Dissertation zur Erlangung des Doktorgrades
der Fakultät für Chemie und Pharmazie
der Ludwig-Maximilians-Universität München

Cryo-EM analysis of the active post-translational Sec translocon

Tsai-Hsuan Weng
aus
Taipei, Taiwan

2021

Erklärung

Diese Dissertation wurde im Sinne von § 7 der Promotionsordnung vom 28. November 2011 von Herrn Prof. Dr. Roland Beckmann betreut.

Eidesstattliche Versicherung

Diese Dissertation wurde eigenständig und ohne unerlaubte Hilfe erarbeitet.

München, 24.02.2021

_____

Tsai-Hsuan Weng

Dissertation eingereicht am: 25.02.2021

1. Gutachter: Prof. Dr. Roland Beckmann

2. Gutachter: Prof. Dr. Daniel Wilson

Mündliche Prüfung am: 28.04.2021

Summary

In eukaryotes, most secretory and membrane proteins are targeted by an N-terminal signal sequence to the endoplasmic reticulum, where the heterotrimeric Sec61 complex serves as protein-conducting channel. In the post-translational mode, fully synthesized proteins are recognized by a specialized channel, called the Sec complex, consisting of the Sec61 complex and additional Sec62, Sec63, Sec71 and Sec72 subunits. Recent structures of this Sec complex in the idle state revealed the overall architecture in a pre-opened state. In this thesis, a cryo-EM structure of the Sec complex bound to a substrate is presented. The signal sequence was inserted into the lateral gate of Sec61α similar to previous structures, yet, with the gate adopting an even more open conformation. The signal sequence was flanked by two Sec62 transmembrane helices, the cytoplasmic N-terminal domain of Sec62 was more rigidly positioned, and the plug domain was relocated. Together, a near complete and integrated model of the active Sec complex was achieved.

Contents

List of Tables

List of Figures

Abbreviations

°C	Degree Celsius
2D	Two-dimensional
3D	Three-dimensional
A (unit)	Absorbance
Å (unit)	Ångstroem
Amp	Ampicillin
ATP	Adenosine triphosphate
bp	Base pair
Brl	Brr2-like
CHS	Cholesteryl hemisuccinate
CK2	Casein kinase 2
cryo-EM	Cryogenic electron microscopy
C-terminus/C-terminal	Carboxy-terminus/carboxy-terminal
CTF	Contrast transfer function
Da	Dalton
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E. coli*/*E.c.	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EMDB	Electron Microscopy Data Bank
ER	Endoplasmic reticulum
FSC	Fourier shell correlation
GDN	Glyco-diosgenin
HDX	Hydrogen-deuterium exchange
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His_6/His_8	Hexahistidine/octahistidine
HRV	Human rhinovirus
IPTG	Isopropyl β-d-thiogalactosid
Ire1α	Inositol-requiring enzyme 1 α
ITC	Isothermal titration calorimetry
Kan	Kanamycin
kb	Kilobase
kDa	Kilodalton
L	Loop
L (unit)	Liter
LB	Luria broth

LMNG	Lauryl maltose neopentyl glycol
M (unit)	Molar
M. jannaschii	*Methanocaldococcus jannaschii*
mEGFP	Monomeric enhanced green fluorescent protein
MES	2-morpholinoethanesulfonic acid
mg	Milligram
ml	Milliliter
mM	Millimolar
MS	Mass spectrometry
nm	Nanometer
nM	Nanomolar
N-terminus/N-terminal	Amino-terminus/amino-terminal
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCC	Protein conducting channel
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PMSF	Phenylmethylsulfonyl fluoride
ppαF	Prepro-α factor
RBS	Ribosomal binding site
RNA	Ribonucleic acid
RNC	Ribosome-nascent chain complex
S. cerevisiae*/*S.c.	*Saccharomyces cerevisiae*
SAD	Single-wavelength anomalous dispersion
SDS	Sodium dodecyl sulfate
SP	Signal peptide
SR	Signal recognition particle receptor
SRP	Signal recognition particle
ss	Signal sequence
SUMO	Small ubiquitin-like modifier
THF	Two-helix finger
TIC	Translocon on the inner chloroplast membrane
TIM	Translocase of the inner membrane
TM	Transmembrane
TOC	Translocon on the outer chloroplast membrane
TOM	Translocase of the outer membrane
Tris	Tris(hydroxymethyl)aminomethane

tRNA	Transfer ribonucleic acid
Ulp1	Ubiquitin-like-specific protease 1
UPR	Unfolded protein response
WT	Wildtype
YPD	Yeast extract peptone dextrose
µg	Microgram
µl	Microliter
µM	Micromolar

Chapter 1: Introduction

Cells are tiny factories enclosed with a lipid bilayer membrane. Eukaryotic cells have many organelles such as nucleus, endoplasmic reticulum (ER), Golgi apparatus, mitochondria, and/or chloroplast, etc., while prokaryotes lack these membrane-bound subcellular compartments. Newly synthesized proteins destined to these compartments need to be targeted to and translocated across membrane boundaries, because most proteins are synthesized in the cytosol. Although mitochondria and chloroplasts have their own ribosomes for protein synthesis, most of mitochondrial or chloroplast proteins are synthesized in and translocated from the cytosol. Thus, it is important for a nascent protein to know where to go and how to get across membrane boundaries. In eukaryotes, unfolded nascent mitochondrial or chloroplast proteins are transported through the TOM/TIM complex or the TOC/TIC complex into mitochondria or chloroplasts, respectively. Although prokaryotes, unlike eukaryotes, miss membrane-bound subcellular compartments, in either cases, newly synthesized secretory and membrane proteins face the same decisive step—they are targeted to and translocated across the ER membrane in eukaryotes or the cytoplasmic membrane in prokaryotes.

1.1 Protein translocation across the ER membrane

The majority of secretory and membrane proteins is targeted by an N-terminal signal sequence to the Sec61 protein conducting channel (PCC), also called translocon, at the endoplasmic reticulum (ER) membrane in eukaryotes or the SecY channel at the cytoplasmic membrane in prokaryotes (Blobel & Dobberstein, 1975; Rapoport *et al*, 2017; Gemmer & Förster, 2020). The Sec61/SecY channel is gated open by the signal sequence and can facilitate either translocation of proteins into the ER lumen or

insertion of membrane proteins by releasing them into the lipid bilayer through a lateral gate. The Sec61/SecY channel can operate in two modes, either co-translationally bound to a translating ribosome or post-translationally translocating proteins through the ER membrane (cytoplasmic membrane in bacteria) (**Figure 1**).

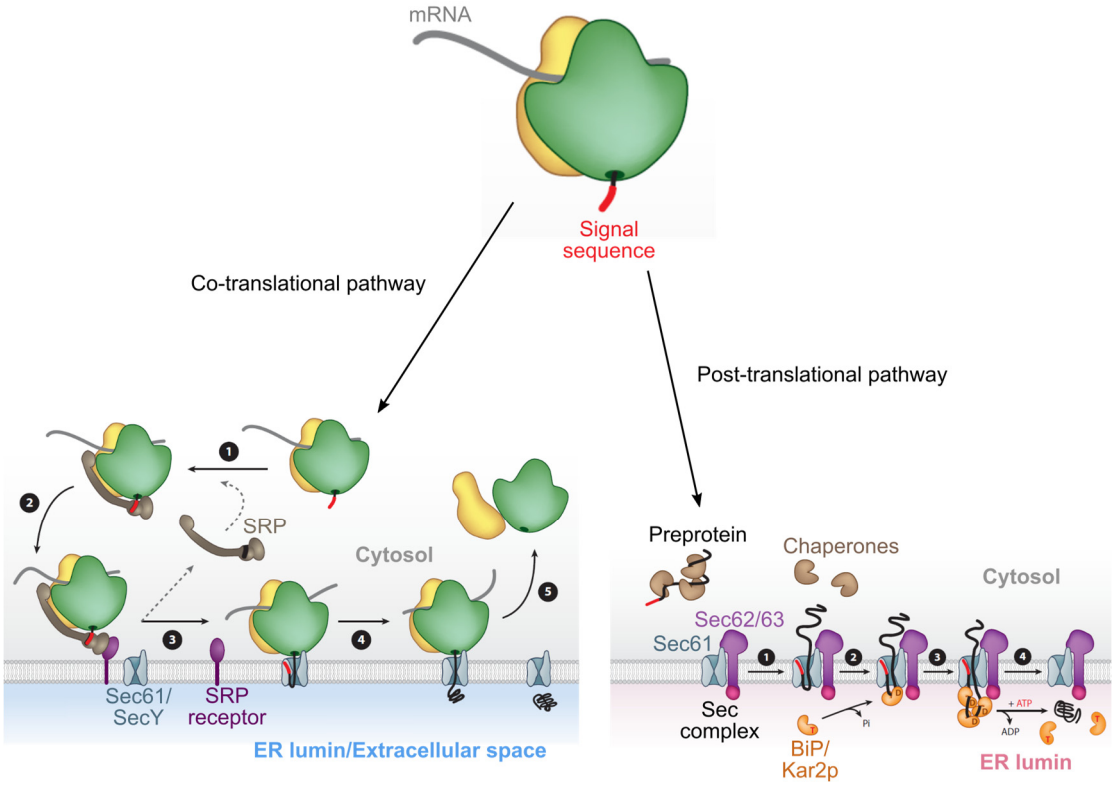


Figure 1. Pathways of newly synthesized protein translocation across the ER membrane.

Scheme showing two pathways for translocation of membrane and secretory proteins across the ER/cytoplasmic membrane. In co-translational pathway (left panel), the signal sequence-containing ribosome-nascent chain complex is recruited to the Sec61/SecY channel by the signal recognition particle (SRP) and the polypeptide is translocated through the channel co-translationally. In the post-translational mode (right panel), in eukaryotes the fully translated polypeptide is protected by chaperons and recruited to the heptameric Sec complex on the ER membrane. The polypeptide is then translocated with the help from ER luminal chaperon BiP (Kar2p in yeast). (Adapted from Park & Rapoport, 2012)

The conserved Sec61/SecY channel is a passive pore where a translocating polypeptide

chain can slide in both directions. To facilitate unidirectional translocation, it requires additional driving force by associating with other factors. In co-translational mode, it is the translating ribosome that provides the driving force; in post-translational mode, an ER luminal Hsp70 protein BiP in eukaryotes prevents the translocating peptides from back sliding (Lyman & Schekman, 1997; Matlack *et al*, 1999), while in prokaryotes an cytoplasmic ATPase SecA pushes the translocating peptide through the channel (Bauer *et al*, 2014; Catipovic *et al*, 2019; Catipovic & Rapoport, 2020).

1.2 Co-translational protein translocation via Sec61 channel

Co-translational protein translocation across the ER membrane in eukaryotes or the cytoplasmic membrane in prokaryotes involves several steps. Biochemical and structural studies have shown insight into the mechanism of co-translational translocation. The first step of co-translational translocation starts with targeting a ribosome-nascent chain complex (RNC) to the Sec61 channel. Second, the Sec61 channel needs to be primed by the binding of a channel partner for translocation. In the co-translational mode, it is the ribosome that activates Sec61/SecY channel. In the third step, the Sec61 channel is opened by signal sequence recognition and translocating polypeptide insertion. Subsequently, the inserted polypeptide is translocated through the central pore, and transmembrane segments—if the translocating substrate is a membrane protein—are integrated into the lipid bilayer via the lateral gate (van den Berg *et al*, 2004; Frauenfeld *et al*, 2011).

1.3 Post-translational protein translocation

In contrast to the co-translational protein translocation, the post-translational translocation is mainly utilized by soluble and secretory proteins. In yeast, these proteins have less hydrophobic signal sequences compared to membrane proteins and

are not recognized by SRP (Ng *et al*, 1996; Ast *et al*, 2013). Instead, these secretory proteins are fully synthesized, released from the ribosome, and kept in an unfolded, translocation competent state by chaperones (Ngosuwan *et al*, 2003). In bacteria, the post-translational protein translocation is carried out by the SecYEG translocon together with the SecA ATPase, which primes the SecY channel open and provides the driving force for polypeptide translocation (Zimmer *et al*, 2008; Bauer *et al*, 2014; Catipovic *et al*, 2019; Catipovic & Rapoport, 2020). However, in eukaryotes, this task is accomplished by the Sec complex, which consists of the Sec61 heterotrimeric complex and the Sec62/Sec63 complex (Deshaies *et al*, 1991; Panzner *et al*, 1995). In addition, instead of SecA, it is the ER luminal chaperone BiP/Kar2p provides the driving force for peptide translocation (Lyman & Schekman, 1997; Matlack *et al*, 1999).

1.4 Architecture of the Sec61 protein conducting channel

In either co-translational or post-translational translocation, Sec61 is the core component in both pathways. The highly conserved heterotrimeric Sec61 channel consists of a large Sec61α subunit (Sec61p in *Saccharomyces cerevisiae*; SecY in *Escherichia coli*) with 10 transmembrane helices (TMs) and two small single-spanning Sec61β (Sbh1p in *S. cerevisiae*; SecG in *E. coli*) and Sec61γ (Sss1p in *S. cerevisiae*; SecE in *E. coli*) (Görlich & Rapoport, 1993; Hanada *et al*, 1994). The structure of the Sec61 channel is well-established in several functional states (Rapoport *et al*, 2017; van den Berg *et al*, 2004; Park *et al*, 2014; Voorhees *et al*, 2014; Voorhees & Hegde, 2016; Kater *et al*, 2019). The first Sec61/SecY crystal structure was solved from an archaea *Methanocaldococcus jannaschii* and represented an idle/closed state (van den Berg *et al*, 2004; **Figure 2**).

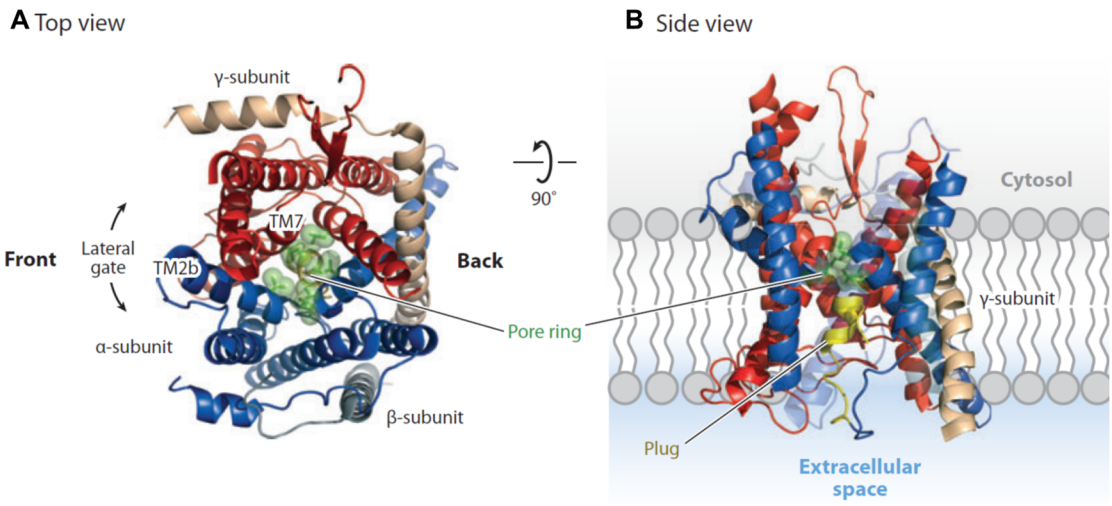


Figure 2. Structure of Sec61/SecY protein conducting channel.
(A) Cytosolic view of the channel. The TMs 1-5 of the N-terminal half of the α subunit are in blue, and the TMs 6-10 of the C-terminal half are in red. The β subunit is in gray and the γ subunit is in beige. The side chains of the pore ring residues are shown as green sticks with spheres. (B) Cut side view of the channel looked from the N-terminal half. The plug domain (yellow) of the α subunit blocks the channel at the luminal/extracellular side of the channel. (Adapted from Park & Rapoport, 2012)

Sec61α

The structure of Sec61 α-subunit is divided into the N-terminal (TMs 1-5) and the C-terminal (TMs 6-10) halves, linked by a loop connecting TM5 and TM6 on the extracellular side (ER luminal side in eukaryotes). When viewed from the cytosolic side, the two halves show a pseudo-symmetry in their TM organization and form a clam-shell-like structure. From the side view, between the two halves Sec61α forms an hourglass-shaped aqueous channel across the membrane with a central constriction dividing cytosolic and luminal sides, and a short helical plug domain between TM1 and TM2 in the luminal region of the channel. The central constriction, also known as pore ring, consists of six residues with their hydrophobic sidechains pointing radially inward. The pore ring was shown to act as the vertical gate for translocating peptides and to

maintain the membrane barrier for small molecules during translocation (Cannon *et al*, 2005; Park & Rapoport, 2011). For channel opening, the plug domain has to relocate to a more peripheral position to allow translocation of the peptide (Zimmer *et al*, 2008). Aside from the channel, there is a gap facing toward the lipid phase, formed between the Sec61α TMs 2 and 3 of the N-terminal half and TMs 7 and 8 of the C-terminal half called the channel's "lateral gate". This lateral gate was observed in different conformations, varying from fully closed in the idle state to partially open when primed with the ribosome or SecA (in prokaryotes) (Zimmer *et al*, 2008; Voorhees *et al*, 2014; Braunger *et al*, 2018; Park *et al*, 2014). Signal sequences and hydrophobic TM segments were found to bind at the lateral gate, resulting in a fully open conformation (Frauenfeld *et al*, 2011; Park *et al*, 2014; Voorhees & Hegde, 2016; Li *et al*, 2016; Ma *et al*, 2019).

Sec61β

The β-subunit of Sec61 (SecG in prokaryotes) consists of an N-terminal cytosolic domain and a C-terminal transmembrane helix. The cytosolic segment is possibly flexible and disordered since it is not visible in available structures. The TM of Sec61β only loosely contacts TMs 1 and 4 of the α-subunit. This may explain why Sec61β is not essential and deletion of it only causes partial translocation defect on the Sec61 channel (Kalies *et al*, 1998).

Sec61γ

The γ-subunit of Sec61 channel (SecE in procaryotes) consists of two α-helical segments. The N-terminal amphipathic helix lies on the cytosolic surface of the membrane and contacts the C-terminal region of the α-subunit of Sec61. It is followed by a short hinge connecting to the long C-terminal transmembrane segment. The TM

spans across the membrane in a diagonal way, contacting Sec61α through the TMs 1, 5, and 6 of the N-terminal half as well as TM10 of the C-terminal half. This makes the γ-subunit clamp together both halves of the α-subunit.

1.5 Targeting polypeptides to Sec61 channel

In co-translational translocation, the process of targeting a RNC to the Sec61 channel is mediated by the signal recognition particle (SRP) and the SRP receptor (SR) on the ER membrane (cytoplasmic membrane in bacteria; **Figure 3**). SRP is a highly conserved ribonucleoprotein composed of SRP RNA and one (in bacteria) or more (in archaea and eukaryotes) SRP proteins (**Figure 3**A). It consists of two domains: The S domain, responsible for signal sequence recognition and SR docking, as well as the less conserved Alu domain, which delays translation elongation (Walter & Blobel, 1981; Mason *et al*, 2000). Several cryo-electron microscopy (cryo-EM) and crystal structures show that the SRP recognizes the hydrophobic signal or signal anchor (SA) sequence as it emerges from the ribosome exit tunnel (Halic *et al*, 2006; Janda *et al*, 2010; Voorhees & Hegde, 2015). The binding of SRP to the signal sequence or the SA is through SRP54, a protein in the universally conserved SRP core in the S domain. SRP54 utilizes a hydrophobic pocket in its methionine-rich (M) domain to recognize the hydrophobic sequence of the nascent chain. While the elongation of polypeptide translation is slowed by the Alu domain of SRP, which competes with the eEF2 (Walter & Blobel, 1981; Voorhees & Hegde, 2015), the RNC-SRP complex is recruited to the ER membrane by the interaction between the GTPase (G) domain of SRP54 and the G domain of SR. GTP-binding on both domains is required for their interaction (Egea *et al*, 2004), which then leads to GTP hydrolysis and transfer of the signal sequence to the Sec61 translocon.

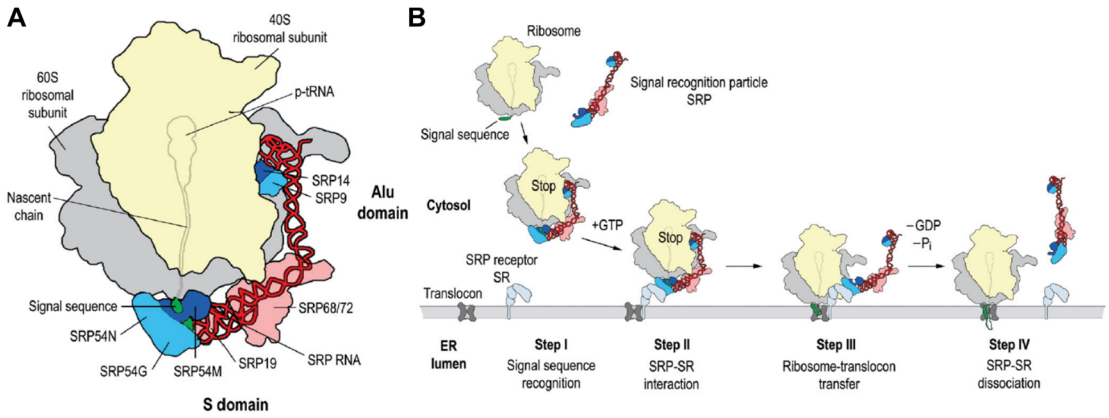


Figure 3.Recognition and targeting of RNC to the ER membrane by SRP. (A) Schematic overview of the mammalian SRP bound to the RNC and the signal sequence emerging from the ribosome exit tunnel. The universally conserved SRP core (SRP54 and SRP RNA helix 8) sits at the exit tunnel and SRP54 binds to the signal sequence with its M domain (SRP54M). The 80S and 40S ribosomal subunits are in gray and yellow, respectively. The outline of the peptidyl-tRNA is shown and the signal sequence is labeled in green. The components of SRP are labeled as follows: SRP RNA, red; SRP54N and G, turquoise; SRP54M, dark blue; SRP19 and SRP68/72, pink; SRP14, dark blue; SRP9, turquoise. (B) Detailed scheme of the co-translational targeting by the SRP system. (Adapted from Wild *et al*, 2004)

Unlike the co-translational pathway, in the post-translational pathway in eukaryotes, the proteins are fully synthesized and released from the ribosome. In yeast, they are protected by cytosolic chaperon Ssa1 and remain competent for translocation (Becker *et al*, 1996; Ngosuwan *et al*, 2003). Biochemical and structural studies indicate that Ssa1-protected peptides may be recruited to the post-translational Sec complex by interacting with Sec71, a subunit of the Sec62/63 complex (Tripathi *et al*, 2017). In some bacteria such as *E. coli*, post-translational targeting of a (partly) unfolded protein to the SecY channel is conducted by the cytosolic chaperon SecB, which binds its substrate in order to prevent it from aggregation (Hartl *et al*, 1990). SecB does not exist in other bacteria and it is still unclear how a polypeptide is targeted post-translationally to the cytoplasmic membrane.

