REVIEW

Signal peptidase I: Cleaving the way to mature proteins

Sarah M. Auclair, Meera K. Bhanu, and Debra A. Kendall*

Department of Pharmaceutical Sciences, The University of Connecticut, Storrs, Connecticut 06269

Received 9 September 2011; Revised 12 October 2011; Accepted 17 October 2011
DOI: 10.1002/pro.757
Published online 26 October 2011 proteinscience.org

Abstract: Signal peptidase I (SPase I) is critical for the release of translocated preproteins from the membrane as they are transported from a cytoplasmic site of synthesis to extracytoplasmic locations. These proteins are synthesized with an amino-terminal extension, the signal sequence, which directs the preprotein to the Sec- or Tat-translocation pathway. Recent evidence indicates that the SPase I cleaves preproteins as they emerge from either pathway, though the steps involved are unclear. Now that the structure of many translocation pathway components has been elucidated, it is critical to determine how these components work in concert to support protein translocation and cleavage. Molecular modeling and NMR studies have provided insight on how the preprotein docks on SPase I in preparation for cleavage. This is a key area for future work since SPase I enzymes in a variety of species have now been identified and the inhibition of these enzymes by antibiotics is being pursued. The eubacterial SPase I is essential for cell viability and belongs to a unique group of serine endoproteases which utilize a Ser-Lys catalytic dyad instead of the prototypical Ser-His-Asp triad used by eukaryotes. As such, SPase I is a desirable antimicrobial target. Advances in our understanding of how the preprotein interfaces with SPase I during the final stages of translocation will facilitate future development of inhibitors that display a high efficacy against SPase I function.

Keywords: signal peptidase; protein transport; signal peptide; serine protease; antibacterial target; leader peptidase

Introduction

Nearly a third of all proteins function outside of the cytosol and, therefore, need to be translocated through or into the cytoplasmic membrane. To accomplish this feat, they are synthesized as preproteins, consisting of a short amino-terminal extension sequence, called the signal peptide region, followed by the mature region of the protein (see Fig. 1). The signal peptide acts as zip-code marking it as a protein destined to reside in an extracytoplasmic location and directing it to a specific secretion pathway. Once the majority of the preprotein is translocated, the signal peptidase (SPase) enzyme is responsible for cleavage of the signal peptide from the preprotein, allowing release from the membrane and correct folding of the mature protein. Accumulation of preproteins at the membrane has been shown to be deleterious for growth.[2] Therefore, the SPase enzymes are vital for cell survival.[3–9]

Abbreviations: SPase, signal peptidase; ER, endoplasmic reticulum; Sec pathway, general secretory pathway; Tat pathway, twin-arginine translocation pathway; SRP, signal recognition particle.

Grant sponsor: National Institutes of Health; Grant number: GM037639.

Correspondence to: Debra A. Kendall, Department of Pharmaceutical Sciences, The University of Connecticut, 69 North Eagleville Road, Storrs, CT 06269-3092.
E-mail: debra.kendall@uconn.edu.

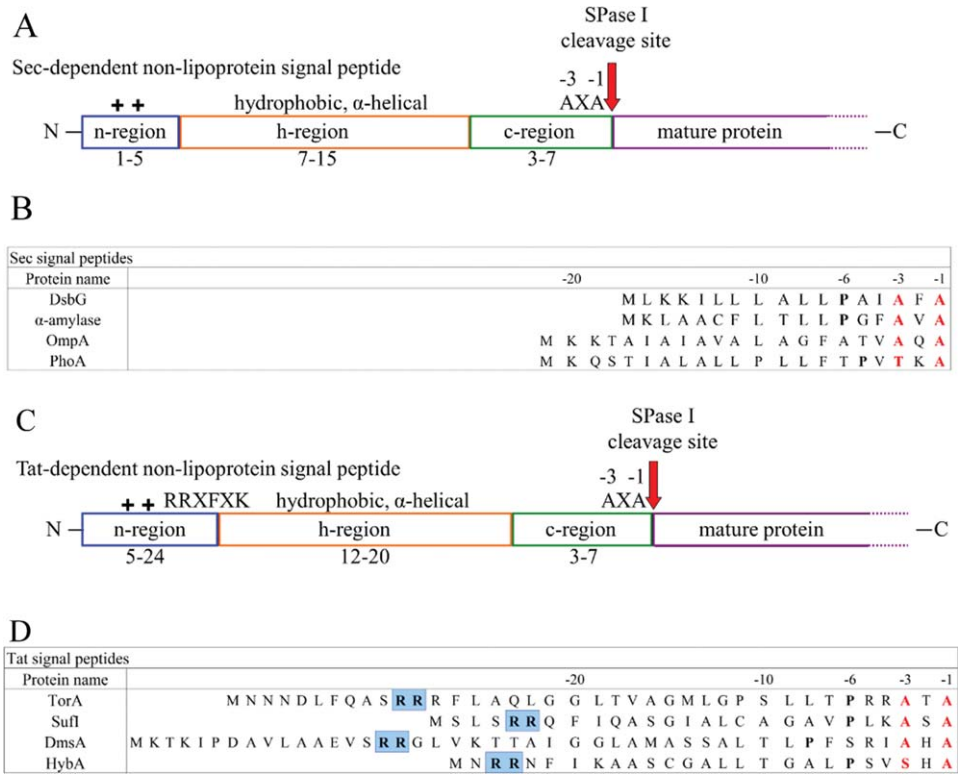

Figure 1. The features and alignment of bacterial signal peptides. The tripartite structure of Sec-dependent non-lipoprotein signal peptides is depicted in (A), where the N-terminus is characterized by the presence of positively-charged residues (blue), the core of the peptide is comprised of hydrophobic residues (orange), and the C-terminus is typically neutral, but polar and contains the cleavage site (green). The red arrow indicates the SPase cleavage site and the amino acid motif common to the cleavage site is given. The mature region (purple) of the preprotein follows the cleavage site. (B) Sec-dependent signal peptide sequence alignment. Signal peptide sequences compiled from the SPdb database.[1] The -1,-3 residues (red) display sequence conservation of small aliphatic residues among a number of Sec- and Tat-dependent signal peptides. If present, the conserved proline is in bold. The site of cleavage is carboxy-terminal to the -1 residue. (C) The Tat-dependent non-lipoprotein signal peptide depicted in the same manner as panel (A). (D) Tat-dependent signal peptide sequence alignment. Depicted as in panel (B) and the residues highlighted in blue are the required arginine residues that give the Tat pathway its name.

SPases have been identified in all orders of life. In eukaryotes, SPase systems are located in the endoplasmic reticulum (ER), the mitochondria, and chloroplasts. In prokaryotes, SPases are classified into three groups: SPase I, II, and IV. SPase II and IV are required for cleaving signal peptides from lipoproteins and prepilin proteins, respectively. This review focuses on SPase I which is responsible for generating mature non-lipoproteins, transported in bacteria by the general secretion (Sec) pathway (see Ref. 10). Recent evidence also indicates transport via the twin-arginine translocation (Tat) pathway.[11] Some bacterial species possess only one essential SPase I enzyme (*Escherichia coli*), while others contain multiple enzymes. Take for example, *Bacillus subtilis*, which has seven closely related SPase I enzymes.[12–14]

SPases have long been recognized as potential antibacterial targets. Recently, there has been a renewed interest in SPase I inhibitors and a number of new studies have identified SPase I in common pathogens[15–17] and tested their inhibition against various antibiotics.[18–25] The bacterial SPase I is part of the SF serine protease clan, for which there is an evolutionary relationship,[26] and belongs to the protease family S26 that utilizes a Ser-Lys catalytic dyad mechanism.[27] It differs from its eukaryotic ER counterparts which belong to the S27 serine protease family and generally use a Ser-His-Asp triad catalytic mechanism. As SPase I exhibits a catalytic mechanism unlike most serine proteases, it is feasible to inhibit the bacterial enzyme without harming the host. At the same time, common serine protease inhibitors are ineffective[28–32] making the development of appropriate inhibitors for this novel target all the more pressing in the face of multidrug resistance.

Protein Secretion Pathways Associated with SPase I

The two most widely used secretion pathways in bacteria, Sec and Tat, are thought to use SPase I at the end stage of translocation. The SPase I enzyme is an essential part of these protein transport pathways. SPase I cleaves non-lipoprotein preproteins

