Signal Peptidases

Mark Paetzel,[†] Andrew Karla,[‡] Natalie C. J. Strynadka,[†] and Ross E. Dalbey*[,‡]

Department of Chemistry, The Ohio State University, Columbus, Ohio 43210, and Department of Biochemistry and Molecular Biology, University of British Columbia, Vancouver VGT IZ3, British Columbia, Canada

Received April 9, 2002

Contents

"Don't follow leaders
Watch the parkin' meters"

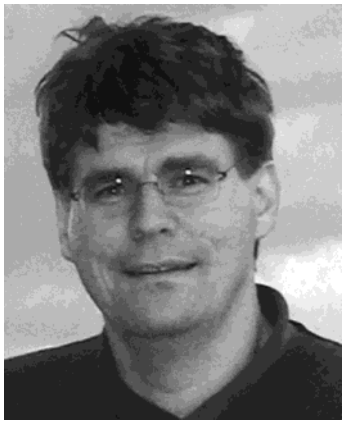
Bob Dylan
Subterranean Homesick Blues

I. Introduction

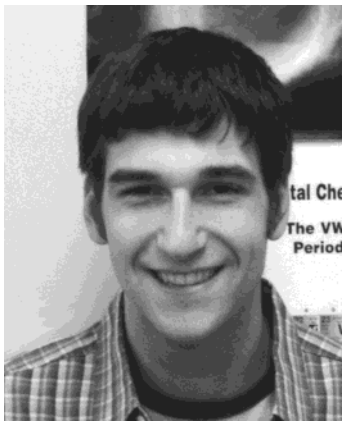
In 1971, Günter Blobel and David Sabatini published the first report of the signal hypothesis where they propose that proteins contain information within their amino acid sequences for protein targeting to the membrane.[1] Soon after this report, Milstein and co-workers discovered that the light chain of IgG from myeloma cells was synthesized in a higher molecular weight form and was converted to its mature form when endoplasmic reticulum vesicles (microsomes) were added to the translation system.[2] Milstein proposed a model based on these results in which microsomes contain a protease that converts the precursor protein form to the mature form by removing the amino-terminal extension peptide. Blobel expanded the signal hypothesis in 1975 with his prediction that proteins localized to different intracellular membranes, such as the mitochondria and chloroplast, would have distinct targeting sequences.[3] These distinct targeting sequences were later found to be cleaved from the exported protein by specific signal peptidases (SPases). In 1999, Günter Blobel received the Nobel Prize in Physiology or Medicine for discovering that proteins have intrinsic signals that govern their transport and localization in the cell.

* To whom correspondence should be addressed. Fax: (614) 292-1532. E-mail: dalbey@chemistry.ohio-state.edu.
[†] University of British Columbia.
[‡] The Ohio State University.

10.1021/cr010166y CCC: $39.75
Published on Web 11/08/2002



Mark Paetzel received a B.Sc. degree in biology in 1984 from Syracuse University and a B.Sc. degree in chemistry from the University of Minnesota in 1991. He obtained his Ph.D. degree in chemistry in 1997 at The Ohio State University. His graduate research involved investigations into the catalytic mechanism and crystallization of bacterial signal peptidase. His work as an MRC postdoctoral fellow at the Department of Biochemistry and Molecular Biology, University of British Columbia (1997–2002), has focused on the crystallographic structure determination of signal peptidases. Major research interests besides the structure and function of signal peptidase include membrane-bound serine proteinases, Ser-Lys dyad proteases and amidases, and protein targeting and translocation systems. He will move to the Department of Molecular Biology and Biochemistry at Simon Fraser University in the fall of 2002 as an assistant professor. He is a recipient of the Michael Smith Foundation for Health Research Scholar Award (2002–2007).



Andrew Karla was born in Cleveland, OH, in 1975. He received a B.S. in biochemistry from the University of Dayton in 1998. He is currently conducting research toward a Ph.D. under Professor Ross Dalbey at The Ohio State University and is participating in the NIH-funded training program for interdisciplinary research at the interface of chemistry and biology. His current work centers on the examination of the determinants of signal peptidase substrate recognition and cleavage.

Natalie C. J. Strynadka received her B.Sc. (Hons/1985) and Ph.D. (1990) in biochemistry at the University of Alberta. As a Gordin Kaplan Memorial postdoctoral fellow her work focused on the crystallographic analysis and inhibition of class A β-lactamases. In 1997 she moved to the Department of Biochemistry and Molecular Biology at the University of British Columbia as an assistant professor and was appointed Associate Professor in 2002. Major research interests include mechanisms of antibiotic resistance, protein secretion systems, and novel antibiotic development at the bacterial membrane. Dr. Strynadka is a Howard Hughes International Scholar, Burroughs Wellcome New Investigator, and CIHR Investigator.

Ross E. Dalbey was born in Martinez, CA. He received his B.S. degree in chemistry at the University of Washington in 1978 and his Ph.D. degree in biochemistry at Washington State University with Ralph Yount in 1983. After postdoctoral work in the laboratory of William Wickner at UCLA, he took a faculty position in 1987 in the Department of Chemistry at The Ohio State University, where he is now a professor. Major research interests are the characterization of proteases that participate in protein export and the understanding of the mechanisms by which proteins insert into biological membranes.

In this review we will focus on SPases, a family of proteases that cleave signal peptides from exported proteins within the cell. We will describe what is known about the targeting sequences that are processed by the proteases. We will review the different substrates for SPases and how these enzymes are assayed for activity. A detailed description of the proteases will be reported. Finally, we will report on what is known about the fate of the signal peptide after it is removed from the exported protein.

A. Overview of Protein Targeting

Protein export to the outer membrane and periplasm of Gram-negative bacteria and the extracellular compartment of Gram-positive bacteria often requires a cleavable signal sequence (Figure 1A). There are at least three distinct SPases involved in cleaving signal peptides in bacteria. SPase I can process nonlipoprotein substrates that are exported by the SecYEG pathway or the twin arginine translocation (Tat) pathway. Lipoproteins that are exported by the Sec pathway are cleaved by SPase II. SPase IV cleaves type IV prepilins and prepilin-like proteins that are components of the type II secretion apparatus (not shown in Figure 1A).

In eukaryotes, proteins that are targeted to the endoplasmic reticulum (ER) membrane are mediated by signal peptides that target the protein either cotranslationally or post-translationally to the Sec61 translocation machinery (Figure 1B). The ER signal peptides have features similar to those of their

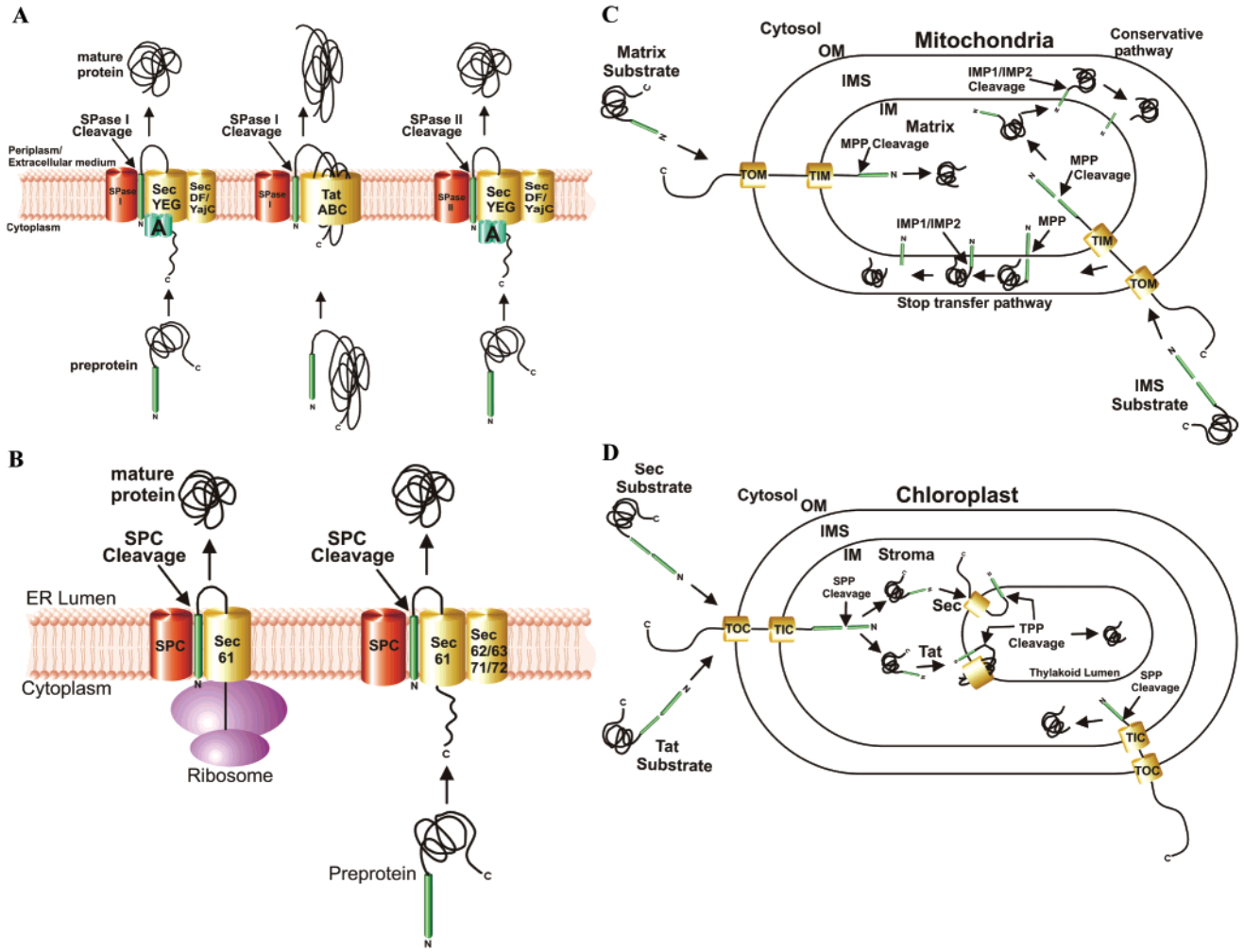


Figure 1. Protein export in bacteria and eukaryotes. (A) Export across the plasma membrane of bacteria. Preproteins are translocated across the plasma membrane using either the Sec translocase or the Tat pathway. The Sec translocase is comprised of two heterotrimeric complexes, SecYEG and SecDFYajC, and SecA, which is a translocation ATPase that drives translocation of hydrophilic regions across the membrane by undergoing cycles of membrane insertion/deinsertion. Nonlipoproteins are exported by the Sec translocase (left side) and are cleaved by SPase I after translocation. Lipoproteins are cleaved by SPase II after Sec-dependent translocation across the plasma membrane (right side). Translocation of proteins that contain twin arginine signal peptides occurs via the Tat translocase comprised of TatA, -B, and -C. After membrane translocation, the preprotein is cleaved by SPase I. (B) Export across the ER membrane. Ribosome-bound preproteins are targeted to the membrane cotranslationally by an SRP-mediated pathway (not shown). The protein is then translocated across the membrane via the Sec61 translocon (Sec61α, Sec61β, and Sec61γ) and TRAM. The signal peptide is cleaved from the preprotein by the SPC. Preproteins are also exported across the ER membranes post-translationally. In yeast, this pathway involves the heterotrimeric Sec61 translocon, the proteins Sec62, Sec63, Sec71, and Sec72, and Kar2p, an ER lumenal chaperone. (C) Protein import into the mitochondrial matrix and mitochondrial intermembrane space. Proteins imported to the mitochondrial matrix typically are synthesized in a higher molecular form with a cleavable matrix signal peptide. The protein is imported across the outer and inner membranes via TOM and TIM. TOM is comprised of the receptors Tom20 and Tom70, and the general import pore (GIP) complex comprised of Tom22, Tom5, Tom40 (the import channels), Tom6, and Tom7. TIM is composed of Tim23, Tim17, and Tim 44. Import into the mitochondrial intermembrane space requires bipartite signal sequences. Most of these intermembrane space proteins are imported by the stop transfer pathway, while a few are imported by the conservative pathway. Proteins imported by the stop transfer pathway cross the outer membrane through the TOM complex and insert into the inner membrane transiently via the TIM23 complex. After membrane insertion, the matrix signal of the imported protein translocates into the matrix where it is cleaved by MPP. The intermembrane space sorting signal acts as a stop transfer, preventing the downstream carboxyl-terminal region from being imported into the matrix. The stop transfer sequence is cleaved by the mitochondrial SPase I (Imp1p/Imp2p). Proteins imported by the conservative pathway are first imported into the matrix through the TOM and TIM complexes. After cleavage of the preprotein by MPP, the protein is then re-exported across the mitochondrial inner membrane. In some cases, this requires Oxa1p. Export is initiated by the intermembrane space sorting signal that resembles the bacterial-like signal peptide. The preprotein is processed by Imp1p/Imp2p after translocation into the intermembrane space. (D) Import into the chloroplast stroma space and thylakoid lumen. Proteins imported into the stromal space require a cleavable chloroplast transit peptide. The protein is imported through the TOC and TIC complexes. The TOC complex is comprised of Toc34, Toc64, Toc160, and Toc75 (the import channel). The TIC complex is comprised of Tic110, Tic55, Tic40, Tic22, and Tic20. After import into the stromal space, the stromal targeting peptide is cleaved from the imported protein by SPP. Proteins imported into the thylakoid space are synthesized with bipartite signal sequences. Following removal of the stroma transit peptide by SPP, the protein is transported by either of two pathways. The Sec pathway requires SecA, SecY, and SecE. The Tat (Δ pH) pathway involves the protein components TatA, Hcf106 (TatB homologue), and TatC and a transmembrane pH gradient. The inserted thylakoid protein is then processed by TPP.

bacterial counterparts. The ER signal peptides are cleaved from the exported protein after export into the ER lumen by the signal peptidase complex (SPC).

Protein targeting to the mitochondria often requires a mitochondrial matrix targeting peptide that directs proteins to the interior of the mitochondria (Figure 1C). The mitochondrial matrix signal is cleaved off the exported protein by the mitochondrial processing peptidase (MPP). Sorting of proteins to the intramitochondrial space (IMS) is often accomplished by a bipartite signal peptide. The first part of the signal is the matrix signal, processed by MPP; the second part is the intramitochondrial sorting signal, which is similar to the bacterial signal peptide. Processing of the intramitochondrial sorting signal is mediated by the mitochondrial SPase Imp1/Imp2. In this import process, the matrix-destined protein passes across the outer and inner membranes via the translocase in the outer mitochondrial membrane (TOM) and the translocase in the inner mitochondrial membrane (TIM), respectively. Most proteins destined to the intramitochondrial space follow the stop transfer pathway although a few are transported by the conservative pathway (see the Figure 1C caption for further details).

Protein import across the chloroplast envelope, comprised of the outer and inner membranes, typically is mediated by a stroma targeting peptide (Figure 1D), and is imported through the translocase in the outer chloroplast membrane (TOC) and the translocase in the inner chloroplast membrane (TIC). After delivery of the protein to the stroma, the signal peptide is cleaved from the imported protein by the stroma processing peptidase (SPP). Sorting of proteins to the thylakoid lumenal compartment requires yet another signal sequence, in addition to the stroma targeting peptide. As seen in proteins targeted to the intramitochondrial space, these thylakoid proteins are synthesized with a bipartite signal. The thylakoid substrates can be imported by either the Sec or Tat pathway. The second sequence immediately following the stroma targeting peptide is similar to the bacterial and ER cleavable signal sequences and is processed by a thylakoid processing peptidase (TPP).

B. Signal Peptides—Structure/Function

The majority of signal peptides play an indispensable role in directing proteins to their correct cellular compartment. It is not surprising that the structural features of these peptides vary as widely as the locations to which they are exported. The specificity of the signal peptide is typically achieved through its interaction with a receptor in the target membrane.

Before describing the targeting signals in eukaryotes, we will first describe signal peptides in bacteria. The four types of bacterial signals include the archetypical signal peptides, lipoprotein signal peptides, Tat signal peptides, and type IV signal peptides.

1. Bacterial Signal Peptides

Geneticists studying bacterial protein export have played a very significant role in advancing the protein export field. They were the first to show that the information for protein export is localized within the

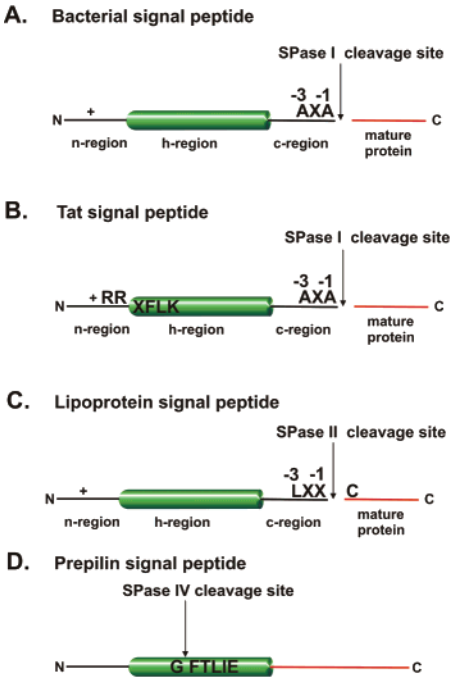


Figure 2. Targeting signal sequences in bacteria. (A) Bacterial signal peptides. The bacterial signal peptides target the protein to the Sec pathway. (B) Tat signal peptide. The Tat signal peptide with the twin arginine consensus motif targets proteins to the Tat pathway. (C) Lipoprotein signal peptide. These signal peptides contain the lipobox sequence and target proteins to the Sec pathway. (D) Prepilin-like signal peptide. These signal peptides are found in type IV piliproteins in certain Gram-negative bacteria as well as in certain prepilin-like proteins involved in type II secretion.

signal peptide. This finding supported Blobel's contention that exported proteins have intrinsic signals that direct proteins to the correct membranes. While signal peptides generally do not have sequence homology, they do have three conserved domains (Figure 2A): a positively charged amino-terminal domain (n), a central hydrophobic domain (h), and a carboxy-terminal hydrophilic domain (c) containing the SPase processing determinants. The discovery that the sorting information within an exported protein is localized within the signal peptide came with the utilization of protein fusions between the exported protein and β-galactosidase. In these studies, Beckwith and colleagues made fusions of the amino-terminal region of the maltose binding protein (encoded by the *malE* gene) with β-galactosidase (encoded by the *lacZ* gene) (for a review see ref 4). Cells bearing these hybrid proteins were maltose-sensitive because the expression of the hybrid protein by the addition of maltose resulted in cell death due to jamming of the secretion machinery by the fusion protein. The fusion proteins were also lac$^-$ because the galactosidase reporter group could not assemble correctly in the cytoplasm when localized at the membrane. Revertants were isolated within the fusion protein by selecting cells possessing the ability to grow in the presence of lactose (maltose resistant). Interestingly, almost all the mutations were located within the MalE portion of the fusion protein, specifically within the signal peptide. Almost all the revertant mutations introduced charged residues into

the hydrophobic core of the signal peptide or resulted in deletions within the hydrophobic region.[5] When these mutations were then introduced back into the wild-type maltose binding protein, the export of the protein was impaired. These results underscore the importance that signal peptide hydrophobicity plays in export. However, the positively charged region of the signal peptide is important for export typically only if the hydrophobicity of the signal sequence is first disrupted.[6] One mechanism proposed to explain the ability of amino-terminal positively charged residues in the signal peptide to promote export is their electrostatic interaction with negatively charged phospholipid headgroups.[7] The hydrophobic core and the C-terminal region of the signal peptide are believed to span the membrane as an α-helix,[8] allowing the cleavage site region to be on the periplasmic side of the membrane. The C-terminal region immediately before the cleavage site has some conserved features with small, uncharged residues at the −1 and −3 positions relative to the cleavage site (or P1 and P3 positions in Schechter and Berger nomenclature[9]). The signal peptide allows the exported protein to interact with the Sec machinery: Signal pepides bind to SecA, the translocation ATPase.[10,11] Signal peptides are also able to interact directly with the lipid bilayer.[12]

There are a number of computational methods for identifying bacterial signal peptides and their cleavage sites from the sequence of proteins. One of the most popular and accessible is the SignalP World Wide Web server (http:// www.cbs.dtu.dk/services/SignalP/), which is based on a combination of several artificial neural networks.[13,14] The signal peptides from Gram-positive bacteria are significantly longer than those from other organisms and have a much longer h-region.[15] The average eukaryotic signal peptide is 22.6 amino acids in length, the average Gram-negative signal peptide is 25.1 amino acids in length, and the average Gram-positive signal peptide contains 32.0 amino acids.[13,14] The signal peptides from Gram-positive bacteria in general have an n-region containing more lysine and arginine residues than that seen in the signal peptides from Gram-negative bacteria or the eukaryotic ER.[16]

2. Tat Signal Peptides

The Tat signal peptides almost always contain two consecutive arginine residues preceding the hydrophobic domain (Figure 2B).[17] Site-directed mutagenesis studies proved that the twin arginines are critical for export.[18] Some of the Tat signal peptides also contain Sec-avoidance sequences. These usually take the form of a positively charged residue in the c-region of the signal peptide that prevents export by the Sec pathway.[19] The c-region also contains the small aliphatic residues at the −1 and −3 positions, allowing the exported protein to be bound and processed by type I SPase.

3. Lipoprotein Signal Peptides

These signal peptides, like the signal peptides of secretory proteins, have a positively charged amino-terminal n-region, a hydrophobic core (h-region), and a carboxyl-terminal specificity c-region (Figure 2C). In addition, the c-region contains the lipobox that has information for lipid modification at the +1 cysteine residue.[20] After lipid modification of the +1 cysteine of the preprotein, the signal peptide is cleaved from the preprotein by SPase II.

4. Prepilin Signal Peptides

The type IV prepilin signal peptides get their name because they were discovered within type IV pillins.[21] They are also found in prepilin-like precursors of proteins that form part of the type II secretion system.[22−25] They are characterized by a short basic region without any hydrophobic domain and are processed by a type IV prepilin SPase (Figure 2D). Precursors containing type IV signal peptides are processed on the amino-terminal side of the hydrophobic region within the mature protein, in contrast to proteins containing bacterial signal peptides, Tat signal peptides, or lipoprotein signal peptides.

5. ER Signal Sequences

The signal peptides that sort proteins to different locations within the eukaryotic cell have to be distinct because these cells contain many different membranous and aqueous compartments. Proteins that are targeted to the ER often contain cleavable signal sequences (Figure 3A). These sequences are similar to bacterial signal peptides in that they have the n-, h-, and c-domains discussed above. Also, like their bacterial counterparts, the hydrophobic character is important for function as a translocation signal.[26] Amazingly, many artificial peptides can function as translocation signals, as demonstrated by Kaiser et al.[27] The most important key feature is hydrophobicity above a certain threshold. ER signal peptides have a higher content of leucine residues relative to bacterial signal peptides.[28] The signal recognition particle (SRP) binds to cleavable signal peptides after they emerge from the ribosome.[29,30] The SRP is required for targeting the nascent protein to the ER membrane. After translocation of the protein to the ER lumen, the exported protein is processed by the SPC.