1.6 Priming Sec61 channel for peptide translocation

The Sec61 channel is primed for translocation upon the binding of a ribosome. Cryo-EM structures of Sec61 bound to an idle ribosome without a translocating peptide show that Sec61 translocon binds to the ribosome by utilizing the cytosolic loops between TMs 6 and 7 (loop 6/7) and those between TMs 8 and 9 (loop 8/9) in the α-subunit and the N-terminal helix of the γ-subunit (Voorhees *et al*, 2014; Gogala *et al*, 2014). The ribosome interacts with the C-terminal half of Sec61 through the backbone of the 28S and ribosomal proteins uL23 and uL29, anchoring the translocon at the exit tunnel. Ribosome binding causes conformational changes at the cytosolic loops 6/7 and 8/9 of Sec61α. The conformational changes propagate through the TMs, causing a slight opening of the lateral gate (**Figure 4**). Notably, only the cytosolic portion of the lateral gate is open, and the plug domain remains unaltered compared to the crystal structure of the idle SecY complex, making Sec61 ready for signal sequence engagement.

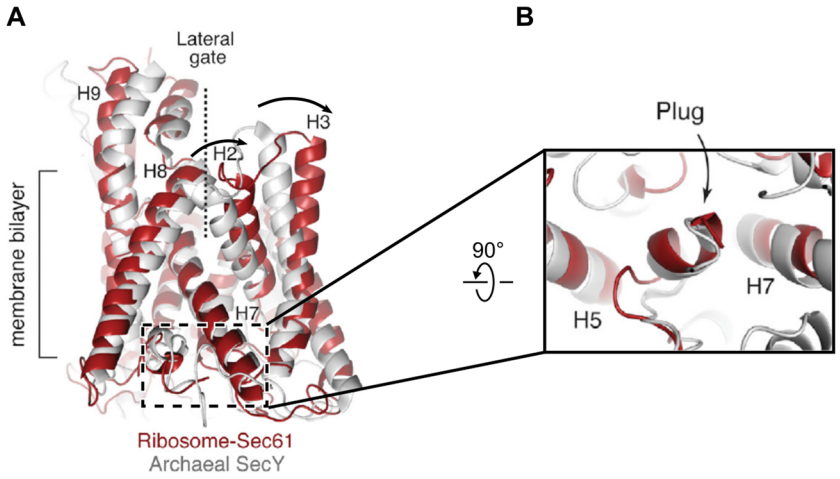


Figure 4. Conformational changes of Sec61α upon ribosome binding.
(A) Structural comparison between the ribosome-bound Sec61 (red) and closed archaeal SecY (gray). The TMs 2 and 3 of the ribosome-bound Sec61 move away from TMs 7 and 8, causing the lateral gate partially open in the cytosolic side. (B) Close-up view of the plug domain shows that in the ribosome-bound Sec61 (red), it remains in the same position at that in the closed archaeal SecY (gray). (Figure adapted from Voorhees *et al*, 2014)

A similar priming event is observed in a complex of SecYEG with the post-translational motor SecA. The crystal structure of the SecY-SecA complex (without translocating substrate) shows how SecA primes the SecY channel for protein translocation by binding to both N- and C-terminal halves of SecY (Zimmer *et al*, 2008). This extensive interaction involves the C-terminal tail and the loop between TMs 8 and 9 in the C-terminal half as well as the loop between TMs 2 and 3 in the N-terminal half of SecY (**Figure 5**A). SecA binding to SecY induces conformational changes of TMs 8 and 9, causing a 5-Å wide opening of the lateral gate, which is significantly more open than that in the ribosome-primed Sec61 channel. The plug domain is also displaced from the center of the channel toward the extracellular side but still closes the channel.

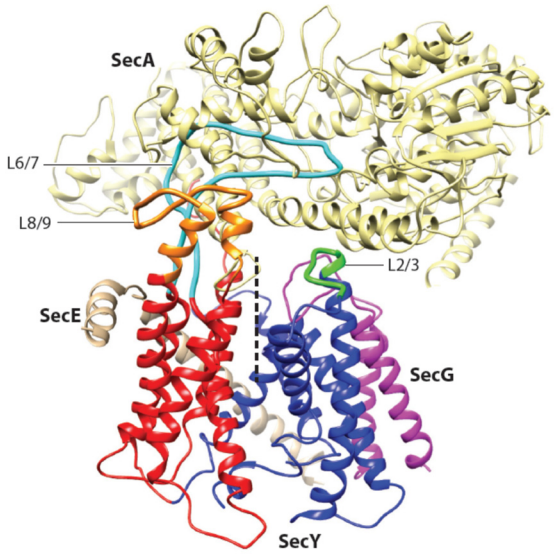


Figure 5. Structure of SecA-SecY complex.
(A) Overview of the crystal structure of the SecA-SecY complex. SecA (yellow) interacts with the cytosolic loops L2/3 (green), L6/7 (cyan) and L8/9 (orange) of SecY. The N-terminal half and the C-terminal half of SecY is in red and blue, respectively. SecG is in magenta and SecE is in beige. (Figure from Rapoport *et al*, 2017)

In the eukaryotic post-translational translocation, the Sec61 channel is associated with the Sec62/63 complex, together called the Sec complex. In association with the Sec62/63 complex, the lateral gate of Sec61α adopts an even more open state as compared to the other primed-state structures (Wu *et al*, 2019; Itskanov & Park, 2019). The architecture of the Sec complex will be discussed in detail later (see Session 1.9).

1.7 Signal sequence recognition and polypeptide insertion

The slightly opened lateral gate of the ribosome-primed Sec61 translocon allows hydrophobic signal sequence recognition and subsequent insertion of a polypeptide. The structural basis of Sec61 engaging a signal sequence in the co-translational mode is revealed by several cryo-EM structures (Frauenfeld *et al*, 2011; Park *et al*, 2014; Voorhees & Hegde, 2016; Kater *et al*, 2019). The so far best-resolved cryo-EM structure from Voorhees & Hegde, 2016 shows a Sec61 translocon interacting with the signal sequence of the secretory protein preprolactin emerging from a translationally stalled RNC. The nascent chain inserts from the ribosome exit tunnel into the channel in a looped configuration with the signal sequence intercalated in the groove at the lateral gate, consistent with the previous cross-linking studies (Mothes *et al*, 1994; Plath *et al*, 1998, 2003). Compared to the ribosome-primed Sec61 structure, the signal sequence-engaged Sec61 undergoes a slight conformational change, with the N-terminal half of the α-subunit moving outward, resulting in a more open lateral gate. In addition, it causes the displacement of the six conserved pore ring residues from their normally planar conformation, as well as the plug domain moving away from the channel. These conformational changes open the Sec61 channel for the translocation of peptides.

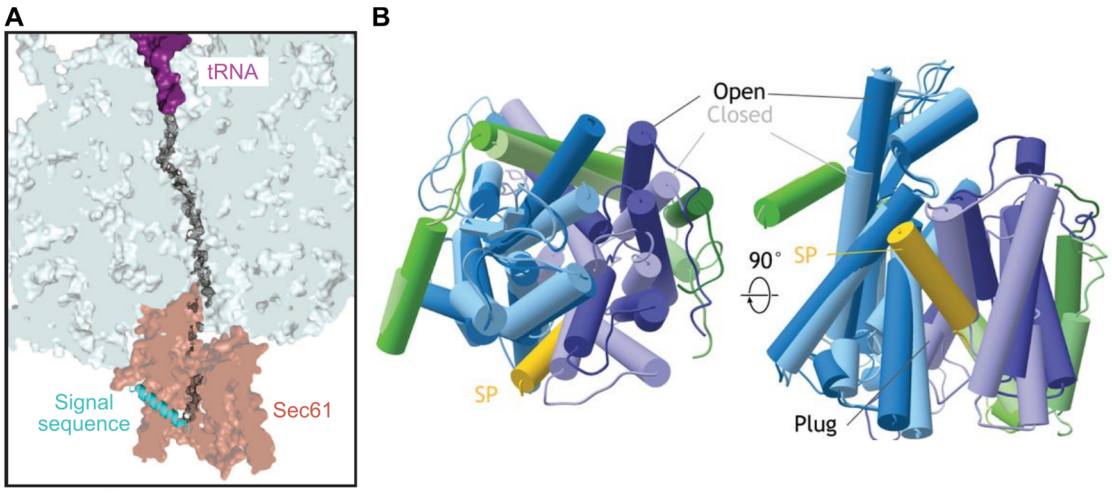


Figure 6. Cryo-EM structure of the signal sequence-engaged ribosome-bound Sec61 translocon.

(A) Cut side view of the cryo-EM structure of the signal sequence-engaged ribosome-bound Sec61. The ribosome (gray), tRNA (dark magenta) and Sec61 (red) are shown in surface representation and the signal sequence (cyan) is shown as ribbon. The density of the nascent chain is shown in mesh. (Figure adapted from Voorhees & Hegde, 2016) (B) Structural comparison of the ribosome-bound Sec61 in its signal sequence-engaged state (dark color) or idle state (light color). The N-terminal half of Sec61 α-subunit is in blue and the C-terminal half in violet. The β and γ subunits are in green and the signal sequence is in yellow. (Figure adapted from Gemmer & Förster, 2020)

The molecular mechanism of the engagement of a polypeptide bearing a N-terminal signal sequence to the SecA-SecY post-translational complex is elucidated first by a crystal structure and later a cryo-EM structure (Li *et al*, 2016; Ma *et al*, 2019). The crystal structure consists of a short secretory peptide sequence fused with SecA and cross-linked to SecY channel with a disulfide bridge. Like the co-translational one, the signal sequence of the secretory peptide binds at the groove outside the lateral gate and the translocating peptide inserts into SecY channel in a loop configuration. The overall structure of SecY is similar to that in the idle SecA-SecY structure except the lateral gate. The lateral gate helices TM3 and TM7 move towards each other while the lateral gate still remains open, providing a pocket for signal sequence recognition.

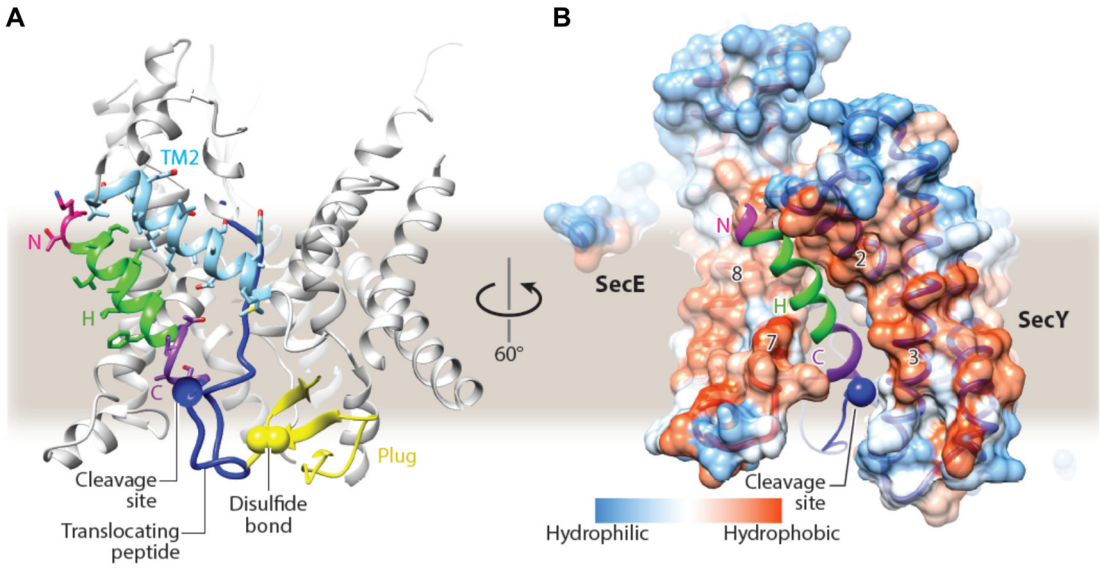


Figure 7. Signal sequence-engaged SecY channel.
(A) Structure of SecY and the signal sequence in the substrate-bound SecA-SecY complex. The translocating peptide (dark blue) is disulfide-crosslinked with the plug domain (yellow) of SecY. The signal sequence (green) binds at the groove at the lateral gate, contacting the TM2 (orange) and TM7 (cyan). The disulfide bond is shown as spheres. For clarity, SecA is omitted. (B) Surface representation of active SecY showing the hydrophobicity of the lateral gate, with hydrophilic and hydrophobic residues in blue and red, respectively. The signal sequence (green and magenta) as well as the translocating peptide (dark blue) are shown in ribbon (Adapted from Rapoport *et al*, 2017)

1.8 Polypeptide translocation

Upon binding of the signal sequence, the plug domain of Sec61α is relocated away from the channel, and the pore ring residues are displaced, resulting in a bigger pore diameter for peptide translocation (Li *et al*, 2016; Voorhees & Hegde, 2016). The hourglass shape of the channel provides minimal contact, except the poring residues, with the translocating peptide, allowing it to pass through smoothly (Cannon *et al*, 2005). In addition, since the signal sequence binds outside the lateral gate and is eventually cleaved off by signal peptidase, it does not hinder the movement of the translocating

peptide. Because a peptide can slide through the Sec61/SecY channel back and forth freely by Brownian motion (Matlack et al, 1999; Bauer et al, 2014), a driving force is required to prevent the translocating peptide from sliding back to the cytosol. In the co-translational pathway, the peptide is passed through the Sec61/SecY channel directly from the exit tunnel of the translating ribosome, which avoids the translocating peptide from moving backward (**Figure 8**A).

In the post-translational translocation pathway in eukaryotes, it is the ER luminal Hsp70 chaperone BiP (Kar2p in yeast) that functions like a molecular ratchet, stopping the backward movement of the translocating peptide (Lyman & Schekman, 1997; Matlack *et al*, 1999). BiP iteratively binds and releases the peptide emerging from the luminal side of the channel, with ATP hydrolysis cycle regulated by the J-domain of Sec63 (**Figure 8**B). On the other hand, the bacterial post-translational translocation relies on the SecA ATPase to accomplish this task in a "push-and-slide" manner (Zimmer *et al*, 2008; Bauer *et al*, 2014; Catipovic *et al*, 2019; Catipovic & Rapoport, 2020). The two-finger-helix (TFH) of SecA inserts into the SecY channel upon ATP binding, providing a power stroke to push the translocating peptide into the channel. ATP hydrolysis allow the THF to reset without pulling the peptide (**Figure 8**C).

Figure 8. Modes of peptide translocation via Sec61/SecY channel. →
(A) In co-translational translocation, the translocating peptide goes into the channel directly from the translating ribosome. This restricts the backward movement of the peptide. (B) In eukaryotic post-translational protein translocation, the ER luminal chaperone BiP (Kar2p in yeast) works as a molecular ratchet by repetitive binding to the translocating peptide emerging at the luminal side to prevent back-sliding. The binding and release of the peptide from BiP require ATP hydrolysis cycle, which is regulated by the J-domain of Sec63. (C) In bacterial post-translational translocation, a "push-and-slide" mechanism performed by the two-helix-finger (THF) of SecA ATPase to push the translocating peptide into the channel. The power stroke of the THF is fueled by ATP hydrolysis. (Adapted from Rapoport *et al*, 2017)

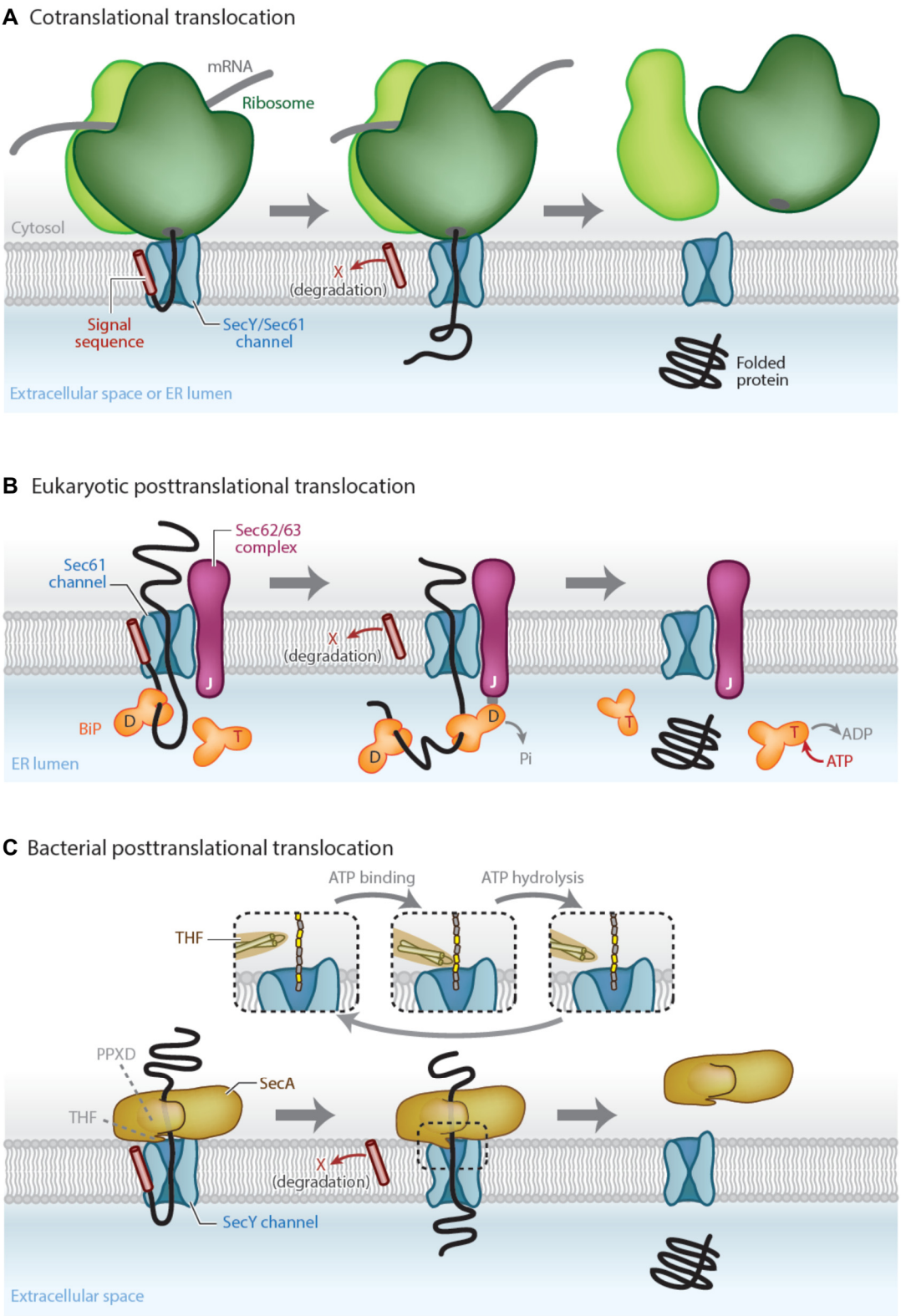


Figure 8. Modes of peptide translocation via Sec61/SecY channel.

1.9 Architecture of the post-translational Sec translocon

1.9.1 Components of the Sec complex

The post-translational Sec complex consists of the heterotrimeric Sec61 complex and the Sec62/63 complex, which is composed of Sec62, Sec63 and additional subunits Sec71 and Sec72 in yeast (Deshaies *et al*, 1991; Panzner *et al*, 1995; Plath *et al*, 1998).

Sec63

Sec63 is an ER-resident membrane protein with three TM helices at its N-terminus and a large cytosolic domain at its C-terminus. On the luminal side of the ER membrane Sec63 harbors between TM2 and TM3 a J-domain that is able to recruit the luminal Hsp70 chaperone BiP (Kar2p in yeast) to the complex. BiP/Kar2p has been shown to act like a molecular ratchet by iterative binding to the translocating peptide, thereby preventing peptides from backsliding and thus providing the driving force for unidirectional translocation (Matlack *et al*, 1999). On the cytoplasmic side, the C-terminal cytosolic domain of Sec63 is in homology with the Brr2-like (Brl) domain, which is found in RNA helicases. In addition, the cytosolic domain of Sec63 harbors an acidic stretch at its ultimate C-terminus that interacts with its partner, Sec62 (Wittke *et al*, 2000). This interaction is shown to be regulated by CK2-dependent phosphorylation at the acidic C-terminus of Sec63 (Wang & Johnsson, 2005).

Sec62

Sec62 is also an essential protein, but its exact function during protein translocation is less well defined. Topologically, Sec62 features two TM helices flanked by two cytosolic domains (Deshaies & Schekman, 1989, 1990). The N-terminal cytosolic domain of Sec62 is the major binding site for the Sec complex by interacting with the acidic C-terminus of Sec63. Deleting this domain impairs translocation activity of the

Sec complex (Wittke *et al*, 2000). The C-terminal cytosolic domain of Sec62 is shown to be a minor binding site for the Sec complex, yet its interaction partner is still unknow (Wittke *et al*, 2000). While acting most likely in concert with Sec63 during translocation, Sec62 was shown by chemical cross-linking to be in direct proximity to inserting signal sequences (Lyman & Schekman, 1997; Matlack *et al*, 1999; Dünnwald *et al*, 1999). Moreover, it was suggested to stabilize binding of the signal sequence to Sec61α and thereby promoting channel gating (Wu *et al*, 2019).

Sec71 and Sec72

In contrast to Sec62 and Sec63, Sec71 and Sec72 are not essential and only exist in fungi. Sec72 is a soluble protein and only Sec71 has a TM helix. But Sec72 is anchored to the ER membrane and integrated into the Sec complex by binding tightly to the C-terminal cytosolic domain of Sec71 to Sec72. Although they are not essential, a structural and biochemical study shows that cytosolic chaperone Ssa1, which protects unfolded peptides and is required in the post-translational pathway, binds to the tetratricopeptide repeat (TPR) domain of Sec72 (Tripathi *et al*, 2017). Together with other studies, Sec71-Sec72 complex is shown to aid in protein translocation by interacting with cytosolic chaperones, thereby facilitating the handover of the secretory peptide from the chaperones to the Sec complex (Feldheim *et al*, 1993; Feldheim & Schekman, 1994; Tripathi *et al*, 2017).

1.9.2 Cryo-EM structure of the Sec complex

The first attempt for solving the structure of the Sec complex was by Harada *et al*, 2011, using cryo-EM. However, because of the limitation of technology at that time, only a featureless 20-Å cryo-EM map could be acquired. The development of direct-electron detector hardware and processing software later led to the “Resolution Revolution” in

the use of cryo-EM to solve protein structures (Kühlbrandt, 2014). By using cryo-EM, it is easier to acquire protein structures at high resolution, especially those that are difficult to crystallize like membrane proteins. In 2019, two groups solved the cryo-EM structure of the Sec complex at near atomic resolution (Wu *et al*, 2019; Itskanov & Park, 2019), showing how this important post-translational translocon is assembled (**Figure 9**).