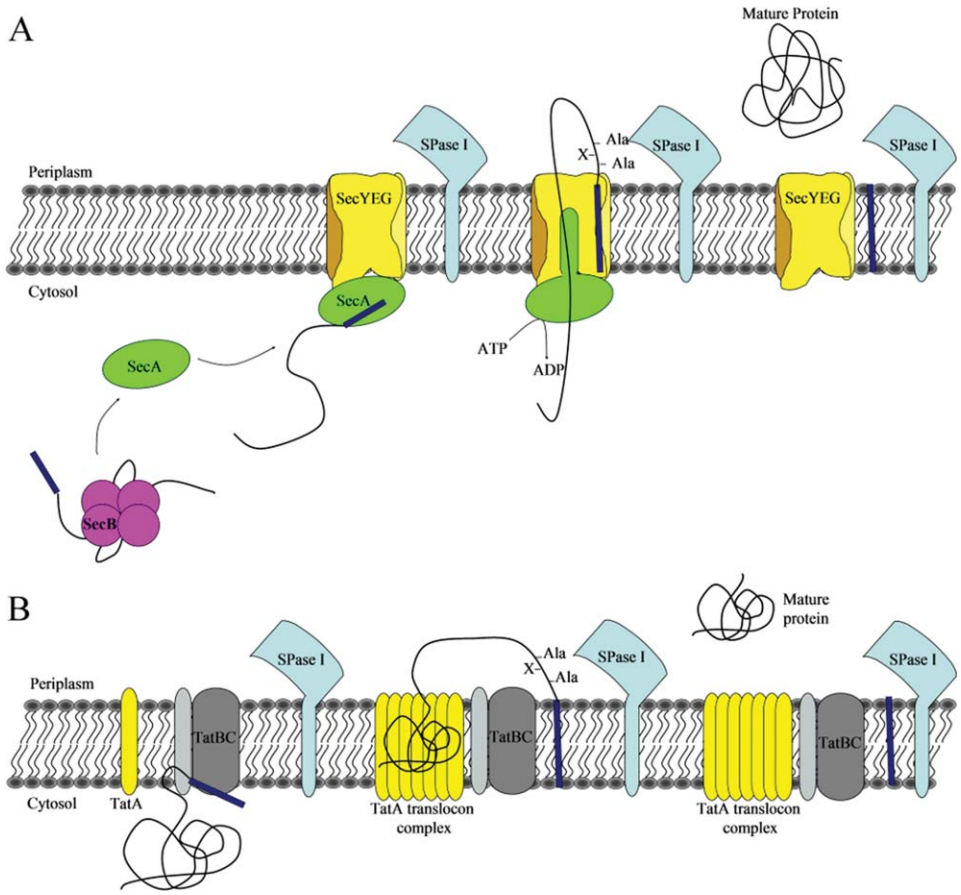

Figure 2. Bacterial protein translocation pathways requiring SPase I. (A) The Sec-dependent general secretory pathway for secretory preproteins via post-translational translocation. Preproteins are bound in the cytosol by SecB or SecA. If bound by SecB, the SecB-preprotein complex then binds to SecA and transfers the preprotein. SecA carrying the preprotein binds the SecYEG translocon channel and using the energy from ATP hydrolysis may propel the preprotein through the channel. Once sufficient preprotein has been translocated to ensure no back-slippage, the SPase I cleaves off the signal peptide, allowing the mature protein to release from the membrane and undergo folding. (B) The Tat-pathway is a post-translational translocation pathway used for secretion of fully folded preproteins. The TatBC complex recognizes and binds the signal sequence of a Tat-dependent preprotein. This causes the recruitment of TatA and the formation of an appropriately sized TatA translocon. The fully-folded preprotein is then secreted into the periplasm, while the signal sequence remains in the membrane. The mature protein is released into the periplasm once the SPase I cleaves the signal peptide.

that are translocated both post-translationally and cotranslationally by the Sec pathway and from the proteins that are translocated through the Tat pathway.[11]

In bacteria, proteins secreted by the Sec pathway typically do so via post-translational translocation in which the protein traverses the inner membrane after it is fully synthesized by the ribosome [see Fig. 2(A)]. The major components of this pathway are an integral membrane translocon complex, SecYEG, and the mobile cytosolic components, the SecA ATPase nanomotor and the chaperone SecB. Preproteins bind directly to SecA in the cytosol of Gram-positive bacteria, which lack a SecB homolog, and in Gram-negative bacteria when a preprotein does not require SecB for secretion. SecB delivers the preprotein to SecA[33–35] and then the SecA-preprotein complex binds to the SecYEG translocon. SecA undergoes repetitive cycles of ATP hydrolysis that provide the energy required for translocation of the preprotein through the SecYEG translocon channel. After at least 80% of the preprotein has been synthesized, SPase I can cleave the preprotein to give rise to the mature protein.[36] The signal peptide is then degraded by the signal peptide peptidase, a membrane-bound enzyme with an active site located in the periplasm that cleaves using a Ser-Lys catalytic dyad mechanism.[37,38]

Co-translational translocation, concomitant synthesis and translocation of a preprotein, occurs similarly with a few exceptions. Typically, this type of translocation is used with integral membrane proteins, however, not all membrane proteins are synthesized with a signal peptide[39] and, therefore, do not require SPase I cleavage. (For a review on co-translational translocation refer to Dalbey and Chen.)[39]

The Tat pathway is a protein transport system with the ability to export proteins in a fully folded

conformation [see Fig. 2(B)] such as some redox-cofactor containing enzymes.[40] The *E. coli* Tat pathway is comprised of three membrane proteins, TatA, TatB, and TatC,[41–44] while the *B. subtilis* pathway lacks the TatB protein.[45,46] Tat-dependent secretory proteins are targeted to the pathway via an N-terminal signal peptide containing an almost invariant twin-arginine motif [Fig. 1(C,D)][47] and are recognized by a complex of TatB and TatC in the membrane.[48,49] Once a signal sequence binds TatBC, TatA is recruited and forms a large protein-conducting channel ranging from 450 to 750 kDa in size.[50–52] The signal peptide region is thought to remain in the membrane where it is cleaved from the preprotein by SPase I, thereby allowing release of the mature protein into the periplasm. However, the instability of the Tat channel has made it difficult to study this pathway and much information is still needed to fully understand how folded enzymes are translocated across the inner membrane and where and how cleavage of the preprotein occurs.

Although the early steps of these pathways have been well studied, our understanding of the preprotein-SPase interaction and cleavage is still in its infancy and a number of questions remain unanswered. What are the steps linking preprotein utilization of the Sec and Tat pathways to cleavage of the signal peptide from preprotein? How does the enzyme find the signal peptide? Does SPase I interact with the Sec or Tat translocon? It is also unknown where the signal peptide is located during cleavage; for example, is it in the periplasm or in the lipid bilayer? If the signal peptide exits the membrane for cleavage, how does this occur and what triggers the cleavage event?

Substrate Specificity

A typical Gram-negative bacterial Sec signal peptide contains 18–30 amino acid residues, while Gram-positive bacteria tend to have significantly longer Sec signal peptides.[53–56] Although signal peptides show no conservation of sequences, they can be divided into three distinct domains. A short N-terminal (n) region containing up to three positively charged residues is followed by a much longer, central, hydrophobic (h) core region after which is a C-terminal (c) polar region containing the consensus cleavage sequence [Fig. 1(A,C)].

The c-region contains crucial residues required for cleavage. Statistical analysis of the amino acid residues in the SPase I cleavage site led to the formulation of the (-1,-3) rule; that is, the residues at the -1 and -3 (also known as P1 and P3) positions, where -1 is immediately before the processing site, are typically made up of small, neutral residues, such as Ala, Gly, Cys, and Ser (Fig. 1).[57–59] The prevalence of Ala in the -1, -3 positions gave rise to the name of the motif, Ala-X-Ala (Fig. 1). Mutational studies carried out by substituting the -1 residue with residues other than those that are preferred at this site such as Asp, Val, and Asn indicated that processing no longer occurred C-terminal to the -1 residue. However, cleavage occurred two residues upstream (at position -3) of the usual site due to the presence of an alternative processing site containing another Ala-X-Ala motif.[60,61] The length of the c-region is also thought to determine cleavage efficiency. When the c-region is longer than nine residues cleavage efficiency decreases and no cleavage is observed when the c-region is 13 residues long.[62] This is likely due to the -1 and -3 residues being too far away from the active site of the enzyme. Another commonly observed feature of signal peptides is the presence of a Pro or Gly residue at the -6 position (see Fig. 1).[63] This small, helix-breaking residue, located between the h- and c-regions, has been suggested to nucleate a short β-conformation at the -5 to -1 region allowing this region to bind to the SPase I active site (see review[64]).[65] It has also been suggested that the peptide binds in a extended conformation with a pronounced backbone twist between -3 and +1.[66] Having a small residue at -6 that halts the extension of the helix may be more important than the potential turn-forming properties. Consistent with this, a recent NMR structure of an alkaline phosphatase peptide bound to SPase I Δ2-75 shows the cleavage region of the signal peptide sequence is unstructured and exposed for cleavage (De Bona *et al.*, personal communication). This unstructured region is also observed when the signal peptide is bound to dodecylphosphocholine micelles (De Bona *et al.*, personal communication). A small helix in the core region is also apparent in the micelle structure. This suggests that the membrane is important in stabilization of the signal peptide and is consistent with the idea that the signal peptide is in contact with lipid and that the enzyme active site functions at the membrane surface. More structural analysis is needed to better define the conformation of the preprotein when bound to SPase I.

The early mature region of a preprotein typically has a net neutral or negative charge, which has been shown to be important in protein secretion.[67,68] Mutants that have positively-charged residues inserted immediately after the signal peptide are not translocated through the membrane.[69–71] Modeling a DsbA peptide (residues 13–25) with the *E. coli* SPase I crystal structure, reveals 13 subsites within the SPase I where the signal peptide and mature portions of the preprotein bind during cleavage.[66] As +1 to +2 in Gram-negative bacteria and +1 to +4 in Gram-positive bacteria display sequence conservation, it is speculated that those regions are involved in binding to the SPase I.[68]

Tat signal peptides tend to be longer (up to 58 amino acids) than Sec signal peptides and they invariably contain two Arg residues in the n-region,

