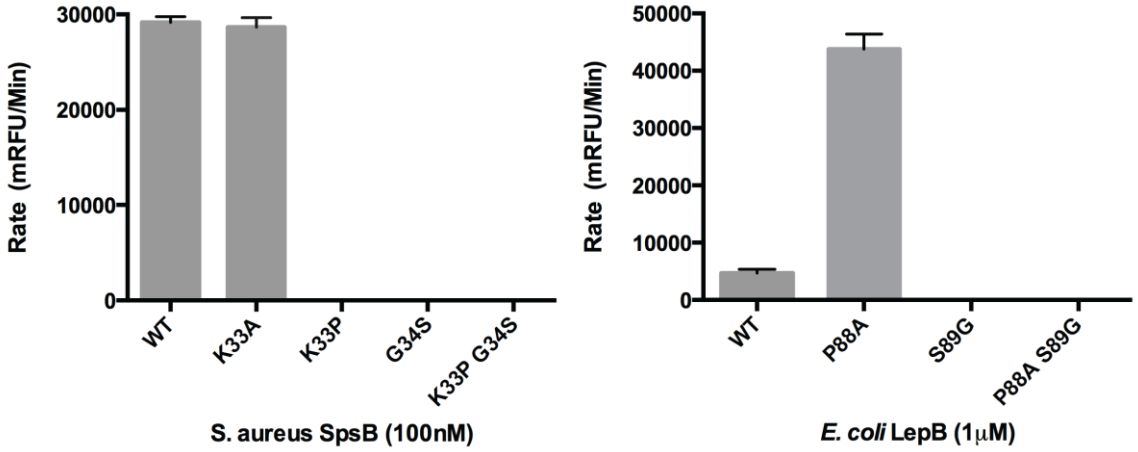


Figure 3.2. S. aureus and E. coli SPase I enzymatic activities.

20μM of dodecanoyl-K(Dabcyl)NGEVAKAAE(EDANS)-NH2 was digested by either 100nM *S. aureus* SpsB and mutants or 1μM *E. coli* LepB and its mutants. Reaction rate in *milli*-relative fluorescent unit per minute (mRFU/Min) measured at Em 340nm and Ex 530nM was plotted for each enzyme. The error bars represent the standard error calculated from triplicate experiments.

S. aureus SpsB K33A k_{cat} and K_M are similar to the wild type enzyme while *E. coli* LepB P88A has 10 times higher k_{cat} and 4.5 times lower K_M compared to the wild type enzyme (Table 3.2).

Table 3.2. ***k_{cat} and K_M values for S. aureus SpsB, E. coli LepB and their mutant enzymes.***

SPase I	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($M^{-1}s^{-1}$)
S. aureus SpsB	$1.3 \times 10^{-1} \pm 0.7 \times 10^{-1}$	0.9 ± 0.2	1.5×10^{5}
S. aureus SpsB K33A	$1.5 \times 10^{-1} \pm 0.8 \times 10^{-1}$	1.0 ± 0.2	1.5×10^{5}
E. coli LepB	$1.5 \times 10^{-3} \pm 0.1 \times 10^{-3}$	32.0 ± 4.5	4.7 × 10
E. coli LepB P88A	$1.5 \times 10^{-2} \pm 0.9 \times 10^{-2}$	7.1 ± 1.2	2.1×10^{3}

See Appendix C (A, F, J, K) for initial velocity vs substrate concentration curves.

3.3.2. E. coli LepB preferentially processes PONA over DANA

Two pre-protein substrates, PONA and DANA, were used to examine whether or not introducing the P88A mutation into *E. coli* LepB causes it to preferentially bind and cleave pre-proteins with a Gram-positive signal peptide. The digestion results showed that both *E. coli* LepB and the P88A mutant preferentially process PONA (the Gram-negative signal peptide) over DANA (the Gram-positive signal peptide) and that the P88A mutant has a higher enzymatic activity than the wild type towards DANA (Figure 3.3A). The *in vivo* processing of PONA and DANA by E. coli SPase I was examined by overexpressing each pre-protein and their non-cleavable counterparts (+1 proline in the P1' position of the signal peptide) in *E. coli* BL21(DE3) (Figure 3.3B). Significant processing of PONA was observed as expected but surprisingly, no significant processing of DANA was observed within two hours of induction. DANA's SDS-PAGE gel analysis could be not distinguished from that of the two non-cleavable PONA and DANA counterparts. The *in vivo* study of BL21(DE3) cultures overexpressing PONA, DANA, and their non-cleavable counterparts showed the cultures behaved differently after ITPG induction (Figure 3.3C). The cultures overexpressing PONA and DANA after IPTG induction both follow normal growth curves with increasing cell density and stabilizing once the cultures reach the stationary phase. The only difference between PONA and DANA expressing culture is that the DANA expressing culture has a lower final cell density during the stationary phase compared to that of the PONA expressing culture. Overexpressing PONA +1 Pro and DANA +1 Pro negatively impacts cell density. After IPTG induction, both cultures show a decrease in cell density due to cell lysis.

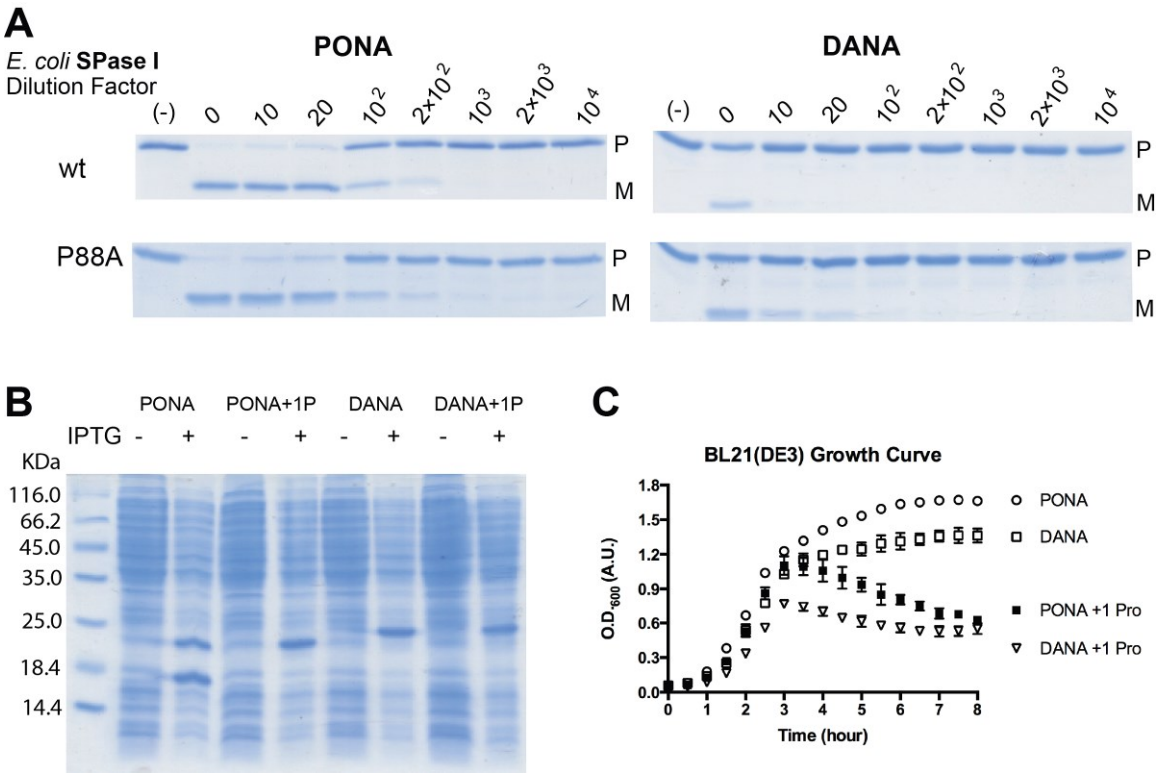


Figure 3.3. PONA and DANA processing by E. coli LepB.

A) Enzymatic reactions were carried out in kinetic buffer containing 3µg PONA or DANA. The reactions were initiated by the addition of 1µL of either *E. coli* LepB (wt) or the P88A mutant at a starting concentration of 20µM, followed by a series of dilutions (10, 20, 100, 200, 1000, 2000, and 10000x dilution). The negative control (-) was set up without *E. coli* SPase I enzymes. Pre-proteins are denoted by a P while the processed mature proteins are denoted by an M. The reactions were terminated by the addition of 6x loading protein sample loading buffer and analyzed on a 15% SDS-PAGE gel stained with PageBlue protein staining solution. B) Expression analysis of PONA and DANA and their non-cleavable counterparts (PONA+1P and DANA+1P) with (+) or without (-) IPTG induction. Samples were taken after 2 hours of induction and run on a 15% SDS-PAGE gel, later stained with PageBlue. C) 100mL *E. coli* BL21(DE3) cultures harbouring either PONA, DANA, the non-cleavable PONA +1 proline or DANA +1 proline were grown at 37°C. Absorbance at 600nm ($O.D._{600}$) readings for each culture were taken every 30 minutes over a course of 8 hours. Each culture was induced between $O.D._{600}$ of 0.6 to 0.7. The error bars indicate the standard errors calculated from triplicate experiments.

3.4. Discussion

Gram-positive and Gram-negative SPase I enzymes share five conserved regions (boxes). However, although many residues in each box are conserved across species but some are only conserved in Gram-negative bacteria and others only in Gram-positive bacteria. Box B houses two such residues; in Gram-negative bacteria, the proposed oxyanion intermediate stabilizing residue is often a serine (threonine is

observed in some Gram-negative bacteria) and the immediately preceding residue is always a proline while in Gram-positive bacteria, it is always a glycine in the place of the oxyanion hole serine (threonine) while the immediately preceding residue does not show any conservation in charge or size but is never a proline. In *E. coli* LepB, the oxyanion hole serine can be replaced with a threonine without having an impact on LepB enzymatic activity but activity is lost when the serine is replaced by an alanine suggesting the side chain hydroxyl participates in stabilizing the tetrahedral oxyanion intermediate (Carlos et al. 2000) (Figure 3.4A). Gram-positive SPase I enzymes have a glycine instead a serine, suggesting the tetrahedral oxyanion intermediate is not stabilized through a side chain interaction. We propose that the tetrahedral oxyanion intermediate may be stabilized by a water molecule recruited by the glycine main chain amide as the main chain amide may be too far away to form a hydrogen bond with the tetrahedral oxyanion intermediate (Figure 3.4B). *S. aureus* SpsB cannot tolerate a serine in the place of a glycine, even though for *E. coli* SPase I, the serine provides the hydroxyl group needed for oxyanion stabilization. In the case of *S. aureus* SpsB, a serine side chain may displace the water molecule and the serine hydroxyl group may not be in an optimal geometry to stabilize the oxyanion.

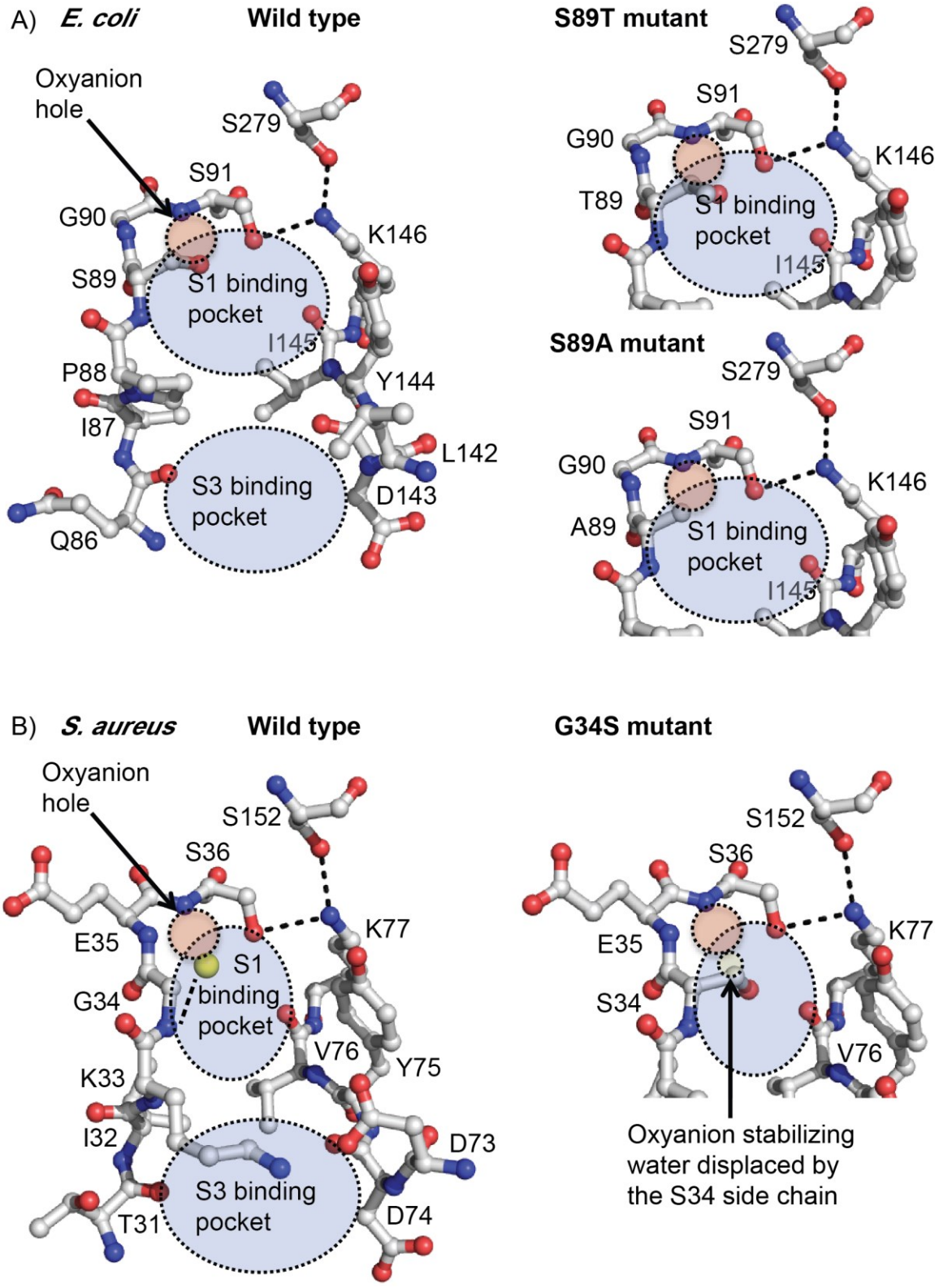


Figure 3.4. ***Substrate binding grooves of the E. coli and S. aureus SPase I enzymes.***

A) *E. coli* SPase I binding groove residues in ball and stick representation. Carbon atoms are white, nitrogen atoms are blue, oxygen atoms red and hydrogen atoms are not shown. Hydrogen bonds are shown as dotted lines, the S1 and S3 substrate binding pockets are shown as dotted blue circles, and the proposed oxyanion hole is shown as a red dotted circle. B) *S. aureus* SPase I binding groove with the same labeling scheme as *E. coli* SPase I. The proposed oxyanion stabilizing water is shown as a yellow sphere.

The role of the residue immediately preceding the oxyanion stabilizing residue was examined using peptide-based and protein-based activity assays. We propose that the main chain amide of this residue may hydrogen bond with a signal peptide substrate when this residue is not a proline (Figure 3.5B). This is supported by the similar enzymatic activity exhibited by both *S. aureus* SpsB and the K33A mutant, showing that the side chain of alanine does not affect enzymatic activity. The K33P mutant did not show measurable activity with the peptide based assays but did show residual enzymatic activity in the PONA digestion assays, although at a much lower level than that of the wild type enzyme (data not shown). This indicates that the enzyme is still active but may have a much lower affinity for its substrates, hence a slower processing rate. That the *E. coli* LepB P88A mutant has an overall higher catalytic efficiency compared to the wild type enzyme towards the same FRET peptide substrate may be due to that the residue immediately preceding the oxyanion stabilizing residue may participate in substrate binding through its main chain amide hydrogen (Figure 3.5A). Overall, our data suggests that even though both Gram-positive and Gram-negative SPase I enzymes process the same substrate (signal peptide of pre-proteins with the AXA cleavage consensus sequence), and share many conserved regions and potential 3-dimentional structures, but they may interact with their substrates forming a different hydrogen bonding network and stabilizing their oxyanion intermediates in a different fashion. Understanding these differences may lead to more effective design of antibiotics towards specific groups of bacteria.

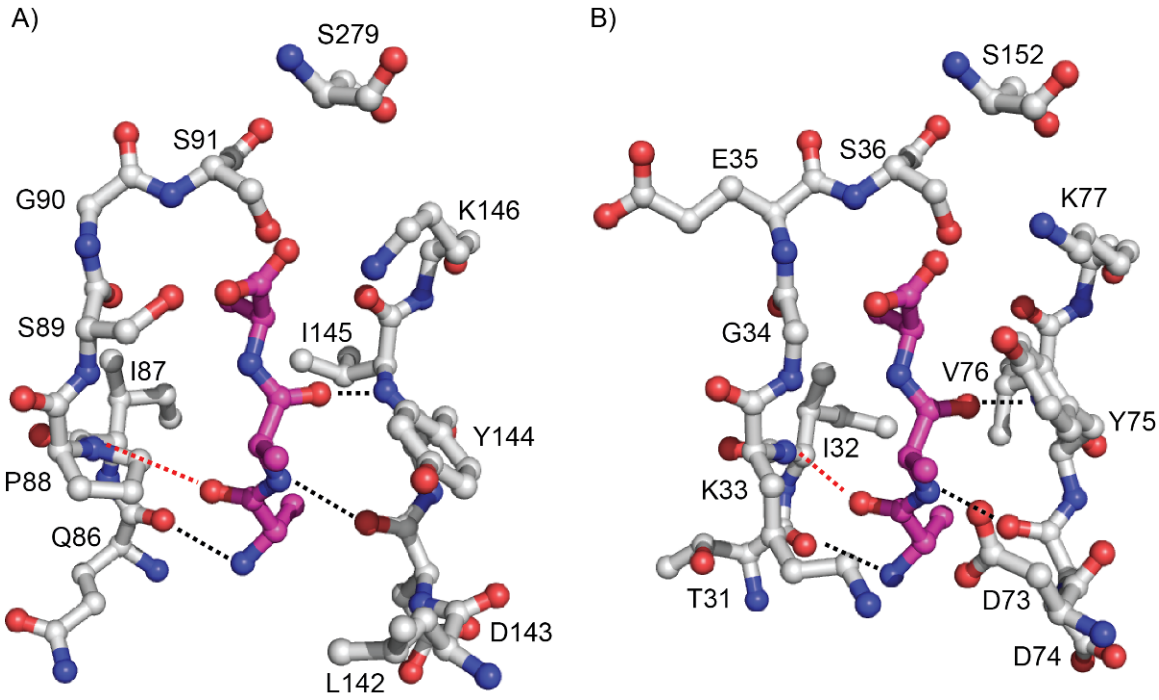


Figure 3.5. E. coli and S. aureus SPase I enzymes modeled with a triple alanine peptide.

A) *E. coli* LepB with a manually docked triple alanine peptide. Nitrogen atoms are blue, oxygen atoms are red, carbon atoms (*E. coli* LepB) are white and carbon atoms (peptide) are magenta. Hydrogen bonds are shown as black dotted lines. The red dotted line is the potential hydrogen bond in the P88A mutant. B) *S. aureus* SpsB with a modeled triple alanine peptide. Same color scheme as panel A. The red dotted line is the potential hydrogen bond missing in the K33P mutant.

4. Engineering of *E. coli* Ecotin with an internal signal peptide C-region as a substrate / competitive inhibitor for *S. aureus* SpsB

4.1. Overview

The goal described in this chapter was to create a protein-based substrate, other than pre-proteins, for *S. aureus* SpsB which can be used as a crystallization additive and as a tool for optimizing the amino acid sequence of future peptide-based substrates for *S. aureus* SpsB. *E. coli* Ecotin was chosen for this task because it is a well characterized protein-based serine protease inhibitor and is an endogenous protein found in the *E. coli* BL21(DE3) strain used for protein overexpression. Ecotin forms a homodimer and inhibits its targets--trypsin, chymotrypsin and other trypsin-fold serine proteases --by inserting the 80s loop into the enzyme's active site like a substrate (McGrath et al. 1994, Perona et al. 1997). Methionine 84 in Ecotin mimics the P1 residue of a substrate (Figure 4.1A). Several modifications have been made to Ecotin to turn it into an *S. aureus* SpsB substrate, SP-Ecotin (Figure 4.1B). Part of the 80s loop has been replaced with a partial signal peptide with a sequence of NGEVAKAAET (described in Section 2.1) and the disulfide bond between cysteine 50 and cysteine 87 has been removed to increase the flexibility of the 80s loop. Activity assays have confirmed that *S. aureus* SpsB cleaves SP-Ecotin while *E. coli* LepB does not cleave SP-Ecotin. SP-Ecotin was turned into a competitive inhibitor by mutating the alanine at the P1' position to a proline. It has been shown in a previous study, when a proline is introduced at the +1 position of maltose binding protein, the protein effectively became a competitive inhibitor of *E. coli* SPase I (Nilsson & von Heijn, 1992). SP-Ecotin +1 proline effectively inhibits *S. aureus* SpsB but not *E. coli* LepB. In summary, we have demonstrated that Ecotin can be turned into a protein-based substrate by replacing the 80s loop with a specific peptide substrate sequence.