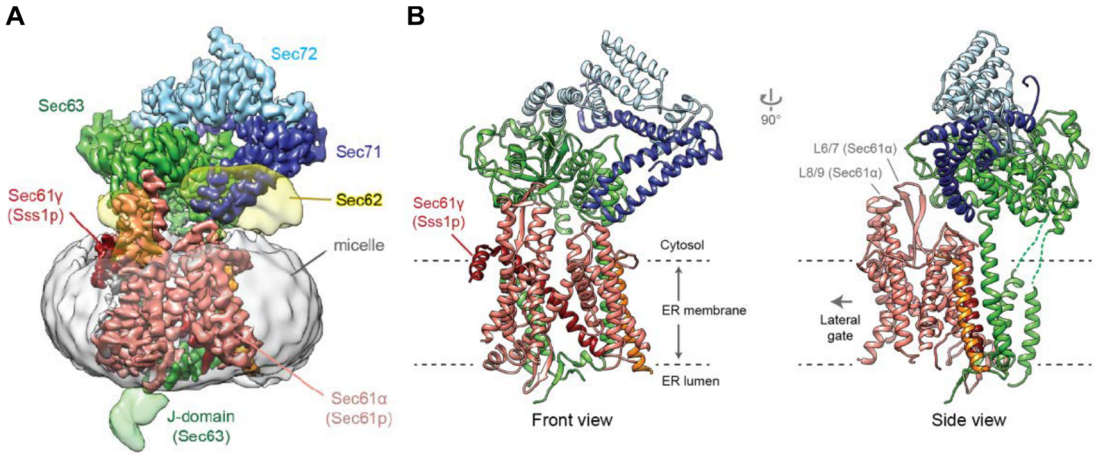


Figure 9. Structure of the heptameric post-translational Sec translocon. (B) The overall structure shows that Sec63, Sec71 and Sec72 sit on the cytosolic side of Sec61. The three TM helices of Sec63 binds to the back side of Sec61. Due to the flexibility and weak density, the model of Sec62 was not built. Sec61 α subunit is labeled in salmon red, β subunit (Sbh1) in orange and γ subunit (Sss1) in red. Sec63 is in green, Sec71 in dark blue and Sec72 in light blue. (Adapted from Itskanov & Park, 2019)

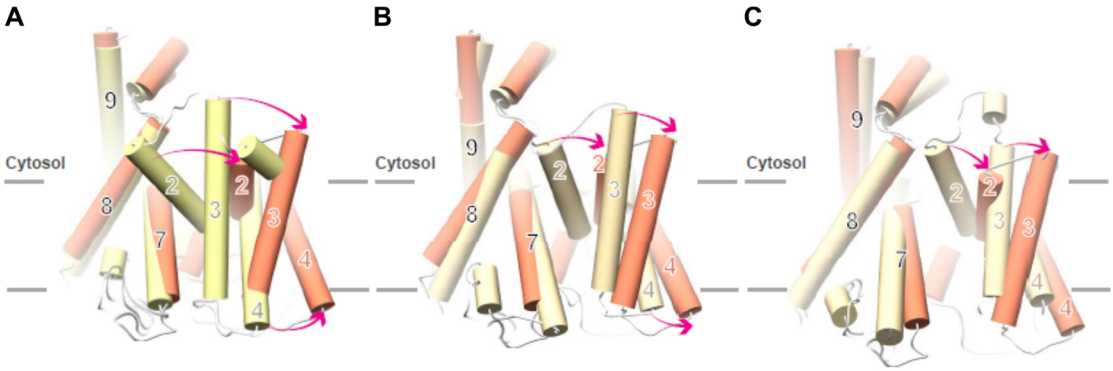


Figure 10. Conformational changes of Sec61α in the Sec complex.
(A) Comparison of Sec61α (salmon red; PDB ID 6ND1; Wu *et al*, 2019) in the Sec complex with idle SecY (yellow) from *M. jannaschii* (PDB ID 1RH5; van den Berg *et al*, 2004), with views from the side (left panel) and from the cytosol (right panel). Red arrows indicate the movement of TMs. (B) As in (A), but in comparison with ribosome-primed Sec61 (PDB ID 3J7Q; Voorhees *et al*, 2014). (C) As in (A), but in comparison with the signal sequence-opened ribosome-bound Sec61 (PDB ID 3JC2; Voorhees & Hegde, 2016). The lateral gate of Sec61α in the Sec complex is more open compared to the above three states. (Adapted from Wu *et al*, 2019)

In these structures, the Sec71-Sec72 heterodimer and the Sec63 C-terminal domain are packed upon the cytosolic face of the Sec61 complex. As a result, the C-terminal cytosolic loops L6/7 and L8/9 of Sec61α are masked in a manner that all ribosome interactions are sterically prohibited, explaining why the Sec complex can exclusively employ the post-translational mode. The lateral gate of the Sec61 channel in the Sec complex is more open than that of the idle, closed SecY crystal structure (**Figure 10**A). Moreover, compared to ribosome-primed (Voorhees et al, 2014; Braunger et al, 2018) or signal sequence-engaged Sec61α (Voorhees & Hegde, 2016), the lateral gate of Sec61α in both structures is stabilized in an even more open conformation (**Figure 10**B and C). This is mainly due to Sec63, which interacts with Sec61α via its three TM helices that bind to both, the N- and C-terminal halves at the back side (opposite from the lateral gate), in addition to the interaction between the Sec63 cytosolic domain and the C-terminal cytosolic loops of Sec61α (Wu et al, 2018; Itskanov & Park, 2018).

These two structures show how the apo Sec complex is assembled, yet without a substrate, the architecture of the Sec complex actively bound to a signal sequence remains to be elucidated.

1.10 Aims

The molecular and structural mechanism of co-translational protein translocation across the ER/cytoplasmic membrane as well as different functional states of the Sec61 translocon in this pathway have been well-studied (van den Berg *et al*, 2004; Frauenfeld *et al*, 2011; Voorhees *et al*, 2014; Park *et al*, 2014; Voorhees & Hegde, 2016; Braunger *et al*, 2018; Kater *et al*, 2019). The structural mechanism of bacterial post-translational translocation by SecA-SecY complex is also well acknowledged (Zimmer *et al*, 2008; Bauer *et al*, 2014; Li *et al*, 2016; Ma *et al*, 2019; Catipovic *et al*, 2019; Catipovic & Rapoport, 2020). By contrast, although functional and biochemical studies have revealed the molecular mechanism of eukaryotic post-translational translocation conducted by the heptameric Sec complex (Matlack *et al*, 1997; Plath *et al*, 1998; Matlack *et al*, 1999; Wittke *et al*, 2000; Plath *et al*, 2003), how the complex is assembled and translocates polypeptides with a less hydrophobic signal sequence into the ER lumen is still unclear due to the lack of structural data. Recently two cryo-EM structures of the idle Sec complex have been resolved, shedding light on the structural assembly of the Sec complex (Wu *et al*, 2019; Itskanov & Park, 2019). Despite some differences between the two structures regarding the plug conformation, the highly similar open conformation of the lateral gate was interpreted to facilitate insertion of and gating by the less hydrophobic signal sequences employed in the post-translational mode. Yet, it is not clear how exactly these signal sequences engage the heptameric complex and how the structurally unknown Sec62 subunit may contribute.

The aims of this thesis were to elucidate i) how a signal sequence engages the Sec complex, and ii) how Sec62 assembles into the complex and contributes to signal sequence recognition. In order to do this, a signal sequence-bound Sec complex should be generated for further structural study by using cryo-EM. As a substrate, prepro-α-factor (ppαF), a well-known cargo protein for the post-translational Sec complex, was chosen. PpαF is a precursor to the α-factor mating pheromone with a cleavable signal sequence. The Sec complex should be purified *ex vivo* from *S. cerevisiae* and reconstituted with ppαF.

For a cryo-EM structure of such a reconstituted signal sequence-bound Sec complex, structural details were expected giving valuable insights into molecular details of signal sequence engagement. In addition, this structure may reveal more information on the placement of Sec62 in this assembly and consequently on its functional role for the translocation process.

Chapter 2: Materials and methods

2.1 Molecular cloning

2.1.1 Plasmids and strains

For purification of the heptameric Sec complex, sequences coding for a 3C cleavage site followed by eight histidines and a triple FLAG sequence (3C-His_8-3×FLAG) tag were inserted downstream of the *SEC62* gene of wild type *Saccharomyces cerevisiae* (*S.c.*) strain W303. For the ppαF-mEGFP construct, the sequence coding for the first 54 amino acids of ppαF followed by mEGFP was cloned into a modified pET28a vector that adds an N-terminal SUMO tag to the translated insert. The N-terminal cytosolic domain of Sec62 from *S.c.* (Sec62 domain; residues 18-145) and a R51E mutant of this domain were cloned by Dr. Jingdong Cheng into the same modified pET28a or into a modified pGEX-6P-1 plasmid. All strains, plasmids and primers used in this study are listed in **Table 1**, **Table 2** and **Table 3**.

Table 1. Plasmids for protein over-expression and purification.

No.	Protein	Vector	Marker	Ori	Description
1	–	pFA.3C.H8.F3	Amp	SP6	3C-His_8-3×FLAG tag integration at C-terminus of genomic *SEC62*
2	ppαF-mEGFP	pET28a	Kan	f1	His_6-FLAG-SUMO-ppαF-mEGFP
3	ppαFm3-mEGFP	pET28a	Kan	f1	His_6-FLAG-SUMO-ppαFm3-mEGFP

Table 2. Yeast and *E. coli* strains.

Strain	Description
S. cerevisiae W303	*MAT*a/*MAT*α {*leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*} [phi^+]
E. coli BL21 (DE3)	Used for protein over-expression and purification.

Table 3. Primers used in PCR.

No.	Primer	Sequence (5'→3')	Description
1	Sec62-pFA-F	GAAACAAGCCAAGAGAGAAAGCAATAAG AAGAAAGCCATCAATGAAAAAGCCGAAC AAAACCTGGAAGTTCTGTTCCAGGG	For generating DNA fragments from pFA.3C.H8.F3 for C-terminal tagging at yeast genomic *SEC62*
2	Sec62-pFA-R	CGAGTTCATGAGGTTACAATATAGAAGG TTTATACAGTAGAGCTATACAGGATAAT GGAAGTATCGATGAATTCGAGCTCG	
3	eGFP-F	ATGGTGAGCAAGGGCGAG	Amplification of EGFP DNA fragments for In-Fusion cloning into pET15b with ppαF
4	eGFP-R_BamHI	AATCGGATCCCTTGTACAGCTCGTCCAT GC	
5	ppαF-F_NdeI	GGGAATTCCATATGAGATTTCCTTCAAT TTTTACTGC	Amplification of ppαF (residues 1-54) DNA fragments for In-Fusion cloning into pET15b with EGFP
6	ppαF54-eGFP-R	CGCCCTTGCTCACCATTGGCAAAACAGC AACATCG	
7	IF_pET15b_BamHI_eGFP-F	GCTGTACAAGGGATCCCACCACCACCAC	For generating linear pET15b for In-Fusion cloning with ppαF and EGFP
8	IF_pET15b_NdeI_ppαF-R	GAAGGAAATCTCATATGTATATCTCCTT CTTAAAGTTAAAC	
9	IF_ppSUMO-F	GAGCTCCGTCGACAAGC	For generating linear pET28a for In-Fusion cloning
10	IF_ppSUMO-R	ACCACCAATCTGTTCTCTGTG	
11	IF_ppαF-ppSUMO-F	GAACAGATTGGTGGTATGAGATTTCCTT CAATTTTTACTGC	For generating ppαF-EGFP fragments for In-Fusion cloning with pET28a
12	IF_eGFP-ppSUMO-R	TTGTCGACGGAGCTCTCACTTGTACAGC TCGTCCATGC	
13	qc_eGFP_A206K_F2	GCTCAGTTTGGACTGGGTGCTCAGGTAG	For a single mutation on EGFP to generate mEGFP
14	qc_eGFP_A206K_R2	CAGTCCAAACTGAGCAAAGACCCCAACG	
15	qc_ppαF-m3_F	AGTTTTATTCGAAGCATCCTCCGCATTA GCTGC	For generating ppαFm3 mutant from wildtype ppαF
16	qc_ppαF-m3_R	GCTTCGAATAAAACTGCAGTAAAAATTG AAGGAAATCTC	

2.1.2 Polymerase chain reaction

Polymerase chain reaction (PCR) was used to generate plasmids containing ppαF-mEGFP fusion constructs and perform site-directed mutagenesis. PCR reactions were set up by using KOD Xtreme hot start DNA polymerase (Merck) or Phusion high-fidelity PCR master mix (Thermo). The reactions were carried out using BioRad 1000 Touch thermocycler. The program used in this study is in **Table 4**.

Table 4. PCR program used for plasmid and target gene amplification as well as mutagenesis.

Step	Temp	Time
Initial denaturation	98°C	10 min
30 cycles	98°C	5 sec
	60°C	10 sec
	72°C	30 sec per kb
Final extension	72°C	2 min
Hold	12°C	

2.1.3 Agarose gel electrophoresis

PCR products were subjected to agarose gel electrophoresis, which can separate DNA based on their sizes. Gels were prepared with 1% (w/v) agarose dissolved in TAE buffer (40 mM Tris pH 8.8, 20 mM acetic acid, 2mM EDTA) by boiling. SYBRSafe (Invitrogen) for DNA visualization was added before the gel solution was cooled. DNA samples mixed 5:1 with 6 × loading dye (NEB) were loaded on the gel with 1 kb or 100 bp DNA ladder (NEB) in one lane as a standard for size comparison. Gels were run at 120 V for 30 min and then visualized by an Inatas GelDoc.

2.1.4 Degradation of parental plasmid

Plasmids isolated from *E. coli* are methylated, thus the restriction enzyme DpnI, which specifically cleaves methylated DNA, can be used to digest parental DNA after PCR

amplification. To degrade parental DNA, 1 µl of DpnI (NEB) was directly added to 50 µl of PCR product and the reaction was incubated at 37°C for 1 h or overnight. The DpnI treated PCR product was purified using the QIAquick PCR purification kit (QIAGEN) according to manufacturer's protocol.

2.1.5 In-Fusion cloning

Once the PCR amplified vectors and inserts were ready, the In-Fusion HD cloning kit (Takara) was used to fuse those DNA fragments according to the manufacturer's protocol. The kit can join two DNA fragments with 15-base pair (bp) homology at their ends. The proprietary In-Fusion enzyme within the kit digests 15 bases from the 3'-end to produce single-stranded ends and promote their pairing. The paired DNA fragments are readily converted into circular DNA after transformed into competent cells.

2.1.6 Plasmid transformation

The products from the In-Fusion reactions were transformed into homemade *E. coli* competent cells DH5α. The competent cells were thawed on ice and 5 µl of the In-Fusion reaction were added into 100 µl of competent cells, incubated on ice for 30 min. The cells were heat-shocked at 42°C for 45 sec and chilled on ice for 3 min. The cells were then recovered by incubation in a shaker at 37°C for 30 min with addition of 900 µl LB medium. The cells were then plated on LB-Agar plate containing kanamycin as a selection marker for positive clones.

2.1.7 Plasmid isolation

Positive colonies were picked and inoculated into 5 ml LB medium containing kanamycin, incubated in a shaker at 37°C overnight. The cells from the overnight culture were pelleted and QIAprep Spin Miniprep Kit (QIAGEN) was used to isolate

plasmids according to the manufacturer's instructions. Isolated plasmids were eluted in 30 µl ddH2O and the concentrations were measured using NanoPhotometer NP80 (Implen). The plasmids were sent to Eurofins Genomics or Sigma for sequencing to confirm coding region sequences.

2.1.8 Site-directed mutagenesis

To generate mEGFP from EGFP or ppαFm3 mutant from wildtype ppαF, primers carrying the mutations were used in whole plasmid PCR with the same protocol in **Table 4**. The PCR products were treated with DpnI at 37°C overnight and transformed into DH5α competent cells and plated on LB-Agar plates containing kanamycin as described before. Positive clones were selected and the plasmids were isolated, sent for sequencing to confirm the mutation.

2.1.9 Cloning of ppαF-mEGFP

DNA fragments of the N-terminal 54 residues of ppαF and EGFP were PCR amplified from plasmids kindly provided by Dr. Birgitta Beatrix with primers 3-6 (**Table 3**). The fragments were originally cloned into plasmid pET15b simultaneously by using In-Fusion cloning kit according to its multiple-insert protocol. The ppαF-EGFP construct was then PCR amplified with primers 11 and 12 (**Table 3**) and cloned into the modified pET28a plasmid containing an N-terminal His_6-FLAG-SUMO tag. In order to avoid EGFP dimerization at high concentration (Phillips, 1997), a single mutation A206K was introduced into EGFP the to generate monomeric EGFP (mEGFP; Zacharias *et al*, 2002), resulting in the final construct ppαF-mEGFP used for binding assay and *in vitro* reconstitution with the Sec complex.

2.1.10 Genomic tag insertion in *S. cerevisiae*

To genomically tag Sec62 at its C-terminus for native pull-out, a DNA sequence containing a 3C-cleavage site, a His_8-3×FLAG tag and nourseothricin-resistant gene (NrsR; nourseothricin acetyltransferase) was PCR amplified from plasmid pFA.3C.H8.F3 (see Appendix Plasmid 1) with primers 1 and 2 (**Table 3**) using the same protocol described in previous session (**Table 4**). Wildtype yeast cells were inoculated and incubated at 30°C overnight. The cells were diluted into 250 ml YPD to 0.2 OD_{600} and then were grown until an OD_{600} of 0.8. After reaching the optimal density, the cells were harvested by centrifugation at 1,160 × *g* for 5 min and split in two tubes. Cell pellets were washed by water once and pelleted as before. The cells were resuspended in 1 ml of 100 mM LiOAc and transferred into Eppendorf tubes. They were then pelleted at 15,700 × *g* for 15 seconds and the LiOAc supernatant was removed. The cells were resuspended into a final volume of 250 µl for each tube by adding about 200 µl of 100 mM LiOAc. In the meantime, single strand salmon-sperm DNA (ssDNA, 2 mg/ml) was prepared by boiling the sample for 5 minutes and immediately transferring it on ice. The cells were pooled and for each reaction, 100 µl of cells were pelleted by 1,500 × *g* for 1 min to remove LiOAc. The transformation mix was added in the following order: 240 µl 50% (w/v) polyethylene glycol (PEG) 3,350, 36 µl 1 M LiOAc, 50 µl ssDNA, and 34 µl PCR product. Each tube was vortexed vigorously until the pellet was dispersed. The cells were incubated at 30°C for 30 minutes and then heat shocked at 42°C for 25 minutes. The cells were pelleted at 2,500 × *g* for 1 min. The supernatant was removed and 300 µl TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA) were added. The cells were gently resuspended and were plated onto YPD agar plates supplemented with 100 µg/ml nourseothricin for selection. After incubation at 30°C for two to three days, positive colonies were validated by PCR and western blot to confirm the tagging was successful.

2.2 Protein analysis

2.2.1 SDS-polyacrylamide gel electrophoresis

Protein samples from purification or pull-down assay were subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE) for separation according to their molecular weight. Fifteen µl of each protein sample were mixed with 5 µl 4× SDS sample buffer (SB; 200 mM Tris-HCl pH 6.8, 8% (w/v) SDS, 40% glycerol, 0.4% (w/v) bromophenol blue and 400 mM 1,4-dithiothreitol) and incubated at 95°C for 5 min. Ten µl of the heat-denatured samples were loaded onto 15% SDS-polyacrylamide (SDS-PAA) gels. PageRuler Prestained/Unstained Protein Ladders (Thermo) were used as protein size markers. The electrophoresis was conducted at 200 V for 50 min in SDS running buffer (25 mM Tris, 192 mM glycine and 0.1% (w/v) SDS). The SDS-PAA gels were then stained with SimplyBlue Coomassie stain (Invitrogen) to visualize protein bands.

2.2.2 Protein concentration measurement

Protein concentration was measured using NanoPhotometer NP80 (Implen) based on the absorbance of the proteins at 280 nm wavelength (A_{280}). The measurement was first blanked by the protein buffer. Around 1-2 µl of samples were applied on the instrument and the concentration is calculated according to A_{280} and the extinction coefficients of the proteins. The extinction coefficients and the molecular weights of the proteins in this thesis were predicted by ProtParam (Gasteiger *et al*, 2005) and listed in **Table 5**.

Table 5. Protein properties.
The molecular weights and extinction coefficients of the proteins in this study.

Protein	Molecular weight (Da)	Extinction coefficient
Sec61α	52,947.9	54,320
Sec61β (Sbh1)	8,719.3	1,490
Sec61γ (Sss1)	8,948.8	4,470
Sec62	31,366.0	38,850
Sec63	75,321.0	75,290
Sec71	24,231.2	26,930
Sec72	21,604.4	13,980
Sec complex	223,138.6	215,330
ppαF-mEGFP	32,580.9	23,380
ppαFm3-mEGFP	32,696.0	23,380

2.3 Protein expression and purification

2.3.1 Expression and purification of ppαF-mEGFP

The wildtype or mutant ppαF-mEGFP constructs were overexpressed in the *E. coli* strain BL21 (DE3). The cells were first inoculated into 50 ml LB medium containing 50 μg/ml kanamycin (LB+Kan) and incubated overnight at 37°C. This small culture was then diluted into a 2 L LB+Kan medium. After 3 h incubation at 37°C, the culture was induced with 0.1 mM IPTG and incubated at 18°C overnight. The cells were harvested by centrifugation at 3,400 × *g* for 10 minutes. The cell pellet was resuspended in ppαF lysis buffer (20 mM HEPES pH7.4, 150 mM NaCl, 1 mM dithiothreitol (DTT) and 5 mM imidazole) and lysed using a microfluidizer (Microfluidics). The lysate was centrifuged in a Ti45 rotor (Beckman Coulter) at 125,440 × *g* at 4°C for one hour, and the supernatant was then subjected to an open column of 4 ml Ni-NTA agarose (QIAGEN). After three washes with 20 ml ppαF-lysis buffer, the column was washed again with cleavage buffer (20 mM HEPES pH 7.4, 150 mM NaCl, 5 mM DTT). On-column cleavage of the SUMO tag by home-made Ulp1 protease was performed in 4

ml cleavage buffer at 4°C overnight. The eluted proteins were further purified using size-exclusion chromatography with a Superdex 200 column (GE Healthcare) in buffer containing 20 mM HEPES pH 7.4, 150 mM NaCl and 1 mM DTT. The fractions containing ppαF-mEGFP was collected, aliquoted and flash-frozen in liquid nitrogen until use.