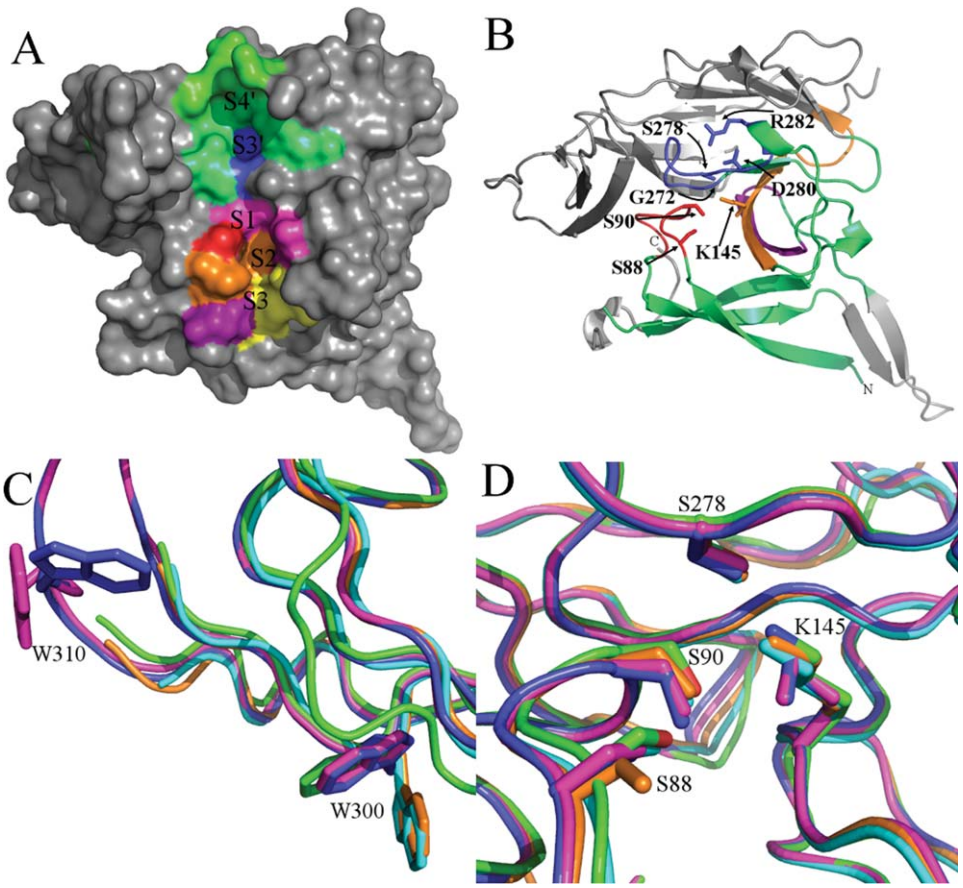

Figure 3. *E. coli* SPase I Δ2-75 apoenzyme crystal structure (PDB ID: 1KN9).[73] (A) A solid surface representation of the SPase I Δ2-75 apoenzyme structure with the modeled signal peptide binding subsites labeled.[66,73] Subsites are colored as follows: green S4', lime green S4' and S3' overlap, blue S3', magenta S1, red S1 and S2 overlap, orange S1, S2, and S3 overlap, purple S2 and S3 overlap, and yellow S3. (B) A ribbon representation of SPase I Δ2-75. The colored portion of the protein represents the conserved domain I, while the gray region is the nonconserved domain II as well as the Gram-negative β-ribbon insertion. Box domains B-D are color coordinated. Box B is shown in red, box C is shown in purple, box D is colored orange, and box E is in blue. The residues that are important for catalysis (S88, S90, K145, and S278) are labeled, colored by box domain, and displayed in ball and stick representation. The placement of G272 and K145 are shown to emphasize the importance of glycine at residue 272. Any other amino acid in that location would result in steric hindrance with K145. D280, and R282 form a salt bridge that is important for structural stability of the enzyme. The N- and C-termini are labeled and the structure is lacking residues 107-124, 136, 176-177, 200-202, and 305-312, which were not resolved in the 3D structure. (C) Superimposition of the SPase I Δ2-75 apoenzyme crystal structure with inhibitor-bound crystal structures to highlight differences in the position of the W300 and W310 side-chains in the various structures. The structures are colored as follows: cyan is the apoenzyme 1KN9 molecule A,[73] green is the apoenzyme 1KN9 molecule D,[73] magenta is the lipopetide-bound structure 1T7D molecule B,[74] orange is the β-lactam-bound structure 1B12 molecule A,[65] and blue is the arylomycin A_2 and β-sultam bound structure 3IIQ molecule A.[75] (D) As in (C) except displaying positioning differences of the side chains for residues S88 and K145 within the active site.

adjacent to the h-region [see Fig. 1(C,D)].[40] They may also carry an Arg or Lys residue in the c-region, which ensures that the preprotein is not translocated through the Sec pathway.[72] The Tat peptides, however, maintain the -1,-3 rule with small aliphatic residues found at the -1 and -3 sites thought to be critical for cleavage by SPase I.[11] Does this indicate that the -1 and -3 residues are all that is important for recognition by SPase I? Or are the modes of recognition of Sec and Tat signal peptides by SPase I different?

The SPase I substrate binding pockets dictate the substrate specificity of the enzyme. Two major substrate-binding subsites in SPase I have been identified and are named S1 and S3 after the -1 and -3 regions of the signal peptide [see Fig. 3(A)].[65] Recently, an additional 11 subsites have been identified by a computational modeling strategy using a peptide containing both signal peptide and mature protein sequences (Table I).[66] These sites named S7, S6, S5, S4, S3, S2, S1, S1', S2', S3', S4', S5', and S6' suggest that the peptide residues -7 to +6 are bound in their respective pockets and that the early mature region of the preprotein is indeed involved in cleavage. Of these sites, six have been identified as having smaller clefts (S3, S2, S1, S1', S3', and S4') and thus play a direct role in the high specificity of the

Table I. *SPase I Residues that Comprise Each Preprotein Binding Subsite*

Subsite	Residues
S7	Glu 82, Pro 83
S6	Pro 83, Phe 84
S5	Phe 84, Gln 85, Asp 142
S4	Phe 84, Gln 85, Pro 87, Asp 142
S3	Phe 84, Gln 85, Ile 86, Pro 87, Ile 101, Val 132, Asp 142, Ile 144
S2	Gln 85, Ile 86, Pro 87, Ser 88, Met 91, Ile 144
S1	Ile 86, Pro 87, Ser 88, Ser 90, Met 91, Leu 95, Tyr 143, Ile 144, Lys 145
S1′	Ser 88, Ser 90, Tyr 143, Ala 279
S2′	Ser 88, Ser 90, Phe 208, Asn 277, Ala 279
S3′	Tyr 50, Met 249, Asp 276, Asn 277, Ala 279, Arg 282, Tyr 283
S4′	Gln 244, Asp 245, Asp 276, Asn 277, Arg 282
S5′	Phe 196, Ser 206, Ala 243, Asp 276, Asn 277
S6′	Phe 196, Ile 242, Ala 243

signal peptide and mature residues at those respective positions.[66]

Enzyme Architecture and Structure

E. coli SPase I is a membrane-embedded polytopic serine endopeptidase. It has a short periplasmic amino-terminal region (residues 1–3), followed by a transmembrane region (residues 4–28), a cytosolic loop (residues 29–58), another transmembrane segment (residues 59–76), and a large C-terminal periplasmic domain carrying the active site (residues 77–323). The two transmembrane helices are packed in a left-handed supercoil.[76] Some variations to this topology exist. For example, in typical Gram-positive bacteria there is only a single transmembrane segment at the N-terminus and a few enzymes are believed to also have a C-terminal transmembrane segment.[13] The transmembrane segments are thought to be required for anchoring the enzyme to the membrane, so that it can position the catalytic domain appropriately for signal peptide cleavage as the preprotein emerges from the translocon. However, the transmembrane domain does not appear to be involved in the recognition of the correct Ala-X-Ala cleavage site.[77]

The periplasmic region (residues 77–323) of SPase I is divided into two antiparallel β-sheet domains. One of the domains (domain I) carries the conserved regions that are found in both Gram-positive and Gram-negative bacteria, while the other domain (domain II) is not conserved [Fig. 3(B)].[78] The residues crucial for catalysis by the *E. coli* enzyme include the nucleophile Ser 90[65,79] and the general base Lys 145.[65,80] Both of these residues reside in the conserved domain [Figs. 3(B) and 4]. Together Ser 90 and Lys 145 form the catalytic dyad. The conserved domain also contains a large exposed hydrophobic region that includes the substrate binding site and the catalytic center.[65] Since a SPase I mutant lacking the transmembrane domains exhibits enhanced *in vitro* cleavage activity in the presence of both phospholipids and the detergent Triton X-100,[81] it is thought that the catalytic domain interacts with the membrane, potentially facilitating signal peptide processing.[65] Additionally, when Trp 300 and Trp 310 are mutated, catalytic activity is reduced [Fig. 3(C)].[82] It is unlikely, given the distance to the active site (>20 Å),[65] that these residues are involved in catalysis; however, they may play an important role in stabilizing the enzyme's structure and its placement at the membrane. In membrane-proteins, aromatic residues are often located at the membrane protein interface.[83] It has been hypothesized that the transmembrane domain residues Trp 20 and Trp 59 are located at the cytoplasmic side of the membrane and Tyr 81 on the periplasmic side of the membrane and all three are thought to be positioned next to the phospholipid headgroups in the membrane.[76]

The nonconserved domain varies in size in different organisms. Gram-negative bacteria tend to have a larger domain than Gram-positive bacteria (see Fig. 4). The physiological role of domain II is unknown; however, evidence suggests that it is not directly involved in catalysis but may play a role in structural stability of the enzyme. For example, *E. coli* SPase I has a disulfide bond between Cys 170 and Cys 176 in the nonconserved domain,[65] which when mutated does not disrupt SPase I catalytic activity.[76,79,84]

Although the overall sequence identity of SPase I from various species is relatively low (see Fig. 4), there are five conserved regions in the catalytic domain that have been identified and are called boxes A, B, C, D, and E.[78] The location of these box domains in the crystal structure is shown in Figure 3(B) and an alignment of the box domains is shown in Figure 4. The transmembrane segments are referred to as box A, while boxes B–E are all located within the periplasmic domain and make up the conserved catalytic domain. Box B consists of residues 88–95 in *E. coli* and is likely positioned near the membrane surface on the periplasmic side. Not surprisingly, the conserved catalytic nucleophile Ser 90 is found in this domain and is positioned on a loop in between two β-strands.[65] In the crystal structure with the bound 5*S*,6*S*-β-lactam (penem) inhibitor, Ser 90 is covalently bonded to the inhibitor.[65] Another conserved residue in this region is Ser 88, which is involved in the formation of the SPase I oxyanion hole and stabilization of the tetrahedral oxyanion intermediate state.[73] Box C consists of residues 127–134 and contains a conserved Gly and Asp. Box D consists of residues 142–153 and contains the general base Lys 145 which is highly conserved in bacterial and mitochondrial enzymes, but is replaced with a conserved His in ER and archaeal