While many proteins are targeted by the signal peptide to the SRP route, some proteins are targeted by an SRP-independent route, at least in yeast, involving the Sec62p−Sec63p complex.[29] There are subtle features within the signal sequence that determine the pathway used during export of the preprotein across the membrane.[31]

6. Mitochondrial Matrix Targeting Signal

Mitochondrial matrix targeting signals are rich in arginine, serine, and alanine residues, and lack aspartic acid and glutamic acid residues.[32] The mitochondrial matrix targeting peptides are amino-terminal and are often cleaved by MPP (Figure 3B). The peptides have a tendency to form amphiphilic helices,[33,34] which is essential for their function.[35] The mitochondrial sequences function in several ways. First, they are recognized directly by surface receptors,[36] allowing the preprotein to be imported. Second, they interact with acidic patches in proteins as they

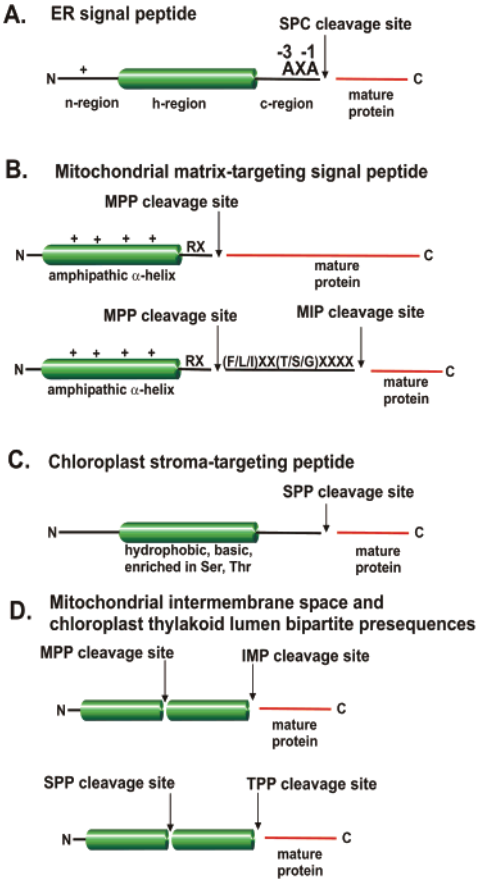


Figure 3. Targeting signal sequences in eukaryotes. (A) ER signal sequences. The signal sequence targets proteins to the Sec61 translocon. (B) Mitochondrial matrix targeting signals. These sequences are recognized by the TOM machinery and are imported across the outer and inner membranes by the TOM−TIM machinery. They are processed either in a one-step process by MPP or by a two-step process first cleaved at the first site by MPP, followed by cleavage at the second site by MIP. (C) Chloroplast stroma targeting peptide. These peptides are recognized first by receptors of the TOC complex and then imported across the envelope through the TOC−TIC complex. SPP removes the stroma targeting peptide from the preprotein to generate the mature chloroplast stroma protein. (D) Mitochondrial intermembrane space and chloroplast thylakoid lumen bipartite signal sequences. The N-terminal portion of the mitochondrial bipartite signal sequence looks like a matrix targeting peptide. The intermembrane space signal resembles the bacterial signal sequence. These bipartite signals get cleaved by both MPP and Imp1p/Imp2p. The amino-terminal part of the chloroplast bipartite signal looks like a chloroplast stroma signal sequence. The C-terminal portion resembles a bacterial-like signal sequence. These presequences are cleaved twice, first by SPP in the stroma space and then by a TPP after translocation of the protein into the chloroplast thylakoid lumen.

are being imported.[37−39] Third, a membrane electrical potential exists across the mitochondrial inner membrane, with the matrix side being negatively charged, which acts directly on the positively charged matrix targeting peptide, thereby actively promoting translocation of the preprotein into the mitochondria.[40,41]

Cleavage of the matrix targeting peptide occurs by either one- or two-step processing. About two-thirds of the preproteins containing matrix targeting peptides are processed by the MPP one-step processing. Cleaved matrix peptides typically have an arginine at either the −2 or −3 position, although the arginine requirement is not absolute.[42] In the two-step processing scheme, maturation in the first step occurs via MPP with an arginine residue typically at the −10 position. Processing at the second step is catalyzed by mitochondrial intermediate peptidase (MIP), eight residues downstream from the first cleavage site.[43,44]

7. Stroma Targeting Peptides

Import into the chloroplast stromal space requires a chloroplast targeting peptide (Figure 3C). These targeting signals are rich in serine and threonine and possess no aspartic acid or glutamic acid residues.[32] Unlike the mitochondrial matrix targeting signals, there is no evidence that the chloroplast targeting peptides can form amphipathic α-helices. Cleavage of these sequences is carried out by an SPP.

8. Mitochondrial Intermembrane Space and Thylakoid Lumenal Space Sorting Signals

Reminiscent of bacterial signal peptides, these sorting sequences have a positively charged n-region and a central hydrophobic h-region with SPase processing determinants localized in the c-region at the −1 and −3 positions (Figure 3D).[45,46] These sorting sequences are preceded by a matrix targeting peptide in mitochondrial intermembrane space proteins or a stroma targeting peptide in chloroplast thylakoid luminal proteins. Mitochondrial intermembrane space proteins are processed by the mitochondrial SPase that has two catalytic subunits (Imp1/Imp2) with different substrate specificities. Imp2 cleavage follows the normal Ala-X-Ala processing rules, while Imp1 does not and cleaves after the asparagine residue at the −1 position.

The chloroplast thylakoid lumenal sorting signal directs proteins that are first imported into the stromal space across the thylakoid membrane. The sorting sequence from the thylakoid lumenal protein is cleaved by the TPP. TPP has a specificity very similar to those of the other type I SPases. The thylakoid sorting sequence can direct proteins across the membrane using either the Sec or Tat (or Δ pH) pathways.

II. Substrates

The natural substrates for SPases are preproteins that are exported within the cell (see Tables 1 and 2). Although not all exported proteins contain cleavable signal peptides, most of them do occur in a precursor form. This includes substrates for type I SPases which process the majority of outer membrane proteins and periplasmic proteins in Gram-negative bacteria and secreted proteins in Gram-positive bacteria (Table 1). The ER type I SPase (SPC) processes many proteins that enter the ER (Table 2) for either retention in the ER or for export to the Golgi apparatus, secretory vesicles, plasma membrane, or vacuole/lysosome. In contrast to plasma membrane proteins in bacteria, which are almost never substrates for type I SPase, many plasma membrane proteins in eukaryotic cells as well as

Table 1. Substrates for Bacterial Signal Peptidases

preprotein substrates used for enzymatic assays	analysis method	refs
	SPase I	
pre-MBP	in vivo and in vitro	64, 75
pro-OmpA	in vivo	64
pre-β-lactamase	in vivo	300
procoat	in vivo and in vitro	64, 69, 103, 301
pro-OmpA nuclease A	in vitro; $k_{cat} = 8.73\ s^{-1}$; $K_m = 16\ \mu M$	76
cytochrome *b* preprotein	in vitro; $k_{cat} = 11\ s^{-1}$; $K_m = 50\ \mu M$	77
	SPase II	
prolipoprotein	in vivo and in vitro	175, 177, 302
pre-PrsA	in vivo	52, 181
	type IV prepilin SPase	
pre-PilA	in vitro	303
pre-XcpT	in vitro	189
pre-TcpA	in vitro	192
PulG	in vivo	304

Table 2. Substrates for Eukaryotic Signal Peptidases

preprotein substrates used for enzymatic assays	analysis method	refs
	ER SPC	
preprolactin	in vitro	71
prepro-α-factor	in vitro	305, 306
human preplacental lactogen (pre-HPL)	in vitro	72
pre-invertase and pre-acid phosphatase	in vivo	59
pre-acid phosphatase	in vivo	59
	MPP	
ornithine carbamoyltransferase precursor	in vitro	216
α-subunit of F1ATPase	in vitro	73
β-subunit of F1ATPase	in vitro and in vivo	63, 73, 74
subunit V cyt *c* oxidase	in vitro	73
citrate synthase preprotein	in vitro and in vivo	73, 225
cytochrome c_1 (two steps)	in vitro	217
Rieske Fe/S protein (two steps)	in vitro	217
cytochrome b_2	in vivo	62, 63, 225
cytochrome *c* oxidase subunit IV	in vivo	307
cytochrome *c* oxidase subunit 3	in vivo	225
Fe/S protein	in vivo	62
	MIP	
human ornithine transcarbamylase (pOTC)	in vitro	308
rat malate dehydrogenase (pMDH)	in vitro	308
neurospora ubiquinol−cytochrome *c* reductase iron sulfur subunit (pFe/S)	in vitro	309
cytochrome *c* oxidase subunit IV	in vitro	251
	SPP	
RBCS (precursor to Ribisco small subunit)	in vitro	266
preprotein to acyl carrier protein	in vitro	257
LHCP precursor	in vitro and in vivo	66, 266
oxygen evolving protein preprotein	in vitro	46
heat shock protein 21 preprotein	in vitro	266, 310
Ribisco activase precursor	in vitro and in vivo	66, 266, 311
plastocyanin precursor	in vitro	266
ferredoxin precursor	in vitro	266, 312

membrane proteins within organelles of the secretory pathway are processed by the ER SPC.

Very recently, there have been some surprises regarding which proteins are substrates for these peptidases. There is evidence that the ER SPC is involved in the processing of the polyproteins of the hepatitis C virus,[47] rubella virus,[48] hantavirus,[49] and foamy virus[50] as well as, in some cases, misfolded membrane proteins.[51]

The type II SPase processes precursors of lipoproteins that are exported across the plasma membrane (Table 1). There are a large number of putative lipoprotein substrates. For example, there are 94 and 114 putative substrates in *E. coli* and *Bacillus subtilis*, respectively.[52,53] Even the bacterium with the smallest known genome has a number of lipoproteins that are predicted to be processed by SPase II.[54]

Type IV prepilin SPase substrates are less common. These substrates include the type IV prepilins, as suggested by the name of the peptidase, as well as prepilin-like proteins that are found to be components of the type II secretion machinery (Table 1).

A large number of MPP and SPP preprotein substrates are imported into the mitochondrial matrix and chloroplast stroma (Table 2), respectively. As mentioned in section I, some of the matrix-localized mitochondrial substrates are cleaved by MIP, after first being cleaved by MPP. Precursor proteins destined to the mitochondrial intermembrane space are often substrates for Imp1p/Imp2p.[55]

Table 3. Kinetic Constants for SPases Measured with Various Peptide Substrates

peptide substrate	method of analysis	k_{cat} (s^{-1})	K_m (M^{-1})	k_{cat}/K_m (M^{-1} s^{-1})	ref
		E. coli SPase I			
9-mer, MBP	HPLC	3.2×10^{-2}	8×10^{-4}	40	78
5-mer, MBP	HPLC	2×10^{-4}	3.7×10^{-3}	0.054	78
11-mer, MBP	FRET-continuous			71	80
26-mer, MBP	FRET-continuous	1.5	0.6×10^{-6}	2500000	313
19-mer	FRET-continuous	0.032	118.1×10^{-6}	270	82
		ER SPC			
31-mer, ppPTH	HPLC	($V_{max} = 1.2$ nM/s)	3.2×10^{-6}		314
		MPP			
30-mer, MDH	HPLC	($V_{max} = 0.13$ pmol/min)	0.73×10^{-6}		212
24-mer, MDH	FRET-continuous			35000000	211

Many thylakoid proteins containing bipartite signals are substrates for TPP.[56]

III. Signal Peptidase Activity Assays

In vivo SPase activity is commonly monitored in intact cells using temperature-sensitive (ts) mutants where the SPase protease is inactivated at the nonpermissive temperature. Preproteins cleaved by the SPase under normal conditions accumulate at the nonpermissive temperature. Temperature-sensitive mutants have been isolated within the *E. coli*[57] and *B. subtilis*[58] SPases. In *Saccharomyces cerevisiae*, ts mutants have been isolated in the Sec11 subunit of the signal peptidase complex[59] and the mitochondrial Imp1 subunit of the mitochondrial SPase.[60] In addition, temperature-sensitive mutants have also been isolated in SPase II[61] and in both subunits of the MPP.[62,63]

Other approaches to measure in vivo activity include controlling the enzymatic activity by constructing a strain which regulates the expression of the SPase[64,65] or controlling activity with specific inhibitors of the protease. For instance, an arabinose inducible SPase I strain was constructed by integrating into the chromosome of a PolA$^-$ *E. coli* mutant, a plasmid encoding an SPase I deletion mutant where its expression is under the control of the araBAD promoter.[64] Additionally, an antisense cDNA method was used to decrease the expression level of the stromal processing enzyme in plants, resulting in the accumulation of precursors to several exported proteins.[66] Using penem inhibitors, Black and co-workers at SmithKline Beecham Pharmaceuticals confirmed that SPase I is the general SPase for processing nonlipoproteins in bacteria.[67] Similarly, the compound globomycin inhibits SPase II, resulting in the accumulation of lipoprotein precursors.[68]

In vitro studies often employ radioactively labeled preproteins that are synthesized using a transcription/translation system.[69−74] Typically, the translation system is an *E. coli* S30 cellular extract, a reticulocyte lysate, or a wheat germ system. To examine signal peptide cleavage of a preprotein, SPase is added to the translation system and the conversion of the radioactive precursor form to the mature form is monitored by SDS−PAGE and autoradiography.

Another useful cell-free strategy is to test whether SPase can cleave chemical amounts of a preprotein substrate. The *E. coli* SPase I, for example, cleaves preMBP,[75] proOmpA nuclease A,[76] and cytochrome b5 preprotein.[77] The precursor and mature forms are separated by SDS−PAGE and visualized by Coomassie brilliant-blue staining. While this preprotein assay is labor intensive since it involves running numerous gels, it has been used to measure the kinetic constants of SPase I cleavage using the proOmpA nuclease A[76] and the cytochrome b_5 preprotein substrates.[77]

A powerful in vitro assay where the kinetic constants of signal peptide cleavage can be more easily measured involves the use of synthetic peptide substrates (Table 3). Here, HPLC is often employed to separate the reactant and product peptides after various times of incubation with the SPase (for example, see ref 78). Some peptide substrates can be used to monitor cleavage using a continuous assay (Table 3). In this assay, intramolecularly quenched fluorescence substrates are employed, where the fluorescent donor is on one side of the peptide and is separated from the acceptor molecule on the other side via the cleavage site. Therefore, substrate cleavage prevents the acceptor molecule from receiving energy from the fluorescent donor via fluorescence energy transfer, resulting in an increase in the measurable fluorescent intensity. This assay is ideal, as it can be used continuously to monitor substrate cleavage using a fluorometer.[79−82]

IV. Type I Signal Peptidase (SPase I)

This family of membrane-bound SPases includes type I SPases in Gram-positive and Gram-negative bacteria,[83,84] the ER,[85] and the mitochondria and chloroplast.[86] We will begin with the type I SPases from bacteria, which have been the most thoroughly studied.

A. Bacterial SPase I

The bacterial type I SPases were the first signal peptidases to be isolated, cloned,[87] sequenced,[88] and characterized.[76,89−92] SPase was purified initially by extracting the isolated membranes with the detergent Triton X-100, followed by ion exchange and chromatofocusing chromatography.[69] The isolation of larger amounts of the SPase was made possible by

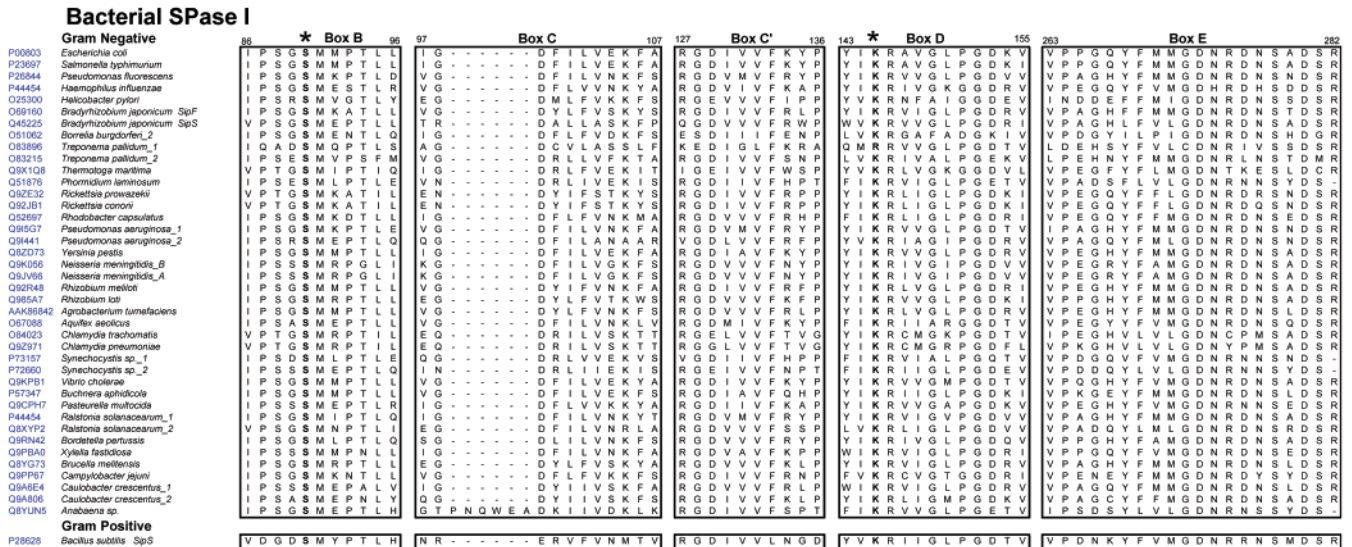


Figure 4. Alignment of the regions of the conserved sequence in bacterial SPase I enzymes. The alignment was generated with the program ClustalW[294] using amino acid sequneces resulting from a BLAST search[295] in the SWISS-PROT database.[296] The Swiss-Prot accesssion numbers are listed to the left of each sequence. The P-type SPases are grouped as described by Tjalsma et al.[297] and Dalbey et al.[137] Residue numbers are shown in the alignment using the *E. coli* SPase nomenclature. The catalytic Ser 90 and Lys 145 residues (*E. coli*) are indicated by bold type and an asterisk. The consensus sequence is shown below each box.

the cloning of SPase under the control of powerful promoters.[64,93] The *E. coli* SPase is now rapidly purified by nickel chelate affinity chromatography using a polyhistidine-tagged form of the protein. The 6His-tagged *E. coli* SPase is active.[91]

Type I SPases have now been cloned and sequenced from a wide variety of bacteria, and many have been isolated and chacterized in pure form. Most bacteria have a single chromosomal copy of the SPase gene; however, there are exceptions. In *B. subtilis*, there are five chromosomal copies of SPase and two plasmid copies.[94,95] The bacterium *Streptomyces lividans* has four copies.[96] *Staphylococcus aureus* has two chromosomal copies of the gene, one of which is expected to encode an inactive enzyme as the catalytic residues are missing.[97]

Figure 4 shows an alignment of the conserved regions within the amino acid sequences of the bacterial SPase I. The consensus sequence is read with the upper case character meaning absolutely conserved, lower case meaning conserved, and x meaning not conserved. The first conserved region (box) of the sequence, A (not shown in Figure 4), is part of the hydrophobic segment that anchors the catalytic domain to the membrane. Region B contains the motif ipsxsMxptlx. This second serine serves as the nucleophile in the catalytic mechanism. The C region contains the dxlxvxkxx consensus in which the conserved leucine residue is part of the substrate binding pocket (S3) that binds the −3 residue of the preprotein substrate. Region C′ contains the motif rgdivvfxxp in which the second valine is part of the S3 pocket. Region D has conserved residues, yikRxxgxpgdtv, with the tyrosine residue forming part of the S1 substrate binding pocket followed by the general base lysine residue and the invariant arginine residue that, in some bacteria, possibly plays a structural role. Finally, region E has many conserved residues, vpxgxyfxmgdnrxnsxdsr. It is thought that the conserved first serine residue (Ser 278 in the *E. coli* enzyme) helps to position the lysine general base toward the nucleophilic serine.

In all the bacteria analyzed so far, SPase I has been shown to be essential for cell viability. Disrupting the *E. coli* lep gene[98] or putting the gene under the control of an inducible promoter[64] proved that SPase is essential. Similar studies have confirmed the essential function of SPase in the human pathogenic bacteria *Streptococcus pneumonia*[99] and *S. aureus*.[97] In the case of *B. subtilis*, both SipS and SipT must be inactivated to cause death.[43,58]

1. Substrate Specificity

The substrate specificity of the *E. coli* SPase is similar to that of the ER SPC. This similarity is demonstrated by the abilities of the *E. coli* enzyme to cleave preproinsulin and the ER SPC to process the M13 procoat protein.[100] This overlap in specificity revealed that both signal peptides and SPase specificity are conserved in evolution.

In bacteria and eukaryotic cells, preproteins have a common pattern in the c-region of the signal peptide at the −1 and −3 positions. This led Gunnar von Heijne to propose that the region preceding the cleavage site constitutes the substrate recognition site for the SPase enzyme.[101] The importance of the residues at the −1 and −3 positions was confirmed by site-directed mutagenesis studies. Residues at the −1 position that were tolerated are alanine, glycine, serine, cysteine, and, in some cases, threonine. Residues tolerated at the −3 position are alanine, glycine, serine, cysteine, isoleucine, valine, and leucine. These results came from studies with a precursor to the maltose binding protein,[102] the M13 procoat protein,[103] and alkaline phosphatase.[104] Analyzing the frequency in which each residue is found at these positions revealed a preference for alanine, or an Ala-X-Ala cleavage motif.[101] Residues at the +1, −2, −4, and −5 positions were not generally critical, as almost any residue was tolerated, except for a proline at the +1 position, which impaired processing.[105,106] Alternatively, with the M13 procoat protein, a proline at the −6 position is very important for SPase processing.[103] Requiring a helix breaker for processing agrees with the observation that, in bacteria, a helix breaker is quite common at the h/c-region border.[107]

Many signal peptides have potential alternative SPase cleavage sites adjacent to the normal −1 and −3 sites. This was shown in studies with preMBP where the residue at the −1 position was mutated to remove the normal processing site.[102] The surprising result was that mutants containing all 20 amino acids at the −1 position were processed. However, mutants with residues other than alanine, glycine, serine, cysteine, and threonine at the −1 position were processed at a new cleavage site two residues upstream.

Potential alternative SPase cleavage sites raise the following question: What is the source of the high fidelity for cleaving the normal site? One possibility is that the fidelity comes from the transmembrane region of SPase that anchors the catalytic domain of the enzyme so it is positioned appropriately to cleave the membrane-anchored substrate at the correct site. Amazingly, constructs of the *E. coli* and *B. subtilis* SPase I which lack the transmembrane regions still process preprotein substrates with high fidelity at the correct site.[108] This shows that the source of the fidelity lies within the C-terminal catalytic domain of SPase.

Proper processing of signal peptides requires the cleavage sites to be located in close proximity to the h/c-region border. Kendall and co-workers discovered this while studying model signal peptides of alkaline phosphatase where the length of the c-region was varied by introducing different numbers of asparagine residues between the cleavage site (Ala-Gln-Ala) and the end of the h-region.[109] Processing occurred in all cases when the introduced Asn residues increased the c-region from 3 to 9 residues in length. However, processing did not occur when Asn residues had lengthened the c-region to 11 and 13 residues. The implication of these studies is that the active site of SPase is positioned at the membrane surface in the periplasmic space and that the cleavage site needs to be positioned such that it is accessible to the active site of SPase. When the length of the c-region exceeds 10 residues, the cleavage site is likely too far from the membrane surface for processing to occur.

2. Structure of the Catalytic Domain

To obtain information about how SPase binds and cleaves its substrates, it was necessary to determine the three-dimensional structure of the catalytic domain. To accomplish this, we made a truncated SPase following the strategy employed by the group at Merck. Knight and co-workers[110] isolated a truncated SPase, Δ2−75, that was missing the amino-terminal membrane-anchoring region (residues 2−75). The Δ2−75 construct proved to be active against peptide substrates[110] and exhibited with the preprotein substrate pro-OmpA-nuclease a k_{cat} of 3 s^{-1} and a K_m of 32 μM.[111] Therefore, the truncated enzyme has a specificity constant roughly 1/20th of that of the wild-type enzyme, which contains the membrane-anchoring region.

The Δ2−75 protein has been crystallized in several space groups. The initial tetragonal crystals ($P4_22_12$) typically diffracted to 3.6 Å resolution and were often mosaic.[112,113] Subtle changes to the crystallization conditions resulted in a new tetragonal crystal form in a different space group ($P4_12_12$) that diffracted to about 2.4 Å resolution.[114] In contrast, an orthorhombic crystal form (space group $P2_12_12$) was considerably more ordered, and the structure was solved to 1.9 Å resolution using multiple isomorphous replacement methods. The structure was solved with an irreversibly covalently bound (5*S*)-penem inhibitor, generously provided by SmithKline Beecham Pharmaceuticals (see Figure 6A for the structure of the inhibitor), attached to the active site serine nucleophile.[92]

The structure (Figure 5) revealed that the proteolytic region of SPase I has a novel protein fold, comprised of two large antiparalled β-sheet domains. One of the β-sheet domains contains all of the conserved regions in the SPase I family (shown by the colored regions); the other β-sheet domain (shown in gray), which varies in size with the different SPases, does not contain any of the conserved sequence motifs in the SPase family. The catalytic domain also contains an extended β-ribbon that protrudes from the conserved β-sheet domain. There is a disulfide bond between Cys 170 and Cys 176 in the nonconserved β-sheet domain.