2.3.2 Purification of the post-translational Sec translocon

The yeast strain expressing C-terminally tagged endogenous Sec62 protein was grown in 5 ml YPD at 30°C overnight and diluted to an OD_{600} of 0.3 in 50 ml YPD. The cells were grown for around 6 h until ~5 OD_{600}. This culture was then again diluted into a large 20-L culture and incubated again at 30°C until reaching an optical density of OD_{600}~5 (in about 20 hours). The cells were pelleted by centrifugation and resuspended in Sec lysis buffer (20 mM HEPES pH 7.4, 100 mM KOAc, 2 mM $Mg(OAc)_2$, 1 mM DTT, 0.5 mM PMSF) supplemented with EDTA-free protease inhibitor cocktail (Roche) and lysed using a microfluidizer. The lysate was centrifuged in an SLA-1500 rotor (Thermo) at 29,800 × *g* at 4°C for 20 min. The supernatant was then centrifuged in a Ti45 rotor (Beckman Coulter) at 185,500 × *g* at 4°C for 1 h. The absorbance A_{280} of the pelleted microsomes was measured by applying the microsomes diluted 1:1000 in 1% SDS on the NanoPhotometer. The microsomes were flash-frozen in liquid nitrogen and stored at -80°C until use.

For purification of the apo heptameric Sec complex, the frozen microsomes (about 12,000 A_{280} in total) were thawed and resuspended with a Dounce homogenizer in solubilization buffer (20 mM HEPES pH 7.4, 0.75 M KOAc, 2.5 mM $Mg(OAc)_2$, 0.4 M sucrose, 0.5 mM EDTA, 1 mM DTT, 0.5 mM PMSF, 3% glyco-diosgenin (GDN, Anatrace) and protease inhibitor cocktail. After incubation with stirring at 4°C for 1 h,

the solubilized microsomes were centrifuged in a Ti45 rotor at 126,000 × *g* at 4°C for one hour. The supernatant was incubated with 300 µl anti-FLAG M2 agarose beads (Sigma) at 4°C for 1 h. The beads were washed three times in 1.5 ml wash buffer (20 mM HEPES pH 7.4, 100 mM KOAc, 2.5 mM $Mg(OAc)_2$, 10% glycerol, 1 mM DTT, 0.02% GDN) and the complex was eluted in 300 µl wash buffer containing 0.5 mg/ml home-made 3C protease at 20°C for 1 h. The eluted sample was diluted two times with Q buffer (20 mM HEPES pH 7.4, 2.5 mM $Mg(OAc)_2$, 1 mM DTT, 0.02% GDN), subjected to an open column of 400 µl Q Sepharose Fast Flow (GE Healthcare) pre-washed with 2 ml Q buffer. After washing the Q column with 2 ml Q buffer, the Sec complex was eluted in 1.2 ml Q buffer with additional 1 M KOAc. The buffer of the eluted complex was exchanged to Sec buffer (20 mM HEPES pH 7.4, 100 mM KOAc, 2.5 mM $Mg(OAc)_2$, 1 mM DTT, 0.02% GDN) and concentrated to ~5 mg/ml using a 100-kDa-cutoff Amicon membrane (GE Healthcare).

2.4 Sec-ppαF pull-down assay

For binding assays with the heptameric Sec complex and ppαF-mEGFP or its translocation defect mutant (ppαFm3-mEGFP), purified Sec complex (0.8 µM) and purified ppαF-mEGFP (0.8, 1.6 or 4 µM) were incubated in 10 µl Sec buffer at 30°C for 20 min. The protein samples were then immobilized on GFP-Trap magnetic agarose beads (Chromotek) at 4°C for 30 min. The GFP-Trap magnetic agarose beads for each reaction were taken from 6 µl slurry, pre-washed with 60 µl Sec buffer two times, and finally in 5 µl Sec buffer until use. The flow-through was collected and mixed with 5 µl 4× sample buffer. The beads were washed three times with 100 µl Sec buffer and mixed with 20 µl 1× sample buffer. All samples were incubated at 95°C for 5 min and subjected to SDS-PAGE analysis followed by Coomassie Brilliant Blue staining.

2.5 Sec62-Sec63 pull-down assay and isothermal titration calorimetry

This was kindly provided by Dr. Jingdong Cheng. In brief, Sec63-C1 and Sec63-C2P peptides with a biotin tag (synthesized by GL Biochem, Shanghai) were incubated with the purified Sec62 domain (residues 18-145) or the R51E mutant Sec62 domain in binding buffer (20 mM HEPES pH 7.4, 100 mM NaCl, 5% glycerol, 0.05% Triton X-100, and 1 mM DTT) for 2 h at 4°C. The protein samples were then immobilized on 25 µl of Streptavidin resin (GE Healthcare) for 20 min at 4°C. The resin was washed three times with binding buffer, and bound proteins were eluted using sample buffer. The final sample was subjected to SDS-PAGE and stained with Coomassie Brilliant Blue.

2.6 On-bead reconstitution of ppαF-mEGFP-bound Sec complex

For on-bead reconstitution of ppαF-bound heptameric Sec complex, the same purification protocol for the apo Sec complex was applied except that, after washing the anti-FLAG M2 beads, the ppαF-mEGFP was added to the complex in a 10:1 molar ratio estimated from previous purification. The sample was incubated at 30°C for 20 minutes followed by three washing steps with Sec buffer. The rest of the purification remained the same as described for the apo heptameric complex. The final purified reconstituted ppαF-Sec complex was concentrated to ~5 mg/ml.

2.7 Cryo-EM analysis and molecular model

2.7.1 Cryo-EM sample preparation and data collection

The reconstituted ppαF-Sec complex (3.5 µl at 5 mg/ml) was applied on glow-discharged Quantifoil R2/2 UltrAuFoil grids, blotted for 2 s at 4°C and 100% humidity, and immediately plunge-frozen in liquid ethane using a Vitrobot Mark IV (FEI). Two

sets of cryo-EM data were acquired on a Titan Krios electron microscope (FEI) using a K2 detector (Gatan) and GIF energy filter. For dataset 1, a total of 6243 dose-fractionated movies were collected with 40 frames, an exposure of 0.9 e^{-}/frame/Å^2, and a magnification resulting in an image pixel size of 1.059 Å per pixel. For dataset 2, a total of 8802 dose-fractionated movies were collected with 40 frames, an exposure of 1.15 e^{-}/frame/Å^2, and a magnification resulting in an image pixel size of 1.059 Å per pixel.

2.7.2 Single particle cryo-EM data analysis

The original movies were first subjected to motion correction and dose weighting using MotionCor2 (Zheng *et al*, 2017) and the CTF parameters were estimated using CTFFIND4 (Rohou & Grigorieff, 2015). The does-weighted micrograph sums were visually inspected to remove bad micrographs. A total of 5112 micrographs of dataset 1 and 5118 micrographs of dataset 2 were selected for further processing in RELION-3 (Zivanov *et al*, 2018) as shown in **Figure 16**. For dataset 1, after auto-picking and several rounds of 2D classifications, 453,116 particles were selected for 3D refinement and then classified into six classes. One class (C6) showed clear secondary structure of the Sec complex. The other classes were subjected to another round of refinement and classification with a mask around Sec61 complex/Sec63/Sec71/Sec72. The class with clear secondary structure density in transmembrane region was merged with C6 and refined with the same mask. This refined map was used as a template for picking particles in dataset 2 using RELION. A total of 117,117 particles were selected after several rounds of 2D classifications. These particles were then merged with the particles from dataset 1 and the same 3D classification process were performed again to obtain classes with clear secondary structure density in transmembrane region. The resulted particles were refined with the same Sec61 complex/Sec63/Sec71/Sec72 mask and

classification was done with a mask around the transmembrane domain of the Sec61 complex without alignment. One class showing extra density of signal sequence and Sec62 TMs was refined with a mask around Sec61 complex/Sec63/Sec71/Sec72 plus signal sequence and Sec62 TMs, yielding a 4.5 Å resolution map after post-processing. The other class showing the heptameric Sec complex in the apo state was refined with the Sec61 complex/Sec63/Sec71/Sec72 mask and post-processed, yielding a map with 4.4 Å resolution. Two rounds of random-phase 3D classification (Gong *et al*, 2016) were performed in cryoSPARC (Punjani *et al*, 2017) to further remove bad particles and particles at the edge of micrographs were also removed. The resulting particle stacks were subjected to 3D refinement in RELION, yielding a 4.4 Å resolution map for the signal sequence-engaged state and 4.3 Å for the apo state. Neither CTF refinement nor Bayesian polishing could further improve the resolution. Local resolution filtered maps were calculated using RELION.

2.7.3 Model building

All models were built in COOT (Emsley et al, 2010). For the signal sequence-bound Sec complex, the structure of yeast heptameric Sec complex (PDB ID 6N3Q) was used to rigid body fit into the map. For the apo Sec complex, the structure of another yeast Sec complex (PDB ID 6ND1) was used to rigid body fit into the map. Figures of models and maps were generated using UCSF ChimeraX (Goddard et al, 2018) and PyMOL (Schrödinger).

2.8 Crystal structure of Sec62-N

This was kindly provided by Dr. Jingdong Cheng. In brief, crystals of the cytosolic N-terminal domain of Sec62 (residues 18-145) were grown at 4°C using the hanging drop vapor diffusion method by mixing equal volumes of the purified protein complex

(20mg/ml) and crystallization buffer (100 mM MES pH 5.7-5.9, 2 M (NH4)2SO4). Crystals were transiently transferred into a cryoprotectant buffer containing reservoir buffer and additional 20% glycerol (v/v) before they were flash frozen in a cold nitrogen stream at -173°C. All data were collected in 0.97958Å wavelength at ESRF (The European Synchrotron Radiation Facility, France). The data were processed using the program XDS package (Kabsch, 2010). Phases were initially determined by the single-wavelength anomalous dispersion (SAD) using the phasing module Autosol; density modification and automatic model building were performed using the AutoBuild of program package PHENIX. The final model was manually built using Coot. All refinements were performed using the refinement module phenix.refine of the PHENIX package(Adams et al, 2010). The model quality was validated using the MolProbity of the PHENIX package, which indicated good stereochemistry according to the Ramachandran plot for the structure (favored: 98.4%, outliers: 0.0%).

Chapter 3: Results

3.1 Construction and purification of substrate ppαF-mEGFP

The N-terminal signal sequence of yeast prepro-α-factor (ppαF), a precursor to the α-factor mating pheromone, is a well-characterized substrate for the Sec complex (Panzner *et al*, 1995; Matlack *et al*, 1997; Plath *et al*, 1998, 2003). Yeast *S. cerevisiae* in its haploid mode exists in either α or **a** mating type. The α cells produce and secrete α-factor to signal their presence to the **a** cells, which respond by growing a mating projection toward the source of the mating signal, and vice versa (for review see Haber, 2012 and Merlini *et al*, 2013). PpαF contains a 19-residue N-terminal signal sequence that directs it into the ER, followed by a propeptide including four tandem repeats of α-factor separated by spacers (**Figure 11**A; Kurjan & Herskowitz, 1982). The propeptide is subsequently cleaved off from the signal sequence and glycosylated in the ER, and proteolytically processed in the Golgi apparatus and finally in secretory vesicles to become mature α-factor (Julius *et al*, 1984; Waters *et al*, 1988). As a secretory precursor, ppαF enters the ER by taking the post-translational pathway (Rothblatt *et al*, 1989; Deshaies & Schekman, 1989). Purified heptameric Sec complex was shown to be capable of translocating ppαF into reconstituted proteoliposomes (Panzner *et al*, 1995; Matlack *et al*, 1997). In addition, cross-linking data showed that the signal sequence of ppαF interacts with the Sec complex through Sec61α lateral gate helices (Plath *et al*, 1998, 2003).

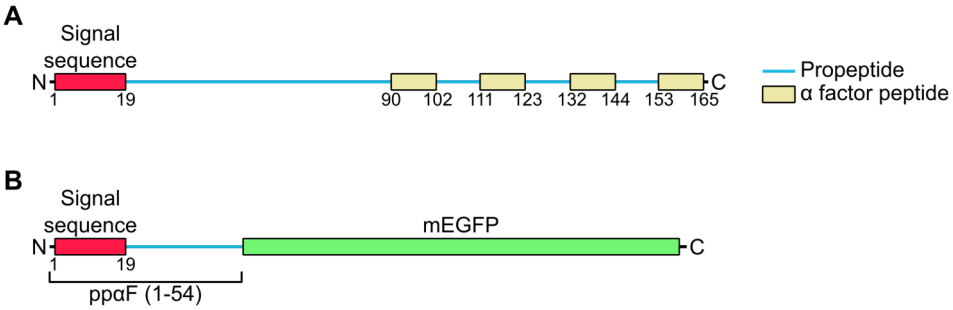


Figure 11. Prepro-α-factor

(A) Schematic diagram of prepro-α-factor. The signal sequence (red) is cleaved while the polypeptide is translocated through the channel. The translocated protein is further processed in the ER and the propeptide (light blue) is removed in the Golgi apparatus, resulting in the mating α-factor. (B) Schematic diagram of the construct of the ppαF-mEGFP fusion protein.

In order to reconstitute a signal sequence-bound heptameric Sec complex *in vitro*, a recombinant protein containing the N-terminal signal sequence of ppαF was constructed. To prevent complete translocation of ppαF and trap the signal sequence in the translocon, the construct contains the N-terminal 54 residues of ppαF fused with mEGFP at the C-terminus (**Figure 11**B). To validate the affinity of the recombinant substrate to the purified Sec complex, the translocation defect mutant ppαFm3 (Allison & Young, 1989) was constructed by introducing a single A13E mutation into the wildtype signal sequence using site-directed mutagenesis. These two constructs were cloned into a modified pET28a vector that carries an N-terminal His_6-FLAG-SUMO tag. The wildtype and mutant ppαF-mEGFP recombinant proteins were over-expressed and purified from *E. coli* cells (**Figure 12**A). The recombinant proteins were pulled down by Ni-NTA agarose and eluted by adding SUMO protease Ulp1 to cleave ppαF-mEGFP off from the resins. The eluates were subjected to gel filtration, resulting in highly purified recombinant substrates (**Figure 12**B and C).

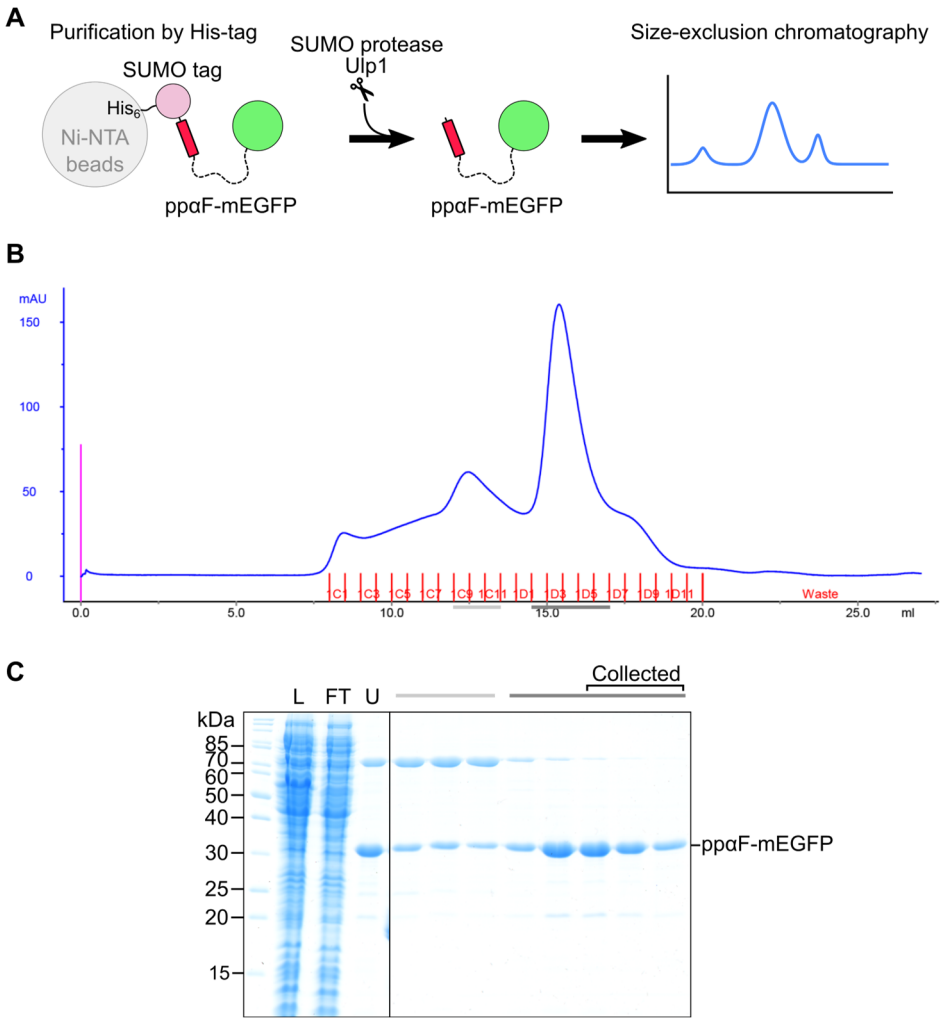


Figure 12. Purification of the ppαF-mEGFP fusion protein.

(A) Schematic depiction of the purification of ppαF-mEGFP. (B) Size-exclusion chromatography (SEC) of the SUMO-protease- eluted ppαF-mEGFP. The profile showed the A_{280} (blue line) of the eluate along the eluted volume. Collected fractions were marked in red. (C) SDS-PAGE analysis of the purified ppαF-mEGFP. (L: clear lystate, 1/6,000 of which was loaded on the gel; FT: flow-through from Ni-NTA, 1/6,000; U: eluate after Ulp1 cleavage, 1/800; light and dark gray bars: SEC fractions from the two peaks highlighted by the respective colored bars, 1/67). Fractions pooled for further assays and reconstitutions were marked with “collected”.

3.2 Purification of the post-translational Sec translocon

Membrane proteins need to be in their native lipid environment to remain stable and active. Due to their hydrophobic characteristic in the transmembrane segments, it is difficult to research membrane proteins in solution compared to soluble, cytosolic proteins. In order to study membrane proteins *in vitro*, however, they are extracted and solubilized from their native lipid environment by detergents, which could possibly cause instability and decrease in activity (Tate, 2010). Although methods have been developed to address these problems by reconstituting membrane proteins into liposomes or nanodiscs (for review see Seddon *et al*, 2004 and Denisov & Sligar, 2016), the detergent-solubilized and purified heptameric Sec complex from *S. cerevisiae* has been shown to remain the activity for substrate-binding and protein translocation in the absence of lipid bilayer (Matlack *et al*, 1997). Thus, for this study detergent-solubilized Sec complex was produced. To enable affinity-purification of the Sec complex, a DNA sequence coding for a His_8-3×FLAG-tag and a 3C-protease cleavage site was introduced downstream of endogenous *SEC62* in *S. cerevisiae* W303. Twenty liters of Sec62-3C-His_8-3×Flag expressing yeast culture were harvested at about 5 OD_{600} and approximately 15 ml of rough microsomes with 1,600 A_{280}/ml were separated from the cell lysates. A total of 12,000 A_{280} of microsomes was solubilized for purification of the heptameric Sec complex via the C-terminally FLAG-tagged endogenous Sec62 (**Figure 13**A). The complex was first pulled down using anti-FLAG M2 agarose and eluted by 3C protease treatment to cleave off the FLAG-tag. The eluate was further purified by anion exchange chromatography and concentrated to around 5 mg/ml. SDS-PAGE analysis showed an intact heptameric Sec complex with all seven components present (**Figure 13**B).

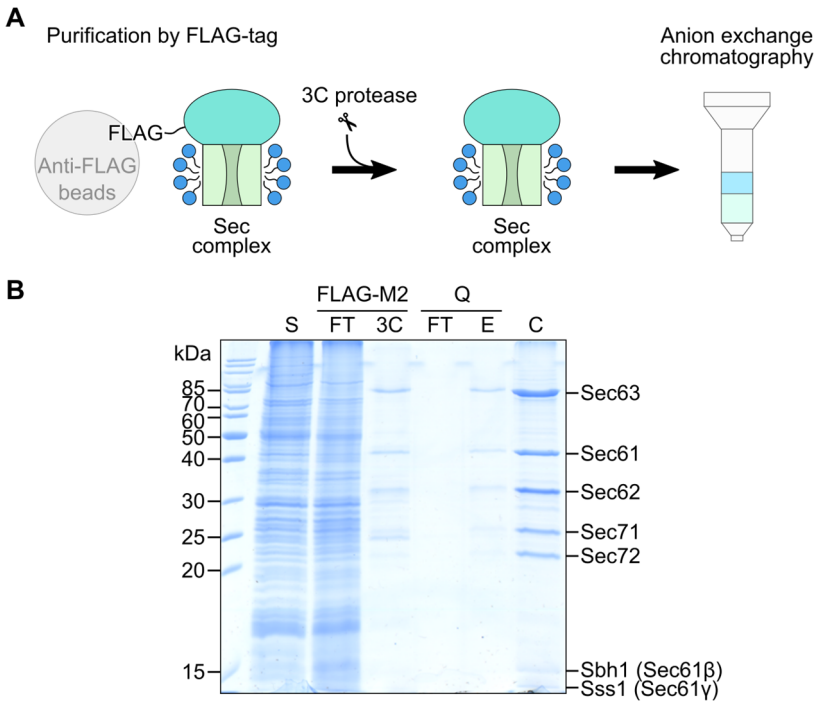


Figure 13. Affinity-purification of the heptameric Sec complex.
(A) Schematic depiction of the purification of the heptameric Sec complex containing a C-terminal FLAG-tagged Sec62. (B) SDS-PAGE analysis of the purification of the Sec complex (S: solubilized microsomes, 1/80,000 of which was loaded on the gel; FLAG-M2-FT: flow-through from the anti-FLAG resin, 1/80,000; FLAG-M2-3C: eluate after 3C cleavage, 1/160; Q-FT: flow-through from Q-sepharose, 1/1,250; Q-E: eluate from Q-sepharose, 1/125; C: concentrated heptameric Sec complex sample, 1/16).

3.3 Reconstitution of the ppαF-mEGFP-bound Sec translocon

In order to test binding of the substrate ppαF-mEGFP to the heptameric Sec complex, a pull-down assay was performed (**Figure 14**A). In this assay, the purified Sec complex was incubated with wildtype or the translocation defect mutant ppαF-mEGFP in different molar ratios. The complex was then immunoprecipitated on GFP-Trap magnetic beads, which were coupled with anti-GFP nanobody that could also bind mEGFP. The precipitated and soluble fractions were subjected to SDS-PAGE to verify the affinity between the Sec complex and the substrates.