Gram-negative bacteria

Species	Enzyme name	MW	% Sequence identity	Box B									Box C												Box D												Box E										
				88							*	95	127								134	142							*	153	272														282
Escherichia coli	LepB	35,960	100.0	S	D	S	M	M	P	T	L		R	G	D	I	V	V	F	K		D	Y	I	K	R	A	V	G	L	P	G	D		G	D	N	R	D	N	S	A	D	S	R
Haemophilus influenzae	Lep	39,734	37.9	S	G	S	M	E	S	T	L		R	G	D	V	I	V	F	K		D	Y	I	K	R	I	V	G	K	G	G	D		G	D	N	R	D	N	S	S	D	S	R
Helicobacter pylori	Lep	33,710	26.5	S	R	S	M	V	G	T	L		R	G	E	V	V	V	F	I		Y	Y	V	K	R	N	F	A	I	G	G	D		G	D	N	R	D	N	S	S	D	S	R
Phormidium laminosum	Lep	22,485	24.9	S	E	S	M	L	P	T	L		R	G	D	I	I	V	F	H		A	F	I	K	R	V	I	G	L	P	G	E		G	D	N	R	N	N	S	Y	D	S	H
Pseudomonas fluorescens	Lep	31,903	41.8	S	G	S	M	K	P	T	L		R	G	D	V	M	V	F	R		N	Y	I	K	R	V	V	G	L	P	G	D		G	D	N	R	D	N	S	N	D	S	R
Rhodobacter capsulatus	Lep	28,878	28.0	S	G	S	M	K	D	T	L		R	G	D	V	V	V	F	R		D	F	I	K	R	L	I	G	L	P	G	D		G	D	N	R	D	N	S	E	D	S	R
Salmonella typhimurium	Lep	35,778	93.5	S	G	S	M	M	P	T	L		R	G	D	I	V	V	F	K		D	Y	I	K	R	A	V	G	L	P	G	D		G	D	N	R	D	N	S	A	D	S	R
Legionella pneumophila	LepB	28,782	38.3	S	G	S	L	E	P	T	L		T	G	E	I	A	V	F	R		D	Y	I	K	R	V	I	G	V	P	G	D		G	D	N	R	D	D	S	A	D	S	R
Bradyrhizobium japonicum	SipF	28,845	30.2	S	G	S	M	K	A	T	L		R	G	D	I	V	V	F	R		D	Y	I	K	R	V	I	G	L	P	G	D		G	D	N	R	D	N	S	T	D	S	R
Bradyrhizobium japonicum	SipS	27,931	29.9	S	G	S	M	E	P	T	L		Q	G	D	V	V	V	F	R		A	W	V	K	R	V	V	G	L	P	G	D		G	D	N	R	D	N	S	A	D	S	R

Gram-positive bacteria

Species	Enzyme name	MW	% Sequence identity	Box B									Box C												Box D												Box E										
				41							*	48	68								75	80							*	91	145														155
Bacillus subtilis	SipS	21,047	20.0	G	D	S	M	Y	P	T	L		R	G	D	I	V	V	L	N		H	Y	V	K	R	I	I	G	L	P	G	D		G	D	N	R	R	N	S	M	D	S	R
Bacillus subtilis	SipT	21,854	21.5	G	S	S	M	Y	P	T	L		R	G	D	I	V	I	I	N		H	Y	V	K	R	L	I	G	K	P	G	E		G	D	N	R	L	N	S	M	D	S	R
Bacillus subtilis	SipU	21,183	21.5	G	S	S	M	A	P	T	L		R	G	D	I	I	V	I	H		S	F	V	K	R	L	I	G	L	P	G	D		G	D	N	R	L	N	S	L	D	S	R
Bacillus subtilis	SipV	18,956	22.2	G	V	S	M	N	P	T	F		R	F	D	I	V	L	F	K		V	L	I	K	R	V	I	G	L	P	G	E		G	D	N	R	I	Y	S	F	D	S	R
Bacillus subtilis	SipW	20,678	14.4	S	G	S	M	E	P	E	F		K	G	D	V	I	T	F	M		A	V	T	H	R	I	V	D	I	T	K	Q		G	D	P	G	V	M	V	Y	A	F	V
Bacillus subtilis	SipP (pTA1015)	21,252	21.5	G	K	S	M	D	P	T	L		R	G	D	I	V	V	L	N		H	Y	V	K	R	L	I	G	L	P	G	D		G	D	N	R	Q	E	S	M	D	S	R
Bacillus subtilis	SipP (pTA1040)	21,568	22.5	G	E	S	M	K	P	T	L		R	G	D	I	V	V	L	N		H	Y	V	K	R	L	I	G	L	P	G	D		G	D	N	R	Q	N	S	M	D	S	R
Staphylococcus aureus	SpsA	20,146	17.3	N	N	D	M	S	P	T	L		N	G	D	I	I	T	Y	R		I	Y	T	S	R	I	I	A	K	P	G	Q		N	D	H	D	N	N	Q	H	D	S	R
Staphylococcus aureus	SpsB	21,692	23.0	G	E	S	M	D	P	T	L		K	G	N	V	V	V	F	H		D	Y	V	K	R	V	I	G	V	P	G	D		G	D	N	R	E	V	S	K	D	S	R
Mycobacterium tuberculosis	Sip	31,880	20.6	S	E	S	M	E	P	T	L		P	G	D	V	I	V	F	R		D	L	V	K	R	V	I	A	V	G	G	Q		G	D	B	R	T	H	S	A	D	S	R

Endoplasmic reticulum

Species	Enzyme name	MW	% Sequence identity	Box B									Box C												Box D												Box E										
				42							*	49	66								73	80							*	91	102														112
Saccharomyces cerevisiae	Sec11	18,762	15.7	S	G	S	M	E	P	A	F		V	G	D	V	V	V	Y	E		P	I	V	H	R	V	L	R	Q	H	N	N		G	D	N	N	A	G	N	D	I	S	L
Homo sapiens sapiens	Spc18	20,625	17.6	S	G	S	M	E	P	A	F		V	G	E	I	V	V	F	R		P	I	V	H	R	V	L	K	I	H	E	K		G	D	N	N	A	V	D	D	R	G	L

Archaea

Species	Enzyme name	MW	% Sequence identity	Box B									Box C												Box D												Box E										
				16							*	23	45								52	117							*	128	138														148
Methanococcus jannaschii	Sip	23,724	15.5	S	D	S	M	Y	P	I	M		V	G	D	I	V	V	Y	K		P	V	I	H	R	V	I	D	K	V	E	F		G	D	N	N	P	I	H	D	P	E	L

Figure 4. Alignment of the conserved box domains for signal peptidase I from different species. The molecular weight (M_w) is given for each protein in g/mol. The % sequence identity is relative to the entire *E. coli* SPase I protein sequence. Box domains B–E are located in the catalytic domain. The serine nucleophile and lysine/histidine general base residues that are involved in signal peptide cleavage are indicated with an asterisk. Box A comprises the transmembrane segments and is not shown.

enzymes (Fig. 4).[78,85] Boxes C and D contain two antiparallel β-strands that may form hydrogen bonds with the signal peptide.[85] Box E consists of residues 272–282 and contains a small 3_{10} helix and part of a small α-helix.[85] Gly 272 is strictly conserved in nature presumably because any other residue's side chain would sterically interfere with the side chain of Lys 145 [Fig. 3(B)].[85] The highly conserved Asp 280 and Arg 282 are involved in a salt bridge, that is, likely necessary for stabilizing the enzyme structure [Fig. 3(B)].[65,85]

Molecular modeling was used to generate a DsbA peptide in complex with SPase I Δ2-75 based on the crystal structure of *E. coli* SPase I in complex with β-lactam[65] and lipopeptide inhibitors.[74] This method identified 13 subsites (S7–S6') in SPase I that are appropriate for the interaction of the residues -7 to +6 of the signal peptide and mature protein.[66] The signal peptide binding subsites are named S1–S7. The -1 residue from the substrate binds in the S1 subsite formed by residues Ile 86, Pro 87, Ser 88, Ser 90, Met 91, Leu 95, Tyr 143, Ile 144, and Lys 145 [Fig. 3(A) and Table I].[66,73] The S2 subsite has the deepest cavity within the binding site which can accommodate residues with large side chains at position -2 and the cavity consists of SPase I residues Gln 85, Ile 86, Pro 87, Ser 88, Met 91, and Ile 144 [Fig. 3(A) and Table I].[66] This site was previously identified as S1, due to the overlap of residues between the sites.[73,85] Substrate residue -2 was initially thought to be solvent exposed,[73] however, in the recent computational modeling of substrate binding, the -2 signal peptide residue was completely buried in the SPase I S2 subsite.[66] The S3 subsite is comprised of residues Phe 84, Gln 85, Ile 86, Pro 87, Ile 101, Val 132, Asp 142, and Ile 144 [Fig. 3(A) and Table I].[65,66] Two Ile (86 and 144) are involved in substrate specificity and when they are mutated display a relaxed substrate specificity at the -1 and -3 positions in the signal peptide.[86,87] Subsite S4 is defined by residues Phe 84, Gln 85, Pro 87, and Asp 142, while subsite S5 includes residues Phe 84, Gln 85, and Asp 142.[66] The S6 and S7 subsites consist of Pro 83 and Phe 84, and Glu 82 and Pro 83, respectively.[66] The early mature region of the substrate binds within subsites S1' to S6' on the SPase I where S1' is formed by Ser 88, Ser 90, Tyr 143, and Ala 279 and subsite S2' is fashioned from residues Ser 88, Ser 90, Phe 208, Asn 277, and Ala 279.[66] Ser 88 and Ser 90 are key residues for catalysis and it is noteworthy that they are positioned within the subsites for the substrate -1 residue and the very early mature region. Subsite S3' is composed of residues Tyr 50, Met 249, Asp 276, Asn 277, Ala 279, Arg 282, and Tyr 283 [Fig. 3(A) and Table I].[66] The S4' subsite includes residues Gln 244, Asp

