

# Co-translational protein targeting and insertion by SecA

Thesis by

Shuai Wang

In Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy

The logo for the California Institute of Technology (Caltech), featuring the word "Caltech" in a bold, orange, sans-serif font.

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2020

(Defended in September 27, 2019)

© 2019

Shuai Wang  
ORCID: 0000-0002-8920-969X

All rights reserved

## Acknowledgement

The graduation of Ph.D marks the end of the systematic education in the broader biological sciences, yet the dawn of a new path towards independent research. I could not have made it to this point without all the people I came across in the past, during when they shaped my personality as a human, my philosophy of doing research, and my connection to the academic world.

I am grateful of my family for their support over the years. I inherited the wisdom of my father Jing Wang and the personality of my mother Xiaolin Mei, who together set the tone of my character. I thank them for providing a supportive environment where I learned to be independent and make my own decisions from early on. I also thank my wife Chen Gui, who has been one of my best friends since I was eleven-years old. I thank her for all her support, company and encouragement over the years, and her love and understanding in our committed relationship. I am grateful that she gave me a comfortable home to enjoy life.

I thank all the great minds who mentored me in schools. My middle school teacher Wei Qi helped to shape my personality. My high school teacher Zegui Li taught me to handle stress, and plan wisely, and he is the first person who said I would be a scholar in the future. My undergraduate advisor Xiangdong Chen introduced me to research and supported me throughout my two rounds of application for graduate school in 2011 and 2013. I thank my research advisor Zhucheng Chen at Tsinghua University who is my first role model in science. He trained me in biochemistry, taught me to distinguish great science from poor research, set an example of a good scientist: think critically, work hard, and fear no failure or authority. I am grateful of my graduate school advisor Shu-ou Shan who cultivated me into a mature researcher. I acquired her traits of being knowledgeable deeply in one field as well as broadly in many related fields. I learned from her to foresee the outcome of experimental designs, and to envision the future direction based on past and current knowledge of a field. Her mentorship gave me not only solid training and guidance, but also ample freedom to pursue my own interest. I thank her for all her advice, her help, and her lab culture. I thank my committee member Bil Clemons, Alex Varshavsky, and David Chan for their advice and helpful discussions over the years. I want to specially thank Bil for connecting me to Justin Chartron, his former student and now a professor at UC-Riverside, to

whom I thank for providing tremendous instructions in the ribosome profiling project. I also thank André Hoelz for illuminating discussions about solving structures and scientific careers in the lunch room over the years. I thank our collaborator Nenad Ban for his supervision in the cryoEM project.

I also thank all the supporting colleagues and friends. I am fortunate to have overlapped with these Shan lab members: David Akopian, George Liang, Sowmya Chandrasekar, Meera Rao, Mike Rome, Aileen Ariosa, Camille McAvoy, Jae ho Lee, Un Seng Chio, Yu-Hsien Hwang Fu, Hyunju Cho, Chien-I Yang, Alex Siegel, Hao Hsuan Hsieh, Arpit Gupta, and Yumeng Liu, and Zikun Zhu, many of who became my source for a second opinion as well as my friends to share the life of graduate school. I thank them for all the advice and company they gave to me. I thank Ahmad Jomaa for fruitful collaboration in the cryoEM project, during which he solved the structure of SecA-RNC at high resolution after extensive data collection and classification. I also thank Yandong Zhang, Haoqing Wang, Peng He, Yapeng Su, Yicheng Luo, Xun Wang, Chengcheng Fan, Kyuhyun Lee for not only being my friends, but also offering help and advice in their fields of expertises. I thank Matthew Zimmer and Tom Miller for helpful discussions in MD simulation. I am also grateful of the experts at Caltech facility centers: Igor Antoshechkin for help in next-generation sequencing, Alasdair McDowall for training me in electron microscopy, Mike Sweredoski for suggesting I learn python, and Annie Moradian for helpful discussions.

## Abstract

Co-translational protein targeting is a conserved process for the biogenesis of membrane proteins. This pathway was generally thought to depend on signal-recognition particle (SRP) for recognition of nascent protein and delivery to the membrane. Recently, SecA was found to also bind ribosomes near the nascent polypeptide exit tunnel, but the function of SecA's ribosome interaction remains unclear.

A combination of *in vitro* reconstitution and *in vivo* targeting assays show that SecA is necessary and sufficient to direct the targeting and translocation of RodZ to the bacterial plasma membrane in an obligatorily co-translational mechanism. The N-terminal extension preceding the transmembrane domain and periplasmic domain sequences immediately downstream of the transmembrane domain of RodZ provide distinguishing features that allow RodZ to engage SecA instead of the SRP machinery. Biochemical and cryoEM analyses further show that the N-terminal amphipathic helix on SecA and the ribosomal protein uL23 together form a composite binding site for the transmembrane domain (TMD) on the nascent chain. This interaction positions additional sites on the ribosome and SecA for recognition of the charged residues on both sides of the TMD, explaining the substrate specificity of SecA recognition. Quantitative kinetic analyses demonstrate that membrane-embedded SecYEG can associate with and remodel the SecA-bound ribosome-nascent chain complex, which together with elongation of the nascent polypeptide facilitates handover of the translating ribosome to the translocase.