The three-dimensional structure of Δ2−75 proved that the essential Ser 90 residue was indeed the nucleophilic residue involved in catalysis because it

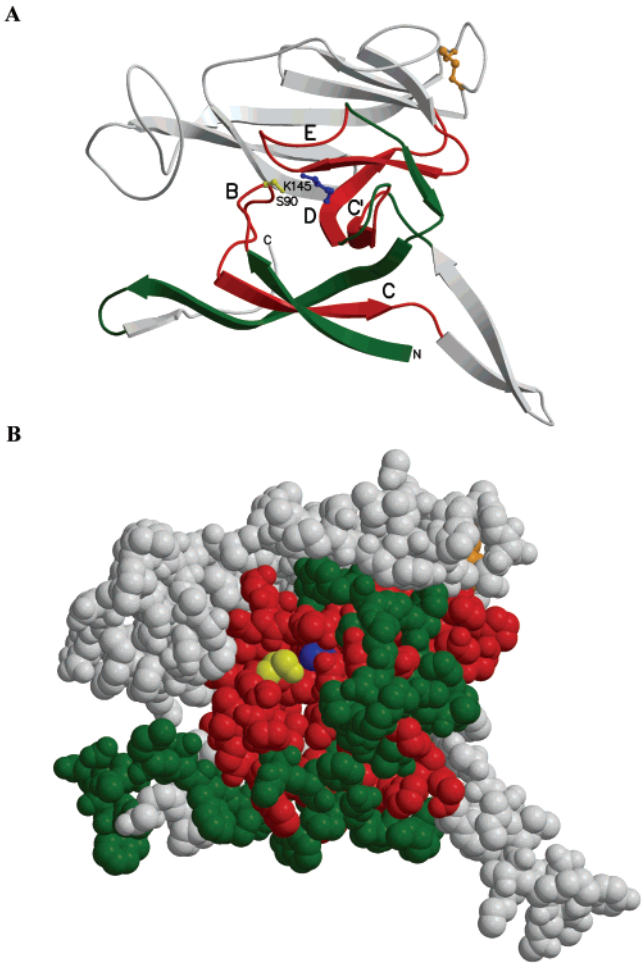


Figure 5. Structure of the catalytic domain of *E. coli* SPase I (Δ2−75). (A) Ribbon diagram of the β-sheet domain with all the conserved regions within the SPase I family shown in color. Nonconserved regions are shown in gray. The active site serine (S90) is shown in yellow, and the active site lysine (K145) is shown in dark blue. The cysteines that form a disulfide bond are shown in orange. (B) van der Waals rendering of Δ2−75. This figure was prepared using the programs MOLSCRIPT[298] and RASTER3D.[299]

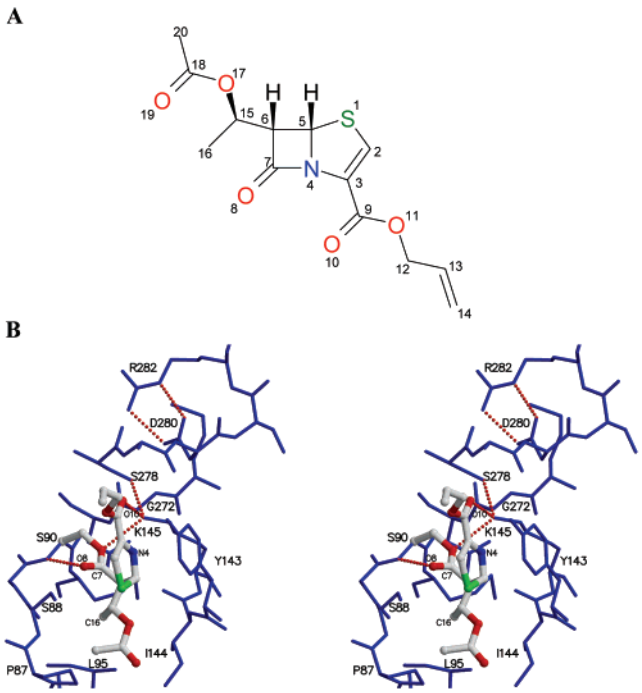


Figure 6. Structure of the active site of *E. coli* SPase I with bound β-lactam-type inhibitor. (A) Structure of the signal peptidase inhibitor allyl (5*S*,6*S*)-6-[(*R*)-acetoxyethyl]-penem-3-carboxylate. (B) The crystal structure revealed the Ser 90 O^{γ} covalently bound to the (5*S*)-penem inhibitor. The Ser 90 O^{γ} was within hydrogen-bonding distance to the ε-amino group of lysine 145. Ser 278 hydrogen bonds to the lysine 145 ε-amino group. The methyl group (C16) of the inhibitor is within the S1 substrate binding pocket. The carbonyl oxygen group (O8) of the cleaved β-lactam ring of the inhibitor is sitting in an oxyanion hole formed by the main chain NH of Ser 90. This figure was prepared using the programs MOLSCRIPT[298] and RASTER3D.[299] Oxygen atoms are shown in red, nitrogen atoms are in blue, and sulfur atoms are in green. The carbon atoms of the inhibitor are drawn in gray. All protein atoms in (B) are shown in blue. The hydrogen bonds are in red.

was covalently attached to the cleaved (5*S*)-penem (Figure 6B). Also, the serine O^{γ} atom was within hydrogen-bonding distance (2.9 Å) to Lys 145, the proposed general base. Lys 145 is partially buried (see Figure 5B), an environment that may be important for its involvement in catalysis. The structure also revealed that Ser 278 forms a hydrogen bond to Lys 145 (Figure 6B). The absolutely conserved glycine at position 272 lies directly adjacent to the Ser-Lys dyad. Any other residue than glycine at this position would most likely interfere with the catalytic machinery. The invariant Asp 280 forms a salt bridge with the conserved Arg 282 and is within hydrogen-bonding distance to Ser 278. This may stabilize the active site region and helps to position the Ser 278 with respect to the Lys 145.[115]

A striking feature of the catalytic domain structure is its extremely large exposed hydrophobic surface. This hydrophobic surface may promote the binding of the catalytic domain to the membrane. Indeed, Δ2−75 binds to lipid vesicles and inserts into phospholipid monolayers.[116] A significant portion of the protease domain is inserted into the lipid phase of the membrane.[117] Interestingly, preprotein processing is stimulated by the addition of detergent or phospholipid to Δ2−75,[111] and the addition of detergent was necessary for optimal crystallization of the protein.[113]

Recently the structure of Δ2−75 was solved in the absence of inhibitor or substrate (apo-enzyme)[114] (Figure 7A). The structure was solved to 2.4 Å resolution and revealed some significant differences in the binding site region when compared to the Δ2−75 inhibitor acyl-enzyme structure. This is seen nicely in Figure 7B where the structures of key amino acids near the active site are compared in the inhibitor acyl-enzyme structure (shown in black) and the apo-enzyme structure (shown in red; the inhibitor is not shown for clarity). First, the Ser 88 O^{γ} side chain is rotated significantly from the penem-bound structure into a position where it could participate in the formation of the oxyanion hole. Second, a water molecule can be observed near the active site region of the apo-enzyme (Figure 7A). This water molecule, which was displaced by the inhibitor, is most likely the deacylating water molecule due to its position (distance and angle) relative to the general base and

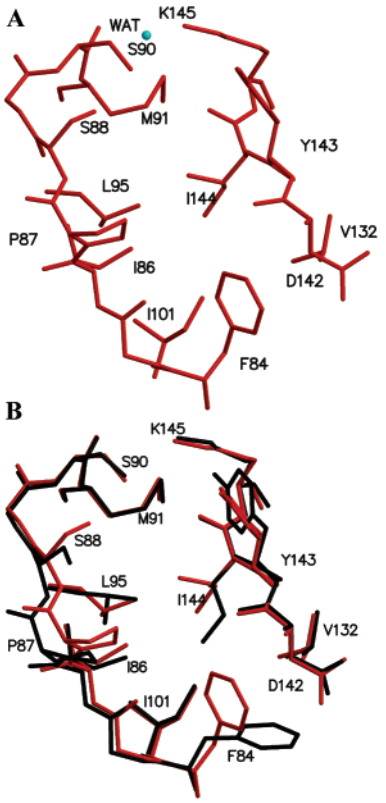


Figure 7. Structural comparison of the active site of *E. coli* SPase I in the absence and presence of bound inhibitor. (A) Residues in the binding pocket of the apo-enzyme of *E. coli* SPase I. The proposed deacylating catalytic water (WAT) is shown in cyan and labeled. (B) Residues in the binding pocket of the apo-enzyme (red) superimposed on the acyl-enzyme binding pocket (black). For clarity the inhibitor in the acyl-enzyme structure is not shown. This figure was prepared using the programs MOLSCRIPT[298] and RASTER3D.[299]

the ester carbonyl carbon of the acyl-enzyme. Third, the S1 binding pocket that binds the −1 (P1) residue of the signal peptide was smaller than that observed in the inhibitor acyl-enzyme structure. Fourth, the S3 pocket that binds the −3 signal peptide residue is partially blocked by the Phe 84 side chain in the apo-enzyme structure as compared to the (5*S*)-penem structure.[114] Interestingly, the structure of SPase shares the same core protein fold as UmuD,[118,119] LexA,[120] and λcI,[121] which are members of the LexA family of proteins that employ a similar serine-lysine dyad mechanism for self-proteolysis. This suggests that these proteases evolved from a common ancestor.

3. Proteolytic Mechanism

Before the structure of the catalytic domain of *E. coli* SPase was solved, mutagenesis studies were performed on the *E. coli* and *B. subtilis* SPases I to explore the catalytic mechanism. With the *E. coli* SPase, Ser 90 was shown to be important for catalysis as mutation to an alanine residue led to an inactive enzyme.[89] Similarly, Lys 145 was shown to be important for activity since mutation to an alanine led to a severe drop in activity.[90,122] No histidines or cysteines are important for enzymatic activity.[89] These studies supported the idea that SPase I carries out catalysis using a Ser-Lys catalytic dyad mechanism (for a review on Ser-Lys dyad proteases and amidases see ref 123).

Protein modification studies, coupled with site-directed cysteine mutants, also provided insightful information about the serine and lysine residues involved in catalysis. Since *E. coli* SPase was fully active when the three cysteine residues were replaced with serine, the putative serine nucleophile was replaced with a cysteine residue. This thio-SPase was enzymatically active. However, the addition of the cysteine-specific reagent *N*-ethylmaleimide (NEM) inactivated the enzyme, while similar treatment of the wild-type SPase had no effect.[90] This result strongly suggested that serine 90 was involved in catalysis.

To explore the role of Lys 145, a similar approach involving cysteine mutants was employed.[91] Mutating the lysine at position 145 to a cysteine residue resulted in an inactive enzyme. However, modification of the Cys 145 with bromoethylamine resulted in a lysine analogue at this position and restored activity to 1% of that of the wild-type enzyme. This γ-thialysine residue has sulfur instead of a methylene group at the γ position. Interestingly, modification of the thio-145 SPase with bromopropylamine also resulted in a significant restoration of activity (0.7% of that of the wild type) by producing the γ-thia-homolysine analogue. While these studies did not provide information about why Lys 145 was important for activity, they did show that lysine analogues could function at the 145 position. The Lys 145 residue could be important for SPase function in either of two ways; it could provide a positive charge possibly for electrostatic interactions with another amino acid side chain within SPase, or it could function as a general base. If the lysine was important in providing a positive charge, then treatment of the thio-SPase with the nontitratable positively charged analogue (2-bromoethyl)trimethylammonium bromide may result in the restoration of activity. However, there was no gain in activity after the cysteine was modified with this reagent. This result is consistent with Lys 145 acting as general base in catalysis.

The above results supported the idea that the catalytic mechanism involved a serine and a lysine residue, and the crystal structure provided direct evidence that Ser 90 was the nucleophile because Ser 90 formed a covalent adduct with the (5*S*)-penem inhibitor. Likewise, the fact that the Ser 90 O^{γ} was within hydrogen-bonding distance to Lys 145 Nζ bolstered the idea that this lysine functions as the general base. In addition, there are no other titratable groups near the vicinity of Ser 90. We suspect that this lysine residue has an unusual pK_a of 8.7, roughly 2 pH units below that of lysine in water. This apparent pK_a value for the lysine was assigned by analyzing pH rate profiles of *E. coli* SPase I. Such profiles displayed a narrow bell-shaped curve with a rising inflection point at pH 8.7 and a desending inflection point at pH 9.3.[91] Optimal activity was observed at pH 9.0. From the crystal structure, one can see that the drop in pK_a for the general base

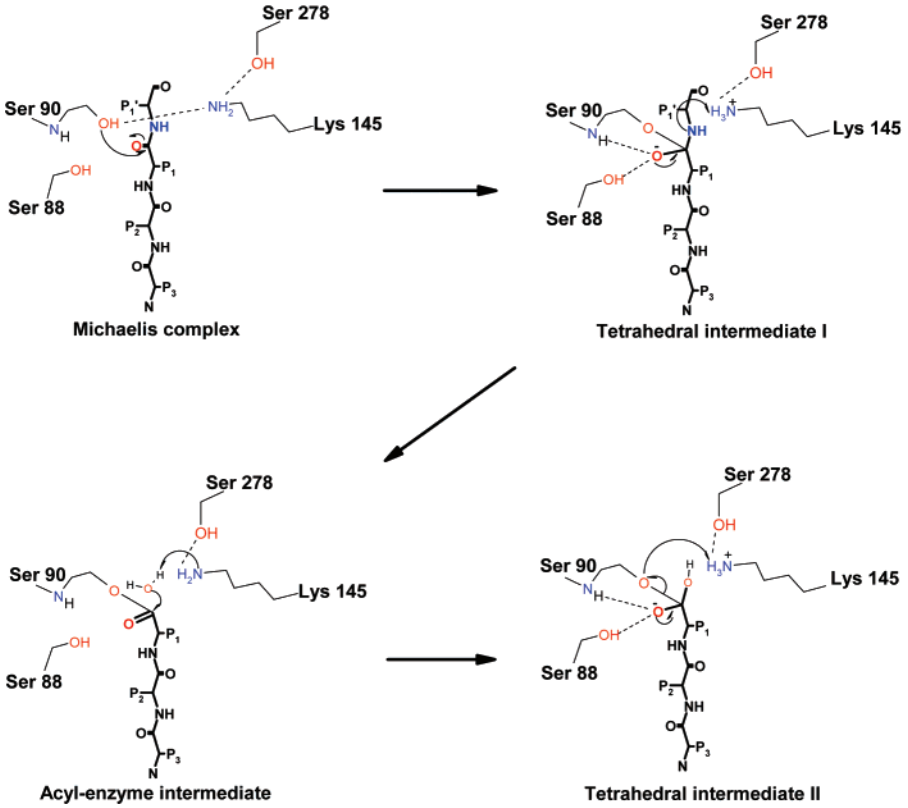


Figure 8. Proposed mechanism for type I signal peptidase. The reaction proceeds through an oxyanion tetrahedral intermediate followed by loss of the mature domain of the substrate, giving rise to an acyl-signal peptide enzyme intermediate. The acyl-enzyme intermediate is attacked by a water, resulting in the release of the cleaved signal peptide.

lysine would result from its burial within the active site upon binding of the substrate. Lys 145 is completely buried in the inhibitor-bound acyl-enzyme structure.[92]

Further details of the catalytic mechanism came from studies exploring the role of conserved amino acids in the vicinity of the active site region.[115] Ser 278 and Gly 272 play important roles since individual mutations of both residues to alanine led to losses of activity of 300-fold and 750-fold, respectively. The structural reasons for the importance of these residues can be seen from the crystal structure of *E. coli* SPase I and are discussed above (Figure 6B). In addition, mutation of the absolutely conserved Arg 146 to alanine resulted in no loss of activity, suggesting that the salt bridge between Arg 146 and Asp 273 is not important for activity. Curiously, mutagenesis studies showed the invariant Asp 280 was not absolutely required for SPase activity.

More recent studies evaluated the catalytic role of Ser 88 in *E. coli* SPase I due to the proposal that this residue may be involved in the formation of the oxyanion hole.[92] Site-directed mutagenesis of Ser 88 to alanine or cysteine led to a large decrease (740−2440-fold, respectively) in k_{cat} with little effect on K_m.[124] This is consistent with Ser 88 participating in the chemical step of the reaction by stabilizing the transition state. The S88T substitution was almost as active as the wild-type enzyme. Modeling studies showed that this substitution could also function in transition-state stabilization. The structure of the SPase apo-enzyme confirmed that Ser 88 was positioned appropriately to function in stabilizing the transition state.[114] A serine at this position is conserved in most SPases in Gram-negative bacteria, whereas aligned sequences in Gram-positive bacteria generally show glycine at this position. Possibly, in the case of glycine, the main chain amide hydrogen participates in transition-state stabilization as is found in the analogous enzyme LexA.[120]

Similar site-directed mutagenesis experiments support a Ser-Lys dyad mechanism in the *B. subtilis* SPase SipS.[125] Van Dijl and co-workers found that Ser 43 (homologous to Ser 90 of the *E. coli* SPase) and Lys 83 (homologous to Lys 145 in *E. coli* SPase) were critical for activity. Likewise, Asp 153 (homologous to Asp 280) was also found to be critical for activity. This is unlike the situation with the *E. coli* enzyme.[115] Therefore, with *B. subtilis* SipS, the absolutely critical residues are Ser 43, Lys 83, and Asp 153. However, in contrast to the *E. coli* SPase, mutations of the homologous residues at Arg 146 and Asp 273 in *B. subtilis* resulted in a decrease in activity most likely due to the protein being structurally impaired.[58] This possible structural role is consistent with the structure of the *E. coli* Δ2−75 where Arg 146 forms a salt bridge with Asp 273.

All the structure/function results to date support the following mechanism of proteolysis for the bacterial type I SPases (Figure 8): Upon binding of the preprotein substrate, the neutral Lys 145 ϵ-amino group, oriented by the O^{γ} of Ser 278, abstracts a proton from the Ser 90 side chain hydroxyl, activating the Ser 90 O^{γ} for attack on the carbonyl of the scissile bond. A tetrahedral intermediate forms which is stabilized by the Ser 88 side chain hydroxyl hydrogen

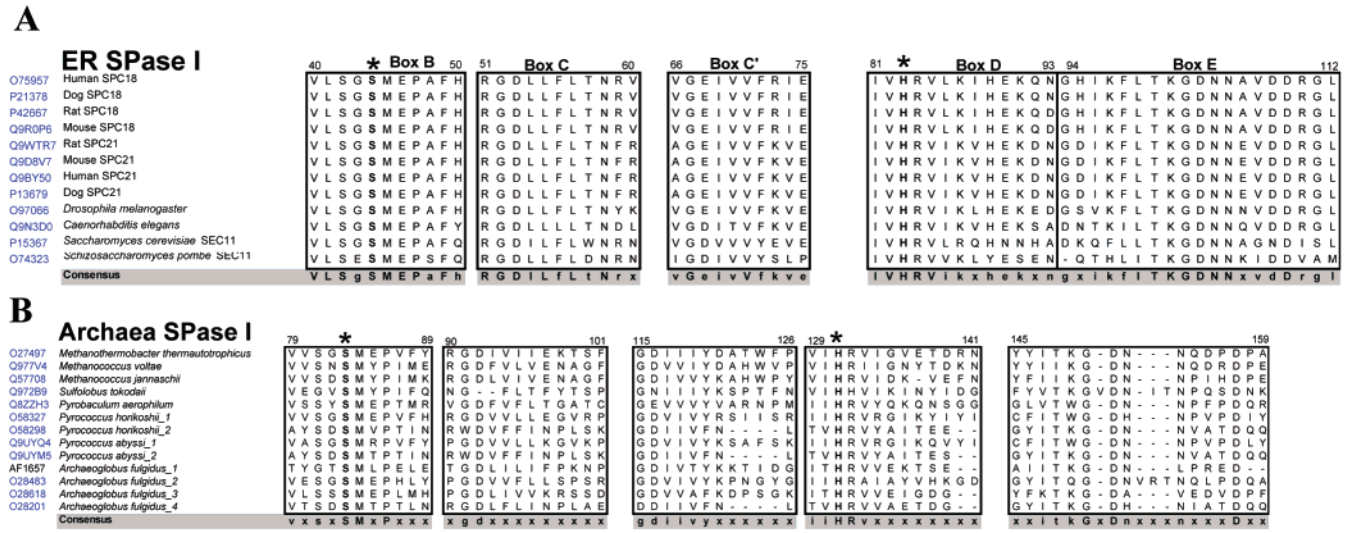


Figure 9. Sequence alignment of type I ER/archaea signal peptidases. (A) Boxes of conserved residues in the homologous SPase subunits from the endoplasmic reticulum of eukaryotic species. The catalytic Ser 44 and His 83 residues (*S. cerevisiae* numbering) are indicated by bold type and with an asterisk. (B) Boxes of conserved residues in type I SPases from archaea bacteria. The numbering is based on the *Methanothermobacter thermautotrophicus* sequence. The alignment was generated with the program ClustalW[294] using amino acid sequences resulting from a BLAST search[295] in the SWISS-PROT database.[296] The Swiss-Prot accesssion numbers are listed to the left of each sequence.

and the main chain amide hydrogen of Ser 90.[92] The lysine then donates its proton to the amide nitrogen of the mature domain (leaving group), and the tetrahedral intermediate breaks down to produce the signal peptide bound acyl-enzyme. The deacylating water then loses its proton to the lysine general base, and the hydroxyl group attacks the ester carbonyl to produce another tetrahedral intermediate. This then breaks down to release the signal peptide and the free enzyme active site.

4. Antibacterial Target

Since type I SPases are essential enzymes for bacterial cell growth and viability, and are located at the membrane surface, these enzymes are considered possible targets for antibacterial compounds. Type I SPases are present in all bacteria except for *Mycoplasma*.[54] The bacterial type I SPases are believed to be monomeric and differ significantly from eukaryotic counterparts, which are multimeric and, in the case of the ER signal peptidase, do not use lysine as a general base. Therefore, it should be possible to isolate or synthesize inhibitors that can selectively inhibit the bacterial enzyme without inhibiting to a great extent the ER signal peptidase or the mitochondrial SPase (Imp1p/Imp2p).

The type I SPases in bacteria and the ER are unusual in that they are resistant to common protease inhibitors against the four protease classes.[110,126,127] Nevertheless, it is possible to synthesize inhibitors. Certain types of β-lactam compounds were found to inhibit SPase.[67,128−130] Of these, the most potent types are (5*S*)-penem compounds. This is the opposite stereochemistry at the 5 position as seen in β-lactam-type antibiotics such as penicillin. The compound allyl (5*S*,6*S*)-6-[(*R*)-acetoxyethyl]-penem-3-carboxylate inhibits the *E. coli* SPase with an IC_{50} of 50 μM.

More effective inhibitors were isolated from natural products by high-throughput screening (K. O'Dwyer, B. Bankosky, M. Black, G. Bruton, F. Fan, R. Macarron, L. McCloskey, B. Orlek, and L. Wang, personal communication). These compounds are lipopeptides, and the most effective one inhibits the *E. coli* SPase with an IC_{50} of 30 nM and the *S. aureus* SPase with a IC_{50} of 1.2 μM. These compounds inhibit processing in vivo and have bactericidal activity against *S. aureus* and *Streptococcus pyogenes*. This work demonstrates that SPase is a possible target for medicinal intervention.

