Crystallization of *Bacilius subtilis* Type I Signal Peptidase S (SipS)

by

Eugean Shin

B.Sc., Simon Fraser University, 2013

Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

in the

Department of Molecular Biology and Biochemistry

Faculty of Science

Approval

Name: **Eugean Shin**

Degree: **Master of Science**

Title: **Crystallization of *Bacilius subtilis* Type I Signal Peptidase S (SipS)**

Examining Committee: **Chair:** Dr. Nancy Forde
                                                  Associate Professor

Dr. Mark Paetzel
Senior Supervisor
Professor

_____

Dr. Lisa Craig
Supervisor
Associate Professor

_____

Dr. Edgar C. Young
Supervisor
Associate Professor

_____

Dr. Frederic Pio
Internal Examiner
Associate Professor

_____

Date Defended/Approved: March 30, 2015 _____

Partial Copyright Licence

SFU

Abstract

Gram-positive *Bacillus subtilis* (*B. subtilis*) signal peptidase I (SPase I) is a membrane-bound endopeptidase that cleaves off the amino-terminal signal peptide from pre-proteins before or after their translocation across the cytoplasmic membrane. *B. subtilis* has five chromosomal SPase I; SipS, SipT, SipU, SipV, SipW, and two plasmid encoded paralogous SipP. SipS is one of major SPase I in this species, which are essential for cell viability. It is also one of the closest of the *B. subtilis* SPase I enzymes in sequence identity to the well characterized Gram-negative *E. coli* SPase I. As a result, SipS was chosen for this research study.

B. subtilis SipS uses a Ser/Lys catalytic dyad for catalytic activity, utilizing Ser43 as the nucleophile and Lys83 as the general base. The constructs - SipS Full-Length (FL), SipS Δ2-35 Wild-Type (WT), SipS Δ2-35 S43A and SipS Δ2-35 K83A – were expressed, purified, and screened for crystallization conditions. Catalytically active SipS Δ2-35 WT formed needle shaped crystal clusters whereas SipS Δ2-35 K83A produced initial hits in crystallization conditions containing lithium sulfate.

Preliminary data for the catalytic activity of *B. subtilis* SipS Δ2-35 WT shows that hexaaminecobalt (III) chloride inhibits the enzyme.

Keywords: signal peptide processing; signal peptidase; Ser/Lys dyad catalysis; x-ray crystallization; inhibitor

To my parents

Acknowledgements

First of all, I would like to express my sincere gratitude towards my supervisor, Dr. Mark Paetzel, for giving me the wonderful opportunity to work and learn in his lab. His passion and his knowledge of crystallography have been very motivating and inspiring. I would also like to thank my committee members, Drs. Lisa Craig and Edgar Young, for their participation in overseeing my research project and their valuable suggestions. I am grateful towards Drs. Frederic Pio and Nancy Forde who agreed to be my internal examiner and chair, respectively.

I do not know how to thank Deidre de Jong-Wong, who is more than just the lab manager. Not only did she make sure all the lab supplies were available when necessary and kept everything in the lab organized, she also provided support and advice when I needed it. This thesis would not have been possible without her help. Thank you Chuanyun Luo for your lovely smile and your happiness that brightened up the lab, and Daniel Chiang for being the best mentor and a great entertainer in the lab. Thanks to Zohreh Sharafianardakani who has shared the Master's experience with me and who constantly provided me with knowledge and hope.

I would like to thank my past lab members and friends, Drs. Kelly Kim, Charlie Stevens, Ivy Chung, Apollos Kim, Sung-Eun Nam, Linda Zhang and Jae-Young Lee. Thanks also to Minfei Fu, Michael Ungerer, Mark Dennison and Kwangjin Park for their friendship and for making the lab a welcoming place. Much appreciation goes to Sunny for his continual support and companionship throughout my studies.

Lastly, I want to deeply thank my family who never gave up on me, taught me their faith and strong love in God, and supported me throughout all these years with love and care. Thank you all.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ATP   Adenosine Triphosphate

AU   Asymmetry Unit

DMSO   Dimethyl Sulphoxide

FL   Full-Length

IPTG   Isopropyl-β-D-thiogalactopyranoside

PAGE   Polyacrylamide Gel Electrophoresis

PCR   Polymerase Chain Reaction

PDB   Protein Data Bank

PEG   Polyethylene Glycol

P_i Inorganic phosphate

RCSB   Research Collaboration for Structural Bioinformatics

RMSD   Root Mean Square Deviation

SDS   Sodium Dodecyl Sulphate

SP   Space Group

SRP   Signal Recognition Particle

WT   Wild-Type

Glossary

Å	Ångströms; 1Å equals to 10^{-10} m.
Apo-structure	A protein structure without substrates or inhibitors bound to the active sites.
Asymmetry unit	The fundamental, smallest unit that is repeated in crystals of macromolecules to which symmetry operations can be applied to create the complete unit cell. Symmetry operations include rotations, translations and screw axes.
Cohex	Hexaaminecobalt (III) chloride
Column volume	The total volume of a chromatography column; the sum of void volume and the matrix volume.
Competitive Inhibitor	A substance that inhibits an enzyme's function by competing with the substrates for the enzyme's active site.
Crystal	An ordered, three-dimensional, repeating molecular array.
Inhibitor	A substance that stops or slow down the catalytic activity of an enzyme.
Irreversible Inhibitor	A type of inhibitor that irreversibly binds to an enzyme making the inhibition permanent.
kDa	Kilodalton, a unit of molecular mass of proteins.
K_M	Concentration of the substrate at half V_{max}. This indicates the binding affinity between an enzyme and a substrate. A low K_M indicates high affinity of an enzyme to a substrate.
Mobile phase	In chromatographic separation, the mobile phase is the fluid phase that carries the mixture of proteins to be purified through the system.
Nucleophile	A chemical molecule with a free pair of electrons which can form a bond with an electrophile by donating these electrons.
Oxyanion hole	A group of residues that stabilize the tetrahedral transition states by providing hydrogen bonds to the negatively charged carbonyl oxygen of the P1 residue in the substrate of a protease.
pK_a	It can be expressed as -$\log_{10}$ of K_a which is an acid dissociation constant. It is a quantitative measure of the strength of an acid in solution and measure how readily and fast the acid give up the protons in solution. The lower the pK_a value is, the stronger the acid is.

Pre-protein	A precursor protein with an N-terminal signal peptide which is to be targeted to the translocation complex in the membrane. The signal peptide is cleaved off by signal peptidase to form the mature protein.
Pro-protein	An inactive precursor protein which is to be activated after post-translational modification (cleavage).
Refinement	The process of changing a model's parameters in order to optimize a function. e.g. Fit the electron density in a crystal structure.
Reversible Inhibitor	A type of inhibitor that binds to enzyme in a reversible manner; it can detach itself from the enzyme after it binds, allowing the enzyme to become functional once again.
Scissile bond	A peptide bond destined to be broken after a catalytic attack, nucleophile.
SEC	Size exclusion chromatography; also known as gel filtration chromatography. This type of column chromatography separates proteins based on their Stoke's radius. The larger the radius the protein has, the faster it is eluted from the column.
SipS	Signal Peptidase Type I S (*B. subtilis*)
SipS full-length S151A	Construct of *B. subtilis* SipS with the serine residue at 151 mutated to alanine
SipS full-length S151C	Construct of *B. subtilis* SipS with the serine residue at 151 mutated to cysteine
SipS Δ2-35	Construct of *B. subtilis* SipS lacking residues 2 to 35 corresponding to the transmembrane domain.
SipS Δ2-35 S43A	Construct of *B. subtilis* SipS Δ2-35 with the catalytic serine residue at 43 mutated to alanine.
SipS Δ2-35 K83A	Construct of *B. subtilis* SipS Δ2-35 with the catalytic lysine residue at 83 mutated to alanine.
Space group	The combination of symmetry operations that characterize a crystal.
SPase I	Type I Signal Peptidase (Bacterial)
Specificity constant	k_{cat}/k_m; measures the efficiency of a catalytic enzyme
Structure factor	The structure factor F_{hkl} is a mathematical function describing the amplitude and phase of a wave diffracted from crystal lattice planes characterised by Miller indices h,k,l.
Tris	Tris(hydroxymethyl) methylamine.
Transition state	The unstable, relatively short-lived state that corresponds to the highest potential energy in the catalyzed reaction coordinates.

Turnover number	k_{cat} the number of substrate molecules that an enzyme can convert to a product per unit time.
Unit cell	The smallest repeating unit of the crystal.
Uncompetitive inhibitor	An inhibitor that demolishes an enzyme's functionality by binding to a site distant from the substrate's binding pocket thereby changing the enzyme's conformation.
V_{max}	The maximum rate at which an enzyme can catalyze a reaction.
Void volume	The elution volume of large molecules that are excluded from the gel filtration medium which pass straight through the packed beads.

Chapter 1.

Introduction

Bacillus subtilis Type I Signal Peptidase S (*B. subtilis* SipS) is a protein in Gram-positive bacteria that catalytically cleaves a signal peptide from the pre-protein. This chapter describes the function and characteristics of the Type I Signal Peptidase (SPase I) enzyme in both Gram-negative and Gram-positive bacteria, the significance of *B. subtilis* SipS and the objectives of this research project.

1.1. The Gram-Positive Bacterial Cell Membrane

Every prokaryotic cell is surrounded by a cell membrane which provides a scaffold and protects the cell's contents from the outside environment. Each layer of the cell membrane is crucial for cell viability because they provide the specific structure and environment necessary for the functioning of important proteins. For example, essential protein complexes associated with protein secretion and transport are embedded in the cell membranes.

The cell membranes of Gram-positive and Gram-negative bacteria are slightly different from each other. Gram-positive bacteria have a cell wall composed of three domains: a plasma membrane that encloses the cytosol (Figure 1-1), a periplasmic space in the middle (Figure 1-1) and a thick peptidoglycan layer that faces the extracellular environment (Figure 1-1) (Auclair et al., 2012). Unlike Gram-negative bacteria, the peptidoglycan layer of Gram-positive bacteria contains anionic polymers such as teichoic and teichuronic acid. Together, the peptidoglycan layer of Gram-positive bacteria maintains structure integrity (Gan et al., 2008; Silhavy et al., 2010;

1

Stewart 2005). However, Gram-negative bacteria have a more complicated membrane structure, with a plasma membrane containing the cytosol (Figure 1-1), a relatively thin peptidoglycan layer (Figure 1-1), and an outer membrane (Figure 1-1) facing the cell's exterior. Gram-positive bacteria are known to produce a larger amount of proteins compared to Gram-negative bacteria, with their simpler membrane structures (Nagarajan, 1993; Simonenl & Palva, 1993).
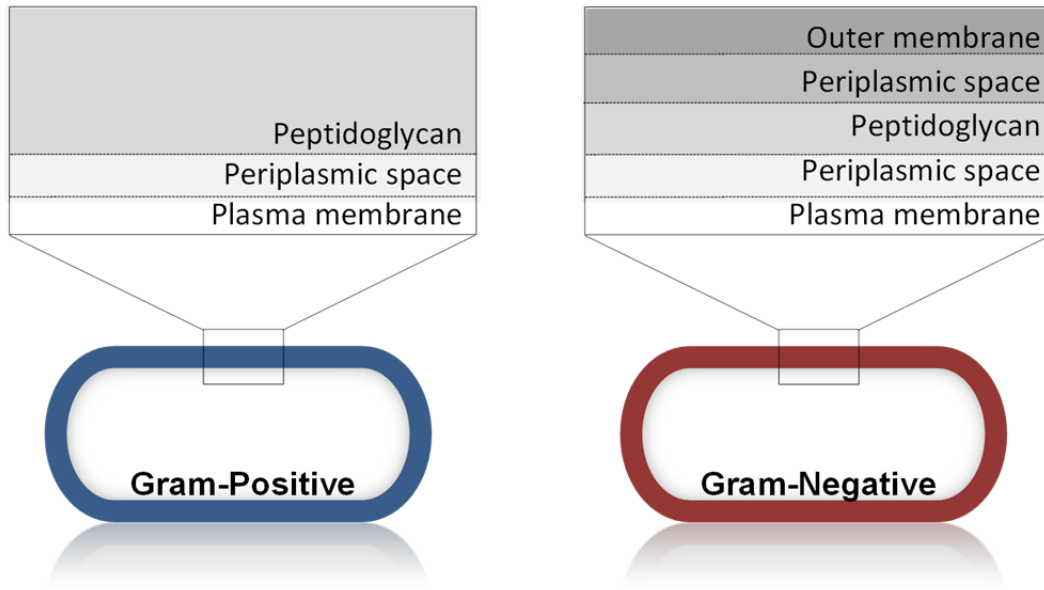
Figure 1-1   Schematic diagram of the compartments of Gram-positive and Gram negative bacterial cell membranes.

The Gram-positive (blue) cell has a cell membrane composed of three layers (A1-A3). The membrane that encloses the cytosol is the plasma membrane (A1), the layer in the middle is the periplasmic space (A2), and the outer layer facing the extracellular environment is a thick peptidoglycan layer (A3). The Gram-negative cell (red) has five separate layers in its cell membrane (B1-B5). The plasma membrane (B1) encloses the cytosol, and the first periplasmic space (B2) encircles the plasma membrane. The next layer is the relatively thin peptidoglycan layer (B3), followed by the second periplasmic space (B4). These stacked layers are protected from the extracellular environment by the outer membrane (B5).

1.2. **Protein Secretion and Export in *B. subtilis***

1.2.1. **Localization and Functions of Secreted Proteins in *B. subtilis***

There are four possible destinations for secreted proteins in *B. subtilis*: the cytoplasm, the plasma membrane, the peptidoglycan cell wall layer and the extracellular environment. Cytoplasmic proteins do not have transport signals and hence are retained in the cytoplasm in their native forms. Proteins that have transport signals are transported across the cytoplasmic membrane to reach their final destinations, either the cell wall or the extracellular environment (von Heijne 1990; Tjalsma et al. 2000; Watson 1984). Proteins with uncleavable transport signal are mainly membrane-embedded proteins. They use their transport signals as targeting signals to direct them to the membrane, but also use them as membrane anchor once they reach the destination(Tjalsma et al., 2000). Because *B. subtilis* does not have the outer membrane that is found in Gram-negative bacteria, the proteins that are transported across the plasma membrane are released into the extracellular environment or remain in the peptidoglycan cell wall (Gould et al., 1975).

Proteins with transport signals usually have amino-terminal signal peptides and are called cell wall proteins or extracellular enzymes depending on their localization. There are 12 resident cell wall proteins in the peptidoglycan layer of *B. subtilis* (Pooley et al., 1996; Tjalsma et al., 2000). Most of them are autolytic enzymes that degrade the cell wall when necessary to allow cell growth and cell division (Blackman et al., 1998; Foster, 1992, 1993; Simonenl & Palva, 1993). The extracellular enzymes that *B. subtilis* secretes into the extracellular environment are primarily degradative enzymes, including proteases, lipases, carbohydrases, DNases and RNases (Simonenl & Palva 1993; Tjalsma et al. 2000). These enzymes are involved in the hydrolysis and depolymerization of natural substances which is why many of them have been marketed commercially. In contrast, membrane-embedded proteins play significant roles in maintaining cell integrity, signal sensing and transduction, transport processes, and energy generation and conservation (Raetz & Dowhan 1990; Silhavy et al., 2010). Despite their significance in cellular function, these proteins have been more difficult to

analyze due to their hydrophobicity. Approximately 700 membrane proteins have been identified in *B. subtilis* (Tjalsma et al., 2000).

Proteins with transport signals are secreted by one of three pathways in *B. subtilis*: the twin-arginine translocation (Tat) pathway, the Sec-dependent pathway, and the pseudopilin (Com) pathway. Protein complexes involved in each pathway are distinguished based on their signal peptide specificity, as well as their translocation mechanism. Of the pathways, the Sec-dependent pathway handles most of the enzymes in *B. subtilis*, approximately 4000 proteins, including extracellular proteins, cell wall proteins, and membrane-bound proteins (Dalbey et al., 2012; Tjalsma et al. 2004).

The *B. subtilis* Sec-dependent pathway utilizes two main protein secretion routes: the co-translational route and the post-translation route. Both of these routes have similar mechanisms utilizing: cytosolic chaperones, the translocation motor and the ATPase SecA, translocation channel SecYEG, signal peptidases (SPases), signal peptide peptidases (SPPases), and protein folding factors that act in the extracellular environment (Figure 1-2) (Driessen & Nouwen, 2008). However, the co-translational route is different from the post-translation route in that it transports the pre-proteins that are in the process of translation rather than the fully synthesized pre-proteins (Park & Rapoport, 2012).

The process of co-translational protein secretion begins in the cytoplasm of the cell as the precursor protein, a protein with a signal peptide attached to its N-terminus, is being translated from ribosomeand synthesized. For proteins that need to be guided out of the cytosol, a cytosolic chaperone called the Signal Recognition Particle (SRP) recognizes the pre-protein during the translation process (Honda et al., 1993) (Figure 1-2). This action prevents the pre-protein from aggregating and misfolding (Zanen et al., 2006). The SRP-pre-protein complex then binds to SecA which is, in turn, binds to one of the translocation channels, SecY (Duong & Wickner, 1997; Lill et al., 1990; Wiech et al., 1991). The ATPase activity of SecA is activated as when it recognizes the SRP-pre-protein complex in the cytoplasm (Hartl et al., 1990). The SRP-pre-protein complex is then carried on to the translocation channels SecYEG where SRP detaches from the

pre-protein once SecA hydrolyzes ATP to ADP and P_i (Economou & Wickner, 1994; Schiebel et al., 1991). SecYEG recognizes the N-terminal signal peptide on the substrate and subsequently transports the pre-protein out of the membrane starting from the N-terminal(Wiech et al., 1991). The signal peptide laterally diffuses out of the channel, so that it can be cleaved by SPase located nearby. The SPase catalytically cleaves a scissile bond, detaching the signal peptide from the pre-protein. The translocation continues and the pre-protein is released onto the periplasmic space as a fully folded, mature form.

The post-translational route is similar to the co-translational one, except that the pre-protein binds to a cytoplasmic chaperone after it has been fully synthesized (Park & Rapoport, 2012). After the fully translated pre-protein leaves the ribosome, it binds to a chaperone CsaA protein, which is a functional homolog of SecB of *E.coli* (Bechtluft et al., 2010; Kawaguchi et al., 2001; Kumamoto, 1989). Csa A plays the same role as SRP in the co-translational secretion pathway.

Studies have found that membrane-embedded proteins are utilized only by Sec-dependent pathway while the extracellular enzymes utilize both the Sec-dependent and the Tat pathway for translocation. Cell wall proteins utilize the Sec-dependent pathway most of the time, but also use Com pathway which was not discussed in this thesis (Tjalsma, Bolhuis, et al., 2000).

While the Sec-dependent pathway translocates only the unfolded pre-proteins, the Tat pathway transports the fully folded pre-proteins in the cytoplasm across or into the cytoplasmic membrane (Gohlke et al., 2005). Only three substrates have been identified for the Tat translocase in *B. subtilis*, which is a very small number compared to the substrates identified for the Sec translocase (Yuan et al., 2010). The Tat translocase in *B. subtilis* is composed of proteins TatA and TatC (Monteferrante et al., 2012). TatC is assumed to be involved in signal peptide recognition whereas TatA acts as a pore/channel for protein translocation across the membrane. *B. subtilis* does not have the TatB protein. In *E. coli*, TatB transfers the pre-protein from TatC to TatA, but the studies suggest that TatA in *B. subtilis* performs both its function and TatB's

(Barnett et al., 2008; Jongbloed et al., 2004; Yuan et al., 2010). The exact mechanism of the Tat pathway in *B. subtilis* is still being studied.
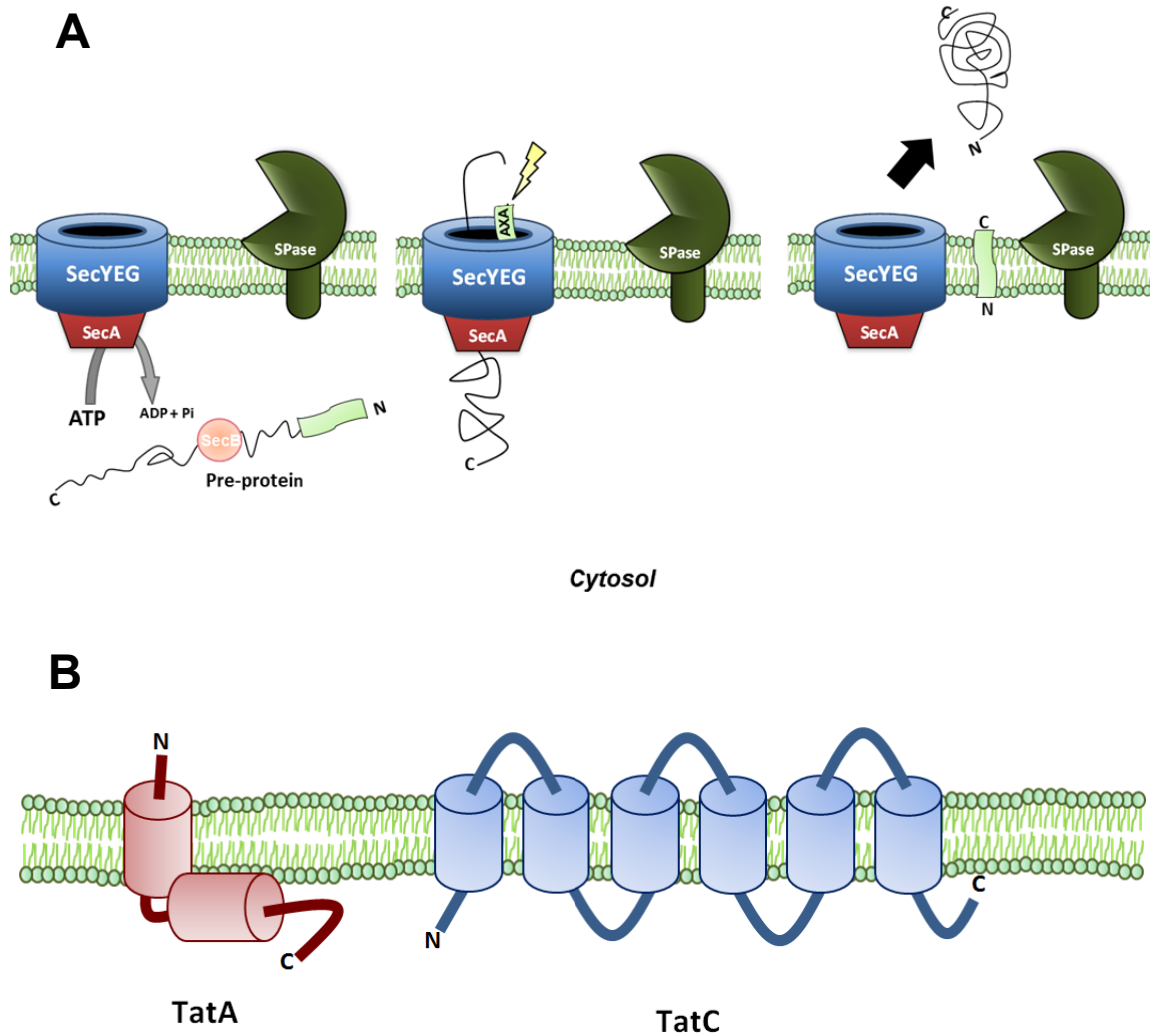
Figure 1-2   The protein translocase complexes

(A)The Sec-dependent pathway secretes and transports the pre-protein across the cytoplasmic membrane using the translocation channels SecYEG. Once SRP binds to the pre-protein during its synthesis, the SRP-pre-protein complex is delivered to SecA which is physically attached to the translocation channels. SecA hydrolyzes ATP into ADP and P_i, releasing the pre-protein from SRP. The pre-protein is then delivered to the channels and is translocated across the membrane. The signal peptide attached at the amino-terminus of the pre-protein is cleaved by SPase during or after the translocation of the pre-protein and the cleaved signal peptide is degraded by SPPase. (B) The Tat translocase in *B. subtilis* is composed of TatA and TatC. TatC recognizes the signal peptide of the pre-protein, which is subsequently transferred to TatA and is transported out of the cytoplasmic membrane.

1.2.2. Signal Peptidase and Signal Peptide Peptidase

Using the transport pathways described in the preceding section, non-cytosolic proteins are transported across the membrane with the help of a short stretch of amino acids at the amino terminus of the polypeptide, called a signal peptide. This signal peptide contains a specific sequence which is recognized by SPase and directs the precursor protein to its final destination (Dalbey et al. 1997; Dalbey & von Heijne 1992). Anchored to the cytoplasmic membrane, SPase recognizes the Ala-X-Ala sequence located at the C-terminal end of the signal peptide and initiates the translocation process of the precursor protein across the membrane (Tjalsma et al., 2000). During or shortly after translocation, the signal peptide is removed from the precursor protein by catalytic cleavage by SPase. As a result of this cleavage, the mature protein is released on the *trans*-side of the membrane in its properly folded, active form (Bolhuis et al., 1996). The remaining signal peptide in the membrane is subsequently degraded by signal peptide peptidase (SPPase) (Hussain et al., 1982).

1.2.3. Signal Peptide

Based on their secretion pathway, there are three main types of signal peptides: Sec-type signal peptides, Tat-type signal peptides and lipoprotein signal peptides(Date & Wickner, 1981; Hussain et al., 1982; Tokunaga et al., 1982). All types of signal peptides have about 28 amino acid residues, containing an N-terminal domain, a hydrophobic core (H-domain), and a C-terminal domain (Briggs et al., 1986; Emr & Silhavy, 1983; von Heijne, 1990). The N-terminal domain consists of about one to five positively charged amino acid residues which interact with the negative charged phospholipids in the lipid bilayer of the *cis*-side of membrane during translocation. The H-domain has a hydrophobic patch of seven to 15 residues in an α-helical conformation. This region interacts with the hydrophobic cores of the membrane to facilitate translocation. Following the H-domain, the C-terminal domain of the signal peptide forms an extended β-sheet structure which helps in binding to the binding pocket of the SPases. As well, the C-terminal domain contains both the recognition and cleavage sites for the SPases.

Among the 180 pre-protein substrates with signal peptides in *B. subtilis*, 166 substrates are predicted to have a Sec-type signal peptide with an average length of 28 residues (Tjalsma et al., 2000). The N-terminal domain of the Sec-type signal peptide contains two or three positively charged amino acids (Lys and Arg) while the hydrophobic core (H-domain) has an average of 19 small, hydrophobic residues. The C-terminal domain contains SPase recognition sites at P1 and P3 (-1 and -3) with a consensus sequence of Ala-X-Ala (Figure 1-3) (Fikes et al., 1990; Simonenl and Palva, 1993).

The Tat-type signal peptide contains an RR-motif in its N-terminal domain with a sequence of Arg-Arg-X-#-#, where # represents a hydrophobic residue. Among all the SPase I substrates in *B. subtilis*, there are 14 predicted pre-proteins with Tat-type signal peptides (Tjalsma et al., 2000). The H-domain of the Tat-type signal peptide is very similar to the Sec-type signal peptide in *B. subtilis*, but it differs from the *E. coli* signal peptide in that it is slightly longer and less hydrophobic (Cristóbal et al., 1999).

The lipoprotein signal peptide is relatively shorter than the non-lipoprotein signal peptides and is unique, having a lipobox at the C-terminal domain with the consensus sequence Leu-(Ala/Ser)-(Ala/Gly)-Cys at P3-P2-P1-P1' (Braun, 1975; Hayashi & Wu, 1990; Sankaran & Wu, 1994; von Heijne, 1990). The lipoprotein signal peptide is similar to the Sec-type signal peptide, with both of the pre-proteins being translocated using the Sec translocase. However, once the pre-protein with the signal peptide is translocated, the catalytic cleavage of the lipoprotein signal peptide is processed by signal peptidase II (SPase II) rather than SPase I.
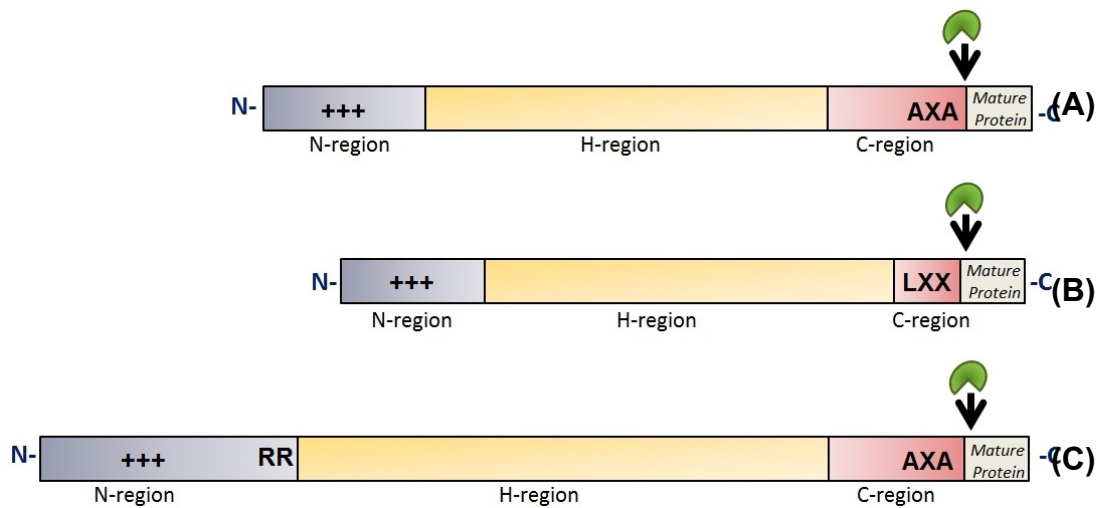
Figure 1-3   The features of a bacterial signal peptide.