## Published Content and Contributions

[1] Wang, S., Yang, C.-I., and Shan, S.-O. 2017. SecA mediates cotranslational targeting and translocation of an inner membrane protein. *Journal of Cell Biology*. 216(11):3639-3653. doi: 10.1083/jcb.201704036.

S. W. designed and did most of the experiments. S. W. and S.-o. S. wrote the paper.

[2] Wang, S.\*, Jomaa, A.\*, Jaskolowski, M., Yang, C.-I., Ban, N., and Shan, S.-O. 2019. Molecular mechanism of co-translational membrane protein recognition and targeting by SecA. *Nat. Struct. Mol. Biol. In press* (\* equal contribution)

S.W. and S.-o. S. conceived the project. S.W. performed most of the biochemical experiments, analyzed the data, and wrote the manuscript with S.-o. S., A. J. and N. B.

## Table of Contents

Acknowledgement .....	iii
Abstract .....	v
Published Content and Contributions .....	vi
Table of Contents .....	vii
List of Illustrations and Tables .....	viii
Nomenclature .....	ix
Chapter 1 SecA mediates co-translational protein targeting and translocation .....	1
1.1 Abstract .....	1
1.2 Introduction .....	1
1.3 Results .....	5
1.4 Discussion .....	22
1.5 Methods .....	25
1.6 Supplementary Figures and Tables .....	31
Chapter 2 Molecular mechanism of co-translational membrane protein recognition and targeting by SecA .....	36
2.1 Abstract .....	36
2.2 Introduction .....	36
2.3 Results .....	39
2.4 Discussion .....	53
2.5 Methods .....	57
2.6 Supplementary Figures and Legends .....	69
References .....	88

## List of Illustrations and Tables

<b>Figures</b>	<b>Page</b>
1.1. Fluorescence measurements of SecA-RNC interactions. ....	6
1.2. Contribution of the ribosome to RNC-SecA affinity. ....	8
1.3. Defining the sequence elements of RodZ for SecA recognition. ....	11
1.4. RodZ is co-translationally targeted and translocated <i>in vivo</i> . ....	14
1.5. Reconstitution of RodZ targeting and translocation in <i>vitro</i> . ....	17
1.6. The NTE and early periplasmic region of RodZ together dictate the selection of a membrane protein into the SecA versus SRP pathway. ....	20
1.7. Diverse targeting pathways deliver nascent proteins to the SecYEG translocon at the inner-membrane. ....	21
S1.1. Fluorescence measurements of SecA-RNC <sub>RodZ</sub> affinity. ....	31
S1.2. Translocation of NTS-TrxA fusion proteins and their dependence on SecA or SRP. ....	32
S1.3. Replicates for reconstitution of RodZ targeting and translocation in <i>vitro</i> . ....	33
S1.4. Equilibrium titrations for RNC binding to SecA or SRP. ....	34
2.1. Thio-specific crosslinking to map SecA surface residues that contact hydrophobic sequences of nascent protein on the ribosome. ....	40
2.2. Cryo-EM structure of the SecA•RNC <sub>RodZ</sub> . ....	42
2.3. Snapshots of the SecA•RNC <sub>RodZ</sub> model. ....	44
2.4. Validation of the interaction sites between RNC <sub>RodZ</sub> and SecA. ....	47
2.5. Effects of nascent chain length on SecYEG binding to and SecA dissociation from RNC <sub>RodZ</sub> . ....	48
2.6. Measurements to distinguish different mechanisms of RNC <sub>RodZ</sub> delivery to SecYEG by SecA. ....	49
2.7 Model of SecA-mediated co-translational recognition and targeting of membrane proteins. ....	52
S2.1. Controls and additional data to map the interaction of SecA with hydrophobic sequences on nascent protein in the co- versus post-translational mode. ....	70
S2.2. Additional data to map the interaction surface of SecA with nascent chains on RNC and with post-translational substrates. ....	73
S2.3. Image classification and refinement of the structure of the RNC <sub>RodZ</sub> •SecA complex. ....	76
S2.4. Local resolution and validation of the cryoEM maps. ....	77
S2.5. Overview of the RodZ TMD binding pocket and the ribosome tunnel region. ....	79
S2.6. Characterization of the samples for fluorescence measurements of RNC <sub>RodZ</sub> transfer from SecA to SecYEG. ....	82
S2.7. Supporting information for the kinetic simulations, and additional data for chase experiments of the RNC <sub>RodZ</sub> •SecA complex. ....	84
S2.8. Comparison of the structure from this work with previous structures. ....	86
<b>Tables</b>	<b>Page</b>
S1.1. Constructs used in RNC preparation and <i>in vivo</i> translocation assays. ....	35
S2.1. Data collection, structure model refinement, and validation statistics. ....	87