B. ER SPase I

In 1986 it became evident that the ER signal peptidase enzyme was more complex than the bacterial enzyme. The purified enzyme from dog pancreas microsomes contained five polypeptides, which varied in molecular weight from 12000 to 25000.[131] The subunits are termed SPC12, SPC18, SPC21, SPC22/23, and SPC25. Baker and Lively, in 1987, purified the chicken ER signal peptidase, which, surprisingly, consisted of only two polypeptides. These subunits, 23 and 19 kDa in size, are called gp23 and p19.[132] In a recent report, however, Lively stated that the chicken SPC contained homologues of SPC18, SPC21, SPC22/23, and SPC25 subunits.[85] The yeast SPase also contained four subunits: 10.8 (SPC1), 18.8 (Sec11), 21.3 (SPC3) and 25 (SPC2)[133,134] kDa in size.

The next significant advance was the cloning of the genes for the mammalian ER SPC. SPC22/23 is a glycoprotein subunit. The canine SPC18[135] and SPC21,[136] which are homologous to each other, were cloned by the Blobel laboratory. Similar genes were cloned for subunits of the SPC from other species (see ref 137).

Analysis of the cloned genes showed that one of the subunits was weakly homologous to the bacterial enzyme (Figure 9). The yeast SPC component Sec11, cloned in 1988, had so little sequence homology it was not recognized as related until later when other

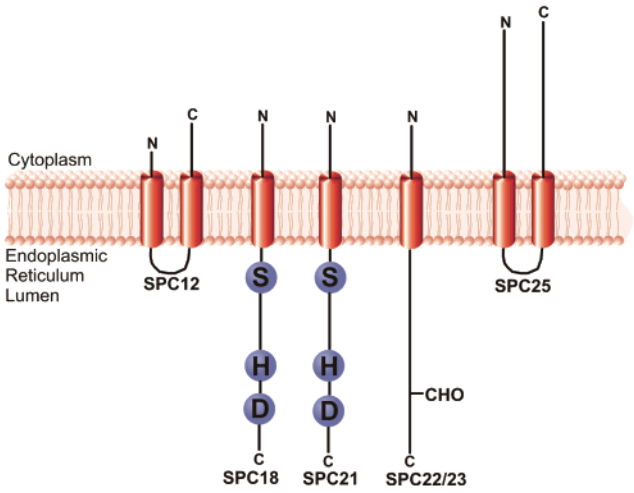


Figure 10. Membrane topologies of SPC subunits. The ER signal peptidase complex consists of five polypeptides, two of them (SPC21 and SPC18) are weakly homologous to the bacterial, chloroplast, and mitochondrial type I SPases. The active site serine is located near the membrane surface on the lumenal side of the membrane.

homologous genes from other organisms were cloned.[138] The Sec11 polypeptide is essential for the SPC function and cell viability in *S. cerevisiae*.[59] Sec11 possesses a serine residue that is homologous to Ser 90 in the *E. coli* SPase (located in box B). However, the general base lysine is missing, and a histidine is found in its place (box D). Human, dog, rat, and mouse have two subunits, termed SPC18 and SPC21, which are highly homologous to Sec11.

The 21 kDa glycoprotein subunit SPC3 is absolutely essential for the function of SPase in *S. cerevisiae*.[134,139] The corresponding protein sequences from dog,[140] chicken,[141] *Caenorhabditis elegans*,[134] and *Schizosaccharomyces pombe* vary in size from 180 to 185 residues and show 40.1% identical residues.[137]

The other subunits of SPC are not essential for SPase activity in yeast.[142] This raises questions regarding the functions of the various additional subunits. Some possibilities are that these subunits in *S. cerevisiae* interact with the translocon complex, they stabilize the SPC, or they are involved in the retention of the complex within the ER. Indeed, it is known that the SPC 25 subunit of the mammalian ER SPC and the yeast SPC 25 homologue interact with the Sec translocon.[143,144]

The membrane topology of the subunits of the dog pancreatic SPC complex is shown in Figure 10. SPC18, SPC21, and SPC 22/23 span the membrane once with their C-terminal domain in the ER lumen.[145] The SPC25 and SPC12 span the membrane twice with their N- and C-termini facing the cytoplasm.[146] The serine, histidine, and aspartic acid residues important for catalysis in SPC18 and SPC21 are located in the ER lumenal domains.

1. Catalytic Mechanism

To define the protease class of the ER SPC, Van Valkenburgh et al. used *S. cerevisiae* to test the effect of making mutations within the Sec11 gene.[147] Mutagenesis studies were performed by changing the conserved residues in the SPase family as well as the histidine (see Figure 9) that appears at the same position as the critical lysine residue in bacteria. Ser 44 (box B), His 83 (box D), and Asp 103 and Asp 109 (box E) were found to be absolutely essential for preprotein processing in vivo. Green and co-workers also mutated all the lysine residues in the Sec11 protein that are conserved in the eukaryotic Sec11 family and in the essential glycosubunit SPC3 and discovered none were essential for processing.[147] From this analysis, they concluded that the ER SPC carried out catalysis either by a Ser-His dyad, analogous to the Ser-Lys dyad of bacterial SPases, or by a Ser-His-Asp triad like other classical serine proteases.

In addition to the ER-type SPases, archaeal SPases have a histidine residue at the position of the critical lysine residue in the bacterial enzyme in box D (Figure 9B). They also contain the conserved serine in box B, an invariant aspartic acid immediately after the absolutely conserved glycine in box E, as well as a second conserved aspartic acid. In certain Gram-positive bacteria, there are SPase genes (for example, SipW in *B. subtilis*; see Figure 4) that contain an ER-like SPase, with the lysine replaced by a histidine. Van Dijl et al. showed that the histidine residue is essential for catalysis.[148] Unlike the yeast SPC subunit Sec11, where the histidine to lysine substitution is inactive, in SipW this mutant is active. Therefore, it can tolerate both residues and remain active.

2. Antiviral Target

Recently, a number of viral polyproteins such as those found in the hepatitis C virus,[47,149,150] hantavirus,[49] flavivirus,[151] rubella virus,[48] and influenza C virus[152,153] were found to contain internal signal peptides that are most likely cleaved by the ER SPC. Mutagenesis of the predicted signal peptidase substrate specificity elements may thus block viral infectivity. The idea of using SPC as a viral target is not new (see refs 47 and 150). Chemists developing successful antiviral drugs will need to make potent inhibitors against the ER SPC. The design of inhibitors would benefit from detailed studies of the substrate specificity and structural studies of the ER SPC. It should be noted that although the ER SPC may be an effective antiviral target, inhibitors of this SPase are likely to cause severe side effects.

These studies on the maturation of polyproteins are also very interesting because they show that SPC can cleave not only amino-terminally located signal peptides, but also after internal signal peptides. Consistent with this is the study by Green and co-workers where mutant membrane proteins were processed by SPC after internal signal peptides.[51] These studies, together, raise the possibility that there are still many other unidentified substrates for SPase that could be processed by SPC, possibly in a regulated fashion.

C. Organellar SPase I

Discovery of a protease in the mitochondria of *S. cerevisiae* that processed proteins destined to the

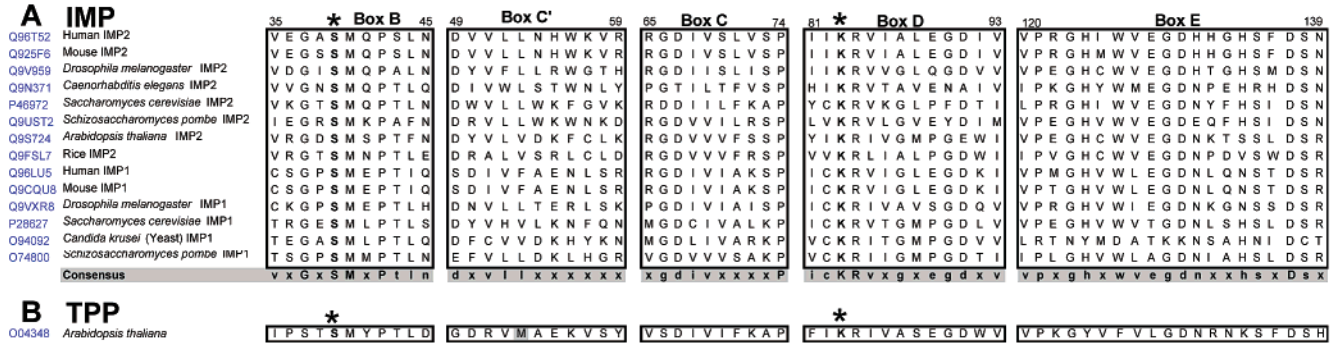


Figure 11. Alignment of conserved regions in organellar type I signal peptidases. (A) Regions of conserved sequence in known mitochondrial Imp's. (B) Sequence of the conserved region of TPP from *Arabidopsis thaliana*. The alignment was generated with the program ClustalW[294] using amino acid sequences resulting from a BLAST search[295] in the SWISS-PROT database.[296] The Swiss-Prot accesssion numbers are listed to the left of each sequence. The residues are numbered acording to the *S. cerevisiae* Imp1 sequence. The nucleophilic serine and general base lysine are indicated with an asterisk.

intermembrane space provided the first clue that the mitochondrial type I SPases may function by a mechanism similar to that of the bacterial SPases.[60] This was discovered when Pratje and co-workers isolated a temperature-sensitive *S. cerevisiae* mutant which showed delayed processing of the cytochrome oxidase subunit II precursor and the intermediate cytochrome b_2.[154,155] They cloned the gene, called Imp1, by complementation of the ts Imp1 mutant at the nonpermissive temperature and discovered that the gene resembled the *E. coli* SPase (see Figure 11). Imp1 is part of a complex[156] and forms a dimer with Imp2.[157] At about the same time, Imp2p was cloned by the Walter laboratory and shown to be highly homologous to Imp1p[158] (Figure 11). Expression of Imp2p was necessary to produce a stable Imp1. There are two SPase subunits that have different substrate specificities. Imp2 has the normal substrate specificity following the canonical −1, −3 processing rules. Imp1 does not follow these rules and cleaves after an asparagine residue at the −1 position. Imp1 and Imp2 are believed to be single transmembrane protein subunits with the active site regions located in the mitochondrial intermembrane space. Recently, it has been shown there is a third component of the Imp complex called Som1.[159−161] Cross-linking and co-immunoprecipitation assays have demonstrated that the Imp1 and Som1 proteins physically interact. It appears that Som1 assists in processing of some Imp1 substrates and helps in stabilizing Imp1.

The chloroplast type I SPase TPP found in the thylakoid membrane is also homologous to the bacterial enzyme and seems to function as a single subunit protein.[162] Sequence gazing among the different species with genes that encode the chloroplast SPase (Figure 11) revealed that, like Imp1 and Imp2, the genes contain the canonical serine and lysine residues in addition to other typical conserved residues found in most bacteria. TPP is predicted to span the thylakoid membrane once, with its catalytic center residing in the thylakoid lumen.

1. Catalytic Mechanism

Catalysis by Imp1p/Imp2p is also carried out using a serine-lysine dyad catalytic apparatus. The first evidence for a critical serine came from studies with Imp2 where mutation of the putative serine nucleophile to an alanine residue inactivated Imp2.[158] Next, a detailed study was performed on both the Imp1p and Imp2p subunits of the mitochondrial SPase complex (Imp1p/Imp2p) in *S. cerevisiae*.[163] Since the type I SPase signature in eubacteria consists of a serine, a lysine, an arginine, and two aspartic acid residues,[163] the analogous residues were mutated. The serine in box B, the lysine in box D, and the second aspartic acid in box E were essential for Imp1 processing of cytochrome b_2. The arginine residue adjacent to the lysine was not critical for activity. Similar results showing an absolute requirement for the serine, lysine, and second aspartic acid were found with Imp2.

The data suggesting that the mechanism of the chloroplast TPP is similar to the bacterial and Imp1p/Imp2p complex is that it contains the conserved SPase signature residues with the serine-lysine dyad.[162] TPP is also inhibited by (5*S*)-penem as is the *E. coli* SPase.[164] As mentioned in the bacterial type I SPase section, the (5*S*)-penem reacts with the active site serine of the *E. coli* SPase. Therefore, TPP most likely belongs to the serine-lysine dyad group of peptidases.

2. Substrate Specificity

The Imp1 subunit of the mitochondrial type I SPase is unusual because it does not follow the −1, −3 rules as do all other type I SPases. Imp1p cleaves substrates with an Asn residue at the −1 position, whereas previous studies on the *E. coli* SPase have shown that SPase cannot cleave substrates with an Asn at the −1 position.[102,103] This is probably because the S1 pocket in the Imp1 peptidase is different from that found in other SPases.

To examine the substrate specificity in more detail, Chen et al. mutated the −1 Asn residue in the Imp1 substrate cytochrome b_2 and studied whether it was still cleaved by Imp1.[163] Surprisingly, Imp1 efficiently processed cytochrome b_2 substrates with alanine, serine, cysteine, leucine, and methionine at the −1 position. The other substrates with different −1 residues were cleaved inefficiently or not all. Thus, there is not an absolute requirement for an Asn residue immediately preceding the cleavage site.

Surprisingly, none of the substrates were recognized by Imp2, even when the −1 residue was predicted to coincide with Imp2's substrate specificity. Clearly, there must be other determinants within the substrate, in addition to the −1 residue, which are recognized by Imp1.

Chloroplast TPP and *E. coli* SPase possess very similar substrate specificity.[165] TPP specificity follows the well-known −1, −3 rule and, in fact, may even more strictly prefer a −1 alanine than most type I SPases. Sequence comparisons of chloroplast thylakoid preproteins with bipartite signal sequences show that an alanine is more frequently found at the −1 and −3 positions. Mutagenesis studies on the pre-33K protein showed that there is a marked preference for an alanine at the −1 position.[166] In processing the pre-33K substrate protein, TPP tolerates a −3 valine instead of an alanine for efficient processing, just like the *E. coli* enzyme. In this instance, however, the leucine substitution significantly inhibited processing.

V. Type II Signal Peptidase (SPase II)

Like the SPase I family, SPase II members are membrane-bound proteases that have their active site exposed to the *trans* side of the membrane so they can cleave exported proteins after translocation across the membrane (for a recent review see ref 53). They are absent in eukaryotic cells and are believed to belong to an unusual class of aspartic acid proteases.

A. Cloning, Purification, and Topological Characterization of SPase II

In Gram-positive and Gram-negative bacteria, SPase II plays a key role in processing lipoproteins. The precursors of these proteins have the lipobox with the consensus sequence Leu-Ala-Gly-Cys. The first lipoprotein discovered was Braun's lipoprotein,[167] which is a major protein in the outer membrane in *E. coli*. The second lipoprotein discovered was the peptidoglycan-associated protein,[168] which is, as the name implies, associated with the peptidoglycan. Adding globomycin, which is a specific inhibitor of SPase II, to the cell results in the accumulation of glyceride-containing precursor forms of the outer membrane lipoproteins in the cytoplasmic membranes of *E. coli*.

The gene encoding SPase II was first cloned in *E. coli* at about the same time by both the Wu laboratory and the Mizushima laboratory. They achieved this either by complementation of a temperature-sensitive SPase II strain[169] or by examining clones for increased globomycin resistance.[170] The gene (*lsp*) was sequenced and found to encode a protein of 164 amino acids.[171,172]

The activity of SPase II can be monitored by the removal of the signal peptide of ^{35}S-labeled diacylglyceryl-modified prolipoprotein, which is isolated by immunoprecipitation from cells treated with globomycin.[173,174] In addition, a rapid and quantitative assay was developed to assay for SPase II where the released signal peptide is recovered by acetone extraction after cleavage.[175] SPase II was purified 35000-fold as follows:[175] membrane isolation, Triton X-100 extraction, heating, DEAE−cellulose chromatography, and chromatofocusing.

SPase II has a pH optimum at 6.0[175] and is inhibited by globomycin in a noncompetitive manner with a K_i of 36 nM.[176] In vitro studies show that nonionic detergents such as Triton X-100 or Nikkol are required for activity.[177] Wu and colleagues, in their studies with SPase II solubilized from membranes, found that the optimal pH for activity was higher than that reported above (pH 7.9) with an optimal temperature from 37 to 45 °C. The optimal pH for glyceryl transferase activity was 7.8 with an optimal temperature of 37 °C.[177] SPase II is not inhibited by any of the conventional protease inhibitors at levels normally needed to inhibit their respective protease class. The membrane topology of the *E. coli* SPase II is predicted to span the membrane four times on the basis of the positive inside rule.[178] Ultimately, Wu and colleagues[179] determined the membrane topology using the gene fusion approach.[180]

B. Conserved Domains in the SPase II Family

Type II SPases are identifiable in a wide variety of organisms, including Gram-negative and Gram-positive bacteria, and *Mycoplasma* (Figure 12). They are not found in archaea or eukaryotes. While some organisms such as *B. subtilis* possess multiple copies of the SPase I gene, all sequenced genomes to date contain only one gene encoding SPase II and all the proteins are predicted to span the membrane four times (Figure 13A). There are five highly conserved domains in these SPases (Figure 12). In the consensus sequence, upper case represents strictly conserved while lower case represents mostly conserved residues. In the first conserved region (A), present within the first transmembrane segment, is the dxxtk motif. Region B, containing the motif NxGaaf, is found in the periplasmic environment in Gram-negative or in cell wall exposed areas of Gram-positive bacteria. Region C contains the motif iiggaxlgNxxDr located partially within transmembrane segment 3. The asparagine 99 and aspartic acid 102 in box C of *B. subtilis* SPase II are critical for activity; aspartic acid 102 is proposed to be one of the catalytic residues at the active site.[181] Region D contains the triplet vvd where the aspartic acid is not required for activity in *B. subtilis*. Region E has the consensus sequence FNxAD where the invariant asparagine, alanine, and aspartic acid play important roles in preprotein processing. The aspartic acid is proposed to be an active site residue located on the lumenal side of transmembrane segment 4 (Figure 13A).

C. Type of Protease

SPase II was suspected to be an unusual protease due to its lack of homology to other proteases and because it was not affected by inhibitors against the serine, cysteine, aspartic acid, and metalloenzyme protease classes.[182] To identify potential active site residues,[181] the 15 residues that were strongly or absolutely conserved in the type II SPase were

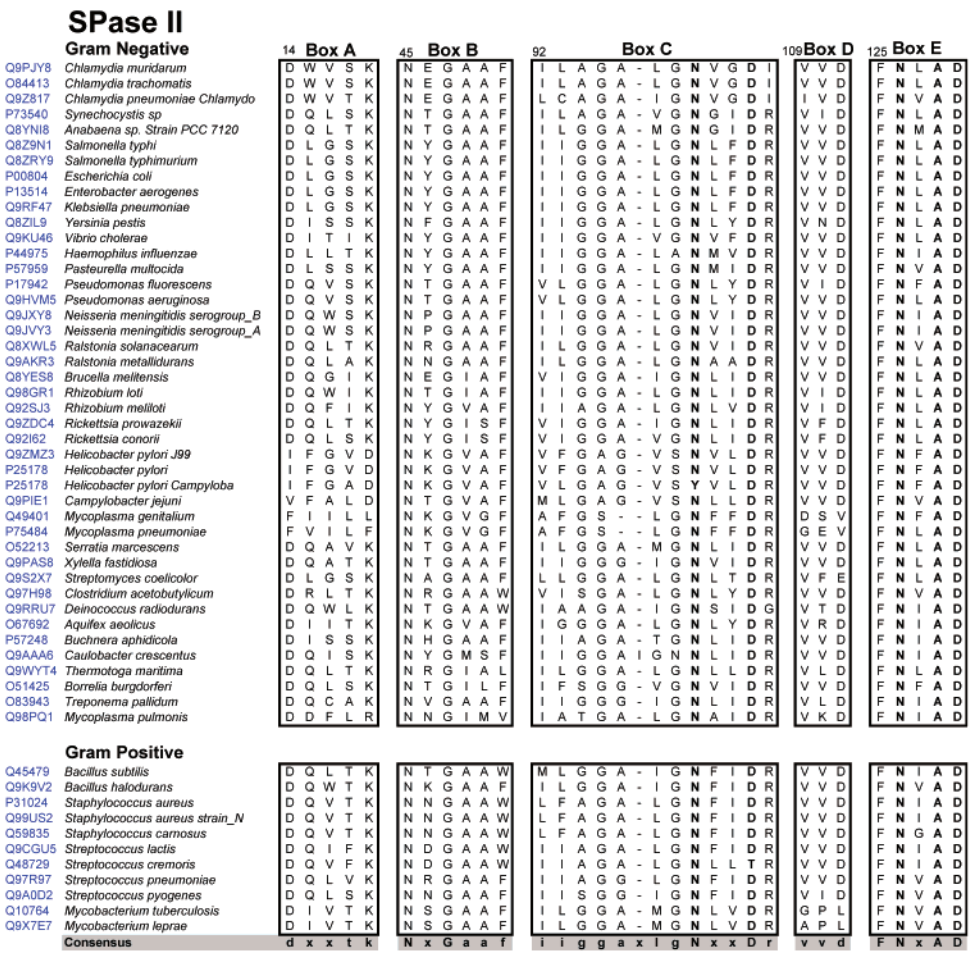


Figure 12. Alignment of bacterial SPase II sequences. Residue numbers are shown in the alignment using the *B. subtilis* SPase II nomenclature. The putative catalytic Asp 102 and Asp 129 residues (*B. subtilis*) are indicated by bold type. The alignment was generated with the program ClustalW[294] using amino acid sequences resulting from a BLAST search[295] in the SWISS-PROT database.[296] The Swiss-Prot accesssion numbers are listed to the left of each sequence.

Table 4. Summary of Mutagenesis Results on the *B. subtilis* SPase II[a]

SPase II mutant	activity	defect	SPase II mutant	activity	defect
D14A	−	structure	R103A	+	
K18A	+		V109A	+	
N45A	+/−		D111A	+/−	
G47A	+		F125A	+	
F50A	+		N126A	−	?
G95A	+		A128V	−	?
N99A	−	?	D129A	−	active site
D102A	−	active site			

[a] From ref 181.

mutated within the *B. subtilis* SPase II. These studies identified Asn 99, Asp 102, Asn 126, Ala 128, and Asp 129 as important residues for SPase II activity (see Table 4). Mutations at these positions completely eliminated proteolytic activity but did not significantly affect SPase II structure on the basis of the observation that the protein was stable within the cell. In addition, Asp 14 is critical for the structure of the SPase II protein, as the D14A mutant protein was misfolded and degraded by proteases within the cell. Further evidence for the aspartate playing an important structural role is that in the *E. coli* SPase II the mutation of the homologous Asp 23 to a glycine residue resulted in a temperature-sensitive enzyme.[183]

Since there are no conserved serine or cysteine residues and no metal requirement for catalysis and the enzyme is inhibited by pepstatin (at high levels), it was proposed that the type II SPase family utilizes an aspartic acid protease mechanism. Most aspartic acid proteases contain the conserved sequence Asp-Thr/Ser-Gly with a conserved Ser or Thr positioned three residues carboxyl-terminal to the aspartic acid. SPase II lacks these features, which led van Dijl and co-workers to propose that SPase II is an unusual aspartic acid protease.