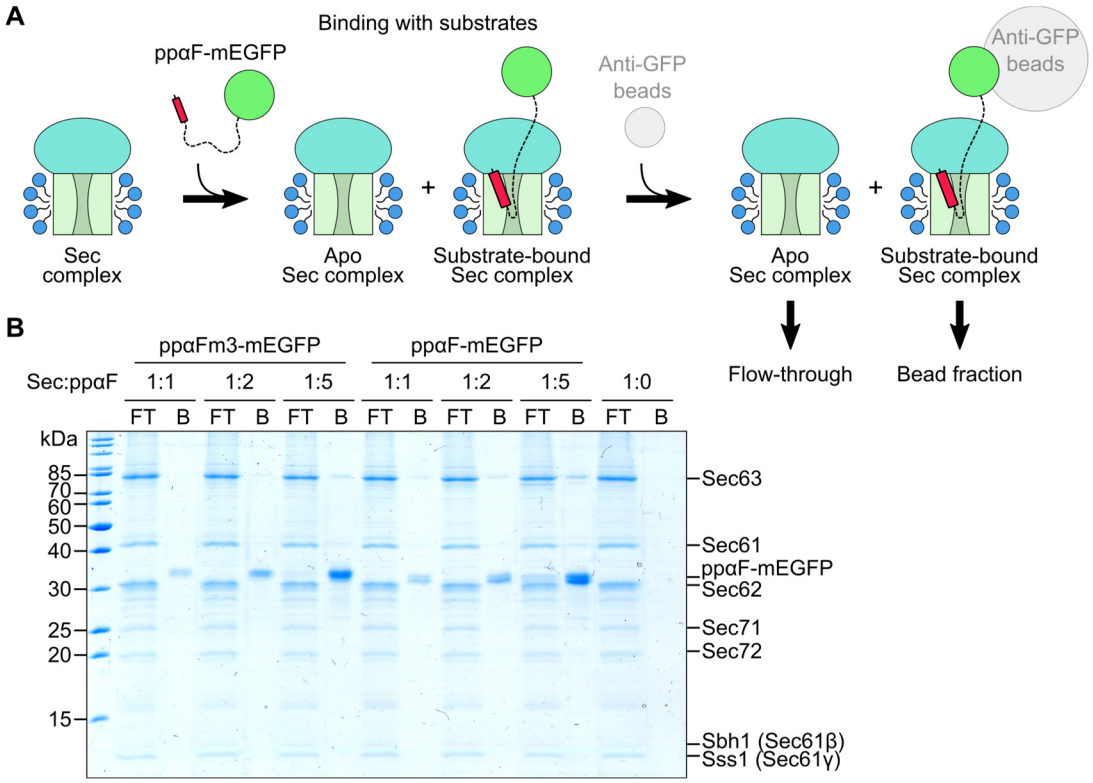


Figure 14. Binding assay assessing formation of Sec-ppαF-mEGFP complexes.

(A) Schematic depiction of the pull-down assay of the Sec complex with the signal sequence-containing mEGFP-tagged ppαF (ppαF-mEGFP) or its mutant on anti-GFP beads. (B) SDS-PAGE analysis of the pull-down assay of the purified Sec complex with ppαF-mEGFP or its mutant m3. Pull-downs were performed at varying molar ratios (B: bead-bound fraction; FT, unbound flow-through).

SDS-PAGE analysis of the pull-down assay indicated that under given conditions a higher fraction of wildtype ppαF-mEGFP was bound to the Sec complex compared to the signal-sequence mutant ppαFm3-mEGFP. This indicates that the observed binding of ppαF-mEGFP to the Sec complex is indeed an effect specific to the signal sequence (**Figure 14**B). This confirmed that the recombinant substrate ppαF-mEGFP is competent for binding to the Sec complex and is suitable for *in vitro* reconstitution. The pull-down assay also showed that only a portion of the Sec complex was pulled down even with ppαF-mEGFP in five-time molar excess of the Sec complex. Thus, a higher

ppαF-mEGFP-to-Sec complex ratio was required in order to increase the fraction of substrate-bound Sec complex. In the following preparation of reconstituted Sec-ppαF-mEGFP complex for single particle cryo-EM analysis, a 1:10 molar ratio of Sec:ppαF-mEGFP was used.

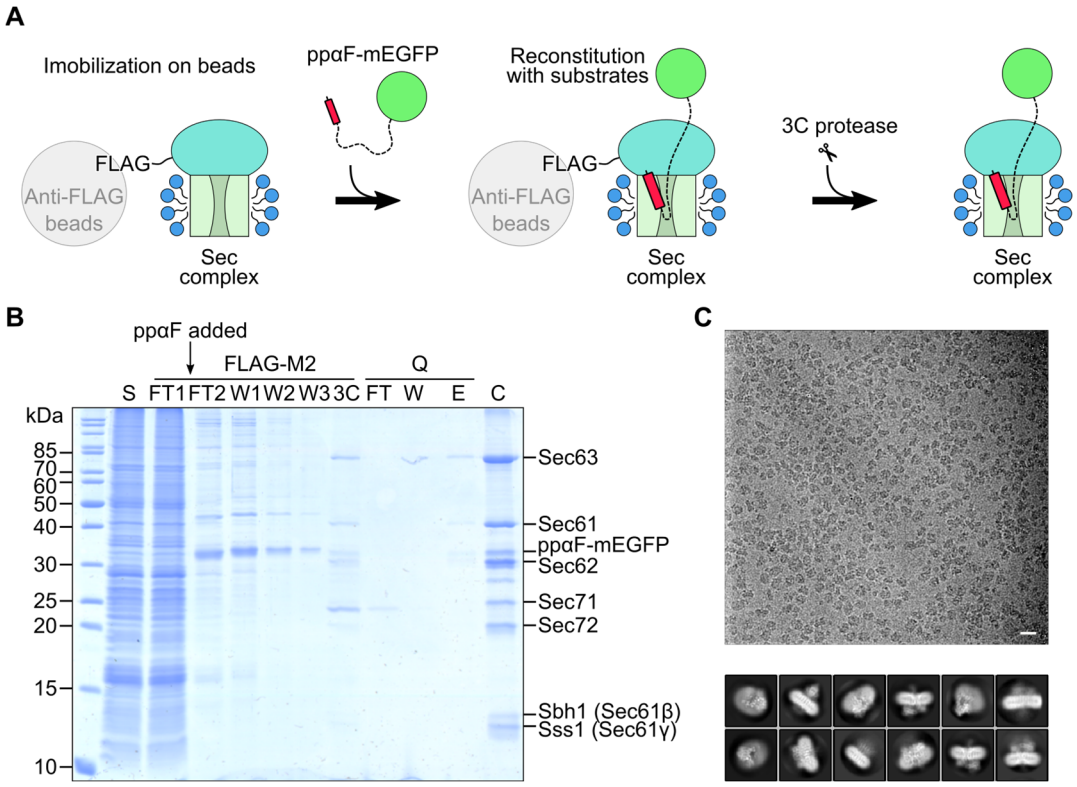


Figure 15. Reconstitution of the signal sequence-bound Sec complex.
(A) Schematic depiction of the purification strategy for the signal sequence-bound Sec complex. (B) SDS-PAGE analysis of the reconstitution of the signal sequence-bound Sec complex (S: solubilized microsomes, 1/80,000 of which was loaded on the gel; FT1: flow-through after binding of the Sec complex on beads, 1/80,000; FT2: flow-through after binding of ppαF-mEGFP to the Sec complex immobilized on beads, 1/100; W1-3: washes after ppαF-mEGFP binding, 1/100; 3C: eluate after 3C cleavage, 1/160; Q-FT: flow-through from Q-sepharose, 1/250; Q-W: wash from Q-sepharose, 1/250; Q-E: eluate from Q-sepharose, 1/125; C: concentrated heptameric Sec complex sample). (C) Left panel: a representative cryo-EM micrograph as obtained from Titan Krios with a K2 direct electron detector. The scale bar represents 20 nm. Right panel: selected 2-dimentional (2D) class averages. The box size was chosen to be 200 × 200 pixels resulting in a box width of 212 Å.

After validating the interaction between the Sec complex and the ppαF-mEGFP construct, the substrate-bound Sec complex was reconstituted and purified by an on-bead method with a tweak on the purification process for the apo Sec complex (**Figure 15**A): The Sec complex was first immobilized on anti-FLAG M2 agarose beads and then incubated with ppαF-mEGFP in 10-time molar excess, which was estimated from above-mentioned binding assays (**Figure 14**); Subsequently, the reconstituted substrate-bound Sec complex was cleaved off from the beads using 3C protease and subjected to anion exchange chromatography for further purification and concentration. Again, SDS-PAGE analysis showed the presence of all seven subunits as well as ppαF-mEGFP (**Figure 15**B). This sample was subsequently subjected to cryo-EM and single particle analysis.

3.4 Cryo-EM analysis of the signal sequence-engaged Sec translocon

Two cryo-EM datasets (in total 10,230 micrographs) of the reconstituted signal sequence-bound Sec complex were collected separately and were merged while processing (**Figure 16**; also see session 2.7 and 2.7.2 in **Materials and Methods**). After several rounds of classification, promising 3-dimentional (3D) classes with clear density for the heptameric Sec complex were further classified using a mask for the TM region of Sec61. This yielded two exclusive 3D classes differing in overall conformation and the presence of an additional distinct density in the open lateral gate, which was assigned to the ppαF signal sequence. In agreement with the roughly estimated stoichiometry between ppαF-mEGFP and the Sec complex on the SDS-PAGE gel (**Figure 15**B), the complexes in the signal sequence-engaged state represented about half of the particles, while the other half represented empty Sec complexes, from here on called the apo state (**Figure 16**A).

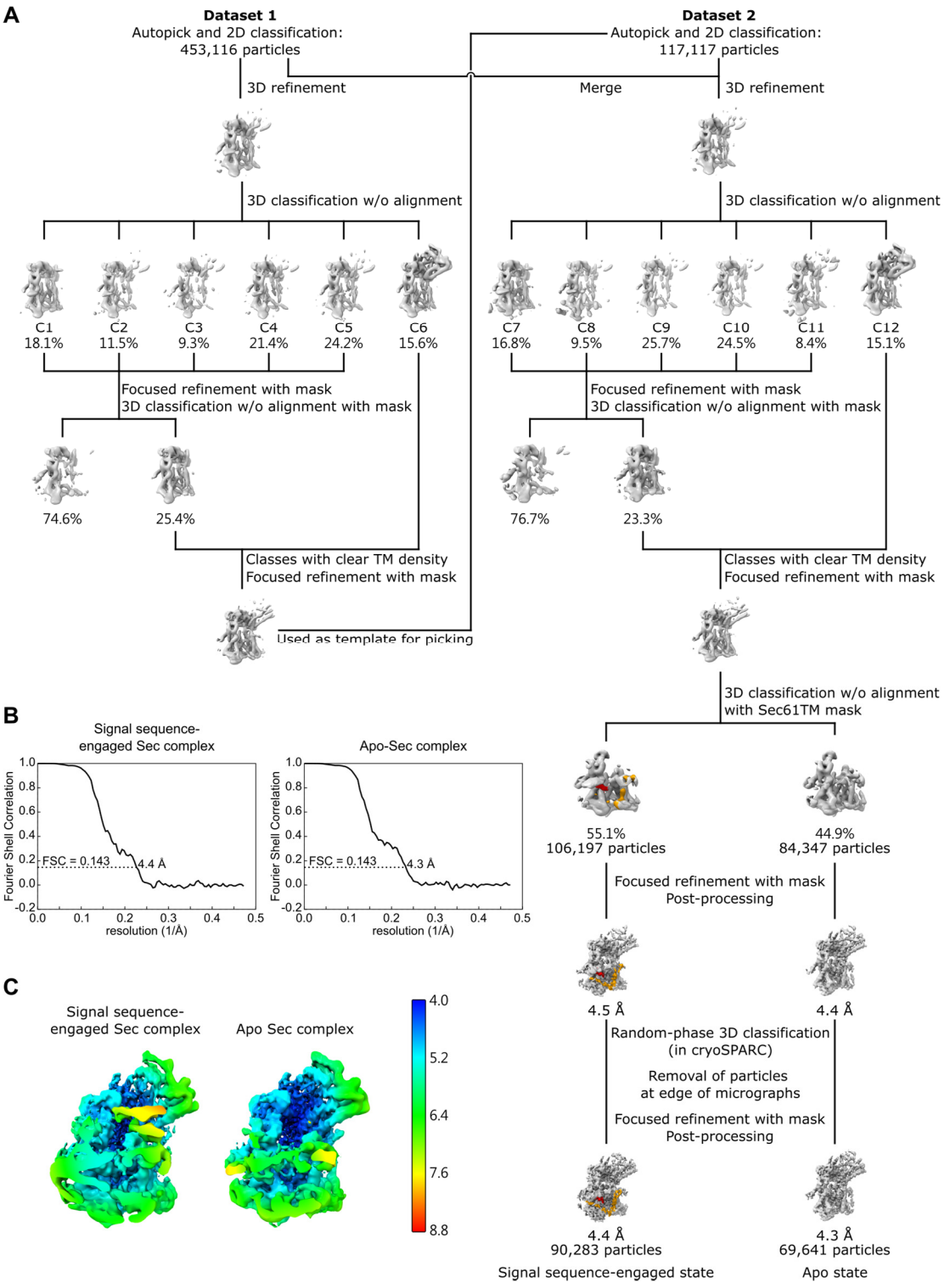


Figure 16. Single particle cryo-EM analysis of the signal sequence-engaged Sec complex.

← **Figure 16. Single particle cryo-EM analysis of the signal sequence-engaged Sec complex.**
(A) 3D classification scheme. A total of 453,116 particles in dataset 1 and 117,117 particles in dataset 2 was selected after 2D classification and refined. The following 3D classification into six classes showed one class (C12) with clear density for the TM region. The other five classes were joined and further 3D classified with a soft binary mask focusing on the TM region. All classes with clear TM density were joined and refined using the mask around the TM region plus the soluble domain of Sec61/63/71/72. A third 3D classification was performed only focusing on the Sec61 complex. This revealed two classes with extra density for the Sec62 TM helices and one class lacking this extra density. The latter was again focused refined resulting in the 4.7 Å resolution structure of the apo heptameric complex. The other classes were joined and subjected to two more refinement rounds, one without applying a mask and one with a mask excluding the Sec62 soluble domain. This resulted in the final 4.8 Å reconstruction of the signal-sequence bound Sec complex. (B) Gold standard Fourier Shell Correlation (FSC) resolution curves of the final 3D reconstructions of the signal sequence-engaged and apo Sec complex. (C) 3D reconstructions of the signal sequence-engaged and apo Sec complexes low-pass filtered and color-coded according to local resolution.

The maps of the apo and the signal sequence-engaged Sec complexes were refined to an overall resolution of 4.3 and 4.4 Å, respectively (**Figure 16**). All TM helices of the Sec complex were clearly resolved, which allowed for the unambiguous rigid body fitting of available structures into both maps with only minor adjustments (see Materials and Methods and **Figure 17**). The overall architecture of the apo Sec complex, which lacks additional density, was very similar to the two previous structures (**Figure 18**A and B; Wu *et al*, 2019; Itskanov & Park, 2019).

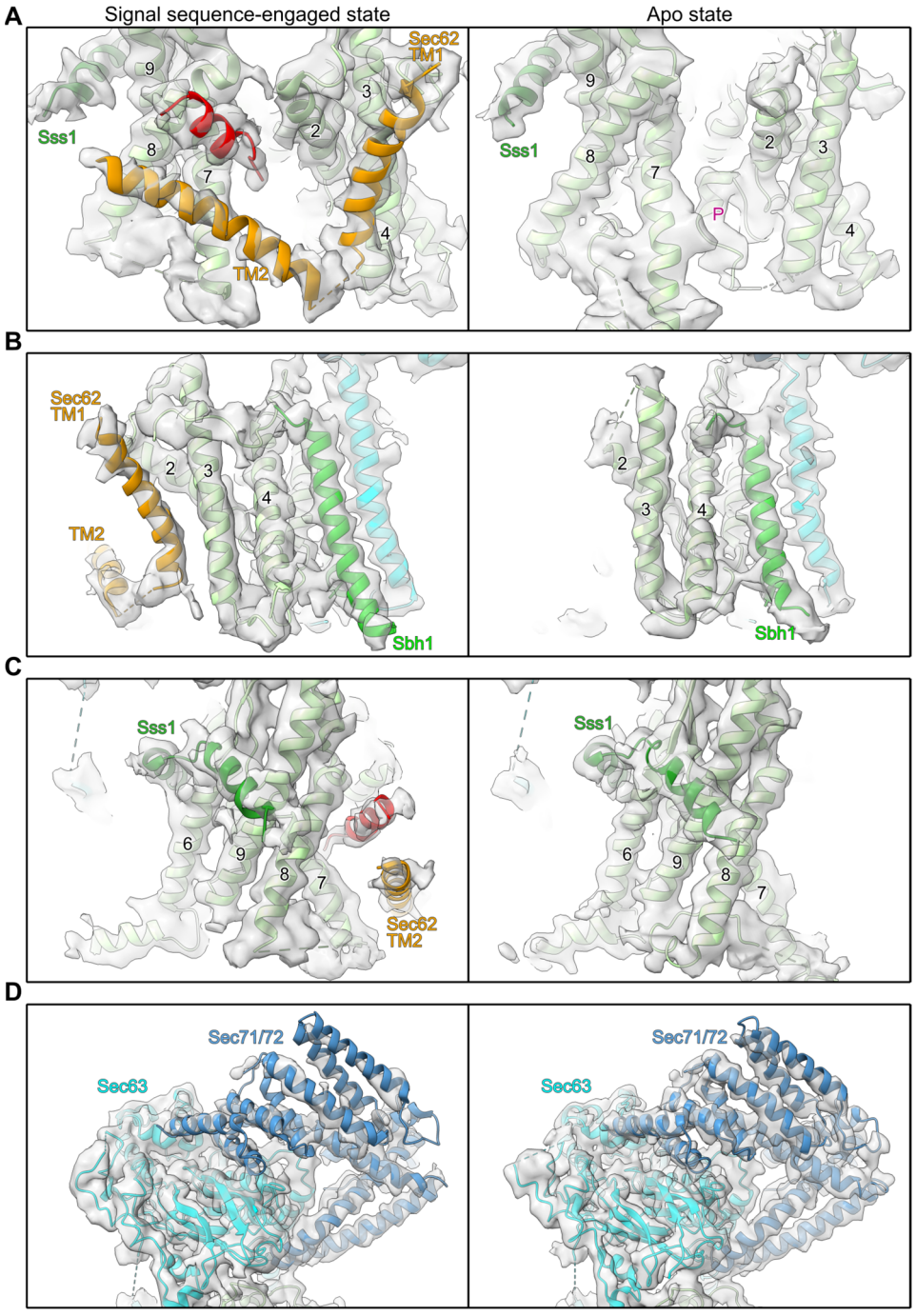


Figure 17. Fitting of molecular models into the cryo-EM maps of ss-bound and apo Sec complex.

(A) View focusing on the lateral gate of Sec61α. In the engaged state (left) the signal sequence (red) is inserted into the lateral gate (TM2 and TM7) and the two Sec62 TMs are stabilized. In the apo state (right) density for the plug (P) is present below the central pore. (B) View focusing on the N-terminal half of Sec61α (TMs 2-4). In the ss-bound state (left) Sec62 TM1 is bound near Sec61α TM2. (C) View focusing on the C-terminal half of Sec61α (TMs 6-9). (D) View focusing on the cytosolic domains of the Sec complex.

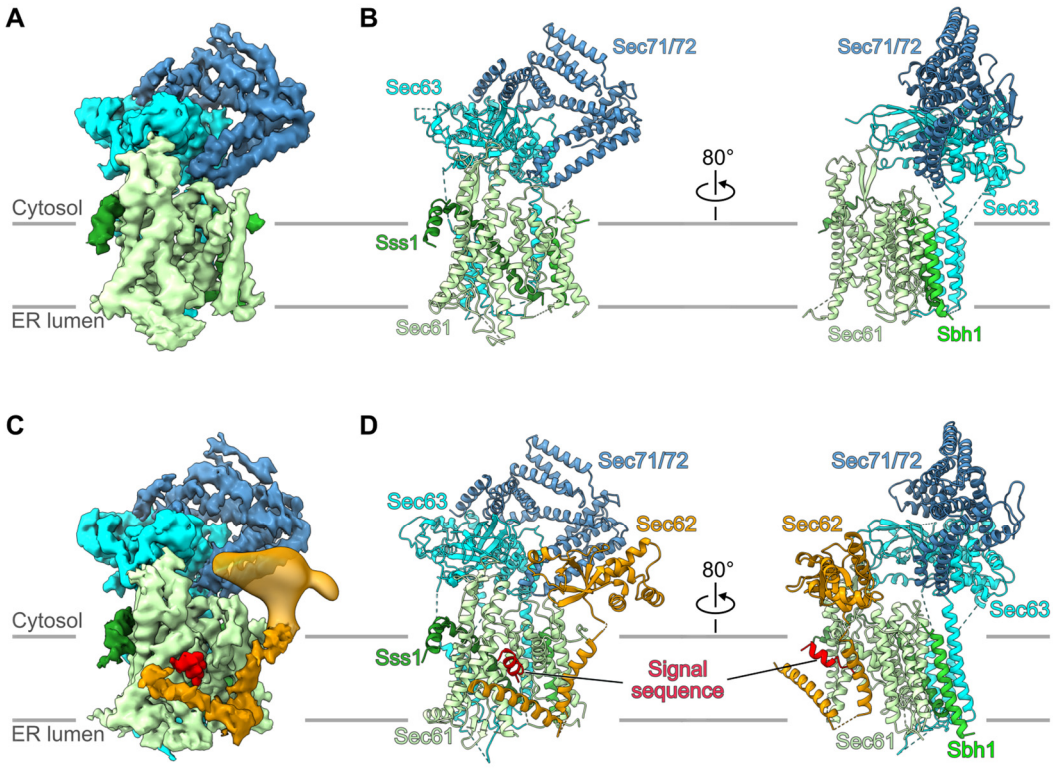


Figure 18. Cryo-EM structures of the apo and signal sequence-engaged Sec complex.

(A) Cryo-EM map of the apo Sec complex shown after focused refinement and post-processing. The map was segmented and color-coded by the individual Sec complex components. For clarity density for the detergent micelle was masked. (B) Molecular model of the apo Sec complex. (C) Cryo-EM map of the signal sequence-bound Sec complex shown after focused refinement and post-processing. For Sec62 N-terminal cytosolic domain the map low-pass filtered to 15 Å is shown. (D) Molecular model of the signal sequence-bound Sec complex. The model for the Sec62 N-terminal cytosolic domain (Sec62 domain) is derived from the crystal structure determined by Dr. Jingdong Cheng.

