이학박사 학위논문

Characterization of Signal Sequences and the Regulatory Subunits of the Signal Peptidase Complex

신호 서열의 특성과 신호 서열 절단 효소 조절 소단위에 관한 연구

2021 년 8 월

서울대학교 대학원

생명과학부

임채원

Characterization of Signal Sequences and the Regulatory Subunits of the Signal Peptidase Complex

by

Chewon Yim

Under the supervision of

Professor Hyun Ah Kim, Ph.D.

A thesis submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

August 2021

School of Biological Sciences

Seoul National University

신호 서열의 특성과 신호 서열 절단 효소 조절 소단위에 관한 연구

지도 교수 김 현 아

이 논문을 이학박사 학위논문으로 제출함
2021 년 8 월

서울대학교 대학원
생명과학부
임채원

임채원의 이학박사 학위논문을 인준함
2021 년 8 월

위 원 장 _____

부위원장 _____

위 원 _____

위 원 _____

위 원 _____

ABSTRACT

Proteins on the secretory pathway comprises of up to 40% of the eukaryotic proteome, illustrating the importance of understanding mechanisms for their biogenesis and quality control. In most cases, the N-terminal signal sequence guides a nascent chain to the Sec61 translocon in the endoplasmic reticulum (ER) membrane and opens the channel to initiate translocation, sometimes aided by the Sec62/Sec63 complex. Meanwhile, signal sequences may undergo the signal peptidase-mediated processing upon the ER translocation.

Signal sequences are surprisingly diverse in the sequence context, yet it remains unclear how these diverse sequences engage and open the general protein-conducting channel, Sec61 translocon. In the first part of the thesis, I undertook to assess signal sequences of varying features for their efficiencies in initiating translocation across the ER using *in vivo* radiolabeling approach in yeast. The data showed that increment in the N-terminal length impaired efficiency of signal sequence for translocation while a more hydrophobic core reestablished the translocation.

Yeast strains of defective translocation components were also assessed for handling of signal sequence variants. The Sec62/Sec63 complex handles moderately hydrophobic signal sequences, implying that these components are specifically required for such types of signal sequences. Hydrophobic internal signal sequences depended on Sec71. These results suggest that individual subunits of the Sec62/Sec63 complex, dynamically contribute to optimal initiation of translocation of proteins with diverse signal sequence features *in vivo*.

In the second part of the thesis, I undertook to explore roles of Spc1 and Spc2 on the processing of substrates by the signal peptidase complex. Using a set of model proteins and *in vivo* radiolabeling approach, I observed enhanced signal peptidase-mediated

processing without Spc1 or Spc2, which was reversely decreased by overexpressed Spc1. Further, Spc1 was shown to associate with membrane proteins and some membrane proteins were selectively reduced in cells lacking Spc1. These data suggest that Spc1 protects transmembrane segments from the signal peptidase-mediated processing, thereby regulates substrate selection for signal peptidase.

Taken together, this thesis shows that multiple translocation components including the signal recognition particle (SRP), the Sec62/Sec63 complex and the signal peptidase complex intricately handle secretory proteins having diverse types of signal sequences and transmembrane segments for proper biogenesis and quality control.

Keywords: Signal sequence, ER translocation, translocon, signal peptidase complex, Spc1, membrane protein

Student number: 2012-20325

TABLE OF CONTENTS

Chapter I - Introduction 1

1. The secretory pathway 2

1. 1. Proteins on the secretory pathway 2

1. 2. ER translocon components 3

1. 2. 1. The Sec61 translocon 3

1. 2. 2. The Sec62/Sec63 complex 4

2. Signal sequences 6

2. 1. Function of signal sequences 6

2. 2. Structural features/characteristics of signal sequences 10

3. Signal peptidase complex 14

4. Aims and experimental approaches 17

Chapter II - Materials and Methods 21

1. Yeast strains and culture conditions 22

2. Plasmid construction 23

3. *in vivo* pulse radiolabeling and pulse-chase 24

4. Immunoprecipitation and autoradiography 24

5. Tunicamycin treatment 25

6. Protein preparation, SDS-PAGE and Western blotting 25

7. Carbonate extraction 26

8. Co-immunoprecipitation 26

9. Prediction programs 27

10. Construction of SPCS1 knockout (KO) cell line and cell culture condition 28

Chapter III - Results 29

1. Systematic assessment of signal sequence characteristics and the required Sec translocon components 30

1. 1. *n* region of signal sequences affects the ER translocation efficiency 30

1. 2. *h* region of signal sequences affects the ER translocation efficiency 36

1. 3. Natural proteins tend to have negative relationship between the length of *n* region and the hydrophobicity of *h* region 40

1. 4. The required translocon components in accordance with signal sequence characteristics 43

1. 5. The requirement of the translocon components was assessed with natural proteins 47

2. Defining roles of the signal peptidase complex subunits in substrate processing 49

2. 1. The absence of Spc1 or Spc2 affects the signal peptidase-mediated processing of model membrane proteins 49

2. 2. Hydrophobicity of a TM segment determines the signal peptidase-mediated processing of model membrane proteins 56

2. 3. The absence of Spc1 or Spc2 increases the signal peptidase-mediated processing of single- and double-spanning model membrane proteins 60

2. 4. The signal peptidase complex lacking Spc1 and Spc2 recognizes the canonical cleavage site 62

2. 5. Overexpressed Spc1 reduces the signal peptidase-mediated processing of model membrane proteins 64

2. 6. Spc1 and Spc2 have overlapping but distinct functions in the signal peptidase-mediated processing 67

2. 7. Spc1 interacts with a TM segment of membrane proteins 69

2. 8. Abundance of some natural membrane proteins was reduced in the *spc1*Δ strain 72

2. 9. Overexpressed Sec11 compromises cell viability 77

2. 10. Human SPCS1 does not rescue the loss of Spc1 in yeast 80

2. 11. Depletion of SPCS1 causes differed processing of membrane proteins in mammalian cells 83

Chapter IV - Discussion 87

REFERENCES..... 105

국문 초록 116

LIST OF FIGURES AND TABLES

Figure 1. The heptameric Sec61 translocon complex in yeast

Figure 2. Two ER targeting modes

Figure 3. A schematic of a signal sequence

Figure 4. A substrate selection by the signal peptidase complex

Figure 5. The signal peptidase complex

Figure 6. A series of CPY-derived model proteins

Figure 7. Extensions to *n* region reduce the initial ER translocation efficiency

Figure 8. Increased hydrophobicity of *h* region increases the ER translocation efficiency

Figure 9. *h* region hydrophobicity determines the cleavage and the localization of proteins

Figure 10. Distribution of *n* region length and *h* region hydrophobicity of signal sequences in yeast proteins

Figure 11. Signal sequences of various characteristics require different combinations of the Sec62/Sec63 components and the SRP for efficient protein translocation

Figure 12. Yeast natural proteins of varying signal sequence requires different combination of the Sec62/Sec63 complex and the SRP

Figure 13. A series of *E. coli* leader peptidase-derived LepH2 model proteins

Figure 14. Spc1 and Spc2 modulate the signal peptidase-mediated processing of double-spanning membrane proteins LepH2

Figure 15. Spc1 and Spc2 modulate the signal peptidase-mediated processing of double-spanning membrane proteins LepH2 in BY4741 cells

Figure 16. A series of *E. coli* leader peptidase-derived LepCC model proteins are recognized and cleaved by the signal peptidase complex

Figure 17. The signal peptidase-mediated processing of LepCC variants is enhanced in the *spc1*Δ and *spc2*Δ strains

Figure 18. The signal peptidase complex lacking its subunits still processes the canonical cleavage site

Figure 19. Overexpressed Spc1 but not Spc2 modulates the signal peptidase-mediated processing of hydrophobic membrane proteins

Figure 20. Overexpressed Spc1 reduces the signal peptidase-mediated processing of LepCC variants

Figure 21. Overexpression of Spc1 or Spc2 does not complement the changed processing of LepH2 variants in deletion strains reciprocally

Figure 22. Overexpressed Spc1 interacts with model membrane proteins

Figure 23. Steady-state levels of natural single-spanning membrane proteins are reduced in the *spc1*Δ strain

Figure 24. Steady-state levels of natural multi-spanning membrane proteins are reduced in the *spc1*Δ strain

Figure 25. Overexpressed Sec11 severely compromises yeast cell viability

Figure 26. Human SPCS1 does not complement the loss of Spc1 in yeast in the processing of LepH2 model proteins

Figure 27. SPCS1 KO cells show differed processing of signal-anchored proteins

Figure 28. Schematics of translocation and processing components mediating signal sequences

Table 1. List of CPY variants and natural proteins used in this study

LIST OF ABBREVIATIONS

DMSO	Dimethyl sulfoxide
Endo H	Endoglycosidase H
ER	Endoplasmic reticulum
HA	Hemagglutinin
IP	Immunoprecipitation
KO	Knockout
Lep	Leader peptidase
OE	Overexprssion
RNC	Ribosome nascent chain complex
SRP	Signal recognition particle
TM	Transmembrane
ts	temperature sensitive
WT	Wild type

Chapter I - Introduction

1. The secretory pathway

1. 1. Proteins on the secretory pathway

The secretory pathway refers to the protein transport route: from the endoplasmic reticulum (ER), Golgi apparatus, lysosomes, vesicles in between and finally to the cell membrane and extracellular space (1). Proteins destined to the secretory pathway constitute up to 40% of the proteome in an organism; these include membrane proteins, luminal proteins in various membranous compartments and proteins secreted out of a cell (2, 3).

The vast majority of proteins on the secretory pathway start their journey entering the ER: a place where maturation, folding and post-translational modifications occur. Such transport to the ER is mostly mediated by the presence of N-terminal signal sequence, an alpha-helical hydrophobic segment of ~20-30 amino acids. Secretory proteins possess a cleavable signal peptide while membrane proteins utilize their first transmembrane (TM) as a signal sequence (4).

1. 2. ER translocon components

Reaching the ER membrane for intracellular transport, a nascent chain is accommodated by a subset of the protein-conducting channel proteins and other auxiliary proteins, in either a simultaneous or a sequential manner.

1. 2. 1. The Sec61 translocon

The Sec61 translocon is the major protein-conducting channel found in the ER membrane (Fig. 1A). This evolutionarily conserved protein complex is composed of three subunits: Sec61(SecY/Sec61α in prokaryotes/mammals), Sbh1(SecE/Sec61β) and Sss1(SecG/Sec61γ). Sec61 is a pore-forming main subunit harboring a lateral gate that opens the channel to the lipid phase. The lateral gate is formed by two helices TM2 and 7, while all 10 TMs form an hourglass shaped channel with a pore in the middle. Traversing through the pore, luminal domains of the ER-delivered polypeptides overcome the energy barrier and translocate across the lipid bilayer; hydrophobic segments laterally insert into the lipid bilayer through the lateral gate.

The lateral gate also functions as a signal sequence-binding site. Previous study has shown that residues across the signal sequence are cross-linked to the lateral gate in a periodic pattern, which implicates that a signal sequence binds to and positions within the lateral gate in a certain manner. Nevertheless, this binding does not reflect all signal sequence-binding cases as the study employed a cleavable, less hydrophobic signal sequence with no N-terminal soluble domain of a secretory protein, because signal sequences greatly vary in sequence context, hydrophobicity and cleavability that would affect binding events. In line with the observation, recent molecular dynamics simulation suggested that the hydrophobic core of less hydrophobic signal sequences positions differently compared to that of more hydrophobic signal sequences. Less hydrophobic

signal sequences are predicted to position themselves inside the pore of the translocon, whereas more hydrophobic signal sequences are in between the lateral gate helices, closer to the lipid phase. Thus, rather than unified binding, signal sequences engage the translocon in more complex context according to their characteristics, however the exact understanding of which is still an open question.

1. 2. 2. The Sec62/Sec63 complex

In *S. cerevisiae*, the Sec61 translocon associates with the tetrameric Sec62/Sec63 complex, consisting of Sec62, Sec63, Sec71 and Sec72. Sec62 and Sec63 are essential subunits in yeast and evolutionarily conserved (Figs. 1B and C) (5-8). They respectively have two and three TMs and associate with each other via their N-terminal and C-terminal domains. Sec63 contains a luminal J domain where the BiP ATPase Kar2 associates with; thereby incoming substrates are captured and prevented from sliding back to the cytoplasm (Fig. 1C) (9-11). By virtue of numerous functional and structural studies, roles of these essential subunits have begun to be revealed.

Sec71 is a single-spanning membrane protein that associates with the Hsp70-interacting Sec72, thereby anchoring cytoplasmic Sec72 in the Sec62/Sec63 complex (12, 13). Sec71 and Sec72 are not essential for growth under normal condition in yeast and do not exist in higher organisms. Albeit not numerous, previous studies concerning Sec71 and Sec72 showed that loss of these subunits causes defective translocation of a subset of secretory proteins, hinting at their yet unveiled roles.

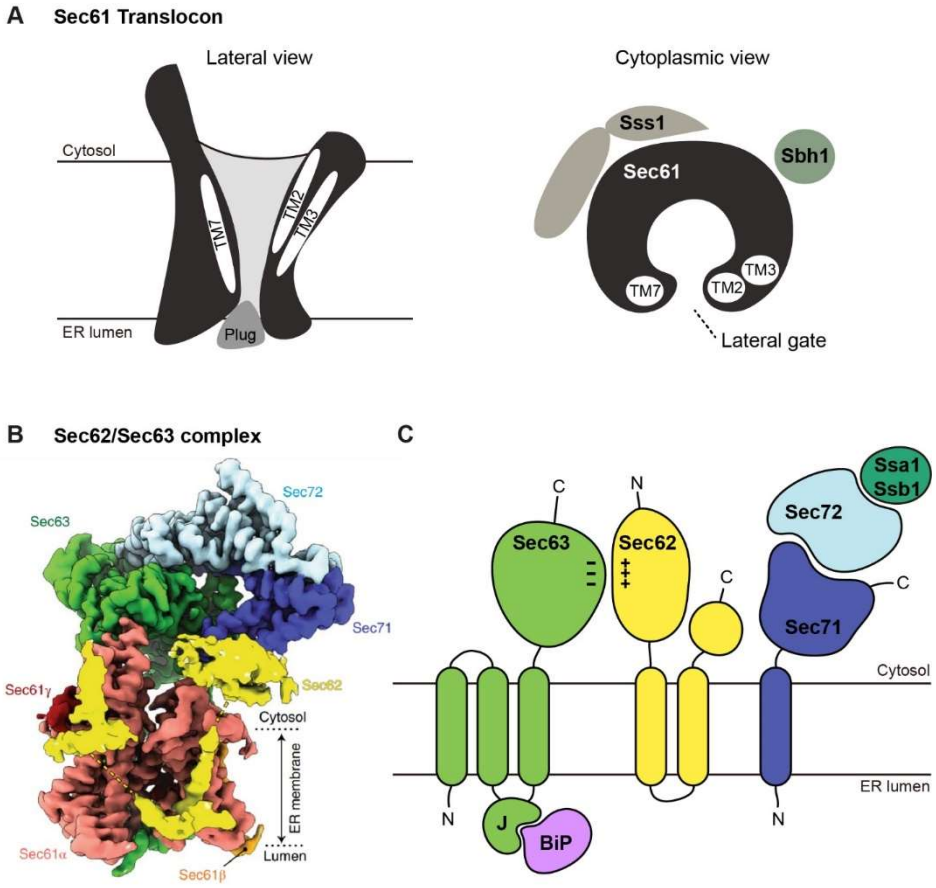

Figure 1. The heptameric Sec61 translocon complex in yeast. (A) A schematic of the trimeric Sec61 translocon. Sec61 is a major pore-forming subunit with a lateral gate composed of TM helices 2 and 7 and a plug domain closing the pore from the luminal side. Sbh1 and Sss1 are embedded in the membrane, adjacent to Sec61. (B) The cryo-EM structure of the Sec61 heptamer (PDB 7KAI) (14). (C) A schematic of the Sec62/Sec63 complex and associating proteins. The N-terminus of Sec62 and the C-terminus of Sec63 forms electrostatic interaction. The J-domain of Sec63 provide BiP binding site in the lumen. Sec72 is anchored to the ER membrane by interacting with Sec71 and forms contact with cytosolic Hsp70s.

2. Signal sequences

2. 1. Function of signal sequences

The pathways of ER targeting

There are two modes of ER targeting (Fig. 2). One is signal recognition particle (SRP)-dependent, where a signal sequence is recognized by the SRP shortly after a signal sequence is synthesized and threaded out of the ribosome (15). While translation transiently stops, the ribosome nascent chain (RNC)-SRP complex docks on the Sec61 translocon in the ER membrane via interaction between the SRP and the SRP receptor and then translation resumes. This SRP-dependent 'co-translational translocation' pathway guides substrates containing a highly hydrophobic signal sequence as hydrophobicity of signal sequence determines SRP binding affinity (16).

Signal sequences of less hydrophobicity are not recognized by SRP and nascent chains are fully translated in the cytoplasm (16). Released from the ribosome, a nascent chain is associated with cytosolic chaperones and guided to the ER membrane, which is termed as 'post-translational translocation' (17-19). Thus, for being a decisive factor for SRP dependence, a signal sequence determines which targeting pathway a nascent chain would take.

However, this classic concept of co- and post-translational translocation may not be delicate enough to illustrate complex translocation processes. Recent studies showed that marginally hydrophobic signal sequences are dependent on both the SRP and the post-translocon components (16, 20, 21). A ribosome profiling study by Jan *et al.*(22) provided evidence that the majority portion of proteins known to be post-translational translocation substrates are translated on the ER membrane-bound ribosomes, resembling the case of co-translational translocation substrates. The SRP targets RNCs to the ER membrane

before signal sequences are translated (23). Thus, the margin between co- and post-translocation pathways is not as clear as it has been perceived.

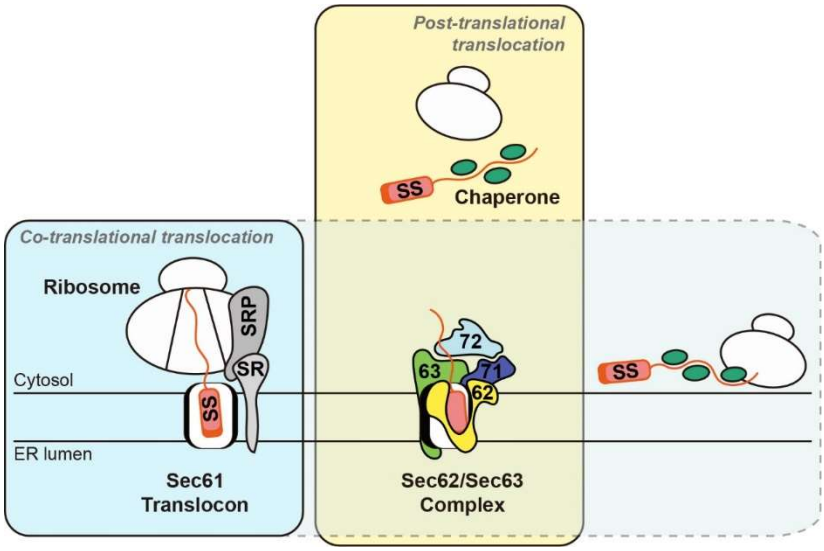

Figure 2. Two ER targeting modes. The co-translational translocation utilize the SRP and the SRP receptor to guide the RNC to the Sec61 translocon (Solid lined blue box). The post-translational translocation employs cytosolic chaperones to guide fully translated nascent chains to the Sec61 translocon, where the Sec62/Sec63 engage for efficient translocation (Yellow box). Yet underappreciated, some proteins require both the SRP and the Sec62/Sec63 complex and the many proteins are synthesized on the ER-bound ribosomes despite being known as post-translational translocation substrates (Dashed lined blue box). SS, signal sequence; SRP, signal recognition particle; SR, signal recognition particle receptor, 62, Sec62; 63, Sec63; 71, Sec71; 72, Sec72.

The initiation of translocation and involved translocon components

The Sec61 translocon remains 'closed' by the plug domain and the lateral gate during inactive idle-state, activation and opening of which necessitates initiation signals for translocation (24, 25). Such signals include intercalation of signal sequences between the lateral gate helices of the Sec61 translocon: the channel begins to open wide, followed by a series of sequential steps toward the 'open' state, involving association of other proteins in the translocation process (25-28).

In case of the co-translational translocation pathway, the ribosome recruited together with a nascent chain docks on the Sec61 translocon and induces channel opening (29). However, the ribosome is not able to open the Sec61 translocon wide enough and in case of the post-translational translocation pathway, the ribosome does not dock on the translocon, calling for other factors to fully open the translocon.

In line with that notion, recent structural studies proved the engagement of Sec62 and Sec63 in the initiation of the translocation process (14, 30). While a signal sequence binds to the lateral gate, Sec63 associates with and opens the Sec61 translocon wider. Moreover, Sec62 locates in close proximity to the lateral gate, inducing further opening (14, 30).

Although Sec71 was not clearly defined in the structure and its role remains vague, the ribosome profiling study implicated a specific role of Sec71 (22, 31, 32). The loss of Sec71 selectively affected internal signal sequences, suggesting that hairpin-forming internal signal sequences depend on Sec71 for their efficient translocation.

Given that signal sequences harbor various characteristics and accordingly require translocon components together and separately, collective evidence points to more exquisite involvement of translocon components than previously thought. Hence, detailed

and mechanistic exploration of translocation initiation process is requested to expand our current understanding.

2. 2. Structural features/characteristics of signal sequences

As aforementioned, a signal sequence is a ~20-30 amino acid-long α-helical hydrophobic segment, mostly found in the N-terminus of a protein (Fig. 3) (4, 33, 34). Signal sequences own distinctive features: N-terminal basic domain (*n* region), a ~7-13 residue-long hydrophobic core (*h* region) and C-terminal polar domain that contains a cleavage site (*c* region) (35).

Albeit such tripartite structure appears to be relatively simple, each region greatly varies in not only amino acid sequence but also the length and the biochemical traits: the hydrophobicity of the *h* region and charged residues in the *n* region. While up to 40% of the proteome contain signal sequences, signal sequences are readily interchangeable between proteins and organisms and even tolerant of a broad range of mutations, surprisingly lacking consensus sequence motif (36-38). An earlier study has shown that ~20% of random sequences was still able to direct secretion in yeast, clearly illustrating signal sequences' diversity in sequence context (38). Although a number of studies have revealed that variables across signal sequences affect signal sequence-mediated translocation, it is poorly understood whether signal sequences of various traits constitute uniform or multiform binding and why some signal sequences require different translocon components to initiate translocation.

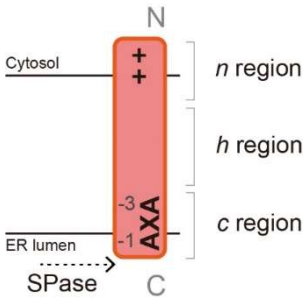

Figure 3. A schematic of a signal sequence. A signal sequence is comprised of tripartite structure: *n* region with positively charged residues, *h* region of hydrophobic core and *c* region containing a cleavage site by the signal peptidase complex. SPase, the signal peptidase complex.

In case of cleavable signal sequences, cleavage occurs by the signal peptidase complex upon ER translocation, a step of which is essential for protein maturation. A cleavage site is found at the C-terminal end of the *c* region, sometimes in tandem, following the '-3, -1 rule': small and neutral amino acids are present at the -3 and -1 positions relative to the cleavage site (Fig. 3) (35, 39). The presence of a cleavage site however does not always lead to cleavage (Fig. 4) (40). A previous study showed that the length of *h* region determines cleavage of the identical cleavage site, implying the decision of cleavage involves more than just the presence of cleavage sites (40). Yet cleavable and membrane-anchored signal sequences still share high homology and membrane-anchored signal sequences may readily possess potential cleavage sites, therefore this raises questions of discrimination principle and the underlying mechanism of sorting between cleavable signal sequences and membrane-anchored (uncleaved) signal sequences.