Listed are (A) Sec-type signal peptide, (B) Lipoprotein signal peptide and (C) Tat-type signal peptide. The N-region is composed of positively charged residues while the H-region has a hydrophobic patch in an α-conformation. The C-region has recognition sites at the -1 and -3 positions with the "Ala-X-Ala" sequence preferred and catalytically cleaved by SPase I. The signal peptide is followed by the mature protein.

1.3. *E. coli* Type I Signal Peptidase

SPase I is a highly conserved enzyme that is commonly found in humans, yeasts, archaea and bacteria. It is essential for most bacteria's viability as its absence results an accumulation of pre-proteins in the cell membranes, leading to cell death (Cregg et al., 1996; Koshland & Sauer 1982; Lammertyn 2004; Nahrstedt et al. 2004; Taheri et al., 2010; Zhbanko et al., 2005)

Bacterial SPase I is different from most serine proteases; it utilizes a Ser/Lys dyad mechanism for catalytic cleavage rather than the classical Ser/His/Asp catalytic triad. Sequence alignments of bacterial SPases I reveal that while the enzymes do not share high levels of identity, they still maintain five similar conserved regions, Boxes A, B, C, D and E (van Dijl et al. 1992; Paetzel et al. 2002; Tjalsma et al. 1998). Box A is an N-terminal transmembrane region while boxes B, C, D and E are soluble, catalytic regions. Box B contains a nucleophilic serine residue (Ser91 in *E. coli* SPase I and

9

Ser43 in *B. subtilis* SipS) and Box D has a general base lysine residue (Lys146 in *E. coli* SPase I and Lys83 in *B. subtilis* SipS) which comprise the catalytic dyad.

E. coli SPase I is the most characterized SPase I in terms of function and structure. Also called leader peptidase (Lep), it is encoded by the lepB gene (Date & Wickner 1981; Wolfes et al., 1983) and is an essential enzyme for cell growth and viability. There is only one chromosomal SPase I in *E. coli* (Dalbey & Wickner, 1985). The enzyme is known to act as a monomer during the cleavage reaction and is susceptible to auto-digestion (Auclair et al., 2012; Tschantz et al. 1995).

This enzyme is a membrane-embedded serine endopeptidase and consists of 323 amino acids (35.9kDa). *E. coli* SPase I has a short periplasmic amino-terminal region (residues 1-3), followed by a transmembrane region (residues 4-28), a cytosolic loop (residues 29-57), another transmembrane segment (residues 58-76), and a large C-terminal periplasmic domain (residues 77-323) which contains the active site (Wolfes et al., 1983). Membrane topology studies have revealed that the C-terminal side of the second transmembrane domain is located in the periplasm as is the short periplasmic amino-terminal region (Figure 1-5) (Luis et al. 1989; Moore & Miurae 1987; Wolfes et al., 1983). The first transmembrane segment and the cytoplasmic domain have been shown to not be involved in the catalytic activity of the enzyme (Bilgin et al., 1990; Carlos et al., 2000).

```
            20        30        40        50        60
SipS       LEWAKAIVIAVVLALLIRNFIFAPYVVDGDSMYPTLHNRERVFVNMTVK...........
SipT       LEWGKAIVIAVLLALLIRNFLFEPYLVEGSSMYPTLHDGERLFVNKTVN...........
SipP_p1015_ LEWGKAIVIAVILALLIRNFLFEPYVVEGKSMDPTLVDSERLFVNKTVK...........
SipP_p1040_ IDWIKAILIALILVFLVRTFLFEPYIVQGESMKPTLFNSERLFVNKFVK...........
SipU       TIVVLSIIMIAALIFTIRLVFYKPFLIEGSSMAPTLKDSERILVDKAVK...........
SipV       .FWFLAGVVSVVLAIQVKNAVFIDYKVEGVSMNPTFQEGNELLVNKFSH...........
E.coli     LETGASVFPVLAIVLIVRSFIYEPFQIPSGSMMPTLLIGDFILVEKFAYGIKDPIYQKTL
SipW       IVLTLVVISTRSSGGEPAVFGYTLKSVLSGSMEPEFNTGSLILVKEITD...........
                                    Box B
```

```
                70        80        90
SipS       .YIGEFDRGDIVVLNGDD...VHYVKRIIGLPGDTVEMK...................
SipT       .YIGELKRGDIVIINGETS.KIHYVKRLIGKGETVQMK...................
SipP_p1015_ .YTGNFKRGDIIILNGKEK.STHYVKRLIGLPGDTVEMK...................
SipP_p1040_ .YTGDFKRGDIVVLNGEEK.KTHYVKRLIGLPGDTIEMK...................
SipU       .WTGGFHRGDIIVIHDKKS.GRSFVKRLIGLPGDSIKMK...................
SipV       .RFKTIHRFDIVLFKGPD..HKVLIKRVIGLPGETIKYK...................
E.coli     IETGHPKRGDIVVFKYPEDPKLDYIKRAVGLPGDKVTYDPVSKELTIQPGCSSGQACENA
SipW       ..VKELQKGDVITFMQDANTAVTHRIVDITKQGDHLLFK...................
            Box C                   Box D
```

```
                100       110       120
SipS       ...............NDQLYINGKKVDEPYLAANKKRAKQDGFDH................
SipT       ...............DDTLYINGKKVAEPYLSKNKKEAEKLGVS................
SipP_p1015_ ...............NDHLFINGNEVKEPYLSYNKENAKKVGIN................
SipP_p1040_ ...............NDNLFVNGKRFNEEYLKENKKDAHDSDLN................
SipU       ...............NDQLYINDKKVEEPYLKEYKQEVKESGVT................
SipV       ...............DDQLYVNGKQVAEPFLKHLKSVSAGSHVTGD...............
E.coli     LPVTYSNVEPSDFVQTFSRRNGGEATSGFFEVPKNETKENGIRLSERKETLGDVTHRILT
SipW       ...............TKGDNNAAADSAPVSDENVRAQYTGFQLP................
```

```
                130       140       150       160       170
SipS       ..............LTDDFGPVKVPDNKYFVMGDNRRNSMDSRNGLGLFTKKQIAGTSKF
SipT       ..............LTGDFGPVKVPKGKYFVMGDNRLNSMDSRNGLGLIAEDRIVGTSKF
SipP_p1015_ ..............LTGDFGPIKVPKDKYFVMGDNRQESMDSRNGLGLFTKDDIQGTEEF
SipP_p1040_ ..............LTGDFGPIKVPKDKYFVMGDNRQNSMDSRNGLGLFNKKDIVGVEEL
SipU       ..............LTGDF.EVEVPSGKYFVMGDNRLNSLDSRNGMGMPSEDDIIGTESL
SipV       ...........FSLKDVTGTSKVPKGKYFVVGDNRIYSFDSRH.FGPIREKNIVG....
E.coli     VPIAQDQVGMYYQQPGQQLATWIVPPGQYFMMGDNRDNSADSRY.WGFVPEANLVGRATA
SipW       ..............YAGYMLHFASQPIGTAVLLIVPGVMLLVYAFVTISSAIREIERKTKA
                                                Box E
```

Figure 1-4   Sequence alignments of SPases I in *B. subtilis* and in *E.coli*

The sequences aligned are *B. subtilis* SipS, SipT, SipU, SipV, SipW, two SipP and *E. coli* SPase I. The highlighted regions with green are the conserved domains between the SPases I in different species. The common residues are highlighted in red. The numbering system corresponds to the *B. subtilis* SipS sequence numbering. The sequence for *E. coli* SPase I in the figure shows residues from 61 to 324. Sequence 2 to 35 is a predicted transmembrane segment in *B. subtilis* SipS whereas residues 4-22 and 59-77 are the two predicted transmembrane segments in *E. coli* SPase I. For *E.coli* SPase I, Box B constitutes residues 89-96, Box C has residues 128-135, Box D has residues 143-154 and Box E is composed of residues from 273 to 283. Domain II in *E. coli* Spase I includes residues 155-263 which mainly covers the regions between Box D and Box E. The counterpart domain II is present in *B. subtilis* SipS, but at a half smaller size than *E. coli* SPase I; the domain is composed of 53 residues covering the region from residue 91 to143. The sequence alignment is analyzed and generated using ESPript.
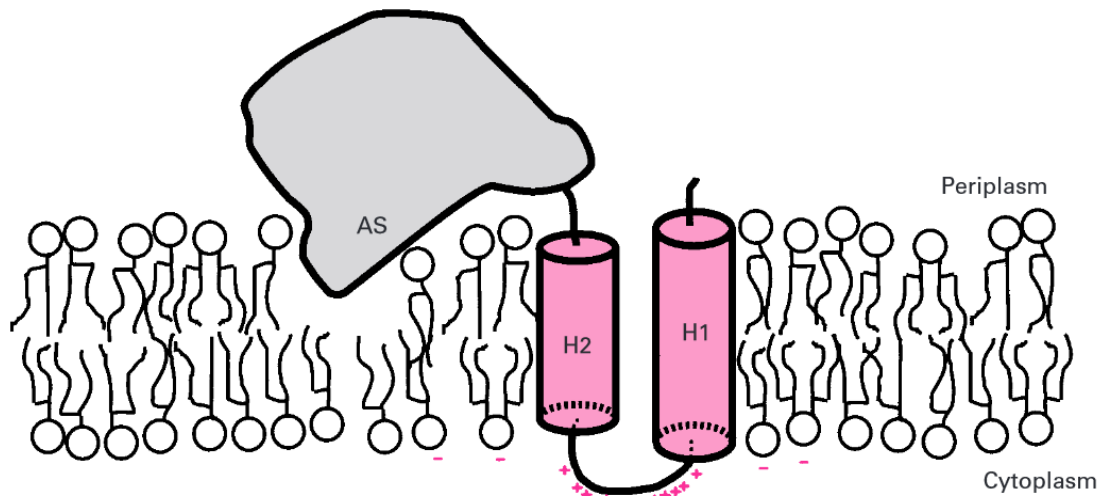
Figure 1-5   Membrane topology of *E.coli* SPase I

Topology of *E.coli* SPase I starts with a short periplasmic amino-terminal region (residues 1-3), followed by a transmembrane region (residue 4-28) (labeled H1), a cytosolic loop (residue 29-57), another transmembrane segment (residue 58-76) (labeled H2), and a large C-terminal periplasmic domain (residue 77-323) containing the active site (labeled AS). The figure is adapted from (van Voorst & de Kruijff, 2000).

1.4. Crystal Structure of *E.coli* SPase I

The first crystal structure of *E. coli* SPase I Δ2-75 with a β-lactam inhibitor was obtained in 1998, and four three-dimensional (3D) crystal structures of *E. coli* SPase I have been solved to date. Structures have been determined for *E. coli* SPase I Δ2-75 with a β-lactam inhibitor, *E. coli* SPase I Δ2-75 apo-enzyme, *E. coli* SPase I Δ2-75 with a lipopeptide inhibitor, and *E. coli* SPase I Δ2-75 with the antibiotic arylomycin lipoglycopeptide. Structural analysis of *E. coli* SPase I has led not only to the discovery of a unique Ser/Lys catalytic dyad mechanism in bacterial SPases I, but also identification of the key residues involved in substrate specificity, activity, and stability.

In resolving the *E. coli* SPase I Δ2-75 structure, it has been determined that the protein is composed of primarily of a β-sheet fold which has two large anti-parallel β-sheet domains: Domains I and II (Paetzel, 1998). Domain I is composed of conserved, essential residues and forms a large, exposed hydrophobic patch extending across the surface of the SPase I Δ2-75 enzyme (Figure 1-7). Domain I contains both

the substrate-binding site and the catalytic dyad (Ser91 and Lys146). It has been suggested that this hydrophobic patch of the enzyme is inserted into the membrane lipid bilayer so that it can optimize the binding interaction with the cleavage site on the signal peptide (Paetzel, 1998). In addition, Domain I contains a β-hairpin extension protruding outward which interacts with the N-terminal strand, giving a conical shape of the enzyme.

Domain II is present only in Gram-negative SPases and contains a disulfide bond which forms between Cys170 and Cys176 (Paetzel, 1998). The function of Domain II has not yet been determined. Site-directed mutagenesis studies have showed that Trp300 and Trp310 severely decrease the catalytic activity of the enzyme (Kim et al., 1995), however, the crystal structure indicates that these residues are far away from the active site (>20 Å). This suggests that Trp300 and Trp310 may play a role in stabilizing the structure of the enzyme and its placement at the membrane (Paetzel, 1998). Also, it was found that the soluble domain of *E. coli* SPase I, lacking the two transmembrane domains and the small cytoplasmic region, still requires a non-ionic detergent, Triton X-100, for optimal catalytic activity (Tschantz et al., 1995) (Paetzel et al., 1995). This is thought to be due to the presence of the large hydrophobic patch extending across the soluble domain of the enzyme, with which Triton X-100 can interact, stabilizing its native structure. Surface analysis has revealed that there are two hydrophobic, depressed binding pockets (S1 and S3) close to the catalytic Ser91 (Paetzel, 1998) (Figure 1-7). Both of the pockets are primarily made up of hydrophobic residues and prefer to interact with Ala residues, but S3 is more shallow and broader than S1 which indicates that it can accommodate a larger substrate residue. This result is consistent with its substrate specificity of Ala-X-Ala at the P3-P2-P1 positions of the substrate (Paetzel et al., 2002; Paetzel et al., 2004). In addition, the electron density map shows that the ε-amino group of Lys146 (Nζ) points toward the Ser91 2.9 Å away and is the only titratable group nearby. This strongly indicates that Lys146 acts as a general base, which abstracts a proton from Ser91, making it an active nucleophile able to attack the scissile bond of the substrate. It was also observed that Lys146 is completely buried inside, surrounded by numerous hydrophobic amino acid residues. This suggests that the high pK_a value of Lys146 (~10.8) is lowered in the micro hydrophobic environment, maintaining its deprotonated

state so that it can receive a proton from Ser91. In addition, determining the structure of *E. coli* SPase I in complex with a β-lactam inhibitor revealed that Ser91 attacks the *si*-face of the scissile peptide bond of the substrates whereas the classical serine protease attacks the *re*-face of the scissile bond (Paetzel et al., 2000). This result was the first evidence that bacterial SPase I uses a unique Ser/Lys catalytic dyad (Paetzel, 1998).
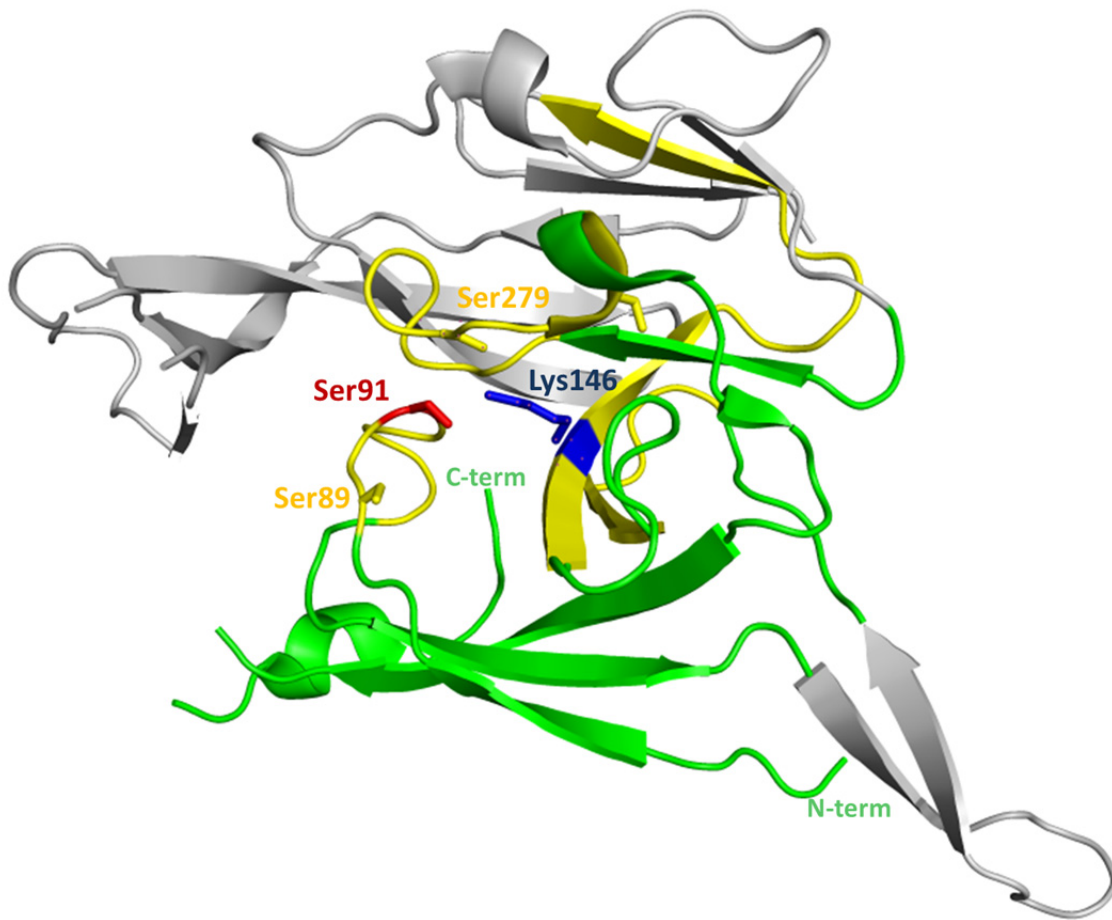
Figure 1-6   Cartoon representation of *E.coli* SPase I Δ2-75

The green and yellow colored portions of the protein represent the conserved domain I, with gray region being the non-conserved domain II. Box B-D, part of the conserved domain I, are show in yellow. In addition, the residues that are important for catalytic reactions, S91 and K146, are labeled and displayed in ball and stick representation. S91 and K146 are identified as S90 and K145 in the old numbering system (UniProt: P00803). This is due to the DNA sequencing error in the original sequence: Arg42 in the original sequence is actually Ala42-Gly43. The original sequence from 1-41 is correct both in register and numbering, whereas the original sequence 44-324 is one residue different in numbering from the new sequence. The counterpart of residues from 110 to 121 in *E. coli* SPase I, which makes up the β-hairpin (gray) in the enzyme, is not found in *B. subtilis* SPases I. In addition, a big part of domain II in *E. coli* SPase I, residues from 160-194 and 226-254, is not found in *B. subtilis* SipS. K146 is located only 2.9 Å away from S91 and is known to abstract a proton from the hydroxyl group of S91. (PDB: 1B12)
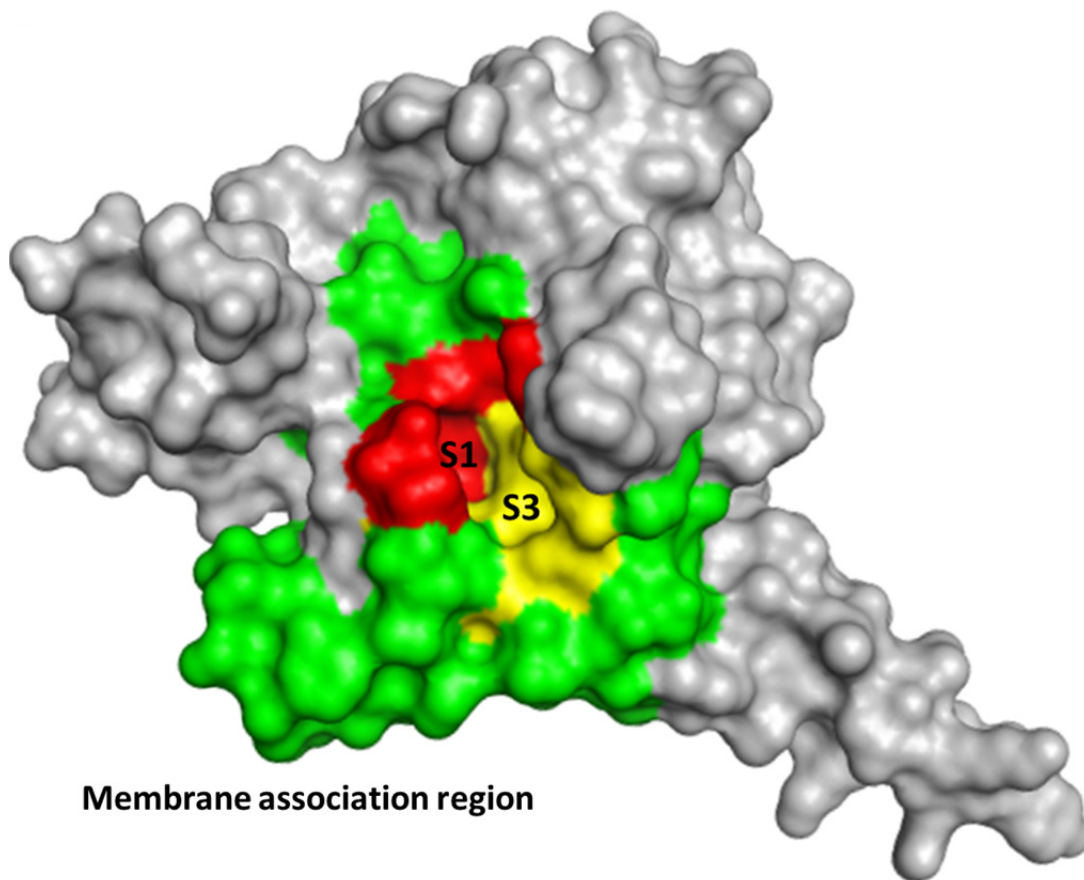
Figure 1-7   Surface representation of *E.coli* SPase I Δ2-75

The surface representation of *E. coli* SPase I Δ2-75 apoenzyme structure is shown with the modeled signal peptide binding sub-sites S1 and S3. The green colored surface represents the large hydrophobic region that lies across the enzyme. Each signal peptide binding sub-sites are labeled. (PDB: 1B12)

1.5. *B. subtilis* Type I Signal Peptidase

There are some Gram-positive bacteria with single SPase I enzymes; for example, *Lactococcus lactis*, *Streptococcus pneumoniae*, *Streptococcus mitis*, *Mycobacterium lep-rae*, *Mycobacterium tuberculosis* and *Aquifex aeolicus.* However, most Gram-positive bacteria have multiple paralogous SPases I. For example, *Bacillus cereus* and *B. subtilis* contain seven paralogous SPases I, *Bacillus anthracis* has six paralogous enzymes, while *Streptomyces lividans, Streptomyces coelicolor, Clostridium per- fringens* and *Bacillus amyloliquefaciens* have four paralogous

enzymes. In addition, *Lactobacillus plantarum, Listeria monocytogenes* and *Staphylococcus epidermidis* have three SPase-encoding genes, while *Bacillus halodurans, Staphylococcus aureus* (spsA and spsB) and *Staphylococcus carnosus* (sipA and sipB) have two SPase-encoding genes (van Roosmalen et al., 2004).

So far, there is no strong evidence supporting the reason why there are multiple SPases I in Gram-positive bacteria. Gram-positive bacteria, especially *B. subtilis*, are well-known for its capacity to produce a large amount of proteins in a variety of conditions. And therefore, there has been a speculation instead, that the multiple SPases I work as back-up SPases I to process the secretory enzymes under any harsh condition in order to be more efficient.

1.5.1.   Major and Minor SPases I

Among the seven paralogous enzymes found in *B. subtilis,* five are chromosomal SPases I; SipS, SipT, SipU, SipV, and SipW. The remaining two are plasmid-encoded SPases I, both named SipP, which are distinguished based on their plasmid types (pTA1015 & pTA1040) (Tjalsma et al., 1997, 1998; van Dijl et al., 1992). They were designated to be SPases I paralogues based on their similarities in sequences. As shown in the sequence alignments, not only they shared the nucleophilic serine and the general base lysine, they also maintained the conserved boxes A-E that are commonly found in SPases I. In addition, they were all able to process one substrate, β-lactamase, which is a common substrate for SPases I (Tjalsma et al., 1997).

SipS and SipT are classified as major chromosomal SPases in *B. subtilis* since they contribute the most to precursor processing and are essential for bacterial survival. Cell strains lacking both SipS and SipT are not viable possibly because their absence may cause an over-accumulation of secretory pre-proteins within the secretion machinery, leading to cell death (Tjalsma et al., 1998). Despite their importance, studies have found SipS and SipT can be functionally replaced with the plasmid-encoded major SPase I, SipP; cell strains lacking SipS and SipT but having SipP can survive (Tjalsma et al., 1999). The controlled gene transcription of the major SPases I

(SipS, SipT and SipP) in concert with the gene expression of most secretory proteins in *B. subtilis* is another indication of their important roles in the species (Bolhuis et al., 1996; Tjalsma et al., 1997).

B. subtilis SipU, SipV and SipW are classified as minor SPases I. Bacterial cells lacking these three SPases I are affected only mildly in terms of their pre-protein processing and secretion(Tjalsma et al., 1998, 1999). It is thought that the functional differences between the major and minor SPases I are due to the differences within their N-terminal domains rather than their sequence identities (van Roosmalen et al., 2001). The N-terminal domains of the major SPases I may position the active sites in the correct orientation so that they can interact with the cleavages sites of essential pre-proteins, the Sec translocase, and another pre-protein translocation. This is currently being investigated.

Other studies have found that the paralogous SPases I in *B. subtilis* have similar but not identical substrate specificity. For example, all of the major and minor SPases I except SipW are able to process a substrate β-lactamase precursor in *B. subtilis,* but the precursor of the α-amylase AmyQ is preferably cleaved by SipT (Meijer et al., 1995; Tjalsma et al., 1997; van Dijl et al., 1992)*. B. subtilis* SipW, on the other hand, is required for the processing of YqxM and TasA, the spore-associated protein (Stover & Driks, 1999; Tjalsma et al., 2000).
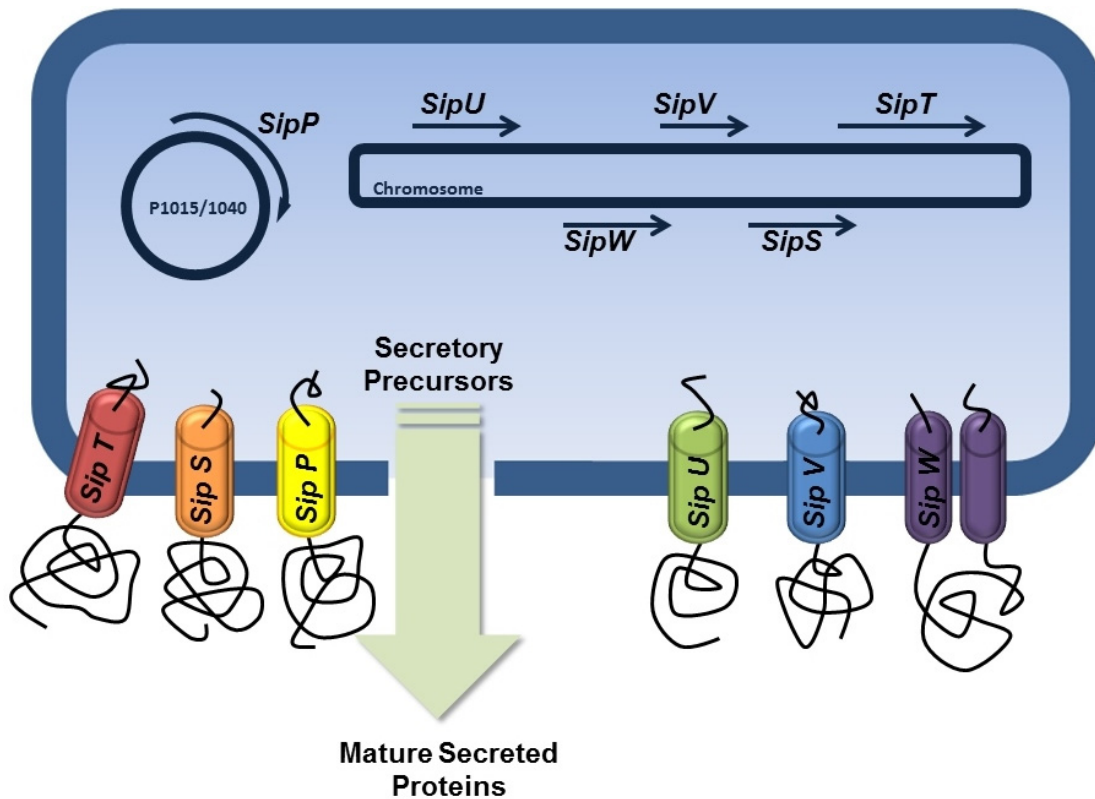
Figure 1-8   Synthesis of the paralogous SPases I in *B. subtilis*.