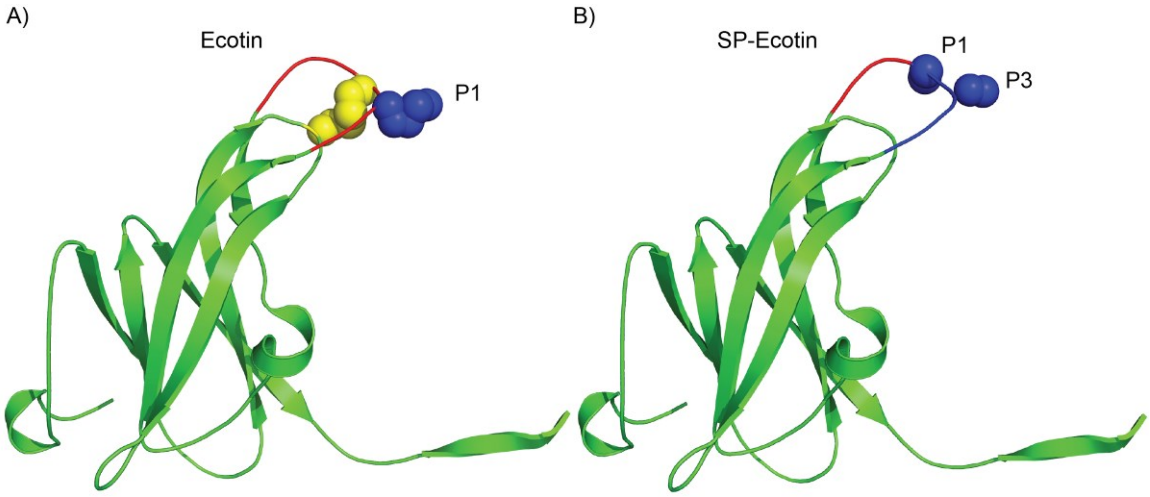


Figure 4.1. E. coli Ecotin and SP-Ecotin 3-dimensional structure.

A) *E. coli* Ecotin in cartoon representation (PDB: 1ECY) (Shin et al. 1996). The 80s loop containing the P1 residue, which inserts into the active site of trypsin, is in red. The P1 residue (M84) is shown as spheres in blue. The disulphide bond between C50 and C87 is shown as yellow spheres. B) SP-Ecotin in cartoon representation. Part of the 80s loop replaced by the peptide substrate, NGEVAKAAET, is in blue. The P1 and the P3 alanines which insert into SPase I's substrate binding pockets are shown as blue spheres.

4.2. Materials and methods

4.2.1. Cloning and mutagenesis

Primers used, and their descriptions, are summarized in Table 4.1. Briefly, the gene encoding *E. coli* Ecotin without its endogenous stop codon was amplified from *E. coli* K12 genomic DNA and ligated into the pET21a(+) vector (Novagen) to introduce a hexahistidine tag to the C-terminus of Ecotin. Ecotin was used as the template to construct SP-Ecotin and SP-Ecotin +1 proline. SP-Ecotin was constructed by deleting residues 81 to 87 and inserting KAAET (P2 to P3') from the peptide substrate into the 80s loop. The PCR products were phosphorylated and blunt end ligated for use as a template for a second insertion. The second insertion introduced NGEVA (P7 to P3) to complete the 10 residues long peptide substrate (Figure 4.2). Cysteine 50 was mutated to a serine to complete the construction of SP-Ecotin while the alanine at the P1' position was mutated to a proline to create the SP-Ecotin +1 proline inhibitor. DNA sequences for all constructs were verified by DNA sequencing (Genewiz Inc.) and no undesired mutations were introduced.

Table 4.1. Primers used to construct Ecotin, SP-Ecotin and SP-Ecotin +1 proline inhibitor

Primer pairs (forward and reverse)	Description
5'-GCCATATGAAGACCATTCTACCTGCAGTATTGTTTGCCGCTTTC-3' 5'-GCCTCGAGGCGAACTACCGCGTTGTCAATTTTCTCTTCCG-3'	Gene specific primer pair for Ecotin lacking a stop codon. Forward and reverse primers carry a NdeI and XhoI restriction sites for subcloning, respectively.
5'-AAAGCAGCAGAAACCGATGGCAAGAAAGAGAAGAAATTT-3' 5'-AACCGGGGAACTGACTTTATCAAA-3'	Deletion of residue 81 to 87 and insertion of P2 to P3' residues of the peptide substrate in the 80s loop of Ecotin.
5'-AACGGGGAAGTAGCAAAAGCAGCAGAAACCGATGGCAAG-3' Reverse primer: same as above	Insertion of P7 to P3 residues of the peptide substrate in the 80s loop of Ecotin.
5'-ACGCTGGAAGTCGATTCCAATTTGCATCGTCTC-3' 5'-GAGACGATGCAAATTGGAATCGACTTCCAGCGT-3'	Mutated cysteine 50 to a serine to remove intermolecular disulfide bond between two SP-Ecotin molecules.
5'-GAAGTAGCAAAAGCAGCAGAAACCGATGGCAAG-3' 5'-CTTGCCATCGGTTTCTGCTGCTTTTGCTACTTC-3'	Mutated the alanine at P1' to a proline in the inserted signal peptide to turn SP-Ecotin into a competitive inhibitor

Restriction sites are underlined.

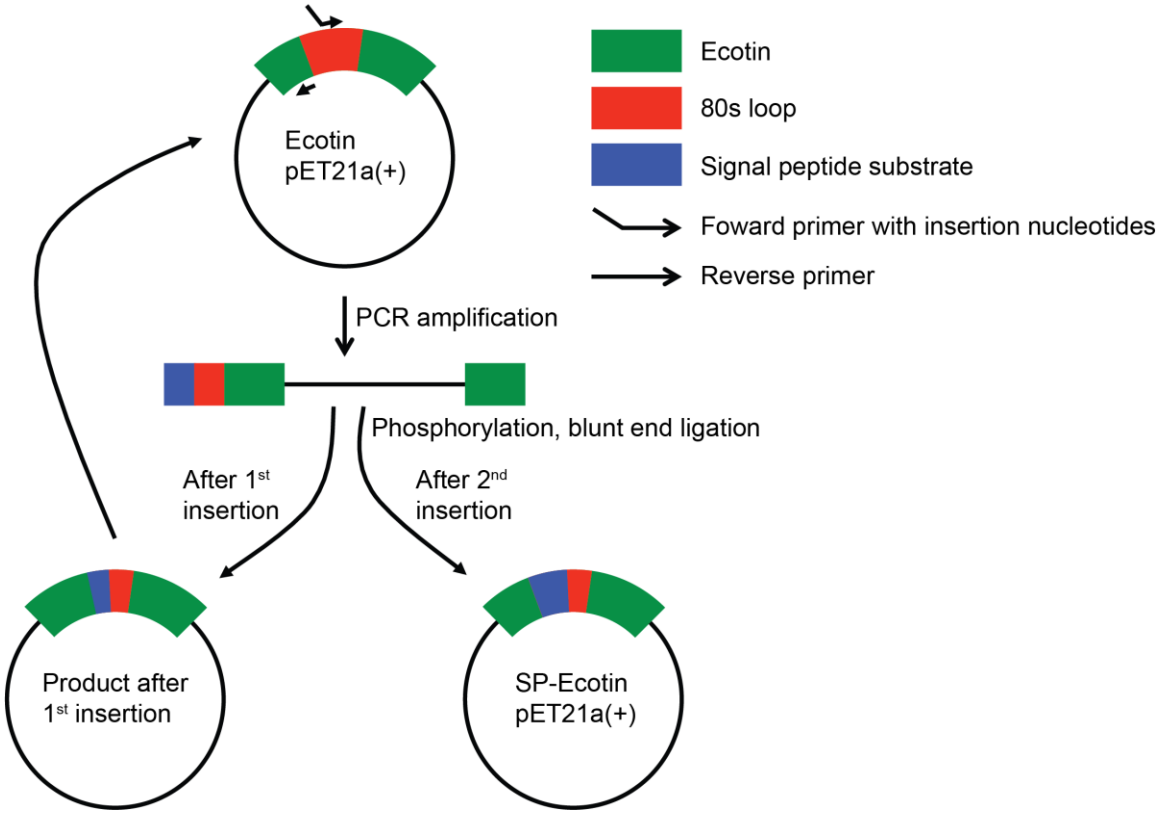


Figure 4.2. Schematic of signal peptide substrate insertion into Ecotin.

SP-Ecotin was constructed using *E. coli* Ecotin in pET21a(+) vector as the template. The complete signal peptide substrate was inserted through two rounds of insertion steps. Each insertion step consisted of a PCR amplification stage using a forward primer carrying nucleotides to introduce the signal peptide substrate followed by phosphorylation and blunt end ligation of the PCR products to generate circular plasmid DNA.

4.2.2. Ecotin and SP-Ecotin overexpression and purification

Plasmids containing Ecotin and Ecotin variants were transformed into *E. coli* BL21(DE3). One liter of LB media supplemented with ampicillin (100µg/mL) was inoculated with 50mL of overnight culture. The cell culture was grown at 37°C with shaking (250rpm) until $O.D._{600}$ reached 0.6. At this point, protein expression was induced with the addition of IPTG (1mM final concentration) and the culture was grown for an additional four hours. Cells were pelleted by centrifugation (6000xg, 5min, 4°C) and cell pellets were stored at -80°C until needed. Cell pellets were resuspended in buffer A (20mM Tris pH8.0, 100mM NaCl, 0.01% DDM) and sonicated with a 15sec pulse and 30sec rest cycle three times using a Sonic Dismembrator Model 500 (Fisher Scientific). The resuspended cells were lysed by passing through an EmsulfiFlex-C3 cell disruptor

(Avestin) for 5min. Cell lysate was clarified by centrifugation (35,000xg, 30min, 4°C) and the supernatant was collected for Ni^{2+}-NTA affinity chromatography. The supernatant was passed through a column with 3mL of Ni^{2+}-NTA resin (Qiagen) equilibrated with buffer A. The column was washed twice with 10 column volume (CV) of buffer A containing 10mM and 20mM imidazole. Recombinant proteins were eluted with 2CV of buffer A containing increasing concentrations of imidazole (100, 250, 500 and 1000mM). Each elution fraction was run on a 15% SDS-PAGE gel, then stained with PageBlue (Thermo Fisher) protein staining solution. Fractions containing Ecotin were concentrated using an Amicon ultracentrifugal filter (Millipore) with a 3-kDa molecular mass cut off. Protein concentration was determined by BCA assay (Thermo Fisher) following the manufacturer's protocol.

4.2.3. SP-Ecotin digestion assays using Gram-positive and Gram-negative SPase I enzymes

SP-Ecotin was digested by *S. aureus* SpsB, *S. epidermidis* SpsB, *B. subtilis* SipS, *B. subtilis* SipT and *E. coli* LepB at a 5:1 molar ratio. The reactions were carried out in buffer A and incubated at 37°C overnight. Each sample was run on a 15% SDS-PAGE gel and stained with PageBlue (Thermo Fisher) protein staining solution.

4.2.4. SP-Ecotin +1 proline inhibition assays using Gram-positive and Gram-negative SPase I enzymes

5μM of dodecanoyl-K(Dabcyl)NGEVAKAAE(EDANS)T-NH_2 was digested by 10nM of *S. aureus* SpsB, *S. epidermidis* SpsB, *B. subtilis* SipS, *B. subtilis* SipT or 500nM *E. coli* LepB, in the presence of 10μM SP-Ecotin +1 proline inhibitor in buffer A. The reaction mixture was set up in a 96 well microplate (Greiner Bio-One, Cat No. 675096) and the reaction progress was monitored by a SpectramMax M5 (Molecular Devices) with an excitation wavelength of 340nm and an emission wavelength of 530nm at 23°C.

4.2.5. Amino terminal sequencing of S. aureus SpsB cleaved SP-Ecotin by the Edman degradation method

Cleaved SP-Ecotin fragments generated by *S. aureus* SpsB digestion were prepared for amino terminal sequencing to determine the *S. aureus* SpsB cleavage site. Cleaved SP-Ecotin fragments in a 15% SDS-PAGE gel were transferred onto a PVDF membrane (Millipore) by electroblotting and stained with Coomassie Brilliant Blue R-250 (Bio-Rad). The membrane was sent to Iowa State University Protein Facility for amino terminal sequencing. Amino terminal sequencing was done using the automated Edman degradation method on a 494 Procise Protein Sequencer/140C Analyzer (Applied Biosystems). The first six residues from the new amino terminus were identified.

4.3. Results

4.3.1. Ecotin and SP-Ecotin purification

Ecotin, SP-Ecotin and SP-Ecotin +1 proline have been successfully purified using the same purification protocol. SP-Ecotin is a secreted protein with an N-terminal signal peptide. The hexahistidine affinity purification tag is attached to the C-terminus of the protein so the secreted SP-Ecotin still retains the tag. Pre-SP-Ecotin has a theoretical molecular mass of 19.51 kDa while SP-Ecotin has a theoretical molecular mass of 17.41 kDa. No significant amount of pre-SP-Ecotin was found in the clarified cell lysate, presumably due to the presence of intact signal peptides which meant they were anchored to the cell membrane and as a result, partitioned into the insoluble fraction. The elution fractions from the Ni^{2+}-NTA affinity column only contained matured SP-Ecotin (Figure 4.3).

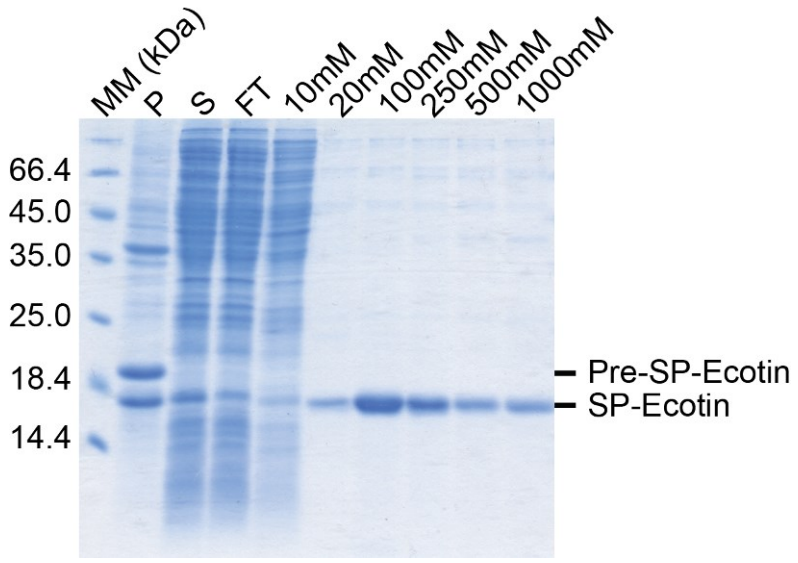


Figure 4.3. SP-Ecotin Ni^{2+}-NTA affinity purification.

Samples at each purification step were analyzed on a 15% SDS-PAGE gel stained with PageBlue protein staining solution. The molecular mass (MM) ladder in kDa is in the left lane, followed by the insoluble pellet (P), the clarified supernatant (S), and the column flow through (FT) to the right. The column was washed twice with buffer A containing 10mM and 20mM imidazole. SP-Ecotin was eluted in buffer A containing 100, 250, 500 and 1000mM imidazole.

4.3.2. SP-Ecotin is cleaved by S. aureus SpsB

SP-Ecotin was treated with various SPase I enzymes to determine whether or not SP-Ecotin is a viable substrate for SPase I. SP-Ecotin was incubated with *S. aureus* SpsB, *S. epidermidis* SpsB, *B. subtilis* SipS, *B. subtilis* SipT and *E coli* LepB. Only *S. aureus* SpsB was able to cleave SP-Ecotin efficiently and produced two cleaved SP-Ecotin fragments while *S. epidermidis* SpsB, *B. subtilis* SipS, and SipT showed little cleavage. *E. coli* LepB showed no cleavage of the substrate (Figure 4.4A). The two cleaved fragments corresponded to the N-terminal and the C-terminal fragments of SP-Ecotin with a molecular mass of 9.83 kDa and 7.60 kDa, respectively. The *S. aureus* SpsB cleavage site was determined, NGEVAKA/AET (the *S. aureus* SpsB cleavage site is indicated by the slash), by amino-terminal sequencing of the C-terminal fragment. SP-Ecotin can be used directly as a co-crystallization additive for catalytically inactive *S. aureus* SpsB S36A and K77A mutants but not with the wild type *S. aureus* SpsB, as SP-Ecotin is readily degraded by it (Figure 4.4B). A proline was then introduced at the P1' position, NGEVAKA**P**ET (P1' proline is in bold lettering), in order to be able to use SP-Ecotin as a co-crystallization additive for *S. aureus* SpsB because SPase I cannot

process signal peptides with a proline at the P1' position (Barkocy-Gallagher & Bassford 1992, Nilsson, von Heijne 1992). The introduction of the P1' proline into SP-Ecotin has successfully prevented processing by *S. aureus* SpsB, thus making it possible to use as a co-crystallization additive with for active *S. aureus* SpsB (Figure 4.4B).

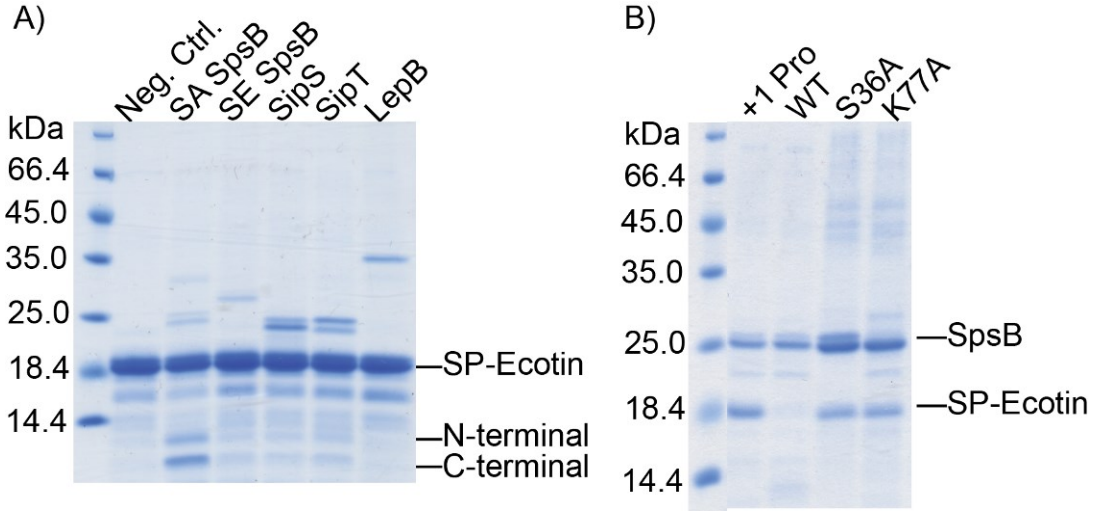


Figure 4.4. SP-Ecotin and SP-Ecotin +1 proline processing by SPase I.

A) SP-Ecotin (30μM) was incubated without (Neg. Ctrl.) or with the addition of one of the following: *S. aureus* SpsB, *S. epidermidis* SpsB, *B. subtilis* SipS, *B. subtilis* SipT or *E. coli* LepB (5μM) in buffer A at 37°C overnight. Samples were run on a 15% SDS-PAGE gel and stained with PageBlue protein staining solution. B) SP-Ecotin +1 proline (+1 Pro, 5μM) was incubated with *S. aureus* SpsB (5μM) (left lane) and SP-Ecotin (5μM) was incubated with *S. aureus* SpsB (WT, 5μM) or one of the two catalytically inactive mutants (S36A and K77A, 5μM) in buffer A at 37°C overnight. Samples were run on a 15% SDS-PAGE gel and stained with PageBlue protein staining solution.

4.3.3. SP-Ecotin +1 proline is a competitive inhibitor for S. aureus SpsB

Activity assays using the dodecanoyl-K(Dabcyl)NGEVAKAAE(EDANS)T-NH_2 peptide substrate in the presence of SP-Ecotin +1 proline were carried out to check the inhibition potency of SP-Ecotin +1 proline on various SPase I enzymes. Wild type Ecotin has no effect on *S. aureus* SpsB's activity while SP-Ecotin +1 proline inhibits *S. aureus* SpsB's activity (Figure 4.5A). *S. epidermidis* SpsB, *B. subtilis* SipS, *B. subtilis* SipT and *E. coli* LepB are not affected by SP-Ecotin +1 proline (Figure 4.5B). This result is in agreement with the previous section's SP-Ecotin digestion assay where *S. aureus* SpsB was the only enzyme capable of efficiently cleaving SP-Ecotin.

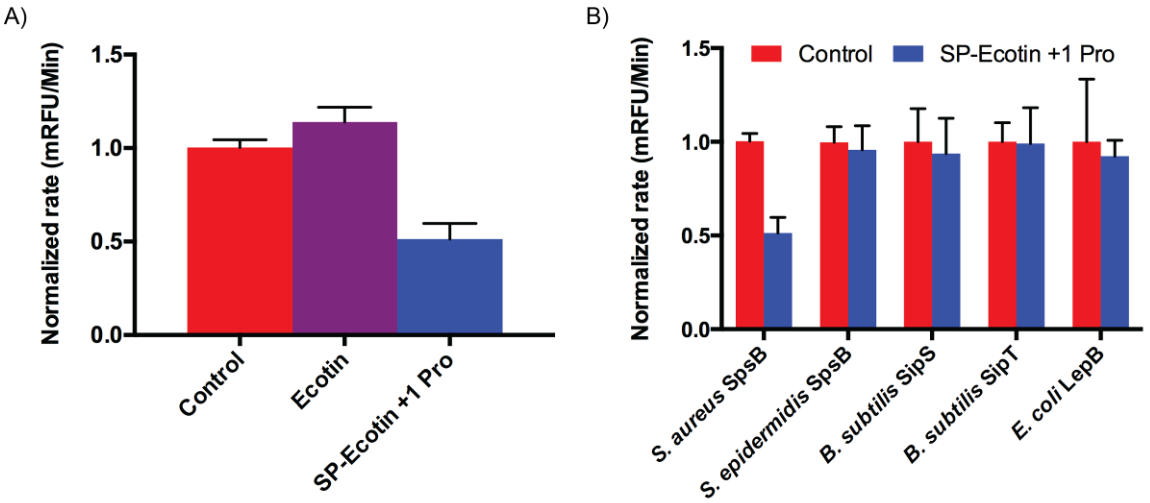


Figure 4.5. SP-Ecotin +1 proline inhibition assays.