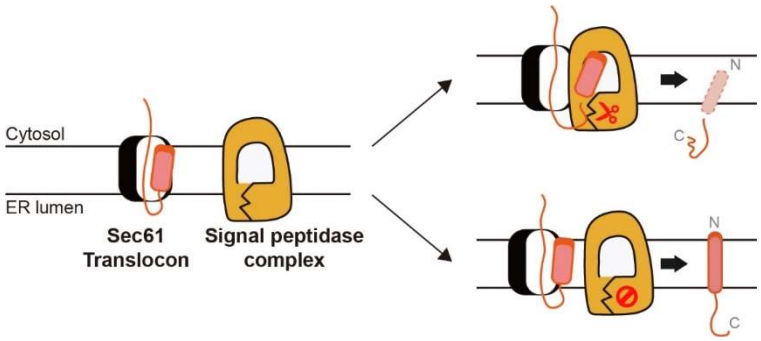

Figure 4. A substrate selection by the signal peptidase complex. Upon engaging the Sec61 translocon and initiating the ER translocation, a signal sequence encounters the signal peptidase complex (*left*). A signal sequence can be recognized and cleaved by the signal peptidase complex (*right, top*) or discriminated from the signal peptidase-mediated processing and inserted into the membrane (*right, bottom*).

3. Signal peptidase complex

Signal peptidase is a membrane-resident protease found across all living organisms; sequence comparison showed homology among leader peptidase of *E. coli*, Sec11 of yeast and SEC11A and SEC11C of mammals (41, 42). While leader peptidase is monomeric and located in the inner membrane, eukaryotic homologs are multimeric, forming a four-subunit complex in the ER membrane (Fig. 5). The four subunits of the signal peptidase complex in eukaryotes include Spc1/SPCS1, Spc2/SPCS2, Spc3/SPCS3 and Sec11/SEC11A and C (yeast/mammal), where Sec11 homologs carry the conserved catalytic domain (43-50). Sec11 and Spc3 are essential proteins in yeast, both of which singly span the ER membrane with a small cytoplasmic and a large luminal domain. A recently revealed cryo-EM structure of the human signal peptidase complex aids molecular understanding of the complex (Fig. 5A) (51). The luminal domain of Spc3 homolog grasps the catalytic core in the luminal domain of Sec11 homolog. This arrangement implies stabilizing role of Spc3, in line with previous data that Spc3 is essential for the maintenance of the functional signal peptidase complex and cell viability, despite a lack of the catalytic activity itself (41, 52).

Spc1 and Spc2 contribute the most of the cytoplasmic side of the signal peptidase complex. These two subunits span the ER membrane twice, harboring both termini facing the cytoplasmic side with a short loop in between (Fig. 5B). The structure showed that while forming a clamp-like structure with each other, Spc1 and Spc2 homologs respectively associate with Spc3 and Sec11 homologs only through TM segments. Together with the observation that Spc1 homolog does not contact with Sec11 homolog, such arrangement suggests that Spc1 and Spc2 might not be directly involved in the catalytic activity of the signal peptidase complex. Yet previous studies still implicate significance of Spc1 and Spc2: overexpressed Spc1 restores the impaired catalytic activity

of Sec11 mutant (49); loss of Spc1 homolog in *Drosophila* causes a developmental defect (53); loss of Spc2 results in reduced catalytic activity at elevated temperature both *in vivo* and *in vitro* (47, 54). Recent studies also showed that some surface receptor proteins and viral proteins necessitate Spc1 homolog for proper biogenesis, implying yet unappreciated roles of these auxiliary subunits (55). Despite so far findings, however, there has been a lack of systematic functional studies mining their roles and underlying mechanisms.

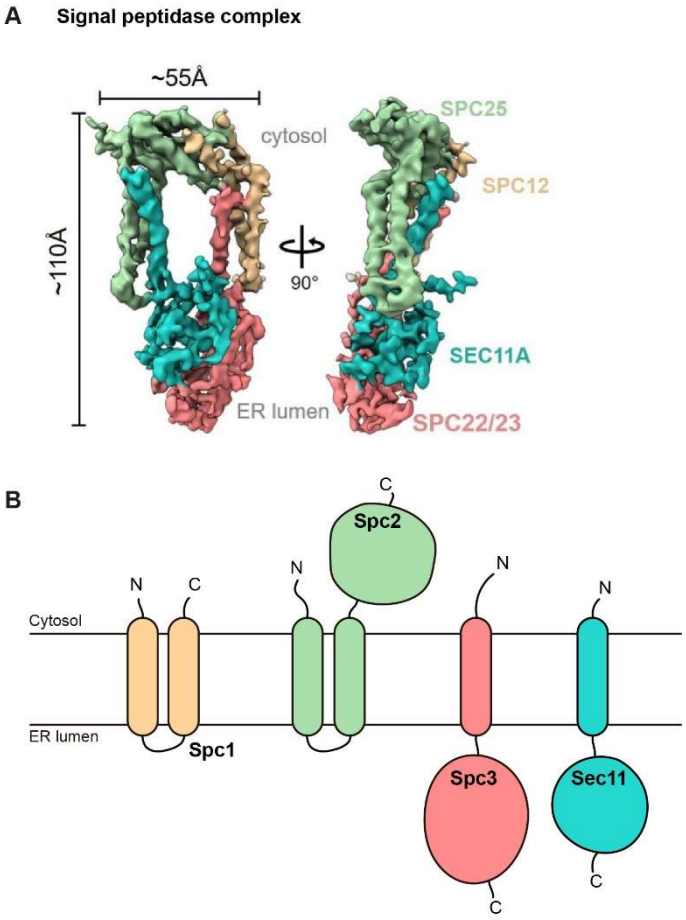

Figure 5. The signal peptidase complex. (A) A cryo-EM structure of the human signal peptidase complex (51). (B) A schematic of the signal peptidase complex subunits.

4. Aims and experimental approaches

Systematic assessment of signal sequence characteristics and the required Sec translocon components

It has been thought that there are two pathways for the ER transport: the SRP-dependent and the Sec62/Sec63 dependent pathways. However, increasing evidence suggests that some signal sequences show dependence on those components together or separately. Thus, the interplay between signal sequences and translocation-involved factors including the SRP and the Sec62/Sec63 complex is more intricate than previously thought. This study aimed to investigate their relationships, regarding the context of varying signal sequence features.

First, a set of model proteins carrying varying signal sequence features were prepared using yeast vacuole protein CPY. Features of signal sequences were systematically varied in two ways: the length of the N-terminus preceding the hydrophobic core (*n* region) and the hydrophobicity of the signal sequence core (*h* region). To assess the initial protein translocation efficiency, yeast cells expressing this series of model proteins were metabolically radiolabeled for 5 min and subjected to IP, SDS-PAGE and autoradiography. Using glycosylation state as a proxy for translocation, I analyzed how varied *n* and *h* regions of the signal sequence affect the initial ER translocation efficiency of CPY cargo protein.

Second, I investigated how signal sequences of varied *n* and *h* regions accordingly require different translocation components. To this end, the same experimental approach was applied: ER translocation efficiency of CPY variants was determined in SRP, Sec62, Sec63, Sec71 and Sec72 defective yeast strains.

I observed that increased *n* region length impaired efficient translocation of moderately hydrophobic signal sequences but not that of highly hydrophobic ones, indicating that moderately hydrophobic signal sequences must have short *n* region to be translocation-competent. Moderately hydrophobic signal sequences also required all four components of the Sec62/Sec63 complex for efficient translocation initiation whereas hydrophobic signal sequences were not except for Sec71. Surprisingly, hydrophobic signal sequences depended on Sec71 for efficient translocation initiation. These results suggest that different targeting and translocation components elaborately cooperate than previously thought, to sort signal sequence of heterogeneous characteristics for proper initiation of protein translocation *in vivo*.

Defining roles of the signal peptidase complex subunits in substrate processing

This study aimed to investigate functions of the signal peptidase complex subunits, mainly two subunits Spc1 and Spc2 (Fig. 5). It was surmised that these subunits might be involved in ER translocation/membrane insertion and signal sequence cleavage, as for the signal peptidase complex being one of the first translocation components and its long-established role in co-translocational processing. To this end, manipulation of each gene was carried out and consequent phenotypes, *i.e.* translocation and/or the signal peptidase-mediated processing, were monitored using a series of model proteins derived from *E. coli* leader peptidase (Lep). These model proteins contain the transmembrane (TM) segment of systematically varied hydrophobicity, strategically positioned C-terminal glycosylation sites and potential cleavage site by the signal peptidase. Thus, easy monitoring of protein targeting, membrane insertion and the signal peptidase-mediated cleavage were possible.

Taking advantage of such unique characteristics of the Lep-based model proteins, I assessed the roles of Spc1 and Spc2 in membrane insertion and signal peptidase-mediated cleavage. Lep model proteins were expressed in yeast strains lacking or overexpressing Spc1 or Spc2 and their membrane insertion and cleavage were assessed by metabolic radiolabeling. My data show that the loss of Spc1 increases the signal peptidase-mediated processing of Lep model proteins whereas overexpression of Spc1 decreases the signal peptidase-mediated processing. The increased signal peptidase-mediated processing in *spc1*Δ cells was not due to cleavage at non-canonical sites, showing that the signal peptidase lacking Spc1 does not process non-canonical cleavage sites; rather, substrate recognition is altered by the loss of Spc1.

Further, natural membrane proteins showed decreased steady state levels in cells lacking Spc1, implicating its role in protecting membrane proteins. Moreover, overexpressed Spc1

co-immunoprecipitated with membrane proteins with potential cleavage site, showing that Spc1 associates with substrates and protects their TM segments from the signal peptidase-mediated processing. Collectively, these data imply that Spc1 recognizes and protects membrane proteins from the signal peptidase-mediated processing, further implicating its role in quality control of membrane proteins.

Chapter II - Materials and Methods

1. Yeast strains and culture conditions

The *S. cerevisiae* haploid W303-1α (*MATα leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*) was used as wild type (WT) strain throughout the thesis, except for BY4741 (*MATa his3*Δ*1 leu2*Δ*0 met15*Δ*0 ura3*Δ*0*) as WT strain in Fig. 15.

The Sec62/Sec63 complex mutant and the signal peptidase mutant strains were constructed based on W303-1α: *sec62 35DDD* (*MATα sec62*Δ::*HIS3 leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15* pRS415 1kb upstream + *sec62 35DDD*), *sec63 A179T* (*MATα sec63*Δ::*HIS3 leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15* pRS415 1kb upstream + *sec63 A179T*), *sec71*Δ (*MATα sec71*Δ::*HIS3 leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*), *sec72*Δ (*MATα sec72*Δ::*HIS3 leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*), *spc1*Δ (*MATα spc1*Δ::*HIS3 leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*), *spc2*Δ (*MATα spc2*Δ::*HIS3 leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*) and *spc1*Δ*spc2*Δ (*MATα spc1*Δ::Hyg^R *spc2*Δ::*HIS3 leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*). The target gene was substituted with *HIS3* cassette amplified from pCgH by homologous recombination (56). For deletion of *SEC62* or *SEC63*, W303-1α was transformed with pRS416 1kb upstream + target gene and resulting transformant was subjected to gene deletion by *HIS3* substitution, followed by removal of pRS416 vector carrying the target gene by 5'-fluoroorotic acid selection.

The SRP *ts* strain *sec65-1* (*MATa leu2-3, -112 ade2 trpl-1 ura3-52 his3-11 sec65-1*) derived from (57). The signal peptidase activity *ts* strain *spc3-4* (*MATa spc3-4 ura3-52 leu2-3,112 his3-*Δ *200 trp1-*Δ*901 suc2-*Δ*9 lys2-80*) derived from (41).

The C-terminal hemagglutinin (HA)-tagging of *SNA3* and *PMP3* in the chromosomes of WT or *spc1*Δ strain was carried out by homologous recombination using the PCR

fragments amplified from the pFA6a-3HA-KanMx6 plasmid (58) with overhangs specific to the upstream and downstream sequences of the stop codon of *SNA3* or *PMP3*.

Yeast cells used in this study were grown in medium appropriately supplemented according to selectable markers at 30˚C except for *ts* mutant *sec65-1* and *spc3-4* strains that were grown at 24˚C and shifted to 37˚C before experimental procedures. For galactose induction of the signal peptidase subunits, cells were grown in 2% glucose medium overnight and spotted onto 2% galactose medium.

2. Plasmid construction

PRC1 (*CPY*) was amplified from W303-1α genome with primers containing homologous regions for pRS424GPD*HA* vector and co-transformed with SmaI-digested vector into W303-1α for homologous recombination (21). On this plasmid (pRS424GPD*CPYHA*) as template, a SmaI site was introduced between the 2nd and the 3rd residues of CPY by site-directed mutagenesis (Toyobo) following manufacture's protocols to fuse in the N-terminal residues of Dap2 or Sec71, amplified from W303-1α genome to generate pRS424GPD*D27CPY(1.9)HA* and pRS424GPD*S30CPY(1.5)HA* vectors. All other CPY variant plasmids were created by site-directed mutagenesis to truncate these two vectors. CPYt constructs were created based on this series of CPY plasmids by truncating after the 323rd residue and signal sequences were further modified by site-directed mutagenesis.

LepH2 variants in a yeast vector were constructed in (59). *E. coli* Lep-derived LepCC constructs (40) were subcloned from the pGEM4z vector into the yeast pRS424GPD*HA* vector by PCR amplification and homologous recombination as above. LepCCt constructs were created by truncating the N-terminal 20 residues except for the start methionine of

pRS424GPD*LepCCHA* constructs. The pRS426GPD*FLAG* vector containing *SPC1* was cloned by the Gibson assembly (60). All plasmids were confirmed by DNA sequencing.

The genes encoding natural membrane proteins (*KTR3*, *GDA1*, *PHO8t* and *ERV46t*) were amplified from W303-1α genome by PCR, subcloned into the p424GPD*HA* (59) and the plasmid pRS415*gHA* with a C-terminal glycosylatable HA tag (*MEP2* under its own promoter).

3. *in vivo* pulse radiolabeling and pulse-chase

For all the yeast strains except for *sec63 A179T*, s*ec65-1* and *spc3-4*, *in vivo* pulse radiolabeling with [^{35}S]Met as follows: cells were grown overnight and 1.5 OD_{600} units of cells were harvested. After washed once with 1 ml of –Met medium, cells were resuspended in 1 ml of –Met medium, starved for 15-30 min at 30˚C and harvested, followed by resuspension with 150 µl and incubation with [^{35}S]Met for indicated time. *sec63 A179T* cells were starved for 15 min at non-permissive temperature 37˚C. *sec65-1* cells were shifted to 37˚C for 30 min prior to starvation, which was done at 37˚C for 15 min. *spc3-4* cells were starved at 37˚C for 30 min and radiolabeled at 37˚C. Radiolabeling was stopped and chased by the addition of 750 µl of ice-cold stop solution buffer.

For pulse-chase experiments, 1.5 OD_{600} units of cells were harvested for each time point and resuspended in –Met medium of the volume corresponding to the number of chase time points. Radiolabeling was stopped and chased by the addition of 50 µl of 200 mM cold-Met solution per 1.5 OD_{600} units of cells for each time point. The reaction was stopped by transferring 1.5 OD_{600} units of cells to 750 µl of ice-cold stop solution buffer.

4. Immunoprecipitation and autoradiography

Radiolabeled cell pellets were resuspended in 100 µl of lysis buffer (20 mM Tris-Cl, pH 7.5, 1% SDS, 1 mM DTT, 1 mM PMSF, and 1X Protease Inhibitor Cocktail) and vortexed with 100 µl of ice-cold glass beads for 2 min twice, keeping the samples on ice for 1 min in-between. Subsequently, the samples were incubated at 65°C for 10 min and centrifuged. Supernatant fractions were mixed with 500 µl of immunoprecipitation buffer (15 mM Tris-Cl, pH 7.5, 0.1% SDS, 1% Triton X-100, and 150 mM NaCl), 1 µl of anti-HA antibody (Biolegend) and 20 µl of prewashed protein G-agarose beads (Thermo Fisher Scientific, Pierce) and rotated at room temperature for 3h or 4˚C overnight. The agarose beads were washed twice with immunoprecipitation buffer, once with ConA buffer (500 mM NaCl, 20 mM Tris-Cl, pH 7.5 and 1% Triton X-100) and once with buffer C (50 mM NaCl and 10 mM Tris-Cl, pH 7.5). The beads were incubated with 40-60 µl of SDS sample buffer (50 mM DTT, 50 mM Tris-Cl (pH 7.6), 5% SDS, 5% glycerol, 50 mM EDTA, 1 mM PMSF, 1X Protease Inhibitor Cocktail (Quartett) and bromophenol blue) at 65°C for 10 min or 95°C for 5 min, followed by Endo H (New England Biolabs) treatment at 37°C for more than 10 min if stated. Protein samples were subjected to SDS-PAGE and gels were dried for autoradiography, visualized by Typhoon™ FLA 7000 and quantified by the MultiGaugeV3.0 software.

5. Tunicamycin treatment

During starvation prior to radiolabeling, cells were supplemented with tunicamycin (Sigma, 100 µg·ml-1) or DMSO as control. The concentration of tunicamycin was maintained throughout the radiolabeling procedure.

6. Protein preparation, SDS-PAGE and Western blotting

For protein preparation of yeast cells, 2-10 OD_{600} unit of cells expressing a protein of interest were harvested after incubation at 30°C overnight unless stated otherwise. Cells

were mixed with 30-150 ul SDS sample buffer depending on cell amount (50 mM Tris-Cl, pH 7.5, 5% SDS, 5% glycerol, 50 mM EDTA, pH 8, 50 mM DTT, 1X protease inhibitor cocktail (Quartett, Germany, PPI1015), 1 mM PMSF) and incubated at 60°C for 10 min or at 95°C for 5 min. Samples were subjected to SDS-PAGE and Western blotting by mouse anti-HA antibody (Biolegend) unless stated otherwise in the figure legends. Western blotting data were visualized using ChemiDoc™ XRS+ and resulting data were processed by Image Lab™ software.

7. Carbonate extraction

Five OD_{600} units of cells were lysed by vortexing with lysis buffer (20 mM Tris-Cl, pH 8.0, 10 mM EDTA, pH 8.0, 100 mM NaCl, 300 mM sorbitol, 1 mM PMSF, 1×protease inhibitor) and glass beads for 10 min at 4°C. Cell lysate was subjected to centrifugation (22,000×*g*, 10 s, 4°C) to remove cell debris and transferred to a new pre-chilled tube (a portion was saved for 'Total' fraction), followed by centrifugation (22,*000*×*g*, 30 min, 4°C). Resulting membrane pellet was washed once with lysis buffer and incubated with 0.1 M Na_2CO_3 (pH 11.5) for 30-60 min on ice. Trichloroacetic acid (TCA)-precipitated 'Total', 'Supernatant' and 'Pellet' fractions were centrifuged (22,*000*×*g*, 15 min, 4°C) and washed with acetone. Samples were resuspended in SDS sample buffer and analyzed by SDS-PAGE and Western blotting.

8. Co-immunoprecipitation

Experimental procedures are based on (Zhang et al., 2017) with following modifications. Crude membrane was isolated from about 15 OD_{600} units of cells and solubilized with 400 µl of lysis buffer (50 mM HEPES-KOH/PBS, pH 6.8, 1% Triton X-100, 150 mM KOAc, 2 mM Mg(OAc)2, 1mM CaCl2, 15% glycerol, 1x PIC, 2 mM PMSF) by rotation at 4°C for 1h. After centrifugation at 14800 rpm for 30 min at 4°C, soluble fraction was

transferred to a tube containing 25 µl of protein G-agraose beads prewashed three times with lysis buffer, followed by rotation for 30 min at 4°C. Beads were removed by quick centrifugation and 15 µl of the lysate was saved for 'Input' fraction while all the remaining supernatant was transferred to a new tube containing 25 µl of prewashed protein G-agraose beads and 1 µl of anti-FLAG mouse antibody (FUJIFILM Wako Pure Chemical Corporation). The immunoprecipitation mixture was rotated for about 4h at 4°C. Beads were washed three times with lysis buffer and sampled by incubation with 40 µl of SDS sample buffer for 15 min at 55°C as 'IP' fraction. 'Input' fraction was mixed with 65 µl of SDS sample buffer and incubated for 15 min at 55°C.

9. Prediction programs

Prediction of the length of *n* regions and the hydrophobicity of *h* regions of signal sequences were calculated by the Δ*G* predictor (http://dgpred.cbr.su.se/) (Hessa, 2007). By submitting full protein sequence of each protein to "full protein scan", the ΔG_{app} value was obtained and the number of N-terminal amino acids preceding the signal sequence was counted. For bioinformatics analysis of yeast secretory and signal-anchored proteins, sequence information of the proteins were extracted from the UniProtKB Protein Knowledgebase by querying "signal peptide" or "signal anchor", respectively, and "*Saccharomyces cerevisiae*". The downloaded data set was combined with the list containing signal peptides and signal anchored sequences in (Reithinger, 2013). Protein sequences of the combined list were run on the Δ*G* predictor, and the position, amino acid sequences and the ΔG_{app} values of signal sequences were collected. Proteins predicted to localize to other organelles, to have more than two TMDs (Kim, 2006) or not to have a predicted signal sequence within the first 60 amino acids were manually sorted and excluded.

Prediction of cleavage sites was carried out by the Signal P-5.0 server (http://www.cbs.dtu.dk/services/SignalP/) (Armenteros, 2019). Full protein sequence of each protein was submitted except natural membrane proteins in Figs. 23 and 24 where sequences starting from TM segments were submitted.

10. Construction of SPCS1 knockout (KO) cell line and cell culture condition

sgRNA constructs for CRISPR/Cas9-mediated genome editing were created by the insertion of synthetic DNA oligos targeting three different regions within *SPCS1* into the BsmB1 site of the lentiGuide-puro vector: sgSPCS1#1, forward 5'-caccgATGGCGCGGGGCGGGGACAC-3' and reverse 5'-aaacGTGTCCCCGCCCCGCGCCATc-3'; sgSPCS1#2, forward 5'-caccgCACGGGCTGTACCGGCCCGT-3' and reverse 5'-aaacACGGGCCGGTACAGCCCGTGc-3'; sgSPCS1#3, 5'- caccgCG GAAGCGGAAGTCTCCGAC-3' and 5'- aaacGTCGGAGACTTCCGCTTCCGc-3'.

Flp-In™ T-REx™ 293 cells (Invitrogen) were used as WT cell line and grown in Dulbecco's modified Eagle medium supplemented with 10% fetal bovine serum in 5% CO_2 at 37˚C. To generate SPCS1 KO cell line, cells were co-transfected with sgSPCS1-bearing lentiGuide-puro vectors and pcDNA-Cas9-G418 using Lipofectamine 2000 (Thermo Fisher Scientific) and selected in medium containing G418. SPCS1 KO was confirmed by Western blotting against SPCS1.