Type I SPases, responsible for the processing of 180 pre-proteins, are divided into two groups: the major SPases--SipS (S), SipT (T), and SipP (P), which are necessary for cell viability, and the minor SPases--SipU (U), SipV (V), and SipW (W), which are not necessary for cell viability *in vitro*. SipP is encoded by the plasmids pTA1015 and pTA1040, which are present in certain natto-producing *B. subtilis* strains. All other SPases are chromosomally encoded. In addition, SipW is the only (ER)-type SPase in *B. subtilis*, showing a high degree of similarity to eukaryotic and archaeal SPases. In contrast to the prokaryotic (P)-type SPases of *B. subtilis*, which have one amino-terminal membrane anchor, SipW appears to have an additional carboxyl-terminal membrane anchor. The figure is adapted from (Tjalsma et al., 2000).

1.5.2.   Membrane Topology

The *B. subtilis* paralogous enzymes can also be classified into the two subgroups: prokaryotic (P)-type and the endoplasmic reticulum (ER)-type. P-type SPases I are found in bacterial and eukaryotic organelles, whereas ER-type SPases are found in all three kingdoms of life. Other than their origins, the differences between the two types are reflected in their sequences and membrane topology. The P-type SPases I utilize a catalytic Ser/Lys dyad while the catalytic lysine is changed to a histidine in the ER-type SPase I. In addition, the ER-type SPases contain a second C

domain, labeled C', located between the C and D domains (Tjalsma, et al., 2000). Lastly, the P-type SPases are known to contain only one C-terminal anchor in a few cases but can have up to three N-terminal anchors, whereas ER-type SPases can have up to three C-terminal anchors but only one N-terminal anchor in rare cases (Tjalsma et al., 2000).

As of now, *B. subtilis* is the only strain that is known to contain both P-type and ER-type SPases I. The P-type SPases I in *B. subtilis* are the chromosomally encoded SipS, SipT, SipU, SipV, and the two plasmid-encoded SipP. SipW is the only ER-type SPase I in *B. subtilis*, showing a high degree of similarity to eukaryotic and archaeal SPases.

1.5.3.   SipS

SipS, one of the major SPase I enzymes in *B. subtilis*, consists of an N-terminal transmembrane domain of about 37-39 residues, followed by a soluble catalytic domain. It has been shown that the transmembrane domain is not necessarily required for catalytic activity of the enzyme (Carlos et al., 2000). However, without the transmembrane domain, the soluble domains of the enzyme are thought to be degraded via self-cleavage at its active site (Tschantz et al., 1995).
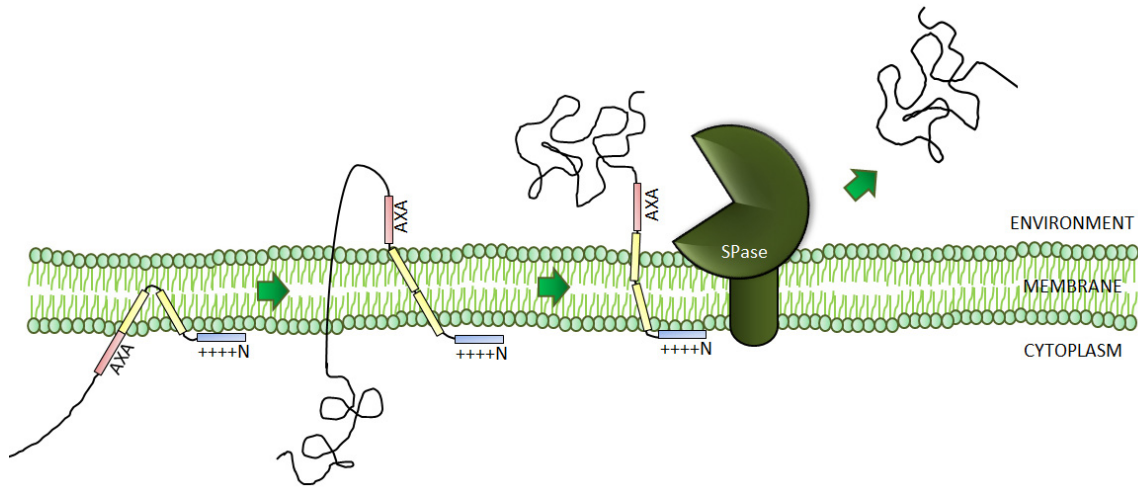
Figure 1-9   Signal peptide insertion and translocation of the pre-protein across the cytoplasmic membrane.

The positively charged N-region of the signal peptide binds to the negative charged phospholipids in the membrane. The H-region then inserts loop wise into the membrane whereupon the mature region of the pre-protein is pulled through. During or shortly after translocation, the signal recognition sequence in the C-region is recognized by membrane bound SPase I and the signal peptide is cleaved. The mature protein is then released into the extracellular environment as a folded, active form.

1.6. A Proposed Catalytic Mechanism of Serine Proteases

Serine proteases are characterized by the presence of serine nucleophiles. Serine proteases such as chymotrypsin, trypsin, and subtilisin utilize the catalytic triad, having three main residues which are crucial for catalytic activity: the serine (Ser) nucleophile, the general base histidine (His), and the coordinating residue aspartic acid (Asp). When the general base His abstracts a proton from the hydroxal group of the Ser nucleophile, the Ser forms an alkoxide which can attack a peptide scissile bond of the substrates. As the His gains an extra proton and becomes positively charged, the coordinating residue Asp nearby stabilizes the positive charge on the His. Serine proteases utilizing the Ser/His/Asp catalytic triad can be inhibited by phenylmethane-sulfonlyfluoride (PMSF), 4-[amidinophenyl]-methane-sulfonylfluoride (APMSF), and 3,4-dichloroisocoumarin (DCI).
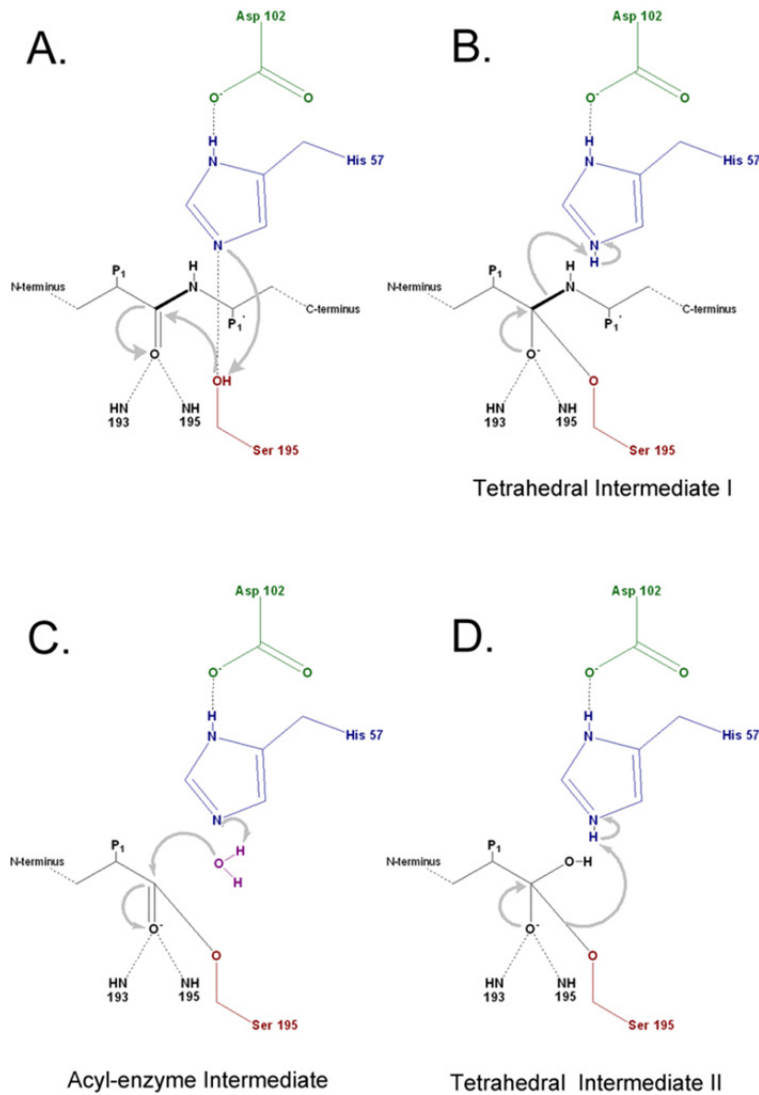
Figure 1-10 **The proposed catalytic mechanism of Ser/His/Asp triad of serine protease.**

A proton on the hydroxyl group (OγH) of Ser 195 is abstracted by the nitrogen group (Nδ1) of His 57 which is in close proximity. The Ser 195 hence becomes a nucleophile which then attacks the carbonyl carbon of the scissile bond of the substrate (A). This leads to a formation of tetrahedral intermediate I. The oxyanion hole is stabilized by the main chain amino groups (NαH). The protonated nitrogen group (Nδ1H) of His 57 donates a proton to the main chain amide group of the scissile bond, detaching product 1 from the substrate (B). It now forms an acyl-enzyme intermediate. The carbonyl carbon of the product 2 is bonded to the hydroxyl group (Oγ) of Ser 195 by an ester bond. The deprotonated nitrogen group (Nδ1) of His 57 activates the deacylating water molecule by abstracting a proton (C). The water becomes a nucleophile and then attacks a carbonyl carbon of product 2, forming tetrahedral intermediate II. This intermediate goes through a rearrangement of the bonds which releases the remaining product. Upon release, the free catalytic enzyme is restored (D).

1.6.1.   Serine/Lysine Catalytic Dyad

Serine proteases such as SPase I, SppA, UmuD and VP4 use a different mechanism from the classical catalytic triad for catalytic activity. They utilize a Ser/Lys dyad. *E. coli* SPase I contains a serine residue (S91) that acts as a nucleophile and has a lysine (Lys146) acting as a general base instead of histidine (Blackt, 1993; Sung & Dalbeys, 1992). The catalytic residues Ser91 and Lys146 in *E. coli,* are highly conserved across all known bacterial SPases I. Site-directed mutagenesis studies have shown that, in *B. subtilis*, Ser43 and Lys83 are the corresponding catalytic dyad residues (van Dijl, 1995). The crystal structure of *E. coli* SPase I in complex with a β-lactam inhibitor has been resolved and the following mechanism of the catalytic dyad was elucidated based on the results. During the substrate recognition and binding process, the side chains of the substrate at P1 and P3 are bound to the enzyme's active site hydrophobic binding pockets S1 and S3 (Figure 1-7). Positions P2, P4 and P5 of the substrate, on the other hand, face away from the enzyme (Paetzel, 1998). The binding interaction between the substrate and the enzyme is likely caused by parallel β-sheet type hydrogen bonding between the C-region of the substrate and the SPase I β-strand 1 with its attached loop leading to the Ser91 (Paetzel, 2014). During cleavage, Lys146 in the active site acts as a general base in which its side chain amino group (N_ζ) abstracts a proton from the hydroxyl group (O_γ) of Ser91. Ser91 then forms an alkoxide that attacks the peptide bond between P1 and P1' of the substrate (Figure 1-7). Normally, Lys146 has a pK_a value of 10.8 which is relatively high. In order to be able to receive a proton from Ser91, the pK_a needs to be lower and the microenvironment surrounding Lys146 achieves this by being rich in hydrophobic amino acid residues (Dao-Pin et al., 1991)**.** After Ser91 forms a covalent bond with the carboxyl group of P1, the complex forms tetrahedral intermediate I. The hydroxyl group of Ser89 and the main chain amide group (N_α) of Ser91 together form an oxyanion hole and stabilizes the tetrahedral intermediate I. Next, the amide group of P1' receives a proton from the protonated amino group (N_ζ) of Lys146. This releases the mature protein from the enzyme leading to formation of an acyl-enzyme intermediate. Lys146 activates the deacylating water that facilitates formation of the tetrahedral intermediate II. Finally, the protonated amino group of Lys146 donates a proton to the Ser91

23

hydroxyl group (O$_\gamma$) and the cleaved signal peptide dissociates, regenerating the SPase I catalytic dyad active site.
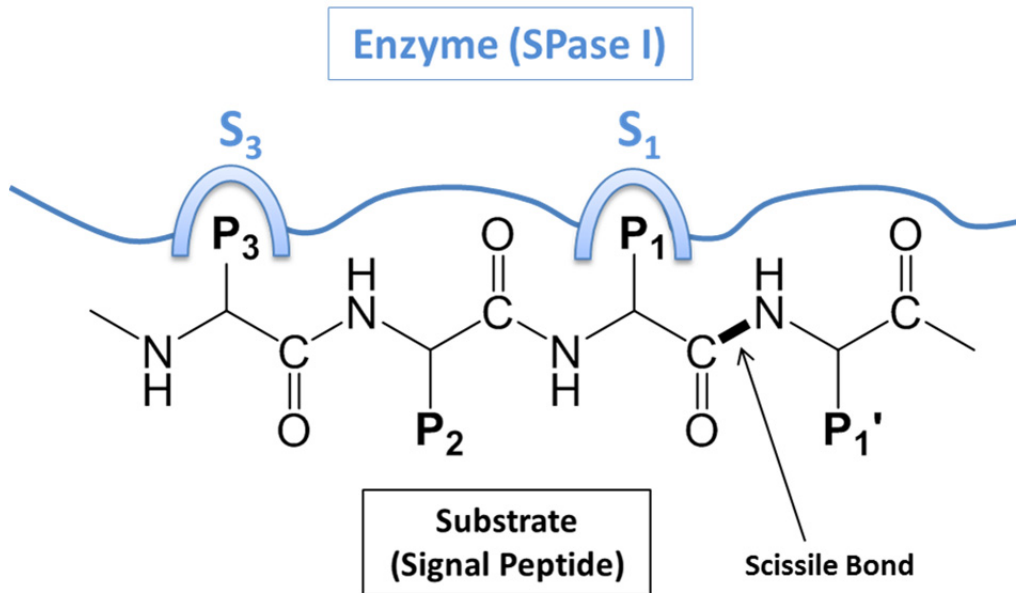
Figure 1-11 **The substrate binding sites on the peptidase using Schechter & Berger nomenclature.**

The residue located on the N-terminal side of the substrate's scissile bond is called P1 or -1 whereas the residue on the C-terminal side is called P1' or +1. The subsequent residues on the N-terminal side of P1 or +1 are labeled with increasing P numbers (e.g. P2, P3, P4) or increasing integers with a minus sign (e.g. -2, -3, -4). The residues on the C-terminal side of P1' or +1 are named with increasing P' numbers or increasing integers with a plus sign. The enzyme's counterpart residues are identified in a similar manner with P being replaced with S. The figure is adapted from (Kim, 2013).
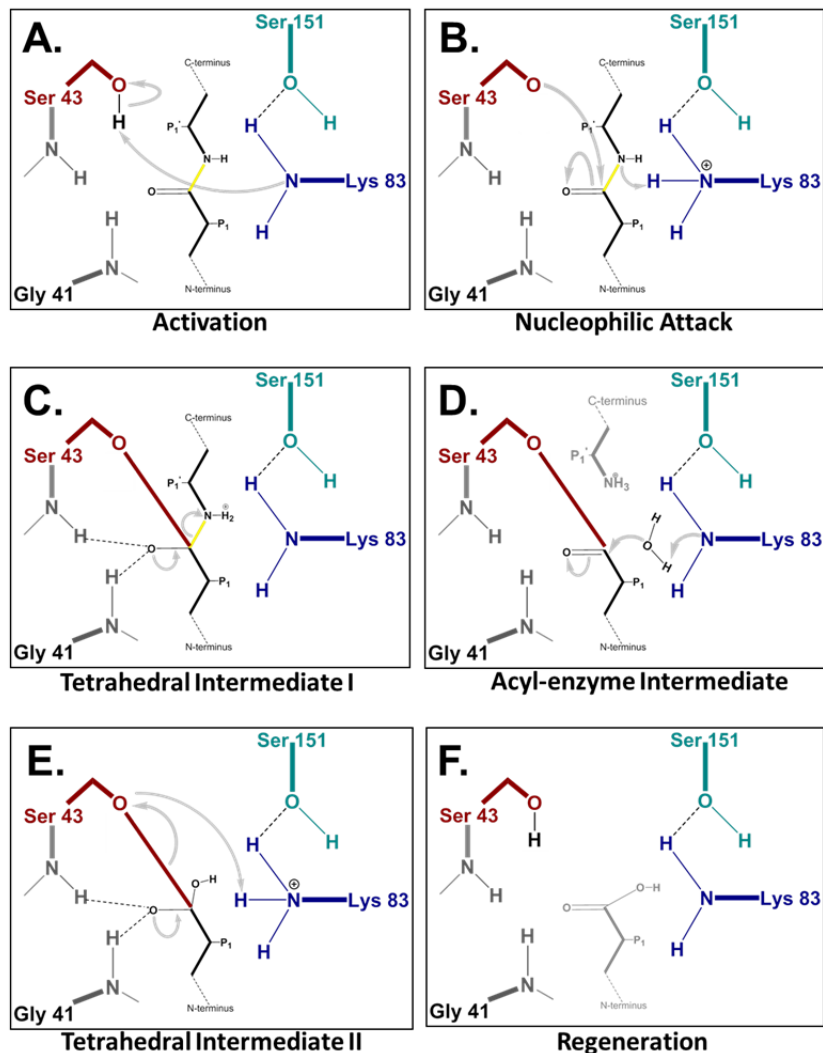
Figure 1-12 **The estimated catalytic mechanism of the Ser/Lys dyad in *B. subtilis* SipS based on the crystal structure of *E. coli* SPase I.**

A proton on the hydroxyl group (Oγ) of Ser43 (labeled in red) is abstracted by the general base Lys83 (labeled in blue) that is located close to the Ser43. The oxyanion hole is created by the amino group (NαH) group of Ser43 and the hydroxyl group (OγH) of Ser43 (A). The hydroxyl group (Oγ) of Ser43 applies a nucleophilic attack on the carbonyl carbon of the scissile bond (shown in yellow) of the substrate (shown in black). This causes the protonated Lys83 amino group (NζH) to donate a proton to the main chain amide nitrogen at the scissile bond (B). A tetrahedral oxyanion transition state I forms, in which the oxyanion is stabilized by hydrogen bonding interactions with the amino group (NαH) group of Ser43 and the backbone amino group (NαH) of Gly41 (C). An acyl-enzyme complex then forms as product 1 (shown in light grey) is released from the substrate. The main carbonyl carbon of the remaining product (product 2) is covalently attached to the hydroxyl group (Oγ) of Ser43 via an ester bond. The deprotonated amino group (Nζ) of Lys83 activates the deacylating water, turning it into a nucleophile, which then attacks the carbonyl carbon of product 2 (D). Tetrahedral oxyanion transition state II is created (E). Product 2 (shown in light gray) is now released (F). This figure is adapted from (Paetzel 2014).

25

1.7. Serine Protease Inhibitors

1.7.1. Classification of Enzyme Inhibitors

Serine protease inhibitors are molecules that bind to the enzyme, inhibiting its catalytic activity. There are two main types of inhibitors; reversible and irreversible inhibitors (Conn et al., 1987). An irreversible inhibitor docks onto the enzyme permanently by forming covalent interactions and irreversibly inhibits the enzyme's catalytic function. A reversible inhibitor, on the other hand, inhibits the enzyme's activity in a reversible manner by forming weak non-covalent interactions. It can freely detach from the enzyme, thus restoring the enzyme's activity.

Reversible inhibitors are divided into three subclasses; competitive, non-competitive and un-competitive inhibitors. A competitive inhibitor has a similar structural shape to the substrate and as such, competes with the substrate for the active site on the enzyme. Competitive inhibition, therefore, can be overcome if the substrate concentration is increased to the point where the competitive inhibitor is overwhelmed. In some cases, the competitive inhibitor can also bind to a site called the inhibitor-binding site, which is usually far removed from the active site on the enzyme; once the competitive inhibitor binds to the inhibitor-binding site, the binding interaction changes the structural conformation of the active site so that the substrate can no longer bind. At this point, the competitive inhibitor can only bind to the enzyme, not the enzyme-substrate complex.

A non-competitive inhibitor does not share structural or chemical similarities to the substrate and hence never competes with the substrate for the enzyme's active site. Instead, it binds to a separate inhibitor-binding site on the free enzyme, as well as to the ES complex. As a consequence, non-competitive inhibition cannot be relieved regardless of the substrate concentration.

An uncompetitive inhibitor can only bind to the ES complex, leading to the formation of an inactive EIS complex. This complex cannot bind to the free enzyme, nor can it form an EI complex.

1.7.2. Signal Peptidase I Inhibitors

Many serine protease inhibitors are not effective against SPase I due to its Ser/Lys catalytic dyad as opposed to the more common Ser/His/Asp catalytic triad that is found in most serine proteases (Paetzel et al., 2000). However, four *E. coli* SPases I inhibitors have been identified.

The first inhibitor is a signal peptide that is composed of 23 amino acid residues. It inhibits the enzyme's activity in a competitive manner, inhibiting the processing of pro-coat and pre-MBP *in vitro* (Wickner et al., 1987). The second inhibitor is a pre-protein with a mutated signal peptide at the +1 position of the cleavage site. When the amino acid residue at +1 is mutated to a proline, it twists the C-region which is critical for fitting into the binding sites of the enzyme (Barkocy-Gallagher & Bassford, 1992). The proline residue disrupts the position of the pre-protein and the orientation of the peptide bond in the binding sites, resulting in catalytic inhibition. The third inhibitor is a Penem-type inhibitor. Penem molecules are a type of β-lactam compounds and are a very effective inhibitors, with an inhibitory concentration of less than 1 μM against *E. coli* SPase I (Allsop et al., 1996; Black & Bruton, 1998; Kuo et al., 1994). This inhibitor has a bond that mimics the scissile bond of the substrate; it forms a covalent bond with the catalytic Ser91 Oγ in and permanently disables the catalytic reaction. The fourth inhibitor is the most recently developed, a novel lipopeptide-based inhibitor named Arylomycin (Kulanthaivel et al., 2004; Schimana et al., 2002). Arylomycin A2, one of Arylomycins, was co-crystallized with *E. coli* SPase I Δ2-75 and structural analysis showed that the C-terminal O45 of the inhibitor forms a hydrogen bond with the catalytic Ser91 Oγ, Lys146 Nζ, and the oxyanion hole Ser89 Oγ (Paetzel et al., 2004). The C30 methyl group of Arylomycin A2 points into the substrate-binding pocket S3 while the C9 methyl group points toward a shallow pocket, which is a possible S5 binding pocket.
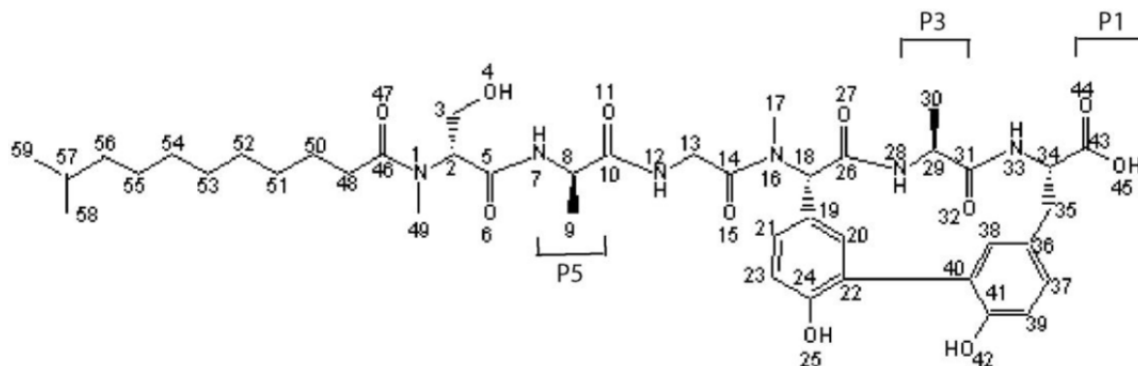
Figure 1-13   Structure of Arylomycin A2

P1, P3, and P5 correspond to the enzyme binding sub-sites S1, S3, and S5, according to Schechter and Burger nomenclature. This figure is adapted from (Luo, 2007)

1.8. Bacterial SPase I: A Novel Antibacterial Drug Target

With clinical pathogens, such as *B. subtilis*, becoming increasingly resistant to traditional antibiotic treatments, developing new strategies to combat these bacteria has a new urgency. SPase I is a promising target for antibacterial drugs that can act against *B. subtilis*. First, it is an enzyme essential for the survival of the *B. subtilis* organism. This characteristic ensures that the antibacterial drug targeted to *B. subtilis* SPase I will kill the organism. Second, the active site of SPase I is located in the periplasmic space of the cytoplasmic membrane, which suggests the drug would have easier access to its target than cytosolic target proteins. This is especially true in Gram-positive bacteria since they do not possess an outer membrane. Third, SPase I, because it utilizes a unique Ser/Lys catalytic dyad rather than the commonly used Ser/His/Asp catalytic triad, is an ideal candidate for the development of a highly specific inhibitor. And fourth, bacterial SPases I are distinctly different from eukaryotic SPases, suggesting there is a good chance for the novel drug to impede bacterial viability without affecting the host. Eukaryotic SPase I acts as a multimeric enzyme using the Ser-His-Asp catalytic triad mechanism whereas bacterial SPase I is a monomer with the Ser/Lys catalytic dyad (Auclair et al., 2012).

If the structure of a protein– its active site, for instance- is known at the atomic level, small molecules can be synthesized to bind to that site and interfere with the

28

enzyme's activity. These small molecules are called inhibitors, or if used therapeutically in humans, drugs. In order to synthesize bioactive molecules in the laboratory, it is necessary to know the geometry and chemical characteristics of the enzyme's active site, its hydrophobic valleys and electrostatic landmarks, and the chemical features and mechanism that endow it with its unique biochemical properties.

1.9. Challenges in Crystallizing Proteins

Growing a protein crystal is a very challenging and complex process since every target protein requires a unique set of chemical, temperature, concentration, ionic strength and pH conditions in order to crystallize. Time is another factor. The complexity, chemical and physical instability and dynamic properties of each protein dictate what is needed for crystallization to occur. This chapter discusses the steps needed for protein crystallization, current challenges and the possible remedies that can be applied.

1.10. X-ray Crystallography: General Steps

In order to obtain a three-dimensional (3D) protein structure, the first step is to generate an ordered protein crystal that is larger than 0.1mm in all dimensions. Once a crystal is obtained, a control needs to be carried out, to determine if the crystal is likely to be protein or inorganic (salt). The control consists of the same conditions under which the original crystal was generated but without the addition of the protein. If there is crystal formation without protein, the original crystal is probably a salt crystal. Another test to differentiate between salt and protein crystals is to view the crystals through a microscope with a polarizing filter. While birefringence is a characteristic shared by all crystals, salt crystals tend to be more highly birefringent than protein crystals. There are other tests that can be carried out to determine if the crystal is likely to be protein but they result in the destruction of the crystal and as such are only carried out if there are multiple crystals. The first test is the crush test. If the crystal is soft and easily crushed, it is probably protein while if the crystal resists crushing, then it is likely salt. There is also a commercial product (Crystal Dye from Hampton Research)

which can be added to a drop containing crystals. If the crystals are protein, the small molecules of the dye can penetrate the crystal's solvent channels and turn the crystal blue. Salt crystals are impenetrable by the dye and remain colourless. Once identified as likely to be protein, the crystal is then looped, mounted and subjected to a focused X-ray beam. The proteins within the crystal diffract the X-ray beam into a characteristic pattern of spots, with unique positions and intensities, called a diffraction pattern. After analysis of the diffraction pattern data, an electron density map of the crystal is built. Using multiple solving strategies, the final 3D structure can then be determined, and deposited in the Protein Data Bank (PDB), an open access repository of 3D structural data of proteins, nucleic acids and other large biological molecules.
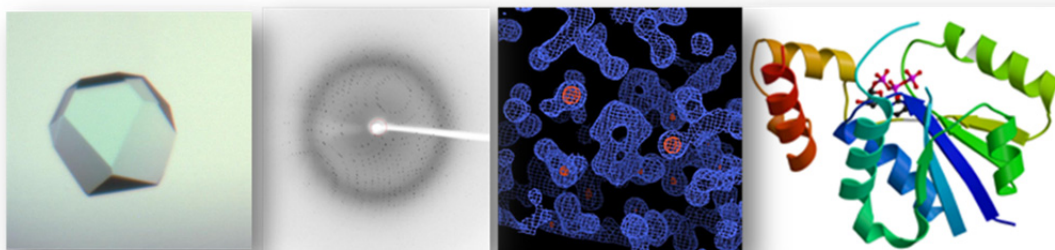
Figure 1-14   General steps in X-ray Crystallography

This diagram shows the workflow for solving the structure of a molecule by X-ray crystallography. First, a protein crystal is obtained. The crystal is then subjected to a focussed, monochrome X-ray beam to create a diffraction pattern. Next, an electron density map is constructed by using the data collected from the diffraction pattern. The final step is determining the three dimensional structure of the molecule. The figure is adapted from (http://crystal.csiro.au/About/Crystallisation.aspx#Formation).