A) Dodecanoyl-K(Dabcyl)NGEVAKAAE(EDANS)T-NH_2 (5μM) was incubated with *S. aureus* SpsB (10nM) without (Control) or with 10μM of Ecotin or SP-Ecotin +1 proline in buffer A at 23°C. The normalized rate in *milli*-relative fluorescence units per minute (mRFU/min) for each reaction is shown. Each coloured bar represents the mean of the normalized rate and the error bars are standard errors calculated from triplicate experiments. B) Dodecanoyl-K(Dabcyl)NGEVAKAAE (EDANS)T-NH2 was incubated with 10nM of *S. aureus* SpsB, *S. epidermidis* SpsB, *B. subtilis* SipS, *B. subtilis* SipT or 500nM of *E. coli* LepB without (Control) or with the presence of SP-Ecotin +1 proline (10μM) in buffer A at 23°C. Each coloured bar represents the mean of normalized rate (mRFU/min) and the error bars shown are standard errors calculated from triplicate experiments.

4.4. Discussion

In this chapter, we have demonstrated that *E. coli* Ecotin can be either a substrate or a inhibitor for *S. aureus* SpsB depending on the amino acid sequence in the 80s loop. Interestingly, none of the other four SPase I enzymes tested had any significant enzymatic activity towards SP-Ecotin within the duration of the experiments. One thing to keep in mind is that SP-Ecotin forms homodimers in solution, which may introduce steric hindrance when the 80s loop is being inserted into the target enzyme. From the previous chapter, we have shown that *S. aureus* SpsB has the highest affinity towards the peptide sequence that has been introduced into SP-Ecotin. Further optimizations can be done to SP-Ecotin such as trying other peptide sequences in the 80s loop and creating monomeric SP-Ecotin by removing the dimerization interaction at its C-terminus (Eggers et al. 2001, Pal, Szilagyi & Graf 1996) which may lessen potential steric hindrance and improving binding affinity for different SPase I enzymes.

Other protein-based inhibitors can also be examined to identify potential scaffold proteins for generating SPase I substrates / inhibitors. For example, *Hirudo medicinalis* Eglin C is a serine protease inhibitor only 70 residues in length (8 kDa). Eglin C has a α/β fold with a central β-sheet flanked on one side by an α-helix and the other side by the 40s loop containing the P1 residue (leucine 45) for its target enzyme (Figure 4.6) (Hyberts et al. 1992).

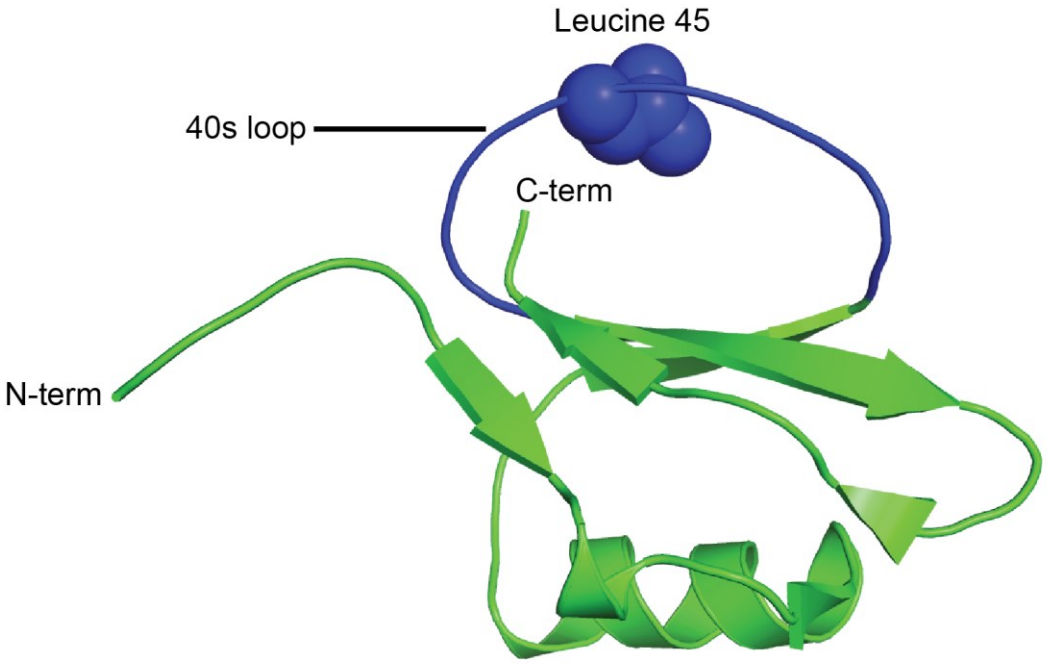


Figure 4.6. Hirudo medicinalis Eglin C structure.

H. medicinalis Eglin C (PDB: 1EGL) in cartoon depiction. The 40s loop containing the P1 residue, leucine 45 with its side chain shown as spheres, is highlighted in blue with the rest of the molecule in green.

Other protein-protein interaction experiments such as isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR) can be carried out to measure thermodynamic binding parameters. A preliminary dissociation constant for the *S. aureus* SpsB and SP-Ecotin +1 proline mutant interaction was determined to be 47.3μM

using a Monolith NT.115 intrument (NanoTemper Technologies) (Figure 4.7). Although the preliminary dissociation curve did not have data points in the saturation range, this approximate K_d determined can still be useful in the futurepreparation of samples for ITC and SPR analyses.

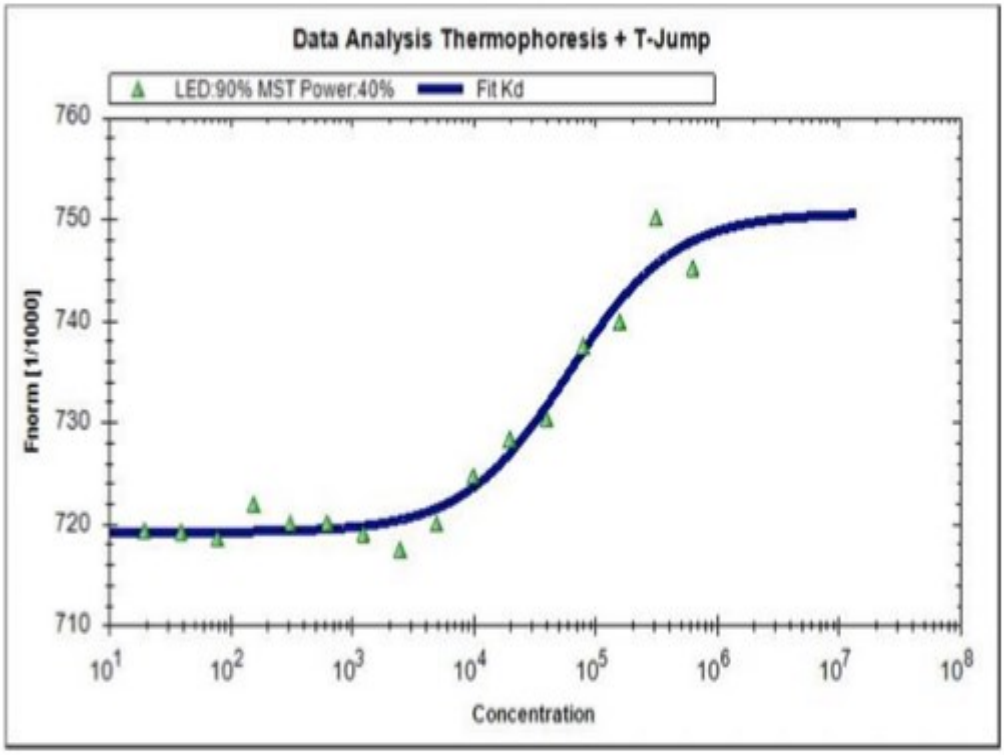


Figure 4.7. Thermophoresis analysis of S. aureus SpsB and E. coli SP-Ecotin +1 proline mutant.

Normalized fluorescence (y-axis) plotted against 15nM to 500μM of *E. coli* SP-Ecotin +1 proline at a fixed concentration of *S. aureus* SpsB on a Monolith NT.115 intrument (NanoTemper Technologies).

5. Biochemical characterization of *Escherichia coli* signal peptide peptidase A2

5.1. Overview

This chapter will describe the designs of several different *E. coli* signal peptide peptidase A2 (SppA2) constructs, as well as their purification, biochemical characterization, kinetic characterization and crystallization. One construct, SppA2 Δ2-37, lacking the predicted single N-terminal transmembrane segment, was found to have a relatively low protein expression level while two SppA2 Δ2-37 mutants, S178A and K230A, were found to have significantly higher expression levels. *E. coli* SppA2 Δ2-37 was digested by trypsin to produce a smaller, trypsin resistant, SppA2 domain (trSppA2) (Figure 5.1) and subsequent amino terminal sequencing and SDS-PAGE analysis of trSppA2 indicated that the N-terminus of SppA is digested by trypsin while little to no digestion occurs at the C-terminus. Systematic truncation from the N-terminus of SppA2 revealed that the N-terminus is required for proper folding of SppA2. Without the protease prone N-terminal region, SppA2 protein expressed as inclusion bodies. After purification, trSppA2 was analyzed using multiangle light scattering (MALS) in conjunction with size exclusion chromatography (SEC) to show that trSppA2 exists as an octamer in solution. Activity assays using both protein-based substrates and peptide-based substrates have demonstrated that trSppA is active while the two mutants, S178A and K230A, are catalytically inactive, adding further support to SppA2 being a Ser/Lys dyad enzyme. Lastly, trSppA K230A was co-crystallized with a dodecanoyl-NGEVAKA-MCA peptide substrate and the resultant crystals diffracted to 3.8Å.

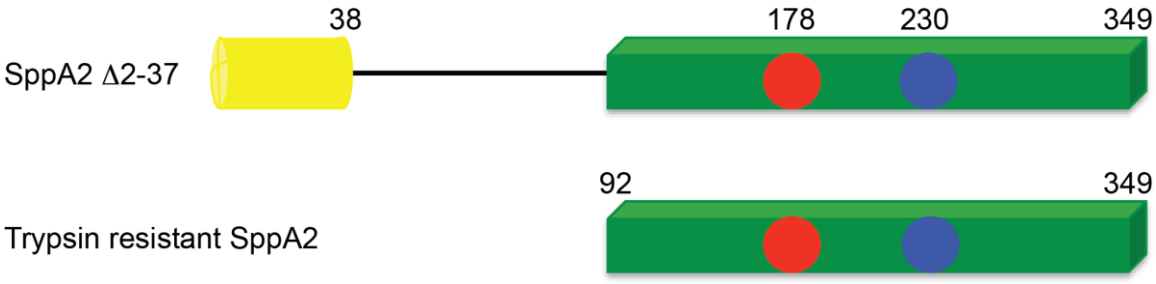


Figure 5.1. Schematic of E. coli SppA2 Δ2-37 and trypsin resistant SppA2.

The soluble domain of SppA2 (residues 38 to 349) was cloned with an N-terminal hexahistidine tag represented by the yellow cylinder. The positions of the proposed serine nucleophile and lysine general base are located at 178 and 230 respectively and represented as red and blue circles. The black line represents the protease prone region and the green box represents the trypsin resistant domain. The trypsin resistant domain (residues 92 to 349) is shown with an intact C-terminus.

5.2. Materials and methods

5.2.1. Cloning and mutagenesis

The *E. coli sppA2* gene comprising the cytosolic domain but lacking residues 2 to 37 (SppA2 Δ2-37) was PCR amplified from *E. coli* strain K12 genomic DNA with a gene specific primer pair and subcloned into the pET28a(+) vector (Novagen) to add an N-terminal hexahistidine tag for Ni^{2+}-NTA affinity purification. SppA2 Δ2-37 was used as a template for further SppA2 constructs. All constructs, their descriptions and primers used, are summarized in Table 5.1. All constructs' DNA sequences were confirmed by DNA sequencing (Genewiz) and the expressed protein sequences matched the SppA2 protein sequence (accession number P0AG14) deposited in the UniProt database.

Table 5.1. Primers used to generated various SppA2 constructs.

Construct name	Description	Primers (forward/reverse)
SppA2 Δ2-37	Catalytically active SppA2 soluble domain	5'-GCCATATGGAAAACCTGTATTTTCAGGGACAGCGTG GCGAGTTACGG-3' 5'-GCCTCGAGTTACATCAATGGCTTATGACCC-3'
SppA2 Δ2-37 S178A	Catalytically inactive SppA2 soluble domain with the serine nucleophile mutated to an alanine	5'-AAAGTCGCTGCCGCCGGCGGTACATGT-3' 5'-ACATGTACCGCCCCGGGCAGCGACTTT-3'

SppA2 Δ2-37 K230A	Catalytically inactive SppA2 soluble domain with the lysine general base mutated to an alanine	5'-GCCGGGCAGTATGCGCGCACGCTGACG-3' 5'-CGTCAGCGTGCGCGCATACTGCCCGGC-3'
SppA2 Δ2-50 K230A	Catalytically inactive SppA2 lacking the first 50 residues	5'-GCCATATGTATAAGGAGATGAAAGAAGAACTGGCC-3' 5'-GCCTCGAGTTACATCAATGGCTTATGACCC-3'
SppA2 Δ2-60 K230A	Catalytically inactive SppA2 lacking the first 60 residues	5'-GCCATATGGCGCTGATGGACTCACATCAG-3' 5'-GCCTCGAGTTACATCAATGGCTTATGACCC-3'
SppA2 Δ2-70 K230A	Catalytically inactive SppA2 lacking the first 70 residues	5'-GCCATATGTGGCACAAAGCGCAGAAGAA-3' 5'-GCCTCGAGTTACATCAATGGCTTATGACCC-3'
SppA2 Δ2-80 K230A	Catalytically inactive SppA2 lacking the first 80 residues	5'-GCCATATGCAAGAAGCGAAAGCAGCAAAAG-3' 5'-GCCTCGAGTTACATCAATGGCTTATGACCC-3'
SppA2 Δ2-91 K230A	Catalytically inactive SppA2 trypsin resistant domain	5'-GCCATATGCTGGGCGAGGTGGCAACT-3' 5'-GCCTCGAGTTACATCAATGGCTTATGACCC-3'

Restriction enzyme sites NdeI and XhoI in the forward and the reverse primers, respectively, are underlined.

5.2.2. E. coli SppA2 overexpression and purification

Plasmids encoding the various SppA2 constructs were transformed into *E. coli* BL21(DE3) for recombinant protein expression. One liter of LB media with kanamycin (25μg/mL) was inoculated with 50mL of overnight culture. The cells were allowed to grow with shaking (250rpm) until mid log phase ($O.D._{600}$ = 0.6) at 37°C. At this point, the cells were induced with isopropyl-β-D-1-thiogalatopyranoside (IPTG, 1mM final concentration) and grown for an additional 4 hours. The cells were pelleted by centrifugation (6000xg, 5min, 4°C) and cell pellets were stored at -80 °C until needed. 5g cell pellet were thawed and resuspended in 25mL of buffer A (20mM Tris pH8.0, 100mM NaCl, 0.01% n-dodecyl-β-D-maltoside). The cells were lysed using a Sonic Dismembrator Model 500 (Fisher Scientific) with a 15sec pulse and 30sec rest cycle, three times at 30% amplitude on ice, followed by passage through an EmulsiFlex-C3 cell disruptor (Avestin) for 5min. The cell lysate was clarified by centrifugation (35,000xg, 45minutes, 4°C) and the clarified cell lysate containing SppA2 was collected for subsequent purification steps. The clarified cell lysate was loaded onto a 3mL Ni^{2+}-NTA (Qiagen) column equilibrated with buffer A. The column was washed twice with 10

column volumes (CV) of buffer A containing first 10mM, and then 25mM imidazole. SppA2 was eluted stepwise with 2CV of buffer A containing increasing concentrations of imidazole (100, 200, 300, 400, 500 and 1000mM). Elution fractions were fun on a 15% SDS-PAGE gel and stained with PageBlue (Thermo Fisher) protein staining solution. Elution Fractions containing SppA2 were pooled together and concentrated to 1mL for size exclusion chromatography using an Amicon ultracentrifugal filter (Millipore) with a 10kDa molecular weight cut off (MWCO). 1mL of SppA2, from the Ni2+-NTA affinity purification step, was loaded onto a Superdex 200 300/10 GL column (Amersham Biosciences) equilibrated with buffer A, connected to an ÄKTA FPLC system (Amersham Biosciences). The column flow rate was set at 0.5mL/min with 1mL fractions. Peak fractions were analyzed on a 15% SDS-PAGE gel and fractions containing purified SppA2 were pooled together and concentrated for subsequent biochemical characterization. Protein concentration was determined by BCA assay (Thermo Fisher) following the manufacturer's protocol.

5.2.3. Trypsin resistant E. coli domain purification

The trypsin resistant SppA2 domain (trSppA2) protein was prepared from *E. coli* SppA2 Δ2-37. During the Ni^{2+}-NTA affinity purification of SppA2 Δ2-37, 10mL of buffer A containing 500µg of trypsin (Sigma catalogue # T1426) was added to the column after the second wash step. The column was then incubated at 37°C on a rocker overnight. The next day, trSppA was collected and the column was washed twice, each time with 10mL of buffer A. These washes were pooled together with the trSppA elution fraction and concentrated to 1mL using an Amicon ultricentrifugal filter (Millipore) with a 100kDa MWCO to remove the trypsin (25.79 kDa). The concentrated trSppA2 was subjected to a SEC purification step as described above in Section 4.2.2 to remove any residual trypsin.

5.2.4. Multiangle light scattering and analytical size exclusion chromatography

100µL of *E. coli* trSppA2 (1mg/mL) was loaded onto a Superdex 200 size exclusion column equilibrated with 20mM Tris pH8.0, 100mM NaCl connected in series with a Dawn 18-angle light-scattering detector coupled to an Optilab rEX interferometric

refractometer and a quasi-elastic light-scattering instrument (Wyatt Technology). It was run at a flow rate of 0.5 mL/min. The molecular mass of the oligomeric trSppA2 was calculated by the ASTRA software (version 5.1, Wyatt Technology) using the Zimm fit method (ZIMM 1948)with a refractive index increment, *dn*/*dc* = 0.185mL/g.

5.2.5. Pro-OmpA nuclease A, aldolase and insulin β-chain digestion assay

Pro-OmpA Nuclease A (PONA), a pre-protein, was treated with *E. coli* SppA2 Δ2-37 protein at a 100:1 (w/w) ratio in buffer A at 37°C. Samples were taken at 0, 5, 10, 15, 30, 60, 120 and 180 minutes. Aldolase (GE Healthcare catalogue # 28-4038-42) was treated with SppA2 Δ2-37, SppA2 Δ2-37 S178A, SppA2 K230A, trSppA2, trSppA2 S178A and trSppA K230A proteins at a 100:1 (w/w) ratio in buffer A at 37°C. Samples were taken at 0, 1 and 16 hours. Insulin β-chain (Sigma catalogue # 30003-72-6) was treated with SppA2 Δ2-37 and trSppA2 proteins at a 100:1 (w/w) ratio in buffer A at 37°C and samples were taken at 0, 1 and 16 hours. All digestion samples were run on a 15% SDS-PAGE gel and stained with PageBlue (Thermo Fisher) protein staining solution.

5.2.6. Activity assay using a lipidated peptide MCA substrate

The peptide, dodecanoyl-NGEVAKA-MCA, was custom synthesized (CanPeptide) and prepared in dimethyl sulfoxide (DMSO) to make a 10mM stock solution. The stock solution was then used to generate different concentrations of peptide by diluting with DMSO for use in kinetic characterization experiments of *E. coli* trSppA2. The reaction mixture contained 100nM trSppA2 in 99µL of buffer A and the reaction was initiated by adding 1µL of the substrate at various concentrations (2.5, 5, 10, 20, 40, 60 and 80µM). The reaction was set up in 96 well microplates (Greiner Bio-One, Cat-No. 675096) and the reaction progress was monitored on a SpectraMax M5 (Molecular Devices) with an excitation wavelength (Ex) of 380nm and an emission wavelength (Em) of 460nm at 23°C. Kinetic constants, k_{cat} and K_M, were determined by fitting data points, collected in triplicates, to the Michaelis−Menten equation using the Prism5 software (GraphPad).

5.2.7. *E. coli trSppA2 K230A protein crystallization*

E. coli trSppA2 K230A (10mg/mL) was incubated with the dodecanoyl-K(Dabcyl)NGEVAKAAE(EDANS)T-NH_2 peptide substrate at a 1:1 molar ratio on ice for two hours before being used in crystallization experiments. The enzyme/substrate complex was crystallized using the hanging drop vapour diffusion method. 1µL of protein mixture was mixed with 1µL of reservoir solution (17.5% w/v PEG3350, 0.15M sodium malonate) and sealed in a well with 1mL of reservoir solution at room temperature. Crystals were observed about 16 hours after the crystallization experiment was set up.

5.2.8. *E. coli SppA2 homology structure model construction*

The full length SppA2 sequence was used as the query sequence for structure model construction using the computer program Protein Homology/analogy Recognition Engine (Phyre) (Kelley, Sternberg 2009). Phyre has returned *E. coli* SppA as the top structural homologue (PDB ID: 3BEZ) (Kim, Oliver & Paetzel 2008) with 100% confidence and 22% sequence identity. A second analysis was done using trSppA2 (residues 92 to 349) which lacks the protease sensitive N-terminal region. *E. coli* SppA was returned as the top structural homologue with 100% confidence and 23% sequence identity. The structure of the region covering residues 92 to 349 for both models are virtually identical so the trSppA2 model was used for analysis without further alteration to the model. The trSppA2 octamer model was constructed by superposing of eight trSppA2 monomers onto the *E. coli* SppA tetramer model using the Align function in the PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC.