Chapter III - Results

1. Systematic assessment of signal sequence characteristics and the required Sec translocon components

1. 1. *n* region of signal sequences affects the ER translocation efficiency

To evaluate the relationship between the length of the *n* region and translocation efficiency, test signal sequences of varying *n* region length were engineered on well-known secretory protein CPY (YMR297W; Fig. 6A). Translocation was monitored by N-linked glycosylation state, process of which occurs in the ER lumen and adds ~2 kDa to the protein mass (Fig. 6B). Glycosylated forms therefore refer to translocated forms and can be easily distinguished from untranslocated and unglycosylated forms on the SDS-PAGE for their slower migration. Glycosylation state can be confirmed by Endoglycosidase H (Endo H) treatment that removes glycan residues from proteins.

Two sets of *n* variants respectively carry gradual extensions fused to the N-terminus of CPY: one extension derived from N-terminal residues of Dap2 (YHR028C; D CPYs) and the other from Sec71 (YBR171W; S CPYs) (Fig. 6A and Table 1). Dap2 and Sec71 are single-spanning ER membrane proteins of opposite membrane topologies, the former orients as N_{cyto}-C_{lumen} and the latter as N_{lumen}-C_{cyto}, hence possible residue-specific effects on translocation could be monitored.

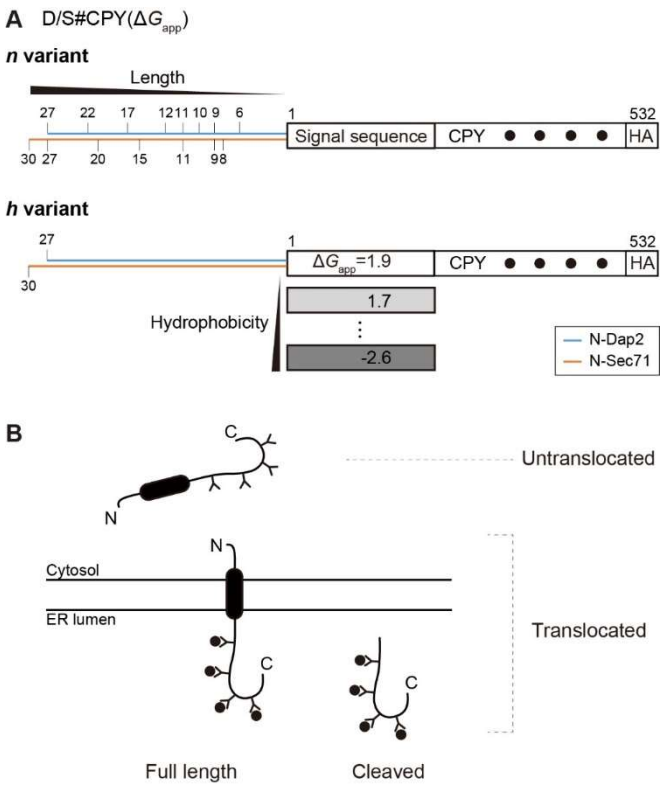

Figure 6. A series of CPY-derived model proteins. (A) A schematic of D/S#CPY(ΔG_{app}) *n* region (*top*) and *h* region (*bottom*) variants (N-Dap2, blue; N-Sec71, orange). The number of *n* region residues and the ΔG_{app} value of *h* region are indicated. Filled circles indicate N-linked glycosylation sequons. Three copies of HA epitope are fused at the C-terminus. (B) Possible membrane topology of D/S#CPY(ΔG_{app}) variants. Untranslocated form is not glycosylated whereas translocated forms are glycosylated and may undergo the signal peptidase-mediated processing. 'Y' indicates glycosylation sites and filled circles indicate glycan moieties.

Name	N-terminal sequence	ΔG_{app}[a]	*n* region length
CPY	MKAFTSLLCGLGLSTTLAKAISL	1.299	0
D27CPY(1.9)	MEGGEEEVERIPDELFDTKKKHLLDKKAF**S**SLLCGLGLSTTLAKAISL	1.921	27
D22CPY(1.9)	MEVERIPDELFDTKKKHLLDKKAF**S**SLLCGLGLSTTLAKAISL	1.921	22
D17CPY(1.9)	MPDELFDTKKKHLLDKKAF**S**SLLCGLGLSTTLAKAISL	1.921	17
D12CPY(1.9)	MDTKKKHLLDKKAF**S**SLLCGLGLSTTLAKAISL	1.921	12
D11CPY(1.9)	MTKKKHLLDKKAF**S**SLLCGLGLSTTLAKAISL	1.921	11
D10CPY(1.9)	MKKKHLLDKKAF**S**SLLCGLGLSTTLAKAISL	1.921	10
D9CPY(1.9)	MKKHLLDKKAF**S**SLLCGLGLSTTLAKAISL	1.921	9
D6CPY(1.9)	MLLDKKAF**S**SLLCGLGLSTTLAKAISL	1.921	6
D1CPY(1.3)	MLKAF**S**SLLCGLGLSTTLAKAISL	1.334	1
S30CPY(1.5)	MSEFNETKFSNNSTFFETEEPIVETKSISKAFTSLLCGLGLSTTLAKAISL	1.465	30
S27CPY(1.5)	MNETKFSNNSTFFETEEPIVETKSISKAFTSLLCGLGLSTTLAKAISL	1.465	27
S20CPY(1.5)	MNSTFFETEEPIVETKSISKAFTSLLCGLGLSTTLAKAISL	1.465	20
S15CPY(1.5)	METEEPIVETKSISKAFTSLLCGLGLSTTLAKAISL	1.465	15
S12CPY(1.5)	MEPIVETKSISKAFTSLLCGLGLSTTLAKAISL	1.465	12
S11CPY(1.5)	MPIVETKSISKAFTSLLCGLGLSTTLAKAISL	1.465	11
S9CPY(1.5)	MPIVESISKAFTSLLCGLGLSTTLAKAISL	1.465	9
S8CPY(1.5)	METKSISKAFTSLLCGLGLSTTLAKAISL	1.465	8
D27CPY(1.7)	MEGGEEEVERIPDELFDTKKKHLLDKKAF**S**SLLC**A**LGLSTTLAKAISL	1.656	27
D26CPY(1.3)	MEGGEEEVERIPDELFDTKKKHLLDKKAF**S**SLLC**L**LGLSTTLAKAISL	1.253	26
D27CPY(0.5)	MEGGEEEVERIPDELFDTKKKHLLDKKAF**S**SLLC**A**L**L**LSTTLAKAISL	0.506	27
D26CPY(0.1)	MEGGEEEVERIPDELFDTKKKHLLDKKAF**S**SLLC**L**L**L**LSTTLAKAISL	0.098	26
D22CPY(-0.5)	MEGGEEEVERIPDELFDTKKKHLLDKLAF**S**SLLCALLLSTTLA	-0.487	22
D22CPY(-1.0)	MEGGEEEVERIPDELFDTKKKHLLDK**L**AF**S**SLLC**L**L**L**LSTTLA	-0.95	22

D22CPY(-2.6)	MEGGEEEVERIPDELFDTKKKHLLDK**LLL**T**L**LLC**LL**LSTTLA	-2.615	22
S30CPY(1.7)	MSEFNETKFSNNSTFFETEEPIVETKSISKAF**S**SLLC**A**LGLSTTLAKAISL	1.656	30
S29CPY(1.3)	MSEFNETKFSNNSTFFETEEPIVETKSISKAF**S**SLLC**L**LGLSTTLAKAISL	1.253	29
S30CPY(0.5)	MSEFNETKFSNNSTFFETEEPIVETKSISKAF**S**SLLC**A**LLLSTTLAKAISL	0.506	30
S29CPY(0.1)	MSEFNETKFSNNSTFFETEEPIVETKSISKAF**S**SLLC**LL**LSTTLAKAISL	0.098	29
S27CPY(-0.9)	MSEFNETKFSNNSTFFETEEPIVETKSIS**L**AF**S**SLLC**LL**LSTTLA	-0.901	27
CPY(-0.6)	MKAFTSLLC**LL**LSTTLAKAISL	-0.551	0
CPY(-1.9)	MK**LL**T**L**LLC**LL**LSTTLAKAI	-1.850	0
Spc3	MFSFVQRFQNVSNQAFSMGIVMVVFIMASSYYQLI	0.337	14
Dap2	MEGGEEEVERIPDELFDTKKKHLLDKLIRVGIILVLLIWGTVLLL	-2.932	25

Table 1. List of CPY variants and natural proteins used in this study. CPY variants used in this study are listed. *n* region length indicates the number of amino acids preceding the signal sequence of CPY whose sequences and ΔG_{app} (kcal/mol) values are also shown. Predicted signal sequence is underlined and mutations within the signal sequence are marked in bold.

The ΔG_{app} and position of the signal sequence in CPY variants were predicted by the ΔG predictor (http://dgpred.cbr.su.se/). Mutated residues are marked in bold. Predicted signal sequences are underlined. Note that ΔG_{app} predicted the position of hydrophobic segment slightly differently depending on the sequence context.

Translocation efficiency of these constructs was assessed *in vivo* by radiolabeling WT yeast transformants with [^{35}S]Met for 5 min, capturing the initial translocation status (Fig. 7A). Glycosylated and unglycosylated species of CPY *n* variants were separated by SDS-PAGE (Fig. 7A) and glycosylation efficiency was plotted against the length of *n* region (Fig. 7B). Glycosylated species were detected with *n* region extension < 12 residues for both D and S CPY sets, indicating that long *n* region extension inhibits the ER translocation of CPY regardless of their original sequence compositions or orientation. Within the *n* region, the length dominates over the contributions of individual or overall sequence composition in the ER translocation of CPY variants. Thus, these results suggest that the length of the *n* region greatly affects the ER translocation.

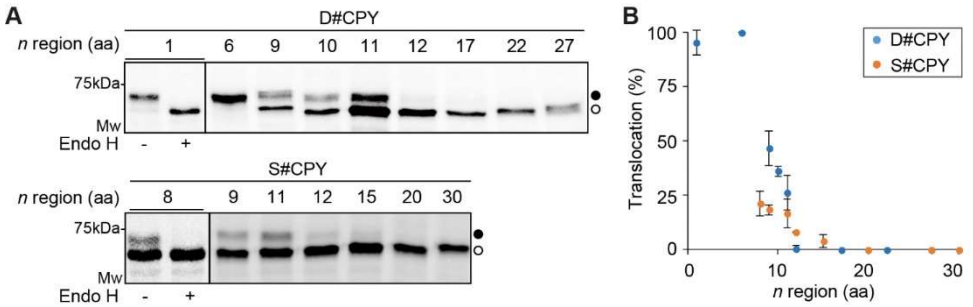

Figure 7. Extensions to *n* region reduce the initial ER translocation efficiency. (A) WT cells expressing indicated D#CPY (*top*) or S#CPY (*bottom*) was radiolabeled with S^{35}[MET] for 5 min. Radiolabeled D/S#CPY was immunoprecipitated using HA antibody, subjected to SDS-PAGE and visualized by autoradiography. D1CPY (1.3) and S8CPY(1.5) samples were incubated in the presence or the absence of endoglycosidase H (Endo H) prior to SDS-PAGE. Open and closed circles indicate glycosylated and unglycosylated bands, respectively. aa, amino acids. (B) Quantified data of (A). Translocation efficiency was calculated as ([glycosylated band (%) / total]) for D#CPY (blue) and S#CPY (orange) and plotted.

1. 2. *h* region of signal sequences affects the ER translocation efficiency

The effect of the *h* region on the ER translocation was assessed by means of systematically varied hydrophobicity. With the longest *n* region extension, the hydrophobicity of signal sequences of D and S CPYs was systematically increased (D27CPY(ΔG_{app}) and S30CPY(ΔG_{app}) (Fig. 6A and Table 1). The hydrophobicity of signal sequences was calculated by the ΔG predictor. The apparent free energy for membrane insertion (ΔG_{app}) denotes how favorable or unfavorable a segment is to be inserted into the membrane via the Sec61 translocon.

The ER translocation of CPY *h* variants were assessed in WT transformants as aforedescribed (Fig. 8A). Less hydrophobic (high ΔG_{app}) *h* variants resulted in no glycosylated species whereas restoration of translocation was observed with both D and S CPY hydrophobic *h* variants, whose signal sequences of ΔG_{app} value lower than −1.0 kcal/mol and 0.5 kcal/mol, respectively. A negative correlation is shown between the translocation efficiency and the ΔG_{app} values (Fig. 8B), indicating that the hydrophobic *h* region could restore translocation defects caused by the long *n* region. Glycosylated species were confirmed by tunicamycin treatment, which inhibits *in vivo* N-linked glycosylation.

Upon tunicamycin treatment, smaller extra band was observed for D22CPY(-1.0), S27CPY(0.1) and S29CPY(-0.9), but not for D22CPY(-2.6) despite the efficient ER translocation (Figs. 8A and 9A; it should be noted that the length of *n* region was differently calculated by the ΔG predictor despite the same fusion, resulting in a little shift in the value of the *n* region). CPY is originally a secretory protein that is cleaved upon the ER translocation and all the CPY variants contain the original cleavage site at the C-terminus of the signal sequence. Hence, the low cleavage efficiency of D22CPY(-2.6) is probably

attributable to its highly hydrophobic *h* region. It was previously shown that less hydrophobic signal sequences were efficiently cleaved by signal peptidase whereas hydrophobic ones in the identical protein context were not, resulting in a membrane-anchored protein. By carbonate extraction, D22CPY(-2.6) was confirmed to be membrane-anchored (Fig. 9B). In sum, these results show that hydrophobic signal sequences can direct the ER translocation in spite of a long *n* region and can become membrane-anchored.

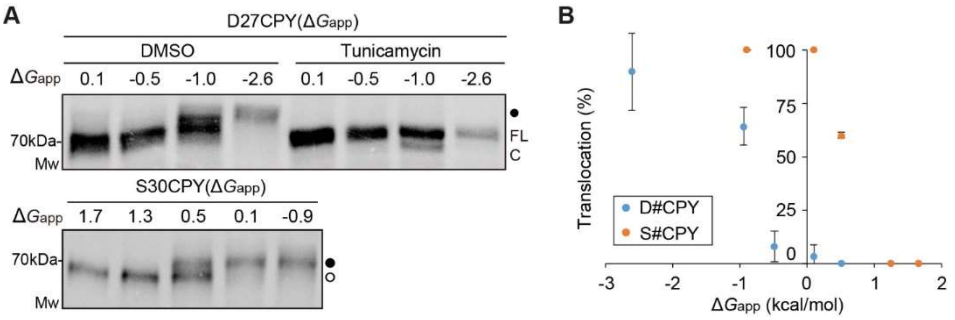

Figure 8. Increased hydrophobicity of *h* region increases the ER translocation efficiency. (A) WT cells expressing D27CPY(ΔG_{app}) of indicated ΔG_{app} values were radiolabeled in the presence or the absence of tunicamycin (*top*). S30CPY (ΔG_{app}) variants expressed in WT were analyzed as in Fig. 8A (*bottom*). FL, full-length form; C, cleaved form; filled circle, glycosylated products. (B) Translocation efficiency for D#CPY and S#CPY was calculated and plotted as in Fig. 8B.

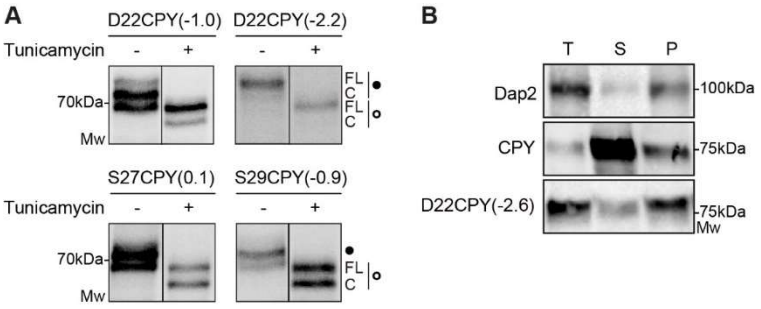

Figure 9. *h* region hydrophobicity determines the cleavage and the localization of proteins. (A) WT cells expressing indicated CPY variants were radiolabeled in the presence or the absence of tunicamycin and analyzed as in Fig. 8A. FL, full-length form; C, cleaved form; a filled circle, glycosylated form; an open circle, unglycosylated form. (B) WT cells expressing indicated proteins were subjected to crude membrane fractionation. Resulting supernatant and membrane fractions were subjected to carbonate extraction, followed by SDS-PAGE and Western blotting.

1. 3. Natural proteins tend to have negative relationship between the length of *n* region and the hydrophobicity of *h* region

For further systematic investigation on the relationship between the *n* region and the *h* region of signal sequences, tested CPY variants showing >15% translocation efficiency at 5 min of *in vivo* radiolabeling were plotted (Fig. 10A). The cutoff value 15% was set to reflect the initial engagement of a nascent chain to the Sec61 translocon. Moderately hydrophobic CPY variants mostly have short *n* region whereas sufficiently hydrophobic variants dispersed over the varied *n* region.

Next, yeast natural proteins were assessed for the *n* region length and the *h* region hydrophobicity. Among proteins predicted to be part of the secretome or SignalP positive (22), selected were 494 proteins containing only one predicted signal sequence within the first 60 residues. Using the ΔG predictor, the *n* region length and the *h* region hydrophobicity were calculated and plotted for each protein. The majority of natural proteins were found under the experimentally measured threshold (Figs. 10A and B, a grey dashed line).

The relationship between the *h* region hydrophobicity and the *n* region length more clearly appears when the natural proteins were sorted in boxes with a 3-residue-window of the *n* region length (Fig. 10C). A negative correlation was observed for the *n* region shorter than 12 residues: as the *n* region lengthens, the ΔG_{app} decreases. When the *n* region exceeds 12 residues, no significant correlation was observed.

The distribution of the *n* region in a given range of the *h* region hydrophobicity was also assessed (Fig. 10D). Proteins with the hydrophobic *h* region (ΔG_{app} value < -0.5 kcal/mol) appear to have a broader range of the *n* region length as experimentally observed. On the other hand, the *n* region is much shorter for the less hydrophobic signal sequences (ΔG_{app}>

0.5 kcal/mol). Collectively, it can be generalized that less hydrophobic signal sequences tend to have a shorter *n* region and that sufficiently hydrophobic signal sequences override a long *n* region.

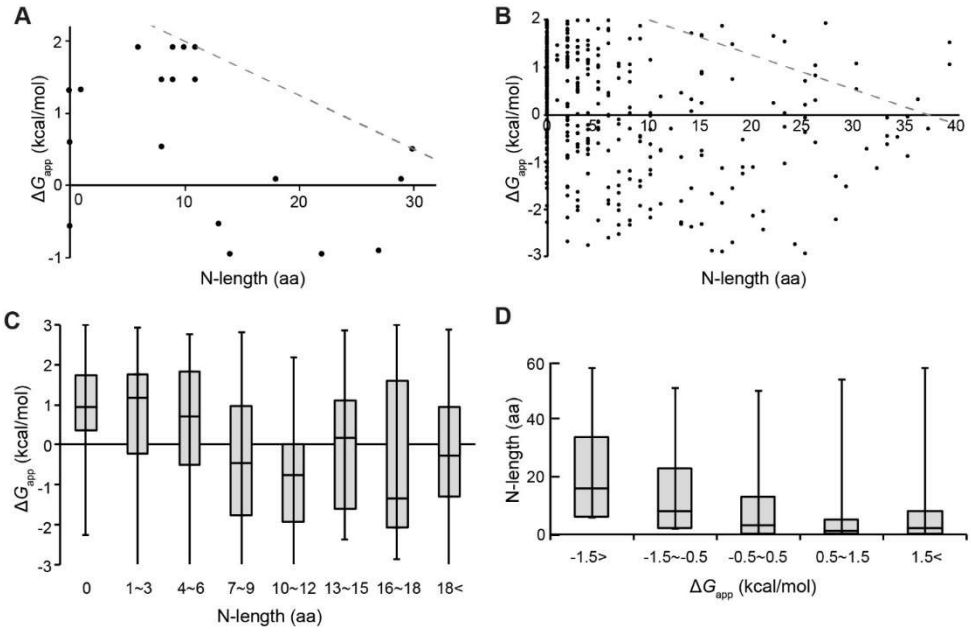

Figure 10. Distribution of *n* region length and *h* region hydrophobicity of signal sequences in yeast proteins. (A) Among profiled CPY-derived variants in Figs. 8 and 9, those with > 15% translocation efficiency at 5 min radiolabeling were plotted: *h* region hydrophobicity (ΔG_{app} value (kcal/mol)) on the y axis and *n* region length (amino acids) on the x axis. (B) Predicted *n* region length and *h* region hydrophobicity of signal sequences of total 494 natural SP and SA proteins in *S. cerevisiae* were plotted as in (A). (C and D) The same set of proteins in (B) are sorted. Boxes and whiskers show 25–75 percentile values with min/max; the median is the central line in each box. (C) Distribution of *h* region hydrophobicity (ΔG_{app} (kcal/mol)) of signal sequences in indicated ranges of *n* region length. (D) Distribution of *n* region length in indicated ranges of *h* region hydrophobicity (ΔG_{app} (kcal/mol)).

1. 4. The required translocon components in accordance with signal sequence characteristics

To elucidate how signal sequences of varied *n* and *h* regions accordingly require which translocon components, translocation efficiency of CPY variants was assessed in yeast strains either defective in or depleted of one of the Sec62/63 complex subunits (*sec62 35DDD*, *sec63 A179T*, *sec71*Δ and *sec72*Δ) or the SRP (*sec65-1*) (Fig. 11).

As Sec62 and Sec63 defective strains, we used *sec62 35DDD* that carries negatively charged Asp substitution in the N-terminus, which impairs electrostatic interaction with Sec63; *sec63 A179T* was prepared by site-directed mutagenesis, where the point mutation in the luminal J domain disrupts interaction with the BiP ATPase Kar2 (Fig. 1C). For nonessential subunits Sec71 and Sec72, single deletion strains were prepared (*sec71*Δ and *sec72*Δ). The well-established *sec65-1* strain was also employed that exhibits a temperature-sensitive defect in SRP function, which impairs co-translocational translocation (Fig. 2).

For simplicity, CPY *n* and *h* variants are indicated according to the length and the hydrophobicity regardless of their origins: N#CPY(ΔG_{app}). First, CPY variants of a short *n* with hydrophobic *h* region, CPY(-0.6) and CPY(-1.9), were expressed in five mutant strains and translocation was compared to that in the WT strain (Fig. 12A). Translocation was defective in *sec71*Δ and *sec65-1* cells whereas other mutant strains did not exhibit any defects.

Secondly tested were N1CPY(1.3) and N6CPY(1.9) variants that contain a short *n* with less hydrophobic h region. While these variants did not require SRP function, they became greatly dependent on all the Sec components (Fig. 11B).

Next, variants containing a signal sequence with a long *n* and hydrophobic *h* region were assessed: N22CPY(-1.0) and N22CPY(-2.6) (Fig. 11C). Both variants showed defective translocation in *sec63 A179T*, *sec71*Δ and *sec65-1* strains where in the two latter strains, translocation was severely compromised, suggesting that such signal sequences greatly depend on Sec71 and SRP function. A relatively less hydrophobic N22CPY(-1.0) also showed considerable translocation defects in *sec62 35DDD* and *sec72*Δ strains whereas a relatively more hydrophobic N22CPY(-2.6) did not, indicating that Sec62 and Sec72 sort the *h* region of low hydrophobicity.

Lastly, variants with a signal sequence of an intermediate *n* and *h* region were analyzed: N18CPY(0.1) and N18CPY(-1.0) (Fig. 11D). While translocation of a relatively more hydrophobic N18CPY(-1.0) was mildly defective in *sec62 35DDD*, *sec63 A179T* and *sec72*Δ strains unlike a relatively less hydrophobic N18CPY(0.1), both variants showed severe translocation defect in *sec71*Δ and *sec65-1* strains. These results suggest that signal sequences of an intermediate *n* and *h* region require interplay of the SRP and the Sec components for proper engagement to the Sec translocon and the initiation process of the ER translocation.