## Nomenclature

BDP. BODIPY-FL

BME. b-mercapto ethanol

BMH. bismaleimidohexane

BMOE. bismaleimidoethane

Cm. 7-hydroxycoumaryl ethylglycine

cryoEM. Cryo-electron microscopy

DDM. N-Dodecyl- $\beta$ -D-Maltopyranoside

DTT. Dithiothreitol

EDTA. Ethylenediaminetetraacetic acid

EMC. ER membrane protein complex

EMD. Electron microscopy databank

ER. endoplasmic reticulum

FRET. Förster resonance energy transfer

HSD. helical scaffold domain

HWD. helical wing domain

IPTG. Isopropyl- $\beta$ -D-thiogalactoside

IVT. *in vitro* translation

MBD. Maltose-binding domain

NBD-I. nucleotide-binding domain-I

NBD-II. nucleotide-binding domain-II

Nd. nanodisc

NTE. N-terminal extension

NTS. N-terminal targeting sequence

PDB. Protein data bank

PPXD. pre-protein crosslinking domain  
RNA. Ribonucleic acid  
RNC. Ribosome nascent chain complex  
RPB. ribosome-associated protein biogenesis factor  
SND. SRP-independent targeting  
SR. Signal recognition particle receptor  
SRP. Signal recognition particle  
SS. Signal sequence  
SUMO. Small Ubiquitin-like Modifier  
TCEP. Tris(2-carboxyethyl) phosphine  
TF. trigger factor  
THF. two-helix finger loop  
tRNA. Transfer-ribonucleic acid  
TrxA. Thioredoxin  
U-IMV. urea-washed inner-membrane vesicles  
WT. Wild type

## Chapter 1 SecA mediates co-translational protein targeting and translocation

A version of this chapter was first published as:

*SecA mediates cotranslational targeting and translocation of an inner membrane protein.*  
Shuai Wang, Chien-I Yang, and Shu-ou Shan. *J Cell Biol.* **2017**, 216(11):3639-3653.

### 1.1 Abstract

Proper localization of proteins to cellular membranes is essential for the folding and function of nascent membrane and secretory proteins. Protein targeting to bacterial plasma membrane was generally thought to occur via two major pathways: co-translational targeting mediated by the signal recognition particle (SRP) and post-translational targeting mediated by SecA and SecB. Recently, SecA was found to also bind ribosomes near the nascent polypeptide exit tunnel, but the function of SecA's ribosome interaction remains unclear. Here, we show that SecA co-translationally recognizes the nascent chain of an inner membrane protein, RodZ, with high affinity and specificity. A combination of *in vitro* reconstitution and *in vivo* targeting assays show that SecA is necessary and sufficient to direct the targeting and translocation of RodZ to the bacterial plasma membrane in an obligatorily co-translational mechanism. The N-terminal extension preceding the transmembrane domain and periplasmic domain sequences immediately downstream of the transmembrane domain of RodZ provide distinguishing features that allow RodZ to engage SecA instead of the SRP machinery. These findings suggest a new route for the targeting of inner membrane proteins in bacteria, and highlight the diversity of protein targeting pathways that may enable an organism to accommodate nascent proteins harboring divergent targeting signals.

### 1.2 Introduction

Roughly 30% of the genome encodes membrane proteins, which are anchored to cellular membranes via at least one transmembrane domain (TMD) and play diverse physiological roles such as signaling, cell shape maintenance, and cell motility. To attain their proper structure and function, newly synthesized membrane proteins must engage dedicated protein targeting

pathways, via which they are delivered to the correct membrane destination in the cell (Zhang and Shan, 2014). Mislocalization of membrane proteins not only deprives cells of functional proteins, but also disrupts cellular protein homeostasis due to improper exposure of the hydrophobic TMDs in the cytosol that could lead to aggregation and misfolding. This demands that the targeting processes for membrane proteins act with high efficiency to minimize the exposure of TMDs in the cytosol.

The co-translational targeting of proteins by the signal recognition particle (SRP) is the most well understood pathway for targeted delivery of integral membrane proteins. SRP recognizes hydrophobic signal sequences or TMDs near the N-terminus of nascent proteins as soon as they emerge from the ribosome exit tunnel (Walter et al., 1981; Schaffitzel et al., 2006; Zhang and Shan, 2014). The TMD on the nascent protein is shielded from the cytosol by the M-domain of SRP. Through the interaction between SRP and the SRP receptor (SR, termed FtsY in bacteria), the nascent protein is delivered to the SecYEG (or Sec61p) protein translocation machinery on the bacterial inner membrane (or the eukaryotic endoplasmic reticulum) (Zhang et al., 2010). SRP-dependent targeting is complete before ~130 amino acids of the nascent polypeptide C-terminal to the signal sequence or TMD is translated (Siegel and Walter, 1988; Ariosa et al., 2014), and releasing nascent proteins from the ribosome abolishes the targeting of SRP-dependent substrates (Kuruma et al., 2008). In bacteria, SRP is generally thought to mediate the targeted delivery of the majority of inner membrane proteins and a number of periplasmic secretory proteins that contain highly hydrophobic signal sequences (Luirink and Sinning, 2004; Schibich et al., 2016).