5.2.9. *Amino terminal sequencing by Edman degradation*

trSppA2 generated from digesting *E. coli* SppA2 Δ2-37 protein with trypsin and SppA2 Δ2-37 digested PONA were prepared for amino terminal sequencing. trSppA2 protein and digested PONA in a 15% SDS-PAGE gel were transferred onto a PVDF membrane (Millipore) by electroblotting and stained with Coomassie Brilliant Blue R-250 (Bio-Rad). The membrane was then sent to the Iowa State University Protein Facility for amino terminal sequencing. Amino terminal sequencing was done using an automated Edman degradation method on a 494 Procise Protein Sequencer/140C Analyzer

(Applied Biosystems). The first six residues from the new amino terminus were identified.

5.3. Results

5.3.1. E. coli SppA2 Δ2-37 purification

SppA2 Δ2-37 protein, with a theoretical molecular mass of 38.56kDa, was successfully isolated using Ni^{2+}-NTA affinity chromatography (Figure 5.2A). SppA2 Δ2-37 showed signs of self degradation as indicated by the smeary bands as opposed to the sharp single bands of the S178A and K230A mutants seen on the SDS-PAGE gel (Figure 5.2B). This result supports the hypothesis that SppA2 is likely a Ser/Lys dyad enzyme with the serine nucleophile and the lysine general base at positions 178 and 230, respectively. SppA2 Δ2-37 K230A was subjected to a size exclusion chromatography analysis. The expected peak elution volume for monomeric SppA2 Δ2-37 K230A is 15.9mL but instead, a peak with a peak elution volume of 9.35mL was observed (Figure 5.3). This elution volume corresponds to a protein 1076kDa in size, which is not within the resolution range of 10kDa to 600kDa for the Superdex 200 column (See Appendix A for the Superdex 200 size exclusion column calibration).

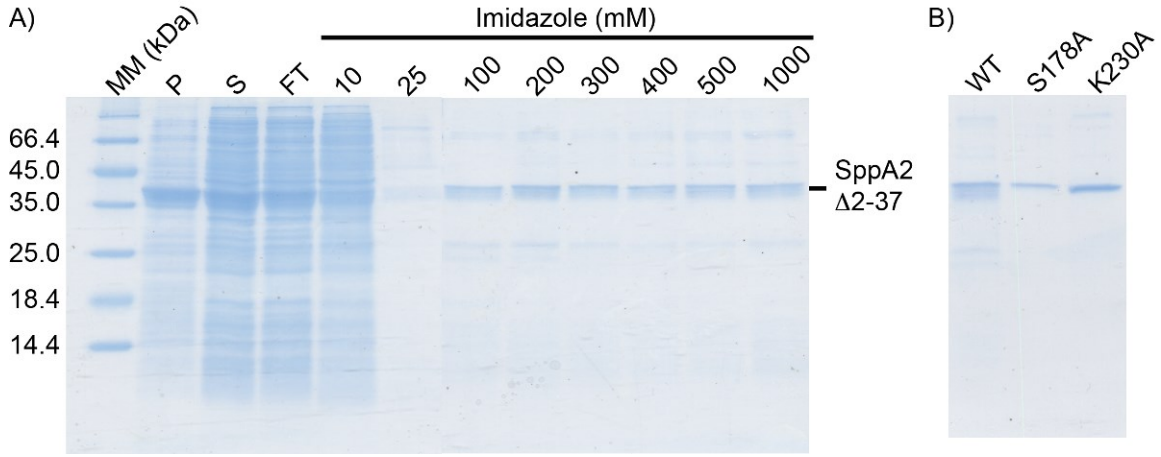


Figure 5.2. E. coli SppA2 Δ2-37 Ni^{2+}-NTA affinity purification.

The SppA2 Δ2-37 protein purification 15% SDS-PAGE gel was stained with PageBlue protein staining solution. A) The molecular mass ladder (MM), in kDa, is in the left lane, followed by pellet (P), supernatant (S), and column flow through (FT) lanes to the right. The column was washed with buffer A containing first 10mM, and then 25mM imidazole. SppA Δ2-37 was eluted stepwise in buffer A containing increasing concentrations of imidazole (100, 200, 300, 400, 500 and 1000mM). B) Purified SppA2 Δ2-37 wild type (WT) and catalytically inactive mutants (S178A and K230A) on a 15% SDS-PAGE gel stained with PageBlue protein staining solution.

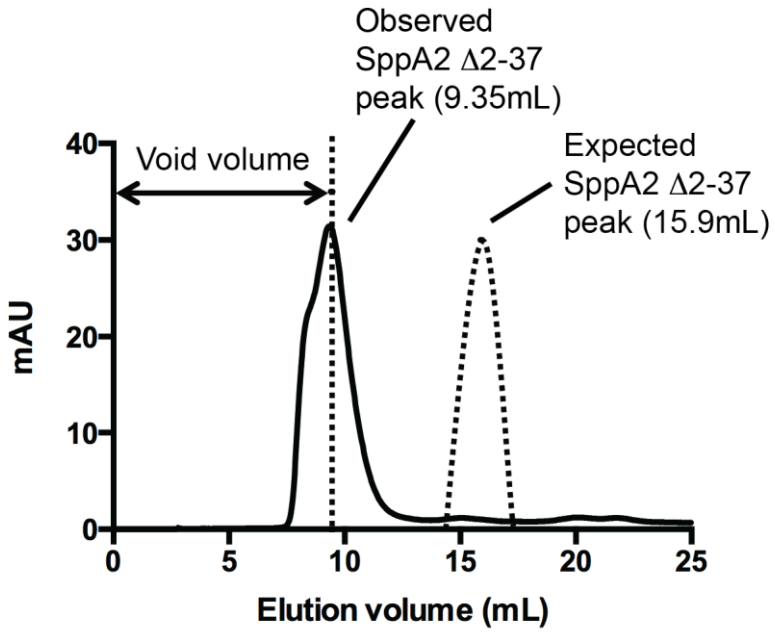


Figure 5.3. E. coli SppA2 Δ2-37 K230A Superdex 200 SEC elution profile.

The elution profile of SppA2 Δ2-37 K230A with the elution volume (mL) on the X-axis and *milli-*absorbance unit (mAU) at 280nm on the Y-axis. 100µL of 1mg/mL purified SppA2 Δ2-37 K230A protein was loaded onto a Superdex 200 column using buffer A as the mobile phase. SppA2 Δ2-37 K230A eluted from the column with a peak volume of 9.35mL, indicating SppA2 Δ2-37 forms a complex in solution. The dotted line peak at 15.9mL is the expected peak for monomeric SppA2 Δ2-37. The void volume (9.5mL) is separated from the elution volume by a vertical dotted line.

5.3.2. E. coli SppA2 has a protease resistant domain and a protease sensitive N-terminal region

SppA2 Δ2-37 K230A protein was treated with various proteases to remove any flexible regions and to generate the smallest possible *E. coli* SppA2 domain, in order to make it more amenable to crystallization. Four readily available proteases, chymotrypsin, trypsin, subtilisin and thermolysin, were used to digest SppA2 Δ2-37 K230A (Figure 5.4). All four proteases produced smaller protease resistant SppA2 fragments. SppA2 Δ2-37 K230A was most resistant to thermolysin as indicated by the presence of residual SppA2 Δ2-37 K230A after an overnight incubation. The resultant protease resistant fragment was larger than the fragments generated by the other three proteases. Subtilisin treatment resulted in a doublet band (representing a heterogeneous protein population) which is not desirable for protein crystallization experiments. Chymotrypsin and trypsin treatments generated SppA2 protease resistant fragments of approximately the same size, with trypsin producing a better final yield than

chymotrypsin. Therefore, trypsin was chosen to be the protease of choice, generating trypsin resistant SppA2 (trSppA2). Based on the relative mobility standard curve generated using the molecular mass ladder in Figure 5.4 (Figure 5.5A), trSppA2 has a relative mobility of 48% and a calculated molecular mass of 28.95kDa. N-terminal sequencing of trSppA2 has confirmed the new N-terminus starts at residue 92, since the preceding residues are rich in lysines which are recognized and cleaved by trypsin (Figure 5.5B). trSppA2, with an assumed intact C-terminus (residues 92-349), has a calculated molecular mass of 29.06kDa. Comparing the molecular mass obtained from both methods we concluded that the removal of the N-terminus is the significant contributor to the change in size observed in trSppA2. The N-terminal region (residues 38-91) was submitted to the PSIPRED secondary structure prediction server (Jones 1999) for analysis. Contrary to the evidence that having random coils usually means that a protein is more protease sensitive, the majority of residues from the removed N-terminal region were predicted to form two α-helices with about one quarter of the residues predicted to form coils (Figure 5.6). These two helices have a theoretical pI of 9.93 (calculated by ProtParam) and a net positive charge at physiological pH with seven negatively charged residues (Asp and Glu) and fourteen positively charged residues (Arg and Lys). These positively charged residues may help in the interaction between the N-terminal region and the negatively charged inner leaflet of the cell membrane. Lipid interaction can be determined through the use of surface plasmon resonance analysis. Briefly, an *E. coli* lipid liposome coated chip can be prepared followed by flowing either *E. coli* SppA2 Δ2-37 or trSppA2 over the chip to measure if there is a difference in interaction depending on the presence or the absence of the protease N-terminal region.

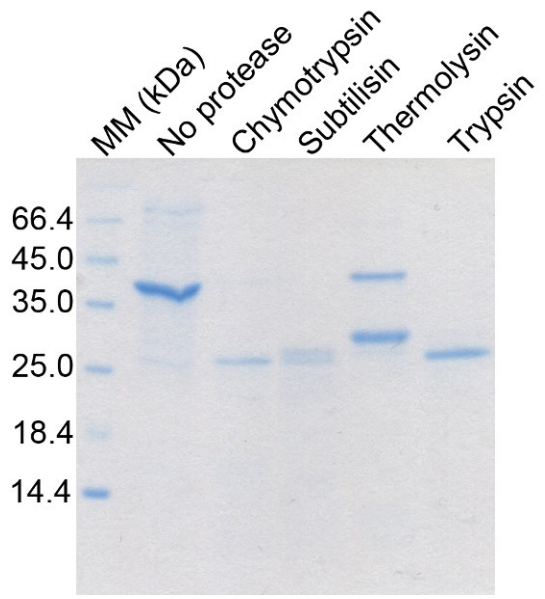


Figure 5.4. E. coli SppA2 Δ2-37 K230A digested by chymotrypsin, subtilisin, thermolysin and trypsin.

5µg of SppA2 Δ2-37 K230A protein was digested by 0.005µg of either chymotrypsin, subtilisin, thermolysin or trypsin in buffer A at 37°C overnight. The negative control was SppA2 Δ2-37 K230A incubated without protease overnight. Each sample was loaded onto a 15% SDS-PAGE gel and stained with PageBlue protein staining solution. The molecular mass ladder in kDa is in the left lane.

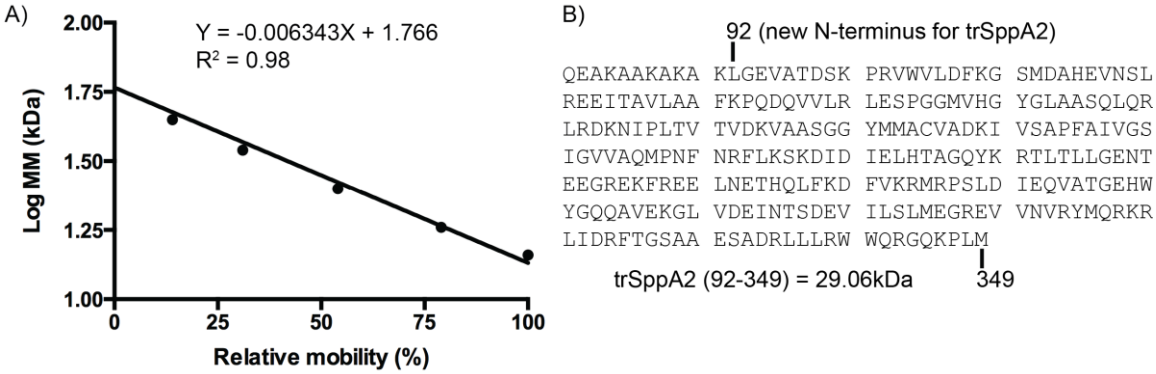


Figure 5.5. E. coli trSppA2 molecular mass estimation by relative mobility method and N-terminal sequencing.

A) The standard curve of relative mobility generated using the molecular mass ladder in Figure 4.4. The equation describing the line of best fit and the correlation are shown on the graph. B) Sequence of trSppA2 with an assumed to be intact C-terminus and its calculated molecular mass in kDa.

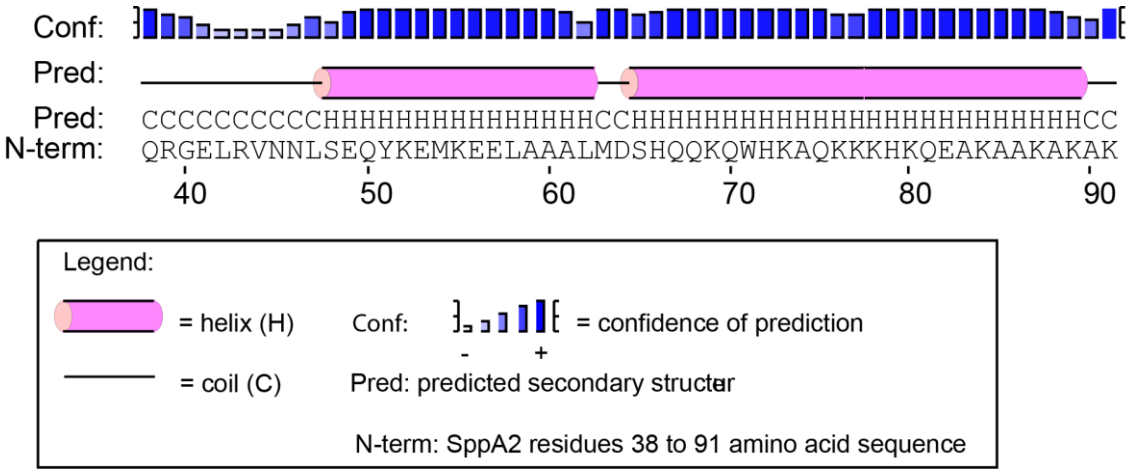


Figure 5.6. Secondary structure prediction of E. coli SppA2 protease sensitive N-terminal region by PSIPRED.

The N-terminal region (residues 38 to 91) of SppA2 removed by trypsin was submitted to the PSIPRED Protein Analysis Workbench server (Buchan et al. 2013). The software used for secondary structure prediction was PSIPRED version 3.3 (Jones 1999).

5.3.3. The protease sensitive N-terminal region is important for protein folding

Systematic truncation from the N-terminus of *E. coli* SppA2 was carried out to determine if trSppA2 can be cloned and overexpressed as an alternative to using SppA2 Δ2-37 as the starting protein followed by trypsin digestion. Six different truncation constructs were tested for their protein overexpression and solubility (Figure 5.7A). Each construct was overexpressed by IPTG induction and prepared as described in Section 5.2.2. The insoluble pellet and the supernatant from each construct were loaded onto a 15% SDS-PAGE gel to analyze the protein distribution of the truncated construct (Figure 5.7B). The original construct, lacking the single predicted N-terminal transmembrane domain (Δ2-37), showed an equal distribution of SppA2 Δ2-37 protein in both the pellet and the supernatant. Any further truncation resulted in the redistribution of SppA2 protein into the pellet, without a significant amount of SppA2 observed in the supernatant. Therefore, for experimental purposes, trSppA2 was produced through trypsin treatment of SppA2 Δ2-37 rather than through direct overexpression from the gene encoding the SppA2 protease resistant domain (Δ2-91). The N-terminal protease sensitive region may act as an intramolecular chaperone, helping the SppA2 protein fold properly but not required for its protease function. *E. coli* SppA2 may have a type I intramolecular chaperone. Type I intramolecular chaperones help in the folding of tertiary

structures and are usually found as an N-terminal extension while type II intramolecular chaperones help in quaternary structure formation and are usually found in near the C-terminus (Chen & Inouye 2008).

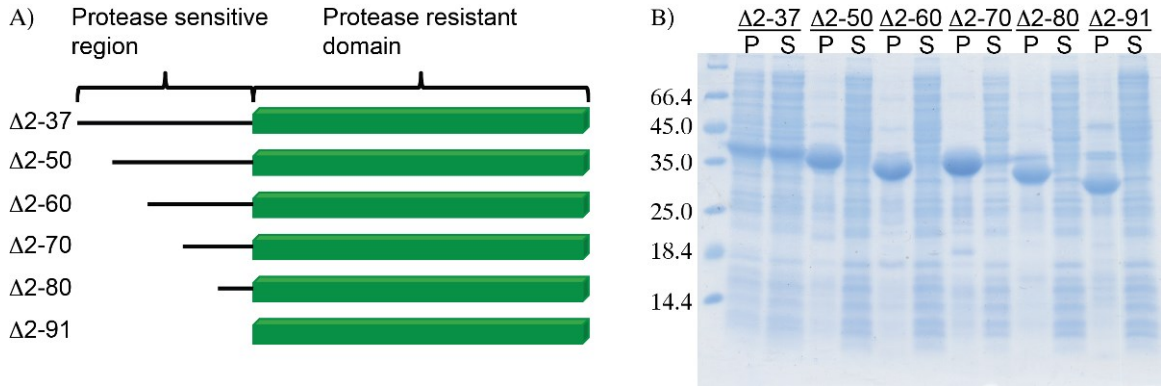


Figure 5.7. E. coli SppA2 K230A N-terminal truncation analysis.

A) Schematic of the systematic truncation of SppA2 K230A constructs. The protease sensitive region removed by trypsin is represented by black lines. The protease resistant domain is represented by green rectangles. The number of residues removed is indicated for each construct. The Δ2-37 construct is the original construct, lacking the predicted N-terminal transmembrane domain. The Δ2-91 construct is the trypsin resistant domain confirmed by N-terminal sequencing. B) The pellet (P) and the supernatant (S) of each truncation construct was loaded onto a 15% SDS-PAGE gel and stained with the PageBlue protein staining solution. The molecular mass ladder, in kDa, is in the left lane.

5.3.4. E. coli trSppA2 purification

Because the protease sensitive N-terminal region is necessary for proper folding of the SppA2 protein, the Ni^{2+}-NTA purification step was modified to produce trSppA2. Instead of eluting with imidazole, SppA2 Δ2-37 bound to the Ni^{2+}-NTA resin was eluted with trypsin, as the N-terminal region of SppA2 Δ2-37 (which is removed by trypsin) contains the hexahistidine purification tag (Figure 5.8A). Trypsin was later removed by filtering the Ni^{2+}-NTA column eluant using a filter with a 100kDa MWCO. The trSppA2 complex was retained while trypsin (23kDa) was removed. Next, the filtered trSppA2 solution was loaded onto a Superdex 200 size exclusion column. trSppA2 has a peak elution volume of 11.9mL which is shifted from the original peak elution volume of 9.35mL for SppA2 Δ2-37 and is well separated from the trypsin peak (17.5mL) (Figure 5.8B) (See Appendix A for the Superdex 200 size exclusion column calibration).

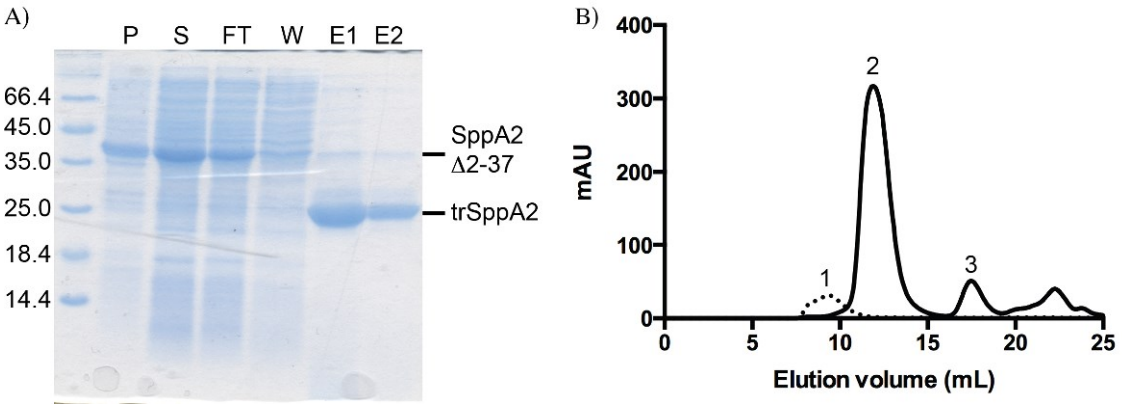


Figure 5.8. E. coli trypsin resistant SppA2 (trSppA2) purification.

A) Samples from the trSppA2 Ni^{2+}-NTA purification process were analyzed on a 15% SDS-PAGE gel stained with PageBlue protein staining solution. The molecular mass ladder, in kDa, is in the left lane, followed by the pellet (P), supernatant (S) and column flow through (FT) to the right. The column was washed (W) once with 10CV of buffer A. trSppA2 was eluted by incubating the column with 2CV of buffer A containing 500µg of trypsin (E1) at 37°C overnight. The column was then washed again with 2CV of buffer A (E2) to collect residual trSppA2. B) Superdex 200 size exclusion profile of SppA2 Δ2-37 (dotted line) overlaid with the elution profile of trSppA2 (solid line). The X-axis is elution volume (mL) and the Y-axis is *milli*-absorbance unit (mAU) at 280nm. Peak 1 (9.35mL) is SppA2 Δ2-37 K230A, peak 2 (11.9mL) is trSppA2 K230A and peak 3 (17.5mL) is trypsin.