In general, both apo and signal sequence-engaged state showed similar conformation (**Figure 18**). The C-terminal cytosolic domains of Sec63 and Sec71/72 dimer are located on top of Sec61α, interacting mainly through cytosolic loops L6/7 and L8/9 of Sec61α, which are also the binding site for ribosome in co-translational mode. The position of Sec63 cytosolic domain prevents the Sec61 channel from ribosome

engagement, thus restricting the Sec complex to post-translational mode. The three TM helices of Sec63 transverse the membrane at the backside of the Sec61 channel. Together, the interactions of Sec63 with the cytosolic loops and the backside of Sec61α make the lateral gate more open, as observed in previous structures (**Figure 19**; also see session 3.5; Wu *et al*, 2019; Itskanov & Park, 2019). Also, as seen in these previous cryo-EM maps, less well resolved density is present for the N-terminal cytosolic domain of Sec62 (Sec62-N) (**Figure 18**C and **Figure 20**A). Although the density for Sec62 cytosolic domain is more rigid in the signal sequence-engaged state, a molecular interpretation was not possible due to limited local resolution in the cryo-EM map. In order to interpret this density, a crystal structure of the Sec62-N domain (residues 18-145) was solved during the course of this thesis by a colleague, Dr. Jingdong Cheng. The crystal structure shows an elongated shape and could be fitted into the density, conforming the initial assignment (see session 3.6).

In the apo structure, the lateral gate is open and the plug is clearly present below the pore ring. Thus, the apo state Sec complex map is closest to the structure described by the Rapoport lab (**Figure 19**; PDB ID 6ND1; Wu *et al*, 2019). The overall structure of the engaged state was similar to the apo state, yet, a more pronounced density for the cytosolic domain of Sec62 and extra densities for the ppαF signal sequence in the lateral gate of Sec61α as well as the two TM helices of Sec62 were observed (**Figure 18**C and D, and **Figure 17**A and B).

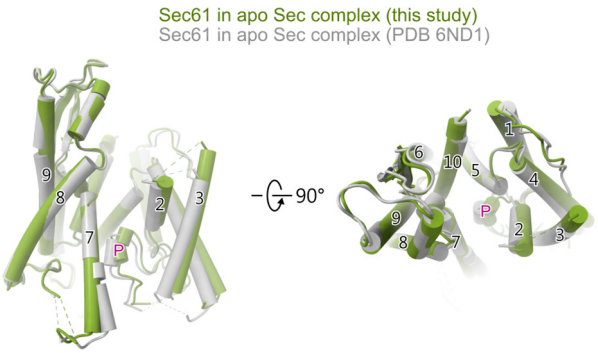


Figure 19. Comparison of Sec61 conformation of the apo Sec complex. Comparison of Sec61α in the apo state (this study; dark green) with Sec61α in the apo Sec complex as shown in the Wu study (PDB ID 6ND1; gray).

3.5 Conformation of ppαF signal sequence, Sec61 and Sec62 TMs

When compared with the reconstruction of the apo state, in the engaged state, extra density was observed within the lateral gate in close contact with TM7 and parallel to TM2 of Sec61α (**Figure 18**C). Due to its unique appearance and position in the lateral gate this density was assigned as the α-helical signal sequence of ppαF (**Figure 18**D and **Figure 20**). Its position is very similar to those signal sequences observed in the structures of co-translationally acting Sec61 or the bacterial SecYEG complex (see session 1.7; Voorhees & Hegde, 2016; Li *et al*, 2016; Ma *et al*, 2019). The position the signal sequence is also fully consistent with previous cross-linking data suggesting that the ppαF signal sequence localizes in close proximity to TMs 2 and 7 of the Sec complex (Plath *et al*, 1998, 2003). Moreover, two additional α-helical densities were observed close to the lateral gate extending towards the Sec62 density at the cytosolic side (Wu *et al*, 2019; Itskanov & Park, 2019) (**Figure 18**C and **Figure 20**, orange). It was noticed that these densities were more pronounced in the engaged state and, since all other TMs of the complex had already been identified, they were assigned to the remaining TM1 and TM2 of Sec62.

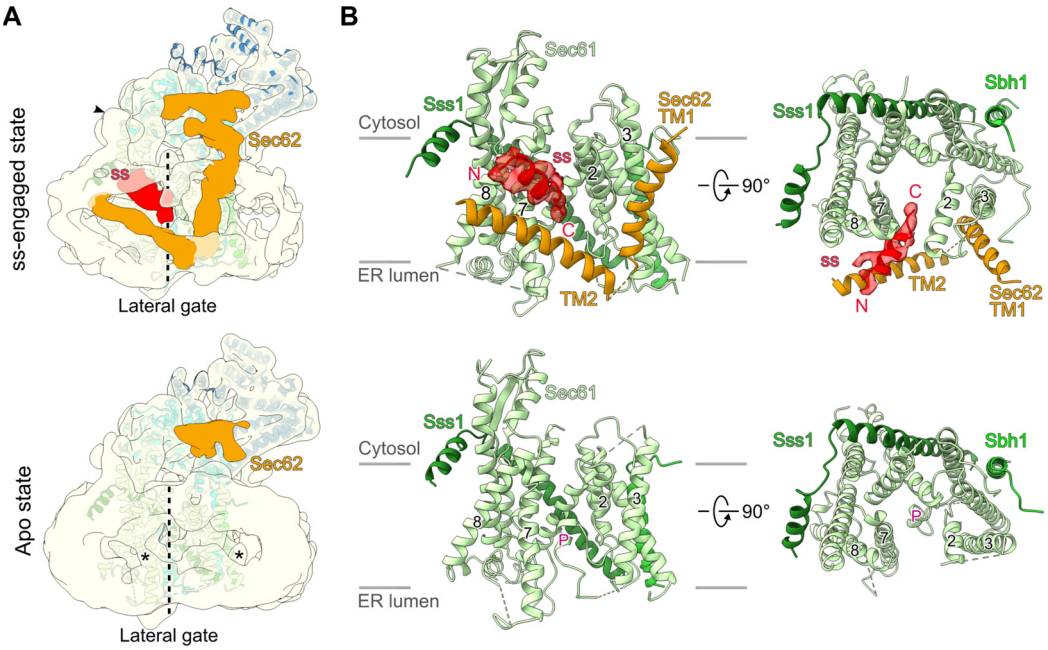


Figure 20. Conformation of the signal sequence-bound Sec complex.
(A) Cut side view of the cryo-EM density of the signal sequence-bound (up) and the apo (bottom) Sec complex highlighting the ppαF signal sequence (ss; red) and Sec62 cytosolic and transmembrane domains (orange). Both maps were low-pass filtered to 8 Å. The density of the Sec62 cytosolic C-terminal domain connected to the Sec62 TM2 is marked by an arrowhead. (B) Two views showing the conformations of the Sec61 complex (green) with the signal sequence (ss; red) bound to the lateral gate (upper panels) and in the apo state. The two Sec62 TMs (orange) are stabilized in ss-bound state. The plug is indicated by "P" in the apo state.

For the conformation of the channel entity itself, the Sec61 complex, a conformational shift of TMs 2 and 3 away from TMs 7 and 8 was observed when compared to the apo state (**Figure 21**A and B) and the ribosome-primed Sec61 (**Figure 21**C; PDB ID 6FTJ; Braunger *et al*, 2018), resulting in an even more open lateral gate. This conformation was most similar to the already more open structure described by the Park lab (**Figure 21**B; PDB ID 6N3Q; Itskanov & Park, 2019), but with an even wider lateral gate (~5.5° relative to TM5). Furthermore, the plug density was not present anymore in its pore closing position and was likely delocalized by the signal sequence guided inserting peptide.

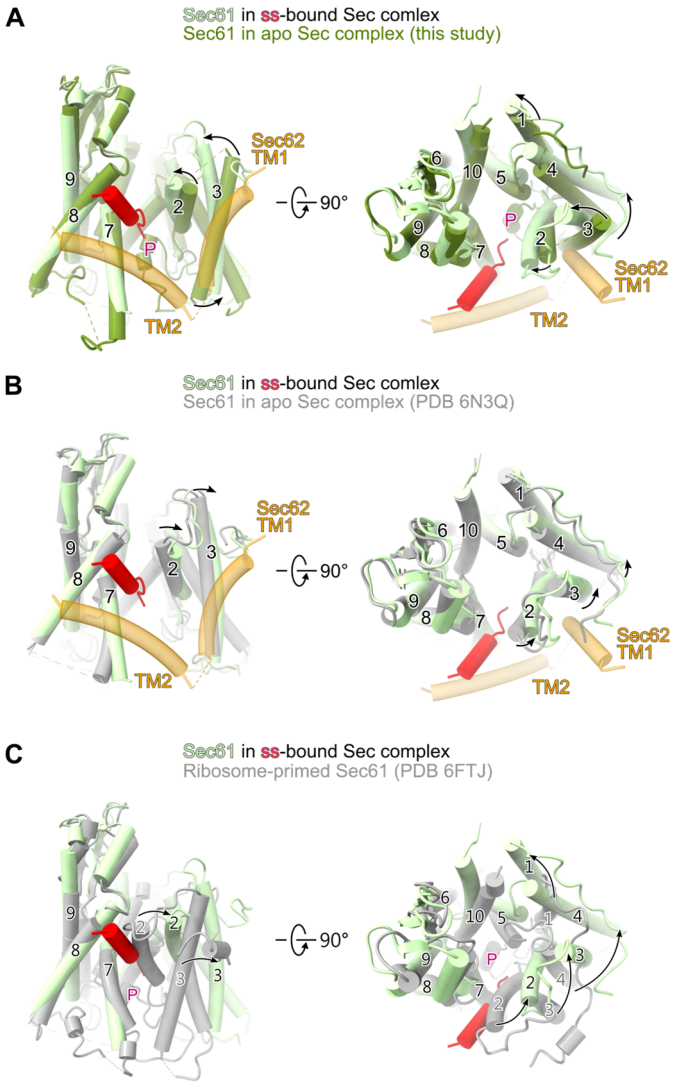


Figure 21. Comparison of Sec61 conformations in the Sec complex structures. Comparison of Sec61α in the ss-bound Sec complex (light green) with the apo Sec complex (dark green) from this study (A), with the apo Sec complex from Itskanov *et al* (PDB ID 6N3Q, B) and with the ss-opened ribosome-bound Sec61 complex (PDB ID 3JC2, C). Structure alignments are based on TMs 7-9 of Sec61α. Black arrows indicate the movements of the helices of the Sec61α. The plug domain is indicated by "P".

Again, the overall position of the ppαF signal sequence in this study is largely consistent with the position of a more hydrophobic co-translational signal sequence of preprolactin (**Figure 22**D; PDB ID 3JC2; Voorhees & Hegde, 2016) and with bacterial post-translational signal sequences observed in SecA-SecY complexes (**Figure 22**E; PDB ID 6ITC; Ma et al, 2019; Li et al, 2016). However, in the structure of the signal sequence-engaged Sec complex, the lateral gate is significantly wider than in any other known structure of signal sequence-bound Sec61/SecY channels (**Figure 22**D and E).

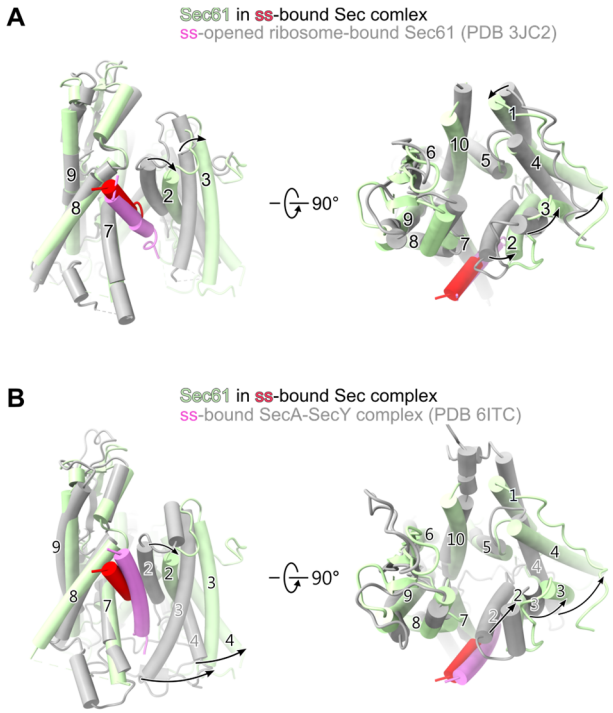


Figure 22. Comparison of signal sequence-engaged Sec61/SecY conformations. Comparison of Sec61α in the ss-bound Sec complex (light green) with the ss-opened ribosome-bound Sec61 complex (PDB ID 3JC2, A) and with SecY in the ss-bound SecA-SecY complex (PDB ID 6ITC, B). All structures are shown from a view focusing on the lateral gate (left) and from the top (cytosolic side; right). Structure alignments are based on TMs 7-9 of Sec61α. Black arrows indicate the movements of the helices of the Sec61α.

Interestingly, the observation that the TM helices of Sec62 were positioned close to the inserted ppαF signal sequence suggested that Sec62 may act by being in direct contact with signal sequences, which would also be in agreement with previous chemical cross-linking data (Plath *et al*, 1998, 2003). In fact, in the signal sequence-engaged state, the still moderately flexible TM2 of Sec62 could directly contact the signal sequence whereas TM1 was in proximity close enough to possibly contact the lateral gate helices TM2 and TM3 of Sec61α (**Figure 20**B, upper panel). Besides, the flexibility of Sec62 TM helices is implied by the weak density, presumably the TMs of Sec62, in the low-pass filtered map of the apo state reconstruction (**Figure 20**A, star in the lower panel), which was also present in the previous structures (Wu *et al*, 2019; Itskanov & Park, 2019). Overall, the more rigid conformation of Sec62 in the engaged state hints at a stabilizing function for the active open channel conformation. This notion is supported by the less pronounced appearance of the N-terminal cytosolic domain of Sec62 and the suspected weak density for the Sec62 TM helices (**Figure 20**A, lower panel) indicating an overall flexibility of Sec62 in the inactive apo state.

3.6 Structural assignment of Sec62-N to the architecture of the Sec complex

The structure of the Sec62-N domain (residues 18-145; **Figure 23**A) was determined by X-ray crystallography to a resolution of 2.5 Å, which was done by Dr. Jingdong Cheng. The structure of Sec62-N displays an elongated shape and can be divided into two tightly interacting lobes: a four-helix bundle (lobe 1) and a β-barrel (lobe 2) (**Figure 23**B). Surprisingly, no structural homologue was found, neither by sequence similarity nor by structure comparison using the DALI server (Holm, 2019). Thus, this novel domain was named the "Sec62 domain".

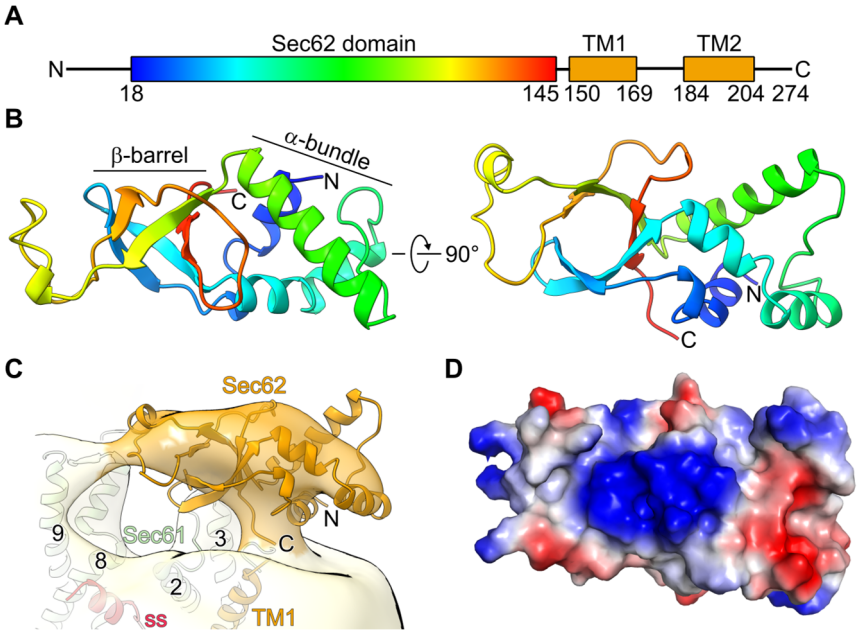


Figure 23. Crystal structure of Sec62 domain.

(A) Schematic view of the domain organization of Sec62. The N-terminal cytoplasmic domain (Sec62 domain) is colored in rainbow. (B) Two views on the crystal structure of the Sec62 domain consisting of an α-helical bundle and a β-barrel domain. (C) Fit of the Sec62 domain crystal structure into the respective cryo-EM density of the ss-engaged heptameric Sec complex as in **Figure 18**C. The map is low-pass filtered to 15 Å. (D) Surface representation of the Sec62 domain colored according to the electrostatic potential. The viewing angle is the same as in the right panel in (B).

The dimension of Sec62-N agreed overall with the cryo-EM density at the given local resolution of about 15 Å, and it was positioned accordingly into the engaged state adjacent to the front side of Sec61α in order to allow for connectivity with the putative TMs of Sec62 (**Figure 23**C). The apparent flexibility of this domain as indicated by the limited resolution did not allow for secondary structure-based fitting and more exact positioning. Notably, in the signal sequence-engaged state map, no large additional or connecting density was present to explain the previously described interaction between the Sec62 and Sec63 cytosolic domains, despite the essential nature of this interaction for translocation activity (Wittke *et al*, 2000). However, the interaction between Sec63 and Sec62-N was previously shown to critically rely on the acidic C-terminus of Sec63, in particular the small region comprising the amino acids 650-663 (Wittke *et al*, 2000). Moreover, the protein kinase CK2-dependent phosphorylation of the threonines 652 and 654 of Sec63 is necessary for the binding of Sec62-N (Wang & Johnsson, 2005). This feature is conserved, because also mammalian Sec62 interacts with the phosphorylated acidic C-terminus of Sec63 *in vitro* (Ampofo *et al*, 2013). Therefore, it is likely that the interaction site of Sec62 for the Sec63 C-terminus harbors a positively charged patch. Indeed, the electrostatic potential surface map of the Sec62 domain revealed a pronounced positively charged surface in the β-barrel lobe (**Figure 23**D) as a candidate region. A sequence alignment of the Sec62 domain showed several positively charged residues which are conserved in fission yeast and three of them (R27, R51 and R104) are even conserved from yeast to humans (**Figure 24**).

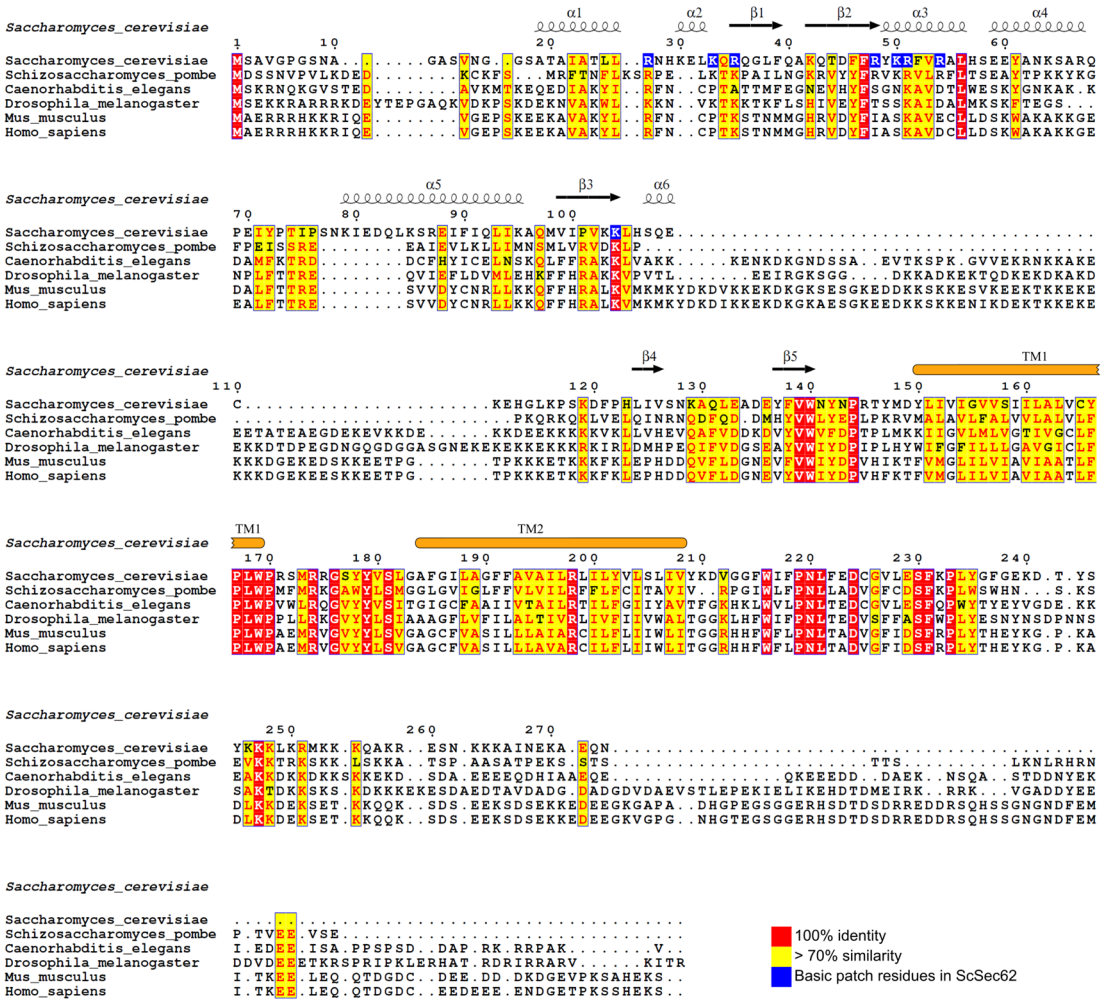


Figure 24. Protein sequence alignment of Sec62 orthologs.
The multiple sequence alignment was performed with T-Coffee program (Notredame *et al*, 2000) and was plotted using ESPript (http://espript.ibcp.fr; Robert & Gouet, 2014). The secondary structure of the Sec62 domain and the predicted TMs (orange bars) are shown above the sequences.

To better characterize the Sec62-Sec63 interaction by biochemical and biophysical means, Dr. Cheng employed two synthetic biotin-labeled peptides representing residues 621-647 (C1) and 648-663 (C2) of Sec63 (**Figure 25**A) for *in vitro* pull-down assays with the purified Sec62-N. In agreement with the previous studies (Wittke *et al*, 2000; Wang & Johnsson, 2005), the pull-down assay showed that the Sec62 domain could be co-purified using both C1 and phosphorylated C2 (C2P), however not with unphosphorylated C2 (**Figure 25**B). Isothermal titration calorimetry (ITC) experiments

further confirmed that Sec63-C2P and Sec62-N interact with nanomolar affinity, whereas no binding was observed for unmodified Sec63-C2 (**Figure 25**C).