In summary, the hydrophobicity of the *h* region determines the requirement for either Sec62/Sec72 or the SRP in line with previous studies; Sec62/Sec72 handles low hydrophobic signal sequences whereas the SRP handles highly hydrophobic ones. It is of note that a long *n* region increases the SRP dependence compared to a short *n* region (Figs. 11A and C). Surprisingly, signal sequences having an intermediate long *n* and hydrophobic *h* require both Sec62/Sec72 and the SRP, providing evidence that they are not mutually exclusive and some signal sequences use both pathways.

Moreover, translocation of all the *n* and *h* variants was defective in *sec71*Δ cells, particularly those with a long *n* and hydrophobic *h* region. Thus, Sec71 handles a wide range of signal sequences for efficient translocation initiation *in vivo*, especially critical for internal signal sequences.

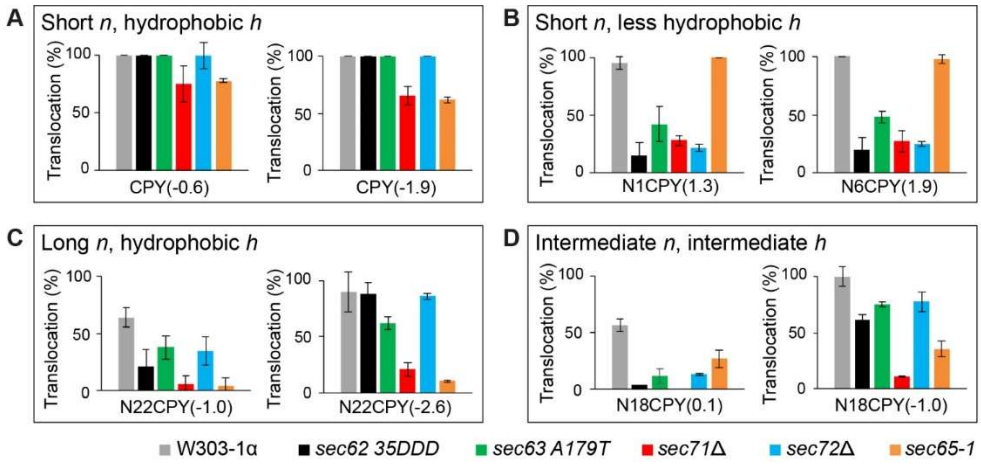

Figure 11. Signal sequences of various characteristics require different combinations of the Sec62/Sec63 components and the SRP for efficient protein translocation. (A-D) Indicated N#CPY(ΔG_{app}) variants (listed in Table 1) in WT and indicated mutant strains were analyzed as in Fig. 8. For *sec65-1* and *sec63 A179T* temperature-sensitive mutant strains, growth/radiolabeling condition was modified as described in materials and methods. The average translocation efficiency ([glycosylated bands × 100 (%) / total]) and standard deviation is indicated (n ≥ 3).

1. 5. The requirement of the translocon components was assessed with natural proteins

Three natural proteins having various *n* and *h* regions, CPY, Dap2 and Spc3 were analyzed for the requirement of the translocon components in their translocation (Fig. 12). All three proteins have the glycosylated C-terminal domain in the ER lumen, which allows measuring translocation efficiency by assessing the glycosylation state.

CPY possess a signal sequence of a short *n* region and less hydrophobic *h* region as the variants in Fig. 12B. Similar to those, CPY showed substantial translocation defects in all the Sec translocon mutant strains except *sec65-1* strain (Fig. 12A). Translocation of Dap2, whose signal sequence has a long *n* and hydrophobic *h* region (the case for those in Fig. 11C), was mildly defective in *sec63 A179T* and *sec71*Δ strains and dramatically compromised in *sec65-1* cells (Fig. 12B). In comparison, Spc3 containing a signal sequence of an intermediate long *n* and *h* region showed defective translocation in all the mutant strains except *sec72*Δ strain (Fig. 12C).

Sec62 and the SRP showed opposite substrate specificity over definite characteristics. CPY, having a short *n* and less hydrophobic *h* region, required functional Sec62 but not the SRP; Dap2, with a long *n* and hydrophobic *h* region, was substantially dependent on the SRP but not Sec62. As intriguing as those shown in Fig. 11D, Spc3 of an intermediate long *n* and *h* region showed dependence on both Sec62 and the SRP, showing the interplay of Sec62 and the SRP for the initiation process of the ER translocation.

Additionally, all three proteins showed translocation defect in *sec63 A179T* and *sec71*Δ strains, the degree of which was greater in the latter, consistent with the observations with the engineered CPY variants. Sec71 is therefore crucial in handling of signal sequences having various characteristics.

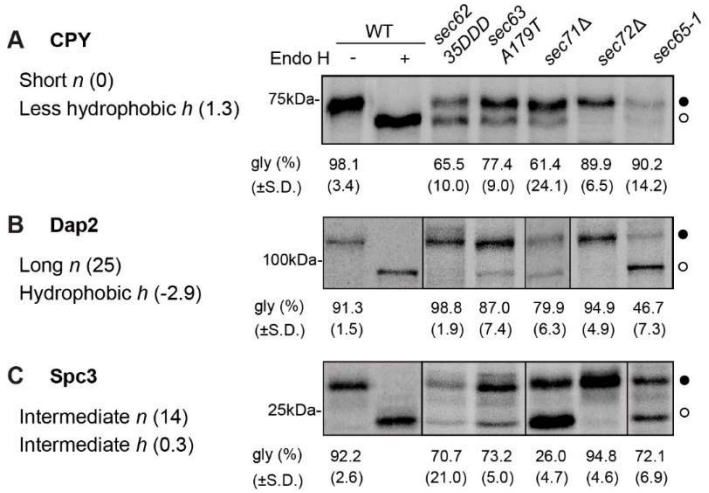

Figure 12. Yeast natural proteins of varying signal sequence requires different combination of the Sec62/Sec63 complex and the SRP. (A-C) CPY (A), Dap2 (B) and Spc3 (C) in WT or mutant strains were analyzed for translocation as in Fig. 8. Those expressed in WT cells were treated with Endo H prior to SDS-PAGE. The average translocation efficiency ([glycosylated bands × 100 (%) / total]) and standard deviation are indicated below (n = 3). Open and closed circles indicate glycosylated and unglycosylated bands, respectively.

2. Defining roles of the signal peptidase complex subunits in substrate processing

2. 1. The absence of Spc1 or Spc2 affects the signal peptidase-mediated processing of model membrane proteins

Spc1 and Spc2 are subunits of the signal peptidase complex and constitute the most of cytoplasmic part of the complex, respectively interacting with other subunits Spc3 and Sec11 through TM segments (Fig. 5). To explore possible functions of Spc1 and Spc2 in protein translocation and the signal peptidase-mediated processing, genetic deletion of each gene was carried out and a series of model membrane protein LepH2 was employed to monitor both membrane insertion and the signal peptidase-mediated processing *in vivo* (Fig. 13).

The model proteins Lep-H2 were based on leader peptidase from *E. coli* (59). Efficiently targeted and inserted into the ER membrane in N_{out}-C_{in} orientation by TM1, LepH2 harbors potential TM2 composed of Leu and Ala residues of different combinations, 19 residues in total (Fig. 13A). Systematically varied number of Leu residues proportionally changes the hydrophobicity of TM2 and the probability for its membrane insertion, which affects topogenesis of LepH2. Topogenesis can be monitored as LepH2 generates multiple isoform species due to two post-translational modifications (Fig. 13B). One is N-linked glycosylation that occurs in the ER lumen and increases protein mass by 2 kDa. LepH2 harbors one glycosylation sequon in the N-terminus of TM1 and the other in the C-terminus of TM2. Therefore, unglycosylated forms refer to untranslocated species remaining in the cytoplasm, singly glycosylated forms refer to species whose TM1 but TM2 is inserted and doubly glycosylated forms refer to species whose TM2 is inserted into the ER membrane.

The other post-translational modification is the signal peptidase-mediated cleavage. An earlier study showed that TM2 contains putative cleavage site presumably due to repeated Leu and Ala that meets cleavage site criterion (the -3, -1 rule). The signal peptidase complex can cleave LepH2, generating smaller species. Hence, LepH2 results in multiple isoform species that are easily separated by SDS-PAGE and reflect each topogenic state including cleavage by the signal peptidase complex. Thereby LepH2 serves as a sensitive and easy molecular indicator for membrane insertion and signal peptidase-mediated cleavage *in vivo*.

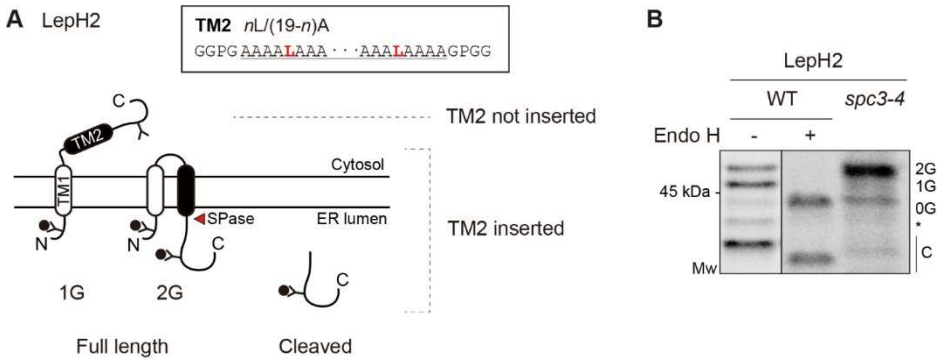

Figure 13. A series of *E. coli* leader peptidase-derived LepH2 model proteins. (A) A schematic of LepH2. TM2 segment of varying hydrophobicity is colored black. TM2 is a 19 residue-long segment of differed ratio between Leu and Ala residues (box). N-linked glycosylation sites and moieties are respectively indicated as Y and filled circles. A red arrowhead points to cleavage by the signal peptidase complex. SPase, the signal peptidase complex. (B) A representative blot of LepH2. LepH2 was expressed in WT and *spc3-4* strains, radiolabeled and subjected to IP, SDS-PAGE and autoradiography. Endo H digestion was carried out prior to SDS-PAGE. 2G, doubly glycosylated form; 1G, singly glycosylated form; 0G, unglycosylated form; *, a fraction cleaved by unknown protease in the cytoplasmic side; C, cleaved form by the signal peptidase complex.

A series of LepH2 includes four variants according to the number of Leu residues and the N-terminal flanking residues of TM2 (Fig. 14A). The more Leu residues are present, the more hydrophobic TM2 becomes, leading to better membrane insertion. Meanwhile, LepH2(K1L) carries positively charged N-terminal flanking residues, which enhances membrane insertion of TM2. Yeast transformants expressing LepH2 were metabolically radiolabeled with [^{35}S]Met for 10 min to monitor early translocation and processing state (Fig. 14B).

By comparing the ratio of TM2 inserted forms (2G and cleaved forms), the membrane insertion efficiency was measured and compared in WT, *spc1*Δ or *spc2*Δ strain (Fig. 14C). The measured membrane efficiency of LepH2 variants did not differ among strains, indicating that Spc1 and Spc2 are not involved in the membrane insertion process of LepH2. It is of note that hydrophobic LepH2(3L) and LepH2(5L) show efficient membrane insertion whereas less hydrophobic LepH2(2L) did not. Lys residues of LepH2(K1L) overcame low hydrophobicity of TM2, leading to enhanced membrane insertion.

Although membrane insertion did not differ among strains, it was prominent that the cleaved form significantly increased in the absence of Spc1 and Spc2 (Fig. 14D). Except LepH2(K1L) of no difference, three other LepH2 variants showed increased level of the cleaved form in *spc1*Δ and *spc2*Δ strains compared to that in WT cells, albeit to a different extent. The relative degree of the cleavage gradually increased proportional to the hydrophobicity of TM2 in the *spc1*Δ strain whereas decreased in the *spc2*Δ cells (Fig. 14E). These data suggest that Spc1 and Spc2 have distinct substrate coverage, perhaps depending on the hydrophobicity of TM segments.

The enhanced cleavage of LepH2 variants were shown not only in the W303-1α background but also in the BY4741-derived strains, implying that the possible regulatory roles of Spc1 and Spc2 subunits are not circumscribed in a certain yeast strain (Fig. 15).

Collectively, the loss of Spc1 and Spc2 induces the signal peptidase-mediated processing of model membrane proteins while membrane insertion was unaffected. These data suggest the regulatory roles of Spc1 and Spc2 on substrates or functions of the signal peptidase complex. In more detail, range of affected substrates differs between *spc1*Δ and *spc2*Δ, suggesting these two subunit may serve distinct roles.

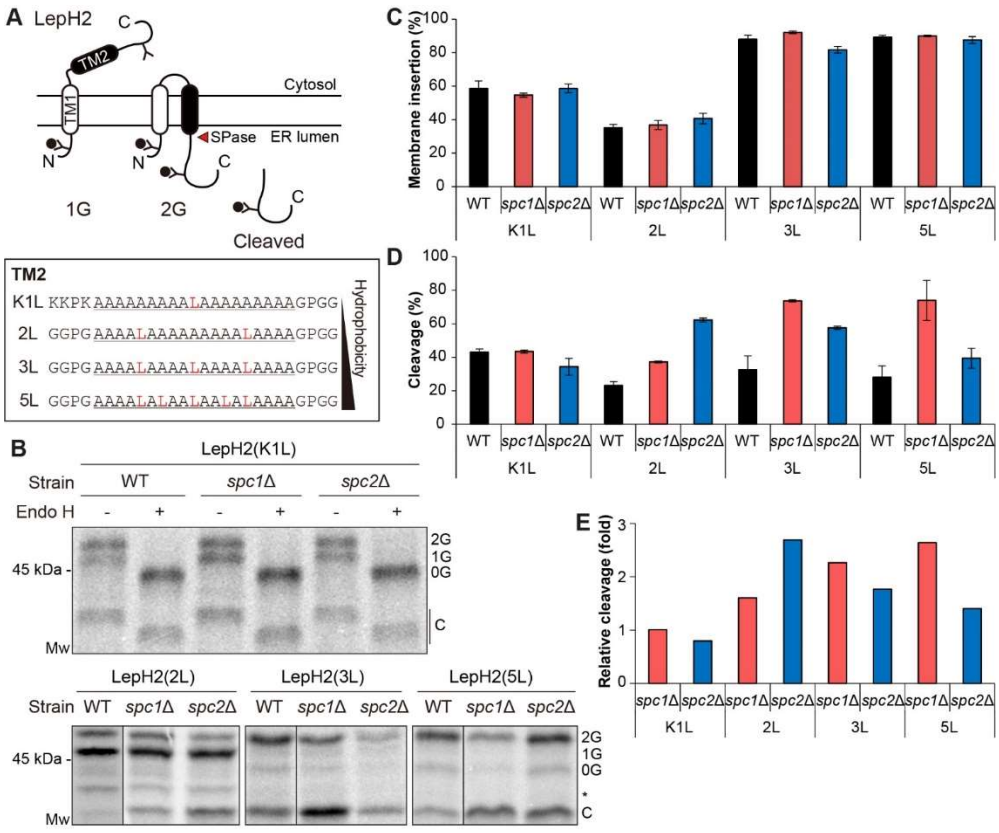

Figure 14. Spc1 and Spc2 modulate the signal peptidase-mediated processing of double-spanning membrane proteins LepH2. (A) A schematic of LepH2. TM2 segment of varying hydrophobicity is colored black, whose amino acid sequences are shown with N- and C-terminal flanking residues (box). N-linked glycosylation sites and moieties are indicated as Y and filled circles. A red arrowhead points to cleavage by the signal peptidase. (B) Indicated LepH2 variants were radiolabeled in WT or deletion strains for 10 min and subjected to IP, SDS-PAGE and autoradiography. LepH2(K1L) variants were treated with Endo H prior to SDS-PAGE (*top*). 2G, doubly glycosylated form; 1G, singly glycosylated form; 0G, unglycosylated form; *, a fraction cleaved by unknown protease in the cytoplasmic side; C, cleaved form by the signal peptidase complex. (C) Membrane insertion of TM2 in LepH2 variants was measured as doubly glycosylated products/(total-0G)*100 (%). (D) Cleavage (%) of TM2 of LepH2 was calculated as cleaved/(2G+cleaved)*100 (%). (E) Relative cleavage (fold) of LepH2 in *spc1*Δ and *spc2*Δ cells compared to that in WT cells was calculated as [cleavage (%) in mutant cells/cleavage (%) in WT cells] and plotted for each LepH2 variant.

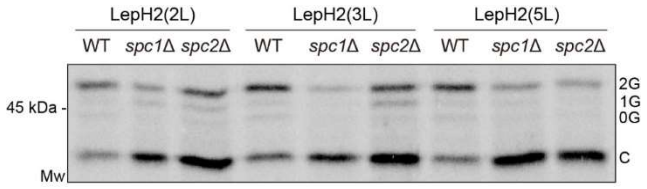

Figure 15. Spc1 and Spc2 modulate the signal peptidase-mediated processing of double-spanning membrane proteins LepH2 in BY4741 cells. LepH2 variants were expressed in BY4741-derived cells, radiolabeled for 10 min and subjected to IP, SDS-PAGE and autoradiography as in Fig. 14B. 2G, doubly glycosylated form; 1G, singly glycosylated form; 0G, unglycosylated form; C, cleaved form by the signal peptidase complex.

2. 2. Hydrophobicity of a TM segment determines the signal peptidase-mediated processing of model membrane proteins

To define decisive factors of the signal peptidase-mediated processing, another series of model proteins were assessed. Derived from *E. coli* leader peptidase as well as LepH2, LepCC variants contain the engineered TM2 composed of Leu residues only, followed by a canonical cleavage cassette (VPSAQA↓A, ↓ indicates the cleavage site, Figure 16A) (40). An earlier study showed that the signal peptidase-mediated cleavage after its TM2 was dependent on the number of Leu residues as determined *in vitro* with dog pancreatic microsomes; TM2 variants with a shorter stretch of Leu residues were cleaved by the signal peptidase complex whereas ones with a longer stretch of Leu residues were not (40). Signal-anchored versions of LepCC were also prepared by deletion of the N-terminus including the first TM of LepCC (LepCCt; Fig. 16B).

To assure whether these variants undergo the signal peptidase-mediated processing in yeast, two strategies were employed: firstly, the cleavage site was destroyed by introducing Pro at the +1 position in 14L variants (61-63); secondly, 14L variants were expressed in the *spc3-4* strain at the non-permissive temperature, which compromises the catalytic activity of the signal peptidase complex (Figs. 16C and D). A slowly migrated full-length band was detected for P+1 variants in WT cells and 14L variants expressed in *spc3-4* strain at the non-permissive temperature whereas a smaller species was predominant for 14L expressed in the WT strain, confirming that LepCC variants are processed by the yeast signal peptidase complex.

Next, three LepCCt variants (14L, 17L and 20L) were metabolically radiolabeled for 5 min in WT cells and their ER translocation and processing states were monitored (Fig. 16E). LepCCt(14L) was efficiently processed by the signal peptidase complex and mostly

generated cleaved form. Albeit less prominent, LepCCt(17L) also generated a cleaved form, indicating that it is a substrate of the signal peptidase complex as well. However, the most hydrophobic 20L variants were not processed, showing that hydrophobic segments escape the signal peptidase-mediated processing as previously observed (40). Endo H treatment confirmed efficient translocation and membrane insertion of LepCC variants in the yeast ER.

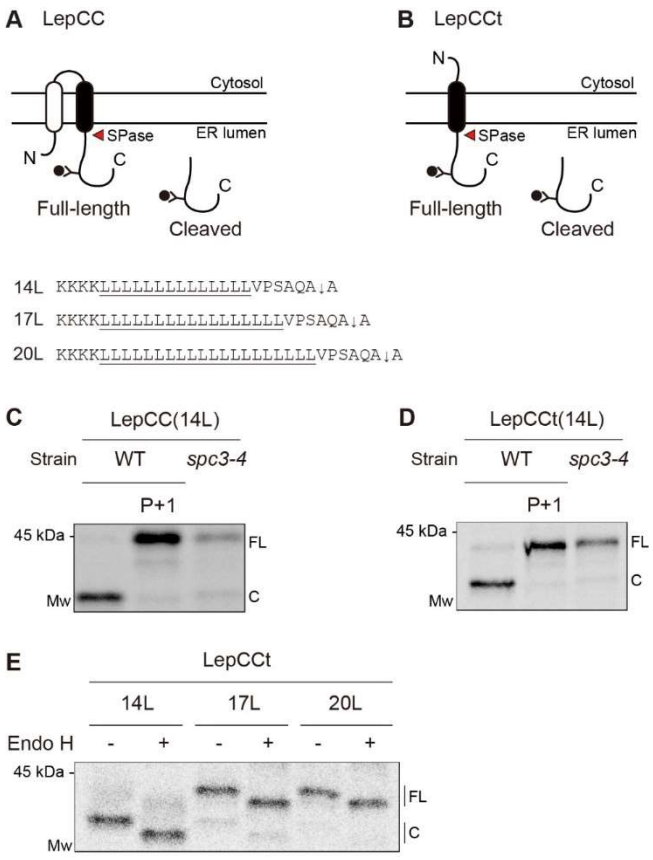

Figure 16. A series of *E. coli* leader peptidase-derived LepCC model proteins are recognized and cleaved by the signal peptidase complex. (A and B) A schematic of LepCC (A) and LepCCt (B). The TM domain is colored black and an N-linked glycosylation site and a moiety are indicated as Y and a filled circle. Flanking and TM sequences including the cleavage site (↓) are shown for three variants. A red arrowhead points to cleavage by the signal peptidase complex. (C and D) 14L variants of LepCC (C) and LepCCt (D) were expressed in WT and *spc3-4* strains, radiolabeled for 5 min and subjected to IP, SDS-PAGE and autoradiography. P+1 indicates Pro substitution at +1 position relative to the cleavage site. (E) Indicated LepCCt variants were analyzed as in

(C). Endo H digestion was carried out prior to SDS-PAGE. FL, full-length form; C, cleaved form.

2. 3. The absence of Spc1 or Spc2 increases the signal peptidase-mediated processing of single- and double-spanning model membrane proteins

In WT cells, the cleavage of LepCC variants depended on hydrophobicity of TM segments and I asked whether such profiles change without Spc1 or Spc2. For this assessment, their processing was monitored in cells depleted of Spc1 or Spc2 by *in vivo* pulse labeling (Figs. 17A and B). Both sets of LepCC and LepCCt showed consistent results: in *spc1*Δ and *spc2*Δ cells, 14L variants were completely cleaved and even 17L variants resulted in substantial amounts of the cleaved form; the most hydrophobic 20L variants were minimally cleaved.

For better kinetic understanding, the processing of LepCCt(17L) variants was traced in WT and the deletion strains by pulse-chase experiments (Fig. 17C). A cleaved product of LepCCt(17L) in the *spc2*Δ strain significantly increased compared to that expressed in the WT strain already at 0 min and even further increased in the following chase time, leaving a very faint full-length form at 15 min time point. This cleavage in *spc1*Δ strain was also gradually enhanced, indicating that the signal peptidase-mediated cleavage continued post-translationally. These data suggest that longer and more hydrophobic TM segments normally evade the signal peptidase-mediated processing to which they are subjected when Spc1 or Spc2 is absent.