5.3.5. E. coli trSppA2 is an octamer in solution

Size exclusion chromatography and multiangle light scattering analysis were used to determine the oligomeric state of trSppA2. trSppA2 was used rather than SppA2 Δ2-37 because trSppA2 has all the flexible regions accessible to trypsin removed. As a result, trSppA2 should yield a better estimate of the true size of trSppA2 in solution. The average molecular mass of trSppA2 was calculated to be 230,000 ± 1,378 g/mol (Figure 5.9) and the mass of the oligomer approximates to eight trSppA2 molecules (each trSppA2 is 29,060 g/mol). This analysis confirmed that trSppA2 exists as an octamer in solution. The oligomeric state of trSppA2 is also in agreement with the oligomeric states of *E. coli* SppA (tetramer) and *B. subtilis* SppA (octamer), as observed from their respective crystal structures (Kim, Oliver & Paetzel 2008, Nam, Kim & Paetzel 2012).

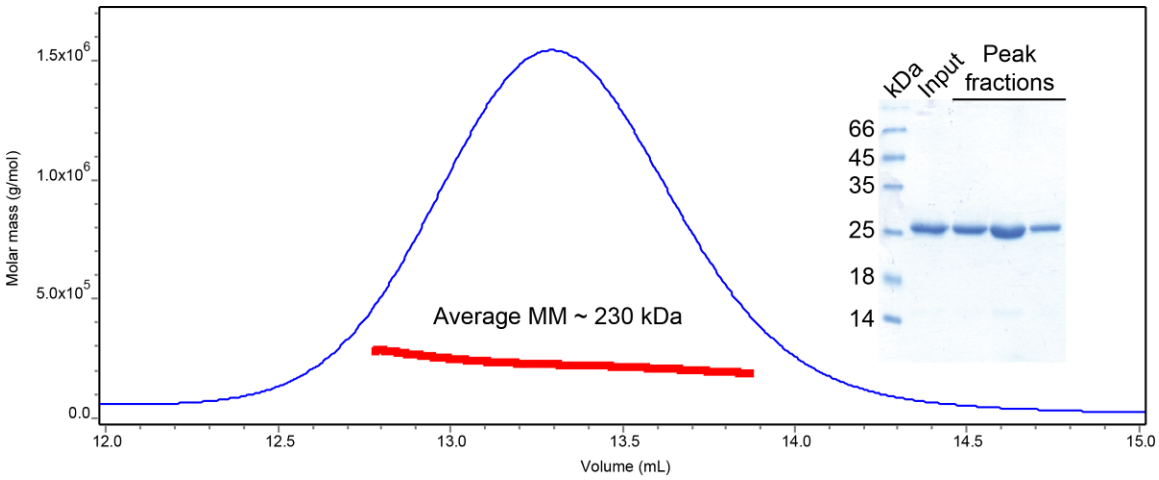


Figure 5.9. Size exclusion chromatography and multiangle light scattering analysis of E. coli trSppA2.

100µL of 1mg/mL trSppA2 K230A was loaded onto a Superdex 200 column connected in series with a Dawn 18-angle light-scattering detector, coupled to an Optilab rEX interferometric refractometer and a quasi-elastic light-scattering instrument (Wyatt Technology). The mobile phased used was 20mM Tris pH8.0 and 100mM NaCl, running at 0.5mL/min with a 0.5mL fraction size. The size exclusion elution profile (blue line) is overlaid with the calculated molecular mass of trSppA2 (red line). trSppA2 under the peak has a average calculated molecular mass of 230kDa. The inset SDS-PAGE gel stained with PageBlue protein staining solution shows the trSppA2 used (Input) and the trSppA2 under the elution peak (12.5mL to 14.0mL)

5.3.6. E. coli SppA2 Δ2-37 cleaves in the hydrophobic region of a signal peptide

A pre-protein substrate previously used to study the activity of *E. coli* SPase I, Pro-OmpA Nuclease A (PONA), was chosen to determine whether or not *E. coli* SppA2 Δ2-37 can digest signal peptides. *E. coli* SPase I liberated nuclease A from the OmpA signal peptide while SppA2 Δ2-37 generated a product slightly bigger than nuclease A (Figure 5.10A). The two catalytically inactive mutants (S178A and K230A) did not cleave PONA again suggesting SppA2 is a Ser/Lys dyad enzyme. The digestion result suggested SppA2 Δ2-37 could possibly cleave within the OmpA signal peptide or within the nuclease A. N-terminal sequencing has confirmed SppA2 Δ2-37 cleaves within the H-region of the OmpA signal peptide (Figure 5.10B). This result is in agreement with *E. coli* SppA and *B. subtilis* SppA both having a substrate preference for hydrophobic residues such as leucine which is abundant in the H-region of signal peptides (Nam, Kim & Paetzel 2012) .

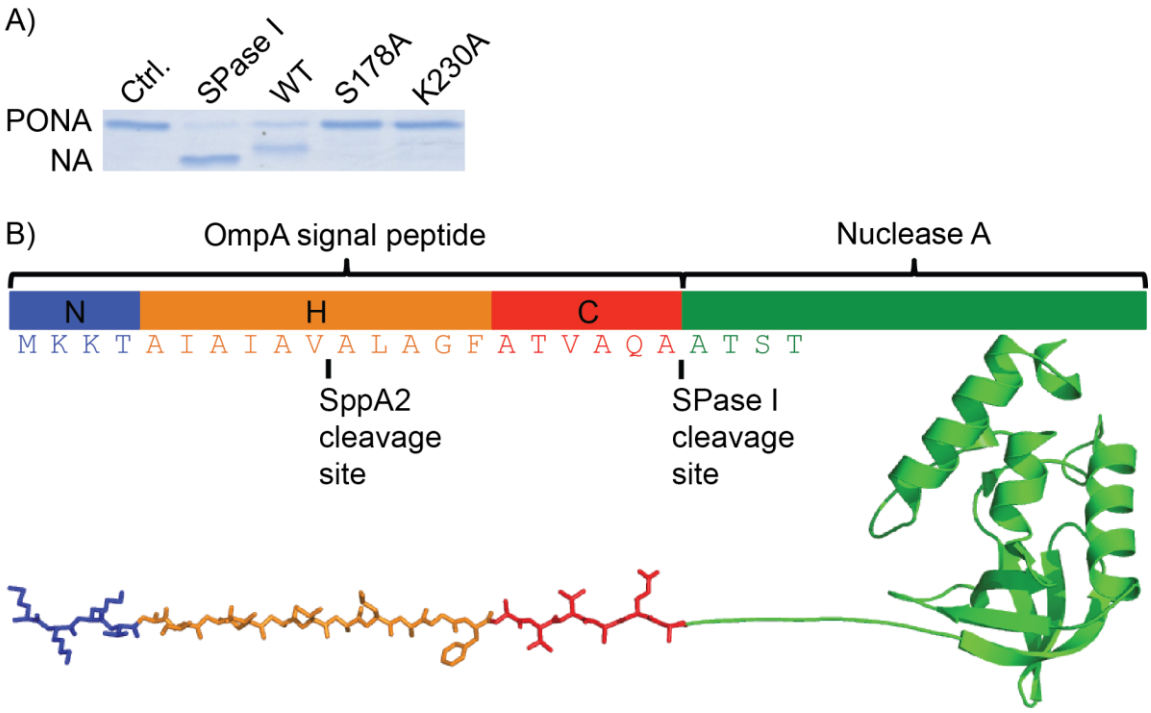


Figure 5.10. PONA digestion by E. coli SppA2 Δ2-37 and cleavage site determination by N-terminal sequencing.

A) 1µg of PONA was digested with 0.01µg of *E. coli* SPase I, SppA2 Δ2-37 (WT) or its two catalytically inactive mutants (S178A and K230A) in buffer A at 37°C overnight. PONA was incubated alone (Crtl.) to show that PONA is stable in the given reaction condition overnight. All samples were loaded on a 15% SDS-PAGE gel and stained with PageBlue protein staining solution. B) Schematic and model of PONA. The schematic shows the OmpA signal peptide's N (blue), H (orange) and C (red) regions and the nuclease A mature protein (green). The nuclease A model (PDB: 1EY0) (Chen et al. 2000) is in cartoon representation while the OmpA signal peptide is shown as sticks with the same color scheme as the PONA schematic. The protein sequence of the OmpA signal peptide is shown in the same color scheme as with the PONA model. Cleavage sites for SPase I and SppA2 are labeled.

5.3.7. The protease sensitive N-terminal region may regulate substrate uptake by E. coli SppA2 Δ2-37

Activity assays were carried out using *E. coli* SppA2 Δ2-37 and trSppA2 to determine if the protease sensitive N-terminal region is important for SppA2's activity. Aldolase, a tetrameric protein 158kDa in size, was chosen as the protein substrate for the activity assays based on its availability in the lab. Aldolase was incubated with SppA2 Δ2-37, trSppA2 and their catalytically inactive mutants (S178A and K230A). trSppA2 cleaved aldolase, however, SppA2 Δ2-37 and the catalytically inactive mutants did not (Figure 5.11A). This experiment revealed that the protease sensitive N-terminal region is not required for SppA2's protease activity, as trSppA2 has demonstrated itself

to be an active protease. From the PONA digestion assay results, it has been shown that SppA2 Δ2-37 cleaves small peptide substrates (such as OmpA signal peptide) but does not cleave aldolase (Figure 5.11A); therefore, we hypothesize that the protease sensitive N-terminal region may act as a selective filter, allowing only small peptide substrates to enter into the SppA2 octamer. To test this hypothesis, insulin β-chain (30 residues, 3.4kDa) was chosen as the substrate since its size is similar to a typical signal peptide. Both SppA2 Δ2-37 and trSppA2 cleaved insulin β-chain (Figure 5.11B). trSppA2 digested the majority of insulin β-chain after 1 hour incubation and complete digestion was seen after 16 hours of incubation. No significant digestion within the first hour was observed with SppA2 Δ2-37, nor did SppA2 Δ2-37 achieve complete digestion after 16 hours of incubation with insulin β-chain. Both the aldolase and insulin β-chain digestion results support the hypothesis that the protease sensitive N-terminal region may act as a selective filter, allowing only small peptides (i.e. remnant signal peptides) to enter into the SppA2 octamer for processing.

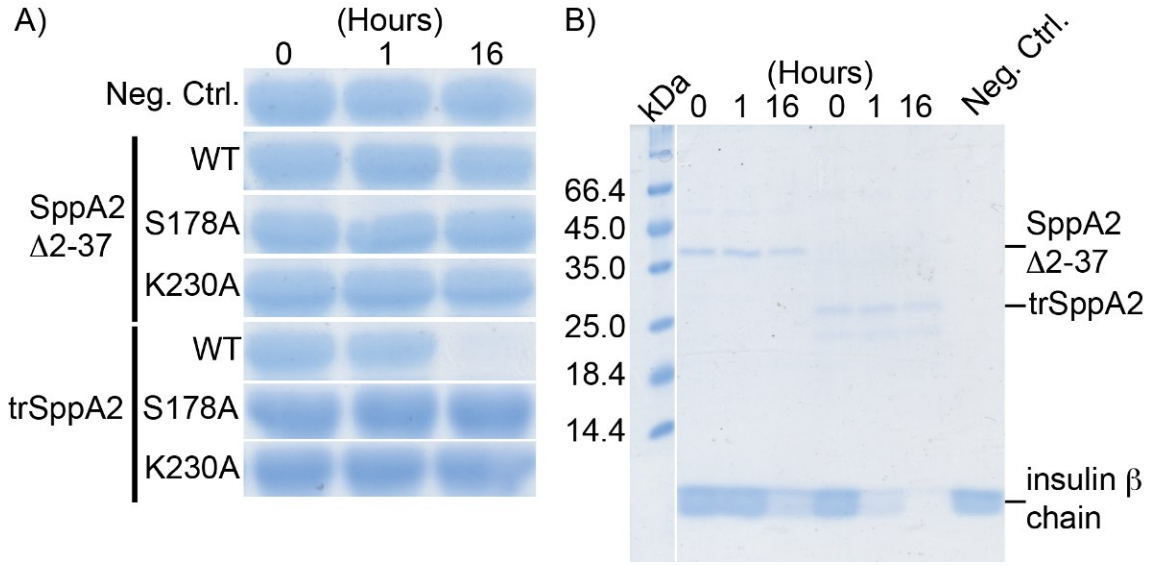


Figure 5.11. Aldolase and insulin β-chain digestion by E. coli SppA2 Δ2-37 and trSppA2.

Aldolase and insulin β-chain digestion samples on 15% SDS-PAGE gels stained with PageBlue protein staining solution. A) 10μg of aldolase was digested with 1μg of SppA2 Δ2-37 WT, S178A, K230A and trSppA2 WT, S178A, K230A in buffer A at 37°C. Aldolase was incubated alone in buffer A to serve as a negative control (Neg. Ctrl.). Samples from each reaction were taken at 0, 1 and 16 hours and analyzed on a 15% SDS-PAGE gel. B) 10μg insulin β-chain was digested separately with 1μg of SppA2 Δ2-37 and 1μg of trSppA2 in buffer A at 37°C. Samples from each reaction were taken at 0, 1 and 16 hours. Insulin β-chain incubated without enzyme for 16 hours was used as a negative control (Neg. Ctrl.). The samples were run on a 15% SDS-PAGE gel.

5.3.8. *Kinetic characterization of E. coli trSppA2 using a lipidated peptide MCA substrate*

Kinetic characterization of *E. coli* trSppA2 was carried out using continuous fluorescence assays in conjunction with two peptide MCA (methylcroumaryl amide) substrates mimicking a signal peptide (Figure 5.12). Both peptides shared the same sequence; with the only difference being that one has a dodecanoyl hydrocarbon chain attached to the N-terminus of the peptide to mimic the H-region of a signal peptide. Increase in fluorescent intensity over time, given off by free MCA liberated by trSppA2, can be measured at Ex 380nm and Em 460nm with a fluorescence spectrometer. Both peptides were treated with trSppA2 but only dodecanoyl-NGEVAKA-MCA was sufficiently digested to provide the fluorescence intensity needed for kinetic analysis (Figure 5.13A). The activity of trSppA2 S178A and K230A were examined using dodecanoyl-NGEVAKA-MCA and both showed no activity, reinforcing that trSppA2 is a Ser/Lys dyad enzyme (Figure 5.13B). The kinetic constants, k_{cat} and K_M, for *E. coli* trSppA2 were determined to be $(3.9 \pm 0.1) \times 10^{-3}\ s^{-1}$ and 4.8 ± 1.2μM, respectively (see Appendix C for the initial velocity vs substrate concentration curve L). Using the same lipidated peptide substrate, *B. subtilis* SppA was found to have a higher turn over number but a lower affinity towards the substrate (Table 5.2). Although *E. coli* SppA's kinetic characterization was carried out using a different substrate, Z-LLL-MCA, *E. coli* SppA has similar kinetic characteristics to *B. subtilis* SppA and *E. coli* SppA2.

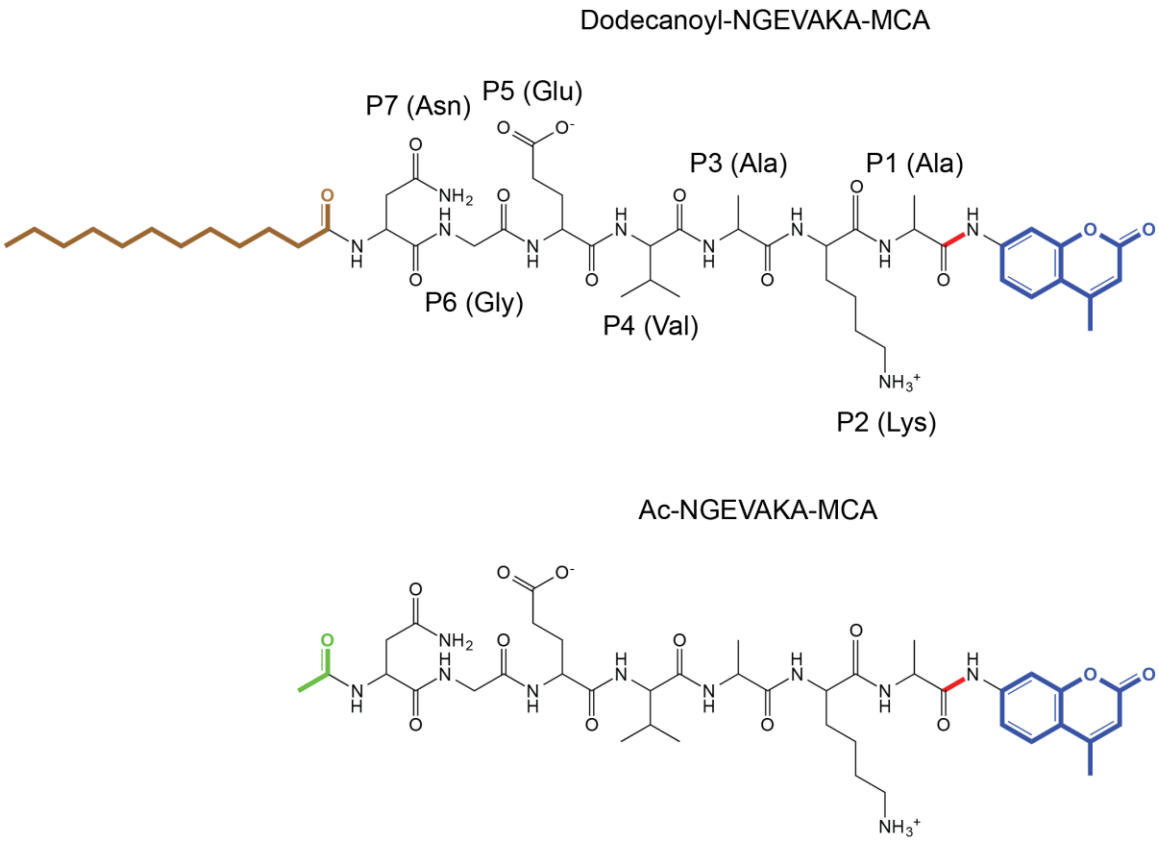


Figure 5.12. Schematic of Ac-NGEVAKA-MCA and dodecanoyl-NGEVAKA-MCA substrates.

Both peptides share the same seven amino acid sequence. The residues from the N-terminus to the C-terminus are labeled as P7 to P1, respectively. The fluorogenic moiety MCA (methylcoumaryl amide) is in blue; the scissile bond to be cleaved by *E. coli* trSppA2 is in red. The lipidated substrate has a dodecanoyl hydrocarbon tail in brown attached at the N-terminus and the non-lipidated substrate has an acetyl group (Ac) in green attached at the N-terminus.

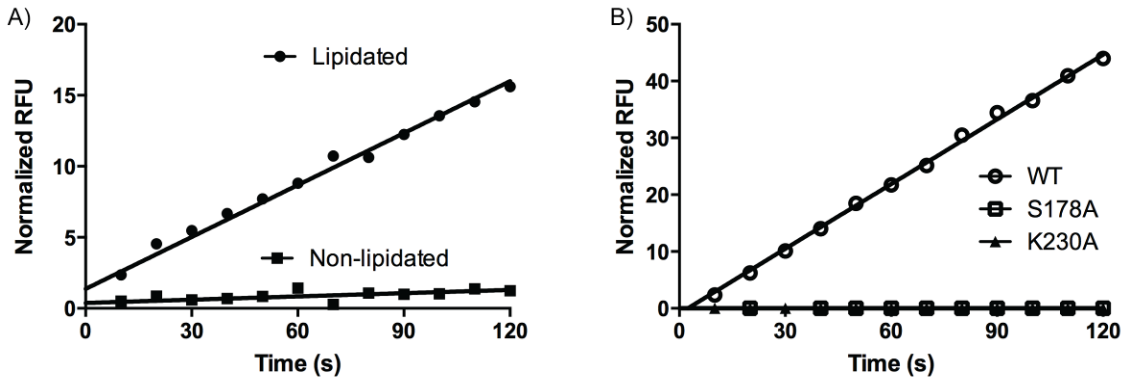


Figure 5.13. E. coli trSppA activity on Ac-NGEVAKA-MCA and dodecanoyl-NGEVAKA-MCA substrates.

A) 100μM of dodecanoyl-NGEVAKA-MCA (lipidated) and Ac-NGEVAKA-MCA (non-lipidated) were incubated separately with 1μM of *E. coli* trSppA2 in buffer A at 23°C. The increase in relative fluorescent unit (RFU) given off by free MCA over time (s) was plotted. B) 100μM of dodecanoyl-NGEVAKA-MCA was incubated separately with 1μM of *E. coli* trSppA2 wild type (WT) and two catalytically inactive mutants (S178A and K230A) in buffer A at 23°C. RFU given off by free MCA over time (s) was plotted.

Table 5.2. Kinetic constants of E. coli SppA, SppA2 and B. subtilis SppA

Enzyme	Substrate	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($s^{-1} \bullet M^{-1}$)
E. coli SppA Δ2-46*	Z-LLL-MCA	$(1.3 \pm 0.2) \times 10^{-3}$	0.4 ± 0.2	3.3×10^{3}
E. coli trSppA	Dodecanoyl-NGEVAKA-MCA	$(3.9 \pm 0.1) \times 10^{-3}$	4.8 ± 1.2	8.1×10^{2}
B. subtilis SppA Δ2-54**	Dodecanoyl-NGEVAKA-MCA	$(7.8 \pm 0.5) \times 10^{-2}$	17 ± 3.2	4.4×10^{3}

*Kim 2013, Ph.D. thesis work, **Nam & Paetzel 2013

5.3.9. E. coli trSppA2 K230A protein crystals

E. coli trSppA2 K230A was co-crystallized with dodecanoyl-K(Dabcyl)NGEVAKA AE(EDANS)T-NH_2 peptide substrate in an initial crystallization condition consisting of 20% w/v PEG3350 and 0.2M sodium malonate, as small crystal clusters without defined shapes (Figure 5.14A). This condition was further optimized to produce fewer crystals with more defined shapes (Figure 5.14B, C, D). Preliminary X-ray diffraction experiments have these crystals diffracting to 3.8Å (Table 5.3). No structure solution has been found using molecular replacement with *E. coli* SppA (PDB: 3BFO) as the search model.