Catalysis is proposed to occur when two aspartic acid residues act as a dyad, promoting proteolysis using general-acid/general-base chemistry. SPase II is believed to contain a catalytic water molecule bound to one of the two aspartic acids in the deprotonated state before it binds substrate. In the first step, the phosphoglyceride-modified preprotein binds to the protease. This results in one of the conserved aspartic acid residues at the *trans* side of the membrane acting as a general base to abstract a proton from a water molecule. This generates a hydroxide, which attacks the scissile peptide bond, generating a tetrahedral intermediate. In the next step, donation of a proton from the aspartic acid to the amino terminus of the mature portion of the lipid-modified preprotein and the donation of a proton from one of

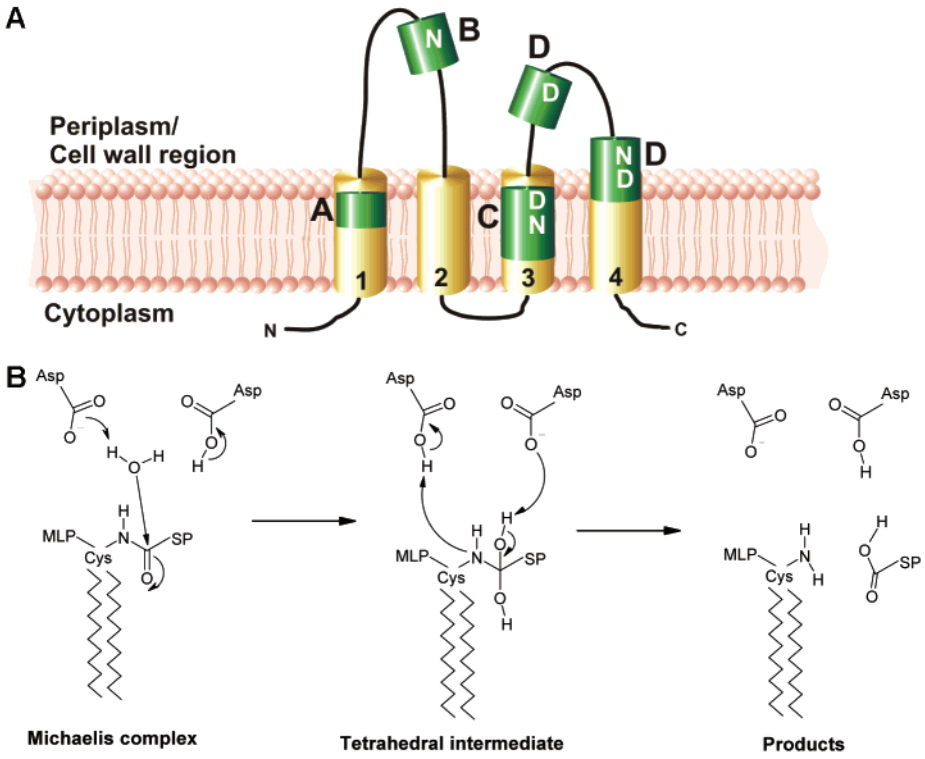


Figure 13. Schematic of the type II signal peptidase structure and proteolytic mechanism. (A) Membrane topology of the SPase II. (B) Proposed mechanism of catalysis of SPase II. SPase II is proposed to contain two active site aspartic acid residues, one of them protonated and the other one deprotonated. Upon binding of substrate, the proton on the water molecule is removed by one of the aspartates, which acts as the general base. The hydroxide then acts as a nucleophile and attacks the carbonyl group of the scissile peptide bond. This produces a tetrahedral intermediate, which then breaks down, and the products are then free to dissociate and produce the free SPase protein. The abbreviation MLP, in the reaction schematic, represents mature lipoprotein; SP represents signal peptide.

the hydroxyl groups of the tetrahedral intermediate to an aspartate result in cleavage of the scissile peptide bond. At this stage both the signal peptide product and the mature lipoprotein product are bound to the SPase II. Dissociation of the signal peptide and the mature protein products from SPase II occurs, and they are replaced by the lytic water. The protease is now ready for another round of catalysis with the two aspartyl residues in the same protonation state as before, prior to forming the Michaelis complex (Figure 13B).

Presently, there is no structural information on SPase II. Visualizing the hydrogen-bonding interactions at the active site at various stages in the catalytic cycle would be exciting. If that is done, one could then determine whether Asp 102 and Asp 129 are located at the active site. The exact roles of the other conserved critical residues are not known at the present time.

D. Substrate Specificity

The hallmark of a lipoprotein signal peptide is the lipobox (in *E. coli*, Leu-Ala-Gly-Cys) located in the C-terminal region of the signal peptide and extending to the +1 position of the mature domain of the protein. This sequence corresponds to the substrate specificity of the glyceryltransferase that modifies the preprotein at the cysteine with the lipid.[184] While the cysteine in the lipobox is absolutely conserved, there can be various amino acids at the other locations. For example, although leucine is found over 65% of the time at the −3 position in *B. subtilis*, *E. coli*, and *Mycoplasma*, other residues are also found. In particular, six other amino acids are found at this position in lipoprotein substrates in *B. subtilis* and *E. coli*. Alanine, serine, and threonine are found at the −2 position in *B. subtilis*, *E. coli*, and *Mycoplasma* over 70% of the time.[53] At the −1 position, alanine and glycine occur over 74% of the time in *B. subtilis* and *E. coli*, whereas in *Mycoplasma*, one observes mainly alanine (80% of the time), with serine and glycine making up 13% and 7% of the −1 residues, respectively.

Because of the requirement for the preprotein to be lipid modified prior to SPase II cleavage, the specificity of this protease is more difficult to decipher. Nevertheless, there have been some studies on this important topic. While the substitution of a valine or leucine at the −1 position led to the accumulation of the unprocessed preprotein, the substitution of a serine had no effect on signal peptide processing.[185] A threonine substitution led to slow lipid modification with no observable signal peptide processing. This suggests that SPase II prefers a small side chain at this position. Interestingly, while amino acid side chains larger than that of serine are rare at the −1 position (less than 3%) in *E. coli*, they are more common in *B. subtilis* (around 20% of the side chains are larger).[53] This suggests that the S1 pocket that binds the −1 residue is larger in *B. subtilis*. In addition to this requirement at the −1 position, the secondary structure at or near the

cleavage site is important for signal peptide processing and lipid modification. There is evidence suggesting that the region just after the cleavage site has a propensity to form a β-turn.[186]

E. Relation to the Type IV Prepilin Signal Peptidase

Like type II SPases, type IV prepilin SPases are integral membrane proteins, not homologous to any known cysteine protease, serine protease, aspartic acid protease, or metalloprotease. The *Pseudomonas aeruginosa* SPase IV spans the membrane eight times with both the amino and carboxy termini in the periplasmic space.[187] SPase IV catalyzes two reactions.[188−190] One reaction cleaves prepilin peptides, and the other methylates the free α-amino group of the newly generated substrate. Signal peptide cleavage is essential for type IV pilus biogenesis and type II protein secretion.[191] However, N-methylation is not required for pilus biogenesis and type II secretion.[191]

La Point and Taylor[192] performed a detailed mutagenesis study examining whether conserved amino acid residues such as cysteine, serine, aspartate, lysine, histidine, and glutamate are important for activity in the type IV prepilin peptidase of *Vibrio cholerae*.[192] Only two aspartate residues are conserved in all the SPase IV enzymes. The two aspartate residues are predicted to be near the cytoplasmic membrane surface where they can participate in the cleavage of prepilin signal sequences. These researchers also showed that the two conserved aspartate residues in the type IV SPase were absolutely critical for activity, and that SPase IV is inhibited by the aspartic acid protease inhibitor 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDAC) glycinamide.[192] All these results are consistent with SPase IV being an aspartic acid protease.

VI. Mitochondrial Processing Peptidase

The existence of MPPs became evident after the discovery that proteins imported into the mitochondria were synthesized in a higher molecular weight precursor form containing a matrix targeting peptide.[193,194] The matrix targeting peptide was proposed to be removed by a protease within the mitochondria. A few years after the preprotein discovery, MPPs were purified and characterized from several sources and shown to be metallopeptidases (there are a number of recent reviews[195−197]).

A. Mitochondrial Matrix Peptides and Cleavage Specificity

Mitochondrial matrix peptides vary in length from 8[198,199] to 58[200] amino acids and are rich in positively charged residues. These sequences alone are sufficient for import and are able to target normally cytosolic proteins into the mitochondrial matrix. These matrix peptides adopt an α-helical conformation in a hydrophobic environment.[35,201−203] A common pattern for substrates processed in one step only by MPP is to have an arginine at the −2 (or −3) position, a small residue (serine or alanine) at the +2 position, and a bulky hydrophobic residue (such as phenylalanine, leucine, isoleucine, or tyrosine) at the +1 position. All of these positions are relative to the MPP cleavage site. The remaining one-third of the substrates are processed by both MPP and MIP (mitochondrial intermediate peptidase), in the two-step process. MPP cleaves the substrate between the −9 and −8 positions (relative to the MIP processing site), most often in the presence of arginine at the −10 position. After MPP cleaves the substrate, MIP cleaves off the eight amino-terminal residues (see Figure 3B).

To determine which portion of the matrix peptide is important for import into the mitochondria and for cleavage, deletion and mutagenesis experiments were conducted. Generally, the N-terminal part of the matrix peptide is important for import.[204−207] The C-terminal part is important for cleavage[205,208,209] and, in some cases, is part of the targeting signal.

The determinants of processing are located primarily at the −2 and +1 positions of the mitochondrial substrate. For example, substitution of arginine at the −2 position in the matrix targeting signal of cytochrome b_2[210] or premalate dehydrogenase (pMDH) matrix peptide abolishes or decreases processing, respectively. Substitution of the +1 aromatic residue in the substrate also inhibited processing of cytochrome b_2[210] and MDH synthetic peptides.[211] In addition to the −2 and +1 positions, an upstream arginine can be important for processing as well.[196,212]

B. Characterization of MPP

The MPP activity has been characterized from many different organisms[74,213−216] and shown to be dependent on a metal ion for activity. Neupert and co-workers were the first to purify the protein to homogeneity.[217] The protein was isolated from *Neurospora* using DEAE−cellulose, zinc chelate, polyethylenimine (PEI)−cellulose, and Ecteola−cellulose chromatography matrix. MPP consisted of two subunits: the processing peptidase and the processing enhancing protein. Isolated in completely pure form, the subunits had no activity. MPP activity requires the presence of both subunits.[210] Soon after, Yang et al. purified the MPP from *S. cerevisiae* and discovered it was also a dimer.[218] The polypeptides are now called α-MPP and β-MPP. In addition, MPP has been purified from rat liver,[219,220] potato,[221] wheat,[222] spinach,[223] and bovine liver.[212] In plants, the MPP is part of the membrane-bound cytochrome bc_1 complex.[221,224]

While the genes of *Neurospora* α-MPP and β-MPP were identified by antibody screening and colony hybridization,[217] the genes encoding the yeast α- and β-subunits were cloned by genetic complementation of temperature-sensitive mutants in their respective subunits.[62,225] The rat and potato MPP genes were identified by polymerase chain reaction (PCR) of cDNA libraries.[220,221,226,227]

The partially purified MPP from *S. cerevisiae* had a pH optimum at 7.5 and was inhibited by metal chelators.[73] However, the activity is restored by adding Zn^{2+}, Co^{2+}, or Mn^{2+}. For the *Neurospora crassa* MPP, the activity is enhanced 3-fold with the addition of manganese and, to a lesser extent, by

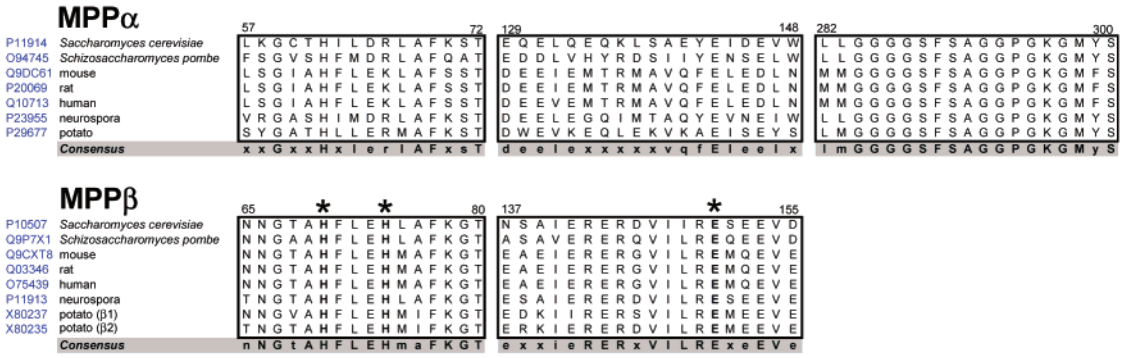


Figure 14. Sequence alignment of select regions of α-MPP and *β*-MPP. Residues involved in metal binding are indicated with an asterisk. The alignment was generated with the program ClustalW[294] using amino acid sequences resulting from a BLAST search[295] in the SWISS-PROT database.[296] The Swiss-Prot accesssion numbers are listed to the left of each sequence.

magnesium. The addition of other metal ions such as Cu^{2+}, Fe^{2+}, Zn^{2+}, and Co^{2+} at high concentration inhibited activity. The yeast MPP and its individual subunits were recently purified using an *E. coli* overexpression system. The purified α-MPP and *β*-MPP possessed close to stoichiometric levels of Zn^{2+} bound.[228] Taken together, the results [224,229,230] are consistent with the protease being a zinc metallopeptidase.

There are three noteworthy regions of the sequence alignment (Figure 14). First, a conserved HXXEH motif is present in the *β*-MPP family. This motif is absent in the α-MPP polypeptide in this region. The HXXEH is the hallmark of the pitrilysin family of proteases.[231] Members in this family include petrilysin and the insulin-degrading enzyme, both of which contain stoichiometric amounts of Zn^{2+}. The HXXEH motif is an inversion of the well-known HEXXH motif that defines the metal binding site in the active site region within the thermolysin family.[232] The second interesting conserved region is enriched in negatively charged residues. The third notable region is a glycine-rich domain present only in the α-subunit of MPP.

C. Site-Directed Mutagenesis and Type of Protease

To prove that the HXXEH in the *β*-MPP is important for catalysis, site-directed mutagenesis was performed. Mutagenesis of the histidines and the glutamic acid in the HXXEH motif in the rat *β*-MPP enzyme completely blocked catalytic activity.[233,234] Also, substitution of the first histidine in the HXXEH motif of the yeast *β*-MPP inhibited activity.[228] Interestingly, this mutation decreased the ability of this subunit to retain Zn^{2+} metal ion, consistent with this residue's involvement in coordinating the metal ion.

The glycine-rich region and the C-terminal 41 amino acids of the α-MPP are important for MPP processing and substrate binding.[235,236] Ito and colleagues concluded that α-MPP contains the substrate binding sites that help MPP recognize, bind, and cleave preproteins with long matrix peptides and basic amino acid residues. There may be a common structure adopted by the substrates that are bound to MPP.[237] A significant contribution to the specificity for the binding of the matrix peptide substrate appears to result from the −2 arginine residue binding to the S2 pocket of MPP. The length of the intervening sequence between the −2 arginine and the upstream (distal) N-terminal arginine of the matrix peptide is unimportant for substrate binding and cleavage because the distance between the two arginines bound to the MPP enzyme is the same. The data suggest that the intervening region is quite flexible.

The current hypothesis is that MPP carries out catalysis using a mechanism similar to that of the pitrilysin family of metallopeptidases that contain the HXXEH motif. Mutagenesis studies on the insulin-degrading enzyme and pitrilysin suggested that the two histidines in the HXXEH sequence and an additional downstream glutamic acid are involved in binding the metal ion.[238,239] Comparison of the sequence of the *β*-MPP with other members of the pitrilysin family revealed some conserved glutamic acid residues. Ito and co-workers studying the rat enzyme showed that one of these glutamic acid residues (Glu 136) was essential for catalysis,[240] and aligned in the sequence with the glutamic acid residue that coordinated a Zn^{2+} in pitrilysin.[238,239,241]

For MPP, the picture that has emerged involves the two histidine residues in the HXXEH motif coordinating the Zn^{2+} metal ion. Glu 136 also participates in the coordination of the metal ion. Glu 59 (in the HXXEH motif) functions as the general base, abstracting the proton from the water molecule and, thereby, facilitating catalysis.

D. Structure of MPP

The first clue to the three-dimensional structure of MPP came from the structure of the homologous core subunits within the mitochondrial cytochrome bc_1 complex from bovine and chicken[242−244] (Figure 15A). The core components, core 1 and core 2, have 21% identity in the bovine system. Each of the core components is comprised of two domains of equal size and of the same fold. The domains contain a five to six stranded *β*-sheet which lies adjacent to three α-helices on one side and one α-helix from the other domain on the opposite side. The two core polypeptides come together, forming a large internal cavity with a small crevace allowing access to the interior. The catalytic metal is not seen in the structure. Using

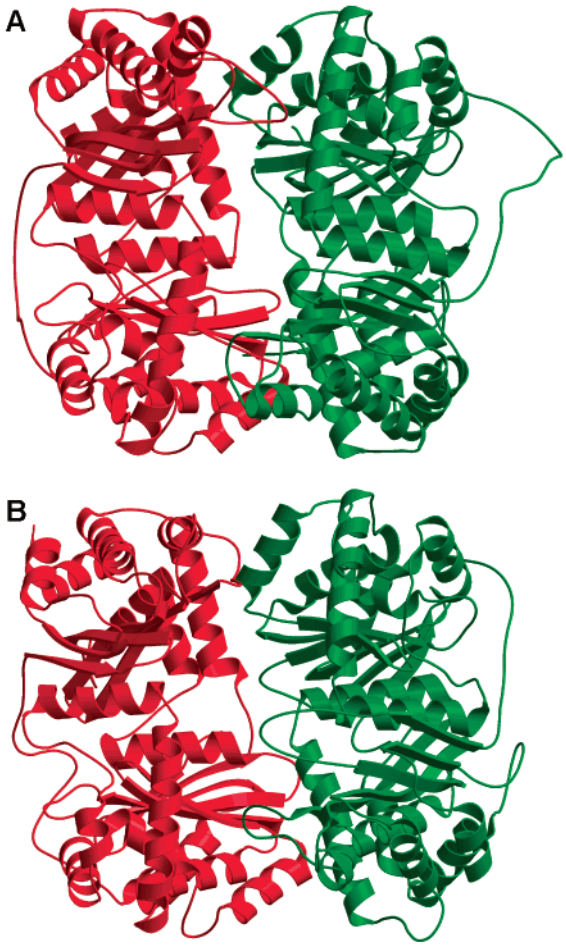


Figure 15. (A) Ribbon diagram of the mitochondrial cytochrome bc_1 complex from bovine. Chain A is shown in green, and chain B is shown in red. The atomic coordinates PDB 1BE3 were used. (B) Ribbon diagram of the yeast mitochondrial processing peptidase. The α-subunit (chain A) is shown in green, and the β-subunit (chain B) is shown in red. The atomic coordinates PDB 1HR6 were used. The programs MOLSCRIPT[298] and RASTER3D[299] were used to produce this figure.

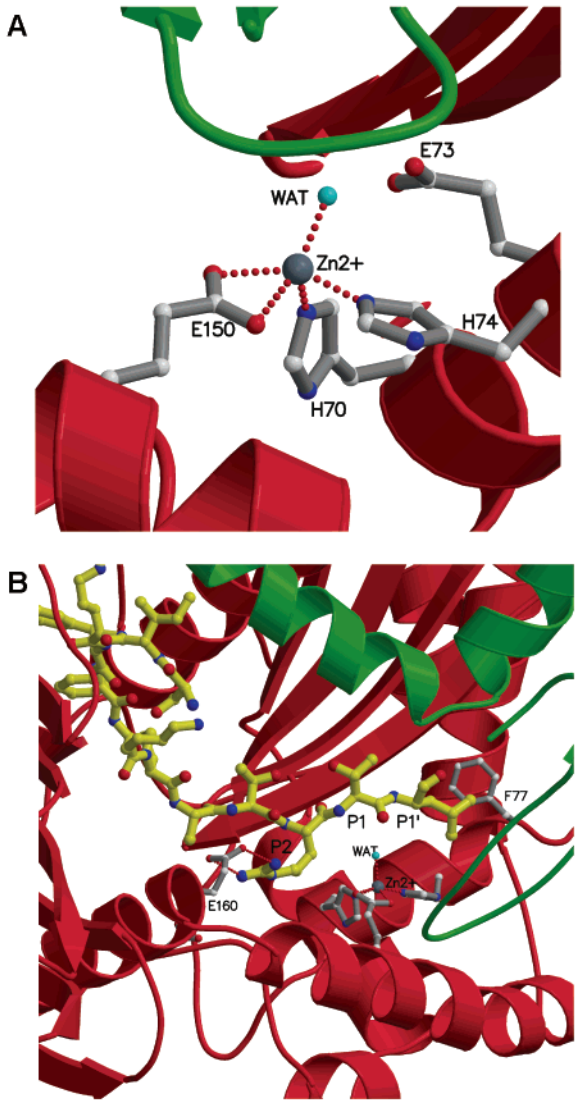


Figure 16. Active site of MPP. (A) Zinc binding site of MPP. (B) Structure of the Cox IV peptide substrate bound to the active site mutant of MPP (α/βE73Q MPP/COX IV). This figure was prepared using the programs MOLSCRIPT[298] and RASTER3D.[299]

the coordinates of the core proteins of the bovine bc_1 core complex, Ito built a model of the rat MPP.[197] The model depicts the metal binding site on β-MPP, the glycine-rich loop on α-MPP near the active site, and the Gly 414/Asp 415 thought to be important in binding long matrix peptides.

In 2001, the structure of the yeast MPP was solved (Figure 15B) and shows a fold similar to that of the bc_1 core components (compare Figure 15B with Figure 15A).[245] Each MPP subunit has two domains like the core components with essentially the same fold. The fold consists of a β-sheet comprised of six strands with three α-helices against one face, and five other α-helices at one end of the sheet.

Figure 15B shows a large central cavity of MPP that contains the active site formed between the α-MPP and β-MPP subunits. The region is quite hydrophilic, containing many acidic residues donated from both subunits. As the signal sequence is rich in basic amino acids, the electrostatic interactions likely promote binding of the substrate to MPP. The highly polar environment is most likely the reason the signal sequence, which has been shown to form an α-helix in a hydrophobic environment, does not form this kind of structure when bound to MPP. Another notable feature is the glycine-rich domain of the α-MPP forming a flexible loop that partially blocks the active site region. This is consistent with it being involved in substrate binding or product release. The zinc binding site of MPP is easily seen in the 2.5 Å map (Figure 16A). Zinc is coordinated to His 74, His 70, and Glu 150 of the yeast β-MPP, corresponding to His 56, His 60, and Glu 136 of the rat β-MPP, which explains why substitution of these residues abolishes activity. The structure of the apo-MPP enzyme shows a water molecule coordinated to the metal ion, within H-bonding distance of the yeast β-MPP general base Glu 73. The structure proves that the inverted zinc binding motif, HXXEH, does indeed function as a zinc-coordinating element.

The structure of MPP with bound peptide substrates was solved to help elucidate the structural features important for substrate binding. To obtain these complexes, a mutant β-MPP was created in

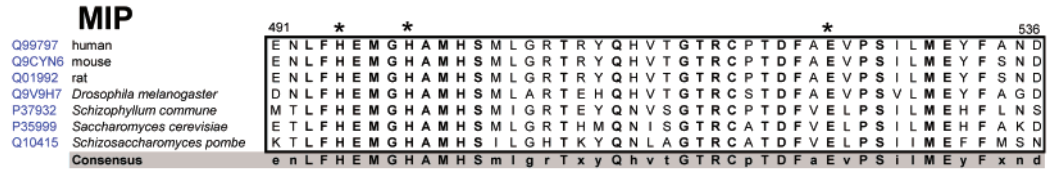


Figure 17. Alignment of sequences near the active site region of MIP. The amino acid sequence of human, mouse, rat, fruit fly, fungus (*Sc. commune*), and yeast (*S. cerevisiae, Sch. pombe*) MIPs are depicted. Identical residues are in bold with the consensus sequences below. Residues predicted to be involved in metal binding are indicated with an asterisk. The alignment was generated with the ClustalW program[294] using amino acid sequences resulting from a BLAST search[295] in the SWISS-PROT database.[296] The Swiss-Prot accesssion numbers are listed to the left of each sequence.

which the general base Glu 73 was changed to a glutamine. The structure of the mutant MPP was solved to 2.7 Å (Figure 16B) bound to a peptide substrate derived from the yeast cytochrome *c* oxidase subunit IV (Cox IV). This substrate is composed of residues 2−25 (2LSLRQSIRFFKPATRT-LCSS-RYLL25; the cleavage site is between Thr 17 and Leu 18) of the CoxIV preprotein. In the structure, the electron density of residues 7−19 of the peptide is observed while the rest of the peptide is not observed since it was disordered in the crystal. Residues 16−18 of the substrate form a short β-strand which hydrogen bonds with a β-strand (residues 98−103 in β-MPP) of MPP in an antiparallel configuration. It passes between two α-helices and possibly interacts with another β-strand in β-MPP. The carbonyl oxygen of the scissile peptide bond is oriented in the direction of the Zn^{2+} metal ion. The −2 (P2) arginine of the substrate signal peptide was found in the S2 binding site formed by Glu 160 and Asp 164 of the β-MPP subunit and interacts with the arginine via a salt bridge. The leucine residue at the +1 (P1′) position of the substrate is located adjacent to Phe 77 within the S1′ binding site of β-MPP.