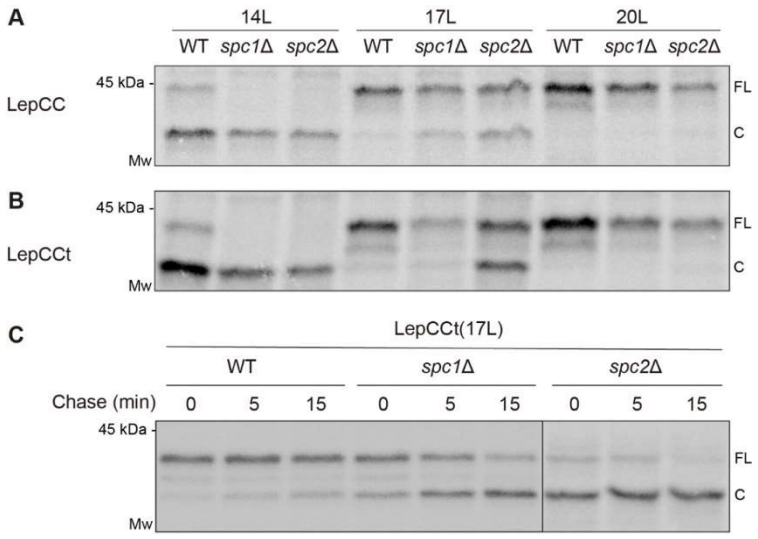

Figure 17. The signal peptidase-mediated processing of LepCC variants is enhanced in the *spc1*Δ and *spc2*Δ strains. (A and B) The indicated LepCC (A) or LepCCt (B) variants in WT, *spc1*Δ or *spc2*Δ cells were radiolabeled for 5 min at 30°C and analyzed as in Fig. 16C. (C) LepCCt(17L) in the WT, *spc1*Δ or *spc2*Δ strain was radiolabeled for 5 min, chased for indicated time points at 30°C and analyzed as in Fig. 16C. FL, full-length form; C, cleaved form.

2. 4. The signal peptidase complex lacking Spc1 and Spc2 recognizes the canonical cleavage site

Observing that the signal peptidase-mediated processing increases in cells lacking Spc1 or Spc2, I asked whether the complex lacking a subunit cleaves at non-canonical site, thereby increasing cleaved forms. To test this possibility, LepCC variants containing Pro at +1 position was assessed in WT and cells lacking both Spc1 and Spc2 (Fig. 18). LepCC Pro+1 variants were not processed irrespective of the presence of Spc1 and Spc2, showing that the signal peptidase complex lacking Spc1 and/or Spc2 still recognizes the canonical cleavage site only.

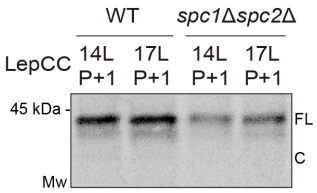

Figure 18. The signal peptidase complex lacking its subunits still processes the canonical cleavage site. Indicated LepCC variants were expressed and radiolabeled for 5 min in WT or *spc1*Δ*spc2*Δ cells and subjected to IP, SDS-PAGE and autoradiography. FL, full length form; C, cleaved form; P+1, Pro at +1 position.

2. 5. Overexpressed Spc1 reduces the signal peptidase-mediated processing of model membrane proteins

Since processing of Lep-derived model membrane proteins by the signal peptidase complex increased in the absence of Spc1 or Spc2, I asked whether overexpression of Spc1 or Spc2 also affects processing, possibly in an opposite manner. A series of LepH2 variants were radiolabeled and chased over 15 min in WT strain containing an empty vector, Spc1 or Spc2 overexpression (OE) vector (Fig. 19A). Cleavage of LepH2 variants in all the strains increased during the chase time, once more indicating that processing continued post-translationally. Among all, clear reduction in processing was observed only for LepH2(5L) in Spc1 OE cells; slight reduction of cleavage at 0 min point markedly increased at 15 min point compared to that seen in WT cells (Figs. 19A and B). Overexpressed Spc1 therefore exerted somewhat dominant role in spite of originally being one subunit of the four subunit-complex. Unlike Spc1, overexpressed Spc2 did not lead to such phenomenon although the loss of Spc2 caused significant change in cleavage profiles of LepH2 (Fig. 14). This observation thus led to an assumption that Spc1 and Spc2 have distinct functions.

To confirm overexpressed Spc1 reduces the signal peptidase-mediated processing, LepCC(17L) and LepCCt(17L) were traced under Spc1 OE conditions (Fig. 20). The cleaved form in Spc1 OE cells decreased for both LepCC (Fig. 20A) and LepCCt (Fig. 20B) 17L variants compared to those expressed in WT cells over 15 min time point. These data suggest that additional Spc1 can protect TM segments from the signal peptidase-mediated cleavage co- and post-translationally.

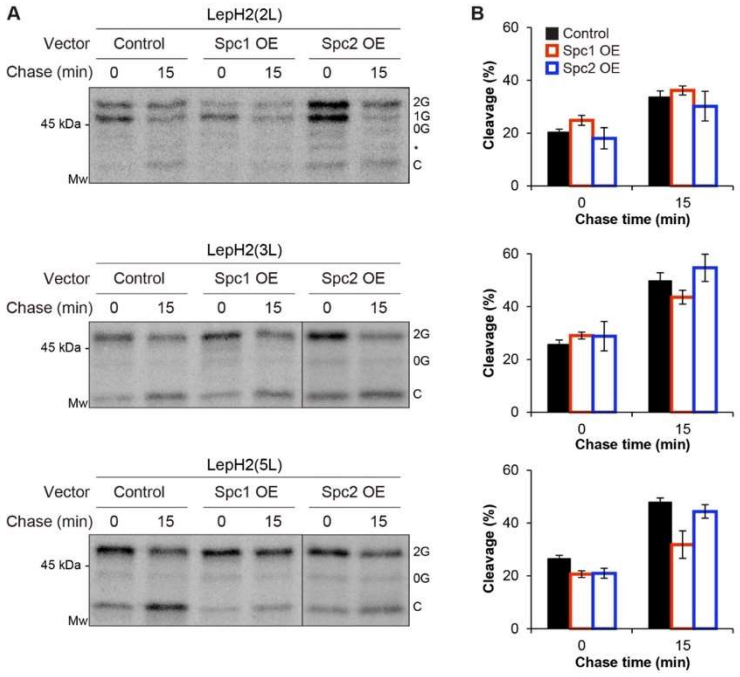

Figure 19. Overexpressed Spc1 but not Spc2 modulates the signal peptidase-mediated processing of hydrophobic membrane proteins. (A) The indicated LepH2 variants in WT cells harboring control, Spc1 OE or Spc2 OE vector were radiolabeled for 10 min chased for indicated time points at 30°C and analyzed as in Fig. 16C. 2G, doubly glycosylated form; 1G, singly glycosylated form; 0G, unglycosylated form; *, a fraction cleaved by unknown protease in the cytoplasmic side; C, cleaved form by the signal peptidase complex. (B) Cleavage (%) in (A) was quantified and plotted. OE, overexpression.

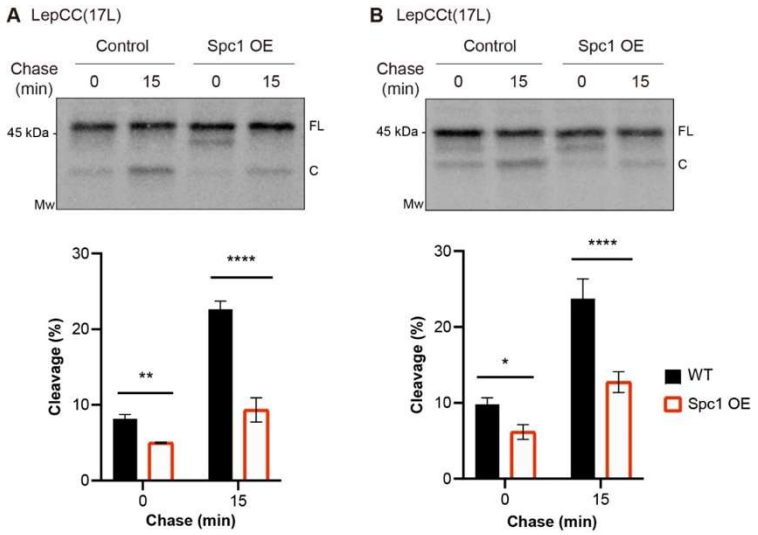

Figure 20. Overexpressed Spc1 reduces the signal peptidase-mediated processing of LepCC variants. (A and B) *Top*, 17L variants of LepCC (A) and LepCCt (B) in WT cells harboring control or Spc1 OE vector were radiolabeled for 5 min chased for indicated time points at 30°C and analyzed as in Fig. 16C. *Bottom*, Cleavage (%) in (A) was quantified and plotted. OE, overexpression.

2. 6. Spc1 and Spc2 have overlapping but distinct functions in the signal peptidase-mediated processing

Although it was implicated above that Spc1 and Spc2 are likely to have distinct functions (Figs. 14 and 19), restoration of one subunit in cells lacking the other was carried out to determine whether they have functional redundancy. For this, LepH2 variants were expressed either with overexpressed Spc1 in the *spc2*Δ strain or overexpressed Spc2 in the *spc1*Δ strain (Fig. 21). The rationale behind this experimental scheme was that if Spc1 and Spc2 function similarly, the processing of the model proteins would be restored by expression of one subunit in the deletion strain of the other subunit. However, rescued processing of model proteins was not observed after 5 min radiolabeling, indicating that Spc1 and Spc2 indeed have overlapping but distinct functions in the signal peptidase-mediated processing.

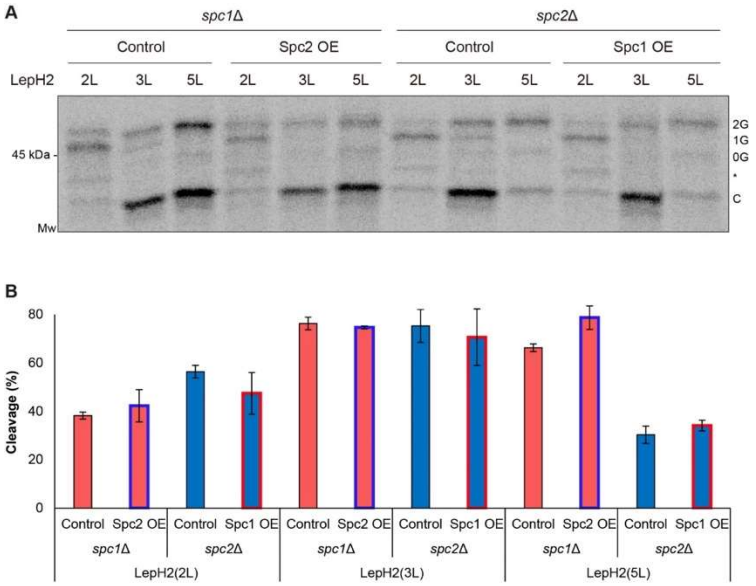

Figure 21. Overexpression of Spc1 or Spc2 does not complement the changed processing of LepH2 variants in deletion strains reciprocally. (A) The indicated LepH2 variants in *spc1*Δ or *spc2*Δ cells harboring control or the counterpart subunit-expressing vector were radiolabeled for 10 min and analyzed as in Fig. 16C. 2G, doubly glycosylated form; 1G, singly glycosylated form; 0G, unglycosylated form; *, a fraction cleaved by unknown protease in the cytoplasmic side; C, cleaved form by the signal peptidase complex. (B) Cleavage (%) in (A) was quantified and plotted. OE, overexpression.

2. 7. Spc1 interacts with a TM segment of membrane proteins

Observations so far demonstrate that Spc1 affects the signal peptidase-mediated processing of model membrane proteins in a consistent manner; the loss of Spc1 caused increased processing, which was repressed by overexpressed Spc1. This led to an assumption that Spc1 might recognize and protect TM segments from the signal peptidase-mediated processing; therefore, co-immunoprecipitation was carried out to detect physical association between Spc1 and membrane proteins (Fig. 22).

First, two LepCC variants were tested; *spc1*Δ cells were transformed with plasmids of the C-terminally FLAG-tagged Spc1 and the HA-tagged LepCC(14L) or LepCC(17L) (Fig. 22A). While a cleaved form of LepCC(14L) was not co-immunoprecipitated with Spc1, a full-length form was, as well as a full-length LepCC(17L). It should be noted that Spc1 was overexpressed under strong GPD promoter, which consequently led to reduction in the overall processing of the model proteins.

An additional series of CPY-derived model proteins were also examined (Fig. 22B). Three N9CPYt variants were modified on D9CPY variant in Fig. 6, whose C-terminus is truncated to ease the separation between the full-length and the cleaved forms. Two *h* region variants were selected: N9CPYt(-1.0) and N9CPYt(-2.6). The more hydrophobic N9CPYt(-2.6) was additionally mutated to destroy its cleavage site, which makes it a membrane-anchored protein (N9CPYt(-2.6) CS_{mut}).

These CPY variants were assessed for Spc1 interaction as above (Fig. 22C). The less hydrophobic N9CPYt(-1.0) was fully cleaved and not recovered in the immunoprecipitated lane, which resembles the case of LepCC(14L) cleaved form. Intriguingly, full-length N9CPYt(-2.6) was selectively pulled down and recovered in the immunoprecipitated lane. Since uncleaved N9CPYt(-2.6) was not detected in the input sample, the full-length form

must be a minor fraction protected from the signal peptidase-mediated processing by association with overexpressed Spc1.

These data show that Spc1 reduces cleavage of membrane proteins and physically associates with TM segments. Indeed, membrane-anchored N9CPYt(-2.6) CS_{mut} was considerably recovered in the immunoprecipitated lane. This result also demonstrates that Spc1 recognizes a hydrophobic segment rather than the cleavage site, since both N9CPYt(-2.6) variants with and without cleavage site were co-immunoprecipitated with overexpressed Spc1 but the less hydrophobic N9CPYt(-1.0) with cleavage site was not.

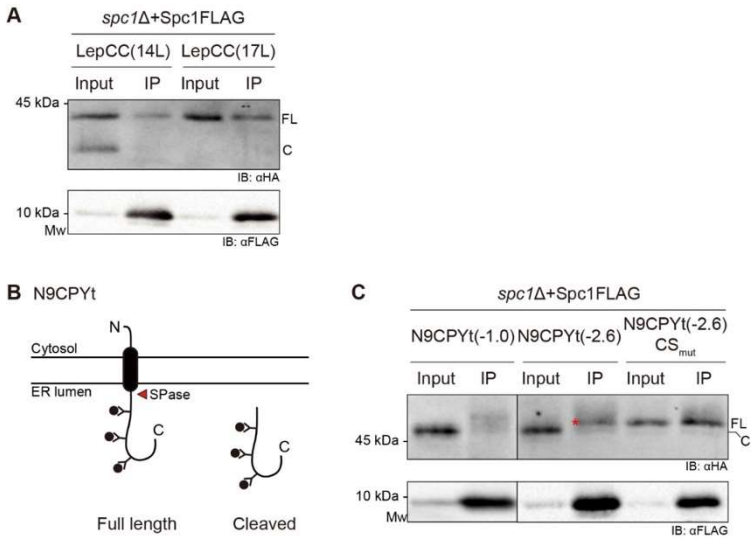

Figure 22. Overexpressed Spc1 interacts with model membrane proteins. (A and C) Co-immunoprecipitation of overexpressed Spc1 and LepCC (A) and N9CPYt (C) variants. *spc1*Δ cells co-expressing Spc1FLAG and the indicated HA-tagged substrates were subjected to crude membrane fractionation. Isolated membranes were solubilized with 1% Triton X-100 lysis buffer, followed by co-immunoprecipitation with anti-FLAG antibodies and visualized by SDS-PAGE and immunoblotting with indicated antibodies. IP; immunoprecipitants; IB, immunoblotting; FL, full-length; C, cleaved. A red asterisk

indicates full-length N9CPYt(-2.6). (B) A schematic of N9CPYt variants. Glycosylation sites and moieties are indicated as Y and filled circles. A red arrowhead points to cleavage by the signal peptidase complex.

2. 8. Abundance of some natural membrane proteins was reduced in the *spc1*Δ strain

Next asked was whether natural proteins are subjected to Spc1-regulated processing *in vivo*. The rationale was that if such membrane proteins contain potential cleavage sites, the loss of Spc1 would fail to protect them from the signal peptidase-mediated processing, leading to cleavage and the cleaved products would be susceptible to degradation. Consequently, amounts of those Spc1-regulated proteins in the *spc1*Δ cell would be reduced compared to those in the WT cell. To this end, a number of proteins were selected 1) whose TM segment(s) possess potential cleavage sites or 2) whose abundance was shown to be reduced in the *spc1*Δ proteome compared to that in the WT proteome by quantitative mass spectrometry analysis. The selected candidates were either subcloned into a vector with HA tag or chromosomally tagged with HA and subjected to Western blotting to assess their processing and steady-state levels in WT and *spc1*Δ strains (Figs. 23 and 24).

Signal-anchored proteins

First, selected signal-anchored proteins were assessed (Fig. 23). Ktr3 and Gda1 are localized in the Golgi membrane, involved in the O-linked glycosylation and the transport of GDP-mannose into the Golgi lumen, respectively (Figs. 23A and B) (64-66). Pho8 is an alkaline phosphatase, localized to the vacuole membrane and involved in protein dephosphorylation and nicotinamide metabolism (Fig. 23C) (67). Pho8t is a C-terminally truncated version of Pho8 for better resolution between full-length and cleaved products on an SDS-gel. Erv46, a component of COPII vesicles involved in transport between the ER and Golgi was C-terminally truncated to remove the C-terminal TM segment to make it a signal-anchored version (Fig. 23D) (68). Although TM segments of these proteins contain potential cleavage sites, the signal peptidase-mediated cleaved fragments were not

detected. However, the amounts of these proteins in the *spc1*Δ strain decreased by 40-60% compared to those in the WT strain (Fig. 23).

Multi-pass membrane proteins

Processing and abundance of a few multi-spanning membrane proteins in *spc1*Δ and WT strains were also compared. Sna3 is in vacuolar intraluminal vesicles and involved in sorting of proteins into multivesicular bodies (Fig. 24A) (69). Pmp3 is involved in cation transport and regulation of membrane potential in the plasma membrane (Fig. 23B) (70). The first TM segments of Sna3 and Pmp3 are in N_{in}-C_{out} orientation and contain potential cleavage sites. Although cleaved fragments were not observed for Sna3 and Pmp3 when assessed by pulse-labeling, steady state levels of these proteins in *spc1*Δ cells were notably decreased compared to those in WT cells (Figs. 24A and B).

Mep2 is a family of ammonium transporter localized to the plasma membrane in yeast (Fig. 24C) (71). It consists of eleven TM segments, two of which in N_{in}-C_{out} orientation contain potential cleavage sites at their C-terminal ends. The amount of full-length Mep2 (~60 kDa) in *spc1*Δ cells also decreased compared to that in WT cells. Notably, potential cleaved forms (two bands around ~30 kDa) were detected whose ratio to the full-length form is much higher in *spc1*Δ cells.

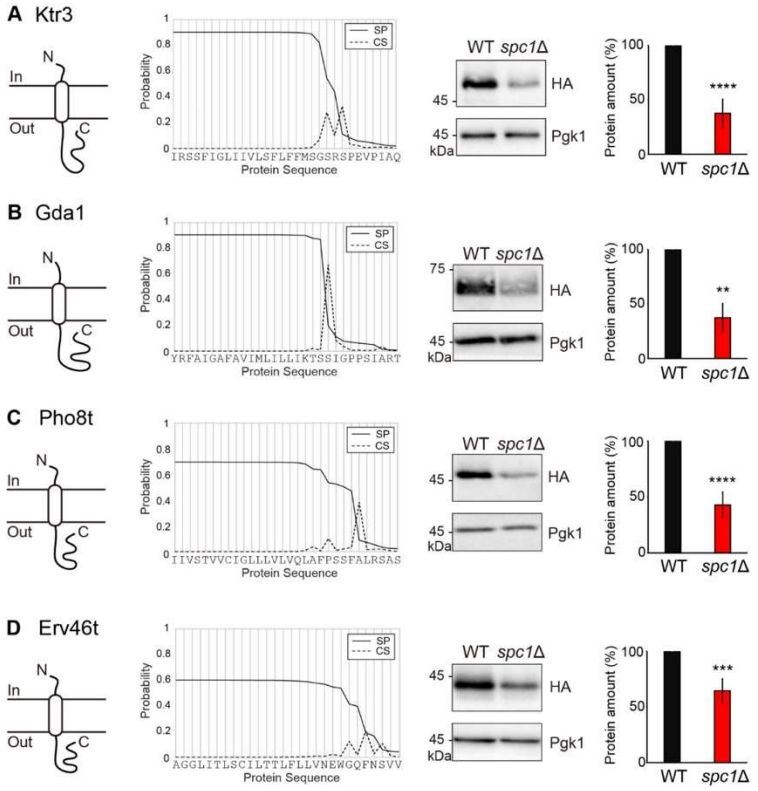

Figure 23. Steady-state levels of natural single-spanning membrane proteins are reduced in the *spc1*Δ strain. (A-D) *Left*, membrane topology and predicted cleavage sites of each protein of interest. *Right*, protein abundance in WT and *spc1*Δ strains was compared by Western blotting. The relative abundance was calculated by comparing the band intensity of the test protein in *spc1*Δ strain to that in WT strain. The ratio of the band intensities of the loading control (Pgk1 or GAPDH) were normalizer for the bands of HA-tagged proteins. p-values were calculated by multiple t-tests; *, p≤0.05; **, p≤0.01; ***, p≤0.01; ****, p≤0.0001. (A) Pho8t, a truncated version of 8 Pho8(YDR481C). (B) Ktr3(YBR205W). (C) Gda1(YEL042W). (D) Erv46t, a truncated version of Erv46(YAL042W).

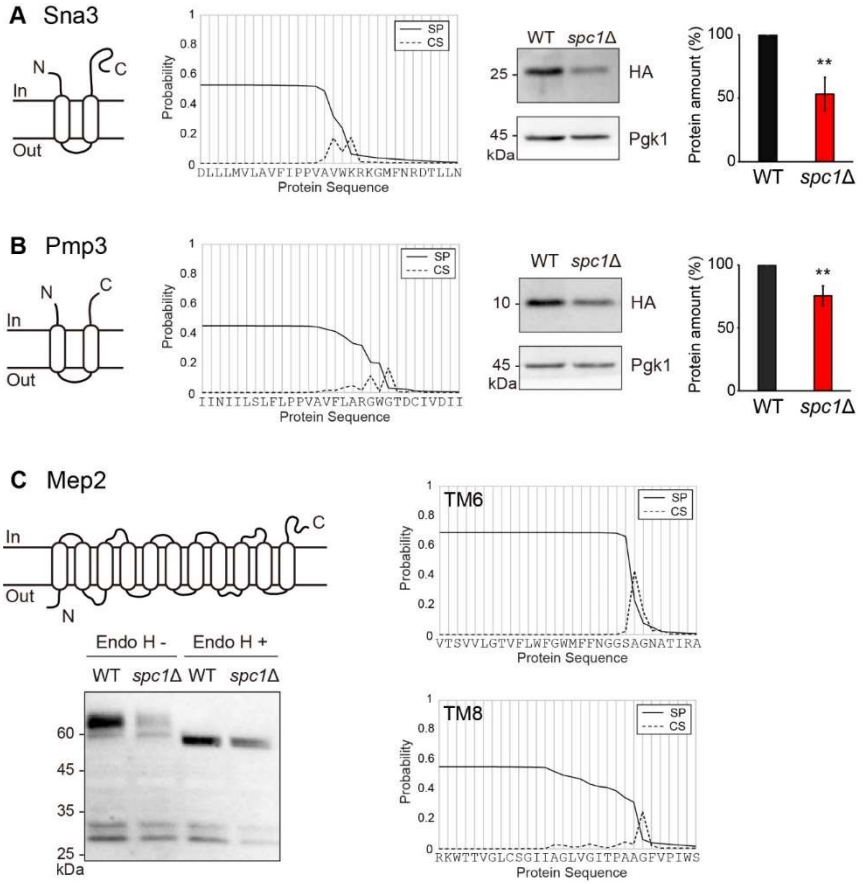

Figure 24. Steady-state levels of natural multi-spanning membrane proteins are reduced in the *spc1*Δ strain. (A and B) *Left*, membrane topology and predicted cleavage sites of each protein of interest. *Right*, protein abundance in WT and *spc1*Δ strains was compared by Western blotting. The relative abundance was calculated by comparing the band intensity of the test protein in *spc1*Δ strain to that in WT strain. The ratio of the band intensities of the loading control (Pgk1 or GAPDH) were normalizer for the bands of HA-tagged proteins. p-values were calculated by multiple t-tests; *, p≤0.05; **, p≤0.01; ***, p≤0.01; ****, p≤0.0001. (A) Sna3(YJL151C). (B) Pmp3(YDR276C). (C) *Top left*, membrane topology of Mep2(YNL142W). *Top bottom*, Mep2 was expressed from WT and

*spc1*Δ cells, subjected to SDS-PAGE and Western blotting. Endo H digestion was carried out prior to SDS-PAGE. *Right*, Predicted cleavage sites in TM segments of Mep2.