245, Asp 276, Asn 277, and Arg 282 [Fig. 3(A) and Table I], while S5' and S6' are composed of Phe 196, Ser 206, Ala 243, Asp 276, and Asn 277, and Phe 196, Ile 242, and Ala 243, respectively.[66]

Another salient feature of the enzyme structure is the burial of the side chain of Lys 145 in a hydrophobic environment formed by several hydrophobic residues (Tyr 143, Phe 133, Met 270, Met 271, Gly 272, and Ala 279) of the conserved domain.[65] This is perhaps why the pK_a of Lys 145 is approximately 8.7;[80,88] nearly two units less than the solvent-exposed pK_a (10.5) of a Lys side chain.[88,89] The side chain is therefore deprotonated and is now capable of acting as a general base for the cleavage reaction. Ser 278 is highly conserved and hydrogen bonds to Lys 145, thereby positioning it correctly relative to Ser 90 [Fig. 3(B,D)].[65,90] The Oγ atom of Ser 90 is $\sim$2.9 Å apart from Lys 145,[65] which is ideal for the formation of a hydrogen bond.

A common characteristic of detergent solubilized SPase I is that it undergoes autocatalysis C-terminal to the sequence Ala38-X-Ala (*E. coli* nomenclature) normally located in the cytoplasm and thus is inaccessible to the catalytic domain *in vivo*.[28] Much of our current understanding of SPase I comes from a soluble form of SPase I that lacks the transmembrane segments. This mutant, Δ2-75, was produced in the early 1990s by Kuo *et al.*[29] and was found to be enzymatically active *in vitro* although it has a specificity constant approximately one-twentieth that of the wild-type enzyme.[81] This mutant is more amenable to classical purification procedures as it is not a membrane protein. It has, therefore, become the ideal candidate for signal peptidase characterization studies, such as crystallization and activity assays. Although this form of the enzyme has helped to elucidate the catalytic mechanism *in vitro*, we still have much to learn about how the enzyme functions in its native *in vivo* membrane environment. With the advent of new technologies that readily enable the study of membrane proteins in a lipid environment, our knowledge of this enzyme and the importance of the transmembrane domain is set for further development in the next few years.

Much of our current knowledge of the structure of SPase I comes from crystal structures of the Δ2-75 mutant apoenzyme [Fig. 3(A,B)],[73] as well as from the Δ2-75 mutant enzyme bound to synthetic inhibitors.[65,74,75] The subtle changes observed between the apoenzyme and inhibitor-bound enzyme have enabled researchers to examine the conformational alterations that take place at the enzyme active site, as it interacts with an inhibitor, and infer from this potential features of the interaction with substrate. A large rotation of the Trp residues, 300 and 310, located at the suspected membrane interface is observed [Fig. 3(C)]. For Trp 300 there appears to be two positions that the side chain may take. The area around Trp 310 is only resolved in the two structures that were solved in the presence of the lipopeptide inhibitor arylomycin A_2;[74,75] however, they display a large change in the orientation of the Trp side chain. Slight changes are also observed in the active site residues Ser 88 and Lys 145 with some inhibitor-bound structures as compared to the apoenzyme structure [Fig. 3(D)]. Residue Ser 88 points in towards residue Ser 90, however, in the β-lactam-bound structure[65] this residue points away from the active site. The Lys 145 side chain is sterically displaced when bound to an arylomycin A_2 inhibitor [Fig. 3(D)] and is no longer able to make a hydrogen bond with Ser 278.[74] These structures have also helped to explain why common serine protease inhibitors do not inhibit SPase I. Typical serine proteases are inhibited by 5*R* β-lactam stereoisomers, while SPase I inhibitors bind with an opposite stereochemistry (5*S*).[65] Although these structures have provided a starting point for the prediction of how preprotein substrates bind and interact with SPase I, to fully understand the interaction we await the breakthrough of a cocrystal structure of substrate-bound SPase I.

Proposed Proteolytic Mechanism

Once the preprotein substrate binds to the active site, it positions itself such that the -1 and -3 residues of the signal peptide become buried inside the SPase I S1 and S3 hydrophobic subsites. The amino group of Lys 145 acts as the general base and deprotonates the hydroxyl group of Ser 90 (Fig. 5). The Ser 90 Oγ atom now acts as the nucleophile, and attacks the substrate P1 residue carbonyl group on the *si*-face of the scissile peptide bond to form the tetrahedral intermediate I. This shift of electrons results in the formation of an oxyanion hole, involving the main chain amide group of Ser 90 and the side chain hydroxyl group of Ser 88,[65,73] which results in the stabilization of the substrate tetrahedral intermediate I (Fig. 5). Lys 145 donates a proton to the amino group of the N-terminus of the mature protein, allowing its release from the enzyme, and generates a signal peptide acyl-enzyme intermediate (Fig. 5). The deacylating water molecule now comes into play, with the loss of one of its protons to the amino group of Lys 145 and the attack of its oxygen atom on the peptide carbonyl group, forming another tetrahedral intermediate (Fig. 5). Again, Ser 88 and Ser 90 serve to stabilize the intermediate via hydrogen bonding. Finally, the amide group of Lys 145 donates a proton to the Oγ atom of Ser 90, leading to the breakdown of the tetrahedral intermediate, release of the signal peptide, and the restoration of the active site to its apoenzyme form.

SPase I as an Antimicrobial Target

SPase I is critical to study because of its potential as a target for novel antibacterial agents. SPase I has

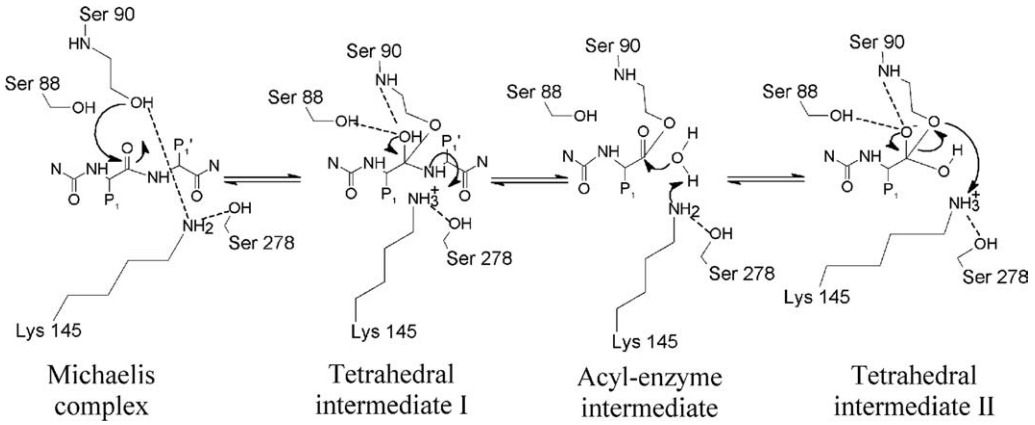

Figure 5. Proposed mechanism for SPase I cleavage of preproteins using a Ser-Lys catalytic dyad. The preprotein substrate (where P_1 is the amino acid at the -1 position of the signal peptide and P_1' is the amino acid in the +1 position of the mature protein) binds to the enzyme active site. The amino group of Lys 145 acts as the general base and deprotonates the hydroxyl group of Ser 90 (Michaelis complex). The Ser 90 Oγ atom now acts as the nucleophile, and attacks the substrate P_1 residue carbonyl group to form a tetrahedral intermediate I. This shift of electrons results in the formation of an oxyanion hole, involving the main chain amide group of Ser 90 and the side chain hydroxyl group of Ser 88,[65,73] which results in the stabilization of the substrate tetrahedral intermediate. Lys 145 donates a proton to the amino group of the N-terminus of the mature protein, allowing its release from the enzyme, and generates a signal peptide acyl-enzyme intermediate. The deacylating water molecule now comes into play, with the loss of one of its protons to the amino group of Lys 145, and the attack of its oxygen atom on the peptide carbonyl group, forming another tetrahedral intermediate. Again, Ser 88 and Ser 90 serve to stabilize the intermediate via hydrogen bonding. Finally, the amide group of Lys 145 donates a proton to the Oγ atom of Ser 90, leading to the breakdown of the tetrahedral intermediate and release of the signal peptide.

been shown to be an essential enzyme for the viability of most bacteria.[3,4,7,17,23] A couple of exceptions are known to exist in bacteria that possess more than one SPase I gene such as *Streptomyces lividans*, where all four SPase I proteins do not appear to be essential,[91] and *B. subtilis*, where at least three SPase I proteins are not essential.[13,92–94] Genarally, when SPase I is inhibited, it leads to an accumulation of secretory preproteins in the inner membrane and eventually cell death.[9,95] Although SPase I is ubiquitous, found in all forms of life from humans to *E. coli*, there are a few key differences between the eukaryotic ER and eubacterial enzymes. Thus, there is the potential to impede bacterial viability without affecting the host. Some key differences are the SPase I cleavage mechanism and oligomeric state; eubacterial SPase I is believed to be monomeric and cleaves using a Ser-Lys dyad, while eukaryotic ER SPase I is multimeric and employs a Ser-His-Asp triad for catalysis. The development of antibiotics against the bacterial SPase I must account for the eukaryotic mitochondrial inner membrane peptidases, Imp1 and Imp2, as they are part of the S26 family. It is unlikely that antibiotics developed for bacterial SPases would affect the Imp proteins as small changes in the SPase I substrate binding site have been shown to greatly reduce the efficacy of an antibiotic. As such, an antibiotic designed for one SPase I may not display a significant inhibition of another SPase I.