1.11. Asymmetric Unit, Space Group and Unit Cell

A protein crystal can grow as a precisely ordered, three-dimensional array of molecules, maintained by interactions between the individual molecules. The remaining space within the crystal consists of large cavities containing water and/or buffer molecules. Each protein crystal can be characterized by a set of parameters that define the arrangement of the fundamental units. This set of parameters includes the asymmetric unit, crystal symmetry and space group, and the unit cell.

The asymmetric unit (AU) represents a basic building block of the crystal structure and is a fundamental unit in protein crystallography. It is called the "asymmetric" unit because there is no symmetry element in the unit. The AU is usually a small integer number, but also it can be a fraction such as ½ if one protein molecule has self-symmetry. The AU is characterized by a set of x, y, z coordinates of the atom.

Once the AU is determined, a unit cell (UC) can be constructed using symmetry operations and its space group. The UC is the smallest repeating translationally repeating unit in a protein crystal lattice and is characterized by three side lengths a, b and c and three angles α, β, and γ (Figure 1-16). These parameters determine which lattice classification the unit cell has, among the seven lattice types (Figure 1-16). The construction of the UC is completed when symmetry operations and the space group of the AU are applied within the assigned lattice types. The symmetry relations between the asymmetric units of chiral molecules within a crystal motive involve only simple rotations and rotations around a screw axis (rotation combined with translation). Space group refers to the combinations of one or more symmetry elements with a specific lattice type and is used to characterize a crystal. There are 65 space groups possible for protein crystals.
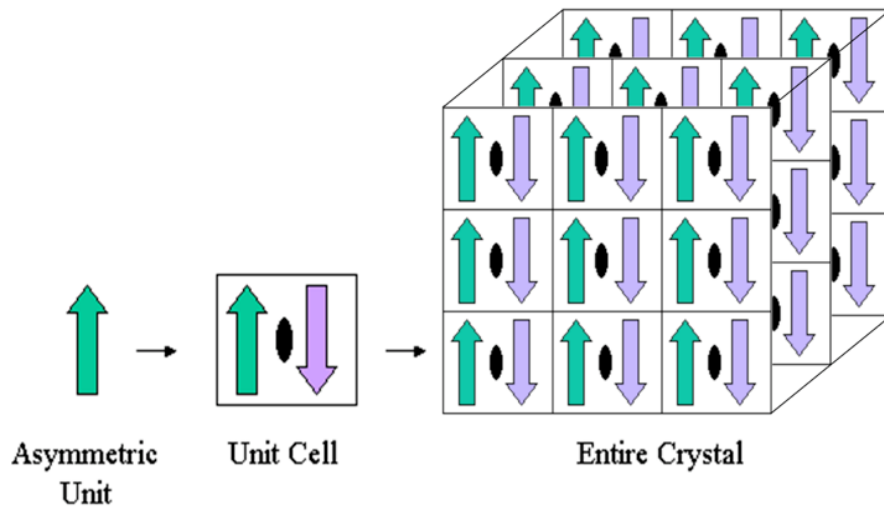
Figure 1-15   The building blocks of a crystal.

In this diagram, the unit cell consists of two asymmetric units (green arrow) that are rotated 180 degrees to each other about a two-fold crystallographic symmetry axis (black oval). This unit cell is repeated and translated, forming a crystal lattice that makes up a crystal. The figure is adapted from (http://www.rcsb.org/pdb/101/static101.do?p=education_discussion/Looking-at-Structures/bioassembly_tutorial.html)
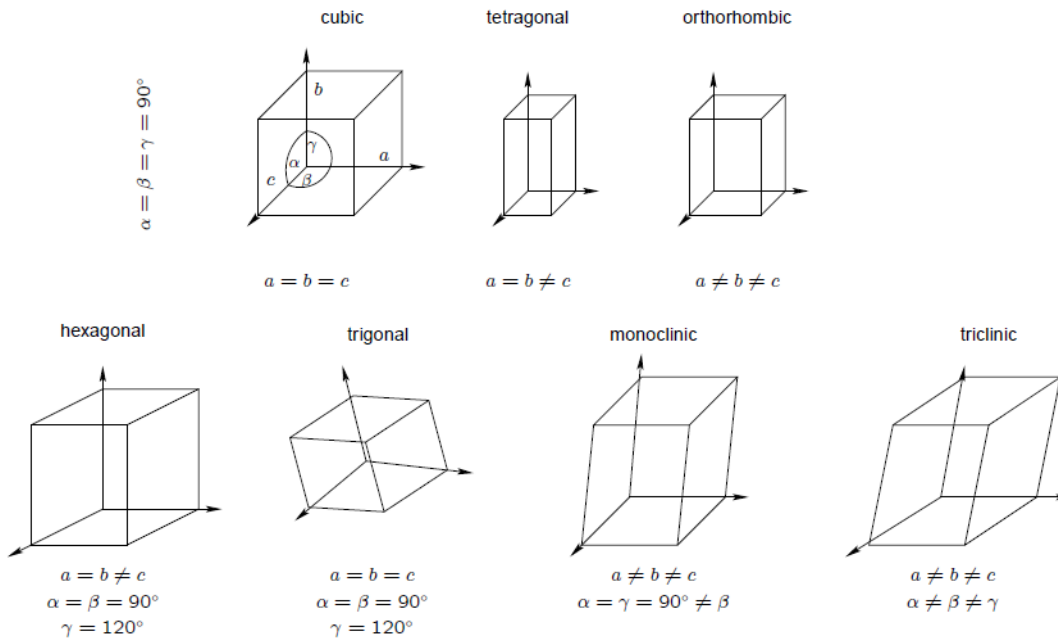
Figure 1-16   Seven basic shapes of unit cell that can be defined with 14 Bravais lattices and 65 space groups.

Extra lattice points can be found in each unit cell, for instance, there can be an atom in the centre of a face or in the centre of the unit cell. If these lattices points are combined with seven basic shapes of unit cell, there will be 14 Bravais lattices and each bravais lattice will fill space without gaps. This figure is adapted from (http://lamp.tu-graz.ac.at/~hadley/psd/L2/L2.php).

1.12. Protein Crystallization

1.12.1.   Vapour Diffusion Method

The first step towards obtaining a 3D protein structure is growing a protein crystal. In order to obtain a protein crystal, there are several physical methods that can be used, including vapour diffusion, free interface diffusion, batch and dialysis. While all these methods work by affecting the super-saturation state of the protein, the vapour diffusion method is the most commonly used technique.

In the vapour diffusion method, the target protein (2-15mg/ml concentration range) is added to a solution containing buffer and precipitants in a 1:1 volume ratio. The solution usually contains the chemicals that are known to promote protein

precipitation without denaturing the protein; these include ionic compounds, organic solvents and the water-soluble polymer polyethylene glycol (PEG). The crystallization droplet containing the purified protein, buffer, and precipitants is sealed within a well, surrounded by but not touching a larger volume of buffer and precipitants, called the reservoir solution. Within this closed environment, the crystallization droplet and reservoir solution equilibrate through the process of evaporation and diffusion (Figure 1-17). As equilibration proceeds, water evaporates from the crystallization droplet and the precipitant and protein concentrations increase, leading to super-saturation of the protein. The equilibrium of the crystallization droplet with the precipitants is restored by the formation of protein crystals (Nemčovičová & Kutá Smatanová, 2012). The vapour diffusion method can be carried out using either of two formats: hanging drop or sitting drop (Hampel et al., 1968). In the sitting drop format, the crystallization droplet is placed upon a pedestal raised above the reservoir solution whereas the hanging drop format suspends the crystallization droplet on a glass slide above the reservoir solution.
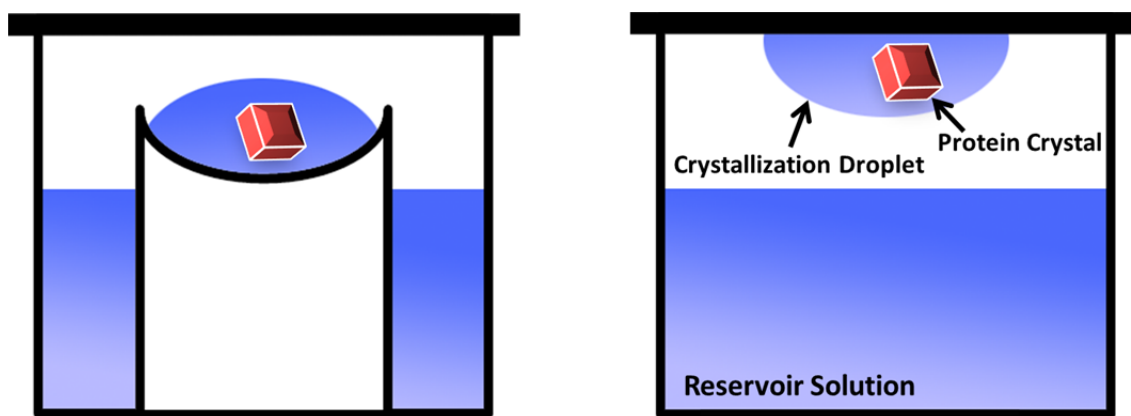
Figure 1-17   Vapour diffusion method: sitting drop and hanging drop.

The image on the left represents the sitting drop and the image on the right is the hanging drop method. The large amount of blue solution in the well represents the reservoir solution. The blue colored solution on the pedestal (sitting drop) or on the bottom of the top glass (hanging drop) represents the protein solution mixed with an aliquot of the reservoir (crystallization droplet).

1.12.2. The Growth of Protein Crystals

Once the target protein forms a crystal in the crystallization droplet, it has successfully passed through the three main steps of protein crystallization: nucleation, growth, and cessation of growth.

Nucleation is the first phase of the crystallization process and is marked by the appearance of very small crystals. This occurs when the now concentrated protein in the crystallization droplet separates from the solution to form crystalline aggregates. Nucleation is known to be the rate-limiting step in the crystallization steps, requiring more energy than the other steps (Nemčovičová & Kutá Smatanová, 2012).

The nucleation step is followed by the growth phase, where the nuclei of the tiny crystals begin to grow in size. The growth of an ordered protein crystal appropriate for use in X-ray crystallography needs to be a slow, controlled precipitation process in an aqueous solution and in conditions where protein denaturation will not occur.

The last stage is the cessation of growth. This happens when the solution reaches equilibrium with the precipitating agent. For X-ray diffraction purposes, the crystals need to grow to greater than 0.1mm in size.

1.13. Protein Crystallization Strategies

1.13.1. Influence of Additives in Protein Crystallization

Once the homogenous, purified and highly concentrated protein of interest is obtained, the vapour diffusion method of crystallization, as introduced in 1.12.1., can be carried out. To increase the chances of crystal growth, additives can be introduced into the crystallization droplet. Additives can be simple ions of common bivalent metal salts, various proteinaceous compounds or natural compounds such as ligands. It has been found that additives can affect every step of the crystallization process from saturation to precipitation.

1.13.2.   Ligands & Inhibitors

Studies have determined that co-crystallization of a protein with its natural ligand or an inhibitor can facilitate protein crystal growth since forming a complex can provide a more stable structure. Having a stable protein structure is beneficial as it prevents disordered protein aggregations. These ligands can be added directly to the crystallization droplet or incubated with the protein for a certain amount of time before being put into the crystallization droplet (Nemčovičová & Kutá Smatanová, 2012).

1.13.3.   Preparation of Proteins to be crystallized

There are many strategies that can be used to obtain crystals suitable for X-ray crystallography studies. The most important requirement for the protein to be used in crystallization trials is homogeneity; the protein needs to be as pure and uniform as possible. Multiple purification methods can be utilized to achieve this (for example, Ni^{2+}-NTA chromatography and size-exclusion chromatography). The solubility and stability of the purified protein is also important. This can be enhanced by finding an appropriate buffer for the protein which can then be adjusted in pH, temperature and by the addition of detergents (Privé, 2007). For crystallization trials, the protein needs to be concentrated as much as possible but not to the point of precipitation. The higher the concentration of the protein, the higher the chance of nucleation and crystal growth.

1.13.4.   Metal Ions

Metal ions are known to promote or contribute to the crystallization of different proteins. Since many proteins need ions for activity, it is thought that the ions stabilize the proteins by helping maintain structures. It has been shown that metal ions stimulate crystal growth; ferritin, α-lactalbumin, and ricin are examples of proteins that could only be crystallized in the presence of divalent metal ions (Nemčovičová & Kutá Smatanová, 2012).

1.14. Data collection

Once the protein of interest forms a crystal in the droplet, the crystal is then subjected to a focussed X-ray beam to test its diffraction. First, the crystal is mounted in a loop and immediately flash-cooled to 100°K. This loop is then attached to a device called a goniometer head, and oriented in the path of the X-ray beam. For cryogenic data collection, a nitrogen gas stream directed on the crystal keeps temperatures at 1000K throughout the experiment. Exposing the crystal to the X-ray beam while maintaining the low temperature is favoured since it minimizes the chances of radiation damage.

Once the crystal is fixed in the X-ray beam line, it is exposed to the X-ray beam. The electrons within the protein crystal scatter the rays of the X-ray beam, which are then detected by an image plate detector. The scattered X-rays are recorded as dark spots called reflections. In order to obtain a complete data set, images are recorded at every rotation of the crystal through a small angle (~1°) at a wavelength of ~1-1.5 Å.

1.15. Structure Solution

1.15.1. Primary Data Analysis

Primary data analysis is composed of indexing, integration, and scaling. In the indexing process, each reflection of the diffraction pattern is analyzed with an optical scanner which extracts two pieces of information: the geometrical arrangement and the intensity of each reflection. The geometrical arrangement of the reflection provides information about the crystal lattice and the symmetry of the crystal, whereas the intensity provides information about the content of the lattice. In this step, each reflection is characterized with integer numbers (h, k, l = Miller Indices) in order to determine the crystal geometry. Obtaining the correct geometry of the crystal is crucial because it determines if the intensities will be integrated and scaled accurately in later steps. In the integration step, the intensity values of these indexed reflections are measured using software programmes. Each reflection is caused by photons that are reflected within the crystal and can be simply calculated by a square of net amplitudes.

Both the intensity value (I) and a background value (σI) are determined as many reflections can be hidden in the background. Indexing and integration can be done by software programmes such as MOSFLM and HKL2000. In the next step, scaling using the software program SCALA, combines the integrated diffraction data of multiple images into one image of the unit cell and its contents. The intensity values of each diffraction image can be slightly different to each other due to errors, and this error can be 'merged' by replacing it with the weighted average intensity value. It has been shown that the closer the intensity values are between the equivalent reflections, the higher the data quality is produced (Chayen & Saridakis, 2008).
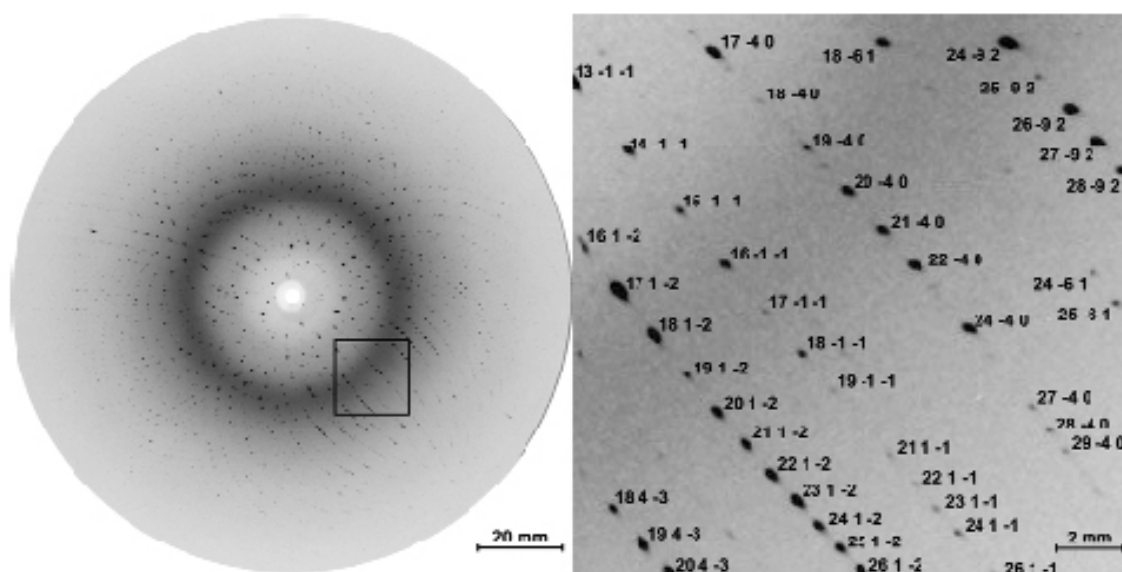
Figure 1-18   Labeled diffraction pattern with Miller indices (hkl)

The geometrical arrangement of the reflections provides the information about the crystal lattice and the symmetry of the crystal, whereas the intensity of the reflection that gives part of the information about the content of the lattice. During indexing, each reflection is labeled with integer numbers (h, k, l = Miller Indices) in order to find the lost phases. The figure is adapted from (http://my.yetnet.ch/dergutemensch/crystallography/indexing.htm).

1.15.2.   Phasing

Each reflection of the diffraction pattern contains an electromagnetic wave having both amplitude and phase. During data collection, amplitude is easily calculated using the integrated intensity value, however, the phase gets lost. Phase is a very important component in solving a 3D structure since it determines the electron density

distribution in the crystal. Phase can be restored through other methods. Current approaches to the phase problem include: direct method, molecular replacement (MR), isomorphous replacement (MIR), and anomalous dispersion (MAD). These methods are used to find initial phases that can subsequently be refined later. If a homologous structure of the target protein is known, MR can be used to estimate the initial phase. This method involves taking the previously solved structure, rotating and translating it into a new crystal until there is a good match to the experimental data. If there is no starting model, then the MIR method, in which one or more heavy atoms are introduced into specific sites within the unit cell without perturbing the crystal lattice, can be used. Heavy atoms are electron dense and give rise to measurable differences in the intensities of the spots in the diffraction pattern. By measuring these differences for each reflection, it is possible to derive an estimate of the phase angle using vector summation methods. A novel protein without any homologous structures can be solved with MAD. This method uses different wavelength of lights to diffract the crystal of protein bound to heavy-atom such as selenomethionine. Scattering factor of heavy-atom bound to protein at different wavelength is measured and then is compared to the anomalous scattering factors of the heavy atom. By collecting data at several wavelengths near the absorption edge of an element in the crystal, one can obtain phase information analogous to that obtained from MIR.

1.15.3. Structure Refinement

Once the initial phase is determined, combined with the calculated amplitudes, the location of each atom of the protein within the asymmetric unit can be estimated. Subsequently, an electron density map can be built with calculated structure factors by using a Fourier transform (Smyth & Martin, 2000). A preliminary model is generated based on the electron density map, with the protein primary sequence fitted into the map. This initial model is then refined against the original data. As the phase improves the resulting electron density maps become clearer and a better model is generated. When the R-factor (the measure of agreement between the model and the data) decreases to 25%, the refinement is usually considered complete.

1.16. Objectives

The main objective of this research project was to gain insight into *B. subtilis* SipS, one of the major SPases I, which is significant for cell's viability. By a solving a crystal structure of *B. subtilis* SipS, a three-dimensional representation of catalytically important regions, for instance, active site and substrate binding pockets, can be visualized. The architecture and the composition of these regions will provide insights into how the substrate binds onto the enzyme. By observing the orientation and the characteristic of the substrate while it is bound on the enzyme, one can design an inhibitor. The primary goal was to generate reproducible, diffraction quality crystals of *B. subtilis* SipS for X-ray crystallographic structure determination by optimizing the overexpression, purification, and crystallization conditions of the enzyme .The primary goal is to determine the 3D structure of the protein using X-ray crystallography techniques. A secondary goal of this project was to identify a potential SipS inhibitor and characterize its effect on the catalytic activity of the enzyme. The following questions were asked:

1. In what buffer condition is the enzyme most soluble while retaining its activity?

Different detergents were used and different purification techniques were carried out in order to determine the optimal environment for the active SipS protein.

2. In what crystallization condition does the enzyme form crystals?

Purified SipS protein samples were crystal plated using all the crystallization screens available in the Paetzel lab. These screens included Crystal Screens 1 and 2, PEG Ion Screen, JCSG+ Screen and PACT screen.

3. What chemical compound can inhibit SipS catalytic activity?

B. subtilis SipS Δ2-35 WT was incubated with different chemical compounds and catalytic activity was measured. To identify the mechanism of inhibition, size exclusion and cross linking experiments were performed.

Chapter 2.

Crystallization of *B. subtilis* SipS

2.1. Introduction

This chapter presents the protocols for *B. subtilis* SipS protein overexpression and purification in preparation for use in crystallization, as well as the crystallization conditions that resulted the initial protein crystals.

2.2. Materials & Methods

2.2.1. Expression and Purification of *B. subtilis* SipS Constructs

Four *B. subtilis* SipS constructs previously made by Daniel Chiang, were available in the Paetzel lab. The constructs include SipS Full-Length (FL), SipS Δ2-35 Wild-Type (WT), SipS Δ2-35 S43A, and SipS Δ2-35 K83A (Table 2-1). The residues 2 to 35 in *B. subtilis* SipS represents a proposed transmembrane segment and the truncated constructs (WT, S43A and K83A) therefore contain a "soluble" catalytic domain which presumably will crystallize more readily without the hydrophobic and flexible transmembrane anchor. The constructs were prepared and purified using the steps described in Figure 2-1. The SipS FL and truncated constructs were prepared and purified in a similar manner except the use of the detergent. Through sonication and Avestin, the SipS FL was first lysed using the buffer without detergent whereas the truncated SipS Δ2-35 constructs were lysed with non-ionic detergent. This is due to the presence of transmembrane domain in the SipS FL construct. The SipS FL were also subsequently dounce homogenized after the first lysis while other constructs were directly subjected to centrifugation. After spin-down, the soluble cell lysate was

obtained for each of the constructs, and then passed through a Ni^{2+}-NTA column, followed by gel filtration to remove unwanted protein contamination. After visualization using SDS-PAGE, the purified protein was concentrated to 15mg/mL and stored at -80º C. The purification protocol for each protein construct is explained in detail in the following section.
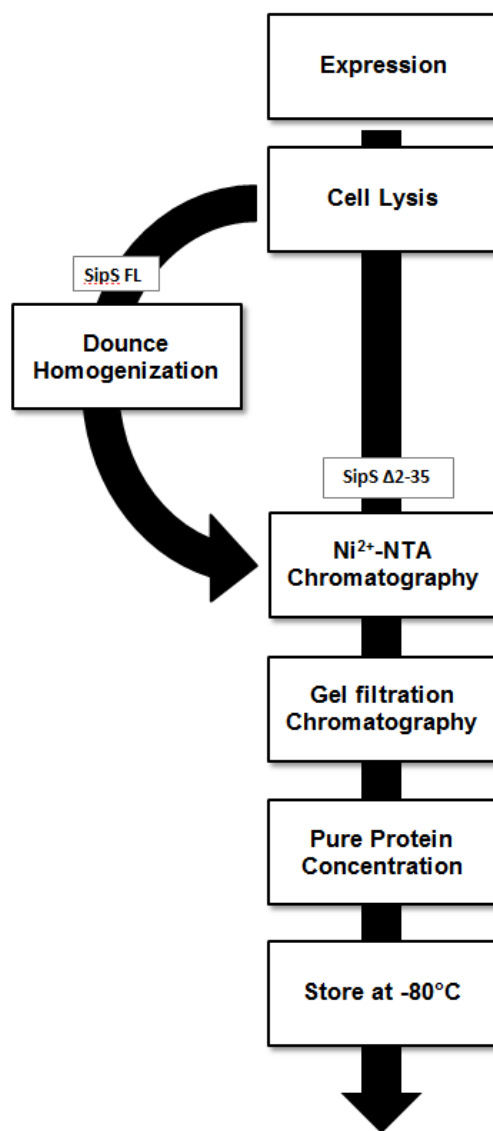
Figure 2-1   **Outline of the experimental procedure for SipS protein isolation and purification**

Table 2-1.   *B. subtilis* SipS constructs that are available at the Paetzel lab

SipS construct	Tag	# (AA)	MW (kDa)	Construct
Full-length (FL)	6XHis	190	21.9	N—[HHHHHH]—[2-35]—[36-184]—C
Δ2-35 Wild-type (WT)	6XHis	157	18.1	N—[HHHHHH]—[36-184]—C
Δ2-35 S43A	6XHis	157	18.1	N—[HHHHHH]—[S43A]—C
Δ2-35 K83A	6XHis	157	18.1	N—[HHHHHH]—[K83A]—C

#AA shows the number of amino acid residues in the enzyme including the N-terminal hexahistidine tag. MW indicates the molecular weight of the protein plus the hexahistidine tag. For the SipS FL construct, the green tag represents the hexahistidine tag, the red box represents a transmembrane domain, and the blue region represents the soluble domain that contains the catalytic active site.

2.2.2.   Overexpression and Ni^{2+}-NTA Purification *B. subtilis* SipS FL

A colony, grown up from a glycerol stock of the *B. subtilis* SipS FL construct in BL21(DE3), was inoculated into 100mL Luria-Bertani (LB) media containing 100µg/ml of the antibiotic kanamycin and grown overnight at 37º C while shaking at 250 rpm. The next morning, 50mL of the overnight culture was transferred into 1L LB media containing 100µg/ml kanamycin. The culture was allowed to grow for 3 hours at 37º C on a 250 rpm shaker. When the optical density (OD_{600}) of the culture reached a value of 0.6, 0.5mM of the inducer isopropyl-β-D-thiogalactopyranoside (IPTG) (BioShop Canada Inc.) was added to facilitate protein overexpression. The culture continued to grow for 5 hours after induction. Following the 5 hour induction, the overexpressed cells were centrifuged at 6,000 rpm for 7 minutes at 4º C and the pellet was harvested while the supernatant was removed and discarded. The pelleted cells were stored at -80º C until required for protein purification.

To isolate the overexpressed *B. subtilis* SipS FL protein from the cell pellet, the pellet was first resuspended in a lysis buffer (TBS X1: 20mM Tris-HCl, pH 8, 100mM NaCl). Detergent was purposely omitted in the buffer so that the cell membranes could

be preserved, maintaining the transmembrane proteins that reside in it. The suspended cells were then sonicated for 10 seconds three times each followed by a 10 second break. Lysis was continued using an Avestin Emuliflex-3C cell homogenizer. After having been passed through the Avestin homogenizer, the lysate was centrifuged at 14,000 rpm for 40 minutes and the pellet was saved while the supernatant was discarded. This time, a different lysis buffer (20mM Tris-HCl, pH 8, 100mM NaCl, 0.1% Triton X-100) was added to the pellet for resuspension. This buffer contained the Triton X-100 detergent to solubilize the protein once the resuspended solution was gently dounce homogenized about 30 times. Following the dounce homogenization, the cell lysate was again centrifuged at 14,000 rpm for 40 minutes. The resulting supernatant was collected and stored at 4° C for later Ni^{2+}-NTA column purification while the pellet was discarded.

Before the purification step, the Ni^{2+}-NTA beads were previously washed with 5mM of imidazole first, and then with 35mM of imidazole, followed by a buffer solution. The supernatant containing the cell lysate protein (1X TBS: 20mM Tris-HCl, pH8, 100mM NaCl, 0.1% Triton X-100) was incubated with Ni^{2+}-NTA beads overnight at 4° C on a slow shaker and then loaded onto a column. The initial flow through was discarded. The column was then washed first with 5mM imidazole in 1X TBS buffer and subsequently with 35mM imidazole in 1X TBS buffer. The flow through was discarded. The protein bound to the Ni^{2+}-NTA beads was then eluted using four serial elution buffers: 100mM, 250mM, 500mM and 1000M of imidazole in 1X TBS buffer and the fractions were collected.