2. 9. Overexpressed Sec11 compromises cell viability

Observations so far suggest that Spc1 and Spc2 affect the signal peptidase-mediated processing of membrane proteins and that especially Spc1 has an important role in substrate recognition for the signal peptidase complex. Thus, Spc1 and Spc2 may be regulatory subunits exerting on Sec11, an evolutionarily conserved catalytic subunit of the signal peptidase complex.

It was surmised that if Spc1 and Spc2 are regulatory subunits of the signal peptidase complex, unregulated catalytic subunit Sec11 by the loss of Spc1 and Spc2 might exhibit impaired phenotype. To test this idea, WT cells were transformed to exogenously express Sec11 to different extent to mimic environment where Sec11 loses possible regulation by Spc1 and Spc2 (Fig. 25). Surprisingly, overexpressed Sec11 under strong GPD promoter compromised cell viability under all tested temperature; exogenously expressed Sec11 under its own promoter did not interrupt cell growth (Fig. 25A).

It was of interest to obtain further insight on the harmful effect of overexpressed Sec11. To that end, the other essential subunit Spc3 was co-overexpressed to test the possibility that Spc3 may stabilize Sec11 and restore cell growth (Fig. 25B). WT cells were transformed to express either Spc3 alone or Spc3 and Sec11 together under Gal-inducible promoter. Although Spc3 alone did not lead to defective growth under all tested temperature, co-overexpression of Spc3 and Sec11 caused severe growth defect, showing that Spc3 was unable to rescue inviability owing to overexpressed Sec11.

Lastly, co-overexpression of Spc1 and Spc2 was conducted in the background of overexpressed Sec11 to test the possibility that rescued growth may happen upon restoration of all four subunits in a stoichiometric manner (Fig. 25C). WT cells were transformed to express either Spc1 and Spc2, Spc3 and Sec11 or all four subunits under

Gal-inducible promoter. Induced Spc1 and Spc2 did not result in growth defect themselves but hardly rescued overexpressed Sec11-caused inviability. Cells expressing all four subunits showed very slight growth compared to cells expressing Spc3 and Sec11 at 24˚C and 30˚C. Although Spc1 and Spc2 are anticipated to be regulatory subunits, these data suggest regulated expression of Sec11 is critical for yeast growth, which cannot be easily restored by mere increased expression of other subunits.

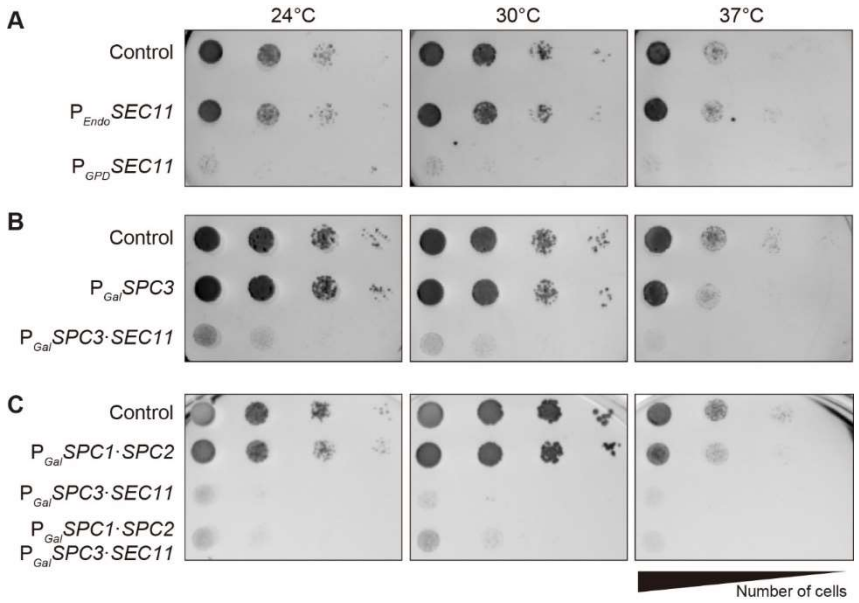

Figure 25. Overexpressed Sec11 severely compromises yeast cell viability. (A-C) W303-1α cells were transformed with indicated vectors. Transformants were spotted on selective medium with serial dilution and grown for 2-3 days under indicated temperature. (A) SC medium –Ura, 2% glucose. (B) SC medium –Ura, 2% galactose. (C) SC medium –Leu –Ura, 2% galactose.

2. 10. Human SPCS1 does not rescue the loss of Spc1 in yeast

SPCS1, the mammalian homolog of yeast Spc1, shares 50% sequence similarity and the same topology with Spc1 (Fig. 26A) (49). Concerning such homology, it was of interest whether SPCS1 functions similarly. To test functional conservation of SPCS1 to Spc1 in yeast, *SPCS1* was subcloned under overexpression promoter for co-expression in *spc1*Δ cells with LepH2 variants and assessed for functional restoration of Spc1 (Figs. 26B and C). The processing state of LepH2(3L) and LepH2(5L) variants was assessed by pulse-chase after radiolabeling (Fig. 26B) and Western blotting (Fig. 26C) as aforedescribed. SPCS1 was co-expressed with LepH2 variants in yeast cells (Fig. 26D). While overexpressed Spc1 successfully rescued the increased processing of LepH2 variants in *spc1*Δ background, SPCS1 did not change cleavage profiles for both LepH2 constructs, showing no trace of functional conservation to Spc1, at least with this series of membrane proteins.

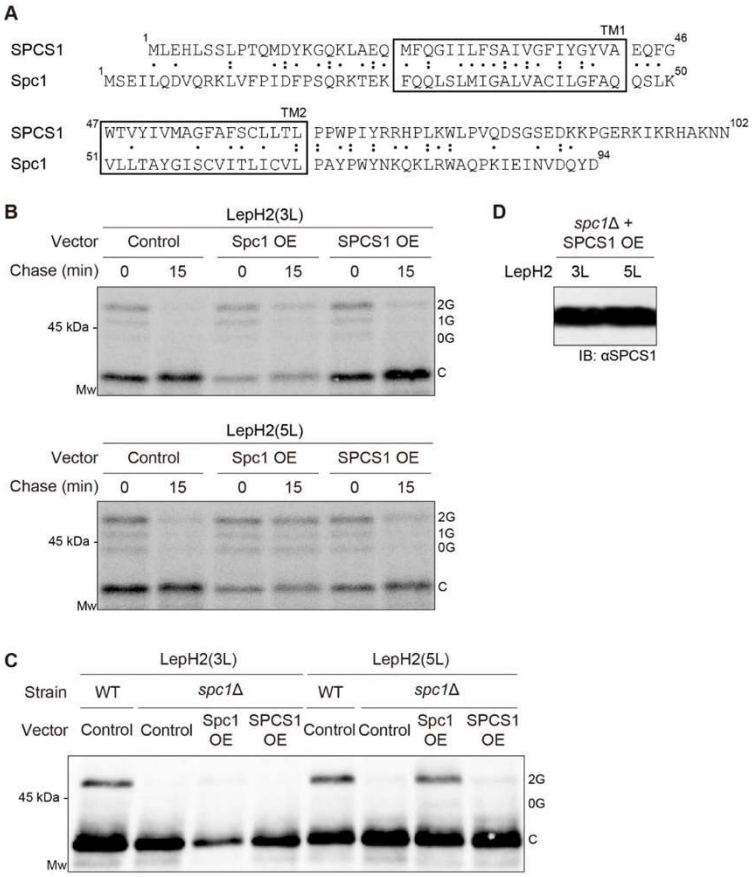

Figure 26. Human SPCS1 does not complement the loss of Spc1 in yeast in the processing of LepH2 model proteins. (A) Amino acid sequences of mammalian SPCS1 and yeast Spc1 are aligned. Identical amino acids are indicated by (:); and similar amino acids are indicated by (•). TM segments are shown in boxes. (B) Indicated LepH2 variants were expressed in *spc1*Δ cells with control, Spc1 OE or SPCS1 OE vector. Transformants were subjected to *in vivo* radiolabeling for 5 min, chased for indicated time points and analyzed as in Fig. 18A. (C) Indicated LepH2 variants were expressed in WT or *spc1*Δ strains with control, Spc1 OE or SPCS1 OE vector and subjected to rapid protein preparation, SDS-PAGE and Western blotting. (D) SPCS1 was overexpressed in *spc1*Δ

cells expressing LepH2 variants in (B), proved by Western blotting against SPCS1. 2G, doubly glycosylated form; 1G, singly glycosylated form; 0G, unglycosylated form; C, cleaved form; OE, overexpression.

2. 11. Depletion of SPCS1 causes differed processing of membrane proteins in mammalian cells

Observing that yeast cells depleted of Spc1 show the increased signal peptidase-mediated processing, it was of interest to examine its mammalian homolog SPCS1 in a similar notion. A previous study reported that mammalian cells lacking SPCS1 exhibits reduced surface expression or secretion of some proteins, suggesting that SPCS1 is closely involved in the biogenesis of proteins in the secretory pathway (55).

A human cell line lacking SPCS1 was constructed by transfection of sgRNA targeting *SPCS1* and the depletion of SPCS1 was confirmed by Western blotting where exogenously induced SPCS1 restored the depletion (Fig. 27A). The steady state level of HLA-E decreased in cells lacking SPCS1, one of the previously reported membrane proteins with reduced surface expression in SPCS1 KO cells (Fig. 27B). HLA-E is a well-studied MHC class 1b antigen and an important modulator of natural killer and cytotoxic T lymphocyte activation (72, 73). A heavy and a light chain constitute HLA-E where the heavy chain contains a signal sequence and a TM segment, which raises possibility of the signal peptidase-mediated processing on this protein. Notably, HLA-E expressed in SPCS1 KO cells migrated slightly slower than that in control cells, which suggests that the protein might have been processed differently due to the loss of SPCS1, perhaps thereby leading to decreased expression.

To monitor phenotype caused by SPCS1 depletion better, a new set of model proteins were employed: a chimera consisting of the N-terminus of human asialoglycoprotein receptor H1 and a part of CPY mature domain in its N- and C-termini, respectively (Fig. 27C). Asialoglycoprotein receptor H1 orients in N_{in}-C_{out} topology, whose TM segment becomes susceptible to the signal peptidase-mediated processing when the 40 residue-long

N-terminal domain is truncated to four residues. Therefore, studies on this protein provided crucial evidence for the length of N-terminus being one of the factors discriminating between cleavable and uncleavable signal sequences. Regarding such previous findings, assessed was the processing of serially truncated N-terminus of asialoglycoprotein receptor H1, with a CPY fusion for easier monitoring.

Four N-terminal variants were induced in either control or SPCS1 KO cell line (Fig. 27C). The shortest variant, A4CPY, efficiently translocated in both cell lines, evidenced by glycosylation state. Nevertheless, the processing differed that the cleaved form decreased upon the depletion of SPCS1. This reduction was seen throughout all tested variants; while the processing decreased in proportion to the N-terminal length increment, considerable amounts of the cleaved form appeared in the control WT cells whereas SPCS1 KO cells did not generate the cleaved form efficiently. Collectively, these data imply that SPCS1 is closely involved in the processing by the signal peptidase complex, yet the underlying mechanism is still elusive and requests further investigation.

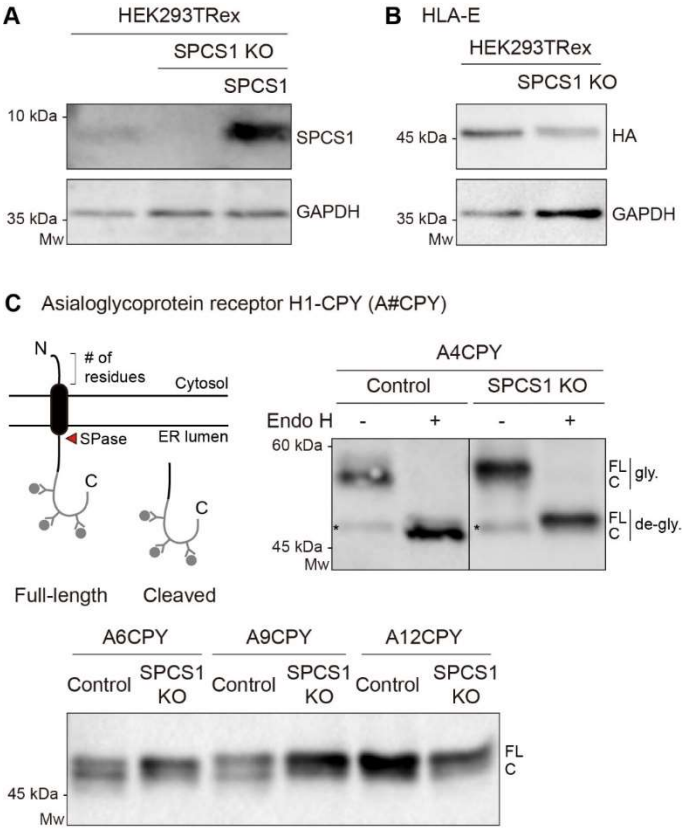

Figure 27. SPCS1 KO cells show differed processing of signal-anchored proteins. (A) HEK293TRex cells were transfected with sgRNA targeting SPCS1 and SPCS1 expression was confirmed by Western blotting against SPCS1. (B) HLA-E-encoding vector was transfected into WT or SPCS1 KO cell lines and analyzed by induction, SDS-PAGE and Western blotting. GAPDH served as a loading control. (C) *Top left*, a schematic of asialoglycoprotein receptor H1-CPY. Asialoglycoprotein receptor H1 part is colored black and CPY part is colored grey. Y and filled circles indicate glycosylation sites. *Top right*, A4CPY were induced in WT or SPCS1 KO cells and analyzed by Endo H digestion, SDS-PAGE and Western blotting. Asterisks indicate untranslocated fraction. *Bottom*, indicated

A#CPY variants were induced and analyzed as A4CPY except for Endo H digestion. FL, full-length form; C, cleaved form; gly., glycosylated form; de-gly., de-glycosylated form.

Chapter IV - Discussion

This thesis presents the research undertaken to explore multiple factors affecting the initial translocation and the early biogenesis of secretory proteins with respect to signal sequences. In particular, *cis*-factors within the signal sequence; the length of *n* region and the hydrophobicity of *h* region and *trans*-factors recognizing and accommodating signal sequences, the SRP, the Sec62/Sec63 complex and the signal peptidase complex have been investigated.

Experimental design and assay systems

The major assay system used in this thesis is an *in vivo* glycosylation assay. Yeast cells expressing test proteins were metabolically radiolabeled for ~5-10 min, short time to capture the early stage of protein biogenesis.

In the first part of the thesis, efficiency of various signal sequence mutants was examined for the ER translocation. Because the ER targeting (*i.e.* the arrival of nascent chains on the ER membrane) and the ER translocation (*i.e.* the opening of the Sec61 channel and C-terminal translocation of test proteins) can be separate, it can be of question whether signal sequence mutants compromised the ER targeting, not the actual translocation. However, recent studies suggest that the ER targeting occurs independent of types of signal sequences, whether being SRP-dependent or -independent in yeast (22). The possibility of compromised ER targeting can be also ruled out concerning the case of LepCC variants (Fig. 16). LepCC variants were all glycosylated upon 5 min of metabolic radiolabeling as shown by Endo H digestion, proving that the ER targeting of these variants were not affected in spite of differed H-segments. Hence, the observed difference of glycosylation states is attributable to differed translocation efficiency of signal sequence variants.

It can be also questioned whether differed glycosylation is due to differed recognition

by glycosylation machinery depending on test proteins. The enzyme complex responsible for the N-linked glycosylation contains two substrate-recognizing subunits, Ost3 and Ost6, which have overlapping but distinct substrate specificity (74). However, test CPY variants examined throughout thesis have always showed a uniform pattern of glycosylation as a single glycosylated band, despite the presence of four glycosylation sequons. If the OST complex recognized any of glycosylation sequon differently, inefficiently glycosylated bands would have been observed, but that was not the case. Hence, the recognition of test proteins by glycosylation machinery is unlikely to differ among the signal sequence mutants.

By virtue of relatively short labeling time, protein degradation is thought to be minimal. This assay system thus provides sensitive and accurate monitoring of protein translocation and processing, which would be difficult to observe by other approaches.

Systematic assessment of signal sequence characteristics and the required Sec translocon components

In the first part of the thesis, I aimed to investigate the elaborate interplay between translocation-involved factors and signal sequences of varying features. For a systematic approach, the signal sequence of a well-characterized yeast protein CPY was modified in the *n* and *h* regions. This variation was to reflect a great diversity in signal sequences where no consensus sequence is found and both the *n* region length and the *h* region hydrophobicity greatly vary. This series of CPY-based model proteins were assessed for translocation competence in WT and translocon mutant strains to unravel the relationship between the signal sequence characteristics and the required components.

The first observation was that lengthened *n* region (>10-12 residues) impairs efficient

translocation, which can be recovered by more hydrophobic *h* region (Figs. 7 and 8). This relationship between *n* and *h* regions was also observed with hundreds of yeast proteins, showing this is a general phenomenon (Fig. 10).

Secondly observed was dynamic requirement of translocation-involved factors (Fig. 28A). The SRP guides hydrophobic signal sequences followed by action of Sec71; less hydrophobic signal sequences are guided by cytosolic chaperones and necessitate all four subunits of the Sec62/Sec63 complex. Some signal sequences of intermediate characteristics required both the SRP and the Sec62/Sec63 complex (Fig. 11). Profiles of natural proteins also support this analysis (Fig. 12).

Signal sequences bind to and open the Sec61 channel to initiate the translocation across the ER membrane (27). Although the lateral gate helices 2 and 7 of the Sec61 translocon was shown to be the signal sequence-binding site, it is more likely that a great diversity of signal sequences confers different binding as suggested in the recent MD simulation study (30). Signal sequences of lower translocation competence (*e.g.* long *n* region length or less hydrophobic *h* region) may be insufficient to open the Sec61 translocon on their own, necessarily requiring the Sec62/Sec63 components for proper binding and subsequent opening of the Sec61 translocon. Recent studies have shown that the opening of the Sec61 translocon and efficient ER translocation involve Sec62, Sec63 and BiP (14, 31, 32). Albeit weak phenotype, the loss of Sec72 also reduced translocation of such signal sequences. As cytosolic Hsp70s bind to Sec72 and hand over a nascent chain, this only cytosolic subunit may serve to stabilize soluble domains within the Sec62/Sec63 complex until translocated (13).

Most intriguingly, Sec71 is critical for all types of signal sequences, including ones with hydrophobic *h* region that were independent of other subunits of the Sec62/Sec63 complex.

A previous study has shown that Sec71 is required for translocation of proteins containing long *n* region that would enter the Sec61 translocon forming a looped conformation. These observations raise some speculations that Sec71 might be involved in the inversion of hydrophobic signal sequences.

Recent researches have shown that most ER-targeted proteins are co-translationally translocate (22, 23). This co-translational translocation is in part achieved by a pioneering round of ribosome targeted to the ER and the Sec61 complex, which allows nascent chains made from following polysomes on the same mRNA to be already present on the ER membrane and to directly contact the SEC complex (Fig. 28A). An earlier structural study showed that a ribosome and Sec62 bind to the same site on Sec61, which makes co-operation of the SRP and the Sec62/Sec63 complex seem unlikely (29). However, recent researches mentioned above explain that most signal sequences can directly contact the SEC complex without direct binding with the SRP. That is, the SRP and the Sec62/Sec63 complex may function in sequential manner that the SRP brings the pioneering RNC to the SEC complex and the Sec62/Sec63 complex subsequently handles the pioneering and following RNCs as polysomes. This was observed for the yeast signal anchor protein Spc3(75).

Supported by emerging evidence showing that most ER-translocating nascent chains are co-translationally targeted to the ER membrane, the role of signal sequences is to initiate the translocation, rather than a mere ER targeting. During the opening and initiation of the Sec translocon, some signal sequences require a subset of the Sec62/Sec63 complex subunits while other signal sequences require another group of subunits (Fig. 28B). That is, heterogeneous signal sequences differently encounter the Sec61 translocon and dynamically associate with the Sec62/Sec63 complex components for efficient translocation initiation *in vivo*.

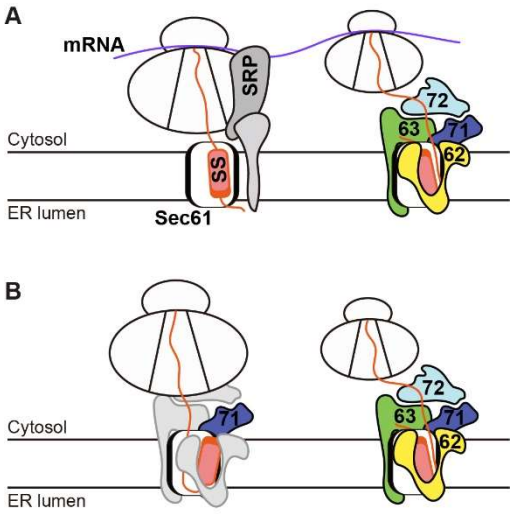

Figure 28. Schematics of translocon components handling signal sequences. (A) Most ER-targeted nascent chains translocate co-translationally, where formation of polysomes on a single transcript contribute in part. A pioneering RNC is guided by the SRP and engage the Sec61 translocon. Following polysomes localize to the ER via mRNA transcript and translate nascent chains whose signal sequence can directly bind to the SEC complex. (B) During signal sequence-mediated channel opening, some signal sequences require the Sec62/Sec63 components to a lesser extent (*left*), while other signal sequences greatly depend on the Sec62/Sec63 components (*right*).

Defining roles of the signal peptidase complex subunits on substrate processing

In the second part of the thesis, I aimed to explore roles of the signal peptidase complex subunits on translocation and processing of ER-incoming substrates. To this end, Spc1 and Spc2 were deleted from yeast genome one at a time and a series of model proteins was expressed, which allows monitoring of protein translocation and the signal peptidase-mediated processing.