Although other proteins involved in protein secretion could also be targets for antibacterial agents, they present the challenge of getting the antibiotics into the inner membrane or cytosol. The catalytic domain of the SPase I enzyme on the other hand is located in the periplasm, making it a much easier target. This is especially true in Gram-positive bacteria which lack an outer membrane. As more pathogens become resistant to our currently employed antibiotics, it is imperative that more work be done to discover new antibiotics to help fight infections.

A number of β-lactam compounds[96–98] and lipopeptides[99] were found to inhibit SPase I. β-lactam compounds have been used to prepare 5*S* penems, which have the appropriate stereochemistry to inhibit SPase I. The most interesting of these compounds are the 6-substituted penems with 5*S*,6*S*, 1'*R* stereochemistry.[100] The penem derivative, (5*S*,6*S*)-6-[(*R*)-acetoxyethyl]-penem-3-carboxylate, has been shown to moderately inhibit SPase I activity of *E. coli* LepB,[97] *Staphylococcus aureus* SpsB,[85] *S. lividans* SipW, SipX, SipY, and SipZ,[101] *Legionella pneumophila* LepB,[17] and cyanobacteria and chloroplast thylakoid SPases.[102] More potent inhibitors of *S. aureus* SpsB have been developed in the form of non-cleavable lipopeptides, based on substrate sequences, but these contain a proline at the -1 position, that result in an IC_{50} of 0.6 μM.[103] In the same manner, a linear peptide based inhibitor has been designed with a proline in the +1 position relative to the signal peptide cleavage site, which allows binding of the peptide, but not cleavage.[104,105] This peptide was shown to inhibit *S. aureus* SpsB activity in a dose-dependent manner, but millimolar concentrations were needed for significant inhibition and thus

is not potent enough to be useful.[105] Recently, another substrate-based peptide aldehyde was shown to inhibit *S. aureus* SpsB with an IC_{50} value of 0.09 μM.[25]

The antibiotic arylomycin lipohexapeptides and lipoglycopeptides initially seemed to inhibit only a few Gram-positive bacteria.[106,107] However, recently it was discovered that arylomycin A_2 (MIC = 1 μg/mL) and a synthetic derivative, C_{16}, (MIC = 0.25 μg/mL) are potent inhibitors of *Staphylococcus epidermidis*,[19,22,108] a Gram-positive bacteria responsible for infections among people with compromised immune systems, as well as *Staphylococcus haemolyticus*, *Staphylococcus hominis*, and *Staphylococcus lugdunensis*.[109] Arylomycin A_2 displays only moderate antibacterial activity against *E. coli*, *Streptococcus pneumoniae*, *S. aureus*, and *Haemophilus influenzae*.[19] The inability of these antibiotics to efficaciously inhibit *E. coli*, *S. aureus*, and *Pseudomonas aeruginosa* SPase I is due in part to the presence of a proline residue in the substrate binding site on SPase I (residue 29 in *S. aureus*, and residue 84 in *E. coli* and *P. aeruginosa*) which seems to confer resistance to arylomycins.[106] The resistance may be due to the inability of the proline residue to form a hydrogen bond with the inhibitor. It has also been suggested that the outer membrane of Gram-negative bacteria prevents the arylomycins from reaching the SPase I.[20] Recently, the derivative arylomycin B-C_{16} was developed and tested for its ability to inhibit growth of a broad range of bacteria.[108] It was found to have nearly identical activities against most bacteria as arylomycin C_{16}, except that the B-C_{16} derivative displays activity against the important pathogen *Streptococcus agalactiae*.[18] As protein secretion is essential in many bacteria for virulence, there is the potential for an arylomycin class or lipopeptide-type antibiotic to inhibit virulence and viability of some Gram-positive bacteria, but these compounds are less likely to be effective, even upon modification, in Gram-negative bacteria.

Structural variations in the active site of enzymes from different species may well explain differences in the specificity of the antimicrobial agents tested. This can be readily observed in the case of the inability of arylomycin to inhibit SPase I in bacteria with a proline residue located in the substrate binding site. These small structural variations may diminish the likelihood for development of broad spectrum antibiotics that treat a number of bacterial infections; however, the selectivity of antibiotics for a particular SPase I can be valuable in specifically treating one type of infection versus another.

Concluding Remarks

The SPase I is a unique serine endoprotease that is essential for bacterial viability. A number of critical features differentiate the bacterial and eukaryotic ER SPase I and, therefore, make the bacterial SPase I a tantalizing target for the development of antibacterial agents. Although SPase I is not inhibited by serine peptidase inhibitors, a number of inhibitors have been identified such as the penem-type inhibitors and lipopeptides. With the development of high-throughput inhibitor screening assays,[22] a number of new potent inhibitors will likely be discovered in the coming years.

Several important questions still remain unanswered about how and where SPase I functions. Does SPase I bind SecYEG and/or TatA in the membrane or is it laterally mobile in the membrane? How does the cleaved signal peptide spatially transition from the SPase I and be presented for degradation by the signal peptide peptidase? Does SPase I interact with the signal peptide peptidase to transfer the signal peptide? What role does the membrane anchor of SPase I play in signal peptide cleavage? Does it bind to the h-region of the signal peptide? Does cleavage of the signal peptide occur in the inner membrane, in the periplasm, or in the translocon? Why do some species have multiple type I signal peptidases? While the three-dimensional structure of the *E. coli* SPase I Δ2-75 has been solved, the structure of the full-length SPase I and a substrate-bound SPase would help resolve these questions. Answers to these questions will be an important step to help with the development of SPase I inhibitors.

Acknowledgments

The authors thank Paolo De Bona for discussion of the signal peptide structure when bound to SPase and dodecylphosphocholine micelles.