Additionally, the structure of the mutant MPP was solved with the bound peptide substrate based on yeast malate dehydrogenase, MDH (2LSRVAKRA-FSSTVANP17). This structure at 3.0 Å resolution confirmed the substrate−MPP interactions. Surprisingly, the peptide was cleaved by the mutant MPP despite the general base Glu 73 being mutated to a glutamine in the β-MPP subunit. The structure of the MPP−product complex revealed that the C-terminal oxygen of the signal peptide substrate was coordinated to the zinc metal ion, taking the place of the water molecule. Similar types of interactions involving the carboxyl group of the product with the metal ion were proposed in the catalytic mechanism of the protease thermolysin.[232] Also, in this structure, residues 7 and 8 of the product form a main chain hydrogen bond to the same β-strand seen in the CoxIV peptide structure of the β-MPP in the antiparallel configuration in the active site region. The structures reveal that β-strands 2−4 of MPP may participate in the substrate recognition event.

E. Relationship to MIP

As mentioned in section I, a number of mitochondrial preproteins containing matrix targeting peptides are processed by MIP, in addition to MPP. These precursors contain the motif R-X↓(F/L/I)-X-X-(T/S/G)-X-X-X-X↓ at the C-terminus of the extension peptide possessing a −10 arginine residue (relative to the MIP processing site).[43,44] MPP cleaves first after the RX sequence, and then MIP cleaves eight residues downstream. A more recent analysis of the two-step cleavage substrates determined a more refined motif for these substrates: RX↓(F/L/I)-(S/X)-(S/T/X)-(T/S/G)-X-X-X-X↓.[246] Cleavage at the second site requires cleavage at the first site. MIP behaves as an octapeptidase in this process.[247]

MIP was purified to homogeneity from the rat mitochondria[248] and the yeast mitochondrial matrix.[249] The function of MIP from rat and yeast was inhibited by the cysteine-reactive reagent NEM. Metal chelator also inhibited the MIP activity, while the activity was restored by the addition of manganese but not zinc. Therefore, the analysis suggested that it was a metallopeptidase.

MIP was cloned from several organisms, including rat,[250] yeast,[251] *Schizophyllum commune*,[252] and human liver.[253] Sequence analysis revealed (Figure 17) that these enzymes have a Zn^{2+} binding motif, H-E-X-X-H, also found in the thermolysin Zn family of proteases and the astacin and collagenase families of proteases.[252,254] In addition, their sequences were homologous to those of the thimet oligopeptidases, which displayed 20−24% sequence identity to MIP. This is logical given that MIP is an oligopeptidase.

Site-directed mutagenesis studies confirmed the importance of the two histidines and the glutamic acid in the Zn^{2+} binding motif for catalysis, since their substitution abolished activity.[249] These results underscore the importance of a functional zinc binding motif for MIP function. Structure/function studies have yet to explain the reason that MIP proteins are inactivated by NEM.[196]

VII. Stroma Processing Peptidase

Blobel and colleagues discovered in 1977 that imported chloroplast proteins were synthesized with a transit peptide.[255] Though the enzyme was partially purified in 1984,[256] the enzyme was not purified to homogeneity until 1992.[257] The protease that cleaved chloroplast preproteins, now called SPP, is a metallopeptidase and belongs to the same protease family as MPPs (see below).

A. Stroma Targeting Sequence and Substrate Specificity

Comparison of stroma targeting sequences shows that they vary widely in size, are enriched in serine and threonine residues, and have few acidic resi-

dues.[258] In contrast to ER signal sequences, they contain no consensus sequence near the cleavage site. The C-terminal sequence generally has a basic residue and lacks acidic residues but can tolerate almost any amino acid side chain at the −1 and +1 positions. However, if one looks carefully, there is a loosely conserved motif surrounding the stroma signal processing site, (I/V)-X-(A/C)↓A, where the arrow indicates the cleavage site.[258]

Several studies were carried out in an attempt to determine the function of the regions of the stroma targeting sequence. Generally, the amino-terminal region is important for import and the carboxy-terminal region is critical for cleavage. Deletions in the carboxy-terminal region of the stroma targeting peptide of the small subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase (Rubisco) altered processing but not import.[259] In the case of the ferredoxin stroma targeting sequence, the amino-terminal portion was important for import whereas the carboxyl-terminal region was important for import and processing both in vitro[260] and in vivo.[261] Further dissection of the ferredoxin transit peptide showed that an amino-terminal deletion affected import at the stage of interaction with the Toc components. A more central deletion affected import by perturbing the interaction of ferrodoxin with the Tic components.[262]

How does substrate cleavage occur with such high fidelity and efficiency given that there are no typical substrate specificity determinants? The answer to this question remains unknown. In some cases, the conserved basic residue is important for processing,[263,264] while in others, it is not.[265] Recently, Richter and Lamppa directly tested the substrate specificity of SPP using the purified enzyme and different preprotein substrates.[266] SPP cleaved six radioactively labeled preproteins without the help of additional chloroplast proteins; strikingly, five of the six preproteins were cleaved after a basic residue. This was also true for the precursor to Rubisco activase that was cleaved by purified SPP between Lys 57 and Ala 58. Previously, it was predicted to cleave between Ala 58 and Ala 59. Therefore, there may be an additional uncharacterized protease in vivo, which is involved in the removal of a few amino-terminal residues from the mature protein. More recently, Emanuelsson et al. used a motif-finding algorithm to find the most common motif in the −20 to +6 region of chloroplast proteins with transit peptides.[267] The motif found was VR↓AAAVXX, where cleavage is predicted (as indicated by the arrow) on the basis of the Richter and Lamppa results.[266] If the specificity of SPP for cleavage after a basic residue is true, then the problem of what determines cleavage specificity is more understood. However, additional proteases would have to play a key role in processing of chloroplast proteins.

B. Purification and Characterization of SPP

To elucidate in more detail signal peptide processing by SPP, the enzyme responsible for processing was partially purified.[256] This protease was called the chloroplast processing enzyme (CPE) (referred to as SPP in this review). SPP has a pH optimum near 9 and an approximate molecular weight of 180000. The enzyme is also inhibited by 1,10-phenanthroline and ethylenediaminetetraacetic acid. More recently, the protein was purified to homogeneity by Oblong and Lamppa using affinity chromatography with the precursor to the light-harvesting chlorophyll binding protein (LHCP) as a ligand, followed by gel filtration and anion exchange chromatography.[257] The purified monomeric protein runs as a 143/145 kDa doublet on an SDS−PAGE gel. This protease processes the precursors to LHCP, acyl carrier protein, and the small subunit of Ribisco.

An immobilized SPP was also prepared using an *E. coli* overexpression system with a biotin-containing peptide at its amino terminus, which was bound to streptavidin-coated paramagnetic particles and extensively washed.[266] This immobilized SPP processed a wide variety of precursors, including the precursor of LHCP, the small subunit of Ribisco, Ribisco activase, ATP synthase γ subunit, oxygen evolving enhancer, plastocyanin, acyl carrier protein, 3-deoxy-D-*arabino*-heptulosonate 7-phosphate synthase, heat shock protein 21, and ferredoxin. According to the data, SPP functions as a general stromal processing peptidase. The activity of SPP is strongly inhibited by 1,10-phenanthroline. The activity, if inhibited by low amounts of 1,10-phenanthroline, can be restored by the addition of Zn^{2+}. These data are consistent with the stroma processing peptidase being a metallopeptidase.

C. Cloning of SPP, Protease Class, and Physiological Function

Antibodies generated against the purified SPP protein were used to isolate a cDNA from pea that encodes the stroma processing peptidase.[268] The cDNA open reading frame predicts a protein of 1259 amino acid residues of approximately 140000 molecular weight. The enzyme contains an amino-terminal region predicted to be a stroma targeting peptide. By examining the SPP sequence, one observes strong homology to the pitrilysin metallopeptidase family of proteins (25−30% identity) including the zinc binding motif (His-X-X-Glu-His) and a 130 amino acid region in the N-terminal part of the SPP (starting at Leu 222) (Figure 18). Therefore, SPP and MPP evolved from a common ancestor. Also, SPP possesses a large C-terminal region of unknown function not found within the petrilysin family.

The sequence homology of the SPP to proteases in the pitrilysin family and its sensitivity to metal chelators indicate that SPP is a metallopeptidase. The results suggest that the histidines in the H-X-X-E-H sequence are involved in coordination of the zinc metal, and that the conserved glutamic acid in the H-X-X-E-H motif would function as the general base. Structural studies, of course, will have to verify this model.

To determine the role of SPP in the biogenesis of imported chloroplast proteins, an antisense SPP RNA was used to decrease the levels of SPP within transgenic tobacco plants.[66] Lamppa and co-workers found that the decreased level of SPP drastically

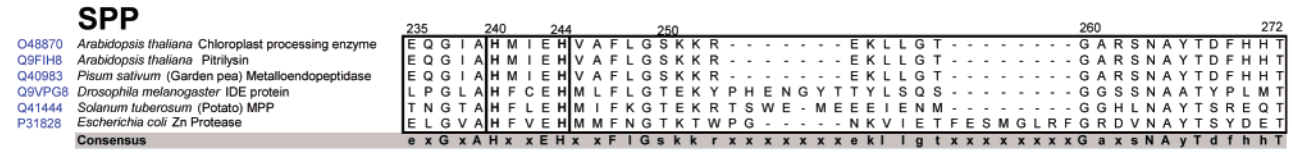


Figure 18. Sequence alignment of the region near the proposed active site of SPP and similar enzymes. The alignment includes the *Arabidopsis thaliana* CPE and pitrilysin, a metalloendopeptidase from the garden pea, IDE protein from the fruit fly, MPP from potato, and finally a zinc protease from *E. coli*. The alignment was generated with the ClustalW program[294] using amino acid sequences resulting from a BLAST search[295] in the SWISS-PROT database.[296] The Swiss-Prot accesssion numbers are listed to the left of each sequence. A consensus sequence for the region is listed at the bottom of the alignment.

affected chloroplast biogenesis, causing retarded shoot and root growth. The number of chloroplasts per cell and the number of thylakoids in each chloroplast also decreased. As expected, the processing activity of the precursor of Rubisco activase was lower in the antisense SPP transgenic tobacco plants. Even more striking was that the import of the precursor to LHCP was inhibited. This suggests that the function of the stroma processing peptidase is required for efficient import. Obviously, SPP plays a significant role in chloroplast biogenesis and protein import.

VIII. Signal Peptide Hydrolases

Signal peptides generated following preprotein cleavage usually disappear in a short time period. This may be due to the fact that signal peptides are harmful to cells since they can inhibit protein translocation and may be lytic if they accumulate in the membrane. The understanding of the signal peptide degradation pathway is only in its infancy and will require a major effort and extensive work in the future.

A. Signal Peptide Degradation in Bacteria

The first clue that there was an enzyme that degraded signal peptides came from work by Mizushima and co-workers studying the membrane biogenesis of the major lipoprotein protein destined to the outer membrane.[174] The addition of globomycin, an SPase II inhibitor, led to the accumulation of prolipoprotein within the inner membrane. By isolating the envelope fraction and incubating in buffer, the processing of prolipoprotein to mature lipoprotein could be observed. While under most conditions the signal peptide was not detected, it could be seen when the envelope fraction containing the radioactive prolipoprotein was incubated at 80 °C for 7 min. Presumably, the reason it could be detected after high-temperature incubation was because the signal peptide degradation activity was heat sensitive.

1. Signal Peptide Peptidase

The signal peptide degradation activity was then shown to correspond to protease IV,[269] an inner membrane protein, which was previously characterized.[270] Protease IV was partially purified and shown to have signal peptide peptidase activity. It was inhibited by the protease inhibitors antipain, chymostatin, elastinal, and leupeptin, suggesting it is a serine protease.

Next, the gene encoding protease IV was isolated by screening the Clark and Carbon collection for clones overexpressing signal peptide peptidase activity. The gene encoding signal peptide peptidase was isolated and found to encode a protein of 618 amino acids with a predicted molecular weight of 67241.[271] This correlated very well with the size of the purified protease IV isolated from cell lysate. The purification of protease IV involved the isolation of the envelope and detergent extraction of the membranes, followed by DE52 ion exchange chromatography and Sephacryl S-300 gel filtration.[271] Finally, cross-linking studies showed that protease IV exists as a tetramer.

To examine the in vivo importance of protease IV, a deletion mutant was made.[272] The *sppA* deletion mutant was constructed, and the position of the gene was determined to be at 38.5 min on the *E. coli* chromosome. The deletion mutant showed that the *sppA* is not an essential gene. The *sppA* mutant displayed slower kinetics in the degradation of the lipoprotein signal peptide, although the signal peptide was still degraded, showing that other proteases function in signal peptide degradation.

In *B. subtilis*, there are two homologues of the *E. coli* signal peptide peptidase:[273] SppA and TepA. SppA is highly homologous to the *E. coli* signal peptide peptidase. TepA has lower homology to SppA and has sequence similarity to the cytosolic ClpP proteases. This protease was named TepA for translocation enhancing protein since *B. subtilis* TepA was required for efficient export of proteins. SppA is an inner membrane protein, while TepA is found in the cytoplasm. Using conditional *B. subtilis* strains, van Dijl and co-workers[273] showed that depletion of SppA impaired the processing of preproteins but only under hypersecretion conditions. In contrast, depletion of TepA in *B. subtilis* demonstrated impaired processing and translocation of secretory proteins even under nonhypersecretion conditions. These studies are consistent with SppA and TepA playing a role in either degrading signal peptides or degrading proteins that are inhibitory to protein export.

2. Oligopeptidase A

In conjunction with the Mizushima studies, Novak et al. isolated another major activity responsible for degrading the lipoprotein signal peptide.[274] The scientists prepared a radiolabeled signal peptide substrate for these studies. The protein activity responsible for signal peptide cleavage was shown to

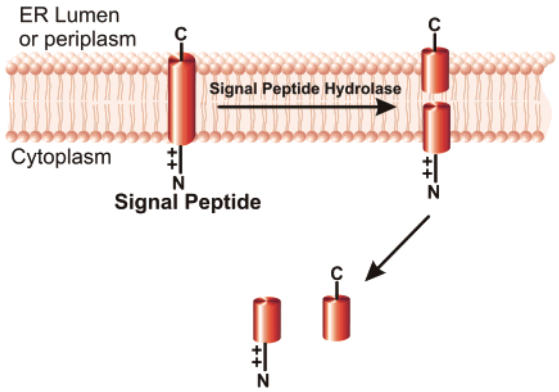


Figure 19. Schematic of the steps in signal peptide hydrolysis. After being cleaved from an exported protein, the signal peptide resides within the lipid bilayer until it is cleaved within its transmembrane segment by a signal peptide hydrolase. The smaller cleaved peptides from this reaction most likely recede back into the cytosol where they are degraded by proteases or become involved in recently described signaling events.[285]

correspond to a protein with a molecular weight of 68000, possessing properties similar to those of the cytoplasmic protease oligopeptidase A from *Salmonella typhimurium*.[275] Both enzymes are inhibited by metal chelators and are stimulated in the presence of Co^{2+}. Additionally, both hydrolyze *N*-acetyl-Ala and have a molecular weight of 68000. Further studies examined the degradation of prolipoprotein signal peptide in the membrane.[276] These studies supported the idea that signal peptide peptidase (protease IV) first cuts the membrane-spanning signal peptide and then oligopeptidase A further degrades the cut portions of the signal peptide.

The gene encoding oligopeptidase A was isolated from *S. typhimurium*[277] and *E. coli*.[278] The genes show that the oligopeptidases have the Zn^{2+} binding motif (HEXXH) found in the thermolysin family. Sequence homology studies reveal oligopeptidase A enzymes are homologous to dipeptidyl carboxypeptidase, metalloendopeptidase EP 24.15,[279] pig angiotensin binding protein,[280] and rat mitochondrial intermediate peptidase.[281] These proteins belong to the thimet protease family that are typically activated by thiols and require a metal ion for activity.[282]

B. ER Signal Peptide Degradation

Recent studies addressed the signal peptide peptidase activity in the ER. These studies were performed by Dobberstein and colleagues,[283] in which they examined the fate of the preprolactin signal peptide. They found that, after SPC cleavage, the liberated signal peptide of the ER-inserted preprotein was quickly cleaved into small pieces. On the basis of size estimation, they proposed that the signal peptide was most likely cleaved in the hydrophobic core region between two leucine residues. In this fashion, the ER signal peptide peptidase activity is similar to that of the bacterial signal peptide peptidase, which also cleaves within the hydrophobic core region of the signal peptide (Figure 19). Zimmermann and co-workers[284] investigated processing of preprolactin signal peptide and the signal peptide of preprocecropin A. While the cleavage of the preprolactin signal peptide required the addition of cytosol, cleavage of the preprocecropin A signal peptide occurred even without the addition of cytosol. Interestingly, the addition of the immunosuppressant cyclosporin A inhibited signal peptide cleavage of both signal peptides. These signal peptides were associated with two ER proteins prior to being processed.

Possibly, in eukaryotic cells the cleaved and released signal peptides may function as cellular signaling molecules. When the cleaved preprolactin signal peptide is released into the cytosol, it is found associated with calmodulin.[285] Similar findings were observed with the signal peptide of the HIV-1 envelope protein p-gp160. Moreover, certain cleaved signal peptides of the human major histocompatibility complex (MHC) can be presented by nonclassical MHC molecules on the cell surface where they allow the cell to resist attack by natural killer (NK) cells.[286] Recently, Lemberg et al. proved that signal peptide cleavage of MHC class I signal peptides is essential for binding of the signal peptide derived region to the nonclassical MHC class I molecule HLA-E.[287] Following cleavage by signal peptide peptidase, the signal peptide is released from the membrane into the cytosol and is then transported into the ER by the TAP transporter where the peptide binds to the HLA-E protein. At this point, they are moved to the cell surface and the epitope is presented to receptors on NK cells. Another interesting finding in this paper is that the generated signal peptide is stably integrated within the lipid bilayer, as the signal peptide remains in the pellet fraction in a carbonate extraction experiment. These data are consistent with signal peptide cleavage occurring within the portion of the signal peptide that is located within the intramembrane region. Cleavage of the signal peptide then releases the peptides from the lipid bilayer.

Clues to understanding the protease responsible for signal peptide degradation came recently through the utilization of protease inhibitors. The addition of the cysteine protease inhibitor 1−3-di(*N*-carboxybenzoyl-L-leucyl-L-leucyl)aminoacetone $(Z\text{-}LL)_2$-ketone, prevented proteolysis of the in vitro generated signal peptides of preprolactin and HIV-1 gp160.[288] The cleavage site within the signal peptide was predicted and determined to be within the center of the membrane-spanning region (Figure 19). Very recently, the human SPP has been cloned and the protein expressed in *S. cerevisiae*.[289] Yeast ER membranes prepared from the yeast expressing the human SPP showed that SPP activity was sensitive to a diazirine-containing derivative of the $(Z\text{-}LL)_2$-ketone inhibitor. Strikingly, SPP may be an aspartic acid protease as it has sequence motifs that are similar to those found in the presenilin-type aspartic acid proteases.

C. Signal Peptide Degradation in Chloroplasts

Following cleavage of the imported chloroplast preprotein by the SPP, the transit peptide is degraded. The proteases carrying out the degradation of the transit peptide are not yet fully characterized.

Interestingly, when SPP (immobilized onto magnetic beads) is incubated with ^{35}S-labeled preproteins,

some subfragments are observed from the preproteins of ferredoxin, heat shock protein 21, LHCP, and Ribisco.[266] Processing of the chloroplast transit peptide to the subfragments by the extensively purified SPP is sensitive to 1,10-phenanthroline, suggesting that SPP itself, and not a contaminating protease, catalyzes the conversion of the transit peptide to the subfragments. While the generated transit peptides bind to SPP with high affinity, the subfragments do not bind to the SPP.

In intact chloroplasts, the pattern of trimming of the transit peptides is similar within the initial time period to that observed with the immobilized SPP.[266] After the subfragments are generated they are rapidly degraded within the chloroplast. As seen in the in vitro system, the degradation of the subfragments in the chloroplast can be prevented by adding metal chelators, 1,10-phenanthroline, or EDTA. The degradation of the subfragments is even more sensitive to these metal chelators than the initial signal peptide cleavage reaction. The addition of 5 mM phenanthroline completely abolishes cleavage of the purified ferredoxin transit peptide subfragment, which can then be restored by the addition of 300 μM Zn^{2+}. The addition of several protease inhibitors against the cysteine, serine, and aspartic acid protease classes produced no affect. The data suggest that the enzyme requires Zn^{2+} for proteolytic activity. Further studies showed that the proteolytic activity also requires ATP.

The present data are consistent with an ATP-dependent metallopeptidase. While the protease has not yet been identified, there are only a few ATP-dependent proteases known to exist within the chloroplast. One of these is the homologue of the *E. coli* FtsH.[290] However, since the subfragment degradation activity is soluble, it is unlikely that the FtsH homologue is involved since this is a membrane-bound protease. Another possibility is that it is the Clp protease,[291] which is a serine protease that is not particularly sensitive to PMSF. Alternatively, it could be a novel, unidentified protease.

IX. Concluding Remarks and Perspective

In this review, we discussed a diverse group of SPases found in bacteria, ER, chloroplasts, and mitochondria. These ubiquitous SPases play a vital role in the cleavage of signal peptides that target the protein to the correct subcellular destination. In some cases, the failure to remove the signal peptide results in the protein being mislocalized either in the cell or in an inactive imported protein.

These SPases include both membrane-bound and soluble members, and several of the peptidases belong to novel classes of protease. The types I and II SPases as well as the type IV SPases are integral membrane proteins located in the plasma membrane of bacteria. Homologues of the type II and type IV proteases are not found in eukaryotic cells. Type I SPases have homologues in the chloroplast thylakoid membrane (TPP) and mitochondrial inner membrane (Imp1/Imp2) that use the same serine-lysine dyad mechanism. In contrast, the ER-localized SPC uses a different catalytic mechanism involving serine and histidine residues. The type II and type IV SPases are unusual aspartic acid proteases that do not possess the conserved Asp-Thr/Ser-Gly motif found in most aspartic acid proteases. Unlike type I and type II SPases that have their catalytic domains on the *trans* side (away from the cytoplasm) of the membrane, the active site of type IV SPases resides on the cytoplasmic membrane surface.

MPP is an SPase that removes the transit peptides from preproteins imported into the mitochondria. MPP is a metallopeptidase utilizing a catalytic zinc atom and is a member of the pitrilysin family of proteases. Some preproteins destined for the mitochondrial matrix are processed not only by MPP, but also by MIP, an octapeptidase. Like MPP, MIP is also a metallopeptidase utilizing zinc. SPP removes transit peptides from chloroplast proteins that are imported into the stroma compartment from the cytoplasm. It shows sequence homology to MPP and is also a member of the pitrilysin metallopeptidase family.[292] Though SPP and MPP most likely evolved from a common ancestor, SPP is active as a monomer while MPP is active only as a dimer.

One of the challenges in the future is to obtain structural information on the membrane-bound proteases. It is absolutely essential to solve the structure of the full-length SPase I, II, and IV to determine the location of the active site with respect to the membrane. It is still not clear whether catalysis takes place within the membrane environment or in the aqueous environment. Knowing where catalysis takes place may shed light on how these SPases are able to cleave their substrates with such remarkable accuracy. The structure of the membrane-bound enzymes would also be useful for designing inhibitors, because it could provide clues regarding the type of compounds having access to the active site. This would be particularly useful for the bacterial type I SPases, which are being considered as antimicrobial targets, and the eukaryotic ER SPC that is a potential antiviral target.