Phaser molecular replacement software (McCoy et al., 2007) was used for structure solution, searching through all potential rotational and translational functions within the I2 space group. Other means of obtaining the phase information can be carried out; for example, single anomalous diffraction (SAD) experiments using seleno-methionine incorporated *E. coli* trSppA2 K230A crystals.

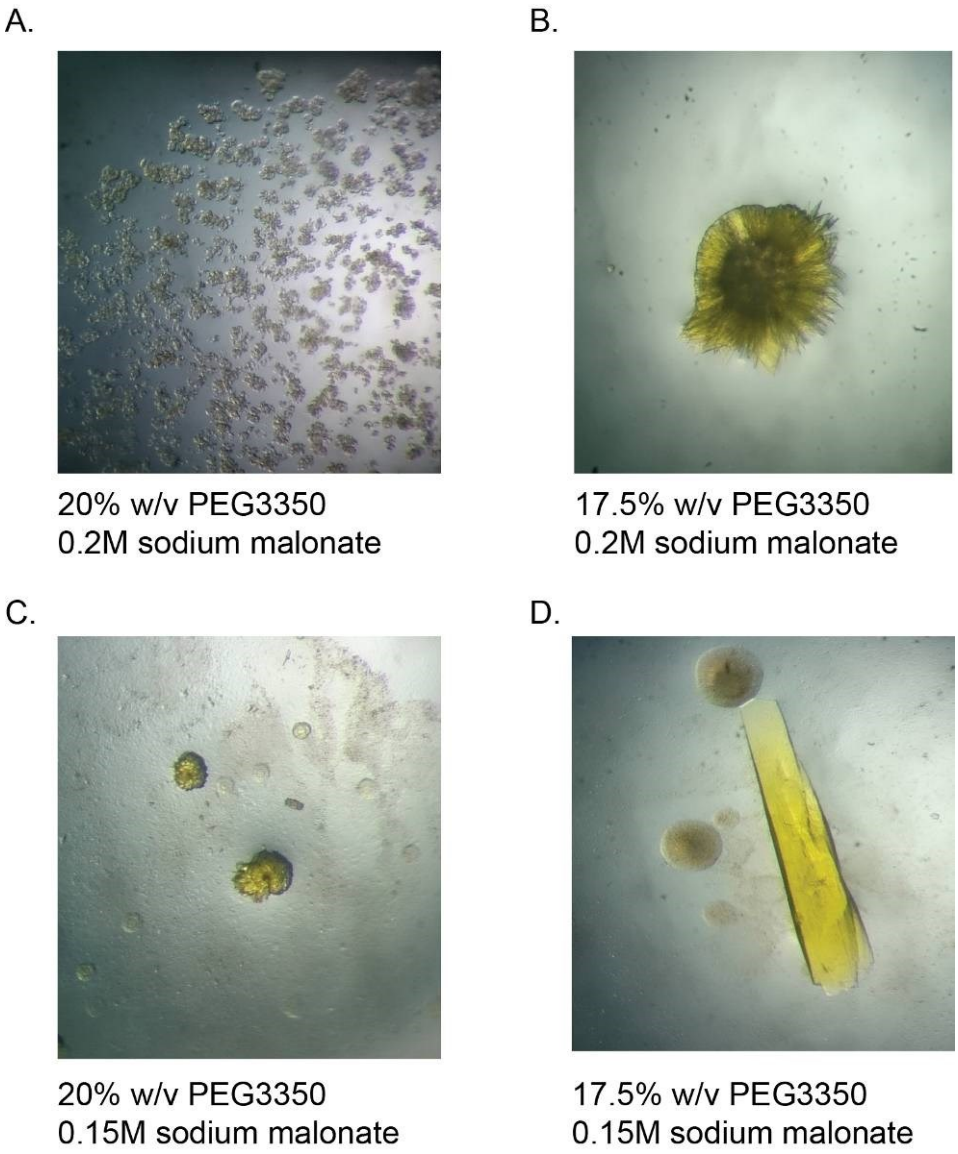


Figure 5.14. E. coli trSppA2 crystals in complex with dodecanoyl-K(Dabcyl)NGEVAKAAE(EDANS)T-NH_2 substrate.

A to D) Crystals of *E. coli* trSppA2 in complex with the peptide substrate with their respective crystallization conditions. Crystals were observed after 16 hours of incubation at room temperature. Photos were taken under 1,000x magnification.

Table 5.3. E. coli trSppA2 K230A in complex with dodecanoyl-K(Dabcyl) NGEVAKAAE(EDANS)T-NH2 peptide substrate diffraction statistics.

Crystal parameters	
Space group	I2
a,b,c (Å)	143.0, 52.1, 199.8
Data collection statistics	
Wavelength (Å)	0.9795
Resolution (Å)	43.98 – 3.8 (4.0 – 3.8)[a]
Total reflections	50001 (7262)
Unique reflections	14700 (2142)
R_{merge}[b]	0.193 (0.388)
Mean (*I*)/σ(*I*)	5.0 (2.6)
Completeness (%)	98.3 (99.4)
Multiplicity	3.4 (3.4)

[a] The data collection statistics in brackets are the values for the highest resolution shell.
[b] $R_{merge} = (\Sigma_{hkl} \Sigma_j |I_{hkl,j} - \langle I_{hkl} \rangle|) / \Sigma_{hkl} \Sigma_j I_{hkl,j}$,where $I_{hkl,j}$ is the intensity of an individual reflection and I_{hkl} is the mean intensity of that reflection.

5.3.10. E. coli SppA2 homology model

We have constructed a trSppA2 structure model using *E. coli* SppA as the template through the Phyre automatic fold recognition server (Kelley, Sternberg 2009). A single trSppA2 molecule contains both the proposed serine (178) and lysine (230) catalytic residues; however, they are separated by a distance of 28.9Å which is far beyond the limit of a typical hydrogen bond length (Figure 5.14A). A potential active site can be formed with two trSppA2 molecules where the serine nucleophile from one trSppA2 molecule comes within hydrogen bonding distance to the lysine general base from the other trSppA2 molecule at the dimerization interface (Figure 5.14B). This suggests that for trSppA2 to function as a protease it must form at least a dimer or higher order oligomers. We propose that the active form of trSppA2 is an octamer with eight potential active sites (Figure 5.14C) based on our SEC and MALS analysis and crystal structures of *E. coli* SppA and *B. subtilis* SppA (Kim, Oliver & Paetzel 2008, Nam, Kim & Paetzel 2012). The trSppA2 octamer is hypothesized to have a dome-

shaped structure with two openings into the interior of the dome (Figure 5.15). According to the model, the axial opening located at the roof of the dome has a diameter of 20.2Å and a wider opening at the base of the dome has a diameter of 87.4Å. There is a constriction on the interior face of the dome having a diameter of 50.9Å and a proposed substrate binding groove along the circumference of the dome which was observed in a peptide bound *B. subtilis* SppA structure (Nam & Paetzel 2013). Based on the dimensions of the trSppA2 octamer and the results from the aldolase digestion with trSppA2, we propose that the protein substrates are likely entering the dome through the base opening rather than the axial opening.

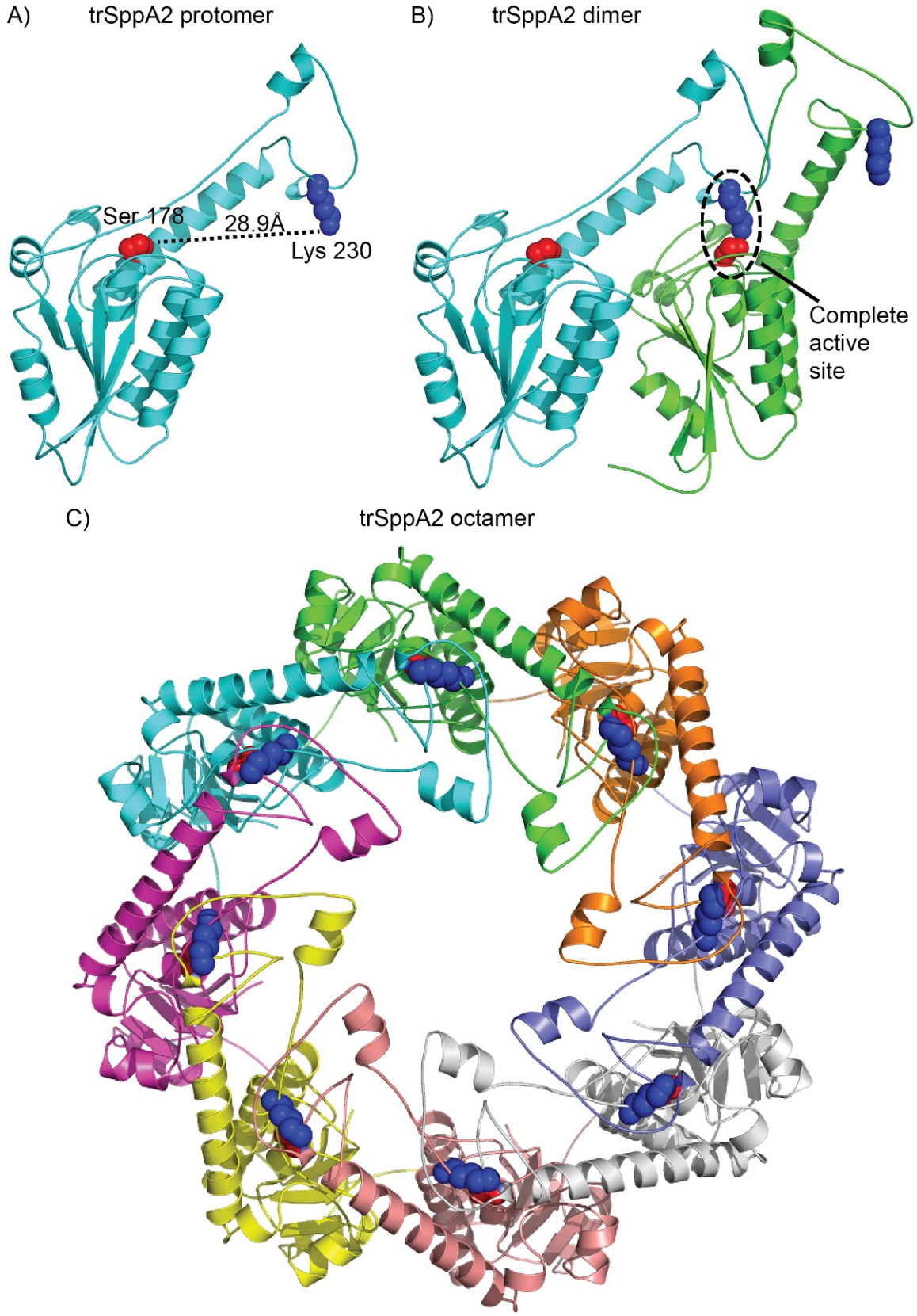

A)
trSppA2 protomer
Ser 178
28.9Å
Lys 230
B)
trSppA2 dimer
Complete
active
site
C)
trSppA2 octamer

Figure 5.15. E. coli trSppA2 structure model.

E. coli trSppA2 homology model constructed by the Phyre server using *E. coli* SppA as the template model. A) A single E. coli trSppA2 molecule depicted in cartoon with side chains of the nucleophile serine and the general base lysine shown as red and blue spheres, respectively. Hydrogen bonds are shown as dotted lines. B) *E. coli* trSppA2 dimer with one molecule in green and the other in cyan and a complete active site formed at the dimer interface. C) *E. coli* trSppA2 octamer with each promoter in a different color.

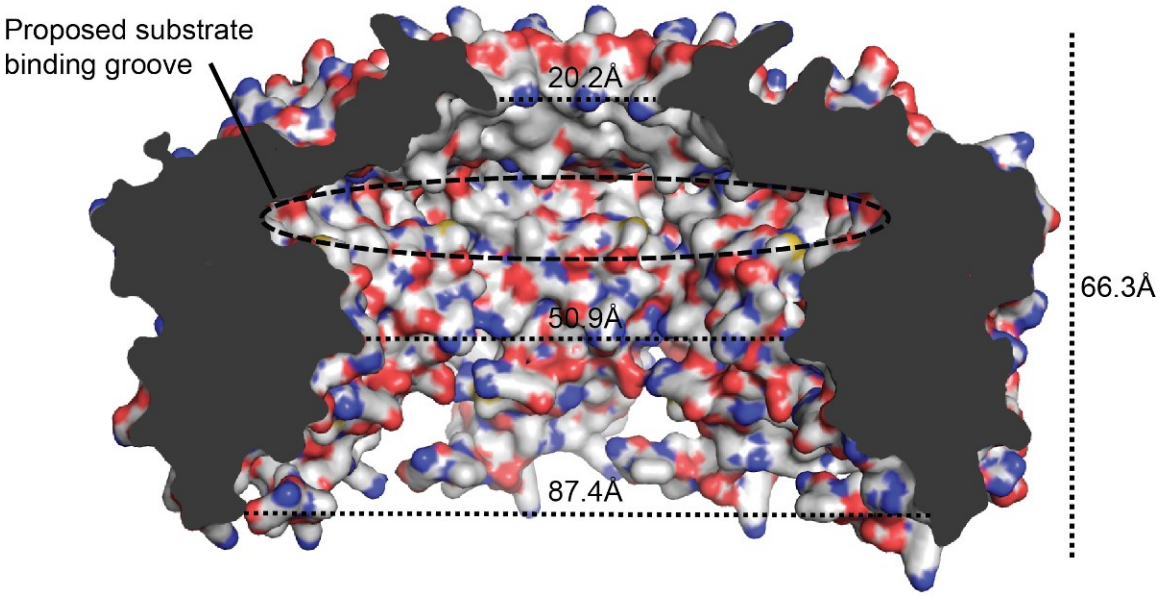


Figure 5.16. E. coli SppA2 octamer dimension.

A cross-section of *E. coli* SppA2 octamer in surface representation with carbon atoms colored in white, nitrogen in blue, oxygen in red and sulphur in yellow. The distance shown in Ångström (Å) was measured using the Measure function in PyMOL.

5.4. Discussion

In this chapter we have presented the first biochemical characterization data of *E. coli* SppA2, both from an enzymatic activity aspect and a structural perspective. Our results strongly suggest that *E. coli* SppA2 is an octameric Ser/Lys dyad enzyme whose quaternary structure is similar to that of *E. coli* and *B. subtilis* SppA. *E. coli* SppA2 shares many traits with to *E. coli* SppA and *B. subtilis* SppA such as similar catalytic residues, domain sizes and oligomeric states but, unlike other SppA enzymes, *E. coli* SppA2 has an unique N-terminal region (Figure 5.17).

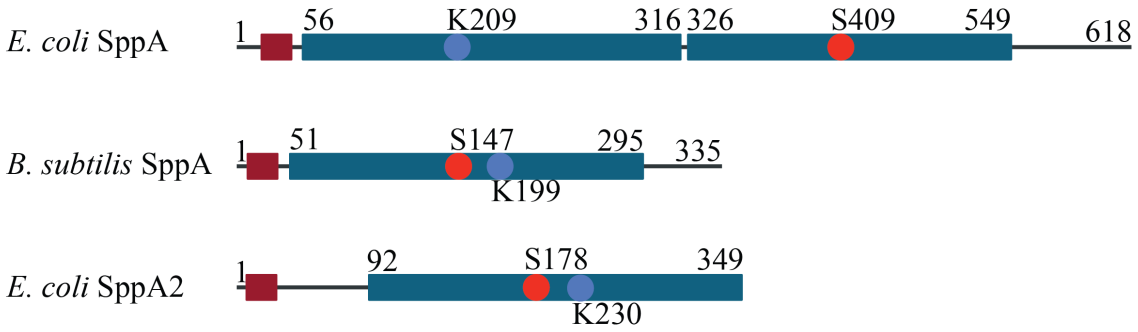


Figure 5.17. E. coli and B. subtilis SppA and E. coli SppA2 domain organization.

The red boxes represent the predicted transmembrane segment. The catalytic domains are represented by blue boxes. The red and light blue spheres represent the proposed catalytic serine and lysine, respectively.

Due to *E. coli* SppA2's classification into the S49 peptidase family, the same as *E. coli* SppA's, *E. coli* SppA2 may also participate in remnant signal peptide degradation. Using a peptide-based activity assay with a substrate mimicking the H-region and the C-region of a typical peptide, we have shown that the H-region having high hydrophobicity contained within the peptide substrate is important for interaction with *E. coli* SppA2. However, there are limitations to what can be hypothesized since the *E. coli* SppA2 used in this study consists only of the soluble domain without the single predicted N-terminal transmembrane anchor and the peptide substrate used is not embedded in a liposome so it does not represent *in vivo* membrane embedded remnant signal peptides. It is not clear how *E. coli* SppA2 extracts membrane embedded peptides and moves them into the central cavity of the octameric dome as *E. coli* SppA2 active sites could be as far as 140Å away from the membrane surface when the protease sensitive N-terminal region takes on an extended α-helical conformation (Figure 5.18A). Perhaps, the protease sensitive N-terminal region can adopt a different conformation or interact with the membrane to pull the octameric dome towards the membrane surface for direct interaction with membrane embedded peptides (Figure 5.18B). Further studies utilizing *E. coli* phospholipid liposomes can be carried out to examine SppA2's role in remnant signal peptide degradation under more physiological conditions. *E. coli* phospholipid liposomes can be reconstituted with *E. coli* SPase I and pre-proteins, as has been previously done with *E. coli* SPase I and M13 viral procoat proteins (Iwashita & Wickner 1982). After a successful protein maturation event, these remnant signal peptide containing liposomes can then be treated with either *E. coli* SppA2 Δ2-37 or *E. coli*

trSppA2 to examine if the N-terminal protease sensitive region participates in membrane interaction and/or remnant signal peptide extraction from the liposomes.

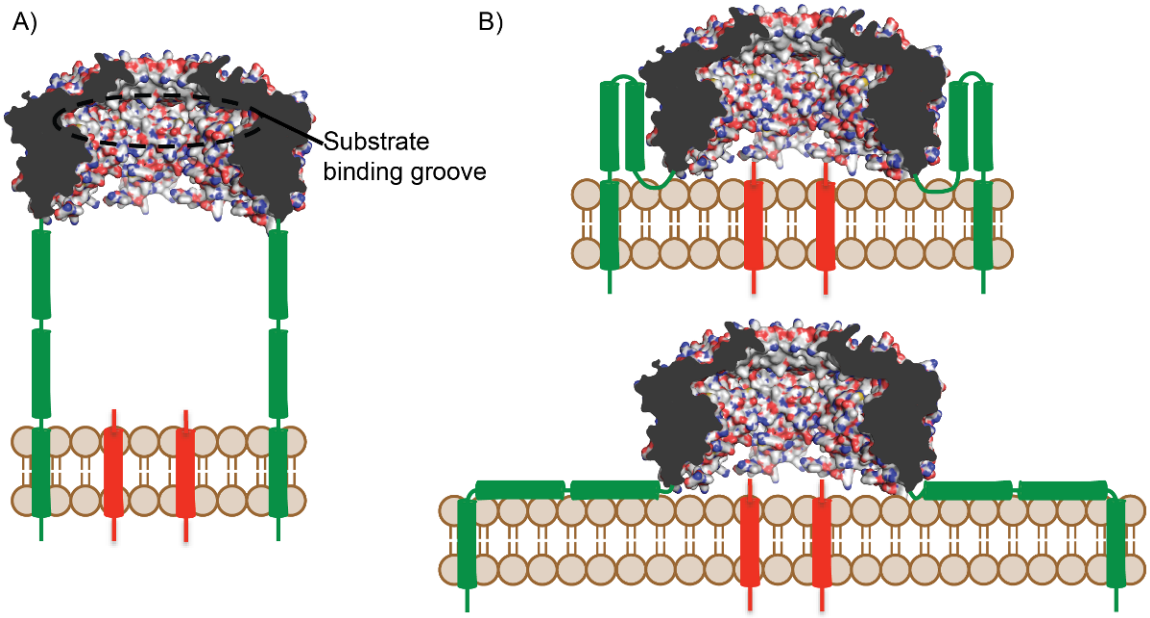


Figure 5.18. E. coli SppA2 and membrane interaction.

A) *E. coli* SppA2 octamer cross section with the single predicted N-terminal transmembrane helix and the two predicted protease sensitive helices as green cylinders for two *E. coli* SppA2 protomers. Remnant signal peptides are depicted as red cylinders. B) Two potential conformations for the N-terminal protease sensitive helices to bring the *E. coli* SppA2 octamer closer to the membrane surface.

In addition to peptidase activity, we have shown that *E. coli* SppA2 also has proteinase activity. We propose that *E. coli* SppA2 may also participate in protein quality control within the cytoplasm. Bacterial proteases involved in protein quality control such as ClpP, HtrA and Lon, are self-compartmentalized which allows them to hide their active sites from the surrounding environment and prevent accidental degradation of native proteins. We have also shown that the protease sensitive N-terminal region of E. coli SppA2 may participate in regulation of the substrate's entry into and exit out of the octameric dome. Regulated entry and exit into compartmentalized proteases is another common strategy utilized by many protein quality control enzymes to ensure that only the desired substrates will gain entry into the inside of the protease compartment where the active sites reside.

To further optimize the current *E. coli* trSppA2 K230A crystals, other general proteases, such as pepsin and proteinase K, can be used to generate protease resistant domains that may be different from the one generated by trypsin. Surface entropy reduction mutagenesis (SERM) can be carried out to create protein surfaces that may be more susceptible to crystallization compared to that of the wild type enzyme. Lastly, detergents other than DDM such as Triton X-100, Tween 20 and Elugent, can also be employed, as well as exposure to other crystallization conditions.