References

1. Choo KH, Tan TW, Ranganathan S (2005) SPdb—a signal peptide database. BMC Bioinformatics 6: 249.
2. Dalbey RE, Wickner W (1985) Leader peptidase catalyzes the release of exported proteins from the outer surface of the *Escherichia coli* plasma membrane. J Biol Chem 260: 15925–15931.
3. Zhbanko M, Zinchenko V, Gutensohn M, Schierhorn A, Klosgen RB (2005) Inactivation of a predicted leader peptidase prevents photoautotrophic growth of *Synechocystis* sp. strain PCC 6803. J Bacteriol 187: 3071–3078.
4. Cregg KM, Wilding I, Black MT (1996) Molecular cloning and expression of the spsB gene encoding an essential type I signal peptidase from *Staphylococcus aureus*. J Bacteriol 178: 5712–5718.
5. Date T (1983) Demonstration by a novel genetic technique that leader peptidase is an essential enzyme of *Escherichia coli*. J Bacteriol 154: 76–83.
6. Klug G, Jager A, Heck C, Rauhut R (1997) Identification, sequence analysis, and expression of the lepB gene for a leader peptidase in *Rhodobacter capsulatus*. Mol Gen Genet 253: 666–673.
7. Taheri T, Salmanian AH, 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. Exp Parasitol 126: 135–145.
8. Zhang YB, Greenberg B, Lacks SA (1997) Analysis of a *Streptococcus pneumoniae* gene encoding signal peptidase I and overproduction of the enzyme. Gene 194: 249–255.
9. Nahrstedt H, Wittchen K, Rachman MA, Meinhardt F (2004) Identification and functional characterization of a type I signal peptidase gene of *Bacillus megaterium* DSM319. Appl Microbiol Biotechnol 64: 243–249.
10. du Plessis DJ, Nouwen N, Driessen AJ (2010) The Sec translocase. Biochim Biophys Acta 1808: 851–865.
11. Luke I, Handford JI, Palmer T, Sargent F (2009) Proteolytic processing of *Escherichia coli* twin-arginine signal peptides by LepB. Arch Microbiol 191: 919–925.
12. Meijer WJ, de Jong A, Bea G, Wisman A, Tjalsma H, Venema G, Bron S, van Dijl JM (1995) The endogenous *Bacillus subtilis* (natto) plasmids pTA1015 and pTA1040 contain signal peptidase-encoding genes: identification of a new structural module on cryptic plasmids. Mol Microbiol 17: 621–631.
13. Tjalsma H, Bolhuis A, van Roosmalen ML, Wiegert T, Schumann W, Broekhuizen CP, Quax WJ, Venema G, Bron S, van Dijl JM (1998) Functional analysis of the secretory precursor processing machinery of *Bacillus subtilis*: identification of a eubacterial homolog of archaeal and eukaryotic signal peptidases. Genes Dev 12: 2318–2331.
14. Tjalsma H, Noback MA, Bron S, Venema G, Yamane K, van Dijl JM (1997) *Bacillus subtilis* contains four closely related type I signal peptidases with overlapping substrate specificities. Constitutive and temporally controlled expression of different sip genes. J Biol Chem 272: 25983–25992.
15. Raynaud C, Charbit A (2005) Regulation of expression of type I signal peptidases in *Listeria monocytogenes*. Microbiology 151: 3769–3776.
16. Powers ME, Smith PA, Roberts TC, Fowler BJ, King CC, Trauger SA, Siuzdak G, Romesberg FE (2011) Type I signal peptidase and protein secretion in *Staphylococcus epidermidis*. J Bacteriol 193: 340–348.
17. Lammertyn E, Van Mellaert L, Meyen E, Lebeau I, De Buck E, Anne J, Geukens N (2004) Molecular and functional characterization of type I signal peptidase from *Legionella pneumophila*. Microbiology 150: 1475–1483.
18. Roberts TC, Smith PA, Romesberg FE (2011) Synthesis and biological characterization of arylomycin B antibiotics. J Nat Prod 74: 956–961.
19. Smith PA, Powers ME, Roberts TC, Romesberg FE (2011) In vitro activities of arylomycin natural-product antibiotics against *Staphylococcus epidermidis* and other coagulase-negative staphylococci. Antimicrob Agents Chemother 55: 1130–1134.
20. Smith PA, Roberts TC, Romesberg FE (2010) Broad-spectrum antibiotic activity of the arylomycin natural products is masked by natural target mutations. Chem Biol 17: 1223–1231.
21. Harris DA, Powers ME, Romesberg FE (2009) Synthesis and biological evaluation of penem inhibitors of bacterial signal peptidase. Bioorg Med Chem Lett 19: 3787–3790.
22. Bockstael K, Geukens N, Rao CV, Herdewijn P, Anne J, Van Aerschot A (2009) An easy and fast method for the evaluation of *Staphylococcus epidermidis* type I signal peptidase inhibitors. J Microbiol Methods 78: 231–237.
23. Bockstael K, Geukens N, Van Mellaert L, Herdewijn P, Anne J, Van Aerschot A (2009) Evaluation of the type I signal peptidase as antibacterial target for biofilm-associated infections of *Staphylococcus epidermidis*. Microbiology 155: 3719–3729.
24. Roberts TC, Schallenberger MA, Liu J, Smith PA, Romesberg FE (2011) Initial efforts toward the optimization of arylomycins for antibiotic activity. J Med Chem 54: 4954–4963.
25. Buzder-Lantos P, Bockstael K, Anne J, Herdewijn P (2009) Substrate based peptide aldehyde inhibits bacterial type I signal peptidase. Bioorg Med Chem Lett 19: 2880–2883.
26. Rawlings ND, Barrett AJ (1993) Evolutionary families of peptidases. Biochem J 290: 205–218.
27. Rawlings ND, Barrett AJ (1994) Families of serine peptidases. Methods Enzymol 244: 19–61.
28. Talarico TL, Dev IK, Bassford PJ,Jr, Ray PH (1991) Inter-molecular degradation of signal peptidase I in vitro. Biochem Biophys Res Commun 181: 650–656.
29. Kuo DW, Chan HK, Wilson CJ, Griffin PR, Williams H, Knight WB (1993) *Escherichia coli* leader peptidase: production of an active form lacking a requirement for detergent and development of peptide substrates. Arch Biochem Biophys 303: 274–280.
30. Kim YT, Muramatsu T, Takahashi K (1995) Leader peptidase from *Escherichia coli*: overexpression, characterization, and inactivation by modification of tryptophan residues 300 and 310 with N-bromosuccinimide. J Biochem 117: 535–544.
31. Black MT, Munn JG, Allsop AE (1992) On the catalytic mechanism of prokaryotic leader peptidase 1. Biochem J 282: 539–543.
32. Zwizinski C, Date T, Wickner W (1981) Leader peptidase is found in both the inner and outer membranes of *Escherichia coli*. J Biol Chem 256: 3593–3597.
33. Breukink E, Nouwen N, van Raalte A, Mizushima S, Tommassen J, de Kruijff B (1995) The C terminus of SecA is involved in both lipid binding and SecB binding. J Biol Chem 270: 7902–7907.
34. Randall LL, Henzl MT (2010) Direct identification of the site of binding on the chaperone SecB for the amino terminus of the translocon motor SecA. Protein Sci 19: 1173–1179.
35. Fekkes P, van der Does C, Driessen AJ (1997) The molecular chaperone SecB is released from the carboxy-terminus of SecA during initiation of precursor protein translocation. EMBO J 16: 6105–6113.
36. Josefsson LG, Randall LL (1981) Processing in vivo of precursor maltose-binding protein in *Escherichia coli* occurs post-translationally as well as co-translationally. J Biol Chem 256: 2504–2507.
37. Wang P, Shim E, Cravatt B, Jacobsen R, Schoeniger J, Kim AC, Paetzel M, Dalbey RE (2008) *Escherichia coli* signal peptide peptidase A is a serine-lysine protease with a lysine recruited to the nonconserved amino-terminal domain in the S49 protease family. Biochemistry 47: 6361–6369.
38. Kim AC, Oliver DC, Paetzel M (2008) Crystal structure of a bacterial signal peptide peptidase. J Mol Biol 376: 352–366.
39. Dalbey RE, Chen M (2004) Sec-translocase mediated membrane protein biogenesis. Biochim Biophys Acta 1694: 37–53.
40. Berks BC (1996) A common export pathway for proteins binding complex redox cofactors?Mol Microbiol 22: 393–404.
41. Weiner JH, Bilous PT, Shaw GM, Lubitz SP, Frost L, Thomas GH, Cole JA, Turner RJ (1998) A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. Cell 93: 93–101.

42. Sargent F, Stanley NR, Berks BC, Palmer T (1999) Sec-independent protein translocation in *Escherichia coli*. A distinct and pivotal role for the TatB protein. J Biol Chem 274: 36073–36082.
43. Bogsch EG, Sargent F, Stanley NR, Berks BC, Robinson C, Palmer T (1998) An essential component of a novel bacterial protein export system with homologues in plastids and mitochondria. J Biol Chem 273: 18003–18006.
44. Sargent F, Bogsch EG, Stanley NR, Wexler M, Robinson C, Berks BC, Palmer T (1998) Overlapping functions of components of a bacterial Sec-independent protein export pathway. EMBO J 17: 3640–3650.
45. Pop O, Martin U, Abel C, Muller JP (2002) The twin-arginine signal peptide of PhoD and the TatAd/Cd proteins of *Bacillus subtilis* form an autonomous Tat translocation system. J Biol Chem 277: 3268–3273.
46. Jongbloed JD, Grieger U, Antelmann H, Hecker M, Nijland R, Bron S, van Dijl JM (2004) Two minimal Tat translocases in *Bacillus*. Mol Microbiol 54: 1319–1325.
47. Rodrigue A, Chanal A, Beck K, Muller M, Wu LF (1999) Co-translocation of a periplasmic enzyme complex by a hitchhiker mechanism through the bacterial tat pathway. J Biol Chem 274: 13223–13228.
48. Bolhuis A, Mathers JE, Thomas JD, Barrett CM, Robinson C (2001) TatB and TatC form a functional and structural unit of the twin-arginine translocase from *Escherichia coli*. J Biol Chem 276: 20213–20219.
49. Alami M, Luke I, Deitermann S, Eisner G, Koch HG, Brunner J, Muller M (2003) Differential interactions between a twin-arginine signal peptide and its translocase in *Escherichia coli*. Mol Cell 12: 937–946.
50. Sargent F, Gohlke U, De Leeuw E, Stanley NR, Palmer T, Saibil HR, Berks BC (2001) Purified components of the *Escherichia coli* Tat protein transport system form a double-layered ring structure. Eur J Biochem 268: 3361–3367.
51. Porcelli I, de Leeuw E, Wallis R, van den Brink-van der Laan E, de Kruijff B, Wallace BA, Palmer T, Berks BC (2002) Characterization and membrane assembly of the TatA component of the *Escherichia coli* twin-arginine protein transport system. Biochemistry 41: 13690–13697.
52. Oates J, Barrett CM, Barnett JP, Byrne KG, Bolhuis A, Robinson C (2005) The *Escherichia coli* twin-arginine translocation apparatus incorporates a distinct form of TatABC complex, spectrum of modular TatA complexes and minor TatAB complex. J Mol Biol 346: 295–305.
53. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Int J Neural Syst 8: 581–599.
54. Nielsen H, Engelbrecht J, Brunak S, von Heijne G (1997) Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 10: 1–6.
55. von Heijne G, Abrahmsen L (1989) Species-specific variation in signal peptide design. Implications for protein secretion in foreign hosts. FEBS Lett 244: 439–446.
56. Tjalsma H, Bolhuis A, Jongbloed JD, Bron S, van Dijl JM (2000) Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. Microbiol Mol Biol Rev 64: 515–547.
57. von Heijne G (1983) Patterns of amino acids near signal-sequence cleavage sites. Eur J Biochem 133: 17–21.
58. von Heijne G (1985) Signal sequences. The limits of variation. J Mol Biol 184: 99–105.
59. Perlman D, Halvorson HO (1983) A putative signal peptidase recognition site and sequence in eukaryotic and prokaryotic signal peptides. J Mol Biol 167: 391–409.
60. Shen LM, Lee JI, Cheng SY, Jutte H, Kuhn A, Dalbey RE (1991) Use of site-directed mutagenesis to define the limits of sequence variation tolerated for processing of the M13 procoat protein by the *Escherichia coli* leader peptidase. Biochemistry 30: 11775–11781.
61. Fikes JD, Barkocy-Gallagher GA, Klapper DG, Bassford PJ,Jr (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. J Biol Chem 265: 3417–3423.
62. Jain RG, Rusch SL, Kendall DA (1994) Signal peptide cleavage regions. Functional limits on length and topological implications. J Biol Chem 269: 16305–16310.
63. von Heijne G (1990) The signal peptide. J Membr Biol 115: 195–201.
64. Izard JW, Kendall DA (1994) Signal peptides: exquisitely designed transport promoters. Mol Microbiol 13: 765–773.
65. Paetzel M, Dalbey RE, Strynadka NC (1998) Crystal structure of a bacterial signal peptidase in complex with a beta-lactam inhibitor. Nature 396: 186–190.
66. Choo KH, Tong JC, Ranganathan S (2008) Modeling *Escherichia coli* signal peptidase complex with bound substrate: determinants in the mature peptide influencing signal peptide cleavage. BMC Bioinformatics 9: S15.
67. Kajava AV, Zolov SN, Kalinin AE, Nesmeyanova MA (2000) The net charge of the first 18 residues of the mature sequence affects protein translocation across the cytoplasmic membrane of gram-negative bacteria. J Bacteriol 182: 2163–2169.
68. Choo KH, Ranganathan S (2008) Flanking signal and mature peptide residues influence signal peptide cleavage. BMC Bioinformatics 9: S15.
69. Geller B, Zhu HY, Cheng S, Kuhn A, Dalbey RE (1993) Charged residues render pro-OmpA potential dependent for initiation of membrane translocation. J Biol Chem 268: 9442–9447.
70. Yamane K, Mizushima S (1988) Introduction of basic amino acid residues after the signal peptide inhibits protein translocation across the cytoplasmic membrane of *Escherichia coli*. Relation to the orientation of membrane proteins. J Biol Chem 263: 19690–19696.
71. Li P, Beckwith J, Inouye H (1988) Alteration of the amino terminus of the mature sequence of a periplasmic protein can severely affect protein export in *Escherichia coli*. Proc Natl Acad Sci USA 85: 7685–7689.
72. Bogsch E, Brink S, Robinson C (1997) Pathway specificity for a delta pH-dependent precursor thylakoid lumen protein is governed by a 'Sec-avoidance' motif in the transfer peptide and a 'Sec-incompatible' mature protein. EMBO J 16: 3851–3859.
73. Paetzel M, Dalbey RE, Strynadka NC (2002) Crystal structure of a bacterial signal peptidase apoenzyme: implications for signal peptide binding and the Ser-Lys dyad mechanism. J Biol Chem 277: 9512–9519.
74. Paetzel M, Goodall JJ, Kania M, Dalbey RE, Page MG (2004) Crystallographic and biophysical analysis of a bacterial signal peptidase in complex with a lipopeptide-based inhibitor. J Biol Chem 279: 30781–30790.
75. Luo C, Roussel P, Dreier J, Page MG, Paetzel M (2009) Crystallographic analysis of bacterial signal peptidase in ternary complex with arylomycin A2 and a beta-sultam inhibitor. Biochemistry 48: 8976–8984.
76. Whitley P, Nilsson L, von Heijne G (1993) Three-dimensional model for the membrane domain of