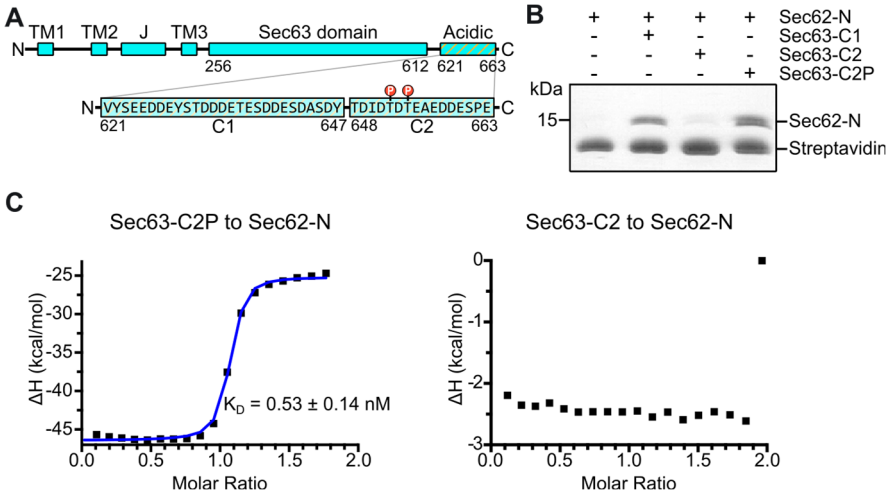


Figure 25. Interaction of Sec62 domain and Sec63 C-terminal peptides. (A) Domain organization of Sec63. The sequence of the acidic patch on the C-terminus—divided into C1 and C2 regions—is shown in the lower panel. The two phosphorylated threonines (T652 and T654) are marked in the C2 region. (B) Pull-down assay of Sec62 domain with different biotin-labeled Sec63 C-terminal peptides. Sec63-C2P denotes di-phosphorylated Sec63-C2. (C) Thermogram obtained from isothermal titration calorimetry (ITC) after titration of Sec62 domain with increasing amounts of Sec63-C2P (left panel) or Sec63-C2 (right panel).

Chapter 4: Discussion

Ever since the first crystal structure of the Sec61/SecY channel has been revealed (van den Berg *et al*, 2004), plenty of structural data on a substrate-engaged Sec61/SecY channel have been reported (Frauenfeld *et al*, 2011; Park *et al*, 2014; Voorhees & Hegde, 2016; Li *et al*, 2016; Ma *et al*, 2019; Kater *et al*, 2019). These structural studies show how the substrates—a signal sequence or a signal anchor—are recognized by an active Sec61/SecY channel. However, these structures are either in co-translational mode with the ribosome, or in bacterial post-translational SecY-SecA complex. Such insights for eukaryotic post-translational translocon were still missing because of the difficulty in acquiring structural data on membrane proteins. Membrane proteins are difficult to crystalize, and there is only one crystal structure in the above-mentioned substrate-engaged structures (Li *et al*, 2016). Although the others came from single particle cryo-EM analysis, cryo-EM was for a long time not capable of resolving structures for small proteins at high resolution due to sample preparation. Membrane proteins often show high affinity to grid support and do not enter hole area. One could prepare their samples at high concentration to force proteins to distribute in grid holes. But to prepare a large amount of membrane proteins is not easy compared to preparation of soluble proteins. Another way is to apply a thin (20-50 Å) amorphous carbon film on grids for better particle distribution (Bernal & Stock, 2004; Passmore & Russo, 2016), but the carbon film increases background noise, resulting in low signal-to-noise ratio for smaller proteins (Zhao *et al*, 2015). Aside from the general challenges for small membrane proteins, another one is to reconstitute a stable substrate-bound Sec complex for single particle cryo-EM analysis. Because a translocating peptide can slide through the channel of the Sec complex, measures must be taken to prevent the substrate from

fully translocated and subsequently dissociated from the complex. To overcome these challenges, in this thesis, the Sec complex was purified in large scale and reconstituted on-bead with the mEGFP-fused ppαF to achieve a highly concentrated sample (5 mg/ml) of the stable substrate-bound Sec complex.

4.1 Comparison of the Sec complex in its apo or signal sequence-engaged state

By using the recombinant ppαF-mEGFP as a substrate and on-bead *in vitro* reconstitution, the cryo-EM structures of the heptameric Sec complex in both substrate-engaged and apo states could be acquired. Although the resolutions are relatively low, processing methods such as Bayesian polishing and CTF refinement in RELION could not provide further improvement. This is likely due to intrinsic flexibility of these highly dynamic complexes, which is an advantage to adapt to various side chain sizes of the translocating peptide. In spite of the comparably low resolutions, all α-helices in the transmembrane domain of the complex could be assigned, allowing a detailed description of the overall architecture and conformation of the signal sequences-engaged state, as well as comparisons with previously described structures.

When comparing all three available apo state Sec complex structures, certain dynamics in Sec61α lateral gate opening and plug position become evident. Notably, the conformation of Sec61α in the apo state in this thesis resembles the one described in Wu *et al*, 2019 (**Figure 19**), whereas the structure of the signal sequence-engaged state with the more open lateral gate is more similar to the apo state Sec complex described in Itskanov *et al*, 2019 (**Figure 21**B). Interestingly, in all three maps of apo states (Wu *et al*, 2019; Itskanov & Park, 2019; this study), extra densities exist within the lateral gate at lower contour levels (**Figure 26**), that may derive from a bound lipid or detergent

molecule. In fact, in the study of Itskanov *et al*, detergents LMNG and CHS were used for solubilization of the Sec complex, whereas digitonin or its analog glyco-diosgenin (GDN) were used in Wu *et al*. or in this thesis, respectively. This explains that the apo state structure in this thesis is more similar to the structure in Wu *et al*. This also indicates that the lateral gate opening of the Sec complex may be affected by size, bulkiness or chemical properties of different detergents or lipids, which may somewhat mimic the signal sequence or act as place holders for the signal sequence. However, upon engagement of the ppαF signal sequence, the lateral gate opens even wider than in other signal sequence-engaged structures (**Figure 21**D and E).

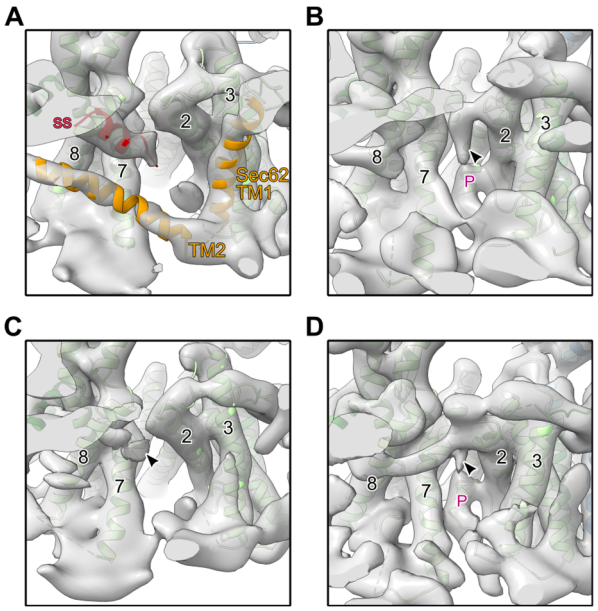


Figure 26. Detergents or lipids bound to the lateral gate of Sec complex. (A) Cryo-EM map and fitted model of the ss-bound Sec complex from this study. (B) Cryo-EM map and fitted model of the apo Sec complex from this study. (C and D) Cryo-EM maps and fitted models of the apo Sec complexes from the Itskanov study (C; PDB ID 6N3Q and EMD-0336) and Wu study (D; PDB ID 6ND1 and EMD-0440). All maps are low-pass filtered to 8 Å adjust at the contour level where detergent micelle density starts to appear. The black triangle marks extra density likely attributing for a bound detergent or lipid molecule and "P" marks the plug.

Lateral gate widening upon signal sequence recognition is known to cause a destabilization and relocation of the plug helix, thereby removing the obstruction from the aqueous peptide channel (Voorhees & Hegde, 2016; Li *et al*, 2016; Ma *et al*, 2019). Also, in the more open signal sequence-engaged state, the plug helix became flexible and relocated in order to open the central pore. Although the translocating peptide in the central pore of Sec61α could not be visualized, this plug relocation was expected because it is required to open the channel and allow translocation of a peptide into the ER lumen.

4.2 Role of Sec62 in the post-translational Sec translocon

In the signal sequence-engaged state, this widened structure of the lateral gate may be further stabilized by the Sec62 TMs. So far, the essential function for Sec62 during translocation remained largely elusive and no densities for the Sec62 TMs could be observed in previous structures due to flexibility (Wu *et al*, 2019; Itskanov & Park, 2019). Previous photo-crosslinking studies showed that the signal sequence of ppαF can be crosslinked to the TMs 2 and 7 of Sec61 as well as either Sec62 or Sec71, which could not be differentiated on SDS-PAGE (Plath *et al*, 1998, 2003). Since Sec71 is structurally not close to the lateral gate, where a signal sequence bind, it is Sec62 that crosslinked with the signal sequence. The structure of the signal sequence-engaged Sec complex clearly shows a relative rigidification of the Sec62 TMs and especially its TM2 in close contact with the signal sequence. This may help stabilize the signal sequence surrounded by lipids at the lateral gate. This stabilization could be necessary for recognition of the less hydrophobic post-translational signal sequences, thereby explaining Sec62's crucial role for post-translational translocation. On the other hand, it should be noted that in the signal sequence-engaged map the density for the Sec62 TM2 does not span completely across the membrane, which might be caused by the

presence of detergent.

Despite a near-complete structure of the heptameric Sec complex bound to a signal sequence, density for the cytosolic N- and C-termini of Sec62 was weak or less resolved than the rest of the complex. Yet, both termini were shown to be binding sites for the Sec complex (Wittke *et al*, 2000). While the interaction partner for the minor binding site in the Sec62 C-terminus is still unknown, Sec62 mainly uses its positively charged N-terminal domain to interact with the acidic C-terminal 14 residues of Sec63, which is regulated by CK2-dependent phosphorylation (Wittke *et al*, 2000; Wang & Johnsson, 2005). Mutations to these basic residues also led to deficiency in translocation of less hydrophobic transmembrane segments (Jung *et al*, 2014). To delineate the interaction between the phosphorylated Sec63 C-terminal tail (Sec63-C2P) and Sec62-N, hydrogen-deuterium exchange mass spectrometry (HDX-MS; Kochert *et al*, 2018) was employed in collaboration with Professor Dr. Gert Bange lab in University of Marburg in the course of this thesis. In brief, regions of most severe HDX reduction mapped to the β-barrel lobe 2 of Sec62-N domain, which essentially represents the highly positively charged patch identified in the crystal structure, while the α-helical bundle centering around the N-terminus of Sec62 domain was less affected (**Figure 27**; compare also to: **Figure 23**D; for more details see Weng *et al*, 2021). Thus, considering the highly negatively charged C-terminus of the Sec63, it is suggested that this patch of Sec62 N-terminal domain with its multitude of arginine and lysine residues would represent the interaction platform for Sec63-C2P.

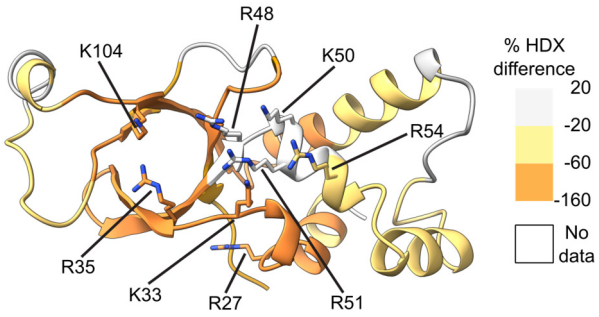


Figure 27. Mapping of interaction site on Sec62 using HDX.
Crystal structure of Sec62 domain color coded according to the difference in HDX between the Sec62 domain in complex with Sec63-C2P and the Sec62 domain alone. The critical residues constituting the interaction interface between Sec63-C2P and Sec62 are labelled.

The crystal structure of the Sec62 domain together with the HDX analysis shows that Sec62 utilizes the basic patch on its N-terminal domain to interact with the ultimate C-terminus (Sec63-C2 comprising the last 16 residues 648-663) of Sec63, and that binding is dependent on threonine phosphorylation of Sec63-C2 (Figs 3D and 4). Besides, the more N-terminal acidic stretch of the Sec63 C-terminus (Sec63-C1; residues 621-647) also shows interaction with the Sec62 domain. However, given that the last Sec63 residue in the model (D612) is more than 40 Å away from the Sec62 basic patch and that the linker between Sec63-C1 and the Sec63 globular domain has only 8 residues, it was assumed that Sec63-C2P and not Sec63-C1 can span the distance and therefore acts as the primary binding partner for the Sec62 domain (**Figure 28**). Alternatively, Sec63-C1 may also interact with the C-terminus of Sec62, which is also positively charged (**Figure 24**). In the maps of both apo and engaged state, a low-resolved, unassigned cytosolic density was observed connecting to the Sec62 TM2, which may attribute for the C-terminus of Sec62 (**Figure 20**A). Accordingly, this density is close to the Sec63 C-terminus (**Figure 20**A and **Figure 28**), further supporting the idea that

Sec63-C1 may interact with the C-terminus of Sec62. Along those lines, deletion of either the N-terminal cytosolic domain or the C-terminal 35 residues of Sec62 not only weakens the interaction with the Sec complex but also causes defects in protein translocation (Wittke *et al*, 2000). Considering that the TM2 of Sec62 stabilizes the signal sequence, the N- and C-termini of Sec62 likely serve as anchors on the Sec complex, bringing the flexible TMs of Sec62 in close proximity to the lateral gate in order to prepare the Sec complex for signal sequence engagement.

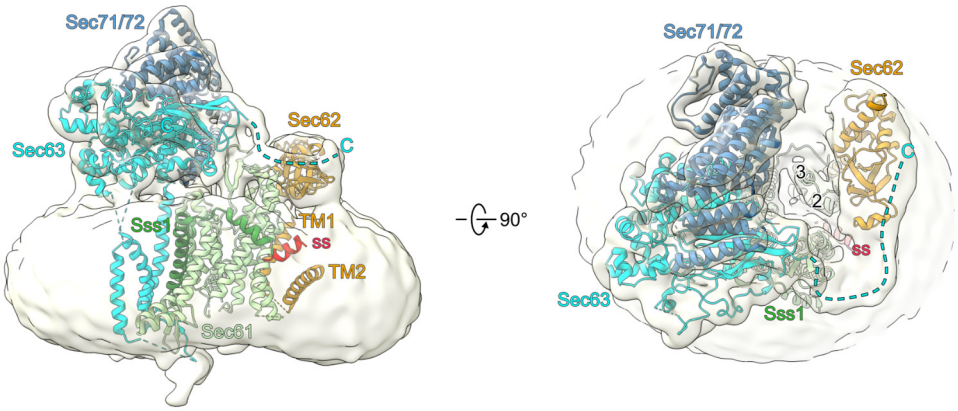


Figure 28. Connection between Sec62 and the Sec complex.
Original cryo-EM reconstruction of the ss-bound Sec complex shown before focused refinement and low-pass filtered to 8 Å to visualize the rather flexible C-terminus of Sec63 (highlighted by a cyan dashed line).

4.3 Model for post-translational protein translocation across the ER membrane

Altogether, the structures of the signal sequence-engaged Sec complex and the Sec62 domain provide a refined model for post-translational protein translocation across the ER membrane (**Figure 29**): To enable translocation of post-translational clients in the ER, the Sec61 complex assembles with Sec62/63 complex, and in yeast with additional subunits Sec71 and Sec72 into the Sec complex (post-translocon). Because of its interaction with Sec63, Sec61α is already in a conformation with a pre-opened lateral

gate as observed in all three available structures of the apo state (Wu *et al*, 2019; Itskanov & Park, 2019; this thesis). In this conformation it is already primed for signal sequence engagement while the plug still blocks the Sec61α channel for the translocating peptide. Within this assembly, Sec62 is very flexible and is anchored on the Sec complex primarily via the acidic C-terminal tail of Sec63 with the basic patch of the N-terminal cytosolic domain of Sec62. Upon substrate engagement, the signal sequence binds to the groove at the lateral gate and is further stabilized by the TM2 of Sec62. Consequently, TM1 of Sec62 is brought closer to the lateral gate and displaces TM3 of Sec61α. This movement causes the lateral gate to be even mor open, followed by plug displacement and translocation of the polypeptide chain.

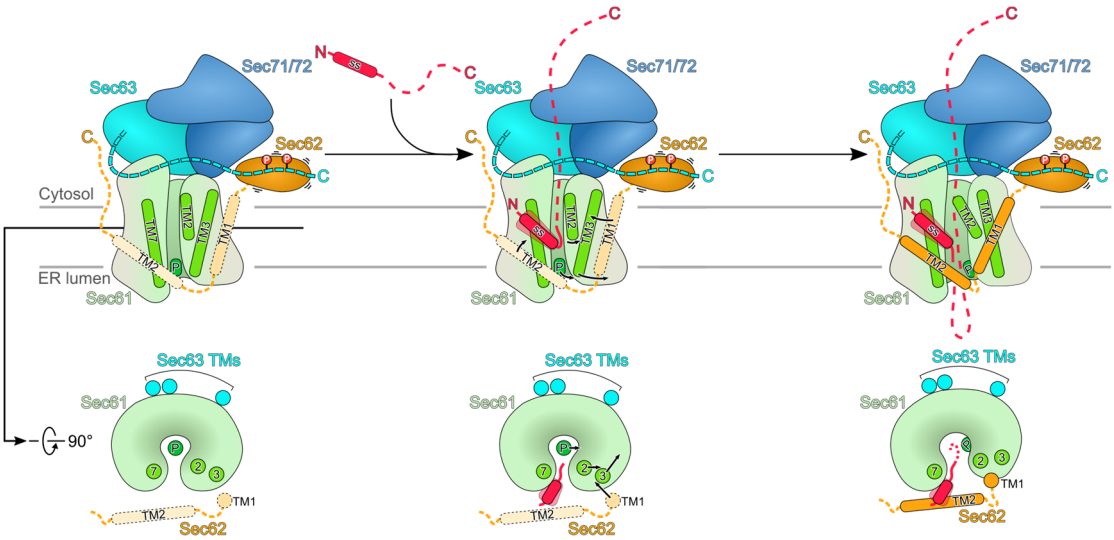


Figure 29. Model for substrate engagement of the post-translational Sec translocon.

In the apo state of Sec complex, the Sec61 channel is closed by the plug while the lateral gate is already open. Sec62 flexibly associates with the Sec complex mainly through a basic surface on its Sec62 domain interacting with the acidic C-terminus of Sec63. During the insertion of a post-translational client, its signal sequence (ss) binds to the groove at lateral gate supported by the Sec62 TM2. At the same time, the Sec62 TM1 pushes the Sec61α TM3 outwards leading to an even more open lateral gate. This also leads to removal of the plug away from the pore ring, allowing, the translocating peptide to be gated through the Sec61α central pore.

Chapter 5: Outlook

This study, by combining multiple structural methods, provides a more complete picture of how a less hydrophobic signal sequence engages the post-translational Sec complex, how Sec62 is assembled in the Sec complex, and the role Sec62 plays in signal sequence engagement. Nevertheless, these results have raised new questions to be addressed in future research.

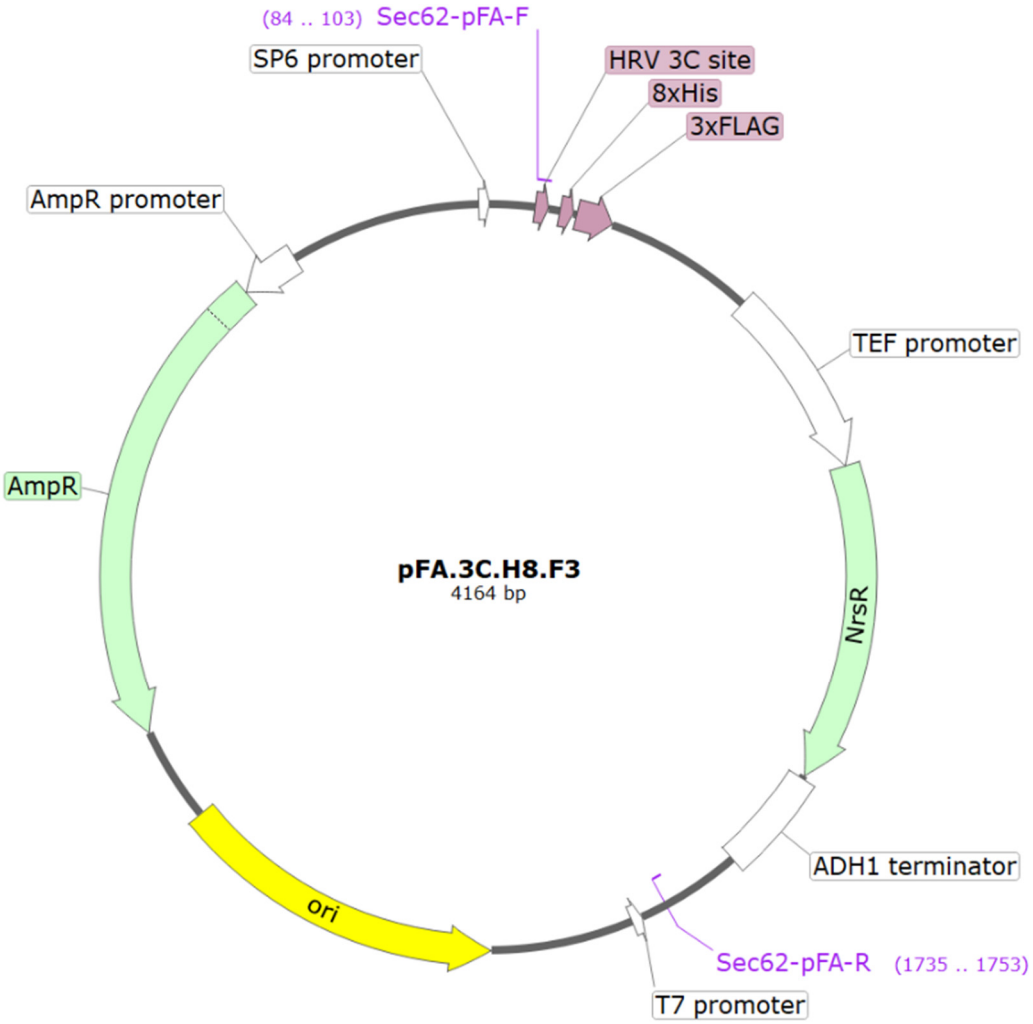
First, the detailed molecular interaction between the signal sequence and the Sec complex is still unclear. Although the structure of the signal sequence-bound Sec complex shows the overall engagement, the side chains of the signal sequence could not be registered because of low local resolution at the lateral gate. Post-translational translocation is part of secretory pathways, which are not only important in fundamental cellular process, but also in virulence of pathogenic fungi (Rollenhagen *et al*, 2020). Identifying the interaction between a signal sequence and the Sec complex in atomic resolution may provide a potential drug target for treating fungal infection.