A second major protein-targeting pathway in bacteria uses SecA, with the participation of the chaperone SecB in some cases. The SecB/A pathway targets the majority of secretory and outer-membrane proteins via a post-translational mechanism (Hartl et al., 1990). Substrates entering this pathway contain signal sequences that are less hydrophobic than those engage SRP/SR (Neumann-Haefelin et al., 2000). These signal sequences are recognized by the PPXD domain of SecA, which couples its ATPase cycle to the translocation of substrate protein across SecYEG (Bauer et al., 2014). The post-translational chaperone, SecB, assists in maintaining preproteins in the unfolded, translocation competent state while also delivering these proteins to SecA bound at the inner membrane (Weiss et al., 1988). The post-translational nature of the SecB/A pathway is supported by the following observations: (1) classic SecB/A-dependent

substrate proteins, such as OmpA and PhoA, can be efficiently inserted into the membrane without coupling the targeting reaction to protein synthesis (Hartl et al., 1990; Gouridis et al., 2009), indicating that a co-translational mode of targeting is not mechanistically obligatory for these substrates; (2) the SecA ATPase cycle and its interaction with SecYEG are enhanced by the mature domain of the nascent protein C-terminal to the signal sequence, suggesting that a substantial length of the nascent protein needs to be exposed before they are targeted by the SecB/A pathway (Gouridis et al., 2009); (3) C-terminal fusion to fast-folding proteins, such as thioredoxin, severely block the translocation of SecA-dependent substrate proteins (Huber et al., 2005a), suggesting that targeting and translocation was not finished before the complete synthesis and folding of the nascent polypeptide.

More recently, however, SecA was found to also interact with the ribosome. SecA binds the 70S bacterial ribosome with a dissociation constant ( $K_d$ ) of 0.9  $\mu$ M (Huber et al., 2011), in part via an interaction with conserved acidic residues on the L23 protein near the ribosome exit site (Singh et al., 2014). Disruption of this ribosomal contact modestly reduces the translocation efficiency of a number of secretory proteins (Huber et al., 2011). Nevertheless, a clear understanding for the role and importance of the SecA-ribosome interaction has been lacking. Although SecA has been observed to contact nascent proteins while they are still bound to the ribosome *in vitro* (Karamyshev and Johnson, 2005; Huber et al., 2016;) and *in vivo* (Randall, 1983), a co-translational requirement has not been demonstrated for the SecA-preprotein contact, nor for the targeting of these secreted proteins, raising questions as to the necessity of recruiting SecA co-translationally.

Up till now, SRP is the only known factor in bacteria that can co-translationally target inner-membrane proteins. Nevertheless, model SRP substrates are still targeted to the bacterial inner membrane, albeit more slowly, under SRP depleted conditions (Wickstrom et al., 2011; Zhang et al., 2012), suggesting the presence of alternative targeting pathways for inner-membrane proteins. In addition, SecA is required for the insertion of multiple inner-membrane proteins that contain large periplasmic domains (Wolfe et al., 1985; Gebert et al., 1988; Sääf et al., 1995; Traxler and Murphy, 1996), which implicates SecA plays a role at some stage of the biogenesis of these membrane proteins. Moreover, some inner-membrane proteins in *E. coli* depend on SecA, rather than SRP, for insertion (Kihara and Ito, 1998; Ulbrandt et al., 1997; Rawat et al., 2015). An inner-membrane protein, AcrB, showed more severe defects in

membrane insertion under SecA-depleted, than SRP-depleted conditions (Qi and Bernstein, 1999). Recently, Rawat et al explored the insertion requirements of two single-span membrane proteins, RodZ and CadC, and suggested the possibility that RodZ is inserted co-translationally by SecA (Lindner and White, 2014; Rawat et al., 2015). A common feature of both proteins is a TMD over 100 residues downstream of the N-terminus and a strict dependence on SecA, but not SRP or FtsY, for insertion. In chloroplast, cpSecA has been speculated to be an alternative targeting factor, as the cpSecA-dependent substrate protein PetA is co-translationally targeted to the thylakoid membrane (Zoschke and Barkan, 2015). These observations compel us to explore the possible role of SecA in mediating a potential co-translational targeting pathway for inner-membrane proteins.