6. Project summary and future directions

6.1. Summary of findings

Protein translocation across cell membrane is an essential cellular process. This thesis work focuses on two enzymes related to this process. Type I signal peptidases (SPase I) directly participate in this process by liberating secreted proteins from their membrane embedded signal peptides. The second enzyme, signal peptide peptidase A2 (SppA2), indirectly participates in protein translocation through its proposed remnant signal peptide removal function in conjunction with other enzymes (e.g. SppA), regenerating the limited membrane space and thus, allowing protein translocation to continue.

6.1.1. S. aureus SpsB project summary

A truncated *S. aureus* SpsB (Δ2-20), still retaining most of its enzymatic activity, was successfully designed, created, overexpressed, purified and refolded. Crystallization amounts (>10mg/mL) of *S. aureus* SpsB Δ2-20 were prepared through the refolding of inclusion bodies. *S. aureus* SpsB Δ2-20 S36A was co-crystallized with arylomycin A_2 inhibitor. Kinetic characterization of *S. aureus*, *S. epidermidis*, *B. subtilis* and *E. coli* SPase I enzymes show that while the Ala-X-Ala consensus sequence of a signal peptide is important for SPase I recognition, the amino acid sequence found in the C-region of a signal peptide also plays an important role in interacting with SPase I. The highly conserved proline in box B (proline 88 in *E. coli* SPase I) of Gram-negative SPase I may contribute to the difference in catalytic efficiency observed between Gram-positive and Gram-negative SPase I. Lastly, Gram-positive SPase I likely uses a water molecule to stabilize the tetrahedral oxyanion intermediate during catalysis rather than a serine which is used by Gram-negative SPase I to fulfill the same role.

6.1.2. *E. coli SppA2 project summary*

Before this thesis project was undertaken, very little was known about *E. coli* SppA2, a proposed Ser/Lys dyad enzyme with a remnant signal peptide removal function based on its sequence similarities to the characterized *E. coli* and *B. subtilis* SppA enzymes. From the results of a series of limited proteolysis experiments with trypsin, we have shown that *E. coli* SppA2 has a flexible N-terminal region and a protease resistant C-terminal catalytic domain. The N-terminal region is required for proper protein folding but is not required for enzymatic activity. Through light scattering spectroscopy and size exclusion analysis, we have shown that *E. coli* SppA2's catalytic domains form an octamer in solution and likely share the same dome shape structure as *E. coli* SppA and *B. subtilis* SppA. From mutagenesis and enzymatic activity studies we have shown that *E. coli SppA2* is, in high probability, a Ser/Lys dyad enzyme with the proposed nucleophile serine and general base lysine at positions 178 and 230, respectively. *E. coli* SppA2 can digest only small peptides while *E. coli* SppA2 without its N-terminal region can digest both small peptides and large folded proteins with a preference for hydrophobic substrates. These results suggest that the N-terminal region of *E. coli* SppA2 may play a regulatory role on its enzymatic activity.

6.2. Future directions

6.2.1. *Future directions for S. aureus SpsB project*

A preliminary X-ray diffraction experiment was carried out using the *S. aureus* SpsB crystals obtained from the first cycle of optimization. Further optimization of crystallization conditions will lead to improved diffraction quality and to the first Gram-positive SPase I structure. In the kinetic characterization of various SPase I, the substrate represents a Gram-positive signal peptide and worked better with Gram-positive SPase I. A substrate having a Gram-negative signal peptide sequence should be designed and used in the next set of kinetic studies. This was attempted and a peptide substrate with the signal peptide sequence from OmpA (an *E. coli* protein) was synthesized but it was too hydrophobic and precipitated in kinetic buffer during the activity analysis. Other Gram-negative signal peptide sequences should be explored to

find a more suitable candidate for kinetic characterization. This thesis has only explored residues found in box B which form one side of the substrate binding groove. Box C and D residues, which form the other side of the substrate binding groove, should be examined to identify if any are Gram-positive or Gram-negative specific. Genetic modification of *E. coli* BL21(DE3) SPase I to introduce the P88A mutation may result in increased secretory protein production as the P88A mutant is better at processing PONA compared to the wild type enzyme, as seen in our *in vitro* PONA digest assay.

6.2.2. Future directions for E. coli SppA2 project

Preliminary X-ray diffraction was observed for crystals of *E. coli* SppA2 in complex with the dodecanoyl-K(Dabcyl)NGEVAKAAE(EDANS)T-NH_2 peptide substrate. Further optimization is required to obtain higher diffraction quality crystals that will lead to determining the structure of the complex. In this thesis, we have shown that the N-terminal region is important for protein folding and may regulate enzymatic activity. We propose that this region may associate with the cell membrane. Circular dichroism analysis in the presence or the absence of liposomes, may provide insight into whether this region undergoes any conformation change to interact with the membrane surface. Surface plasmon resonance technology, using *E. coli* SppA2 Δ2-37 and trSppA2 with a liposome coated chip, can be used to determine if the N-terminal region contributes to the interaction with the cell membrane.

References

Auclair, S.M., Bhanu, M.K. & Kendall, D.A. 2012, "Signal peptidase I: cleaving the way to mature proteins", *Protein science : a publication of the Protein Society,* vol. 21, no. 1, pp. 13-25.

Baird, L., Lipinska, B., Raina, S. & Georgopoulos, C. 1991, "Identification of the Escherichia coli sohB gene, a multicopy suppressor of the HtrA (DegP) null phenotype", *Journal of Bacteriology,* vol. 173, no. 18, pp. 5763-5770.

Barkocy-Gallagher, G.A. & Bassford, P.J.,Jr 1992, "Synthesis of precursor maltose-binding protein with proline in the +1 position of the cleavage site interferes with the activity of Escherichia coli signal peptidase I in vivo", *The Journal of biological chemistry,* vol. 267, no. 2, pp. 1231-1238.

Black, M.T. 1993, "Evidence that the catalytic activity of prokaryote leader peptidase depends upon the operation of a serine-lysine catalytic dyad", *Journal of Bacteriology,* vol. 175, no. 16, pp. 4957-4961.

Bruton, G., Huxley, A., O'Hanlon, P., Orlek, B., Eggleston, D., Humphries, J., Readshaw, S., West, A., Ashman, S., Brown, M., Moore, K., Pope, A., O'Dwyer, K. & Wang, L. 2003, "Lipopeptide substrates for SpsB, the Staphylococcus aureus type I signal peptidase: design, conformation and conversion to alpha-ketoamide inhibitors", *European journal of medicinal chemistry,* vol. 38, no. 4, pp. 351-356.

Buchan, D.W., Minneci, F., Nugent, T.C., Bryson, K. & Jones, D.T. 2013, "Scalable web services for the PSIPRED Protein Analysis Workbench", *Nucleic acids research,* vol. 41, no. Web Server issue, pp. W349-57.

Buzder-Lantos, P., Bockstael, K., Anne, J. & Herdewijn, P. 2009, "Substrate based peptide aldehyde inhibits bacterial type I signal peptidase", *Bioorganic & medicinal chemistry letters,* vol. 19, no. 10, pp. 2880-2883.

Carlos, J.L., Klenotic, P.A., Paetzel, M., Strynadka, N.C. & Dalbey, R.E. 2000, "Mutational evidence of transition state stabilization by serine 88 in Escherichia coli type I signal peptidase", *Biochemistry,* vol. 39, no. 24, pp. 7276-7283.

Chen, J., Lu, Z., Sakon, J. & Stites, W.E. 2000, "Increasing the thermostability of staphylococcal nuclease: implications for the origin of protein thermostability", *Journal of Molecular Biology,* vol. 303, no. 2, pp. 125-130.

Chen, L., Tai, P.C., Briggs, M.S. & Gierasch, L.M. 1987, "Protein translocation into Escherichia coli membrane vesicles is inhibited by functional synthetic signal peptides", *The Journal of biological chemistry,* vol. 262, no. 4, pp. 1427-1429.

Chung, C.H., Ives, H.E., Almeda, S. & Goldberg, A.L. 1983, "Purification from Escherichia coli of a periplasmic protein that is a potent inhibitor of pancreatic proteases", *The Journal of biological chemistry,* vol. 258, no. 18, pp. 11032-11038.

Dalbey, R.E., Lively, M.O., Bron, S. & van Dijl, J.M. 1997, "The chemistry and enzymology of the type I signal peptidases", *Protein science : a publication of the Protein Society,* vol. 6, no. 6, pp. 1129-1138.

Dalbey, R.E. & Wickner, W. 1985, "Leader peptidase catalyzes the release of exported proteins from the outer surface of the Escherichia coli plasma membrane", *The Journal of biological chemistry,* vol. 260, no. 29, pp. 15925-15931.

Date, T. 1983, "Demonstration by a novel genetic technique that leader peptidase is an essential enzyme of Escherichia coli", *Journal of Bacteriology,* vol. 154, no. 1, pp. 76-83.

Date, T. & Wickner, W. 1981, "Isolation of the Escherichia coli leader peptidase gene and effects of leader peptidase overproduction in vivo", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 78, no. 10, pp. 6106-6110.

Degering, C., Eggert, T., Puls, M., Bongaerts, J., Evers, S., Maurer, K.H. & Jaeger, K.E. 2010, "Optimization of protease secretion in Bacillus subtilis and Bacillus licheniformis by screening of homologous and heterologous signal peptides", *Applied and Environmental Microbiology,* vol. 76, no. 19, pp. 6370-6376.

Duong, F. & Wickner, W. 1997a, "Distinct catalytic roles of the SecYE, SecG and SecDFyajC subunits of preprotein translocase holoenzyme", *The EMBO journal,* vol. 16, no. 10, pp. 2756-2768.

Duong, F. & Wickner, W. 1997b, "The SecDFyajC domain of preprotein translocase controls preprotein movement by regulating SecA membrane cycling", *The EMBO journal,* vol. 16, no. 16, pp. 4871-4879.

Economou, A., Christie, P.J., Fernandez, R.C., Palmer, T., Plano, G.V. & Pugsley, A.P. 2006, "Secretion by numbers: Protein traffic in prokaryotes", *Molecular microbiology,* vol. 62, no. 2, pp. 308-319.

Economou, A. & Wickner, W. 1994, "SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion", *Cell,* vol. 78, no. 5, pp. 835-843.

Eggers, C.T., Murray, I.A., Delmar, V.A., Day, A.G. & Craik, C.S. 2004, "The periplasmic serine protease inhibitor ecotin protects bacteria against neutrophil elastase", *The Biochemical journal,* vol. 379, no. Pt 1, pp. 107-118.

Eggers, C.T., Wang, S.X., Fletterick, R.J. & Craik, C.S. 2001, "The role of ecotin dimerization in protease inhibition", *Journal of Molecular Biology,* vol. 308, no. 5, pp. 975-991.

Ekici, O.D., Paetzel, M. & Dalbey, R.E. 2008, "Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration", *Protein science : a publication of the Protein Society,* vol. 17, no. 12, pp. 2023-2037.

Grudnik, P., Bange, G. & Sinning, I. 2009, "Protein targeting by the signal recognition particle", *Biological chemistry,* vol. 390, no. 8, pp. 775-782.

Guex, N., Peitsch, M.C. & Schwede, T. 2009, "Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective", *Electrophoresis,* vol. 30 Suppl 1, pp. S162-73.

Hussain, M., Ichihara, S. & Mizushima, S. 1982, "Mechanism of signal peptide cleavage in the biosynthesis of the major lipoprotein of the Escherichia coli outer membrane", *The Journal of biological chemistry,* vol. 257, no. 9, pp. 5177-5182.

Hussain, M., Ozawa, Y., Ichihara, S. & Mizushima, S. 1982, "Signal peptide digestion in Escherichia coli. Effect of protease inhibitors on hydrolysis of the cleaved signal peptide of the major outer-membrane lipoprotein", *European journal of biochemistry / FEBS,* vol. 129, no. 1, pp. 233-239.

Hyberts, S.G., Goldberg, M.S., Havel, T.F. & Wagner, G. 1992, "The solution structure of eglin c based on measurements of many NOEs and coupling constants and its comparison with X-ray structures", *Protein science : a publication of the Protein Society,* vol. 1, no. 6, pp. 736-751.

Ichihara, S., Beppu, N. & Mizushima, S. 1984, "Protease IV, a cytoplasmic membrane protein of Escherichia coli, has signal peptide peptidase activity", *The Journal of biological chemistry,* vol. 259, no. 15, pp. 9853-9857.

Ichihara, S., Suzuki, T., Suzuki, M. & Mizushima, S. 1986, "Molecular cloning and sequencing of the sppA gene and characterization of the encoded protease IV, a signal peptide peptidase, of Escherichia coli", *The Journal of biological chemistry,* vol. 261, no. 20, pp. 9405-9411.

Jones, D.T. 1999, "Protein secondary structure prediction based on position-specific scoring matrices", *Journal of Molecular Biology,* vol. 292, no. 2, pp. 195-202.

Karamanou, S., Gouridis, G., Papanikou, E., Sianidis, G., Gelis, I., Keramisanou, D., Vrontou, E., Kalodimos, C.G. & Economou, A. 2007, "Preprotein-controlled catalysis in the helicase motor of SecA", *The EMBO journal,* vol. 26, no. 12, pp. 2904-2914.

Kelley, L.A. & Sternberg, M.J. 2009, "Protein structure prediction on the Web: a case study using the Phyre server", *Nature protocols,* vol. 4, no. 3, pp. 363-371.

Kihara, A., Akiyama, Y. & Ito, K. 1995, "FtsH is required for proteolytic elimination of uncomplexed forms of SecY, an essential protein translocase subunit", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 92, no. 10, pp. 4532-4536.

Kim, A.C., Oliver, D.C. & Paetzel, M. 2008, "Crystal structure of a bacterial signal Peptide peptidase", *Journal of Molecular Biology,* vol. 376, no. 2, pp. 352-366.

Kim, Y.T., Muramatsu, T. & Takahashi, K. 1995, "Leader peptidase from Escherichia coli: overexpression, characterization, and inactivation by modification of tryptophan residues 300 and 310 with N-bromosuccinimide", *Journal of Biochemistry,* vol. 117, no. 3, pp. 535-544.

Klenotic, P.A., Carlos, J.L., Samuelson, J.C., Schuenemann, T.A., Tschantz, W.R., Paetzel, M., Strynadka, N.C. & Dalbey, R.E. 2000, "The role of the conserved box E residues in the active site of the Escherichia coli type I signal peptidase", *The Journal of biological chemistry,* vol. 275, no. 9, pp. 6490-6498.

Krieger, E., Joo, K., Lee, J., Lee, J., Raman, S., Thompson, J., Tyka, M., Baker, D. & Karplus, K. 2009, "Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8", *Proteins,* vol. 77 Suppl 9, pp. 114-122.

Krieger, M., Kay, L.M. & Stroud, R.M. 1974, "Structure and specific binding of trypsin: comparison of inhibited derivatives and a model for substrate binding", *Journal of Molecular Biology,* vol. 83, no. 2, pp. 209-230.

Kuo, D., Weidner, J., Griffin, P., Shah, S.K. & Knight, W.B. 1994, "Determination of the kinetic parameters of Escherichia coli leader peptidase activity using a continuous assay: the pH dependence and time-dependent inhibition by beta-lactams are consistent with a novel serine protease mechanism", *Biochemistry,* vol. 33, no. 27, pp. 8347-8354.

Liu, J., Luo, C., Smith, P.A., Chin, J.K., Page, M.G., Paetzel, M. & Romesberg, F.E. 2011, "Synthesis and characterization of the arylomycin lipoglycopeptide antibiotics and the crystallographic analysis of their complex with signal peptidase", *Journal of the American Chemical Society,* vol. 133, no. 44, pp. 17869-17877.

Luirink, J., ten Hagen-Jongman, C.M., van der Weijden, C.C., Oudega, B., High, S., Dobberstein, B. & Kusters, R. 1994, "An alternative protein targeting pathway in Escherichia coli: studies on the role of FtsY", *The EMBO journal,* vol. 13, no. 10, pp. 2289-2296.

Luo, C., Roussel, P., Dreier, J., Page, M.G. & Paetzel, M. 2009, "Crystallographic analysis of bacterial signal peptidase in ternary complex with arylomycin A2 and a beta-sultam inhibitor", *Biochemistry,* vol. 48, no. 38, pp. 8976-8984.

Matsuyama, S., Fujita, Y. & Mizushima, S. 1993, "SecD is involved in the release of translocated secretory proteins from the cytoplasmic membrane of Escherichia coli", *The EMBO journal,* vol. 12, no. 1, pp. 265-270.

Maurizi, M.R. 1992, "Proteases and protein degradation in Escherichia coli", *Experientia,* vol. 48, no. 2, pp. 178-201.

McGrath, M.E., Erpel, T., Bystroff, C. & Fletterick, R.J. 1994, "Macromolecular chelation as an improved mechanism of protease inhibition: structure of the ecotin-trypsin complex", *The EMBO journal,* vol. 13, no. 7, pp. 1502-1507.

Milstein, C., Brownlee, G.G., Harrison, T.M. & Mathews, M.B. 1972, "A possible precursor of immunoglobulin light chains", *Nature: New biology,* vol. 239, no. 91, pp. 117-120.

Nam, S.E., Kim, A.C. & Paetzel, M. 2012, "Crystal structure of Bacillus subtilis signal peptide peptidase A", *Journal of Molecular Biology,* vol. 419, no. 5, pp. 347-358.

Nam, S.E. & Paetzel, M. 2013, "Structure of signal peptide peptidase A with C-termini bound in the active sites: insights into specificity, self-processing, and regulation", *Biochemistry,* vol. 52, no. 49, pp. 8811-8822.

Nilsson, I. & von Heijne, G. 1992, "A signal peptide with a proline next to the cleavage site inhibits leader peptidase when present in a sec-independent protein", *FEBS letters,* vol. 299, no. 3, pp. 243-246.

Nishiyama, K., Suzuki, T. & Tokuda, H. 1996, "Inversion of the membrane topology of SecG coupled with SecA-dependent preprotein translocation", *Cell,* vol. 85, no. 1, pp. 71-81.

Novak, P. & Dev, I.K. 1988, "Degradation of a signal peptide by protease IV and oligopeptidase A", *Journal of Bacteriology,* vol. 170, no. 11, pp. 5067-5075.

Pacaud, M. 1982, "Purification and characterization of two novel proteolytic enzymes in membranes of Escherichia coli. Protease IV and protease V", *The Journal of biological chemistry,* vol. 257, no. 8, pp. 4333-4339.

Paetzel, M. 2013, "Structure and mechanism of Escherichia coli type I signal peptidase", *Biochimica et biophysica acta,* .

Paetzel, M., Dalbey, R.E. & Strynadka, N.C. 2000, "The structure and mechanism of bacterial type I signal peptidases. A novel antibiotic target", *Pharmacology & therapeutics,* vol. 87, no. 1, pp. 27-49.

Paetzel, M., Dalbey, R.E. & Strynadka, N.C. 1998, "Crystal structure of a bacterial signal peptidase in complex with a beta-lactam inhibitor", *Nature,* vol. 396, no. 6707, pp. 186-190.

Paetzel, M., Goodall, J.J., Kania, M., Dalbey, R.E. & Page, M.G. 2004, "Crystallographic and biophysical analysis of a bacterial signal peptidase in complex with a lipopeptide-based inhibitor", *The Journal of biological chemistry,* vol. 279, no. 29, pp. 30781-30790.

Paetzel, M. & Strynadka, N.C. 1999, "Common protein architecture and binding sites in proteases utilizing a Ser/Lys dyad mechanism", *Protein science : a publication of the Protein Society,* vol. 8, no. 11, pp. 2533-2536.

Paetzel, M., Strynadka, N.C., Tschantz, W.R., Casareno, R., Bullinger, P.R. & Dalbey, R.E. 1997, "Use of site-directed chemical modification to study an essential lysine in Escherichia coli leader peptidase", *The Journal of biological chemistry,* vol. 272, no. 15, pp. 9994-10003.

Pal, G., Szilagyi, L. & Graf, L. 1996, "Stable monomeric form of an originally dimeric serine proteinase inhibitor, ecotin, was constructed via site directed mutagenesis", *FEBS letters,* vol. 385, no. 3, pp. 165-170.

Papanikou, E., Karamanou, S. & Economou, A. 2007, "Bacterial protein secretion through the translocase nanomachine", *Nature reviews.Microbiology,* vol. 5, no. 11, pp. 839-851.

Perona, J.J., Tsu, C.A., Craik, C.S. & Fletterick, R.J. 1997, "Crystal structure of an ecotin-collagenase complex suggests a model for recognition and cleavage of the collagen triple helix", *Biochemistry,* vol. 36, no. 18, pp. 5381-5392.

Randall, L.L., Crane, J.M., Lilly, A.A., Liu, G., Mao, C., Patel, C.N. & Hardy, S.J. 2005, "Asymmetric binding between SecA and SecB two symmetric proteins: implications for function in export", *Journal of Molecular Biology,* vol. 348, no. 2, pp. 479-489.

Randall, L.L., Topping, T.B., Hardy, S.J., Pavlov, M.Y., Freistroffer, D.V. & Ehrenberg, M. 1997, "Binding of SecB to ribosome-bound polypeptides has the same characteristics as binding to full-length, denatured proteins", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 94, no. 3, pp. 802-807.

Rao, S., Bockstael, K., Nath, S., Engelborghs, Y., Anne, J. & Geukens, N. 2009, "Enzymatic investigation of the Staphylococcus aureus type I signal peptidase SpsB - implications for the search for novel antibiotics", *The FEBS journal,* vol. 276, no. 12, pp. 3222-3234.

Rawlings, N.D., Waller, M., Barrett, A.J. & Bateman, A. 2013, "MEROPS: the database of proteolytic enzymes, their substrates and inhibitors", *Nucleic acids research,* .

Regnier, P. 1981, "The purification of protease IV of E. coli and the demonstration that it is an endoproteolytic enzyme", *Biochemical and biophysical research communications,* vol. 99, no. 4, pp. 1369-1376.