Another area deserving much more attention is the study of the proteases that degrade the signal peptides. Particularly intriguing are the peptidases that hydrolyze the membrane-spanning signal peptides after they are generated by SPase cleavage of the preproteins. On the basis of the limited results collected so far, it seems likely that these enzymes attack the signal peptide within the membrane, thereby releasing the fragments into the aqueous environment for further degradation. Such a proposal, where cleavage takes place within a membrane, has been made with γ-secretase that cleaves the membrane-localized β-amyloid precursor protein within the membrane, associated with Alzheimer's disease.[293] Another possibility is that the signal peptides are removed first from the membrane and then cleaved by the peptidase. The mechanism by which a hydrophobic signal peptide is released from the membrane would be very interesting to know, as it most likely would be a protein-catalyzed mechanism.

X. Abbreviations

SPase I	type I signal peptidase, leader peptidase
SPase II	lipoprotein signal peptidase
SPase IV	prepilin signal peptidase
ER SPC	endoplasmic reticulum signal peptidase complex
Imp	inner membrane protease (located in the mitochondria)
MPP	matrix processing peptidase (located in the mitochondria)
MIP	mitochondrial intermediate peptidase
SPP	stromal processing peptidase
TPP	thylakoid processing peptidase
Tat	twin arginine translocation

XI. Acknowledgment

This work was supported by National Science Foundation Grant MCB-9808843 and National Institutes of Health Grant GM63862 (to R.E.D.) and by grants from the Canadian Institute of Health Research, the Canadian Bacterial Diseases Network of Excellence, the Burroughs Wellcome Foundation, and the Howard Hughes Medical Institute to N.C.J.S. M.P. is supported by a CIHR postdoctoral fellowship and N.C.J.S. by a CIHR scholarship. We thank Dr. Andreas Kuhn for reading the manuscript.

XII. References

(1) Blobel, G., Sabatini, D. D. *Biomembranes* **1971**, *2*, 193.
(2) Milstein, C.; Brownlee, G. G.; Harrison, T. M.; Mathews, M. B. *Nat. New Biol.* **1972**, *239*, 117.
(3) Blobel, G.; Dobberstein, B. *J. Cell Biol.* **1975**, *67*, 852.
(4) Schatz, P. J.; Beckwith, J. *Annu. Rev. Genet.* **1990**, *24*, 215.
(5) Bedouelle, H.; Bassford, P. J., Jr.; Fowler, A. V.; Zabin, I.; Beckwith, J.; Hofnung, M. *Nature* **1980**, *285*, 78.
(6) Puziss, J. W.; Fikes, J. D.; Bassford, P. J., Jr. *J. Bacteriol.* **1989**, *171*, 2303.
(7) Nesmeyanova, M. A.; Karamyshev, A. L.; Karamysheva, Z. N.; Kalinin, A. E.; Ksenzenko, V. N.; Kajava, A. V. *FEBS Lett.* **1997**, *403*, 203.
(8) Briggs, M. S.; Cornell, D. G.; Dluhy, R. A.; Gierasch, L. M. *Science* **1986**, *233*, 206.
(9) Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157.
(10) Miller, A.; Wang, L.; Kendall, D. A. *J. Biol. Chem.* **1998**, *273*, 11409.
(11) Wang, L.; Miller, A.; Kendall, D. A. *J. Biol. Chem.* **2000**, *275*, 10154.
(12) Hoyt, D. W.; Gierasch, L. M. *J. Biol. Chem.* **1991**, *266*, 14406.
(13) Nielsen, H.; Engelbrecht, J.; Brunak, S.; von Heijne, G. *Int. J. Neural Syst.* **1997**, *8*, 581.
(14) Nielsen, H.; Engelbrecht, J.; Brunak, S.; von Heijne, G. *Protein Eng.* **1997**, *10*, 1.
(15) von Heijne, G.; Abrahmsen, L. *FEBS Lett.* **1989**, *244*, 439.
(16) Edman, M.; Jarhede, T.; Sjostrom, M.; Wieslander, A. *Proteins* **1999**, *35*, 195.
(17) Berks, B. C. *Mol. Microbiol.* **1996**, *22*, 393.
(18) Stanley, N. R.; Palmer, T.; Berks, B. C. *J. Biol. Chem.* **2000**, *275*, 11591.
(19) Bogsch, E.; Brink, S.; Robinson, C. *EMBO J.* **1997**, *16*, 3851.
(20) Hayashi, S.; Wu, H. C. *J. Bioenerg. Biomembr.* **1990**, *22*, 451.
(21) Strom, M. S.; Lory, S. *Annu. Rev. Microbiol.* **1993**, *47*, 565.
(22) Reyss, I.; Pugsley, A. P. *Mol. Gen. Genet.* **1990**, *222*, 176.
(23) Lindeberg, M.; Collmer, A. *J. Bacteriol.* **1992**, *174*, 7385.
(24) Nunn, D. N.; Lory, S. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 47.
(25) Bally, M.; Filloux, A.; Akrim, M.; Ball, G.; Lazdunski, A.; Tommassen, J. *Mol. Microbiol.* **1992**, *6*, 1121.
(26) Cioffi, J. A.; Allen, K. L.; Lively, M. O.; Kemper, B. *J. Biol. Chem.* **1989**, *264*, 15052.
(27) Kaiser, C. A.; Preuss, D.; Grisafi, P.; Botstein, D. *Science* **1987**, *235*, 312.
(28) Nielsen, H.; Engelbrecht, J.; von Heijne, G.; Brunak, S. *Proteins* **1996**, *24*, 165.
(29) Zheng, N.; Gierasch, L. M. *Cell* **1996**, *86*, 849.
(30) Ng, D. T.; Brown, J. D.; Walter, P. *J. Cell Biol.* **1996**, *134*, 269.
(31) Martoglio, B.; Dobberstein, B. *Trends Cell Biol.* **1998**, *8*, 410.
(32) von Heijne, G.; Steppuhn, J.; Herrmann, R. G. *Eur. J. Biochem.* **1989**, *180*, 535.
(33) von Heijne, G. *EMBO J.* **1986**, *5*, 1335.
(34) Roise, D.; Horvath, S. J.; Tomich, J. M.; Richards, J. H.; Schatz, G. *EMBO J.* **1986**, *5*, 1327.
(35) Roise, D.; Theiler, F.; Horvath, S. J.; Tomich, J. M.; Richards, J. H.; Allison, D. S.; Schatz, G. *EMBO J.* **1988**, *7*, 649.
(36) Abe, Y.; Shodai, T.; Muto, T.; Mihara, K.; Torii, H.; Nishikawa, S.; Endo, T.; Kohda, D. *Cell* **2000**, *100*, 551.
(37) Honlinger, A.; Keil, P.; Nelson, R. J.; Craig, E. A.; Pfanner, N. *Biol. Chem. Hoppe-Seyler* **1995**, *376*, 515.
(38) Bolliger, L.; Junne, T.; Schatz, G.; Lithgow, T. *EMBO J.* **1995**, *14*, 6318.
(39) Dietmeier, K.; Honlinger, A.; Bomer, U.; Dekker, P. J.; Eckerskorn, C.; Lottspeich, F.; Kubrich, M.; Pfanner, N. *Nature* **1997**, *388*, 195.
(40) Schatz, G. *Nature* **1997**, *388*, 121.
(41) Voos, W.; Martin, H.; Krimmer, T.; Pfanner, N. *Biochim. Biophys. Acta* **1999**, *1422*, 235.
(42) Schneider, G.; Sjoling, S.; Wallin, E.; Wrede, P.; Glaser, E.; von Heijne, G. *Proteins* **1998**, *30*, 49.
(43) Gavel, Y.; von Heijne, G. *Protein Eng.* **1990**, *4*, 33.
(44) Hendrick, J. P.; Hodges, P. E.; Rosenberg, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 4056.
(45) Hartl, F. U.; Neupert, W. *Science* **1990**, *247*, 930.
(46) Ko, K.; Cashmore, A. R. *EMBO J.* **1989**, *8*, 3187.
(47) Blight, K. J.; Kolykhalov, A. A.; Reed, K. E.; Agapov, E. V.; Rice, C. M. *Antiviral Ther.* **1998**, *3*, 71.
(48) Law, L. M.; Duncan, R.; Esmaili, A.; Nakhasi, H. L.; Hobman, T. C. *J. Virol.* **2001**, *75*, 1978.
(49) Lober, C.; Anheier, B.; Lindow, S.; Klenk, H. D.; Feldmann, H. *Virology* **2001**, *289*, 224.
(50) Lindemann, D.; Pietschmann, T.; Picard-Maureau, M.; Berg, A.; Heinkelein, M.; Thurow, J.; Knaus, P.; Zentgraf, H.; Rethwilm, A. *J. Virol.* **2001**, *75*, 5762.
(51) Mullins, C.; Lu, Y.; Campbell, A.; Fang, H.; Green, N. *J. Biol. Chem.* **1995**, *270*, 17139.
(52) Tjalsma, H.; Kontinen, V. P.; Pragai, Z.; Wu, H.; Meima, R.; Venema, G.; Bron, S.; Sarvas, M.; van Dijl, J. M. *J. Biol. Chem.* **1999**, *274*, 1698.
(53) Tjalsma, H.; Zanen, G.; Bron, S.; van Dijl, J. M. In *Co-translational and Posttranslational Proteolysis of Proteins*, 3rd ed.; Dalbey, R. E., Sigman, D. S., Eds.; Academic Press: San Diego, CA, 2001; Vol. XXII.
(54) Fraser, C. M.; Gocayne, J. D.; White, O.; Adams, M. D.; Clayton, R. A.; Fleischmann, R. D.; Bult, C. J.; Kerlavage, A. R.; Sutton, G.; Kelley, J. M.; et al. *Science* **1995**, *270*, 397.
(55) Pratje, E.; Esser, K.; Michaelis, G. In *Signal Peptidases*; Heijne, G. v., Ed.; R. G. Landes Co.: Austin, TX, 1994.
(56) Robinson, C. In *Signal Peptidases*; Heijne, G. v., Ed.; R. G. Landes Co.: Austin, TX, 1994.
(57) Inada, T.; Court, D. L.; Ito, K.; Nakamura, Y. *J. Bacteriol.* **1989**, *171*, 585.
(58) Bolhuis, A.; Tjalsma, H.; Stephenson, K.; Harwood, C. R.; Venema, G.; Bron, S.; van Dijl, J. M. *J. Biol. Chem.* **1999**, *274*, 15865.
(59) Bohni, P. C.; Deshaies, R. J.; Schekman, R. W. *J. Cell Biol.* **1988**, *106*, 1035.
(60) Behrens, M.; Michaelis, G.; Pratje, E. *Mol. Gen. Genet.* **1991**, *228*, 167.
(61) Yamagata, H.; Ippolito, C.; Inukai, M.; Inouye, M. *J. Bacteriol.* **1982**, *152*, 1163.
(62) Pollock, R. A.; Hartl, F. U.; Cheng, M. Y.; Ostermann, J.; Horwich, A.; Neupert, W. *EMBO J.* **1988**, *7*, 3493.
(63) Yaffe, M. P.; Schatz, G. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 4819.
(64) Dalbey, R. E.; Wickner, W. *J. Biol. Chem.* **1985**, *260*, 15925.
(65) van Dijl, J. M.; van den Bergh, R.; Reversma, T.; Smith, H.; Bron, S.; Venema, G. *Mol. Gen. Genet.* **1990**, *223*, 233.
(66) Wan, J.; Bringloe, D.; Lamppa, G. K. *Plant J.* **1998**, *15*, 459.
(67) Black, M. T.; Bruton, G. *Curr. Pharm. Des.* **1998**, *4*, 133.
(68) Hussain, M.; Ichihara, S.; Mizushima, S. *J. Biol. Chem.* **1980**, *255*, 3707.
(69) Zwizinski, C.; Wickner, W. *J. Biol. Chem.* **1980**, *255*, 7973.
(70) Tokunaga, M.; Loranger, J. M.; Wolfe, P. B.; Wu, H. C. *J. Biol. Chem.* **1982**, *257*, 9922.
(71) Jackson, R. C. *Methods Enzymol.* **1983**, *96*, 784.
(72) Strauss, A. W.; Zimmerman, M.; Boime, I.; Ashe, B.; Mumford, R. A.; Alberts, A. W. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 4225.
(73) Bohni, P. C.; Daum, G.; Schatz, G. *J. Biol. Chem.* **1983**, *258*, 4937.
(74) McAda, P. C.; Douglas, M. G. *J. Biol. Chem.* **1982**, *257*, 3177.
(75) Wickner, W.; Moore, K.; Dibb, N.; Geissert, D.; Rice, M. *J. Bacteriol.* **1987**, *169*, 3821.
(76) Chatterjee, S.; Suciu, D.; Dalbey, R. E.; Kahn, P. C.; Inouye, M. *J. Mol. Biol.* **1995**, *245*, 311.
(77) Gallagher, J.; Kaderbhai, N. N.; Kaderbhai, M. A. *Biochim. Biophys. Acta* **2001**, *1550*, 1.
(78) Dev, I. K.; Ray, P. H.; Novak, P. *J. Biol. Chem.* **1990**, *265*, 20069.

(79) Yaron, A.; Carmel, A.; Katchalski-Katzir, E. *Anal. Biochem.* **1979**, *95*, 228.
(80) Zhong, W.; Benkovic, S. J. *Anal. Biochem.* **1998**, *255*, 66.
(81) Rosse, G.; Kueng, E.; Page, M. G.; Schauer-Vukasinovic, V.; Giller, T.; Lahm, H. W.; Hunziker, P.; Schlatter, D. *J. Comb. Chem.* **2000**, *2*, 461.
(82) Peng, S. B.; Zheng, F.; Angleton, E. L.; Smiley, D.; Carpenter, J.; Scott, J. E. *Anal. Biochem.* **2001**, *293*, 88.
(83) Paetzel, M.; Dalbey, R. E.; Strynadka, N. C. *Pharmacol. Ther.* **2000**, *87*, 27.
(84) Carlos, J. L.; Paetzel, M.; Klenotic, P. A.; Strynadka, N. C.; Dalbey, R. E. In *Co-translational and Posttranslational Proteolysis of Proteins*, 3rd ed.; Dalbey, R. E., Sigman, D. S., Eds.; Academic Press: San Diego, CA, 2001; Vol. XXII.
(85) Green, N.; Fang, H.; Miles, S.; Lively, M. O. In *Co-translational and Posttranslational Proteolysis of Proteins*, 3rd ed.; Dalbey, R. E., Sigman, D. S., Eds.; Academic Press: San Diego, CA, 2001; Vol. XXII.
(86) Howe, C. J.; Floyd, K. A. In *Co-translational and Posttranslational Proteolysis of Proteins*, 3rd ed.; Dalbey, R. E., Sigman, D. S., Eds.; Academic Press: San Diego, CA, 2001; Vol. XXII.
(87) Date, T.; Wickner, W. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 6106.
(88) Wolfe, P. B.; Wickner, W.; Goodman, J. M. *J. Biol. Chem.* **1983**, *258*, 12073.
(89) Sung, M.; Dalbey, R. E. *J. Biol. Chem.* **1992**, *267*, 13154.
(90) Tschantz, W. R.; Sung, M.; Delgado-Partin, V. M.; Dalbey, R. E. *J. Biol. Chem.* **1993**, *268*, 27349.
(91) Paetzel, M.; Strynadka, N. C.; Tschantz, W. R.; Casareno, R.; Bullinger, P. R.; Dalbey, R. E. *J. Biol. Chem.* **1997**, *272*, 9994.
(92) Paetzel, M.; Dalbey, R. E.; Strynadka, N. C. *Nature* **1998**, *396*, 186.
(93) Wolfe, P. B.; Silver, P.; Wickner, W. *J. Biol. Chem.* **1982**, *257*, 7898.
(94) Tjalsma, H.; Noback, M. A.; Bron, S.; Venema, G.; Yamane, K.; van Dijl, J. M. *J. Biol. Chem.* **1997**, *272*, 25983.
(95) Meijer, W. J.; de Jong, A.; Bea, G.; Wisman, A.; Tjalsma, H.; Venema, G.; Bron, S.; van Dijl, J. M. *Mol. Microbiol.* **1995**, *17*, 621.
(96) Parro, V.; Schacht, S.; Anne, J.; Mellado, R. P. *Microbiology* **1999**, *145*, 2255.
(97) Cregg, K. M.; Wilding, I.; Black, M. T. *J. Bacteriol.* **1996**, *178*, 5712.
(98) Date, T. *J. Bacteriol.* **1983**, *154*, 76.
(99) Zhang, Y. B.; Greenberg, B.; Lacks, S. A. *Gene* **1997**, *194*, 249.
(100) Watts, C.; Wickner, W.; Zimmermann, R. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 2809.
(101) von Heijne, G. *Eur. J. Biochem.* **1983**, *116*, 17.
(102) Fikes, J. D.; Barkocy-Gallagher, G. A.; Klapper, D. G.; Bassford, P. J., Jr. *J. Biol. Chem.* **1990**, *265*, 3417.
(103) Shen, L. M.; Lee, J. I.; Cheng, S. Y.; Jutte, H.; Kuhn, A.; Dalbey, R. E. *Biochemistry* **1991**, *30*, 11775.
(104) Karamyshev, A. L.; Karamysheva, Z. N.; Kajava, A. V.; Ksenzenko, V. N.; Nesmeyanova, M. A. *J. Mol. Biol.* **1998**, *277*, 859.
(105) Nilsson, I.; von Heijne, G. *FEBS Lett.* **1992**, *299*, 243.
(106) Barkocy-Gallagher, G. A.; Bassford, P. J., Jr. *J. Biol. Chem.* **1992**, *267*, 1231.
(107) von Heijne, G. *Nucleic Acids Res.* **1986**, *14*, 4683.
(108) Carlos, J. L.; Paetzel, M.; Brubaker, G.; Karla, A.; Ashwell, C. M.; Lively, M. O.; Cao, G.; Bullinger, P.; Dalbey, R. E. *J. Biol. Chem.* **2000**, *275*, 38813.
(109) Jain, R. G.; Rusch, S. L.; Kendall, D. A. *J. Biol. Chem.* **1994**, *269*, 16305.
(110) Kuo, D. W.; Chan, H. K.; Wilson, C. J.; Griffin, P. R.; Williams, H.; Knight, W. B. *Arch. Biochem. Biophys.* **1993**, *303*, 274.
(111) Tschantz, W. R.; Paetzel, M.; Cao, G.; Suciu, D.; Inouye, M.; Dalbey, R. E. *Biochemistry* **1995**, *34*, 3935.
(112) Paetzel, M.; Strynadka, N. J. C.; James, M. N. G.; Dalbey, R. E. In *Proteolysis in Cell Functions*; Hopsu-Havu, V. K., Jarvinen, M., Kirschke, H., Eds.; IOS Press: Amsterdam, 1997; Vol. 13.
(113) Paetzel, M.; Chernaia, M.; Strynadka, N.; Tschantz, W.; Cao, G.; Dalbey, R. E.; James, M. N. *Proteins* **1995**, *23*, 122.
(114) Paetzel, M.; Dalbey, R. E.; Strynadka, N. C. *J. Biol. Chem.* **2002**, *277*, 9512.
(115) Klenotic, P. A.; Carlos, J. L.; Samuelson, J. C.; Schuenemann, T. A.; Tschantz, W. R.; Paetzel, M.; Strynadka, N. C.; Dalbey, R. E. *J. Biol. Chem.* **2000**, *275*, 6490.
(116) van Klompenburg, W.; Paetzel, M.; de Jong, J. M.; Dalbey, R. E.; Demel, R. A.; von Heijne, G.; de Kruijff, B. *FEBS Lett.* **1998**, *431*, 75.
(117) van den Brink-van der Laan, E.; Dalbey, R. E.; Demel, R. A.; Killian, J. A.; de Kruijff, B. *Biochemistry* **2001**, *40*, 9677.
(118) Paetzel, M.; Strynadka, N. C. *Protein Sci.* **1999**, *8*, 2533.
(119) Peat, T. S.; Frank, E. G.; McDonald, J. P.; Levine, A. S.; Woodgate, R.; Hendrickson, W. A. *Nature* **1996**, *380*, 727.
(120) Luo, Y.; Pfuetzner, R. A.; Mosimann, S.; Paetzel, M.; Frey, E. A.; Cherney, M.; Kim, B.; Little, J. W.; Strynadka, N. C. *Cell* **2001**, *106*, 585.
(121) Bell, C. E.; Frescura, P.; Hochschild, A.; Lewis, M. *Cell* **2000**, *101*, 801.
(122) Black, M. T. *J. Bacteriol.* **1993**, *175*, 4957.
(123) Paetzel, M.; Dalbey, R. E. *Trends Biochem. Sci.* **1997**, *22*, 28.
(124) Carlos, J. L.; Klenotic, P. A.; Paetzel, M.; Strynadka, N. C.; Dalbey, R. E. *Biochemistry* **2000**, *39*, 7276.
(125) van Dijl, J. M.; de Jong, A.; Venema, G.; Bron, S. *J. Biol. Chem.* **1995**, *270*, 3611.
(126) Zwizinski, C.; Date, T.; Wickner, W. *J. Biol. Chem.* **1981**, *256*, 3593.
(127) Black, M. T.; Munn, J. G.; Allsop, A. E. *Biochem. J.* **1992**, *282*, 539.
(128) Kuo, D.; Weidner, J.; Griffin, P.; Shah, S. K.; Knight, W. B. *Biochemistry* **1994**, *33*, 8347.
(129) Allsop, A. E.; Brooks, G.; Bruton, G.; Coulton, S.; Edwards, P. D.; Hatton, I. K.; Kaura, A. C.; McLean, S. D.; Pearson, N. D.; Smale, T. C.; Southgate, R. *Biorg. Med. Chem. Lett.* **1995**, *5*, 443.
(130) Allsop, A.; Brooks, G.; Edwards, P. D.; Kaura, A. C.; Southgate, R. *J. Antibiot. (Tokyo)* **1996**, *49*, 921.
(131) Evans, E. A.; Gilmore, R.; Blobel, G. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 581.
(132) Baker, R. K.; Lively, M. O. *Biochemistry* **1987**, *26*, 8561.
(133) YaDeau, J. T.; Klein, C.; Blobel, G. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 517.
(134) Meyer, H. A.; Hartmann, E. *J. Biol. Chem.* **1997**, *272*, 13159.
(135) Shelness, G. S.; Blobel, G. *J. Biol. Chem.* **1990**, *265*, 9512.
(136) Greenburg, G.; Shelness, G. S.; Blobel, G. *J. Biol. Chem.* **1989**, *264*, 15762.
(137) Dalbey, R. E.; Lively, M. O.; Bron, S.; van Dijl, J. M. *Protein Sci.* **1997**, *6*, 1129.
(138) van Dijl, J. M.; de Jong, A.; Vehmaanpera, J.; Venema, G.; Bron, S. *EMBO J.* **1992**, *11*, 2819.
(139) Fang, H.; Mullins, C.; Green, N. *J. Biol. Chem.* **1997**, *272*, 13152.
(140) Shelness, G. S.; Kanwar, Y. S.; Blobel, G. *J. Biol. Chem.* **1988**, *263*, 17063.
(141) Newsome, A. L.; McLean, J. W.; Lively, M. O. *Biochem. J.* **1992**, *282*, 447.
(142) Fang, H.; Panzner, S.; Mullins, C.; Hartmann, E.; Green, N. *J. Biol. Chem.* **1996**, *271*, 16460.
(143) Kalies, K. U.; Rapoport, T. A.; Hartmann, E. *J. Cell Biol.* **1998**, *141*, 887.
(144) Antonin, W.; Meyer, H. A.; Hartmann, E. *J. Biol. Chem.* **2000**, *275*, 34068.
(145) Shelness, G. S.; Lin, L.; Nicchitta, C. V. *J. Biol. Chem.* **1993**, *268*, 5201.
(146) Kalies, K. U.; Hartmann, E. *J. Biol. Chem.* **1996**, *271*, 3925.
(147) Van Valkenburgh, C.; Chen, X.; Mullins, C.; Fang, H.; Green, N. *J. Biol. Chem.* **1999**, *274*, 11519.
(148) Tjalsma, H.; Stover, A. G.; Driks, A.; Venema, G.; Bron, S.; van Dijl, J. M. *J. Biol. Chem.* **2000**, *275*, 25102.
(149) Lin, C.; Lindenbach, B. D.; Pragai, B. M.; McCourt, D. W.; Rice, C. M. *J. Virol.* **1994**, *68*, 5063.
(150) Wu, J. Z. *IUBMB Life* **2001**, *51*, 19.
(151) Amberg, S. M.; Rice, C. M. *J. Virol.* **1999**, *73*, 8083.
(152) Hongo, S.; Sugawara, K.; Muraki, Y.; Matsuzaki, Y.; Takashita, E.; Kitame, F.; Nakamura, K. *J. Virol.* **1999**, *73*, 46.
(153) Pekosz, A.; Lamb, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 13233.
(154) Pratje, E.; Mannhaupt, G.; Michaelis, G.; Beyreuther, K. *EMBO J.* **1983**, *2*, 1049.
(155) Pratje, E.; Guiard, B. *EMBO J.* **1986**, *5*, 1313.
(156) Schneider, A.; Behrens, M.; Scherer, P.; Pratje, E.; Michaelis, G.; Schatz, G. *EMBO J.* **1991**, *10*, 247.
(157) Schneider, A.; Oppliger, W.; Jeno, P. *J. Biol. Chem.* **1994**, *269*, 8635.
(158) Nunnari, J.; Fox, T. D.; Walter, P. *Science* **1993**, *262*, 1997.
(159) Jan, P. S.; Esser, K.; Pratje, E.; Michaelis, G. *Mol. Gen. Genet.* **2000**, *263*, 483.
(160) Bauerfeind, M.; Esser, K.; Michaelis, G. *Mol. Gen. Genet.* **1998**, *257*, 635.
(161) Esser, K.; Pratje, E.; Michaelis, G. *Mol. Gen. Genet.* **1996**, *252*, 437.
(162) Chaal, B. K.; Mould, R. M.; Barbrook, A. C.; Gray, J. C.; Howe, C. J. *J. Biol. Chem.* **1998**, *273*, 689.
(163) Chen, X.; Van Valkenburgh, C.; Fang, H.; Green, N. *J. Biol. Chem.* **1999**, *274*, 37750.
(164) Barbrook, A. C.; Packer, J. C.; Howe, C. J. *FEBS Lett.* **1996**, *398*, 198.
(165) Halpin, C.; Elderfield, P. D.; James, H. E.; Zimmermann, R.; Dunbar, B.; Robinson, C. *EMBO J.* **1989**, *8*, 3917.
(166) Shackleton, J. B.; Robinson, C. *J. Biol. Chem.* **1991**, *266*, 12152.
(167) Braun, V. *Biochim. Biophys. Acta* **1975**, *415*, 335.
(168) Mizuno, T. *J. Biochem. (Tokyo)* **1979**, *86*, 991.
(169) Yamagata, H.; Daishima, K.; Mizushima, S. *FEBS Lett.* **1983**, *158*, 301.
(170) Tokunaga, M.; Loranger, J. M.; Wu, H. C. *J. Biol. Chem.* **1983**, *258*, 12102.
(171) Yu, F.; Yamada, H.; Daishima, K.; Mizushima, S. *FEBS Lett.* **1984**, *173*, 264.