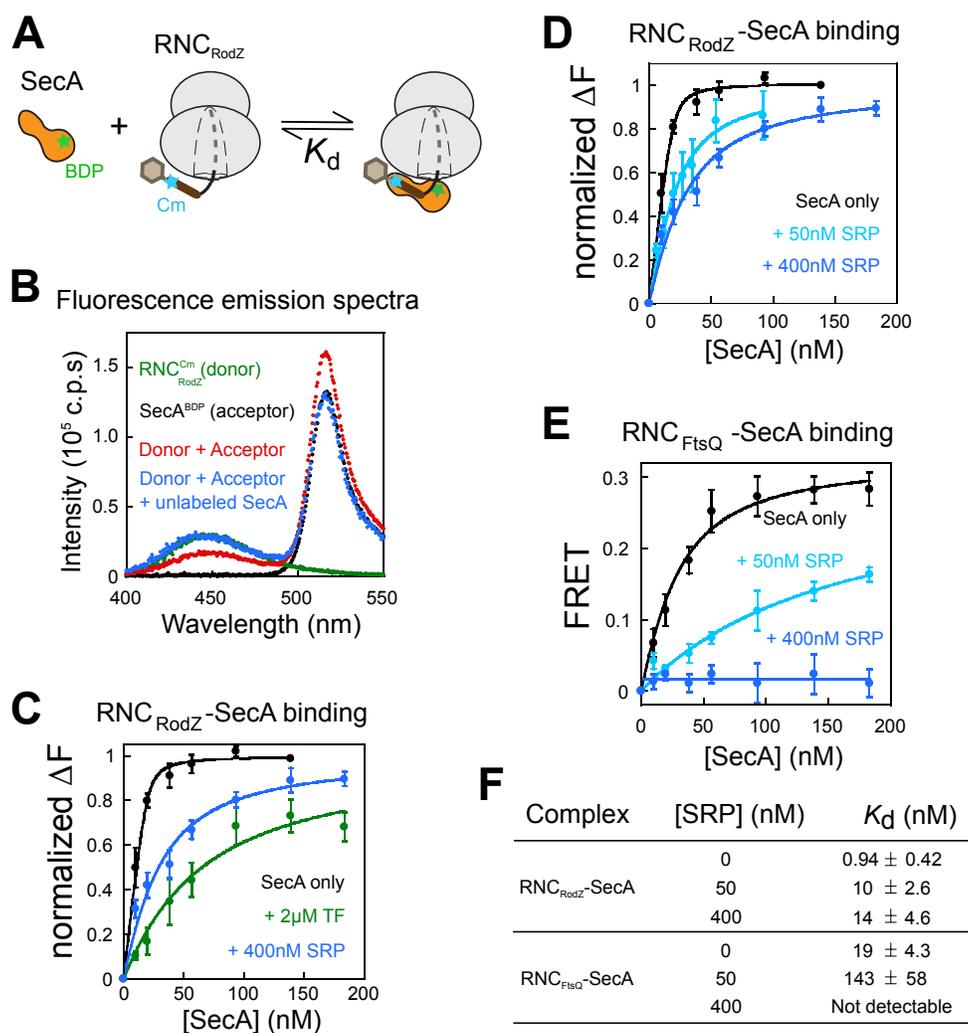
Using a combination of quantitative binding measurements, *in vitro* reconstitutions and *in vivo* targeting assays, we demonstrate here that SecA co-translationally recognizes and targets an inner-membrane protein, RodZ. The extended N-terminal element preceding the internal TMD of RodZ and periplasmic sequences immediately following the TMD enable the selection of RodZ by the SecA, rather than the SRP machinery, for membrane targeting. This work uncovers a new role of SecA and provides evidence for an SRP-independent co-translational targeting pathway for a subset of inner membrane proteins in bacteria.

### 1.3 Results

#### *SecA co-translationally interacts with the RodZ nascent chain*

As a candidate substrate that could co-translationally interact with and be targeted by SecA, we tested RodZ (Rawat et al., 2015). RodZ is a single pass type II membrane protein comprised of an N-terminal cytoplasmic domain (residues 1-111), a TMD (residues 112-132) anchored on the bacterial plasma membrane, and a C-terminal periplasmic domain (residues 133-337). The *in vivo* biogenesis of RodZ was shown to be dependent on SecA and SecYEG, but had no dependence on the bacterial SRP protein Ffh, the SRP receptor FtsY, or the post-translational chaperone SecB (Ulbrandt et al., 1997; Rawat et al., 2015). As discussed by Rawat et al, a co-translational mechanism of targeting would be beneficial for minimizing the cytosolic exposure of the RodZ-TMD and the premature folding of the RodZ periplasmic domain in the cytosol; we therefore hypothesized that SecA could be recruited to ribosome-nascent chain complexes (RNCs) bearing newly synthesized RodZ.

To detect the interaction between RNC<sub>RodZ</sub> and SecA, we used Förster resonance energy transfer (FRET). As the FRET donor, we used amber suppression technology (Saraogi et al., 2011) to incorporate a fluorescent amino acid, 7-hydroxycoumaryl ethylglycine (Cm), at residue 111 upstream of the RodZ TMD (Figure 1.1 A and Table S1.1). As the FRET acceptor, we labeled SecA at residue 12 with BODIPY-FL (BDP). The mutations and fluorescence labeling did not perturb the activity of SecA in mediating post-translational protein translocation (Figure S1 A), nor the interaction of RNC with targeting factors (Saraogi et al., 2011). For initial binding measurements, we purified RNC<sub>RodZ</sub> displaying the N-terminal 180 amino acids of RodZ; the RodZ nascent chain is followed by a 34-residue SecM stalling sequence (Nakatogawa and Ito, 2002), which occupies most of the ribosome exit tunnel (Zhang et al., 2015). When purified RNC<sub>RodZ</sub> was incubated with SecA<sup>BDP</sup>, we observed 44% reduction in Cm fluorescence and a corresponding increase in BDP fluorescence, indicating FRET (Figure 1.1 B, red). As expected from the competition between labeled and unlabeled SecA, addition of a 10-fold excess of unlabeled SecA removed the FRET signal (Figure 1.1 B, blue). This result also ruled out the environmental sensitivity of Cm as a contributor to the FRET signal, and indicated a reversible binding equilibrium between RNC<sub>RodZ</sub> and SecA.