Aliquots were taken from each fraction and individually combined with 2 x loading buffer in a 1:1 ratio (10μL:10μL). The mixtures were heated at 90° C for 3 minutes and then 10μL of each sample was loaded onto a 15% SDS-PAGE gel to check the presence and quality of the purified protein. The elution fractions that contained the protein were kept at 4° C for later purification by size-exclusion chromatography.

2.2.3. Overexpression and Ni^{2+}-NTA Purification *B. subtilis* SipS Δ2-35 WT

B. subtilis SipS Δ2-35 WT protein was overexpressed as described in the protocol presented in the first paragraph in Section 2.2.2. However, the induction of SipS Δ2-35 WT by IPTG took 4 hours rather than 5 hours.

To isolate the overexpressed *B. subtilis* SipS Δ2-35 WT protein from the cell pellet, the pellet was resuspended in lysis buffer (20mM Tris-HCl pH 8, 100mM NaCl, 0.025% Tween-20). The resuspended cells were then sonicated for 10 seconds three times each followed by a 10 seconds break. Lysis was continued using the Avestin Emuliflex-3C cell homogenizer. After having been passed through the Avestin homogenizer, the lysate was centrifuged at 14,000 rpm for 50 minutes. The pellet was discarded and the supernatant was reserved for purification.

The Ni^{2+}-NTA beads in the column was first washed with 5mM of imidazole and then subsequently with 35mM imidazole which was finished by a buffer wash. The supernatant containing the cell lysate protein (20mM Tris-HCl, pH8, 100mM NaCl, 0.025% Tween-20) was then passed through the column and the initial flow through was collectedthen reloaded onto the same column. This step was repeated to assure that all of the target proteins are bound to the Ni^{2+}-NTA beads in the column. The protein bound to the Ni^{2+}-NTA beads was then eluted using four serial elution buffers: 100mM, 250mM, 500mM and 1000M of imidazole in 1X TBS buffer and the fractions were collected.

2.2.4. Overexpression and Ni^{2+}-NTA Purification *B. subtilis* SipS Δ2-35 S43A

The SipS Δ2-35 S43A construct was overexpressed as described in Section 2.2.2. except the induction by IPTG took 4 hours rather than 5 hours. The overexpressed protein was then purified using Ni^{2+}-NTA chromatography as described in Section 2.2.3.

2.2.5. Expression and Ni²⁺-NTA Purification *B. subtilis* SipS Δ2-35 K83A

The construct SipS Δ2-35 K83A was overexpressed as described in Section 2.2.2 except the induction by IPTG took 4 hours rather than 5 hours. The overexpressed proteins were subsequently purified by Ni²⁺-NTA chromatography using the protocol described in Section 2.2.3. However, the lysis and the purification buffers were different – instead of 1X TBS buffer with 0.025% Tween-20, the buffer for this construct contained 20mM Tris-HCl pH 8, 100mM NaCl, 0.025% Tween-20 and 0.005% DDM.

2.2.6. Purification Using Size Exclusion Column Chromatography

The Ni²⁺-NTA purified *B. subtilis* SipS constructs, suspended in 5 mL fractions of 1X TBS buffer with their respective detergents and containing 100mM, 250mM, 500mM or 1M imidazole, were applied individually to a pre-packed S-100 size exclusion column (HiPrep TM 26/60, Sephacryl TM S-100, Amersham Biosciences) connected to an ÄKTA Prime system (Pharmacia). The column was pre-equilibrated with the same buffer each construct used for lysis. The presence and purity of each isolated protein was displayed as peaks on a chromatogram. The fractions under these peaks were collected and run on 15% SDS-PAGE gels. The appropriate sample fractions were combined and then concentrated to 10mg/mL-15mg/mL using a stirred cell (Amicon) with a YM-10 membrane (Amicon, 10 KDa molecular weight cut off ([MWCO], Millipore). The concentration of the purified protein was verified in triplicate using the BCA Assay.

2.2.7. Limited Proteolysis

The SipS Δ2-35 K83A (1 mg/mL) was digested with four different proteases – subtilisin, trypsin, chymotrypsin and thermolysin (0.1 mg/mL) – in order to create a smaller protease resistant core. The proteolysis reaction was carried out at room temperature (25º C) with the ratio of the enzyme to the protease being 1000:1. To follow the proteolysis progress over time, aliquots of the proteolysis reaction were stopped at specific time points by adding SDS-loading dye containing β-

mercaptoethanol. The time points where the reaction was arrested were 0, 5, 10, 15, 30 minutes, 1, 2 and 3 hours. The reaction-loading dye mixture was then heated for 10 minutes run on 15% SDS-PAGE gel to visualize the digestion progress.

2.2.8.   Protein Crystallization, Optimization and Negative Control

Crystallization conditions were created using the vapour diffusion method in the sitting drop format. 0.5μL of *B. subtilis* SipS protein (11mg/mL-15mgmL) was plated on the pedestal with 0.5μL of reservoir solutions from PACT, JCSG+, PEG/ION, CS1, and CS2 Screens, respectively. 1ml of reservoir solution was pipetted into each well surrounding the pedestal and then each condition was completely sealed with transparent plastic tape. Screening the crystal plates through a microscope was done on a weekly basis. If initial hits or crystal needles were found, the condition of the reservoir solution was optimized in order to create a crystal of better quality.

2.2.9.   Co-crystallization

Arylomycin A2 stock solution was made at 20mM in 100% DMSO and was stored at -80º C After crystal needles were observed for the SipS Δ2-35 WT construct, the protein was combined with the inhibitor arylomycin A2 in the purification buffer before being set up in crystal plates. The mixture of the protein and the inhibitor was 1:1 molar ratio. The complex was kept for 30 minutes on ice and then was crystal plated under the same conditions that generated the original needles.

2.2.10.   Additive Screening

After needles were found for the SipS Δ2-35 WT construct, the protein was screened with additives under the same condition that generated the orginal crystals. Once crystallization droplets were made (1uL of protein (14.2mg/mL) + 1uL of reservoir), 1uL of each additive reagent (10mM) from Additive Screen (Hampton Research) was added to individual wells. The wells were then sealed and screening for hits was done on a daily basis.

2.3. Results

2.3.1. Ni^{2+}-NTA Affinity Chromatography & Size Exclusion Chromatography

In order to purify a protein of the construct, Ni^{2+}-NTA affinity chromatography was used as a first step in the purification process. The soluble cell lysate of the SipS construct (SipS FL, SipS Δ2-35 WT, SipS Δ2-35 S43A and SipS Δ2-35 K83A) was applied to a Ni^{2+}-NTA column and the flow through was collected. The protein bound to the beads was washed with 10 mM and 35 mM imidazole and then eluted in the 100mM, 250 mM, 500 mM and 1000 mM imidazole fractions (Figure 2-2). Although the protein was relatively pure comparing to the cell lysate, other unwanted proteins were also shown on the lane.

In order to purify the protein further, the proteins collected in imidazole fractions were loaded onto size-exclusion chromatography subsequently. The maximum volume of sample that the size-exclusion column can withstand is 5 mL, and hence the four imidazole fractions (~20 mL) were combined and then concentrated to 5 mL before loading onto the size-exclusion column. After the purification is done, one large peak was found near to a void volume with an approximate size of 76 kDa on the chromatogram. Fractions under the peak were subsequently collected and were run on SDS-PAGE gel to check the presence and quality (Figure 2-3). The SDS-PAGE gel analysis showed that the fractions contain the protein, plus the contaminants, implying the purification did not work.

Therefore, a new batch of cell lysate containing the SipS construct was prepared. The cell lysate was run on Ni^{2+}-NTA column and the fractions were collected in a same way. However, before the fractions were directed on a size-exclusion column, they were neither combined nor concentrated this time; each fraction was run on the column individually, resulting four runs per batch of cell lysate. All of four fractions yielded three universal peaks this time. Elution volume and location of each peak were recorded and used in estimating molecular weight of the protein. In addition, samples from each peak were collected and run on a 15% SDS-PAGE gel in order to confirm the presence and the quality of the protein. Figrue 2-4 represents a chromatogram of

SipS Δ2-35 WT eluted with 500 mM imidazole, and the rest of the fractions had the same result and therefore are not shown. The protein represented by Peak 1 in the chromatogram was around 50kDa in size and the samples run on the SDS-PAGE gel show that the peak represents aggregates of SipS Δ2-35 WT and an unknown protein contaminant. The protein in Peak 2 was the purified, monomeric form of SipS Δ2-35 WT with molecular weight being approximately 18kDa. The last peak was measured to be around 3kDa although when the fractions of this peak were run on an SDS-PAGE gel, no band was seen. The third peak is assumed to be aggregates of imidazole. The truncated mutants SipS Δ2-35 S43A and SipS Δ2-35 K83A showed two major peaks in the chromatogram, the second peak being the purified monomeric protein (Figure 2-5 and Figure 2-6). For the SipS FL, two peaks were observed in the chromatogram. The first peak was the purified protein in a tetrameric form (Figure 2-7).

The fractions representing the purified protein peaks were collected, combined together and concentrated to 10–15 mg/mL. SipS Δ2-35 WT was concentrated to 14.2 mg/mL whereas SipS Δ2-35 S43A and SipS Δ2-35 K83A were concentrated to 11 mg/mL. SipS FL concentrated to only 3 mg/mL. Each of the concentrated protein samples were run on a 15% SDS-PAGE gelfor visualization. Approximately 500µL of each concentrated protein was stored at -80º C until needed for crystal plating.

It was learnt that SipS truncated construct prefer not to be concentrated when it is not completely pure, otherwise it would create a large oligomer containing the unwanted proteins.
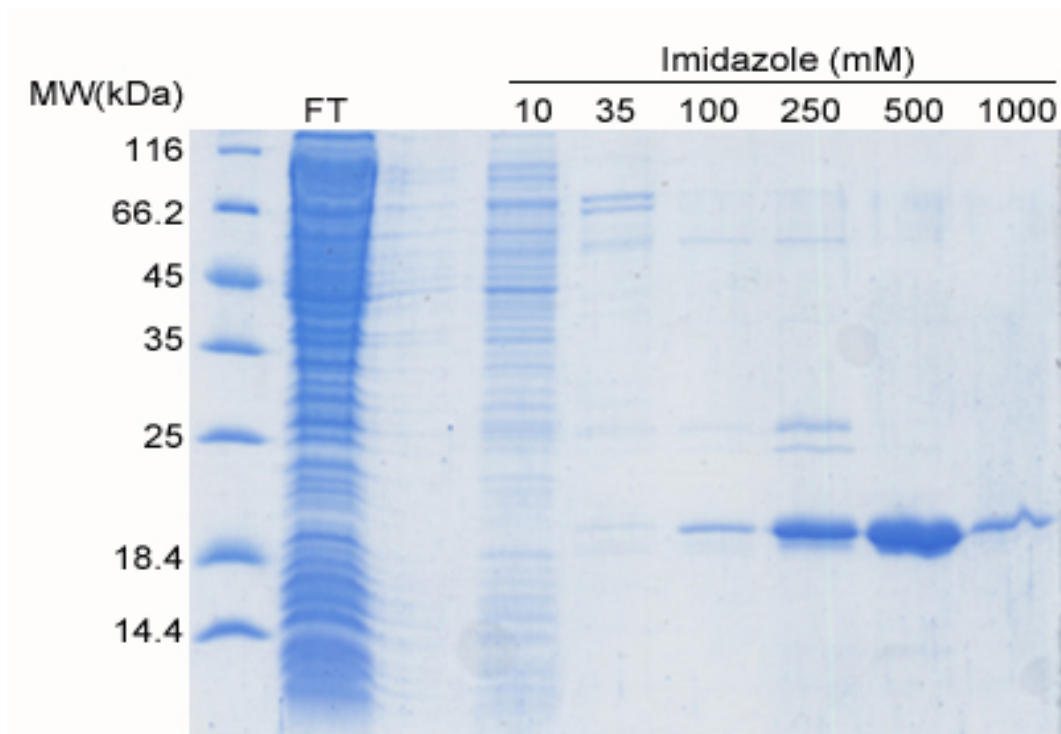
Figure 2-2   SDS-PAGE gel showing results of SipS Δ2-35 WT Ni^{2+}-NTA Column Purification

The figure shows the results of SipS Δ2-35 WT protein (from the cell pellet) after Ni^{2+}-NTA column purification. The supernatant (not shown on the gel) was loaded on to the column, and the flow through was collected. The protein bound to the beads in the column were then eluted out using 1 x TBS buffer with 0.025% Tween-20 with increasing concentrations of imidazole (10 mM, 35 mM, 100 mM, 250 mM, 500 mM and 1000 mM). Most of the SipS Δ2-35 WT proteins eluted in the imidazole concentrations of 100 mM, 250 mM, 500 mM and 1000 mM.
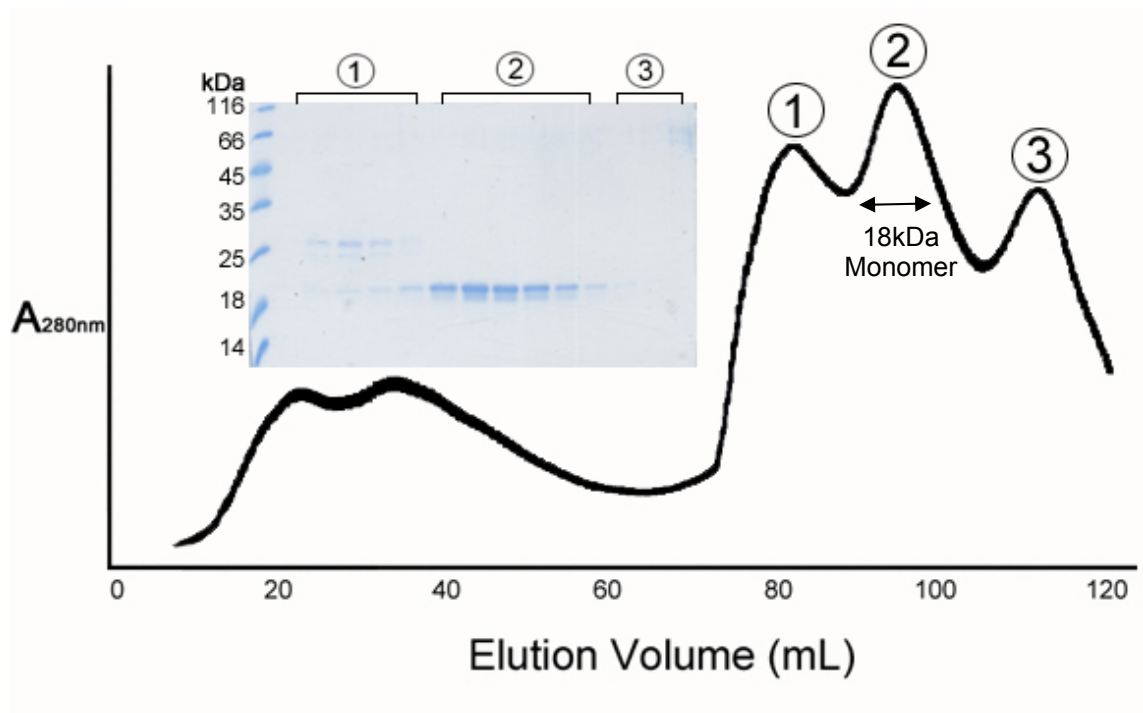
Figure 2-3   SDS-PAGE gel analysis of size-exclusion chromatogram of SipS Δ2-35 WT

The chromatogram shows one, large main elution peak. The fractions of each peak were collected and run on SDS-PAGE gel which is shown on the right on the chromatogram. The peak contains the protein and the contaminant which implies the sample did not get purified.

Column	Hi-Load Superdex 200 16/60
Void Volumn	45 mL
Column Volumn	120 mL
Flow Rate	1 mL/min
Mobile Phase	20 mM Tris-HCl (pH 8), 100 mM NaCl, 0.025% Tween-20

Figure 2-4   SDS-PAGE gel analysis of size-exclusion chromatogram of SipS Δ2-35 WT

The chromatogram of SipSΔ2-35 WT eluted with 500 mM imidazole shows three main elution peaks. The fractions of each peak were collected and run on SDS-PAGE gel which is shown on the left on the chromatogram. The first peak is found to be a contaminant, the second peak is the target protein and the third peak is presumed to be the imidazole from the previous buffer.
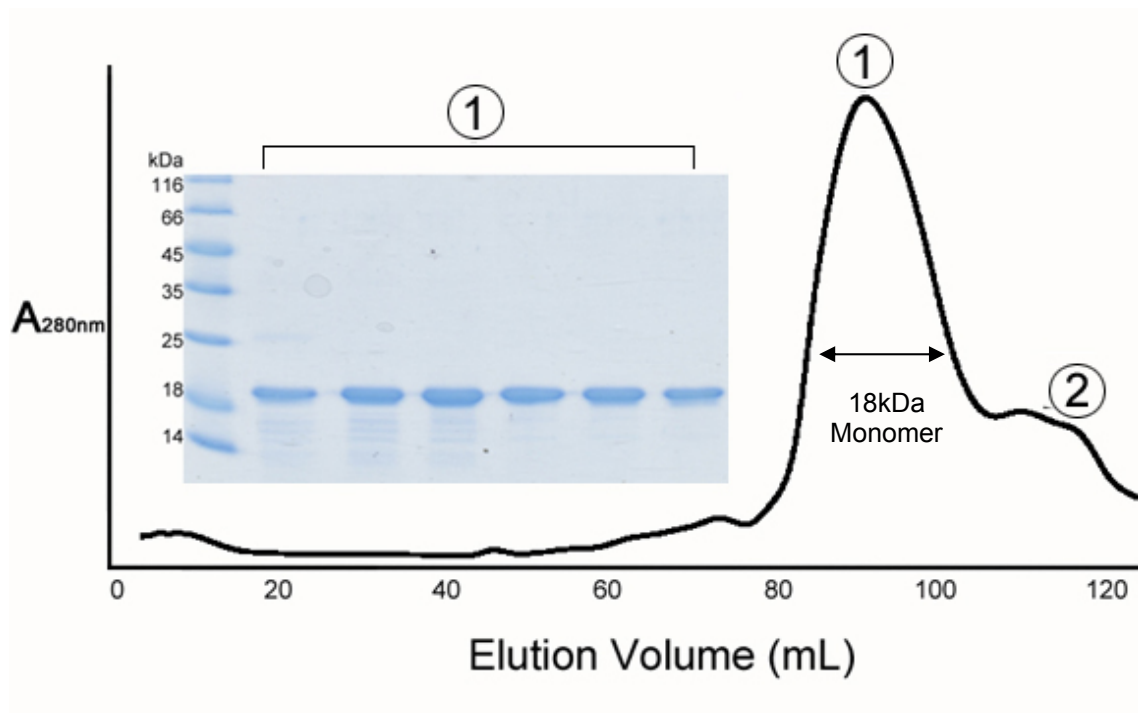
Column	Hi-Load Superdex 200 16/60
Void Volumn	45 mL
Column Volumn	120 mL
Flow Rate	1 mL/min
Mobile Phase	20 mM Tris-HCl (pH 8), 100 mM NaCl, 0.025% Tween-20, 0.01% DDM

Figure 2-5   SDS-PAGE gel analysis of size-exclusion chromatogram of SipSΔ2-35 K83A

The chromatogram shows two main elution peaks. The fractions of each peak were collected and run on SDS-PAGE gel which is shown on the left on the chromatogram. The first peak is the purified target protein and the second peak is presumed to be the imidazole from the previous buffer.
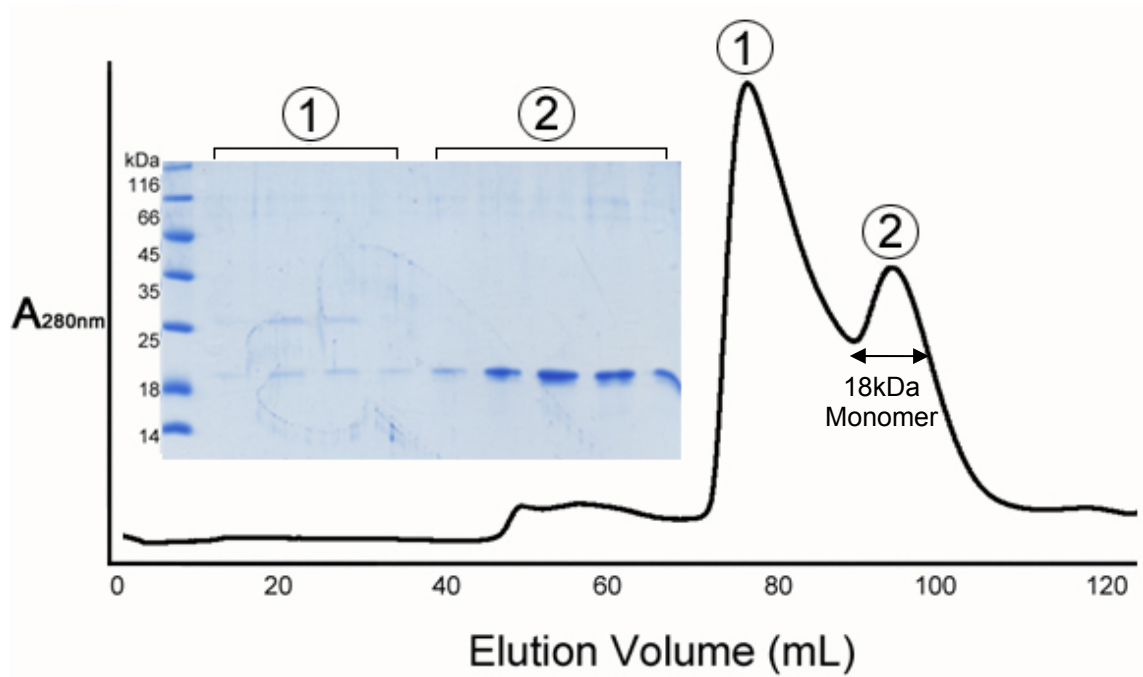
Column	Hi-Load Superdex 200 16/60
Void Volumn	45 mL
Column Volumn	120 mL
Flow Rate	1 mL/min
Mobile Phase	20mM Tris-HCl (pH 8), 100mM NaCl, 0.025% Tween-20

Figure 2-6   SDS-PAGE gel analysis of size-exclusion chromatogram of SipS∆2-35 S43A

The chromatogram shows three main elution peaks. The fractions of each peak were collected and run on SDS-PAGE gel which is shown on the left on the chromatogram. The first peak is a contaminant, the second peak is the target protein and the third peak is presumed to be the imidazole from the previous buffer.
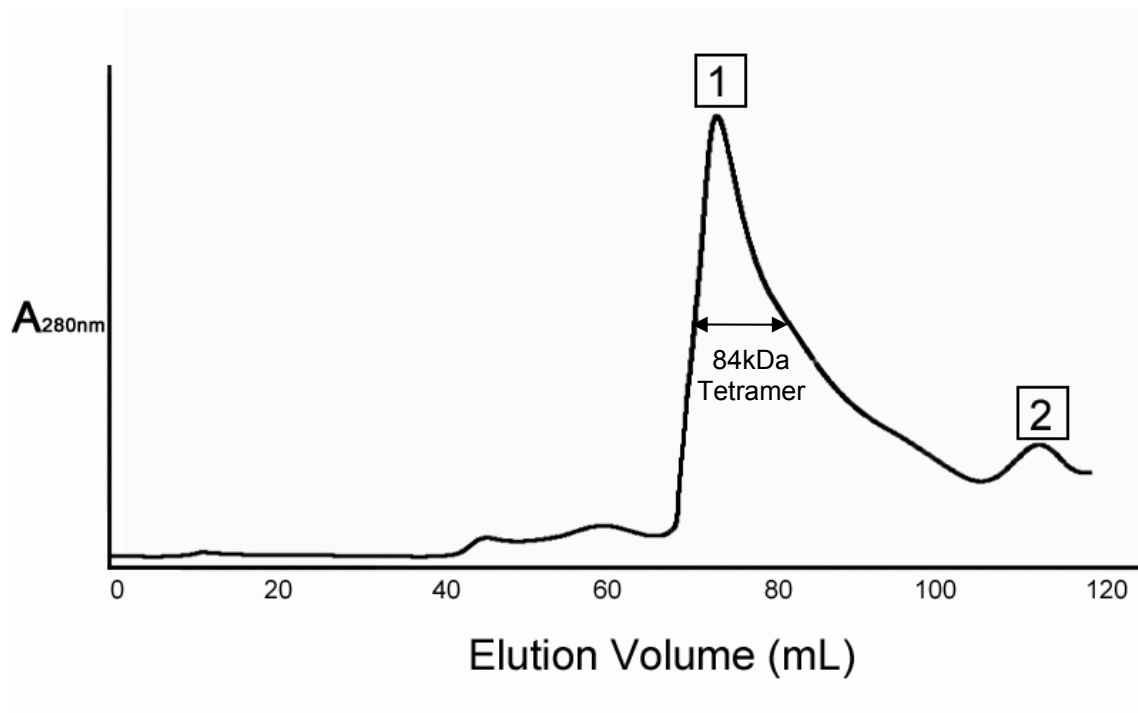
Column	Hi-Load Superdex 200 16/60
Void Volumn	45 mL
Column Volumn	120 mL
Flow Rate	1 mL/min
Mobile Phase	20 mM Tris-HCl (pH 8), 100 mM NaCl, 0.1% Triton X-100

Figure 2-7   SDS-PAGE gel analysis of size-exclusion chromatogram of SipS FL

The chromatogram shows two main elution peaks. The fractions of each peak were collected and run on SDS-PAGE gel which is shown on the left on the chromatogram. The first peak is the purified target protein and the second peak is presumed to be the imidazole from the previous buffer.
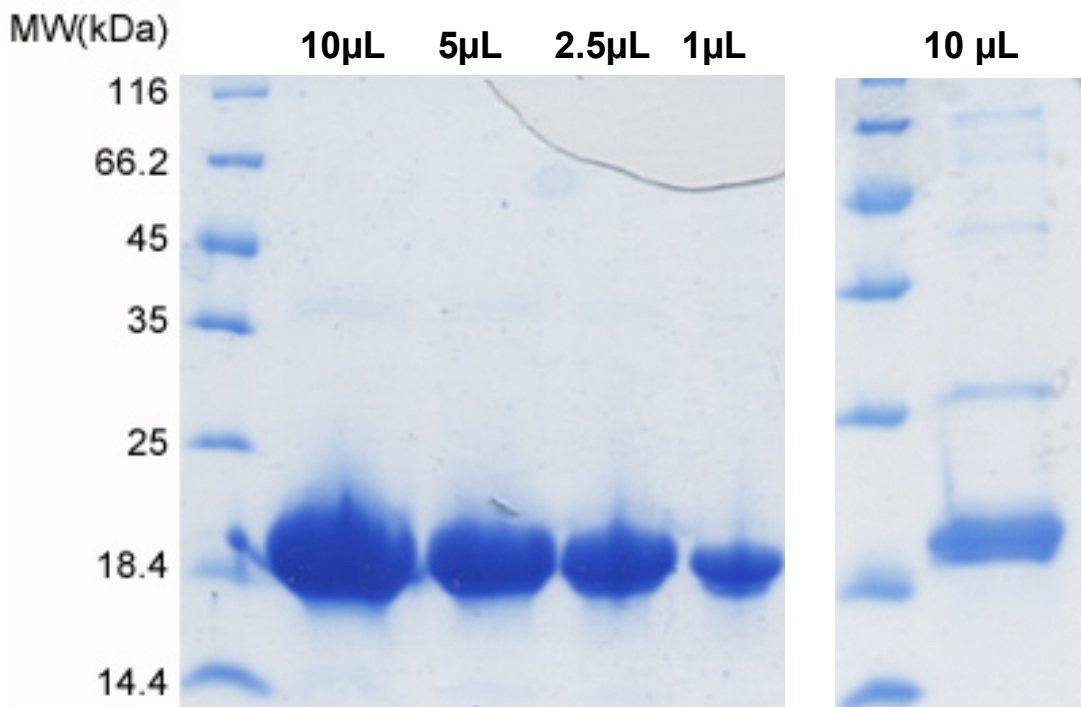
Figure 2-8 **SDS-PAGE gel showing the purity of column purified, concentrated SipS Δ2-35 WT with concentration of 14.2 mg/mL and 2 mg/mL**

The SDS-PAGEgel on the left shows the protein purified in buffer containing 0.025% Tween-20 buffer (14 mg/mL) whereas the gel on the right shows the protein purified in buffer with 0.1% Triton X-100 (2 mg/mL). The proteins in buffer with 0.1% Triton X-100 could not be concentrated further due to aggregation and precipitation. The bands shown at ~28kDa and at higher molecular weights on the right gel are the pronounced contaminants after concentration step.