Neither membrane insertion nor translocation was affected by the loss of Spc1 or Spc2, demonstrating that these subunits are not involved in translocation of tested proteins. However, the signal peptidase-mediated processing of membrane proteins greatly increased in cells lacking Spc1 or Spc2.

Since Spc1 and Spc2 harbor similar topology and similar effect on model proteins in one's absence, their redundancy was tested by overexpressing one subunit in cells lacking the other (Fig. 21). However, the cleavage profiles of tested proteins were not altered, showing that Spc1 and Spc2 cannot compensate the loss of each other, suggesting that they carry out distinct functions. In fact, *spc1*Δ and *spc2*Δ cells exhibited somewhat different substrate range (Fig. 14). Importantly, this increased processing was not due to aberrant cleavage at non-canonical cleavage sites (Fig. 18).

I focused on Spc1, the absence of which resulted in increased processing while overexpression led to decreased processing of hydrophobic model proteins. Observing that Spc1 protects hydrophobic segments from the signal peptidase-mediated processing, I hypothesized that Spc1 might interact with hydrophobic segments and co-immunoprecipitation experiment was carried out (Fig. 22). The data show that Spc1 interacts with a TM segment of membrane proteins, suggesting that Spc1 may shield TM segments from being presented to the active site of the signal peptidase complex, thereby

protecting them. Assuming that Spc1 protects TM segments, it was one of the major interests to find natural proteins affected by Spc1. Some single-spanning and multi-spanning membrane proteins indeed showed markedly reduced steady state levels in cells lacking Spc1, despite cleaved fragments were not detected (Figs. 23 and 24). This implicates that Spc1 affects biogenesis and stability of membrane proteins.

This study suggests that Spc1 functions to recognize and protect TM segments from the signal peptidase-mediated processing, thereby contributing to accurate substrate selection for the signal peptidase complex and quality control of membrane proteins. Deletion of Spc1 does not impair the signal peptidase activity nor cell growth in yeast, indicating that Spc1 may have a limited role in yeast. However, deletion of SPC12 (Spc1 homolog) in *D. melanogaster* led to developmental lethality (53). The number of single-pass membrane proteins is increased in multicellular organisms where cell-cell interactions are mediated through the receptors in the plasma membrane and many of them are single-pass membrane proteins (76). Hence, protecting TM segments from misrecognition by the signal peptidase complex may be more crucial and the role of Spc1 may be more prominent in higher eukaryotes.

In line with this, it has been observed that expression of surface receptors was reduced in SPCS1-deficient human cell-lines (55). Indeed, HLA-E, one of the surface proteins shown in Fig. 28 exhibited greatly reduced steady state level in SPCS1 KO cell lines compared to that in control cells. Moreover, model single-spanning proteins showed differed processing in SPCS1 KO human cell lines, illustrating regulatory roles of SPCS1 on the signal peptidase-mediated processing, yet the results are still preliminary.

Perspectives

Evolutionarily acquired surprising diversity of signal sequences and physiological significance

Since the establishment of the signal hypothesis (4), the concept of 'signal sequence equivalence' was widely accepted that a unified and general function of signal sequences is to mediate the ER transport of a nascent chain, despite their great diversity in length, hydrophobicity, amino acid composition and charged residues (33). This notion was supported by the discoveries where some signal sequences are interchangeable among proteins and organisms and tolerant of a broad range of mutations (34, 36-38). However, if all signal sequences are equivalent and functionally degenerate, why have they evolved with such diversity?

In fact, signal sequences are not equivalent. Signal sequences vary in functional efficiency of mediating translocation and are not always interchangeable (77-79). Previous studies showed that some protein isoforms produced by alternative splicing or alternative translation start sites differ only in their signal sequences, hinting that signal sequences are the key regulatory factors that alter final products (80-82).

Thus, it is likely that each signal sequence carries its own information and therefore that cells might have acquired signal sequence-dependent regulatory mechanisms during the course of evolution. For instance, signal sequences control engaged translocon and associating complexes and regulate folding and maturation of proteins (83-86). Some signal sequences depend on the SRP whereas others do not, so that thereby signal sequences determine the route of the ER transport. The relevant importance is well illustrated in studies with prokaryotic cells, where interchange of signal sequences changes the targeting route and alters production efficiency of functional proteins (87). Protein secretion efficiency depends on signal sequences in eukaryotic organisms, implying that

each mature protein would require a proper signal sequence for desired production (88-91). Meanwhile, signal sequences determine engaging chaperones for successful protein maturation (92, 93). These findings illustrate how cells have evolved to exploit signal sequence-dependent modulations for efficient protein biogenesis.

Some signal sequences are inefficiently cleaved by the signal peptidase complex, despite the ER translocation (94). This inefficient processing can be rather beneficial for cells; cells can produce both soluble and membrane-bound isoforms from a single gene transcript, thereby lowering cost for efficient protein production. Differed in final localization, those isoforms may also have distinct functions from each other. For example, a soluble form of carboxypeptidase E processes hormones and neuropeptides while a membrane-bound form acts as a sorting receptor for prohormones and proneuropeptides toward the regulated secretory pathway (95-98). Therefore, signal sequences can induce efficient protein biogenesis paradoxically by inefficient processing, resulting in multiple isoforms differing in localization and function. Further, cells utilize signal sequences as independent proteins after cleavage from corresponding mature proteins. For example, a cleaved signal peptide of HLA-E is presented in the cell surface and functions as an indicator of intracellular status, signaling against nature killer cells (72, 73).

These so far observations clearly show that signal sequences have evolved diversely to modulate the association with translocation components, folding and secretion of proteins, production of multiple protein isoforms and generation of functional fragments after cleavage, and therefore provide explanations for their surprising complexity acquired over evolution.

Translocational misregulation of proteins caused by defective translocation machinery

and relevant diseases

As roughly one third of the proteins in the cell necessitate the ER for maturation and localization to their final destinations, cells maintain balanced ER homeostasis and competent translocation state. Hence, the ER translocation machinery is one of the major factors that affect cellular physiology. For being a general translocation machinery, disrupted Sec complex can exhibit pleiotropic effects resulting from improper translocation of proteins whose translocation depends on the Sec complex. Sec61, Sec62 and Sec63 are conserved and are linked to some human diseases. One of such examples is related to human Sec63.

The onset of Sec63-dependent autosomal-dominant polycystic liver and kidney diseases is attributed to failed ER translocation of polycystins (99-103). Like its yeast homolog, human Sec63 mediates the ER translocation of some proteins along with the luminal chaperone BiP (104, 105). However, another subset of proteins translocate independent of Sec63 function (106). This difference in Sec63-dependence might result from differences among signal sequences; the underlying mechanism of Sec63 action has been elusive to date. The disease-responsible polycystins-1 and -2 are multi-spanning membrane proteins where polycystin-1 carries a cleavable signal sequence whereas polycystin-2 carries a non-cleavable signal sequence (107-109). It was observed that the loss of Sec63 severely compromised the biogenesis of polycystin-1 but only moderately affected that of polycystin-2 (110). Hence, Sec63 may differently handle polycystins depending on their signal sequence characteristics, in line with the findings of this thesis. Altering Sec63-dependence might be a therapeutic concept for such diseases. Although how mutations of Sec63 affect the biogenesis of polycystins and the onset of disease has not been clearly identified, if mutated Sec63 impairs its interaction with Sec61, it may lead to unsuccessful translocation of polycystins. If that is the case, enhancing the interaction between Sec63

and Sec61 using small molecules may repress such defective ER translocation, alleviating symptoms.

Incomplete processing of signal sequences generates isoforms differed in localization and function

Many proteins have been reported to exist in both membrane and soluble compartments in mammalian cells (94, 111). The signal peptidase-mediated processing provides one mechanism for such dual existence; examples include carboxypeptidase E (112), tumor necrosis factor (TNF) (113, 114) and Golgi sialyltransferase (115, 116). The signal peptidase complex cleaves a TM segment of these proteins and releases soluble domain, resulting in differently localizing isoforms. Interestingly, these proteins are functionally distinct depending on soluble or membrane-bound state. In addition to carboxypeptidase E mentioned above, in case of TNF, a soluble form is secreted by activated monocyte and involved in response to septic shock and cachexia whereas a membrane-bound form is expressed on the cell surface and kills tumor cells and virus-infected cells by cell-to-cell contact (113, 117).

Asialoglycoprotein receptor (ASGP R) subunit H2 also generates multiple isoforms depending on the signal peptidase-mediated processing, which consequently differ in localization and function (118). This heterogeneity first arises from alternative splicing: mRNA of H2a isoform encodes five extra ectodomain residues located three residues away from the TM segment, which somehow results in the signal peptidase-mediated processing and secretion out of cells. Conversely, H2b isoform does not undergo the processing and form surface receptor complex with another subunit H1 in the plasma membrane (118). ASGP R subunits are exclusively expressed in liver cells and hepatoma cells and ASGP R

complex functions in the selective binding and internalization of Gal/GalNAc glycoproteins. Related to this function, a recent model describes a role of a secreted form of H2a in binding of Gal/GalNAc glycoproteins from the circulation and transport to liver (119). Because cleavable H2a and membrane-anchored H2b differ in the ectodomain region away from the TM segment and the cleavage site, it is surprising to find that this discrepancy results in differed processing state. Thus, this interesting finding illustrates that proteins contain regulatory signals for the signal peptidase-mediated processing not only within a signal sequence or a cleavage site, but also within a mature domain, mechanism of which deserves further investigation.

These examples show that proteins with both soluble and membrane-bound forms own distinct function, implying the significance of regulated signal peptidase-mediated formation of the two different isoforms. Regarding that the signal peptidase complex is responsible for production of isoforms, it is important to understand exact mechanisms concerning the action of the signal peptidase complex.

Roles of the signal peptidase complex subunits for viral protein biogenesis

Many viruses hijack and utilize host machineries for their protein synthesis, assembly and replication. Viral proteome comprises several structural and nonstructural proteins including Spike, Envelope and Capsid coat proteins, translated as a single polypeptide. This viral polypeptide is processed by viral and host proteases to release multiple viral proteins; the signal peptidase complex is one of the exploited host machineries.

A chemical named cavinafungin was found to inhibit growth of dengue and Zika virus in infected cells where the signal peptidase complex was identified as the target, implicating the importance of the signal peptidase complex on viral protein production (120). More

specifically, several studies showed that SPCS1 and SPCS2 (mammalian homologs of Spc1 and Spc2, respectively) are involved in biogenesis of some viral proteins. SPCS1 was found to interact with viral membrane proteins of hepatitis C virus (HCV) and Japanese encephalitis virus (JEV) through TM segments and this interaction was critical for viral replication (121, 122). Flaviviruses including JEV and Zika significantly depend on SPCS1 and SPCS2 for replication (55, 123, 124).

Unique and intriguing feature of the signal peptidase-mediated processing of viral signal sequences is that the cleavage is suboptimal and delayed, unlike the rapid, co-translocational signal peptidase-mediated processing in principle (86, 125-131). Intriguingly, "optimally" delayed processing was shown necessary for virus assembly and/or release: both increased and decreased processing of viral proteins by the signal peptidase complex compromised viral propagation (83, 125, 128, 132-135). For example, enhancing the signal peptidase-mediated processing of C-prM of flavivirus reduced incorporation of the viral proteins into the viral particles (132, 135, 136), implicating that C-prM protein should associate with the host signal peptidase long enough to ensure the assembly among C-prM and other viral proteins. Similar to the case of flavivirus, the delayed cleavage by the host signal peptidase seen in HCV proteins also appears to contribute to viral assembly (132, 133, 135).

Based on these findings and regarding the roles of Spc1 shown in yeast cells, SPCS1 might influence viral protein biogenesis and assembly by affecting the kinetics and the efficiency of the signal peptidase-mediated processing. As the depletion of Spc1 resulted in enhanced processing of membrane proteins, especially those with an internal signal sequence, the human signal peptidase lacking SPCS1 may show increased processing of such proteins as well. Regarding that both full-length and mature forms of C protein in flavivirus are necessary, increased processing might result in loss or lack of full-length C

and compromise viral formation. Further, the assembly of functional nucleocapsid requires the simultaneous recruitment of multiple viral proteins in the same assembly site, which is aided by the delayed signal peptidase-mediated processing. The depletion of SPCS1, and also SPCS2, might shift this controlled processing of viral proteins and thereby disrupt the formation of functional viral nucleocapsid.

The observation that SPCS1 interacts with viral membrane proteins and is critical for their biogenesis (55, 121, 122) suggest that SPCS1 might act as a membrane-embedded chaperoning factor for some viral signal sequences and TM segments. NS2B protein of JEV was found to associate with SPCS1 via their TM segments and this association is critical for viral formation (121, 122). This is in line with the findings with the yeast homolog Spc1 in this thesis that TM segments of membrane proteins associate with Spc1 and insert into the lipid bilayer, otherwise they may undergo the signal peptidase-mediated processing or degradation.

Hinted from the yeast homolog Spc1 researches, depletion of SPCS1 might result in increased processing of TM segments of viral proteins and/or enhanced degradation of viral proteins, both of which compromise functional viral proteome. SPCS1 could be a good therapeutic target because it is a noncatalytic subunit of the signal peptidase complex (55, 124).

Physiological significance of Spc2

Roles of Spc2 are relatively less investigated in this thesis. Although less unveiled, functions of Spc2 have been implicated in various physiological aspects. Spc2 is not only involved in the signal peptidase-mediated processing ((47) and this thesis), but also in the ER translocation of proteins (137) and unfolded protein response (UPR) (138, 139). A

recent study by Hosomi *et al*. showed that CPY without its signal sequence could translocate across the ER membrane and reach the target organelle vacuole, which is greatly enhanced in the absence of Spc2 (137). This finding suggests that the ER translocon machinery somehow allows translocation of incorrect substrates lacking a proper signal sequence where Spc2 acts as a gatekeeper against such aberrant translocation. This finding can also be related to previous high throughput researches, which showed the connection of Spc2 and UPR: cells lacking Spc2 exhibited elevated UPR (138); meanwhile, activated UPR upregulates Spc2 (139). Collectively, these observations imply that Spc2 might function to sort ER-incoming substrates properly; otherwise, aberrant substrates enter the ER, undermining the ER homeostasis and provoking ER stress and UPR.

Outlook

Taken together, various translocation components are involved to handle and accommodate signal sequences properly. Since many diseases are caused by defective biogenesis and quality control of numerous proteins on the secretory pathway, researches concerning this basic cellular mechanism are of great significance.

It was fascinating to find that some proteins are generated in multiple isoforms, differed in localization and function; *e.g.* carboxypeptidase E, TNF, asialoglycoprotein receptor and viral structural proteins. Utilizing incomplete cleavage as one of the strategies to produce multiple proteins from a single gene may help the cell to diversify its proteome. For example, liver cell-exclusive asialoglycoprotein receptor H2 is produced as a membrane-bound H2b and a cleaved H2a. A cleaved, secreted version H2a circulates in the blood system and recognizes substrates, thereby liver can handle Gal/GalNAc glycoproteins all over the body. I found it interesting that separate organs and systems maintain overall

homeostasis via controlled proteolytic processing of proteins. I would like to expand and apply so-far findings on such controlled processing of differently functional isoforms.

REFERENCES

1. G. Palade, Intracellular aspects of the process of protein synthesis. *Science* **189**, 347-358 (1975).
2. S. Mitaku, M. Ono, T. Hirokawa, S. Boon-Chieng, M. Sonoyama, Proportion of membrane proteins in proteomes of 15 single-cell organisms analyzed by the SOSUI prediction system. *Biophys Chem* **82**, 165-171 (1999).
3. E. Wallin, G. von Heijne, Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. *Protein Sci* **7**, 1029-1038 (1998).
4. G. Blobel, B. Dobberstein, Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *J Cell Biol* **67**, 835-851 (1975).
5. S. Wittke, M. Dunnwald, N. Johnsson, Sec62p, a component of the endoplasmic reticulum protein translocation machinery, contains multiple binding sites for the Sec-complex. *Mol Biol Cell* **11**, 3859-3871 (2000).
6. R. J. Deshaies, S. L. Sanders, D. A. Feldheim, R. Schekman, Assembly of yeast Sec proteins involved in translocation into the endoplasmic reticulum into a membrane-bound multisubunit complex. *Nature* **349**, 806-808 (1991).
7. K. Plath, T. A. Rapoport, Spontaneous release of cytosolic proteins from posttranslational substrates before their transport into the endoplasmic reticulum. *J Cell Biol* **151**, 167-178 (2000).
8. J. Tyedmers *et al.*, Homologs of the yeast Sec complex subunits Sec62p and Sec63p are abundant proteins in dog pancreas microsomes. *Proc Natl Acad Sci U S A* **97**, 7214-7219 (2000).
9. S. K. Lyman, R. Schekman, Binding of secretory precursor polypeptides to a translocon subcomplex is regulated by BiP. *Cell* **88**, 85-96 (1997).
10. B. Misselwitz, O. Staeck, K. E. Matlack, T. A. Rapoport, Interaction of BiP with the J-domain of the Sec63p component of the endoplasmic reticulum protein translocation complex. *J Biol Chem* **274**, 20110-20115 (1999).
11. S. S. Vembar, M. C. Jonikas, L. M. Hendershot, J. S. Weissman, J. L. Brodsky, J domain co-chaperone specificity defines the role of BiP during protein translocation. *J Biol Chem* **285**, 22484-22494 (2010).
12. H. Fang, N. Green, Nonlethal sec71-1 and sec72-1 mutations eliminate proteins associated with the Sec63p-BiP complex from S. cerevisiae. *Mol Biol Cell* **5**, 933-942 (1994).

13. A. Tripathi, E. C. Mandon, R. Gilmore, T. A. Rapoport, Two alternative binding mechanisms connect the protein translocation Sec71-Sec72 complex with heat shock proteins. *J Biol Chem* **292**, 8007-8018 (2017).
14. S. Itskanov, K. M. Kuo, J. C. Gumbart, E. Park, Stepwise gating of the Sec61 protein-conducting channel by Sec63 and Sec62. *Nat Struct Mol Biol* **28**, 162-172 (2021).
15. P. Walter, G. Blobel, Purification of a membrane-associated protein complex required for protein translocation across the endoplasmic reticulum. *Proc Natl Acad Sci U S A* **77**, 7112-7116 (1980).
16. D. T. Ng, J. D. Brown, P. Walter, Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J Cell Biol* **134**, 269-278 (1996).
17. M. G. Waters, G. Blobel, Secretory protein translocation in a yeast cell-free system can occur posttranslationally and requires ATP hydrolysis. *J Cell Biol* **102**, 1543-1550 (1986).
18. L. L. Randall, Translocation of domains of nascent periplasmic proteins across the cytoplasmic membrane is independent of elongation. *Cell* **33**, 231-240 (1983).
19. T. Ast, G. Cohen, M. Schuldiner, A network of cytosolic factors targets SRP-independent proteins to the endoplasmic reticulum. *Cell* **152**, 1134-1145 (2013).
20. J. H. Reithinger, J. E. Kim, H. Kim, Sec62 protein mediates membrane insertion and orientation of moderately hydrophobic signal anchor proteins in the endoplasmic reticulum (ER). *J Biol Chem* **288**, 18058-18067 (2013).
21. S. J. Jung, J. E. Kim, J. H. Reithinger, H. Kim, The Sec62-Sec63 translocon facilitates translocation of the C-terminus of membrane proteins. *J Cell Sci* **127**, 4270-4278 (2014).
22. C. H. Jan, C. C. Williams, J. S. Weissman, Principles of ER cotranslational translocation revealed by proximity-specific ribosome profiling. *Science* **346**, 1257521 (2014).
23. J. W. Chartron, K. C. Hunt, J. Frydman, Cotranslational signal-independent SRP preloading during membrane targeting. *Nature* **536**, 224-228 (2016).
24. B. Van den Berg *et al.*, X-ray structure of a protein-conducting channel. *Nature* **427**, 36-44 (2004).
25. E. Park, T. A. Rapoport, Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annu Rev Biophys* **41**, 21-40 (2012).
26. B. Jungnickel, T. A. Rapoport, A posttargeting signal sequence recognition event in the endoplasmic reticulum membrane. *Cell* **82**, 261-270 (1995).
27. K. Plath, W. Mothes, B. M. Wilkinson, C. J. Stirling, T. A. Rapoport, Signal sequence recognition in posttranslational protein transport across the yeast ER

membrane. *Cell* **94**, 795-807 (1998).
28. R. M. Voorhees, R. S. Hegde, Structure of the Sec61 channel opened by a signal sequence. *Science* **351**, 88-91 (2016).
29. R. M. Voorhees, I. S. Fernandez, S. H. Scheres, R. S. Hegde, Structure of the mammalian ribosome-Sec61 complex to 3.4 A resolution. *Cell* **157**, 1632-1643 (2014).
30. P. Bhadra, L. Yadhanapudi, K. Romisch, V. Helms, How does Sec63 affect the conformation of Sec61 in yeast? *PLoS Comput Biol* **17**, e1008855 (2021).
31. X. Wu, C. Cabanos, T. A. Rapoport, Structure of the post-translational protein translocation machinery of the ER membrane. *Nature* **566**, 136-139 (2019).
32. S. Itskanov, E. Park, Structure of the posttranslational Sec protein-translocation channel complex from yeast. *Science* **363**, 84-87 (2019).
33. G. von Heijne, Signal sequences. The limits of variation. *J Mol Biol* **184**, 99-105 (1985).
34. L. M. Gierasch, Signal sequences. *Biochemistry* **28**, 923-930 (1989).
35. G. von Heijne, Patterns of amino acids near signal-sequence cleavage sites. *Eur J Biochem* **133**, 17-21 (1983).
36. K. Talmadge, S. Stahl, W. Gilbert, Eukaryotic signal sequence transports insulin antigen in Escherichia coli. *Proc Natl Acad Sci U S A* **77**, 3369-3373 (1980).
37. M. Muller, I. Ibrahimi, C. N. Chang, P. Walter, G. Blobel, A bacterial secretory protein requires signal recognition particle for translocation across mammalian endoplasmic reticulum. *J Biol Chem* **257**, 11860-11863 (1982).
38. C. A. Kaiser, D. Preuss, P. Grisafi, D. Botstein, Many random sequences functionally replace the secretion signal sequence of yeast invertase. *Science* **235**, 312-317 (1987).
39. G. von Heijne, How signal sequences maintain cleavage specificity. *J Mol Biol* **173**, 243-251 (1984).
40. I. Nilsson, P. Whitley, G. von Heijne, The COOH-terminal ends of internal signal and signal-anchor sequences are positioned differently in the ER translocase. *J Cell Biol* **126**, 1127-1132 (1994).
41. H. Fang, C. Mullins, N. Green, In addition to SEC11, a newly identified gene, SPC3, is essential for signal peptidase activity in the yeast endoplasmic reticulum. *J Biol Chem* **272**, 13152-13158 (1997).
42. C. VanValkenburgh, X. Chen, C. Mullins, H. Fang, N. Green, The catalytic mechanism of endoplasmic reticulum signal peptidase appears to be distinct from most eubacterial signal peptidases. *J Biol Chem* **274**, 11519-11525 (1999).
43. C. Zwizinski, W. Wickner, Purification and characterization of leader (signal)

peptidase from Escherichia coli. *J Biol Chem* **255**, 7973-7977 (1980).
44. R. E. Dalbey, G. Von Heijne, Signal peptidases in prokaryotes and eukaryotes--a new protease family. *Trends Biochem Sci* **17**, 474-478 (1992).
45. J. T. YaDeau, C. Klein, G. Blobel, Yeast signal peptidase contains a glycoprotein and the Sec11 gene product. *Proc Natl Acad Sci U S A* **88**, 517-521 (1991).
46. G. S. Shelness, G. Blobel, Two subunits of the canine signal peptidase complex are homologous to yeast SEC11 protein. *J Biol Chem* **265**, 9512-9519 (1990).
47. W. Antonin, H. A. Meyer, E. Hartmann, Interactions between Spc2p and other components of the endoplasmic reticulum translocation sites of the yeast Saccharomyces cerevisiae. *J Biol Chem* **275**, 34068-34072 (2000).
48. G. Greenburg, G. S. Shelness, G. Blobel, A subunit of mammalian signal peptidase is homologous to yeast SEC11 protein. *J Biol Chem* **264**, 15762-15765 (1989).
49. H. Fang, S. Panzner, C. Mullins, E. Hartmann, N. Green, The homologue of mammalian SPC12 is important for efficient signal peptidase activity in Saccharomyces cerevisiae. *J Biol Chem* **271**, 16460-16465 (1996).
50. E. A. Evans, R. Gilmore, G. Blobel, Purification of microsomal signal peptidase as a complex. *Proc Natl Acad Sci U S A* **83**, 581-585 (1986).
51. A. M. Liaci *et al.*, Structure of the Human Signal Peptidase Complex Reveals the Determinants for Signal Peptide Cleavage. *bioRxiv* 10.1101/2020.11.11.378711, 2020.2011.2011.378711 (2020).
52. H. A. Meyer, E. Hartmann, The yeast SPC22/23 homolog Spc3p is essential for signal peptidase activity. *J Biol Chem* **272**, 13159-13164 (1997).
53. E. Haase Gilbert, S. J. Kwak, R. Chen, G. Mardon, Drosophila signal peptidase complex member Spase12 is required for development and cell differentiation. *PLoS One* **8**, e60908 (2013).
54. C. Mullins, H. A. Meyer, E. Hartmann, N. Green, H. Fang, Structurally related Spc1p and Spc2p of yeast signal peptidase complex are functionally distinct. *J Biol Chem* **271**, 29094-29099 (1996).
55. R. Zhang *et al.*, A CRISPR screen defines a signal peptide processing pathway required by flaviviruses. *Nature* **535**, 164-168 (2016).
56. K. Kitada, E. Yamaguchi, M. Arisawa, Cloning of the Candida glabrata TRP1 and HIS3 genes, and construction of their disruptant strains by sequential integrative transformation. *Gene* **165**, 203-206 (1995).
57. C. J. Stirling, J. Rothblatt, M. Hosobuchi, R. Deshaies, R. Schekman, Protein translocation mutants defective in the insertion of integral membrane proteins into the endoplasmic reticulum. *Mol Biol Cell* **3**, 129-142 (1992).
58. M. S. Longtine *et al.*, Additional modules for versatile and economical PCR-based

gene deletion and modification in Saccharomyces cerevisiae. *Yeast* **14**, 953-961 (1998).