**Figure 1.1. Fluorescence measurements of SecA-RNC interactions.**

(A) Scheme of the FRET assay to detect the interaction of SecA with the RodZ nascent chain on the ribosome.

(B) Fluorescence emission spectra for indicated samples. Where indicated, reactions contained 20 nM

RNC<sub>RodZ</sub><sup>Cm</sup>, 40 nM SecA<sup>BDP</sup>, and 400 nM unlabeled SecA. (C) Representative equilibrium titrations to measure

the  $K_d$  values of the SecA-RNC<sub>RodZ</sub> complex. Reactions contained 20 nM RNC<sub>RodZ</sub><sup>Cm</sup> without (black) or with

SRP (blue) or TF (green) present. The titration curves before normalization are shown in Figure S1.1 B. Lines

are fits of the data to Eq 3. (D-E) Representative equilibrium titrations to measure the  $K_d$  values of the

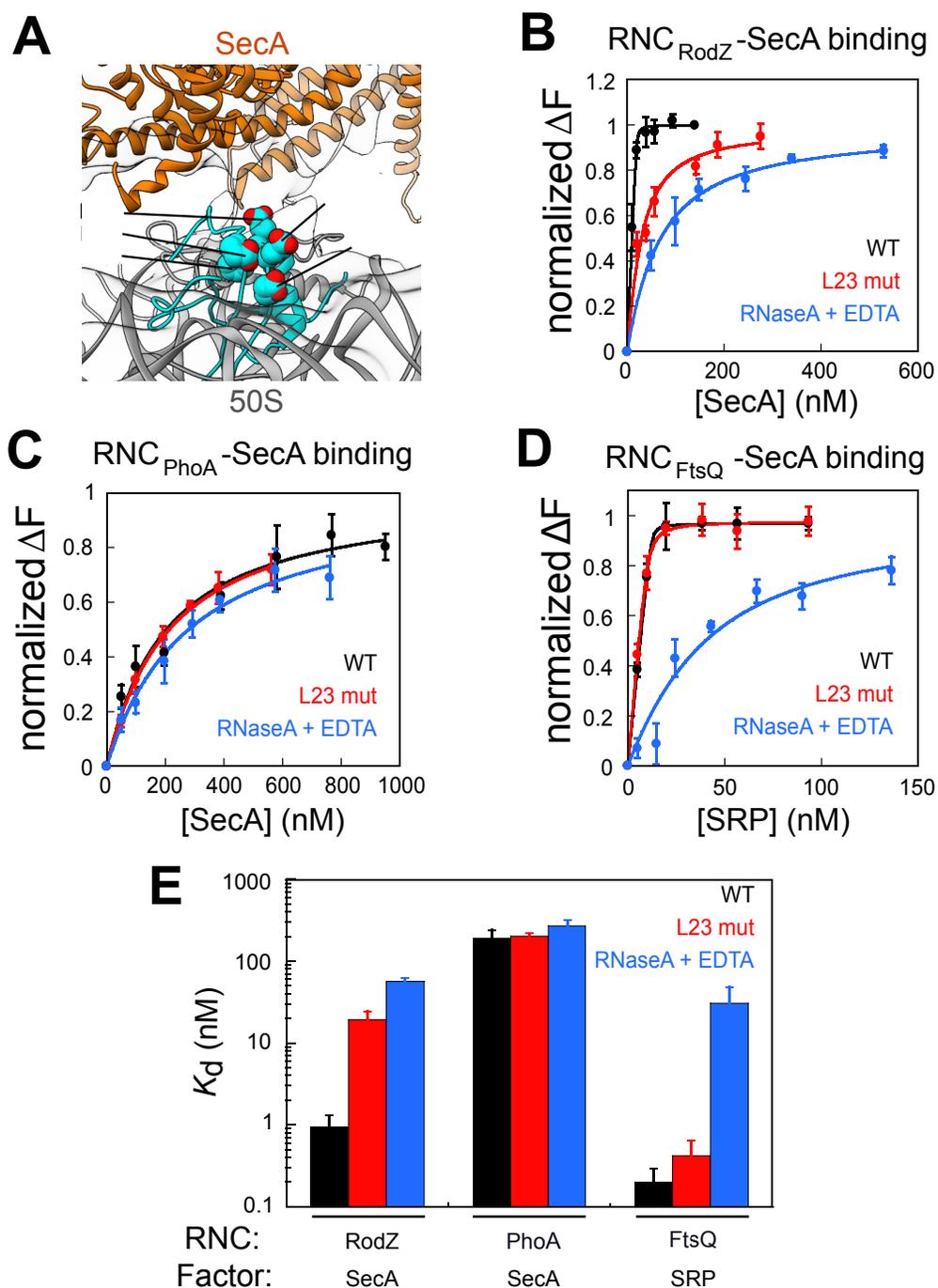
SecA•RNC<sub>RodZ</sub> (D) and SecA•RNC<sub>FtsQ</sub>. (E) complexes at increasing concentrations of SRP. Lines are fits

of the data to Eq 3. (F) Summary of the  $K_d$  values of SecA-RNC complexes, obtained from the data in parts C-E

and their replicates. Values represent mean  $\pm$  S.D., with  $n = 3$ .