(172) Innis, M. A.; Tokunaga, M.; Williams, M. E.; Loranger, J. M.; Chang, S. Y.; Chang, S.; Wu, H. C. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 3708.
(173) Tokunaga, M.; Tokunaga, H.; Wu, H. C. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 2255.
(174) Hussain, M.; Ichihara, S.; Mizushima, S. *J. Biol. Chem.* **1982**, *257*, 5177.
(175) Dev, I. K.; Ray, P. H. *J. Biol. Chem.* **1984**, *259*, 11114.
(176) Dev, I. K.; Harvey, R. J.; Ray, P. H. *J. Biol. Chem.* **1985**, *260*, 5891.
(177) Tokunaga, M.; Loranger, J. M.; Wu, H. C. *J. Cell Biochem.* **1984**, *24*, 113.
(178) von Heijne, G. *J. Mol. Biol.* **1992**, *225*, 487.
(179) Munoa, F. J.; Miller, K. W.; Beers, R.; Graham, M.; Wu, H. C. *J. Biol. Chem.* **1991**, *266*, 17667.
(180) Manoil, C.; Beckwith, J. *Science* **1986**, *233*, 1403.
(181) Tjalsma, H.; Zanen, G.; Venema, G.; Bron, S.; van Dijl, J. M. *J. Biol. Chem.* **1999**, *274*, 28191.
(182) Neurath, H. In *Proteolytic Enzymes. A Practical Approach*; Beynon, R. J., Bond, J. S., Eds.; IRL Press: Cary, NC, 1989.
(183) Sankaran, K.; Wu, H. C. *J. Biol. Chem.* **1994**, *269*, 19701.
(184) Sankaran, K.; Wu, H. C. In *Signal Peptidases*; Heijne, G. v., Ed.; R. G. Landes Co.: Austin, TX, 1994.
(185) Pollitt, S.; Inouye, S.; Inouye, M. *J. Biol. Chem.* **1986**, *261*, 1835.
(186) Inouye, S.; Duffaud, G.; Inouye, M. *J. Biol. Chem.* **1986**, *261*, 10970.
(187) Strom, M. S.; Bergman, P.; Lory, S. *J. Biol. Chem.* **1993**, *268*, 15788.
(188) Strom, M. S.; Nunn, D. N.; Lory, S. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 2404.
(189) Nunn, D. N.; Lory, S. *J. Bacteriol.* **1993**, *175*, 4375.
(190) Pugsley, A. P. *Mol. Microbiol.* **1993**, *9*, 295.
(191) Pepe, J. C.; Lory, S. *J. Biol. Chem.* **1998**, *273*, 19120.
(192) LaPointe, C. F.; Taylor, R. K. *J. Biol. Chem.* **2000**, *275*, 1502.
(193) Maccecchini, M. L.; Rudin, Y.; Blobel, G.; Schatz, G. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 343.
(194) Maccecchini, M. L.; Rudin, Y.; Schatz, G. *J. Biol. Chem.* **1979**, *254*, 7468.
(195) Glaser, E.; Dessi, P. *J. Bioenerg. Biomembr.* **1999**, *31*, 259.
(196) Adamec, J.; Kalousek, F.; Isaya, G. In *Co-translational and Posttranslational Proteolysis of Proteins*, 3rd ed.; Dalbey, R. E., Sigman, D. S., Eds.; Academic Press: San Diego, CA, 2001; Vol. XXII.
(197) Ito, A. *Biochem. Biophys. Res. Commun.* **1999**, *265*, 611.
(198) Matsushita, Y.; Kitakawa, M.; Isono, K. *Mol. Gen. Genet.* **1989**, *219*, 119.
(199) Van der Pas, J. C.; Rohlen, D. A.; Weidner, U.; Weiss, H. *Biochim. Biophys. Acta* **1991**, *1089*, 389.
(200) Okamura, T.; John, M. E.; Zuber, M. X.; Simpson, E. R.; Waterman, M. R. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 5705.
(201) Karslake, C.; Piotto, M. E.; Pak, Y. K.; Weiner, H.; Gorenstein, D. G. *Biochemistry* **1990**, *29*, 9872.
(202) Thornton, K.; Wang, Y.; Weiner, H.; Gorenstein, D. G. *J. Biol. Chem.* **1993**, *268*, 19906.
(203) Hammen, P. K.; Gorenstein, D. G.; Weiner, H. *Biochemistry* **1994**, *33*, 8610.
(204) Hurt, E. C.; Muller, U.; Schatz, G. *EMBO J.* **1985**, *4*, 3509.
(205) Hurt, E. C.; Pesold-Hurt, B.; Suda, K.; Oppliger, W.; Schatz, G. *EMBO J.* **1985**, *4*, 2061.
(206) Haldi, M.; Guarente, L. *J. Biol. Chem.* **1989**, *264*, 17107.
(207) Nishi, T.; Nagashima, F.; Tanase, S.; Fukumoto, Y.; Joh, T.; Shimada, K.; Matsukado, Y.; Ushio, Y.; Morino, Y. *J. Biol. Chem.* **1989**, *264*, 6044.
(208) Maarse, A. C.; Van Loon, A. P.; Riezman, H.; Gregor, I.; Schatz, G.; Grivell, L. A. *EMBO J.* **1984**, *3*, 2831.
(209) Hurt, E. C.; Allison, D. S.; Muller, U.; Schatz, G. *J. Biol. Chem.* **1987**, *262*, 1420.
(210) Arretz, M.; Schneider, H.; Guiard, B.; Brunner, M.; Neupert, W. *J. Biol. Chem.* **1994**, *269*, 4959.
(211) Ogishima, T.; Niidome, T.; Shimokata, K.; Kitada, S.; Ito, A. *J. Biol. Chem.* **1995**, *270*, 30322.
(212) Niidome, T.; Kitada, S.; Shimokata, K.; Ogishima, T.; Ito, A. *J. Biol. Chem.* **1994**, *269*, 24719.
(213) Bohni, P., Gasser, S., Schatz, G., Kroon, A. M., Saccone, C., Eds. *The Organization and Expression of the Mitochondrial Genome*; Elsevier: Amsterdam, 1980; pp 423−433.
(214) Schmidt, B.; Wachter, E.; Sebald, W.; Neupert, W. *Eur. J. Biochem.* **1984**, *144*, 581.
(215) Conboy, J. G.; Fenton, W. A.; Rosenberg, L. E. *Biochem. Biophys. Res. Commun.* **1982**, *105*, 1.
(216) Miura, S.; Mori, M.; Amaya, Y.; Tatibana, M. *Eur. J. Biochem.* **1982**, *122*, 641.
(217) Hawlitschek, G.; Schneider, H.; Schmidt, B.; Tropschug, M.; Hartl, F. U.; Neupert, W. *Cell* **1988**, *53*, 795.
(218) Yang, M.; Jensen, R. E.; Yaffe, M. P.; Oppliger, W.; Schatz, G. *EMBO J.* **1988**, *7*, 3857.
(219) Ou, W. J.; Ito, A.; Okazaki, H.; Omura, T. *EMBO J.* **1989**, *8*, 2605.
(220) Kleiber, J.; Kalousek, F.; Swaroop, M.; Rosenberg, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 7978.
(221) Braun, H. P.; Emmermann, M.; Kruft, V.; Schmitz, U. K. *EMBO J.* **1992**, *11*, 3219.
(222) Braun, H. P.; Schmitz, U. K. *Trends Biochem. Sci.* **1995**, *20*, 171.
(223) Eriksson, A. C.; Sjoling, S.; Glaser, E. *J. Bioenerg. Biomembr.* **1996**, *28*, 285.
(224) Eriksson, A. C.; Glaser, E. *Biochim. Biophys. Acta* **1992**, *1140*, 208.
(225) Witte, C.; Jensen, R. E.; Yaffe, M. P.; Schatz, G. *EMBO J.* **1988**, *7*, 1439.
(226) Paces, V.; Rosenberg, L. E.; Fenton, W. A.; Kalousek, F. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5355.
(227) Emmermann, M.; Braun, H. P.; Arretz, M.; Schmitz, U. K. *J. Biol. Chem.* **1993**, *268*, 18936.
(228) Luciano, P.; Tokatlidis, K.; Chambre, I.; Germanique, J. C.; Geli, V. *J. Mol. Biol.* **1998**, *280*, 193.
(229) Adamec, J.; Gakh, O.; Spizek, J.; Kalousek, F. *Arch. Biochem. Biophys.* **1999**, *370*, 77.
(230) Braun, H. P.; Schmitz, U. K. *Eur. J. Biochem.* **1992**, *208*, 761.
(231) Rawlings, N. D.; Barrett, A. J. *Methods Enzymol.* **1995**, *248*, 183.
(232) Matthews, B. W. *Acc. Chem. Res.* **1988**, *21*, 333.
(233) Kitada, S.; Shimokata, K.; Niidome, T.; Ogishima, T.; Ito, A. *J. Biochem. (Tokyo)* **1995**, *117*, 1148.
(234) Striebel, H. M.; Rysavy, P.; Adamec, J.; Spizek, J.; Kalousek, F. *Arch. Biochem. Biophys.* **1996**, *335*, 211.
(235) Nagao, Y.; Kitada, S.; Kojima, K.; Toh, H.; Kuhara, S.; Ogishima, T.; Ito, A. *J. Biol. Chem.* **2000**, *275*, 34552.
(236) Shimokata, K.; Kitada, S.; Ogishima, T.; Ito, A. *J. Biol. Chem.* **1998**, *273*, 25158.
(237) Kojima, K.; Kitada, S.; Ogishima, T.; Ito, A. *J. Biol. Chem.* **2001**, *276*, 2115.
(238) Becker, A. B.; Roth, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 3835.
(239) Perlman, R. K.; Gehm, B. D.; Kuo, W. L.; Rosner, M. R. *J. Biol. Chem.* **1993**, *268*, 21538.
(240) Kitada, S.; Kojima, K.; Shimokata, K.; Ogishima, T.; Ito, A. *J. Biol. Chem.* **1998**, *273*, 32547.
(241) Becker, A. B.; Roth, R. A. *Biochem. J.* **1993**, *292*, 137.
(242) Xia, D.; Yu, C. A.; Kim, H.; Xia, J. Z.; Kachurin, A. M.; Zhang, L.; Yu, L.; Deisenhofer, J. *Science* **1997**, *277*, 60.
(243) Iwata, S.; Lee, J. W.; Okada, K.; Lee, J. K.; Iwata, M.; Rasmussen, B.; Link, T. A.; Ramaswamy, S.; Jap, B. K. *Science* **1998**, *281*, 64.
(244) Zhang, Z.; Huang, L.; Shulmeister, V. M.; Chi, Y. I.; Kim, K. K.; Hung, L. W.; Crofts, A. R.; Berry, E. A.; Kim, S. H. *Nature* **1998**, *392*, 677.
(245) Taylor, A. B.; Smith, B. S.; Kitada, S.; Kojima, K.; Miyaura, H.; Otwinowski, Z.; Ito, A.; Deisenhofer, J. *Structure (Camb)* **2001**, *9*, 615.
(246) Branda, S. S.; Isaya, G. *J. Biol. Chem.* **1995**, *270*, 27366.
(247) Isaya, G.; Kalousek, F.; Rosenberg, L. E. *J. Biol. Chem.* **1992**, *267*, 7904.
(248) Kalousek, F.; Isaya, G.; Rosenberg, L. E. *EMBO J.* **1992**, *11*, 2803.
(249) Chew, A.; Rollins, R. A.; Sakati, W. R.; Isaya, G. *Biochem. Biophys. Res. Commun.* **1996**, *226*, 822.
(250) Isaya, G.; Kalousek, F.; Rosenberg, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 8317.
(251) Isaya, G.; Miklos, D.; Rollins, R. A. *Mol. Cell Biol.* **1994**, *14*, 5603.
(252) Isaya, G.; Sakati, W. R.; Rollins, R. A.; Shen, G. P.; Hanson, L. C.; Ullrich, R. C.; Novotny, C. P. *Genomics* **1995**, *28*, 450.
(253) Chew, A.; Buck, E. A.; Peretz, S.; Sirugo, G.; Rinaldo, P.; Isaya, G. *Genomics* **1997**, *40*, 493.
(254) Bode, W.; Gomis-Ruth, F. X.; Huber, R.; Zwilling, R.; Stocker, W. *Nature* **1992**, *358*, 164.
(255) Dobberstein, B.; Blobel, G.; Chua, N. H. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 1082.
(256) Robinson, C.; Ellis, R. J. *Eur. J. Biochem.* **1984**, *142*, 337.
(257) Oblong, J. E.; Lamppa, G. K. *EMBO J.* **1992**, *11*, 4401.
(258) Gavel, Y.; von Heijne, G. *FEBS Lett.* **1990**, *261*, 455.
(259) Ostrem, J. A.; Ramage, R. T.; Bohnert, H. J.; Wasmann, C. C. *J. Biol. Chem.* **1989**, *264*, 3662.
(260) Pilon, M.; Wienk, H.; Sips, W.; de Swaaf, M.; Talboom, I.; van't Hof, R.; de Korte-Kool, G.; Demel, R.; Weisbeek, P.; de Kruijff, B. *J. Biol. Chem.* **1995**, *270*, 3882.
(261) Rensink, W. A.; Pilon, M.; Weisbeek, P. *Plant Physiol.* **1998**, *118*, 691.
(262) Rensink, W. A.; Schnell, D. J.; Weisbeek, P. J. *J. Biol. Chem.* **2000**, *275*, 10265.
(263) Archer, E. K.; Keegstra, K. *Plant Mol. Biol.* **1993**, *23*, 1105.
(264) Bassham, D. C.; Creighton, A. M.; Karnauchov, I.; Herrmann, R. G.; Klosgen, R. B.; Robinson, C. *J. Biol. Chem.* **1994**, *269*, 16062.
(265) Clark, S. E.; Lamppa, G. K. *J. Cell Biol.* **1991**, *114*, 681.
(266) Richter, S.; Lamppa, G. K. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 7463.
(267) Emanuelsson, O.; Nielsen, H.; von Heijne, G. *Protein Sci.* **1999**, *8*, 978.

(268) VanderVere, P. S.; Bennett, T. M.; Oblong, J. E.; Lamppa, G. K. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 7177.
(269) Ichihara, S.; Beppu, N.; Mizushima, S. *J. Biol. Chem*. **1984**, *259*, 9853.
(270) Pacaud, M. *J. Biol. Chem*. **1982**, *257*, 4333.
(271) Ichihara, S.; Suzuki, T.; Suzuki, M.; Mizushima, S. *J. Biol. Chem*. **1986**, *261*, 9405.
(272) Suzuki, T.; Itoh, A.; Ichihara, S.; Mizushima, S. *J. Bacteriol*. **1987**, *169*, 2523.
(273) Bolhuis, A.; Matzen, A.; Hyyrylainen, H. L.; Kontinen, V. P.; Meima, R.; Chapuis, J.; Venema, G.; Bron, S.; Freudl, R.; van Dijl, J. M. *J. Biol. Chem*. **1999**, *274*, 24585.
(274) Novak, P.; Ray, P. H.; Dev, I. K. *J. Biol. Chem*. **1986**, *261*, 420.
(275) Vimr, E. R.; Green, L.; Miller, C. G. *J. Bacteriol*. **1983**, *153*, 1259.
(276) Novak, P.; Dev, I. K. *J. Bacteriol*. **1988**, *170*, 5067.
(277) Conlin, C. A.; Miller, C. G. *J. Bacteriol*. **1992**, *174*, 1631.
(278) Conlin, C. A.; Trun, N. J.; Silhavy, T. J.; Miller, C. G. *J. Bacteriol*. **1992**, *174*, 5881.
(279) Parsell, D. A.; Lindquist, S. *Annu. Rev. Genet*. **1993**, *27*, 437.
(280) Georgopoulos, C.; Welch, W. J. *Annu. Rev. Cell Biol*. **1993**, *9*, 601.
(281) Yura, T.; Nagai, H.; Mori, H. *Annu. Rev. Microbiol*. **1993**, *47*, 321.
(282) Tisljar, U.; Barrett, A. J. *Biochem. J*. **1990**, *267*, 531.
(283) Lyko, F.; Martoglio, B.; Jungnickel, B.; Rapoport, T. A.; Dobberstein, B. *J. Biol. Chem*. **1995**, *270*, 19873.
(284) Klappa, P.; Dierks, T.; Zimmermann, R. *Eur. J. Biochem*. **1996**, *239*, 509.
(285) Martoglio, B.; Graf, R.; Dobberstein, B. *EMBO J*. **1997**, *16*, 6636.
(286) Braud, V. M.; Allan, D. S.; O'Callaghan, C. A.; Soderstrom, K.; D'Andrea, A.; Ogg, G. S.; Lazetic, S.; Young, N. T.; Bell, J. I.; Phillips, J. H.; Lanier, L. L.; McMichael, A. J. *Nature* **1998**, *391*, 795.
(287) Lemberg, M. K.; Bland, F. A.; Weihofen, A.; Braud, V. M.; Martoglio, B. *J. Immunol*. **2001**, *167*, 6441.
(288) Weihofen, A.; Lemberg, M. K.; Ploegh, H. L.; Bogyo, M.; Martoglio, B. *J. Biol. Chem*. **2000**, *275*, 30951.
(289) Weihofen, A.; Binns, K.; Lemberg, M. K.; Ashman, K.; Martoglio, B. *Science* **2002**, *296*, 2215.
(290) Ostersetzer, O.; Adam, Z. *Plant Cell* **1997**, *9*, 957.
(291) Shanklin, J.; DeWitt, N. D.; Flanagan, J. M. *Plant Cell* **1995**, *7*, 1713.
(292) Bolhuis, A.; Koetje, E.; Dubois, J. Y.; Vehmaanpera, J.; Venema, G.; Bron, S.; van Dijl, J. M. *Mol. Biol. Evol*. **2000**, *17*, 198.
(293) Selkoe, D. J. *J. Biol. Chem*. **1996**, *271*, 18295.
(294) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Res.* **1994**, *22*, 4673.
(295) Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucleic Acids Res*. **1997**, *25*, 3389.
(296) Bairoch, A.; Apweiler, R. *Nucleic Acids Res*. **2000**, *28*, 45.
(297) Tjalsma, H.; Bolhuis, A.; van Roosmalen, M. L.; Wiegert, T.; Schumann, W.; Broekhuizen, C. P.; Quax, W. J.; Venema, G.; Bron, S.; van Dijl, J. M. *Genes Dev*. **1998**, *12*, 2318.
(298) Kraulis, P. G. *J. Appl. Crystallogr*. **1991**, *24*, 946.
(299) Meritt, E. A.; Bacon, D. J. *Methods Enzymol*. **1997**, *277*, 505.
(300) Koshland, D.; Botstein, D. *Cell* **1982**, *30*, 893.
(301) Chang, C. N.; Blobel, G.; Model, P. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 361.
(302) Pragai, Z.; Tjalsma, H.; Bolhuis, A.; van Dijl, J. M.; Venema, G.; Bron, S. *Microbiology* **1997**, *143*, 1327.
(303) Nunn, D. N.; Lory, S. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 3281.
(304) Pugsley, A. P.; Dupuy, B. *Mol. Microbiol*. **1992**, *6*, 751.
(305) Waters, M. G.; Evans, E. A.; Blobel, G. *J. Biol. Chem*. **1988**, *263*, 6209.
(306) YaDeau, J. T.; Blobel, G. *J. Biol. Chem*. **1989**, *264*, 2928.
(307) Yaffe, M. P.; Ohta, S.; Schatz, G. *EMBO J*. **1985**, *4*, 2069.
(308) Kalousek, F.; Hendrick, J. P.; Rosenberg, L. E. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 7536.
(309) Isaya, G.; Kalousek, F.; Fenton, W. A.; Rosenberg, L. E. *J. Cell Biol*. **1991**, *113*, 65.
(310) Vierling, E.; Nagao, R. T.; DeRocher, A. E.; Harris, L. M. *EMBO J*. **1988**, *7*, 575.
(311) Werneke, J. M.; Zielinski, R. E.; Ogren, W. L. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 787.
(312) Smeekens, S.; van Binsbergen, J.; Weisbeek, P. *Nucleic Acids Res*. **1985**, *13*, 3179.
(313) Stein, R. L.; Barbosa, M. D.; Bruckner, R. *Biochemistry* **2000**, *39*, 7973.
(314) Caulfield, M. P.; Duong, L. T.; Baker, R. K.; Rosenblatt, M.; Lively, M. O. *J. Biol. Chem*. **1989**, *264*, 15813.

CR010166Y