59. C. Lundin, H. Kim, I. Nilsson, S. H. White, G. von Heijne, Molecular code for protein insertion in the endoplasmic reticulum membrane is similar for N(in)-C(out) and N(out)-C(in) transmembrane helices. *Proc Natl Acad Sci U S A* **105**, 15702-15707 (2008).
60. D. G. Gibson, Enzymatic assembly of overlapping DNA fragments. *Methods Enzymol* **498**, 349-361 (2011).
61. I. Nilsson, G. von Heijne, A signal peptide with a proline next to the cleavage site inhibits leader peptidase when present in a sec-independent protein. *FEBS Lett* **299**, 243-246 (1992).
62. J. Cui *et al.*, Competitive Inhibition of the Endoplasmic Reticulum Signal Peptidase by Non-cleavable Mutant Preprotein Cargos. *J Biol Chem* **290**, 28131-28140 (2015).
63. G. A. Barkocy-Gallagher, P. J. Bassford, Jr., Synthesis of precursor maltose-binding protein with proline in the +1 position of the cleavage site interferes with the activity of Escherichia coli signal peptidase I in vivo. *J Biol Chem* **267**, 1231-1238 (1992).
64. M. Lussier, A. M. Sdicu, F. Bussereau, M. Jacquet, H. Bussey, The Ktr1p, Ktr3p, and Kre2p/Mnt1p mannosyltransferases participate in the elaboration of yeast O- and N-linked carbohydrate chains. *J Biol Chem* **272**, 15527-15531 (1997).
65. L. Mallet, F. Bussereau, M. Jacquet, Nucleotide sequence analysis of an 11.7 kb fragment of yeast chromosome II including BEM1, a new gene of the WD-40 repeat family and a new member of the KRE2/MNT1 family. *Yeast* **10**, 819-831 (1994).
66. P. Berninsone, J. J. Miret, C. B. Hirschberg, The Golgi guanosine diphosphatase is required for transport of GDP-mannose into the lumen of Saccharomyces cerevisiae Golgi vesicles. *J Biol Chem* **269**, 207-211 (1994).
67. A. Donella-Deana, S. Ostojic, L. A. Pinna, S. Barbaric, Specific dephosphorylation of phosphopeptides by the yeast alkaline phosphatase encoded by PHO8 gene. *Biochim Biophys Acta* **1177**, 221-228 (1993).
68. S. Otte *et al.*, Erv41p and Erv46p: new components of COPII vesicles involved in transport between the ER and Golgi complex. *J Cell Biol* **152**, 503-518 (2001).
69. F. Reggiori, H. R. Pelham, Sorting of proteins into multivesicular bodies: ubiquitin-dependent and -independent targeting. *EMBO J* **20**, 5176-5186 (2001).
70. C. Navarre *et al.*, Purification and complete sequence of a small proteolipid associated with the plasma membrane H(+)-ATPase of Saccharomyces cerevisiae.

J Biol Chem **267**, 6425-6428 (1992).
71. A. M. Marini, S. Soussi-Boudekou, S. Vissers, B. Andre, A family of ammonium transporters in Saccharomyces cerevisiae. *Mol Cell Biol* **17**, 4282-4293 (1997).
72. F. Borrego, M. Ulbrecht, E. H. Weiss, J. E. Coligan, A. G. Brooks, Recognition of human histocompatibility leukocyte antigen (HLA)-E complexed with HLA class I signal sequence-derived peptides by CD94/NKG2 confers protection from natural killer cell-mediated lysis. *J Exp Med* **187**, 813-818 (1998).
73. V. M. Braud *et al.*, HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. *Nature* **391**, 795-799 (1998).
74. B. L. Schulz *et al.*, Oxidoreductase activity of oligosaccharyltransferase subunits Ost3p and Ost6p defines site-specific glycosylation efficiency. *Proc Natl Acad Sci U S A* **106**, 11061-11066 (2009).
75. S. J. Jung, J. E. H. Kim, T. Junne, M. Spiess, H. Kim, Cotranslational Targeting and Posttranslational Translocation can Cooperate in Spc3 Topogenesis. *J Mol Biol* **433**, 167109 (2021).
76. I. D. Pogozheva, A. L. Lomize, Evolution and adaptation of single-pass transmembrane proteins. *Biochim Biophys Acta Biomembr* **1860**, 364-377 (2018).
77. S. J. Kim, D. Mitra, J. R. Salerno, R. S. Hegde, Signal sequences control gating of the protein translocation channel in a substrate-specific manner. *Dev Cell* **2**, 207-217 (2002).
78. C. G. Levine, D. Mitra, A. Sharma, C. L. Smith, R. S. Hegde, The efficiency of protein compartmentalization into the secretory pathway. *Mol Biol Cell* **16**, 279-291 (2005).
79. D. T. Rutkowski, V. R. Lingappa, R. S. Hegde, Substrate-specific regulation of the ribosome- translocon junction by N-terminal signal sequences. *Proc Natl Acad Sci U S A* **98**, 7823-7828 (2001).
80. K. Clark, E. Hammond, P. Rabbitts, Temporal and spatial expression of two isoforms of the Dutt1/Robo1 gene in mouse development. *FEBS Lett* **523**, 12-16 (2002).
81. M. Damodarasamy, M. Zhang, K. Dienger, F. X. McCormack, Two rat surfactant protein A isoforms arise by a novel mechanism that includes alternative translation initiation. *Biochemistry* **39**, 10189-10195 (2000).
82. T. Nakajima *et al.*, Functional analysis of a mutation occurring between the two in-frame AUG codons of human angiotensinogen. *J Biol Chem* **274**, 35749-35755 (1999).
83. E. L. Snapp *et al.*, Structure and topology around the cleavage site regulate post-translational cleavage of the HIV-1 gp160 signal peptide. *Elife* **6** (2017).

84. R. S. Hegde, S. W. Kang, The concept of translocational regulation. *J Cell Biol* **182**, 225-232 (2008).
85. R. S. Hegde, H. D. Bernstein, The surprising complexity of signal sequences. *Trends Biochem Sci* **31**, 563-571 (2006).
86. B. Martoglio, B. Dobberstein, Signal sequences: more than just greasy peptides. *Trends Cell Biol* **8**, 410-415 (1998).
87. R. Freudl, Signal peptides for recombinant protein secretion in bacterial expression systems. *Microb Cell Fact* **17**, 52 (2018).
88. N. Roongsawang *et al.*, A Novel Potential Signal Peptide Sequence and Overexpression of ER-Resident Chaperones Enhance Heterologous Protein Secretion in Thermotolerant Methylotrophic Yeast Ogataea thermomethanolica. *Appl Biochem Biotechnol* **178**, 710-724 (2016).
89. G. P. Lin-Cereghino *et al.*, The effect of alpha-mating factor secretion signal mutations on recombinant protein expression in Pichia pastoris. *Gene* **519**, 311-317 (2013).
90. J. A. Rakestraw, S. L. Sazinsky, A. Piatesi, E. Antipov, K. D. Wittrup, Directed evolution of a secretory leader for the improved expression of heterologous proteins and full-length antibodies in Saccharomyces cerevisiae. *Biotechnol Bioeng* **103**, 1192-1201 (2009).
91. T. Kjeldsen *et al.*, Secretory expression of human albumin domains in Saccharomyces cerevisiae and their binding of myristic acid and an acylated insulin analogue. *Protein Expr Purif* **13**, 163-169 (1998).
92. Z. Frenkel, M. Shenkman, M. Kondratyev, G. Z. Lederkremer, Separate roles and different routing of calnexin and ERp57 in endoplasmic reticulum quality control revealed by interactions with asialoglycoprotein receptor chains. *Mol Biol Cell* **15**, 2133-2142 (2004).
93. S. Sun, X. Li, M. Mariappan, Signal Sequences Encode Information for Protein Folding in the Endoplasmic Reticulum. *bioRxiv* 10.1101/2020.06.04.133884, 2020.2006.2004.133884 (2020).
94. M. R. Ehlers, J. F. Riordan, Membrane proteins with soluble counterparts: role of proteolysis in the release of transmembrane proteins. *Biochemistry* **30**, 10065-10074 (1991).
95. L. D. Fricker, S. H. Snyder, Purification and characterization of enkephalin convertase, an enkephalin-synthesizing carboxypeptidase. *J Biol Chem* **258**, 10950-10955 (1983).
96. D. R. Cool *et al.*, Carboxypeptidase E is a regulated secretory pathway sorting receptor: genetic obliteration leads to endocrine disorders in Cpe(fat) mice. *Cell*

88, 73-83 (1997).
97. H. Lou *et al.*, Sorting and activity-dependent secretion of BDNF require interaction of a specific motif with the sorting receptor carboxypeptidase e. *Neuron* **45**, 245-255 (2005).
98. R. McGirr, L. Guizzetti, S. Dhanvantari, The sorting of proglucagon to secretory granules is mediated by carboxypeptidase E and intrinsic sorting signals. *J Endocrinol* **217**, 229-240 (2013).
99. S. V. Fedeles, A. R. Gallagher, S. Somlo, Polycystin-1: a master regulator of intersecting cystic pathways. *Trends Mol Med* **20**, 251-260 (2014).
100. M. J. Janssen, J. Salomon, R. H. Te Morsche, J. P. Drenth, Loss of heterozygosity is present in SEC63 germline carriers with polycystic liver disease. *PLoS One* **7**, e50324 (2012).
101. E. Waanders, R. H. te Morsche, R. A. de Man, J. B. Jansen, J. P. Drenth, Extensive mutational analysis of PRKCSH and SEC63 broadens the spectrum of polycystic liver disease. *Hum Mutat* **27**, 830 (2006).
102. J. P. Drenth, J. A. Martina, R. van de Kerkhof, J. S. Bonifacino, J. B. Jansen, Polycystic liver disease is a disorder of cotranslational protein processing. *Trends Mol Med* **11**, 37-42 (2005).
103. S. Davila *et al.*, Mutations in SEC63 cause autosomal dominant polycystic liver disease. *Nat Genet* **36**, 575-577 (2004).
104. N. Schauble *et al.*, BiP-mediated closing of the Sec61 channel limits Ca2+ leakage from the ER. *EMBO J* **31**, 3282-3296 (2012).
105. S. Lang *et al.*, Different effects of Sec61alpha, Sec62 and Sec63 depletion on transport of polypeptides into the endoplasmic reticulum of mammalian cells. *J Cell Sci* **125**, 1958-1969 (2012).
106. D. Gorlich, T. A. Rapoport, Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell* **75**, 615-630 (1993).
107. J. L. Bakeberg *et al.*, Epitope-tagged Pkhd1 tracks the processing, secretion, and localization of fibrocystin. *J Am Soc Nephrol* **22**, 2266-2277 (2011).
108. T. Mochizuki *et al.*, PKD2, a gene for polycystic kidney disease that encodes an integral membrane protein. *Science* **272**, 1339-1342 (1996).
109. P. D. Wilson, Polycystin: new aspects of structure, function, and regulation. *J Am Soc Nephrol* **12**, 834-845 (2001).
110. S. V. Fedeles *et al.*, A genetic interaction network of five genes for human polycystic kidney and liver diseases defines polycystin-1 as the central determinant of cyst formation. *Nat Genet* **43**, 639-647 (2011).

111. J. Wegrzyn, J. Lee, J. M. Neveu, W. S. Lane, V. Hook, Proteomics of neuroendocrine secretory vesicles reveal distinct functional systems for biosynthesis and exocytosis of peptide hormones and neurotransmitters. *J Proteome Res* **6**, 1652-1665 (2007).
112. D. Parkinson, Two soluble forms of bovine carboxypeptidase H have different NH2-terminal sequences. *J Biol Chem* **265**, 17101-17105 (1990).
113. C. Perez *et al.*, A nonsecretable cell surface mutant of tumor necrosis factor (TNF) kills by cell-to-cell contact. *Cell* **63**, 251-258 (1990).
114. R. Ishisaka *et al.*, A part of the transmembrane domain of pro-TNF can function as a cleavable signal sequence that generates a biologically active secretory form of TNF. *J Biochem* **126**, 413-420 (1999).
115. J. C. Paulson, J. Weinstein, E. L. Ujita, K. J. Riggs, P. H. Lai, The membrane-binding domain of a rat liver Golgi sialyltransferase. *Biochem Soc Trans* **15**, 618-620 (1987).
116. K. J. Colley, E. U. Lee, J. C. Paulson, The signal anchor and stem regions of the beta-galactoside alpha 2,6-sialyltransferase may each act to localize the enzyme to the Golgi apparatus. *J Biol Chem* **267**, 7784-7793 (1992).
117. M. Kriegler, C. Perez, K. DeFay, I. Albert, S. D. Lu, A novel form of TNF/cachectin is a cell surface cytotoxic transmembrane protein: ramifications for the complex physiology of TNF. *Cell* **53**, 45-53 (1988).
118. S. Tolchinsky, M. H. Yuk, M. Ayalon, H. F. Lodish, G. Z. Lederkremer, Membrane-bound versus secreted forms of human asialoglycoprotein receptor subunits. Role of a juxtamembrane pentapeptide. *J Biol Chem* **271**, 14496-14503 (1996).
119. J. Hu, J. Liu, D. Yang, M. Lu, J. Yin, Physiological roles of asialoglycoprotein receptors (ASGPRs) variants and recent advances in hepatic-targeted delivery of therapeutic molecules via ASGPRs. *Protein Pept Lett* **21**, 1025-1030 (2014).
120. D. Estoppey *et al.*, The Natural Product Cavinafungin Selectively Interferes with Zika and Dengue Virus Replication by Inhibition of the Host Signal Peptidase. *Cell Rep* **19**, 451-460 (2017).
121. L. Ma *et al.*, Host Factor SPCS1 Regulates the Replication of Japanese Encephalitis Virus through Interactions with Transmembrane Domains of NS2B. *J Virol* **92** (2018).
122. R. Suzuki *et al.*, Signal peptidase complex subunit 1 participates in the assembly of hepatitis C virus through an interaction with E2 and NS2. *PLoS Pathog* **9**, e1003589 (2013).
123. N. J. Barrows *et al.*, Dual roles for the ER membrane protein complex in flavivirus infection: viral entry and protein biogenesis. *Sci Rep* **9**, 9711 (2019).

124. M. Rother, M. Naumann, Signal peptidase complex subunit 1 is an essential Zika virus host factor in placental trophoblasts. *Virus Res* **296**, 198338 (2021).
125. S. Shanmugam, M. Yi, Efficiency of E2-p7 processing modulates production of infectious hepatitis C virus. *J Virol* **87**, 11255-11266 (2013).
126. S. Carrere-Kremer *et al.*, Regulation of hepatitis C virus polyprotein processing by signal peptidase involves structural determinants at the p7 sequence junctions. *J Biol Chem* **279**, 41384-41392 (2004).
127. C. Lin, B. D. Lindenbach, B. M. Pragai, D. W. McCourt, C. M. Rice, Processing in the hepatitis C virus E2-NS2 region: identification of p7 and two distinct E2-specific products with different C termini. *J Virol* **68**, 5063-5073 (1994).
128. C. E. Stocks, M. Lobigs, Signal peptidase cleavage at the flavivirus C-prM junction: dependence on the viral NS2B-3 protease for efficient processing requires determinants in C, the signal peptide, and prM. *J Virol* **72**, 2141-2149 (1998).
129. C. E. Stocks, M. Lobigs, Posttranslational signal peptidase cleavage at the flavivirus C-prM junction in vitro. *J Virol* **69**, 8123-8126 (1995).
130. S. M. Amberg, A. Nestorowicz, D. W. McCourt, C. M. Rice, NS2B-3 proteinase-mediated processing in the yellow fever virus structural region: in vitro and in vivo studies. *J Virol* **68**, 3794-3802 (1994).
131. G. Blobel, B. Dobberstein, Transfer of proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. *J Cell Biol* **67**, 852-862 (1975).
132. M. Lobigs, E. Lee, M. L. Ng, M. Pavy, P. Lobigs, A flavivirus signal peptide balances the catalytic activity of two proteases and thereby facilitates virus morphogenesis. *Virology* **401**, 80-89 (2010).
133. S. Denolly *et al.*, The amino-terminus of the hepatitis C virus (HCV) p7 viroporin and its cleavage from glycoprotein E2-p7 precursor determine specific infectivity and secretion levels of HCV particle types. *PLoS Pathog* **13**, e1006774 (2017).
134. J. Rana *et al.*, Role of Capsid Anchor in the Morphogenesis of Zika Virus. *J Virol* **92** (2018).
135. M. Lobigs, E. Lee, Inefficient signalase cleavage promotes efficient nucleocapsid incorporation into budding flavivirus membranes. *J Virol* **78**, 178-186 (2004).
136. E. Lee, C. E. Stocks, S. M. Amberg, C. M. Rice, M. Lobigs, Mutagenesis of the signal sequence of yellow fever virus prM protein: enhancement of signalase cleavage In vitro is lethal for virus production. *J Virol* **74**, 24-32 (2000).
137. A. Hosomi *et al.*, The ER-associated protease Ste24 prevents N-terminal signal peptide-independent translocation into the endoplasmic reticulum in

Saccharomyces cerevisiae. *J Biol Chem* **295**, 10406-10419 (2020).
138. M. C. Jonikas *et al.*, Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. *Science* **323**, 1693-1697 (2009).
139. K. J. Travers *et al.*, Functional and genomic analyses reveal an essential coordination between the unfolded protein response and ER-associated degradation. *Cell* **101**, 249-258 (2000).

국문 초록

진핵 세포 단백체의 약 40%는 분비 경로를 따르는 분비 및 막관통 단백질로 구성되어 있어, 이러한 단백질들의 생합성과 품질 관리 기작에 대한 이해의 중요성을 시사한다. 대부분의 경우, 단백질 아미노기 말단의 수송 신호서열이 작용하여 단백질을 소포체막의 단백질 수송 채널(Sec61 트랜스로콘)로 이끌고 수송 과정을 시작하며 이 일련의 수송 과정은 Sec62/Sec63 복합체를 필요로 하기도 한다. 한편, 수송 신호 서열은 소포체 내로 수송되는 도중 신호 서열 절단 복합체의 작용으로 절단되기도 한다.

수송 신호서열은 매우 놀랍도록 다양한 서열을 가지는데, 이렇게 다양한 서열을 갖는 수송 신호서열들이 어떻게 한 개의 보편적인 단백질 수송 채널 Sec61 에 의해 인식되고 수송되는지는 명확히 알려져 있지 않다. 이 질문에 답하기 위해 본 연구의 첫 번째 부분에서는 수송 신호서열이 갖는 다양한 특성과 그에 따른 소포체 수송 효율을 효모 세포 내 방사성동위원소 표지법을 이용하여 측정하였다. 이를 통해, 수송 신호서열의 아미노기 말단의 길이가 증가함에 따라 수송 효율이 감소하며, 반대로 수송 신호서열의 소수성이 증가하면 수송 효율이 회복되는 것을 관찰하였다.

또한 수송 복합체 돌연변이체에서 다양한 수송 신호서열의 효율을 관찰함으로써 각 소단위의 기능을 살펴보고자 하였다. 소수성이 높지 않은 수송 신호 서열들은 Sec62/Sec63 복합체에 대한 의존성을 보이는 가운데, 소수성이 높은 수송 신호서열은 특이적으로 Sec71 소단위에 대한 의존성을

보였다. 이러한 결과는 Sec62/Sec63 복합체의 각 소단위들이 다양한 특성의 수송 신호서열의 효율적인 수송을 위해 역동적으로 기능함을 보여 준다.

본 연구의 두 번째 부분에서는 신호서열 절단 복합체의 두 소단위인 Spc1 과 Spc2 가 기질 절단 과정 중 수행하는 역할에 대하여 탐구하였다. 모델 단백질과 세포 내 방사성 동위원소 표지법을 이용하여, Spc1 과 Spc2 가 결실되는 경우 신호서열 절단이 증가하며, 반대로 Spc1 이 과발현되는 경우에는 신호서열 절단이 감소하는 것을 관찰하였다. 또한 Spc1 과 막단백질의 물리적 상호작용과 다수의 막단백질의 양이 Spc1 결실세포주에서 감소하는 것을 관찰하였다. 이러한 결과는 Spc1 이 신호서열 절단 복합체로부터 막관통영역을 보호함으로써 신호서열 절단 복합체의 기질 선택을 조절함을 시사한다.

종합적으로 본 연구는 신호인식입자, Sec62/Sec63 복합체 및 신호서열 절단 복합체 등의 복합적인 수송 매개체들이 다양한 특성의 신호서열과 막관통영역을 인식하고 수송 및 절단을 매개함으로써 올바른 생합성과 품질 관리에 기여함을 보여준다.

주요어: 신호서열, 소포체 수송, 단백질 수송 채널, 신호서열 절단 복합체, Spc1, 막단백질

학번: 2012-20325