Equilibrium titrations based on the FRET assay showed that SecA binds RNC<sub>RodZ</sub> tightly, with a  $K_d$  value of  $\sim 1$  nM (Figure 1 C); this affinity is  $\sim 900$ -fold higher than that of SecA for empty ribosomes (Huber et al., 2011), suggesting additional interactions of SecA with the RodZ nascent chain. As other ribosome-associated protein biogenesis factors, such as SRP and trigger factor (TF), could compete for binding to the ribosome and RodZ nascent chain under physiological conditions (Ariosa et al., 2014; Gamerdinger et al., 2015), we further tested if the SecA-RNC<sub>RodZ</sub> interaction survives the presence of these factors. Equilibrium titrations in the presence of near-physiological concentrations of SRP (400 nM) or TF (2  $\mu$ M) showed that the SecA-RNC<sub>RodZ</sub> interaction was weakened by these factors but remained strong, with  $K_d$  values of  $\sim 19$  nM and  $\sim 55$  nM, respectively (Figure 1.1 C and Figure S1.1 B). In addition, raising the SRP concentration beyond 50 nM did not significantly weaken the binding between SecA and RNC<sub>RodZ</sub> (Figure 1.1 D and F; and Figure S1.1 C). As a negative control, we used RNC<sub>FtsQ</sub>, a well-characterized SRP substrate (Estrozi et al., 2011). Although the interaction of SecA with RNC<sub>FtsQ</sub> could be detected, this interaction was  $\sim 20$ -fold weaker than that with RNC<sub>RodZ</sub> and did not withstand the competition from physiological concentration of SRP (Figure 1.1 E and F). These data strongly suggest that the nascent chain of RodZ can efficiently and specifically recruit SecA during translation.

We next asked whether the ribosome contributes to the recruitment of SecA onto nascent RodZ. To this end, we disassembled the ribosomes in purified RNCs by RNase A and EDTA treatment (Figure S1.1 D). As an independent and more specific means to perturb the SecA-ribosome interaction, we mutated an acidic patch (F51A/E52A/E54A/E56A/E89A) on the ribosomal protein L23 that contacts the N-terminus of SecA (Figure 1.2 A; (Huber et al., 2011; Singh et al., 2014)). Both perturbations significantly weakened the interaction of SecA with the RodZ nascent chain. The L23 mutations weakened the binding affinity of SecA for RNC<sub>RodZ</sub> over 20-fold (Figure 1.2 B and E, red). The binding defect was larger,  $\sim 60$ -fold, with EDTA- and RNase A-treated RNC<sub>RodZ</sub> (Figure 1.2 B and E, blue). As a negative control, we tested RNC bearing the nascent chain of PhoA, a post-translational SecA substrate (Gouridis et al., 2009). Although an interaction between SecA and RNC<sub>PhoA</sub> could be detected, neither the L23



**Figure 1.2. Contribution of the ribosome to RNC-SecA affinity.**

(A) Structure of SecA bound to the 70S ribosome (EMD 2565). The crystal structures of SecA (orange, PDB 1m6n) and ribosome (grey, PDB 2aw4) were docked into the EM density. Acidic residues on L23 (cyan) that contact SecA are in *spacefill*. (B-D) Equilibrium titrations to measure the affinity of SecA for wildtype and modified RNC<sub>RodZ</sub> (B) and RNC<sub>PhoA</sub> (C), as well as the affinity of SRP for RNC<sub>FtsQ</sub> (D). Lines are fits of the