In addition, regarding the interaction between Sec62 and Sec63, questions on how this is regulated still remain. Studies have shown that, independent of Sec62, Sec63 is also involved in nuclear fusion and some co-translational translocation (Ng & Walter, 1996; Young *et al*, 2001). Therefore, regulation of Sec62-Sec63 interaction, such as CK2 phosphorylation on Sec63 C-terminus, is probably necessary for Sec63 to perform multiple functions. However, while removing phosphorylation on Sec63 C-terminus disrupts interaction between Sec62-N and the ultimate C-terminus of Sec63, Sec62 still remain loosely bound to the Sec complex through its cytosolic, basic C-terminus. Given that the binding partner for the Sec62 C-terminus is possibly the acidic C-terminal tail of Sec63, further regulation may be required.

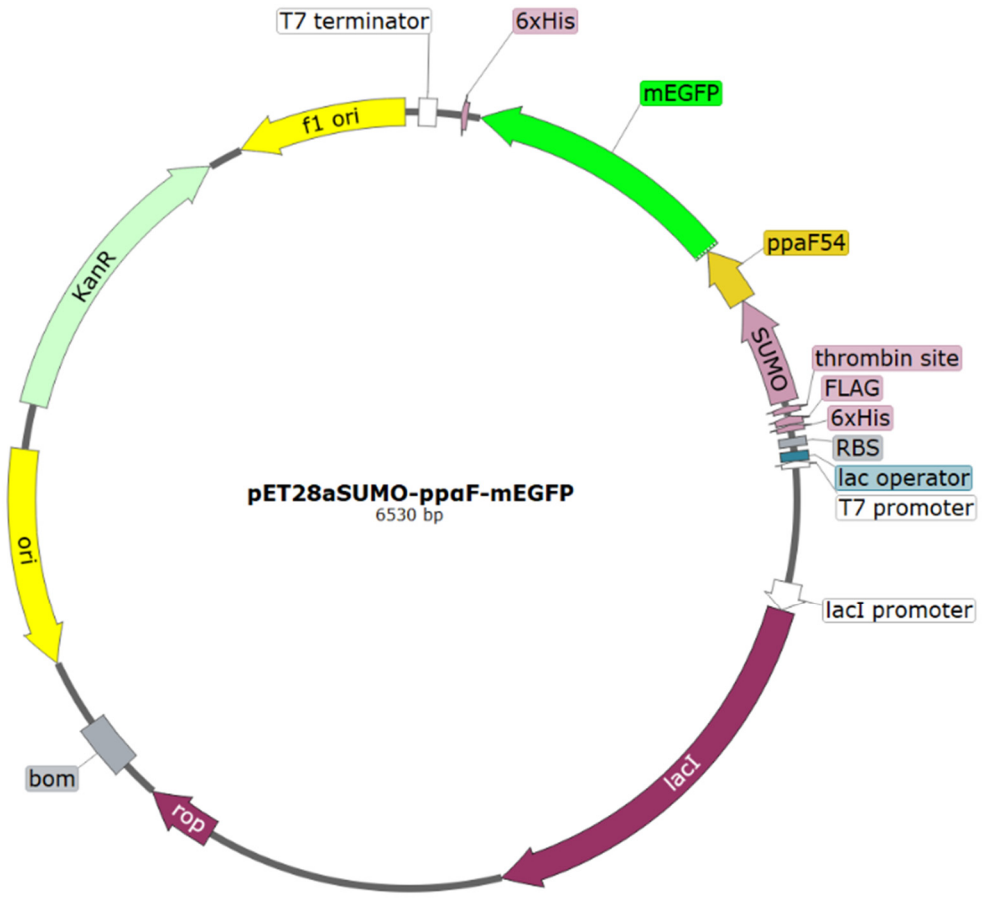
Finally, considering that Sec71 and Sec72 only exist in fungi, structural investigation on post-translational Sec complex in higher eukaryotes is crucial. Compared to the Sec proteins in yeast, the residue number of cytosolic C-terminal tail of human Sec62 is doubled (~70 residues in yeast), and the cytosolic C-terminal domain of human Sec63 has ~100 more residues (424 residues in yeast). In terms of function, in contrast to the yeast Sec complex the mammalian one (Sec61/62/63) serves as a fail-safe pathway for shorter secretory proteins (Lakkaraju *et al*, 2012). Additionally, the co-translational Sec61/63 complex in mammals was found to be involved in unfolded protein response by interacting with Ire1α (Plumb *et al*, 2015; Sundaram *et al*, 2017). Further structural study on mammalian complex could possibly reveal mechanisms that different from that in yeast.

Appendix

Plasmid 1, used for genomic tagging at the C-terminus of Sec62 in *S. cerevisiae*. Positions of the primers used in PCR amplification are marked in pink. NsrR: nourseothricin acetyltransferase.



Plasmid 2, used for over-expressing ppαF-mEGFP or its mutant ppαFm3-mEGFP in E. coli. A stop codon is at the end of mEGFP sequence.



References

Allison DS & Young ET (1989) Mutations in the signal sequence of prepro-alpha-factor inhibit both translocation into the endoplasmic reticulum and processing by signal peptidase in yeast cells. *Mol Cell Biol* 9: 4977–4985

Ampofo E, Welker S, Jung M, Müller L, Greiner M, Zimmermann R & Montenarh M (2013) CK2 phosphorylation of human Sec63 regulates its interaction with Sec62. *Biochim Biophys Acta* 1830: 2938–2945

Ast T, Cohen G & Schuldiner M (2013) A network of cytosolic factors targets SRP-independent proteins to the endoplasmic reticulum. *Cell* 152: 1134–1145

Bauer BW, Shemesh T, Chen Y & Rapoport TA (2014) A "Push and Slide" Mechanism Allows Sequence-Insensitive Translocation of Secretory Proteins by the SecA ATPase. *Cell* 157: 1416–1429

Becker J, Walter W, Yan W & Craig EA (1996) Functional interaction of cytosolic hsp70 and a DnaJ-related protein, Ydj1p, in protein translocation in vivo. *Mol Cell Biol* 16: 4378–4386

van den Berg B, Clemons WM, Collinson I, Modis Y, Hartmann E, Harrison SC & Rapoport TA (2004) X-ray structure of a protein-conducting channel. *Nature* 427: 36–44

Bernal RA & Stock D (2004) Three-Dimensional Structure of the Intact Thermus thermophilus H+-ATPase/Synthase by Electron Microscopy. *Structure* 12: 1789–1798

Blobel G & Dobberstein B (1975) 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

Braunger K, Pfeffer S, Shrimal S, Gilmore R, Berninghausen O, Mandon EC, Becker T, Förster F & Beckmann R (2018) Structural basis for coupling protein transport and N-glycosylation at the mammalian endoplasmic reticulum. *Science* 360: 215–219

Cannon KS, Or E, Clemons WM, Shibata Y & Rapoport TA (2005) Disulfide bridge formation between SecY and a translocating polypeptide localizes the

translocation pore to the center of SecY. *J Cell Biol* 169: 219–225

Catipovic MA, Bauer BW, Loparo JJ & Rapoport TA (2019) Protein translocation by the SecA ATPase occurs by a power-stroke mechanism. *EMBO J* 38: e101140

Catipovic MA & Rapoport TA (2020) Protease protection assays show polypeptide movement into the SecY channel by power strokes of the SecA ATPase. *EMBO Rep* 21: e50905

Denisov IG & Sligar SG (2016) Nanodiscs for structural and functional studies of membrane proteins. *Nat Struct Mol Biol* 23: 481–486

Deshaies RJ, Sanders SL, Feldheim DA & Schekman R (1991) Assembly of yeast Sec proteins involved in translocation into the endoplasmic reticulum into a membrane-bound multisubunit complex. *Nature* 349: 806–808

Deshaies RJ & Schekman R (1989) SEC62 encodes a putative membrane protein required for protein translocation into the yeast endoplasmic reticulum. *J Cell Biol* 109: 2653–2664

Deshaies RJ & Schekman R (1990) Structural and functional dissection of Sec62p, a membrane-bound component of the yeast endoplasmic reticulum protein import machinery. *Mol Cell Biol* 10: 6024–6035

Dünnwald M, Varshavsky A & Johnsson N (1999) Detection of Transient In Vivo Interactions between Substrate and Transporter during Protein Translocation into the Endoplasmic Reticulum. *Mol Biol Cell* 10: 329–344

Egea PF, Shan S, Napetschnig J, Savage DF, Walter P & Stroud RM (2004) Substrate twinning activates the signal recognition particle and its receptor. *Nature* 427: 215–221

Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66: 486–501

Feldheim D & Schekman R (1994) Sec72p contributes to the selective recognition of signal peptides by the secretory polypeptide translocation complex. *J Cell Biol* 126: 935–943

Feldheim D, Yoshimura K, Admon A & Schekman R (1993) Structural and functional characterization of Sec66p, a new subunit of the polypeptide translocation apparatus in the yeast endoplasmic reticulum. *Mol Biol Cell* 4: 931–939

Frauenfeld J, Gumbart J, van der Sluis EO, Funes S, Gartmann M, Beatrix B, Mielke T, Berninghausen O, Becker T, Schulten K, *et al* (2011) Cryo–EM structure of the ribosome–SecYE complex in the membrane environment. *Nat Struct Mol Biol* 18: 614–621

Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD & Bairoch A (2005) Protein Identification and Analysis Tools on the ExPASy Server. In *The Proteomics Protocols Handbook*, Walker JM (ed) pp 571–607. Totowa, NJ: Humana Press

Gemmer M & Förster F (2020) A clearer picture of the ER translocon complex. *J Cell Sci* 133

Goddard TD, Huang CC, Meng EC, Pettersen EF, Couch GS, Morris JH & Ferrin TE (2018) UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci Publ Protein Soc* 27: 14–25

Gogala M, Becker T, Beatrix B, Armache J-P, Barrio-Garcia C, Berninghausen O & Beckmann R (2014) Structures of the Sec61 complex engaged in nascent peptide translocation or membrane insertion. *Nature* 506: 107–110

Gong X, Qian H, Zhou X, Wu J, Wan T, Cao P, Huang W, Zhao X, Wang X, Wang P, *et al* (2016) Structural Insights into the Niemann-Pick C1 (NPC1)-Mediated Cholesterol Transfer and Ebola Infection. *Cell* 165: 1467–1478

Görlich D & Rapoport TA (1993) Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell* 75: 615–630

Haber JE (2012) Mating-Type Genes and MAT Switching in Saccharomyces cerevisiae. *Genetics* 191: 33–64

Halic M, Blau M, Becker T, Mielke T, Pool MR, Wild K, Sinning I & Beckmann R (2006) Following the signal sequence from ribosomal tunnel exit to signal recognition particle. *Nature* 444: 507–511

Hanada M, Nishiyama KI, Mizushima S & Tokuda H (1994) Reconstitution of an efficient protein translocation machinery comprising SecA and the three membrane proteins, SecY, SecE, and SecG (p12). *J Biol Chem* 269: 23625–23631

Harada Y, Li H, Wall JS, Li H & Lennarz WJ (2011) Structural Studies and the

Assembly of the Heptameric Post-translational Translocon Complex. *J Biol Chem* 286: 2956–2965

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

Holm L (2019) Benchmarking fold detection by DaliLite v.5. *Bioinforma Oxf Engl* 35: 5326–5327

Itskanov S & Park E (2019) Structure of the posttranslational Sec protein-translocation channel complex from yeast. *Science* 363: 84–87

Janda CY, Li J, Oubridge C, Hernández H, Robinson CV & Nagai K (2010) Recognition of a signal peptide by the signal recognition particle. *Nature* 465: 507–510

Julius D, Schekman R & Thorner J (1984) Glycosylation and processing of prepro-alpha-factor through the yeast secretory pathway. *Cell* 36: 309–318

Jung S, Kim JEH, Reithinger JH & Kim H (2014) The Sec62–Sec63 translocon facilitates translocation of the C-terminus of membrane proteins. *J Cell Sci* 127: 4270–4278

Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66: 125–132

Kalies K-U, Rapoport TA & Hartmann E (1998) The β Subunit of the Sec61 Complex Facilitates Cotranslational Protein Transport and Interacts with the Signal Peptidase during Translocation. *J Cell Biol* 141: 887–894

Kater L, Frieg B, Berninghausen O, Gohlke H, Beckmann R & Kedrov A (2019) Partially inserted nascent chain unzips the lateral gate of the Sec translocon. *EMBO Rep* 0: e48191

Kochert BA, Iacob RE, Wales TE, Makriyannis A & Engen JR (2018) Hydrogen-Deuterium Exchange Mass Spectrometry to Study Protein Complexes. *Methods Mol Biol Clifton NJ* 1764: 153–171

Kühlbrandt W (2014) Cryo-EM enters a new era. *eLife* 3: e03678

Kurjan J & Herskowitz I (1982) Structure of a yeast pheromone gene (MF alpha): a putative alpha-factor precursor contains four tandem copies of mature alpha-

factor. *Cell* 30: 933–943

Lakkaraju AKK, Thankappan R, Mary C, Garrison JL, Taunton J & Strub K (2012) Efficient secretion of small proteins in mammalian cells relies on Sec62-dependent posttranslational translocation. *Mol Biol Cell* 23: 2712–2722

Li L, Park E, Ling J, Ingram J, Ploegh H & Rapoport TA (2016) Crystal structure of a substrate-engaged SecY protein-translocation channel. *Nature* 531: 395–399

Lyman SK & Schekman R (1997) Binding of Secretory Precursor Polypeptides to a Translocon Subcomplex Is Regulated by BiP. *Cell* 88: 85–96

Ma C, Wu X, Sun D, Park E, Catipovic MA, Rapoport TA, Gao N & Li L (2019) Structure of the substrate-engaged SecA-SecY protein translocation machine. *Nat Commun* 10: 2872

Mason N, Ciufo LF & Brown JD (2000) Elongation arrest is a physiologically important function of signal recognition particle. *EMBO J* 19: 4164–4174

Matlack KE, Misselwitz B, Plath K & Rapoport TA (1999) BiP acts as a molecular ratchet during posttranslational transport of prepro-alpha factor across the ER membrane. *Cell* 97: 553–564

Matlack KES, Plath K, Misselwitz B & Rapoport TA (1997) Protein Transport by Purified Yeast Sec Complex and Kar2p Without Membranes. *Science* 277: 938–941

Merlini L, Dudin O & Martin SG (2013) Mate and fuse: how yeast cells do it. *Open Biol* 3

Mothes W, Prehn S & Rapoport TA (1994) Systematic probing of the environment of a translocating secretory protein during translocation through the ER membrane. *EMBO J* 13: 3973–3982

Ng DT, Brown JD & Walter P (1996) Signal sequences specify the targeting route to the endoplasmic reticulum membrane. *J Cell Biol* 134: 269–278

Ng DT & Walter P (1996) ER membrane protein complex required for nuclear fusion. *J Cell Biol* 132: 499–509

Ngosuwan J, Wang NM, Fung KL & Chirico WJ (2003) Roles of cytosolic Hsp70 and Hsp40 molecular chaperones in post-translational translocation of presecretory

proteins into the endoplasmic reticulum. *J Biol Chem* 278: 7034–7042

Notredame C, Higgins DG & Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* 302: 205–217

Panzner S, Dreier L, Hartmann E, Kostka S & Rapoport TA (1995) Posttranslational protein transport in yeast reconstituted with a purified complex of Sec proteins and Kar2p. *Cell* 81: 561–570

Park E, Ménétret J-F, Gumbart JC, Ludtke SJ, Li W, Whynot A, Rapoport TA & Akey CW (2014) Structure of the SecY channel during initiation of protein translocation. *Nature* 506: 102–106

Park E & Rapoport TA (2011) Preserving the membrane barrier for small molecules during bacterial protein translocation. *Nature* 473: 239–242

Park E & Rapoport TA (2012) Mechanisms of Sec61/SecY-Mediated Protein Translocation Across Membranes. *Annu Rev Biophys* 41: 21–40

Passmore LA & Russo CJ (2016) Chapter Three - Specimen Preparation for High-Resolution Cryo-EM. In *Methods in Enzymology*, Crowther RA (ed) pp 51–86. Academic Press

Phillips GN (1997) Structure and dynamics of green fluorescent protein. *Curr Opin Struct Biol* 7: 821–827

Plath K, Mothes W, Wilkinson BM, Stirling CJ & Rapoport TA (1998) Signal sequence recognition in posttranslational protein transport across the yeast ER membrane. *Cell* 94: 795–807

Plath K, Wilkinson BM, Stirling CJ & Rapoport TA (2003) Interactions between Sec Complex and Prepro-α-Factor during Posttranslational Protein Transport into the Endoplasmic Reticulum. *Mol Biol Cell* 15: 1–10

Plumb R, Zhang Z-R, Appathurai S & Mariappan M (2015) A functional link between the co-translational protein translocation pathway and the UPR. *eLife* 4: e07426

Punjani A, Rubinstein JL, Fleet DJ & Brubaker MA (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 14: 290–296

Rapoport TA, Li L & Park E (2017) Structural and Mechanistic Insights into Protein Translocation. *Annu Rev Cell Dev Biol*

Robert X & Gouet P (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* 42: W320–W324

Rohou A & Grigorieff N (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* 192: 216–221

Rollenhagen C, Mamtani S, Ma D, Dixit R, Eszterhas S & Lee SA (2020) The Role of Secretory Pathways in Candida albicans Pathogenesis. *J Fungi* 6

Rothblatt JA, Deshaies RJ, Sanders SL, Daum G & Schekman R (1989) Multiple genes are required for proper insertion of secretory proteins into the endoplasmic reticulum in yeast. *J Cell Biol* 109: 2641–2652

Seddon AM, Curnow P & Booth PJ (2004) Membrane proteins, lipids and detergents: not just a soap opera. *Biochim Biophys Acta BBA - Biomembr* 1666: 105–117

Sundaram A, Plumb R, Appathurai S & Mariappan M (2017) The Sec61 translocon limits IRE1α signaling during the unfolded protein response. *eLife* 6: e27187

Tate CG (2010) Practical Considerations of Membrane Protein Instability during Purification and Crystallisation. In *Heterologous Expression of Membrane Proteins: Methods and Protocols*, Mus-Veteau I (ed) pp 187–203. Totowa, NJ: Humana Press

Tripathi A, Mandon EC, Gilmore R & Rapoport TA (2017) Two alternative binding mechanisms connect the protein translocation Sec71-Sec72 complex with heat shock proteins. *J Biol Chem* 292: 8007–8018

Voorhees RM, Fernández IS, Scheres SHW & Hegde RS (2014) Structure of the mammalian ribosome-Sec61 complex to 3.4 Å resolution. *Cell* 157: 1632–1643

Voorhees RM & Hegde RS (2015) Structures of the scanning and engaged states of the mammalian SRP-ribosome complex. *eLife* 4: e07975

Voorhees RM & Hegde RS (2016) Structure of the Sec61 channel opened by a signal sequence. *Science* 351: 88–91

Walter P & Blobel G (1981) Translocation of proteins across the endoplasmic

reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. *J Cell Biol* 91: 557–561

Wang X & Johnsson N (2005) Protein kinase CK2 phosphorylates Sec63p to stimulate the assembly of the endoplasmic reticulum protein translocation apparatus. *J Cell Sci* 118: 723–732

Waters MG, Evans EA & Blobel G (1988) Prepro-alpha-factor has a cleavable signal sequence. *J Biol Chem* 263: 6209–6214

Weng T-H, Steinchen W, Beatrix B, Berninghausen O, Becker T, Bange G, Cheng J & Beckmann R (2021) Architecture of the active post-translational Sec translocon. *EMBO J* 40: e105643

Wild K, Halic M, Sinning I & Beckmann R (2004) SRP meets the ribosome. *Nat Struct Mol Biol* 11: 1049–1053

Wittke S, Dünnwald M & Johnsson N (2000) Sec62p, A Component of the Endoplasmic Reticulum Protein Translocation Machinery, Contains Multiple Binding Sites for the Sec-Complex. *Mol Biol Cell* 11: 3859–3871

Wu X, Cabanos C & Rapoport TA (2019) Structure of the post-translational protein translocation machinery of the ER membrane. *Nature* 566: 136–139

Young BP, Craven RA, Reid PJ, Willer M & Stirling CJ (2001) Sec63p and Kar2p are required for the translocation of SRP-dependent precursors into the yeast endoplasmic reticulum in vivo. *EMBO J* 20: 262–271

Zacharias DA, Violin JD, Newton AC & Tsien RY (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. *Science* 296: 913–916

Zhao M, Wu S, Zhou Q, Vivona S, Cipriano DJ, Cheng Y & Brunger AT (2015) Mechanistic insights into the recycling machine of the SNARE complex. *Nature* 518: 61–67

Zheng SQ, Palovcak E, Armache J-P, Verba KA, Cheng Y & Agard DA (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* 14: 331–332

Zimmer J, Nam Y & Rapoport TA (2008) Structure of a complex of the ATPase SecA

and the protein-translocation channel. *Nature* 455: 936–943

Zivanov J, Nakane T, Forsberg B, Kimanius D, Hagen WJH, Lindahl E & Scheres SHW (2018) RELION-3: new tools for automated high-resolution cryo-EM structure determination. *bioRxiv*: 421123

Acknowledgements

Finally, it is all done! Now it is time to do the thing:

First, I would like to thank Prof. Dr. Dr. Roland Beckmann, for having me in the lab and funding me for my PhD. To Prof. Dr. Daniel Wilson, Prof. Dr. Klaus Förstemann, Prof. Dr. Johannes Stigler, Prof. Dr. Julian Stingele and Dr. Dietmar Martin, for being on my examination board. To my graduate program, IMPRS-LS, for organizing many seminars, workshops and retreats. To Prof. Dr. Daniel Wilson, Prof. Dr. Irmgard Sinning and Dr. Birgitta Beatrix, for being in my thesis advisory committee and giving me advice on my projects. To Dr. Thomas Becker, without whom all this may not become true, for advising me from the begining, pushing me forward, and helping me to perfect my manuscript and thesis. To Dr. Jingdong Cheng, for his constant teaching on processing data and, of course, finishing the paper together. To the EM team, Dr. Otto Berninghausen, and the Ice Queens, Sussane Rieder and Charlotte Ungewickell, for freezing samples, collecting data and keeping the cryo-EMs working day and night without problems. To Heidemarie Sieber, Joanna Musial and Andrea Gilmozzi, without whom the lab will be in a big big trouble, for organizing the lab and providing technical support. To Aleksandra Sarman-Grilc, for all the administration work. To the IT team, Dr. André Heuer, Dr. Lukas Kater, Dr. Michael Ameismeier and Katharina Best, for maintaining the processing pipeline and the never-enough storage for cryo-EM data. To Robert Buschauer, for helping me freeze my sample so many times. To Dr. Ting Su, for guiding me when I started processing. To Hanna Kratzat, for helping me deal with the evil phone contract and so many other stuffs that needed to speak German. To Dr. Jennifer Wells, Vivekanandan Shanmuganathan, Timur Mackens-Kiani, Dr. Alexandra Knorr and Dr. Katharina Braunger, for their accompany and discussion, no matter

scientific or not, which made the life in the lab more colorful. To my other friends in Munich, for hanging around with me sometimes so that I was not always in the lab. To my dear, Wan-Yu, for your love and accompany. Lastly, To my family, my parents and my sister, for their unconditional support and always making me feel so warm when I went back home.