2.3.2. Truncated SipS is Both Soluble and Active in Tween-20

Previously, all of the SipS constructs were purified in buffer containing 0.1% Triton X-100 using the purification method presented in 2.3.1. Ni-NTA column purification and size-exclusion column purification had the same results for both of the protein batches purified with 0.1% Triton X-100 and 0.025% Tween-20 but the purified protein with 0.1% Triton X-100 could not be concentrated over 2 mg/mL without precipitating (Figure 2-8). Increasing the salt concentration (20 mM Tris-HCl pH 8, 1 M NaCl, 0.1% Triton X-100) in the buffer allowed the protein to be concentrated to 10mg/mL, but as shown in the results of the catalytic activity assay, both SipS FL and SipS Δ2-35 WT lost its activity in the presence of such high salt. Purifying the truncated constructs in buffer containing the detergent 0.025% Tween-20 allowed concentration over 10mg/mL to be reached and while the mutants lost their enzymatic activity in the

presence of Tween-20, SipS Δ2-35 WT did not. The SipS FL protein could not be purified in buffer containing Tween-20, but retained its activity when dialyzed in buffer with Tween-20 (20 mM Tris-HCl pH 8, 100 mM NaCl, 0.025% Tween-20) after purification in the same buffer with 0.1% Triton X-100 .

2.3.3.   SipS is Vulnerable to Proteases

The SipS Δ2-35 K83A protein was subjected to digestion with different proteases to create a protease-resistant core that would improve crystallization outcomes. It is known that creating a clean protease-resistant core helps the packing of the proteins in the crystallization process. The proteases used for digestion were subtilisin, trypsin, thermolysin and chymotrypsin. After digestion and visualization on an SDS-PAGE gel, no singular, smaller bands were observed, indicating that the proteases did create a protease-resistant core but instead completed digested the SipS Δ2-35 K83A protein after 3 hours of digestion (Figure 2-9).
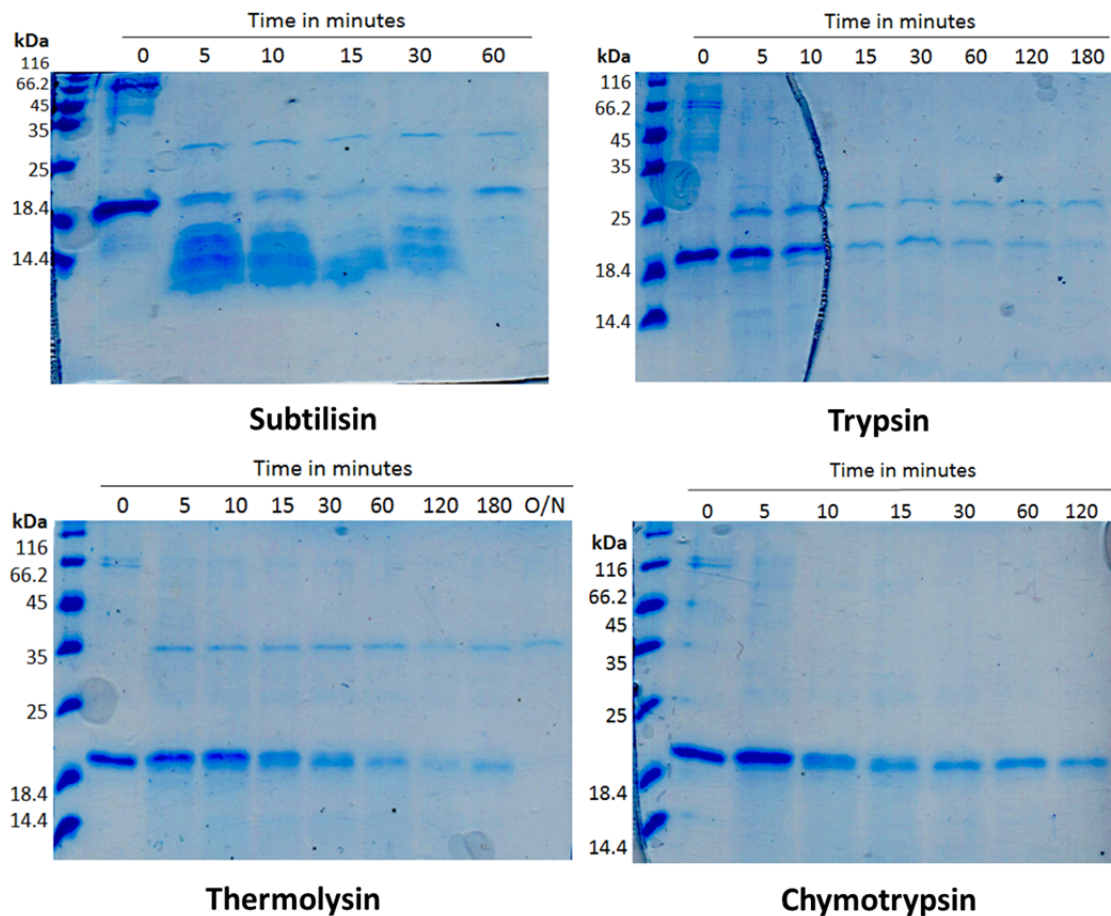
Figure 2-9 Limited proteolysis of SipS Δ2-35 K83A with different proteases

Four SDS-PAGE gels show the digestion progression of the SipS Δ2-35 K83A by different proteases. The protein was digested rapidly by subtilisin and slowly by trypsin and thermolysin. The protein seemed to be slightly resistant to chymotrypsin but no protease resistant core was generated by any of the proteases.

2.3.4. Crystallization of SipS Δ2-35 WT

The SipS Δ2-35 WT (14.2 mg/mL) purified in Tris buffer with Tween-20 (20 mM Tris-HCl 8, 100 mM NaCl, 0.025% Tween-20) yielded an initial hit in the well D7 of the JCSG$^+$ screen; a condition made up of 0.2 M ammonium sulfate, 40% PEG 400 and Tris-HCl pH 8. The crystal plates were incubated at 20º C and it took approximately 1 week for the crystal hits to appear. The condition was optimized to see if more clearly defined and larger protein crystals could be generated. The best crystal needles were obtained when the reservoir was optimized to contain 0.7 M lithium sulfate, 45% PEG

400 and Tris-HCl pH 8.5 (Figure 2-9). Additive screening and co-crystallization with Arylomycin A2 were carried out but no protein crystals have formed even after 12 weeks.
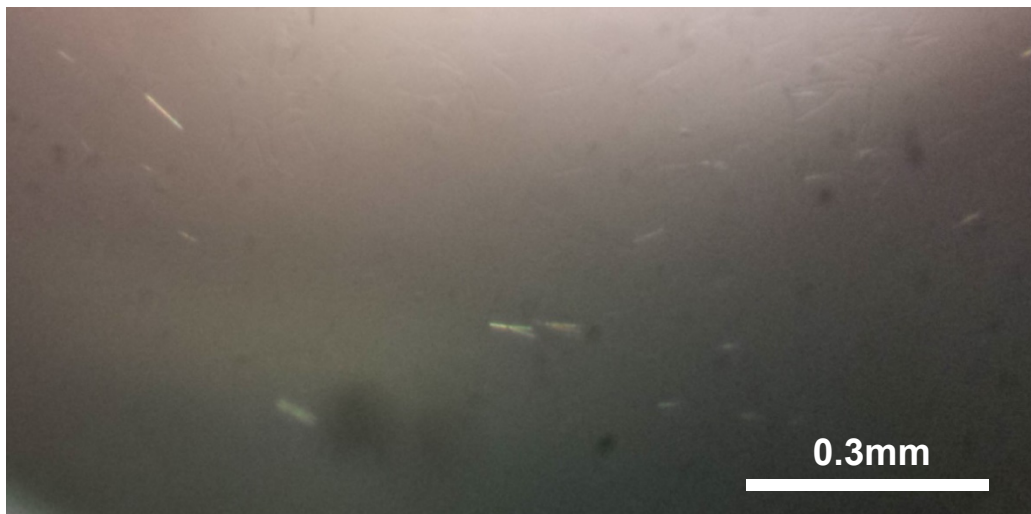
Figure 2-10 **Crystal needles of *B. subtilis* SipS Δ2-35 WT**

B. subtilis SipS Δ2-35 WT (14.2 mg/mL) formed needles after one week at room temperature when the reservoir condition was optimized to contain 0.7 M lithium sulfate, 45% PEG 400 and 0.1 M Tris pH 8.5. The crystallized needles were formed from a crystallization droplet containing 1uL of the protein sample and 1uL of the reservoir solution. The crystallization droplet was set up using the sitting drop format.

After the needle-shaped crystals were obtained, additive screening was subsequently performed: the crystallization droplet (2 µL) was introduced to 1 µL of each additive. It was found that the crystal needles changed in morphology and grew in size after the addition of 1 mM of hexaminecobalt (III) chloride (Cohex). The crystals did not seem quite ordered and consistent in size. More crystal plates with different concentration of Cohex were set up.

2.3.5. Crystallization of *B. subtilis* SipS K83A

Purified SipS Δ2-35 K83A protein (11 mg/mL) in 20 mM Tris pH 8, 100 mM NaCl, 0.025% Tween-20, 0.01% DDM yielded an initial hit in well # 31 of the PEG/ION Screen. The crystal plate was incubated at 20º C for 12 weeks before the crystals were observed. The condition contained 0.2 M lithium sulfate monohydrate, 20% w/v PEG 3350 (Figure 2-10).

Figure 2-11   Initial hits of SipS ∆2-35 K83A

An Initial hit was found for *B. subtilis* SipS ∆2-35 K83A (11 mg/mL) in the reservoir condition containing 0.2 M lithium sulfate monohydrate and 20% PEG 3,350 after 12 weeks at room temperature. The crystals formed in a crystallization droplet containing 1uL of the protein sample and 1uL of the reservoir solution. The crystallization droplet was set up using the sitting drop format.

2.4. Discussion

This section summarizes the crystallization conditions that generated needles of SipS ∆2-35 WT and initial hits for SipS ∆2-35 K83A. These promising crystallization hits can be further optimized in future studies.

Several optimizations were attempted during the purification step. When the fractions of protein were combined and concentrated together prior to the size-exclusion chromatography, the protein did not get purified and were eluted in the volume very close to void volume. This is because the protein formed aggregates with other contaminant during the concentration step. It is assumed that the detergent gets concentrated during the concentration step, causing the protein and the contaminant aggregate together. When the fractions did not get combined and concentrated, the protein was successfully purified as a monomer during the size-exclusion chromatography.

This optimized purification method was applied to the two batches of the protein. One was SipS Δ2-35 WT purified with buffer containing Tween-20 (20 mM Tris-HCl pH 8, 100 mM NaCl and 0.025% Tween-20) and the other one was the same construct in buffer with Triton X-100 (20 mM Tris-HCl pH 8, 100 mM NaCl and 0.1% Triton X-100). It was found that the construct purified with 0.1% Trtion X-100 could only reach up to 2 mg/mL during the concentration step. The protein got precipitated if concentrated further. The same protein in 0.025% Tween-20 buffer was able to be concentrated over 14.2 mg/mL. This is explained by the structure and the molecular mass of monomeric state of the detergents used. As shown in Appendix A, Tween-20 has a larger molecular mass both as a monomer and as a micelle. This seems to be caused by its larger hydrophobic tail than the one in Triton X-100, which can mask the hydrophobic surface area on the enzyme more efficiently. This effect not only made the enzyme reach high concentration, but allowed the enzyme to maintain its native structure and be catalytically active. A different detergent such as n-Dodecyl-β-D-Maltoside (DDM) was also shown to make the enzyme highly active. However, this was not specifically chosen for the further experiments since Tween-20 had the same effect.

Different optimization experiments were carried out to enhance crystal growth. A limited proteolysis study was carried out on the SipS Δ2-35 K83A protein to see if any protruding residues that could potentially disrupt the crystallization process, could be removed, leaving only a protease resistant core in a process called polishing. The resultant protease resistant core would be more rigid than the original protein and hence more likely to crystallize. However, all four proteases tested completely digested the protein. The four proteases—trypsin, subtilisin, chymotrypsin and thermolysin-- prefer to catalytically cleave peptide bonds formed with hydrophobic amino acids. *E. coli* SPase I has been shown to have a large, exposed hydrophobic surface and does not resist complete digestion by these proteases, and from the results of this experiment, it is likely that *B. subtilis* SipS has the same structural property. This result can be due to the relatively high concentration of the protease in the limited proteolysis reaction or the broad range of time points. In the future, a lower concentration of the proteases could be used in the reaction to see if it results a core of the protein, also the time points can be taken more frequent.

The optimized condition that generated needles had 0.7 M Lithium sulfate, 45% PEG 400, 0.1 M Tris-HCl pH 8 whereas the initial hits were formed in the presence of 0.2 M lithium sulfate monohydrate and 20% PEG 3,350. A negative control was carried out to see if the reservoir solution interacts with the buffer, forming a salt crystal. The crystal droplet of negative control was clear, suggesting that the initial hits were proteins crystals. Lithium sulfate was found in both conditions that generated SipS Δ2-35 WT and SipS Δ2-35 K83A crystals. The lithium sulfate molecules may be immobilized between the protein molecules during crystal lattice formation, forming specific interactions. Varying the concentration of lithium sulfate may pinpoint a condition that promotes formation of crystals suitable for X-ray diffraction.

The inhibitor, Arylomycin A2, was co-crystallized with the SipSΔ2-35 WT protein in the condition that generated the initial hits, yielding precipitation but no initial hits nor crystals. In the future, the enzyme inhibitor complex should be screened, using crystallization screens available to see if another condition may generate crystals.

Chapter 3.

The Inhibitory Effect of Hexaaminecobalt (III) Chloride (Cohex) on the Catalytic Activity of SipS Δ2-35 WT

3.1. Introduction

My previous experiments show that hexaaminecobalt (III) chloride (Cohex) changes the shape of the needle shaped crystal of SipS Δ2-35 WT during crystallization optimization experiments. In order to test if Cohex promotes crystal growth by physically binding to SipS, catalytic activity assays on the enzyme was carried out with other chemical compounds. In this chapter, different chemical compounds – $NiCl_2$, $CoCl_2$, and hexaaminecobalt (III) chloride (Cohex) – were tested for possible inhibitory effects on the catalytic activity of the SipS Δ2-35 WT protein. Subsequent experiments – cross-linking experiments SipS Δ2-35 WT and size-exclusion chromatography – explored possible mechanisms involved in the inhibition of SipS Δ2-35 WT.

3.2. Material and Methods

3.2.1. Materials

The FRET peptide was synthesized by the Peptide Institute. The FRET peptide used in the kinetic experiment is Dodecanoyl-K(Dabcyl)NGEVAKAAE(EDANS)T-NH$_2$ which contains the recognition and cleavage site of *B.subtilis* SipS: A-X-A at P3-P2-P1.

3.2.2.   Measurements of Kinetic Constants

The FRET peptide "Dodecanoyl-K(Dabcyl)NGEVAKAAE(EDANS)T-NH$_2$" was available in the Paetzel lab as a 1 mM stock solution dissolved in 100% DMSO. The FRET peptide contains a fluorescent molecule called EDANS, covalently attached to glutamic acid located at P2' on the substrate. EDANS absorbs light at a wavelength between 300 and 400 nm and emits light at a wavelength range of 400 to 600 nm. A quencher molecule Dabcyl is also found in the FRET peptide, covalently attached to lysine located right after Dodecanoyl. Dabcyl absorbs light that is emitted by EDANS but can absorb any light that has wavelength within the range of 350 to 600nm. If an enzyme is not present or severely damaged, the FRET peptide does not get cleaved and therefore stays intact. In this case, the fluorescence that is emitted by EDANS gets quenched by Dabcyl that is located nearby and hence, there will be no fluorescence detected. If there is a fully functional enzyme that successfully cleaves the FRET peptide, the Dabcyl molecule will get detached from the peptide and will no longer be close to the EDANS. This will stop the fluorescence quenching, so that when EDANS emits light, a molecular device will detect and quantify the fluorescence. For this experiment, EDANS was hit by light at 340 nm and emitted light at 540 nm. Each cleavage reaction contained a 10 μM substrate peptide and 5 nM SipS Δ2-35 WT protein. The final reaction volume was made up to 100 μL with buffer (20 mM Tris pH 8, 100 mM NaCl and 0.025% Tween-20). A SpectraMax M5 (Molecular Devices) spectrofluorometer with a 96-well microplate format was used to measure the increase in fluorescence over the course of the reaction. The increase in fluorescence was recorded continuously for 5-10 minutes, although initial rates were obtained in the first 30 seconds to 1 minute of the reaction. A linear graph of Fluorescence (RFU) vs. Time (s) was created by a software program. The slope of the linear line was calculated to get an initial velocity of rate of the reaction with a unit of RFU/min.

3.2.3.   Crosslinking Experiment

The purified SipS Δ2-35 WT protein (14.2 mg/mL) in buffer (20 mM Tris-HCl pH 8, 100 mM NaCl and 0.025% Tween-20) was first diluted to 1.42 mg/mL with a phosphate buffer (100 mM phosphate pH 8, 100 mM NaCl, 0.025% Tween-20) and

then was dialyzed with the same phosphate buffer overnight at 4° C to remove any remaining Tris. For a negative control experiment, 40 μL from the dialyzed SipS Δ2-35 WT protein (1.42 mg/mL) was incubated with 5 μL 2% of glutaraldehyde in a 100μL of reaction. The final reaction volume contained 0.57 mg/mL of the SipS Δ2-35 WT protein and 0.1% glutaraldehyde filled up with phosphate buffer. 10 μL aliquots were removed from the ongoing reaction at different time points added to 10 μl 2 X SDS loading dye and then heated for 3 minutes at 90 °C to stop the reaction. For an experiment testing the role of Cohex, the reaction volume was set up the same except that 50 mM of Cohex was also put into it. 10 μL aliquots were taken out from the reaction in a same manner as the negative control experiment.

3.2.4. Size-exclusion Chromatography

For a negative control experiment, the column (Superdex 75 HR 10/30) was pre-equilibrated with buffer (20 mM Tris-HCl pH 8, 100 mM NaCl, and 0.025% Tween-20). The SipS Δ2-35 WT protein (500 μL, 1.42 mg/mL) was run on the column and the peaks were observed. Once the negative control experiment was completed, the column was pre-equilibrated with buffer (20 mM Tris-HCl pH 8, 100 mM NaCl, 0.025% Tween-20, 25 mM Cohex) containing Cohex. Before loading on to the column, the SipS Δ2-35 WT protein (500 μL, 1.42 mg/mL) was incubated with 25 mM Cohex for 30 minutes on ice. The sample mixture was then applied to the column which generated the elution peaks. The location of the resulting peaks on the chromatogram was compared to the ones on the chromatogram generated by the negative control experiment. This way, one could estimate if there is a change in the oligomeric state of the protein by the addition of Cohex.

3.2.5. Homology Model

The *B. subtilis* SipS homology model was created using a software program PHYRE with the *E. coli* SPase I Δ2-75 crystal structure (PDB: 1B12) as a template (Kelley & Sternberg, 2009).

3.3. Results

3.3.1. Hexaaminecobalt (III) Chloride (Cohex) Inhibits the Catalytic Activity of SipS Δ2-35 WT

To see if the catalytic activity of SipS Δ2-35 enzyme can be affected by additives, SipS Δ2-35 WT protein was incubated individually with $NiCl_2$, $CoCl_2$, and Cohex. $NiCl_2$ and $CoCl_2$ were chosen as control chemical compounds since they do not possess the hexamine groups found in Cohex. Yet, they all have the inert metal ion. This way, one could observe the effect of hexamine does onto the enzyme by comparing the results. It was expected that nickel and cobalt would have no effect as they are unable to form any electrostatic or hydrogen bonds. The initial rate of catalysis was measured using the SpectraMax5 with a FRET peptide substrate. The enzyme without additive showed a V_{init} value around 12,500 RFU/min; this value did not change in the presence of 10 mM $CoCl_2$ and slightly decreased in the presence of 10 mM $NiCl_2$. However, when 10 mM Cohex was incubated with the enzyme, SipS Δ2-35 WT activity decreased by ~75%. Follow-up experiments found that the inhibition is concentration-dependent (Figure 3-1B). The effect of inhibition on the enzyme increased as the concentration of Cohex increased.

Cohex contains six amine groups (Figure 3-2), which may have increased the pH of the reaction condition. The concern was that this change in pH may have inhibited the catalytic activity of the SipS WT rather than a direct chemical interaction between Cohex and the enzyme. In order to test if the inhibition is due to the pH of the reaction condition, the pH of 10 mM $NiCl_2$, $CoCl_2$, and Cohex in buffer (20 mM Tris pH 8, 100 mM NaCl and 0.025% Tween-20) was measured. The pH of 10 mM $NiCl_2$ was 7.64, 10 mM $CoCl_2$ was 7.84 and 10 mM Cohex was 8.26. The original experiment, which measured SipΔ2-35 WT enzyme's V_{init} at12,500 RFU/min was carried out in buffer at pH 8.0.
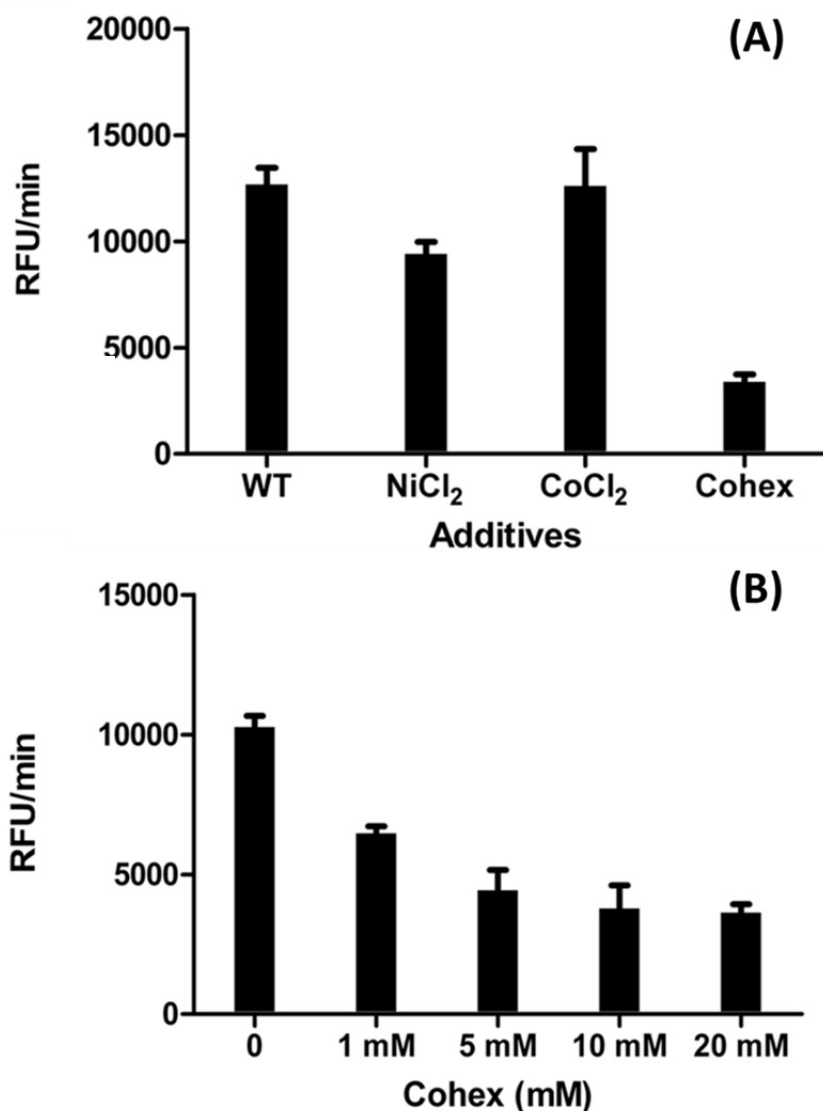
Figure 3-1   Activity profile of SipS Δ2-35 WT with different additives

(A) The catalytic reaction contained the SipS Δ2-35 WT with a concentration of 100 nM, the peptide of 1 μM and additives of 10 mM. The buffer for the reaction was and V_{init} for each reaction was measured in quadruplets, one outlier value was omitted and then the average value was calculated. The first bar represents average V_{init} value when no additive was added, and the rest of the bars are labeled with the additives added. The error bars represent standard deviations. (B) The catalytic reaction contained the SipS Δ2-35 WT (5 nM), the peptide (1 μM) and Cohex of 1 mM, 5 mM, 10 mM and 20 mM. The same reaction buffer was used and the measurement method was carried out the same way. Each bar represents an identical catalytic reaction with a different concentration of Cohex. The first 60 seconds were measured for V_{init} value.

67

Table 3-1.   List of chemical compounds used in the kinetic assay

Chemical Compound	Concentration (in 20 mM Tris pH 8, 100 mM NaCl, 0.025% Tween-20)	pH
$NiCl_2$	10 mM	**7.64**
$CoCl_2$	10 mM	**7.84**
Cohex	10 mM	**8.26**

3.3.2.   Cohex Does Not Facilitate the Dimeric Form of SipSΔ2-35 WT

The Size-exclusion chromatography experiment indicated that the truncated *B.subtilis* SipS construct (SipS Δ2-35) that lacks the transmembrane domain purifies in the monomeric form. Since Cohex inhibits the catalytic activity of the SipS Δ2-35 WT and promoted crystal growth of SipS Δ2-35 WT, it was hypothesized that Cohex may play a role in dimerizing the SipSΔ2-35 WT protein. A cross-linking experiment was designed to test this hypothesis. Results showed that the monomeric protein could form dimers and trimers after 5 minutes incubation, in both the presence and absence of 50 mM Cohex (Figure 3-2). This indicates that Cohex does not facilitate dimerization of the protein.