Saito, A., Hizukuri, Y., Matsuo, E., Chiba, S., Mori, H., Nishimura, O., Ito, K. & Akiyama, Y. 2011, "Post-liberation cleavage of signal peptides is catalyzed by the site-2 protease (S2P) in bacteria", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 108, no. 33, pp. 13740-13745.

Schechter, I. & Berger, A. 1967, "On the size of the active site in proteases. I. Papain", *Biochemical and biophysical research communications,* vol. 27, no. 2, pp. 157-162.

SCHOELLMANN, G. & SHAW, E. 1963, "Direct evidence for the presence of histidine in the active center of chymotrypsin", *Biochemistry,* vol. 2, pp. 252-255.

Shen, L.M., Lee, J.I., Cheng, S.Y., Jutte, H., Kuhn, A. & Dalbey, R.E. 1991, "Use of site-directed mutagenesis to define the limits of sequence variation tolerated for processing of the M13 procoat protein by the Escherichia coli leader peptidase", *Biochemistry,* vol. 30, no. 51, pp. 11775-11781.

Shin, D.H., Song, H.K., Seong, I.S., Lee, C.S., Chung, C.H. & Suh, S.W. 1996, "Crystal structure analyses of uncomplexed ecotin in two crystal forms: implications for its function and stability", *Protein science : a publication of the Protein Society,* vol. 5, no. 11, pp. 2236-2247.

Sianidis, G., Karamanou, S., Vrontou, E., Boulias, K., Repanas, K., Kyrpides, N., Politou, A.S. & Economou, A. 2001, "Cross-talk between catalytic and regulatory elements in a DEAD motor domain is essential for SecA function", *The EMBO journal,* vol. 20, no. 5, pp. 961-970.

Silhavy, T.J., Kahne, D. & Walker, S. 2010, "The bacterial cell envelope", *Cold Spring Harbor perspectives in biology,* vol. 2, no. 5, pp. a000414.

Simor, A.E., Gilbert, N.L., Gravel, D., Mulvey, M.R., Bryce, E., Loeb, M., Matlow, A., McGeer, A., Louie, L., Campbell, J. & Canadian Nosocomial Infection Surveillance Program 2010, "Methicillin-resistant Staphylococcus aureus colonization or infection in Canada: National Surveillance and Changing Epidemiology, 1995-2007", *Infection control and hospital epidemiology : the official journal of the Society of Hospital Epidemiologists of America,* vol. 31, no. 4, pp. 348-356.

Smith, P.A., Roberts, T.C. & Romesberg, F.E. 2010, "Broad-spectrum antibiotic activity of the arylomycin natural products is masked by natural target mutations", *Chemistry & biology,* vol. 17, no. 11, pp. 1223-1231.

Suzuki, T., Itoh, A., Ichihara, S. & Mizushima, S. 1987, "Characterization of the sppA gene coding for protease IV, a signal peptide peptidase of Escherichia coli", *Journal of Bacteriology,* vol. 169, no. 6, pp. 2523-2528.

Tettelin, H., Masignani, V., Cieslewicz, M.J., Eisen, J.A., Peterson, S., Wessels, M.R., Paulsen, I.T., Nelson, K.E., Margarit, I., Read, T.D., Madoff, L.C., Wolf, A.M., Beanan, M.J., Brinkac, L.M., Daugherty, S.C., DeBoy, R.T., Durkin, A.S., Kolonay, J.F., Madupu, R., Lewis, M.R., Radune, D., Fedorova, N.B., Scanlan, D., Khouri, H., Mulligan, S., Carty, H.A., Cline, R.T., Van Aken, S.E., Gill, J., Scarselli, M., Mora, M., Iacobini, E.T., Brettoni, C., Galli, G., Mariani, M., Vegni, F., Maione, D., Rinaudo, D., Rappuoli, R., Telford, J.L., Kasper, D.L., Grandi, G. & Fraser, C.M. 2002, "Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V Streptococcus agalactiae", *Proceedings of the National Academy of Sciences of the United States of America,* vol. 99, no. 19, pp. 12391-12396.

Tjalsma, H., Noback, M.A., Bron, S., Venema, G., Yamane, K. & van Dijl, J.M. 1997, "Bacillus subtilis contains four closely related type I signal peptidases with overlapping substrate specificities. Constitutive and temporally controlled expression of different sip genes", *The Journal of biological chemistry,* vol. 272, no. 41, pp. 25983-25992.

Tokuda, H. 2009, "Biogenesis of outer membranes in Gram-negative bacteria", *Bioscience, biotechnology, and biochemistry,* vol. 73, no. 3, pp. 465-473.

Tschantz, W.R., Paetzel, M., Cao, G., Suciu, D., Inouye, M. & Dalbey, R.E. 1995, "Characterization of a soluble, catalytically active form of Escherichia coli leader peptidase: requirement of detergent or phospholipid for optimal activity", *Biochemistry,* vol. 34, no. 12, pp. 3935-3941.

UniProt Consortium 2014, "Activities at the Universal Protein Resource (UniProt)", *Nucleic acids research,* vol. 42, no. 1, pp. D191-8.

Valent, Q.A., Scotti, P.A., High, S., de Gier, J.W., von Heijne, G., Lentzen, G., Wintermeyer, W., Oudega, B. & Luirink, J. 1998, "The Escherichia coli SRP and SecB targeting pathways converge at the translocon", *The EMBO journal,* vol. 17, no. 9, pp. 2504-2512.

Van den Berg, B., Clemons, W.M.,Jr, Collinson, I., Modis, Y., Hartmann, E., Harrison, S.C. & Rapoport, T.A. 2004, "X-ray structure of a protein-conducting channel", *Nature,* vol. 427, no. 6969, pp. 36-44.

van der Wal, F.J., Oudega, B., Kater, M.M., ten Hagen-Jongman, C.M., de Graaf, F.K. & Luirink, J. 1992, "The stable BRP signal peptide causes lethality but is unable to provoke the translocation of cloacin DF13 across the cytoplasmic membrane of Escherichia coli", *Molecular microbiology,* vol. 6, no. 16, pp. 2309-2318.

van Roosmalen, M.L., Geukens, N., Jongbloed, J.D., Tjalsma, H., Dubois, J.Y., Bron, S., van Dijl, J.M. & Anne, J. 2004, "Type I signal peptidases of Gram-positive bacteria", *Biochimica et biophysica acta,* vol. 1694, no. 1-3, pp. 279-297.

von Heijne, G. 1985, "Signal sequences. The limits of variation", *Journal of Molecular Biology,* vol. 184, no. 1, pp. 99-105.

Vrontou, E. & Economou, A. 2004, "Structure and function of SecA, the preprotein translocase nanomotor", *Biochimica et biophysica acta,* vol. 1694, no. 1-3, pp. 67-80.

Waite, R.D., Rose, R.S., Rangarajan, M., Aduse-Opoku, J., Hashim, A. & Curtis, M.A. 2012, "Pseudomonas aeruginosa possesses two putative type I signal peptidases, LepB and PA1303, each with distinct roles in physiology and virulence", *Journal of Bacteriology,* vol. 194, no. 17, pp. 4521-4536.

Wang, P., Shim, E., Cravatt, B., Jacobsen, R., Schoeniger, J., Kim, A.C., Paetzel, M. & Dalbey, R.E. 2008, "Escherichia coli signal peptide peptidase A is a serine-lysine protease with a lysine recruited to the nonconserved amino-terminal domain in the S49 protease family", *Biochemistry,* vol. 47, no. 24, pp. 6361-6369.

Weihofen, A. & Martoglio, B. 2003, "Intramembrane-cleaving proteases: controlled liberation of proteins and bioactive peptides", *Trends in cell biology,* vol. 13, no. 2, pp. 71-78.

ZIMM, B.H. 1948, "The dependence of the scattering of light on angle and concentration in linear polymer solutions", *The Journal of physical and colloid chemistry,* vol. 52, no. 1, pp. 260-267.

Zwizinski, C., Date, T. & Wickner, W. 1981, "Leader peptidase is found in both the inner and outer membranes of Escherichia coli", *The Journal of biological chemistry,* vol. 256, no. 7, pp. 3593-3597.

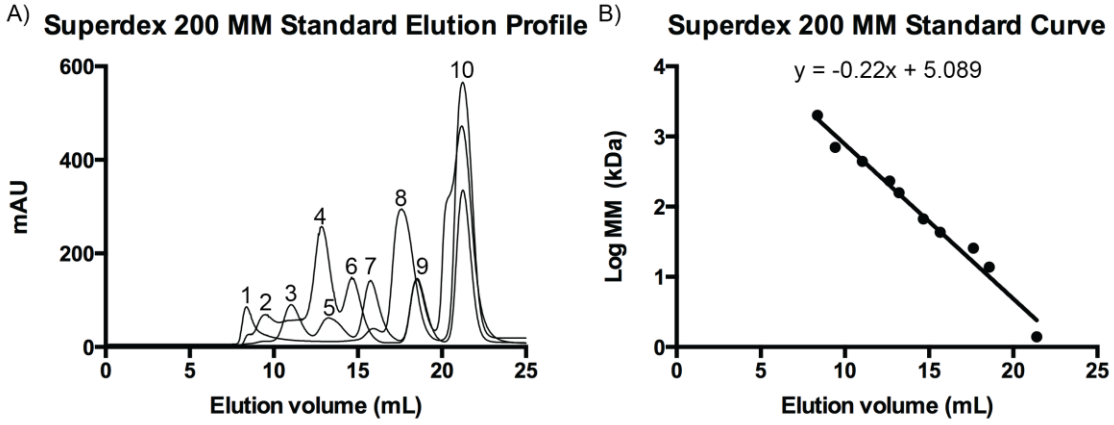
Zwizinski, C. & Wickner, W. 1980, "Purification and characterization of leader (signal) peptidase from Escherichia coli", *The Journal of biological chemistry,* vol. 255, no. 16, pp. 7973-7977.

Appendices

Appendix A.

Superdex 200 size exclusion column calibration

The Superdex 200 column was calibrated using high molecular weight (HWM) gel filtration calibration kit and low molecular weight (LWM) gel filtration calibration kit (GE Healthcare catalogue #: 28-4038-41 and 28-4038-41). Each protein standard was dissolved in calibration buffer (10mM Tris pH7.4, 150mM NaCl, 1mM EDTA, 2mM DTT) to a concentration of 5mg/mL. Protein standard solutions were centrifuged (16,500xg, 20min, 4°C) to remove any protein aggregates and 100uL of protein standard solution was loaded onto the Superdex 200 column. The Superdex 200 column was connected to an ÄKTA FPLC system and run with a flow rate of 0.5mL/min, 0.5mL fraction size and calibration buffer as the mobile phase. The table below listed the different protein standards used to calibrate the Superdex 200 column and their respective molecular mass. Panel A shows the absorbance at 280nm in *milli*-absorbance unit (mAU) for each standard protein as they come off the Superdex 200 column. Each peach is labeled with a number from 1 to 10 which corresponds to each individual standard protein. Panel B shows the standard curve from the Log of molecular mass (kDa) of each stand protein versus their respective elution volume (mL). The equation for the line of best fit is shown with a R^2 value of 0.98.

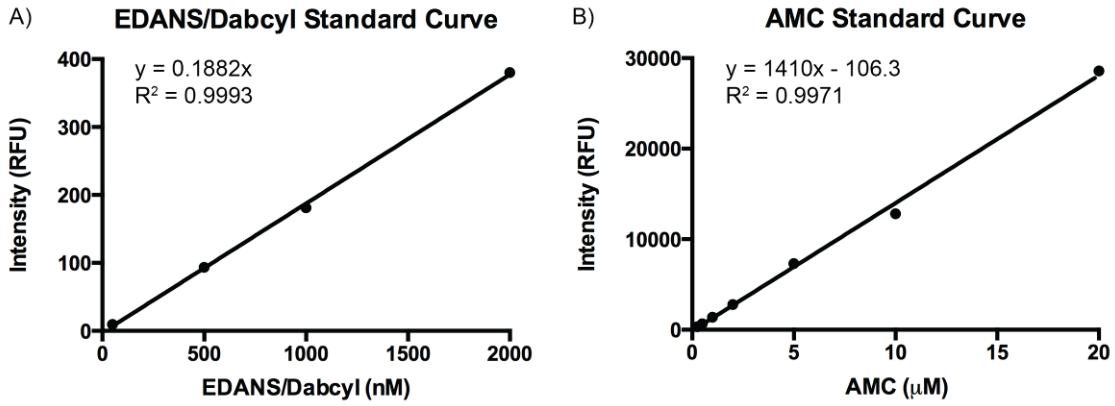


Molecular mass standard (kDa)		Elution volume (mL)
1	Dextran (2000)	8.37
2	Thyroglobulin (669)	9.42
3	Ferritin (440)	11.04
4	Catalase (232)	12.66
5	Aldolase (158)	13.22
6	Albumin (67)	14.66
7	Ovalbumin (43)	15.66
8	Chymotrypsinogen (25.7)	17.63
9	Ribonuclease A (13.7)	18.57
10	Vitamin B12 (1.4)	21.4

Appendix B.

Standard curve for converting fluorescence into molar units

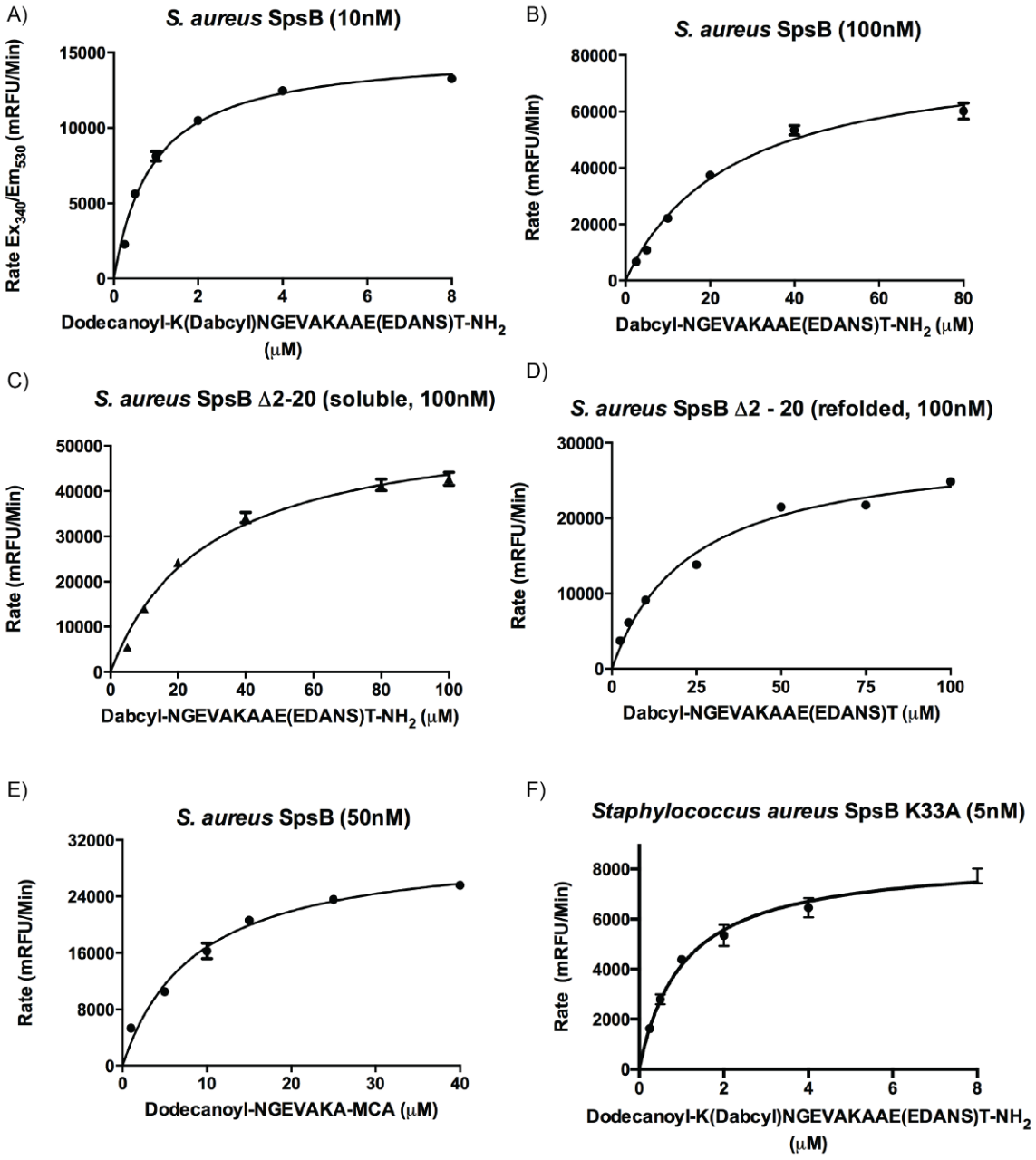
These following standard curves were generated with EDANS and Dabcyl or MCA at various concentrations in kinetic buffer (20mM Tris pH8.0, 100mM NaCl, 0.01% DDM). Relative fluorescence intensity (RFU) for EDANS/Dabcyl was measured at 50, 500, 1000 and 2000nM using Ex 340nm and Em 530nm on a SpectraMax M5 (Molecular Devices) instrument. RFU for AMC was taken at 0.25, 0.5, 1, 2.5, 5, 10 and 20µM using Ex 380 and Em 460. The equation for the line of best fit and the R-square value for each curve are shown.

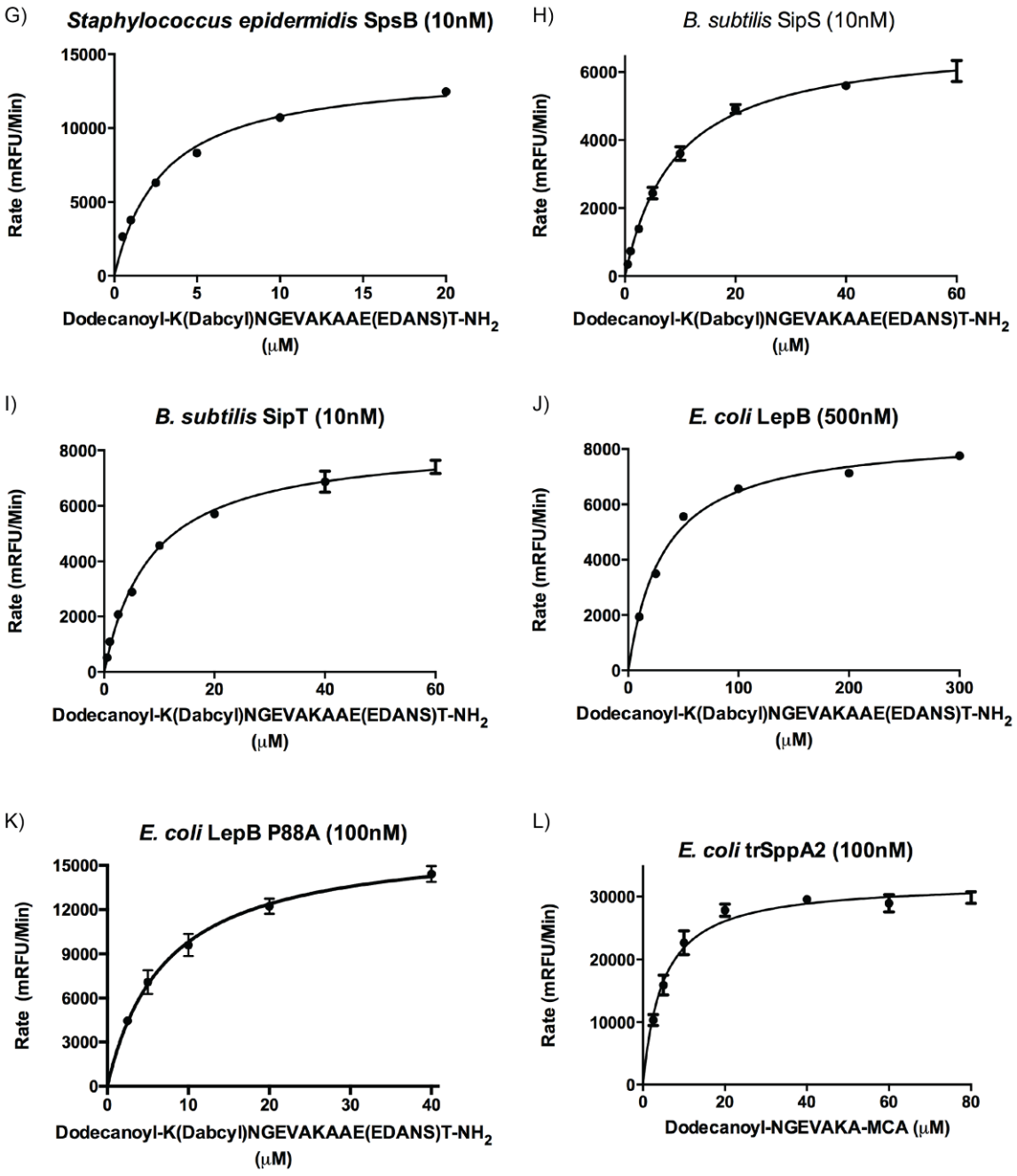


Appendix C.

Initial velocity vs [S] kinetic curves for each enzyme and substrate combination examined in this thesis work

Each graph shown is generated from triplicate experiments plotted using Prism5 software (GraphPad). A sample kinetic constants calculation for A) *S. aureus* SpsB with decanoyl-K(Dabcyl)NGEVAKAAE(EDANS)T-NH_2 is shown.





Vmax and K_M can be obtained by doing a non-linear regression (Michaelis-Menten) analysis on graph A) using Prism 6 (GraphPad), Vmax of 15124 mRFU/Min and a K_M 0.92μM were obtained. k_{cat} was calculated from Vmax using the standard curve from Appendix B.

$$\frac{\frac{15124mRFU}{Min} x \frac{1RFU}{1000mRFU} x \frac{1Min}{60s} x \frac{1nM}{0.1882RFU}}{10nM\ (total\ enzyme)} = 0.13s^{-1}$$