Escherichia coli leader peptidase based on disulfide mapping. Biochemistry 32: 8534–8539.
77. Carlos JL, Paetzel M, Brubaker G, Karla A, Ashwell CM, Lively MO, Cao G, Bullinger P, Dalbey RE (2000) The role of the membrane-spanning domain of type I signal peptidases in substrate cleavage site selection. J Biol Chem 275: 38813–38822.
78. Dalbey RE, Lively MO, Bron S, van Dijl JM (1997) The chemistry and enzymology of the type I signal peptidases. Protein Sci 6: 1129–1138.
79. Sung M, Dalbey RE (1992) Identification of potential active-site residues in the *Escherichia coli* leader peptidase. J Biol Chem 267: 13154–13159.
80. Paetzel M, Strynadka NC, Tschantz WR, Casareno R, Bullinger PR, Dalbey RE (1997) Use of site-directed chemical modification to study an essential lysine in *Escherichia coli* leader peptidase. J Biol Chem 272: 9994–10003.
81. Tschantz WR, Paetzel M, Cao G, Suciu D, Inouye M, Dalbey RE (1995) Characterization of a soluble, catalytically active form of *Escherichia coli* leader peptidase: requirement of detergent or phospholipid for optimal activity. Biochemistry 34: 3935–3941.
82. Kim YT, Muramatsu T, Takahashi K (1995) Identification of Trp300 as an important residue for *Escherichia coli* leader peptidase activity. Eur J Biochem 234: 358–362.
83. Landolt-Marticorena C, Williams KA, Deber CM, Reithmeier RA (1993) Non-random distribution of amino acids in the transmembrane segments of human type I single span membrane proteins. J Mol Biol 229: 602–608.
84. Whitley P, von Heijne G (1993) The DsbA-DsbB system affects the formation of disulfide bonds in periplasmic but not in intramembraneous protein domains. FEBS Lett 332: 49–51.
85. Paetzel M, Dalbey RE, Strynadka NC (2000) The structure and mechanism of bacterial type I signal peptidases. A novel antibiotic target. Pharmacol Ther 87: 27–49.
86. Karla A, Lively MO, Paetzel M, Dalbey R (2005) The identification of residues that control signal peptidase cleavage fidelity and substrate specificity. J Biol Chem 280: 6731–6741.
87. Ekici OD, Karla A, Paetzel M, Lively MO, Pei D, Dalbey RE (2007) Altered -3 substrate specificity of *Escherichia coli* signal peptidase 1 mutants as revealed by screening a combinatorial peptide library. J Biol Chem 282: 417–425.
88. Planas A, Kirsch JF (1991) Reengineering the catalytic lysine of aspartate aminotransferase by chemical elaboration of a genetically introduced cysteine. Biochemistry 30: 8268–8276.
89. Paetzel M, Dalbey RE (1997) Catalytic hydroxyl/amine dyads within serine proteases. Trends Biochem Sci 22: 28–31.
90. Klenotic PA, Carlos JL, Samuelson JC, Schuenemann TA, Tschantz WR, Paetzel M, Strynadka NC, Dalbey RE (2000) The role of the conserved box E residues in the active site of the *Escherichia coli* type I signal peptidase. J Biol Chem 275: 6490–6498.
91. Palacin A, Parro V, Geukens N, Anne J, Mellado RP (2002) SipY Is the *Streptomyces lividans* type I signal peptidase exerting a major effect on protein secretion. J Bacteriol 184: 4875–4880.
92. Stover AG, Driks A (1999) Control of synthesis and secretion of the *Bacillus subtilis* protein YqxM. J Bacteriol 181: 7065–7069.
93. Stover AG, Driks A (1999) Secretion, localization, and antibacterial activity of TasA, a *Bacillus subtilis* spore-associated protein. J Bacteriol 181: 1664–1672.
94. Tjalsma H, van den Dolder J, Meijer WJ, Venema G, Bron S, van Dijl JM (1999) The plasmid-encoded signal peptidase SipP can functionally replace the major signal peptidases SipS and SipT of *Bacillus subtilis*. J Bacteriol 181: 2448–2454.
95. Koshland D, Sauer RT, Botstein D (1982) Diverse effects of mutations in the signal sequence on the secretion of beta-lactamase in *Salmonella typhimurium*. Cell 30: 903–914.
96. Allsop A, Brooks G, Edwards PD, Kaura AC, Southgate R (1996) Inhibitors of bacterial signal peptidase: a series of 6-(substituted oxyethyl)penems. J Antibiot 49: 921–928.
97. Black MT, Bruton G (1998) Inhibitors of bacterial signal peptidases. Curr Pharm Des 4: 133–154.
98. Kuo D, Weidner J, Griffin P, Shah SK, Knight WB (1994) Determination of the kinetic parameters of *Escherichia coli* leader peptidase activity using a continuous assay: the pH dependence and time-dependent inhibition by beta-lactams are consistent with a novel serine protease mechanism. Biochemistry 33: 8347–8354.
99. Paetzel M, Karla A, Strynadka NC, Dalbey RE (2002) Signal peptidases. Chem Rev 102: 4549–4580.
100. Allsop A, Brooks G, Bruton G, Coulton S, Edwards PD, Hatton IK, Kaura AC, Maclean SD, Pearson ND, Smale TC, Southgate R (1995) Penem inhibitors of bacterial signal peptidase. Bioorg Med Chem Lett 5: 443–448.
101. Geukens N, Lammertyn E, Van Mellaert L, Engelborghs Y, Mellado RP, Anne J (2002) Physical requirements for in vitro processing of the *Streptomyces lividans* signal peptidases. J Biotechnol 96: 79–91.
102. Barbrook AC, Packer JC, Howe CJ (1996) Inhibition by penem of processing peptidases from cyanobacteria and chloroplast thylakoids. FEBS Lett 398: 198–200.
103. Bruton G, Huxley A, O'Hanlon P, Orlek B, Eggleston D, Humphries J, Readshaw S, West A, Ashman S, Brown M, Moore K, Pope A, O'Dwyer K, Wang L (2003) Lipopeptide substrates for SpsB, the *Staphylococcus aureus* type I signal peptidase: design, conformation and conversion to alpha-ketoamide inhibitors. Eur J Med Chem 38: 351–356.
104. Barkocy-Gallagher GA, Bassford PJ,Jr (1992) Synthesis of precursor maltose-binding protein with proline in the +1 position of the cleavage site interferes with the activity of *Escherichia coli* signal peptidase I in vivo. J Biol Chem 267: 1231–1238.
105. Kavanaugh JS, Thoendel M, Horswill AR (2007) A role for type I signal peptidase in *Staphylococcus aureus* quorum sensing. Mol Microbiol 65: 780–798.
106. Kulanthaivel P, Kreuzman AJ, Strege MA, Belvo MD, Smitka TA, Clemens M, Swartling JR, Minton KL, Zheng F, Angleton EL, Mullen D, Jungheim LN, Klimkowski VJ, Nicas TI, Thompson RC, Peng SB (2004) Novel lipoglycopeptides as inhibitors of bacterial signal peptidase I. J Biol Chem 279: 36250–36258.
107. Schimana J, Gebhardt K, Holtzel A, Schmid DG, Sussmuth R, Muller J, Pukall R, Fiedler HP (2002) Arylomycins A and B, new biaryl-bridged lipopeptide antibiotics produced by *Streptomyces* sp. Tu 6075. I. Taxonomy, fermentation, isolation and biological activities. J Antibiot 55: 565–570.
108. Roberts TC, Smith PA, Cirz RT, Romesberg FE (2007) Structural and initial biological analysis of synthetic arylomycin A2. J Am Chem Soc 129: 15830–15838.
109. Cheung GY, Otto M (2010) Understanding the significance of *Staphylococcus* epidermidis bacteremia in babies and children. Curr Opin Infect Dis 23: 208–216.