Two separate experiments were carried out to confirm the role of Cohex in the dimerization of SipS Δ2-35 WT protein: a negative control experiment to test the protein dimerization in the absence of Cohex and an experiment to screen for the protein dimerization with 25 mM Cohex. For the negative control experiment, the SipS Δ2-35 WT protein (0.5 mg/mL) was applied to a SEC column that was pre-equilibrated with buffer (20 mM Tris-HCl pH8, 100 mM NaCl, and 0.025% Tween-20). The three main peaks were observed in the chromatogram where the first peak is considered to be the monomeric SipS Δ2-35 WT protein (18 kDa) based on the standard curve (Figure 3-3). This proves that the SipS Δ2-35 WT protein is normally in the monomeric state unless disrupted by the chemical compounds. For testing the role of Cohex in dimerizing the enzyme, the sample mixture containing the SipS Δ2-35 WT protein (0.5 mg/mL) and 25 mM of Cohex was incubated for 30 minutes on ice before being loaded onto the

column. The result shows that, despite the presence of Cohex, SipS Δ2-35 WT is eluted out of the column as a monomeric form (Figure 3-3B). The first peak in the chromatogram was found at the exactly the same elution volume as in the negative control, representing the unchanged monomeric state of the enzyme. This indicates that there is no dimerization of the protein in the presence of Cohex, and hence Cohex does not facilitate the dimerization of the enzyme when it inhibits its catalytic activity. In addition, its role in promoting protein crystal growth is not due to the dimerization of the enzyme.
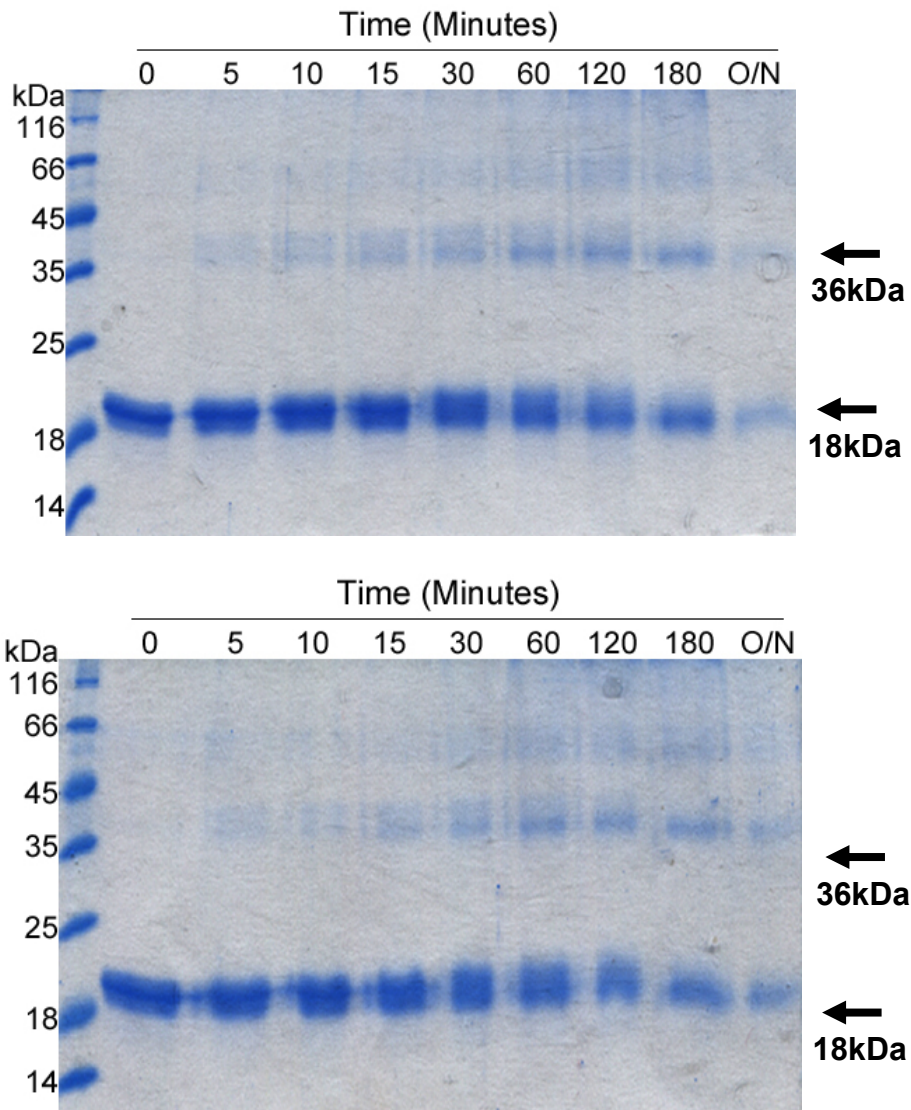
Figure 3-2 **SDS-PAGE gel showing SipS Δ2-35 WT forming dimers in the absence and the presence of 50 mM hexaaminecobalt (III) chloride (Cohex).**

The first SDS-PAGE gel shows the tendency of the protein to oligomerize in the   absence of Cohex. The protein started to dimerize just after 5 minutes, and by 60 minutes, a distinct band of dimerized protein band was shown. The second SDS-PAGE gel, representing a dimerization of protein in the presence of 50 mM Cohex, shows the exactly same oligomerization pattern as the negative control.
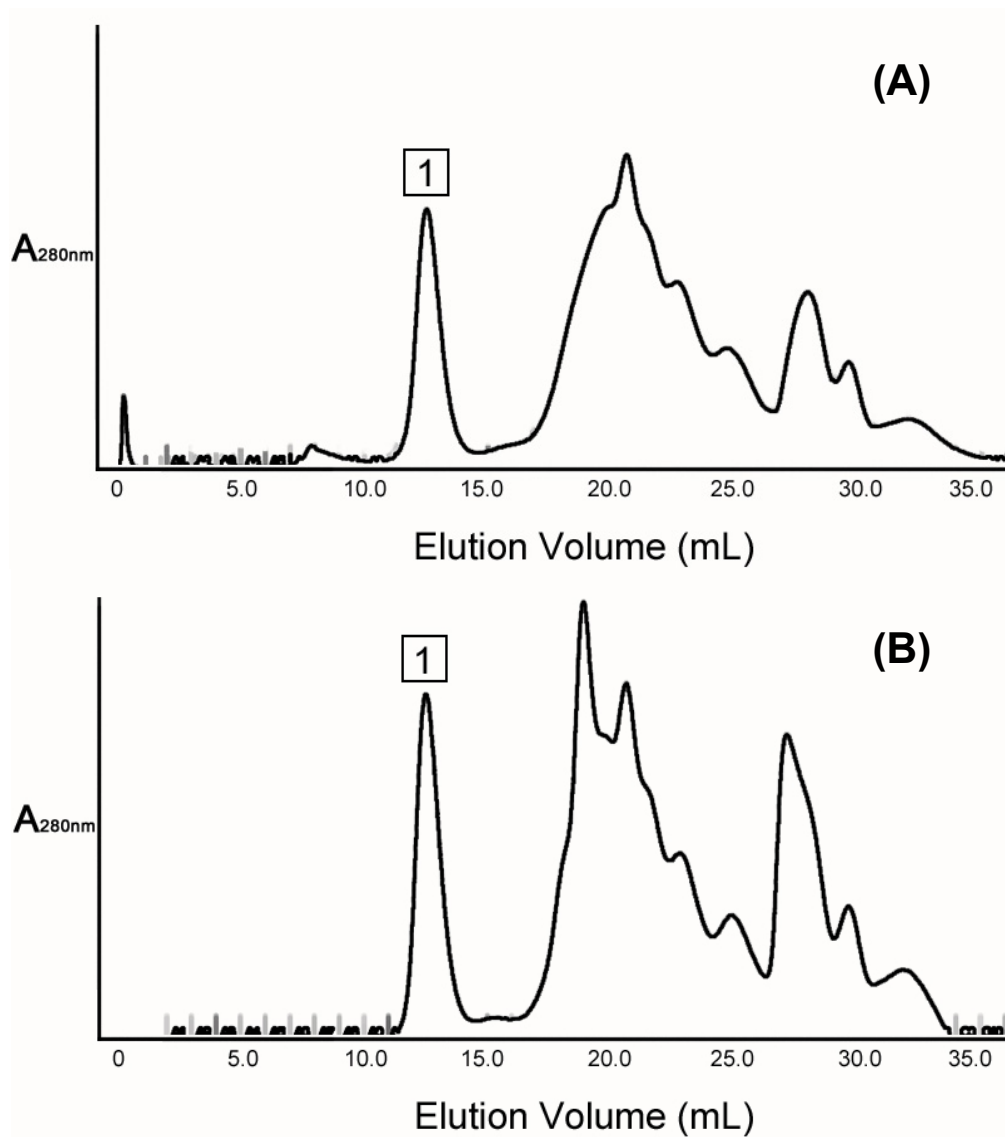
Column	Superdex 75 HR 10/30
Void Volumn	7 mL
Column Volumn	23.6 mL
Flow Rate	0.5 mL/min
Mobile Phase	**(A)** 20 mM Tris-HCl (pH 8), 100 mM NaCl, 0.025% Tween-20
	(B) 20 mM Tris-HCl (pH 8), 100 mM NaCl, 0.025% Tween-20, 25 mM Cohex

Figure 3-3   The chromatograms of SipSΔ2-35 WT in the absence and the presence of Cohex

Legend is on next page.

The first chromatogram (A) shows SipS Δ2-35 WT purification without Cohex whereas the chromatogram at the bottom (B) represents the same enzyme with 1 mM of Cohex. Among the three curves, the first peak represents the protein which is found at the same elution volume for both the experiments. This indicates that the protein is at the same molecular weight (18 kDa) and at the monomeric state.

3.4. Discussion

The catalytic activity of *B. subtilis* SipS Δ2-35 WT enzyme was affected by the addition of 10 mM Cohex but unaffected by $NiCl_2$ or $CoCl_2$. Cohex has six amine groups (Figure 3-6) so to test if its presence increased the pH of the reaction mixture, pH readings were taken of Cohex and the other two additives, separately in the reaction buffer. pH screening revealed that the pH change brought about by the addition of Cohex to the buffer is minimal, with an increase of 0.26 pH units. The pH change brought about by 10 mM $CoCl_2$ was a decrease of 0.16 units and 10 mM $NiCl_2$ decreased the pH by 0.36. The catalytic activity of the enzyme was unchanged in the presence of each of these two additives. Although the addition of Cohex to the buffer increased the pH to 8.26, it is difficult to conclude that pH caused the significant decrease in V_{init} since the pH increase was only 0.26 units. Overall, this implies that the inhibition of the enzyme by Cohex is not due to the pH change in the environment.

The SipS Δ2-35 WT protein normally exists in the monomeric state but there was a change in the crystal needles of the protein when Cohex was added into the crystallization drop. Since the most of SPases I have been crystallized as a dimer in the past, it was asked if Cohex induces dimerization of the protein, hence promotes the crystal growth of the enzyme. A cross-linking experiment and a size-exclusion chromatography experiment were carried out to determine if Cohex causes oligomerization of the enzyme and results indicated that the enzyme remains in the monomeric form regardless of whether or not Cohex is present. This implies that Cohex catalytically inhibits the enzyme by a different mechanism, possibly by physically binding to it.

Previous research has shown that a protein channel TolC is inhibited by Cohex through its electrostatic interaction with negatively charged aspartic acid residues residing in the entry of the channel. In addition to the electrostatic interactions, Cohex

can also form a numerous, effective hydrogen-bonding interactions with anions such as O^- and Cl^- (Sharma et al., 2006). Several mechanisms of interaction can be speculated based on the characteristic of the SipS Δ2-35 WT and Cohex.

Cohex can possibly form a hydrogen bond with polar molecules or form an electrostatic interaction with negatively charged residues on the enzyme. To find the potential residues on the enzyme that can interact with Cohex, a homology model of *B. subtilis* SipS Δ2-35 WT was created based on the structure of *E. coli* SPase I Δ2-75 using PHYRE (Figure 3-4). The catalytic domain of *B. subtili*s SipS has 30.67% sequence identity to the catalytic domain of *E. coli* SPase I, and showed 100.0% confidence and 99% coverage in the homology model. The cartoon representation of the homology model represents that the enzyme contains numerous negatively charged residues on the surface and the active site contains the two polar serine residues and the general base lysine (Figure 3-4). The surface representation of the predicted model was subsequently generated with vacuum electrostatics in order to characterize the regions by charges (Figure 3-5). The red-colored surfaces on the enzyme represent the negatively charged residues, which can possibly associate with Cohex via electrostatic interaction, altering the structure of the enzyme, thereby inhibiting it.

The homology model also shows the architecture of the active site of the enzyme, which contains polar residues such as Ser43 and Ser151 (Figure 3-4). These polar serine residues, with their hydroxyl groups, can form hydrogen bonds with Cohex. However, it is unlikely due to the presence of Lys83 in the active site that may repel Cohex and the steric hindrance in the active site that can block Cohex entry.

It is also possible that Cohex decreases the catalytic activity by interacting with the substrate, the FRET peptide. The FRET peptide contains negatively charged glutamic acid residues which could potentially be involved in an electrostatic interaction with Cohex. Future experiments will be performed to test the mode of inhibition observed by Cohex on SipS.
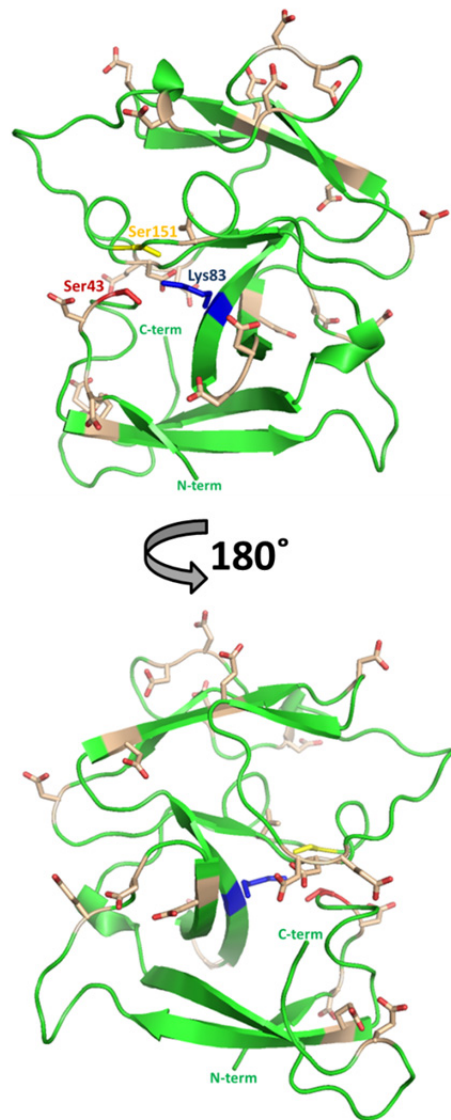
Figure 3-4   Cartoon representation of a homology model of *B. subtilis* SipS based on the *E. coli* SPase I Δ2-75 structure (PDB: 1B12).

The structure of *B. subtilis* SipS was predicted using the *E. coli* SPase I Δ2-75 as a template (Kelley & Sternberg, 2009, PDB:1B12). The SipS nucleophilic Ser43 (red stick) and general base Lys83 (blue stick) are located close to each other in the active site. The enzyme contains numerous negatively charged residues (Asp and Glu) on the surface of the enzyme which are shown as sticks in the figure.
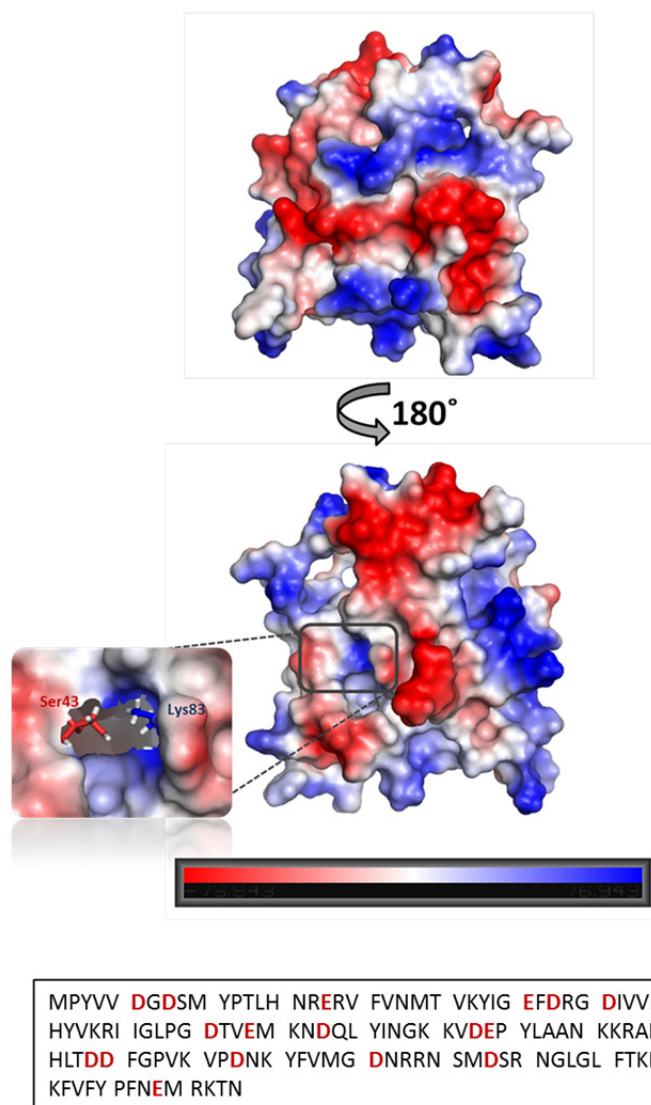
Figure 3-5   Surface representation of homology model of *B. subtilis* SipS with vacuum electrostatics

The homology model of *B. subtilis* SipS is recreated with vacuum electrostatics in the figure above. Red colored regions indicate the patch covered by negatively charged residues while blue colored regions are composed of positively charged residues. The negatively charged patch on the enzyme likely had electrostatic interactions with Cohex. The active site on the enzyme is labeled with a gray rectangle and enlarged to show the nucleophilic serine (Ser43) and the general base lysine (Lys83). The sequence for the residues in the homology model is shown at the bottom which lacks the transmembrane segment (residue 2-35). The pI for the sequence is 8.69. Negatively charged residues Asp and Glu are colored red in the sequence.
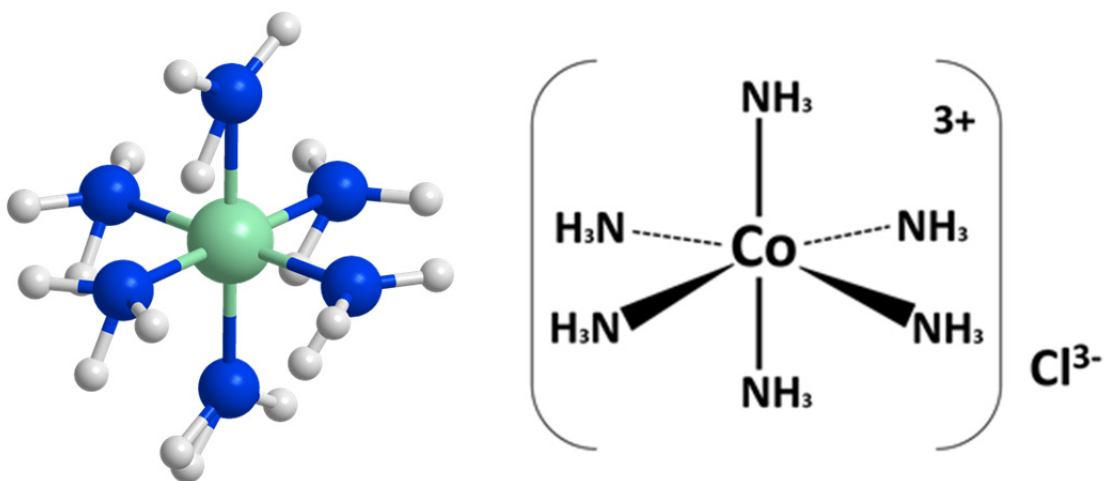
Figure 3-6   Octahedral coordination sphere and chemical structure of Cohex

The diagram represents 6-coordinate complexes with an octahedral geometry. The octahedron Cohex is represented as a planar square with each amine group above and below the plane. All positions on an octahedron are geometrically equivalent.

Chapter 4.

Summary and Future Directions

B. subtilis is a very mild human pathogen among the Gram-positive bacteria, especially compared to *Bacillus anthracis* (*B. anthracis*) and *Bacillus Cereus* (*B. cereus*). However, *B. subtilis* can cause bacteremia, endocarditis, pneumonia, and septicemia. In addition, *B. subtilis* produces a high quantity of the toxic extracellular enzyme, subtilisin. Subtilisin is known to cause allergic or hypersensitivity reactions when individuals are exposed to it repeatedly in high concentrations. Determining the structure of *B. subtilis* SipS is important as it is a potential drug target. The structure of *B. subtilis* SipS will provide insight into the active site of the protein and other binding sites which can potentially be inhibited by novel antibacterial drugs.

Future research directions could include validation and optimization of the SipS Δ2-35 WT and SipS Δ2-35 K83A crystals, crystallization of other *B. subtilis* SPases I, testing of other chemical compounds for enzyme inhibition and studying the role of Ser151 in *B. subtilis* SipS.

4.1. Validation and Optimization of Protein Crystals

For the SipS Δ2-35 K83A mutant, the nature of initial hits need to be determined; are they protein crystals (protein) or inorganic (salt) crystals? The negative control results indicated that salts were not crystallizing within the time period that the wells were under regular observation so therefore protein is the likely source of the crystals. However, this method is not 100% reliable since it is not known when equilibrium between the crystallization droplet and the reservoir is reached; it is possible that crystallization in the negative control occurred after the observation period ended. This

leaves the make-up of the initial hits still in question. There are several other techniques that can be used to distinguish between organic and inorganic crystals. A direct method is simply to mount the crystal and irradiate it with a beam of monochromatic X-rays. The diffraction pattern will identify whether the crystal is salt or protein. A salt crystal is composed of very small unit cells and will generate very few reflections whereas a protein crystal will produce many reflections. Another method which could be carried out in future is SDS-PAGE visualization wherein the crystals are collected, washed, dissolved and run on a 15% SDS-PAGE gel. If a band of the expected molecular weight appears on the gel, there is a good chance that the crystals are protein. If no band is observed, the crystals were likely salt. In addition, the crystals can also be physically manipulated. Salt crystals are densely packed with little solvent between the salt molecules, making the salt crystal very difficult to crush, whereas protein crystals are fragile due to their loose packing and large solvent channels between the protein molecules within the crystal. Both the SipS Δ2-35 WT and SipS Δ2-35 K83A crystals had lithium sulfate (Li_2SO_4) in their conditions, indicating that lithium sulfate may interact with the SipS constructs, facilitating crystal growth. Optimization of the two conditions, by decreasing and increasing the concentration of lithium sulphate, may promote better crystal formation and growth.

Another series of future experiments could be modification of the purification protocols, with the goal of achieving higher protein concentrations. Experiments have shown that SipS Δ2-35 K83A can be concentrated more by using a buffer that contains two detergents: 0.025% Tween-20 and 0.005% DDM. Combining two or three detergents together can increase the solubility and stability of the protein, increasing the chances of generating higher quality protein crystals. In future experiments, the percentages of Tween-20 and DDM can be adjusted and other detergents could be tested to see if protein concentration can be increased even more. Another approach could be to attempt to co-crystallize the truncated SipS Δ2-35 constructs with a substrate – for example, PONA – which is also known to stabilize protein structure.

4.1.1.   Rescue Methods

Other than the conventional method of crystallization introduced in Chapter 2 - for instance, vapour diffusion - there are several rescue methods that were recently introduced and which could be attempted in future experiments. One method utilizes a large carrier protein fused to the protein of interest. These carrier proteins include maltose binding protein (MBP), glutathione-S-transferase (GST), thioredoxin (TRX) and lysozyme (Smyth et al., 2003). It has been shown that, in some instances, the presence of the large carrier protein increases the target protein's expression and solubility (Moon et al., 2010). In addition, the carrier protein's position at the N-terminus can have a chaperone-like effect, helping the protein fold correctly and as a result, increasing the number of active protein molecules. This method can be used with another rescue strategy called surface entropy reduction (SER) technique. Protein crystallization is an entropy-driven process; the protein crystal is formed during the equilibration reaction between the ordering of protein molecules and the release of water molecules ordinarily bound to the surface of those protein molecules. SER attempts to compensate for the entropic behavior of large hydrophilic side chains on the protein surface. Random motions by these side chains can subtly increase the entropy of the system to the point that forming positive lattice contacts is inhibited (Moon et al., 2010). In the SER method, patches of surface-exposed large, charged residues are mutated to small nonpolar amino acids, reducing surface entropy and allowing formation of crystal contacts featuring backbone amide and carbonyl groups.

4.2. Crystallization of Other *B. subtilis* SPases I

B. subtilis has six paralogous SPases I, other than SipS. Among these, SipT and SipP are known as the major SPases I, and together with SipS, are essential for survival of the bacterium. Because of this shared property, future research should also focus on the determination of their structures. Successfully targeting any of the major SPases I could lead to new, effective therapeutic approaches to treating *B. subtilis* infections.

4.3. Chemical Compounds Inhibiting *B. subtilis* SPases I

Experimental results suggested that Cohex inhibits the catalytic activity of *B. subtilis* SipS but the mechanism is not clear. In order to test if the inhibition is due to the binding between the enzyme and Cohex, but not the substrate and Cohex, one can use Isothermal Titration Chromatography (ITC) and Biacore technology. These are binding assays that can measure the interaction between the chemical compound and the enzyme. Once the binding interaction is confirmed, the next step would be to determine the nature of the inhibition - competitive or non-competitive. This could be done using the same catalytic activity assay, but varying the concentrations of substrate and enzyme. Different catalytic curves could be generated from a series of experiments, and interpreted using the fit-line calculated by the instrument's software. If the V_{init} is only minimally decreased when K_m increased significantly, the software program would consider it as a competitive inhibition. If K_m stays the same while V_{init} decreases significantly, the inhibition is mostly likely a non-competitive inhibition. Another avenue of future research could be to test other chemical compounds with the *B. subtilis* SPase I enzymes to identify potential inhibitors.

4.4. The Third Coordinating Residue in *B. subtilis* SipS

It has been assumed that the Ser/Lys dyad enzymes do not possess residues equivalent to the aspartate of the Ser/His/Asp triad. However, recent studies and structures have shown that *E. coli* SPase I contains a third serine residue in close proximity to the active site, which may play a role equivalent to that of the Asp (Klenotic et al., 2000). Ser278 is only 2.9 Å away from the lysine residue in *B. subtilis* SPase I. The recently solved *B. anthracis* SPase I has been found to have a third serine residue 2.9 Å away from the lysine (Minasov, n.d.). This close proximity indicates there may be a hydrogen-bonding interaction between the two residues, as well as a possible role for the third serine in depressing the pK_a value of the lysine. It was demonstrated that *E. coli* signal peptide peptidase (SPPase), utilizing the same Ser/Lys dyad, contains Ser431 which acts as a coordinating residue with the general base Lys209 (Kim et al., 2008). It was found that the pK_a of the mutant S431A SPPase is higher than that of the wild type S431 SPPase, proposing that the optimal catalytic activity is achieved at a

different pH for the mutant. This may be due to the ability of the Ser431 residue to titrate the Lys209, making it easier for the serine to release its proton at a certain pH.

In future, by carrying out a pH-profiling experiment on *B. subtilis* SipS, the role of the third serine residue (Ser151), along with mutations at position 151, in potentially coordinating with the general base Lys83 may be clarified. The mutant full-length enzyme, SipS FL S151A, could be generated through site-directed mutagenesis. The activity assay could be done using a FRET peptide and a spectrophotometer/plate reader such as the SPectraMax M5. Each catalytic reaction could be carried out at a different pH, and the pK_a value could be determined at the ascending inflection point of pH-rate profile curve. The pK_a values between the SipS FL WT and the SipS FL S151A could be compared, illuminating the role of Ser151 in depressing the pK_a value of Lys83 in *B. subtilis*.

4.5. Conclusion

Finding an optimal crystallization condition for SipS and identifying a buffer that increases the solubility and therefore concentration of the protein are two steps in the experimental process leading to elucidation of its structure using x-ray crystallography techniques. In this thesis, four *B. subtilis* SipS constructs were overexpressed, purified and concentrated. These purified proteins were then crystal plated and crystals were obtained in the presence of lithium sulfate. In addition, preliminary data showed that the catalytic activity of *B. subtilis* SipS was inhibited by hexaminecobalt (III) chloride. Using this molecule as an additive in SipS crystallization may improve the order of the resulting crystals.

References

Allsop, A., Brooks, G., Edwards, P. D., Kaura, A. C., Southgate, R., Beechampharmaceuticals, S., & Park, B. (1996). Inhibitors of Bacterial Signal Peptidase : A Series of 6- (Substituted oxyethyl) penems. *The Journal of Antibiotics*, *49*(9), 921–928.

Auclair, S. M., Bhanu, M. K., & Kendall, D. (2012). Signal peptidase I: cleaving the way to mature proteins. *Protein Science : A Publication of the Protein Society*, *21*(1), 13–25.

Barkocy-Gallagher, G. A., & Bassford, P. J. J. (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*, *267*(2), 1231–1238.

Barnett, J. P., Eijlander, R. T., Kuipers, O. P., & Robinson, C. (2008). A minimal Tat system from a gram-positive organism: a bifunctional TatA subunit participates in discrete TatAC and TatA complexes. *The Journal of Biological Chemistry*, *283*(5), 2534–2542.

Bechtluft, P., Nouwen, N., Tans, S. J., & Driessen, A. J. M. (2010). SecB-A chaperone dedicated to protein translocation. *Molecular BioSystems*, *6*(4), 620–627.

Bilgin, N., Lee, J., Zhu, H., Dalbey, R., & von Heijne, G. (1990). Mapping of catalytically important domains in Escherichia coli leader peptidase. *The EMBO Journal*, *9*(9), 2717–2722.

Black, M. T., & Bruton, G. (1998). Inhibitors of bacterial signal peptidases. *Current Pharmaceutical Design*, *4*(2), 133–154.

Blackman, S. A., Smith, T. J., Foster, S. J., & Court, F. (1998). The role of autolysins during vegetative growth of Bacillus subtilis 168. *Microbiology*, *144*, 73–82.

Blackt, 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*, *175*(16), 4957–4961.

Bolhuis, A., Sorokin, A., Azevedo, V., Ehrlich, S. D., Braun, P. G., de Jong, A., … van Dijl, J. M. (1996). Bacillus subtilis can modulate its capacity and specificity for protein secretion through temporally controlled expression of the sipS gene for signal peptidase I. *Molecular Microbiology*, *22*(4), 605–18.

Braun, V. (1975). Covalent lipoprotein from the outer membrane of Escherichia coli. *Biochimica et Biophysica Acta*, *415*(3), 335–377.

Briggs, M. S., Cornell, D. G., Dluhy, R. A., & Gierasch, L. M. (1986). Conformations of signal peptides induced by lipids suggest initial steps in protein export. *Science (New York, N.Y.)*, *233*(4760), 206–208.

Carlos, J. L., Paetzel, M., Brubaker, G., Karla, A., Ashwell, C. M., Lively, M. O., … Dalbey, R. E. (2000). The role of the membrane-spanning domain of type I signal peptidases in substrate cleavage site selection. *The Journal of Biological Chemistry*, *275*(49), 38813–22.

Chayen, N. E., & Saridakis, E. (2008). Protein crystallization: from purified protein to diffraction-quality crystal. *Nature Methods*, *5*(2), 147–53.

Conn, E., Stumpf, P. K., Bruening, G., & Doi, R. Y. (1987). *Outlines of Biochemistry, 5th Edition* (Vol. 05, p. 1987).

Cregg, K. M., Wilding, E. I., & Black, M. T. (1996). Molecular Cloning and Expression of the spsB Gene Encoding an Essential Type I Signal Peptidase from Staphylococcus aureus. *Journal of Bacteriology*, *178*(19), 5712–5718.

Cristóbal, S., de Gier, J. W., Nielsen, H., & von Heijne, G. (1999). Competition between Sec- and TAT-dependent protein translocation in Escherichia coli. *The EMBO Journal*, *18*(11), 2982–2990.

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*, *6*(6), 1129–38.

Dalbey, R. E., & von Heijne, G. (1992). Signal peptidases in prokaryotes and eukaryotes - a new protease family. *Trends in Biochemical Sciences*, *17*(11), 474–478.

Dalbey, R. E., Wang, P., & van Dijl, J. M. (2012). Membrane proteases in the bacterial protein secretion and quality control pathway. *Microbiology and Molecular Biology Reviews : MMBR*, *76*(2), 311–330.

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*, *260*(29), 15925–15931.

Dao-Pin, S., Anderson, D. E., Baase, W. A., Dahlquist, F. W., & Matthews, B. W. (1991). Structural and thermodynamic consequences of burying a charged residue within the hydrophobic core of T4 lysozyme. *Biochemistry*, *30*(49), 11521–11529.

Date, T., & Wickner, W. (1981a). 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*, *78*(10), 6106–6110.

Date, T., & Wickner, W. (1981b). Leader Peptidase Is Found in Both the Inner and Outer Membranes of Escherichia coli. *The Journal of Biological Chemistry*, *256*(7), 3593–3597.

Driessen, A. J. M., & Nouwen, N. (2008). Protein translocation across the bacterial cytoplasmic membrane. *Annual Review of Biochemistry*, *77*, 643–667.

Duong, F., & Wickner, W. (1997). Distinct catalytic roles of the SecYE, SecG and SecDFyajC subunits of preprotein translocase holoenzyme. *The EMBO Journal*, *16*(10), 2756–2768.

Economou, A., & Wickner, W. (1994). SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. *Cell*, *78*(5), 835–843.

Emr, S. D., & Silhavy, T. J. (1983). Importance of secondary structure in the signal sequence for protein secretion. *Proceedings of the National Academy of Sciences of the United States of America*, *80*(15), 4599–4603.

Fikes, J. D., Barkocy-Gallagher, G. A., Klapper, D. G., & Bassford, P. J. J. (1990). Maturation of Escherichia coli maltose-binding protein by signal peptidase I in vivo. Sequence requirements for efficient processing and demonstration of an alternate cleavage site. *The Journal of Biological Chemistry*, *265*(6), 3417–3423.

Foster, S. J. (1992). Analysis of the autolysins of Bacillus subtilis 168 during vegetative growth and differentiation by using renaturing polyacrylamide gel electrophoresis. *Journal of Bacteriology*, *174*(2), 464–470.

Foster, S. J. (1993). Molecular analysis of three major wall-associated proteins of Bacillus subtilis 168: evidence for processing of the product of a gene encoding a 258 kDa precursor two-domain ligand-binding protein. *Molecular Microbiology*, *8*(2), 299–310.

Gan, L., Chen, S., & Jensen, G. J. (2008). Molecular organization of Gram-negative peptidoglycan. *Proceedings of the National Academy of Sciences*.

Gohlke, U., Pullan, L., McDevitt, C., Porcelli, I., de Leeuw, E., Palmer, T., … Berks, B. C. (2005). The TatA component of the twin-arginine protein transport system forms channel complexes of variable diameter. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(30), 10482–6.

Gould, A. R., May, B. K., & Elliott, W. H. (1975). Release of Extracellular Enzymes from Bacillus amyloliquefaciens. *Journal of Bacteriology*, *122*(1), 34–40.

Hampel, A., Labanauskas, M., Connors, P. G., Kirkegard, L., RajBhandary, U. L., Sigler, P. B., & Bock, R. M. (1968). Single crystals of transfer RNA from formylmethionine and phenylalanine transfer RNA's. *Science (New York, N.Y.)*, *162*(3860), 1384–1387.

Hartl, F. U., Lecker, S., Schiebel, E., Hendrick, J. P., & Wickner, W. (1990). The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the E. coli plasma membrane. *Cell*, *63*(2), 269–279.

Hayashi, S., & Wu, C. (1990). MINI-REVIEW Lipoproteins in Bacteria. *Journal of Bioenergetics and Biomembranes*, *22*(3), 451–471.

Honda, K., Nakamura, K., Nishiguchi, M., & Yamane, K. (1993). Cloning and Characterization of a Bacillus subtilis Gene Encoding a Homolog of the 54-Kilodalton Subunit of Mammalian Signal Recognition Particle and Escherichia coli Ffh. *Journal of Bacteriology, 175*(15), 4885–4894.

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*, *257*(9), 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*, *129*(1), 233–239.

Jongbloed, J. D. H., Grieger, U., Antelmann, H., Hecker, M., Nijland, R., Bron, S., & van Dijl, J. M. (2004). Two minimal Tat translocases in Bacillus. *Molecular Microbiology*, *54*(5), 1319–25.

Kawaguchi, S. I., Mu, È., Linde, D., Kuramitsu, S., Shibata, T., Inoue, Y., & Vassylyev, D. G. (2001). The crystal structure of the ttCsaA protein : an export-related chaperone from Thermus thermophilus. *The EMBO Journal*, *20*(3), 562–569.

Kelley, L. a, & Sternberg, M. J. E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nature Protocols*, *4*(3), 363–71. doi:10.1038/nprot.2009.2

Kim, A. C., Oliver, D. C., & Paetzel, M. (2008). Crystal structure of a bacterial signal Peptide peptidase. *Journal of Molecular Biology*, *376*(2), 352–66.

Kim, Y. T., Muramatsu, T., & Takahashi, K. (1995). Identification of Trp300 as an Important Residue for Escherichia Coli Leader Peptidase Activity. *European Journal of Biochemistry*, *234*(1), 358–362.

Klenotic, P. A., Carlos, J. L., Samuelson, J. C., Schuenemann, T. A., Tschantz, W. R., Paetzel, M., … 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*, *275*(9), 6490–6498.

Koshland, D., & Sauer, R. T. (1982). Diverse Effects of Mutations in the Signal Sequence on the Secretion of β-Lactamase in Salmonella typhimurium. *Cell*, *30*(October), 903–914.

Kulanthaivel, P., Kreuzman, A. J., Strege, M., Belvo, M. D., Smitka, T., Clemens, M., … Peng, S. (2004). Novel lipoglycopeptides as inhibitors of bacterial signal peptidase I. *The Journal of Biological Chemistry*, *279*(35), 36250–8.

Kumamoto, C. A. (1989). Escherichia coli SecB protein associates with exported protein precursors in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(14), 5320–5324.

Kuo, D., Weidner, J., Griffin, P., Shah, S. K., & Jj, W. B. K. (1994). Determination of the Kinetic Parameters of Escherichia coli Leader Peptidase Activity Using a Continuous Assay : The pH Dependence and Time-Dependent Inhibition by β-Lactams Are Consistent with a Novel Serine Protease Mechanism. *Biochemistry*, *33*(27), 8347–8354.

Lammertyn, E. (2004). Molecular and functional characterization of type I signal peptidase from Legionella pneumophila. *Microbiology*, *150*(5), 1475–1483.

Lill, R., Dowhan, W., & Wickner, W. (1990). The ATPase activity of SecA is regulated by acidic phospholipids, SecY, and the leader and mature domains of precursor proteins. *Cell*, *60*(2), 271–280.

Luis, J., Millan, S., Boyd, D., Dalbey, R., & Wickner, W. (1989). Use of phoA Fusions To Study the Topology of the Escherichia coli Inner Membrane Protein Leader Peptidase. *Journal of Bacteriology*, *171*(10), 5536–5541.

Meijer, W. J. J., de Jong, A., Bea, G., Wisman, A., Tjalsma, H., Venema, G., … van Dijl, J. M. (1995). The endogenous Bacillus subtilis (natto) plasmids pTA1015 and pTA1040 contain signal peptidase- encoding genes: identification of a new structural module on cryptic plasmids. *Molecular Microbiology*, *17*(4), 621–631.

Minasov, G. . S. L. . D. I. . W. J. . S. S. . K. K. . A. W. F. . C. for S. G. of I. D. (CSGID). (n.d.). 1.8 Angstrom Crystal Structure of Signal Peptidase I from Bacillus anthracis. *TO BE PUBLISHED*. Retrieved from http://www.rcsb.org/pdb/explore/explore.do?structureId=4NV4

Monteferrante, C. C., Baglieri, J., Robinson, C., & van Dijl, J. M. (2012). TatAc, the third TatA subunit of Bacillus subtilis, can form active twin-arginine translocases with the TatCd and TatCy subunits. *Applied and Environmental Microbiology*, *78*(14), 4999–5001.

Moon, A. F., Mueller, G., Zhong, X., & Pedersen, L. (2010). A synergistic approach to protein crystallization: combination of a fixed-arm carrier with surface entropy reduction. *Protein Science : A Publication of the Protein Society*, *19*(5), 901–13.

Moore, K. E., & Miurae, S. (1987). A Small Hydrophobic Domain Anchors Leader Peptidase to the Cytoplasmic Membrane of Escherichia coli. *The Journal of Biological Chemistry*, *262*(18), 8806–8813.

Nagarajan, V. (1993). *Bacillus subtilis and other gram-positive bacteria: Protein secretion*. (A. Sonenshein, J. A. Hoch, & R. Losick, Eds.) (pp. 713–7726). Washington, D.C.: American Society for Microbiology.

Nahrstedt, H., Wittchen, K., Rachman, M., & Meinhardt, F. (2004). Identification and functional characterization of a type I signal peptidase gene of Bacillus megaterium DSM319. *Applied Microbiology and Biotechnology*, *64*(2), 243–9.

Nemčovičová, I., & Kutá Smatanová, I. K. (2012). *Alternative Protein Crystallization Technique : Cross-Influence Procedure (CIP)*. (E. Borisenko, Ed.)*InTec* (Crystalliz.). InTech.

Paetzel, M. (1998). Crystal structure of a bacterial signal peptidase in complex with a β-lactam inhibitor. *Nature*, *396*(12 November), 186–190.

Paetzel, M. (2014). Structure and mechanism of Escherichia coli type I signal peptidase. *Biochimica et Biophysica Acta*, *1843*(8), 1497–508.

Paetzel, M., Chernaia, M., Strynadka, N., Tschantz, W., Cao, G., Dalbey, R. E., & James, M. N. (1995). Crystallization of a soluble, catalytically active form of Escherichia coli leader peptidase. *Proteins*, *23*(1), 122–125.

Paetzel, M., Dalbey, R. E., & Strynadka, N. C. J. (2000). The structure and mechanism of bacterial type I signal peptidases. A novel antibiotic target. *Pharmacology & Therapeutics*, *87*(1), 27–49.

Paetzel, M., Dalbey, R. E., & Strynadka, N. C. J. (2002). Crystal structure of a bacterial signal peptidase apoenzyme: implications for signal peptide binding and the Ser-Lys dyad mechanism. *The Journal of Biological Chemistry*, *277*(11), 9512–9.

Paetzel, M., Goodall, J. J., Kania, M., Dalbey, R. E., & Page, M. G. P. (2004). Crystallographic and biophysical analysis of a bacterial signal peptidase in complex with a lipopeptide-based inhibitor. *The Journal of Biological Chemistry*, *279*(29), 30781–90.

Paetzel, M., Karla, A., Strynadka, N. C. J., & Dalbey, R. E. (2002). Signal peptidases. *Chemical Reviews*, *102*(12), 4549–4580.

Park, E., & Rapoport, T. A. (2012). Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annual Review of Biophysics*, *41*, 21–40.

Pooley, H. M., Merchante, R., & Karamata, D. (1996). Overall protein content and induced enzyme components of the periplasm of Bacillus subtilis. *Microbial Drug Resistance (Larchmont, N.Y.)*, *2*(1), 9–15.

Privé, G. G. (2007). Detergents for the stabilization and crystallization of membrane proteins. *Methods (San Diego, Calif.)*, *41*(4), 388–97.

Raetz, C. R., & Dowhan, W. (1990). Biosynthesis and function of phospholipids in Escherichia coli. *The Journal of Biological Chemistry*, *265*(3), 1235–1238.

Sankaran, K., & Wu, H. C. (1994). Lipid modification of bacterial prolipoprotein. Transfer of diacylglyceryl moiety from phosphatidylglycerol. *The Journal of Biological Chemistry*, *269*(31), 19701–19706.

Schiebel, E., Driessen, A. J., Hartl, F. U., & Wickner, W. (1991). Delta mu H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. *Cell*, *64*(5), 927–939.

Schimana, J., Gebhardt, K., Holtzel, A., Schmid, D. G., Sussmuth, R., Muller, J., … Fiedler, H. P. (2002). Arylomycines A and B, New Biaryl-bridged Lipopeptide Antibiotics Produced by Streptomyces sp. Tu 6075. *The Journal of Antibiotics*, *55*(6), 565–570.

Sharma, R., Bala, R., Sharma, R., Singh, K., & Ferretti, V. (2006). Second-sphere coordination in anion binding : Synthesis and characterization of hexaamminecobalt (III) salts and X-ray structure determination of [Co (NH 3) 6] Cl (L) 2. *Journal of Molecular Structure*, *784*, 117–123.

Silhavy, T. J., Kahne, D., & Walker, S. (2010). The bacterial cell envelope. *Cold Spring Harbor Perspectives in Biology*, *2*(5), a000414.

Simonenl, M., & Palva, I. (1993). Protein Secretion in Bacillus Species. *Microbiological Reviews*, *57*(1), 109–137.

Smyth, D. R., Mrozkiewicz, M. K., McGrath, W. J., Listwan, P., & Kobe, B. (2003). Crystal structures of fusion proteins with large-affinity tags. *Protein Science : A Publication of the Protein Society*, *12*(7), 1313–22.

Smyth, M. S., & Martin, J. H. J. (2000). Review x Ray crystallography. *Mol Path*, *53*, 8–14.

Stewart, G. C. (2005). Taking shape: control of bacterial cell wall biosynthesis. *Molecular Microbiology*, *57*(5), 1177–1181.

Stover, A. G., & Driks, A. (1999). Secretion, Localization, and Antibacterial Activity of TasA, a Bacillus subtilis Spore-Associated Protein. *Journal of Bacteriology*, *181*(5), 1664–1672.

Sung, M., & Dalbeys, R. E. (1992). Identification of Potential Active-site Residues in the Escherichia coli Leader Peptidase. *Journal of Biological Chemistry*, *267*(19), 13154–13159.

Taheri, T., Salmanian, A., Gholami, E., Doustdari, F., Zahedifard, F., & Rafati, S. (2010). Leishmania major: disruption of signal peptidase type I and its consequences on survival, growth and infectivity. *Experimental Parasitology*, *126*(2), 135–45.

Tjalsma, H., Antelmann, H., Jongbloed, J. D. H., Braun, P. G., Darmon, E., Dorenbos, R., … van Dijl, J. M. (2004). Proteomics of protein secretion by Bacillus subtilis: separating the "secrets" of the secretome. *Microbiology and Molecular Biology Reviews : MMBR*, *68*(2), 207–233.

Tjalsma, H., Bolhuis, A., Jongbloed, J. D., Bron, S., & van Dijl, J. M. (2000). Signal peptide-dependent protein transport in Bacillus subtilis: a genome-based survey of the secretome. *Microbiology and Molecular Biology Reviews : MMBR*, *64*(3), 515–47.

Tjalsma, H., Bolhuis, A., van Roosmalen, M. L., Wiegert, T., Schumann, W., Broekhuizen, C. P., … van Dijl, J. M. (1998). Functional analysis of the secretory precursor processing machinery of Bacillus subtilis: identification of a eubacterial homolog of archaeal and eukaryotic signal peptidases. *Genes & Development*, *12*(15), 2318–31.

Tjalsma, H., Noback, M., Venema, G., Yamane, K., Bron, S., & van Dijl, J. M. (1997). Bacillus subtilis Contains Four Closely Related Type I Signal Peptidases with Overlapping Substrate Specificities. *The Journal of Biological Chemistry*, *272*(41), 25983–25992.

Tjalsma, H., Stover, a G., Driks, A., Venema, G., Bron, S., & van Dijl, J. M. (2000). Conserved serine and histidine residues are critical for activity of the ER-type signal peptidase SipW of Bacillus subtilis. *The Journal of Biological Chemistry*, *275*(33), 25102–8.

Tjalsma, H., van den Dolder, J., Meijer, W. J., Venema, G., Bron, S., & van Dijl, J. M. (1999). The plasmid-encoded signal peptidase SipP can functionally replace the major signal peptidases SipS and SipT of Bacillus subtilis. *The Journal of Bacteriology*, *181*(8), 2448–54.

Tokunaga, M., Tokunaga, H., & Wu, H. C. (1982). Post-translational modification and processing of Escherichia coli prolipoprotein in vitro. *Proceedings of the National Academy of Sciences of the United States of America*, *79*(7), 2255–2259.

Tschantz, W. R., Paetzel, M., Cao, G., Suciu, D., Inouye, M., & Dalbeys, R. E. (1995). Characterization of a Soluble, Catalytically Active Form of Escherichia coli Leader Peptidase: Requirement of Detergent or Phospholipid for Optimal Activity. *Biochemistry*, *34*(12), 3935–3941.

Van Dijl, J. M., de Jong, A., Vehmaanperä, J., Venema, G., & Bron, S. (1992). Signal peptidase I of Bacillus subtilis: patterns of conserved amino acids in prokaryotic and eukaryotic type I signal peptidases. *The EMBO Journal*, *11*(8), 2819–28.

Van Dijl, J. M., de Jong, A., Venema, G., & Bron, S. (1995). Identification of the Potential Active Site of the Signal Peptidase SipS of Bacillus subtilis. *The Journal of Biological Chemistry*, *270*(9), 3611–3618.

Van Roosmalen, M. L., Geukens, N., Jongbloed, J. D. H., Tjalsma, H., Dubois, J. F., Bron, S., … Anné, J. (2004). Type I signal peptidases of Gram-positive bacteria. *Biochimica et Biophysica Acta*, *1694*(1-3), 279–97.

Van Roosmalen, M. L., Jongbloed, J. D., Dubois, J. Y., Venema, G., Bron, S., & van Dijl, J. M. (2001). Distinction between major and minor Bacillus signal peptidases based on phylogenetic and structural criteria. *The Journal of Biological Chemistry*, *276*(27), 25230–5.

Van Voorst, F., & de Kruijff, B. (2000). Role of lipids in the translocation of proteins across membranes. *Biochemical Journal*, *347*, 601–612.

Von Heijne, G. (1990). The Signal Peptide. *The Journal of Membrane Biology, 115*, 195–201.

Watson, M. E. (1984). Compilation of published signal sequences. *Nucleic Acids Research, 12*(13), 5145–5164.

Wickner, W., Moore, K., Dibb, N., Geissert, D., & Rice, M. (1987). Inhibition of Purified Escherichia coli Leader Peptidase by the Leader (Signal) Peptide of Bacteriophage M13 Procoat. *Journal of Bacteriology*, *169*(8), 3821–3822.

Wiech, H., Klappa, P., & Zimmerman, R. (1991). Protein export in prokaryotes and eukaryotes Theme with variations. *FEBS Letters*, *285*(2), 182–188.

Wolfes, P. B., Wickners, W., & Goodman, M. (1983). Sequence of the Leader Peptidase Gene of Escherichia coli and the Orientation of Leader Peptidase in the Bacterial Envelope. *The Journal of Biological Chemistry*, *258*(19), 12073–12080.

Yuan, J., Zweers, J. C., van Dijl, J. M., & Dalbey, R. E. (2010). Protein transport across and into cell membranes in bacteria and archaea. *Cellular and Molecular Life Sciences : CMLS*, *67*(2), 179–99.

Zanen, G., Antelmann, H., Meima, R., Jongbloed, J. D. H., Kolkman, M., Hecker, M., … Quax, W. J. (2006). Proteomic dissection of potential signal recognition particle dependence in protein secretion by Bacillus subtilis. *Proteomics*, *6*(12), 3636–3648.

Zhbanko, M., Zinchenko, V., Gutensohn, M., Schierhorn, A., & Klo, R. B. (2005). Inactivation of a Predicted Leader Peptidase Prevents Photoautotrophic Growth of Synechocystis sp. Strain PCC 6803. *Journal of Bacteriology*, *187*(9), 3071–3078.

Appendix A.

Table of Detergents

Detergent	Type	Agg #	MW mono (micelle)	CMC mM (%w/v)
Triton X-100	Nonionic	140	647g (90kg)	0.24 (0.0155)
Tween 20	Nonionic	-	1228g (-)	0.06 (0.0074)
n-Dodecyl-β-D-Maltoside (DDM)	Nonionic	98	510.6g (50kg)	0.17 (0.009)

Detergent	Chemical Structure
Triton X-100	
Tween-20	
n-Dodecyl-β-D-Maltoside (DDM)	

Appendix B.

Molecular Mass Calculation Using the Standard Curve and Elution Volume

1. Superdex 200 Hiload 16/60

For Superdex 200 HiLoad 16/60 Column, the total column volume is 120mL when the void volume is 45mL. Before the target protein was applied onto the column, standard proteins with known molecular weight were run through this column. This way, one can create a standard curve for calculating a molecular weight of any protein in the future. The partition coefficient, K_{av}, of each standard protein was calculated according to their elution volume, using the following equation:

$$K_{av} = \frac{V_e - V_o}{V_t - V_o}$$ **(Equation 1)**

V_e represents elution volume of each standard protein, V_o is the void volume (45mL) and V_t represents the total column volume (120mL).

The obtained K_{av} value for each of the standard protein was then used to create a standard curve: K_{av} vs Log(MW) (Figure B1). Once the standard curve is created, one could extract the equation from the linear line in the plot and use it for calculating the size of unknown protein purified under the column. The linear line equation was regenerated as following equation to ease the calculation of protein molecular mass.

$$MW = 10^{(-2.8078 \times K_{av} + 3.0314)}$$ **(Equation 2)**

B. subtilis SipS Δ2-35 WT protein, for instance, was calculated to be the monomer (~18kDa) using the standard curve. The protein was eluted in 92.5mL during the SEC purification and therefore, K_{av} for the protein is 0.63416 according to Equation 1. The

Kav was then used to calculate molecular weight using Equation 2, which results 17.8915kDa.
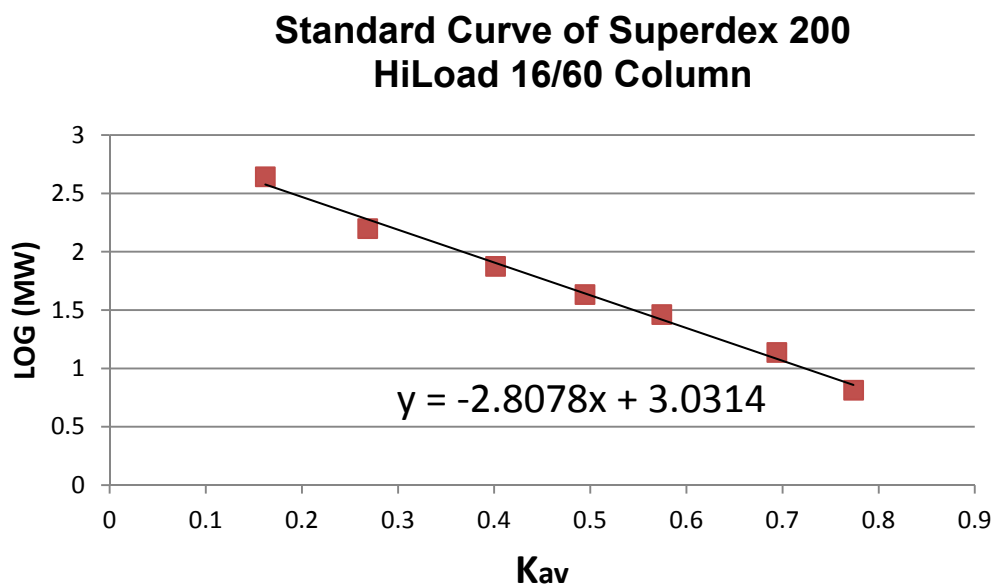
Standard Curve of Superdex 200 HiLoad 16/60 Column

$$y = -2.8078x + 3.0314$$

Figure B1.   Standard Curve of Superdex 200 HiLoad 16/60 Column
The standard curve was generated to estimate the molecular weights of eluted proteins.

Table B1. List of Standard Proteins

Standard protein	MW (kDa)	Ve (mL)	Kav
Ferritin	440	57	0.1619
Aldolase	158	65	0.26833
Conalbumin	75	75	0.40136
Ovalbumin	43	82	0.49448
Carbonic Anhydrase	29	88	0.5743
Ribonuclease A	13.7	97	0.69403
Aprotinin	6.5	103	0.77385

2. Superdex 75 HR 10/30

The total column volume for Superdex 75 HR 10/30 column is 23.6mL when the void volume is 7mL. Before the target protein was applied onto the column, standard proteins with known molecular weight were run through this column. This way, one can create a standard curve for calculating a molecular weight of any protein in the future. The partition coefficient, K_{av}, of each standard protein was calculated according to their elution volume, using the following equation:

$$K_{av} = \frac{V_e - V_o}{V_t - V_o}$$ **(Equation 1)**

V_e represents elution volume of each standard protein, V_o is the void volume (7mL) and V_t represents the total column volume (23.6mL).

The obtained K_{av} value for each of the standard protein was then used to create a standard curve: K_{av} vs Log(MW) (Figure B2). Once the standard curve is created, one could extract the equation from the linear line in the plot and use it for calculating the size of unknown protein purified under the column. The linear line equation was regenerated as following equation to ease the calculation of protein molecular mass.

$$MW = 10^{(-4.778 \times K_{av} + 2.8053)}$$ **(Equation 3)**

B. subtilis SipS Δ2-35 WT protein, for instance, was calculated to be the monomer (~18kDa) using the standard curve. The protein was eluted in 12.5mL during the SEC purification and therefore, K_{av} for the protein is 0.327273 according to Equation 1. The K_{av} was then used to calculate molecular weight using Equation 3, which results 18.24kDa.

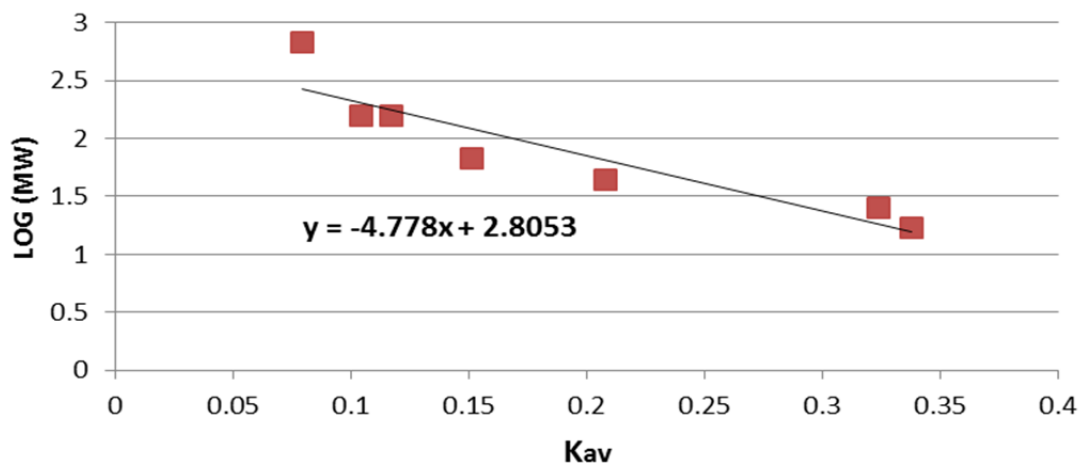## Standard Curve of Superdex 75 HR 10/30 Column

Figure B2.   Standard Curve of Superdex 75 HR 10/30 Column
The standard curve was generated to estimate the molecular weights of eluted proteins.

Table B2. List of Standard Proteins

Standard protein	MW (kDa)	Ve (mL)	Kav
Thyroglobulin	670	8.41	0.07939
g-Globulin	158	8.82	0.10424
Aldolase	158	9.03	0.11697
Albumin	67	9.6	0.15152
Ovalbumin	44	10.53	0.20788
Chymotrypsinogen A	25	12.44	0.32364
Myoblobulin	17	12.68	0.33818
Ribonuclease A	13.7	13.55	0.39091
Vitamin B-12	1.35	18.35	0.68182