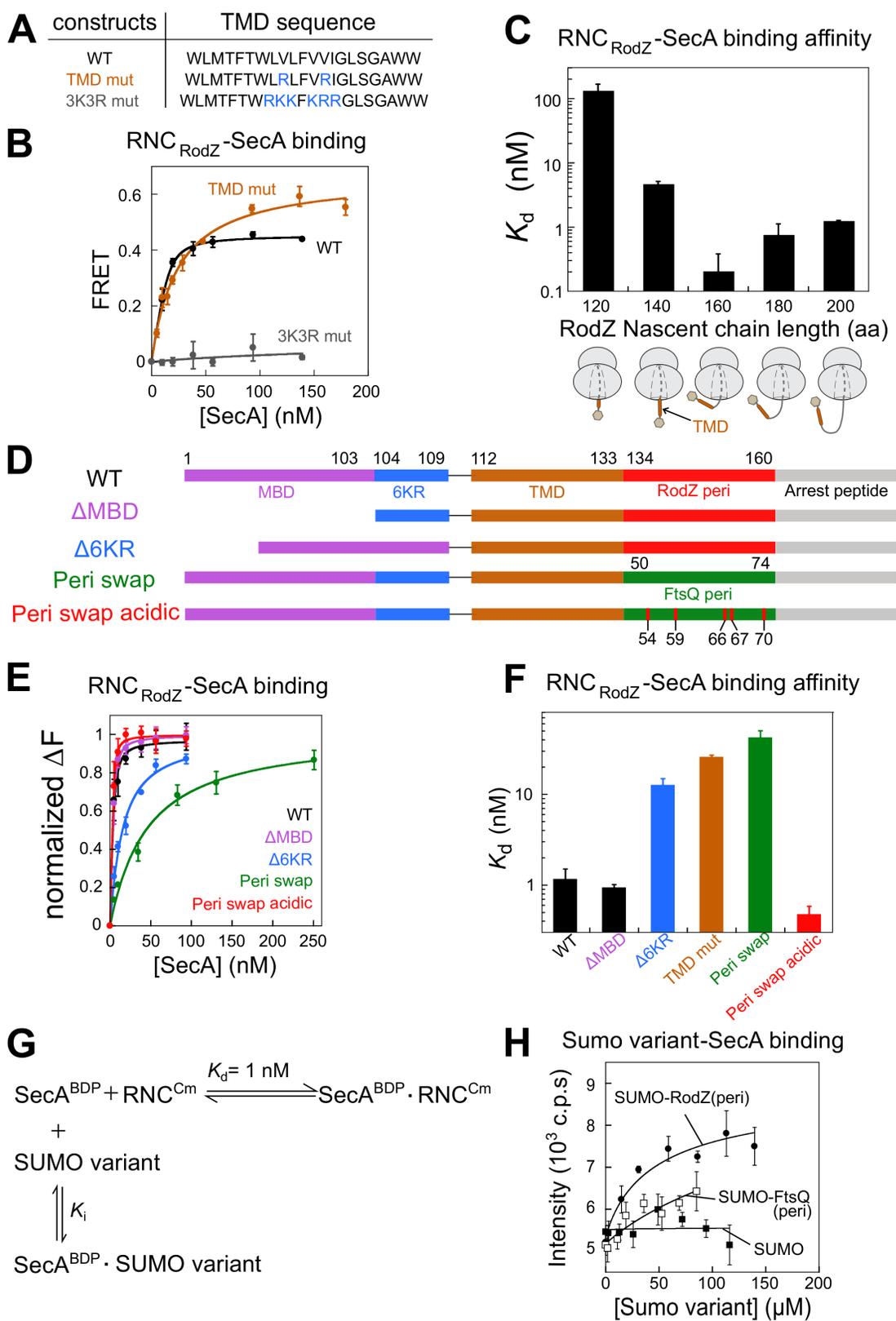
data to Eq 3. (E) Summary of the  $K_d$  values, derived from the data in parts B-D. Values represent mean  $\pm$  S.D., with  $n = 3$ .

mutations nor the EDTA-RNase A treatment affected this interaction (Figure 1.2 C and E), indicating that SecA binds the PhoA nascent chain independently of the ribosome. As a positive control, the interaction of SRP with its substrate, RNC<sub>FtsQ</sub>, was also disrupted by the RNase A and EDTA treatment (Figure 1.2 D and E, blue). However, SRP-RNC<sub>FtsQ</sub> binding was unaffected by the L23 mutations (Figure 1.2 D and E, red), indicating that this acidic patch on L23 provides a specific docking site for SecA. These results showed that efficient recruitment of SecA to the RodZ nascent chain requires specific contacts of SecA with the ribosomal protein L23.

### ***SecA recognizes multiple sequence elements on the RodZ nascent chain.***

To probe the sequence elements on the RodZ nascent chain required for SecA recognition, we first tested the role of the RodZ TMD (Figure 1.3 A). Introduction of two arginines weakened the SecA-RodZ interaction ~26-fold, raising the  $K_d$  value to ~26 nM (Figure 1.3 B, TMD mut). Introduction of six basic residues into the RodZ-TMD abolished detectable interaction of SecA with the nascent chain (Figure 1.3 B, 3K3R mut). These results suggest that the hydrophobic TMD on RodZ provides an important recognition element for SecA.

If SecA recognizes the TMD on RNC<sub>RodZ</sub>, then the SecA-RNC<sub>RodZ</sub> interaction would be sensitive to the length of the nascent polypeptide, as complete exposure of the TMD on the ribosome would require at least 133 amino acids of the RodZ nascent chain to be displayed in the stalled RNC. We therefore systematically varied the length of the RodZ nascent chain (length does not count SecM arrest sequence). As expected, SecA binding was barely detectable when the RodZ nascent chain is 120 amino acids, at which length only a portion of the TMD is available (Figure 1.3 C and Figure S1.1 E inset). Significantly stronger SecA binding was observed at longer nascent chain lengths, with the tightest binding observed when the RodZ



### Figure 1.3. Defining the sequence elements of RodZ for SecA recognition.

(A) Sequences of TMD in wildtype RodZ and RodZ TMD mutants. (B) Equilibrium titrations to measure the affinity of SecA for RNC<sub>RodZ</sub> bearing wildtype and mutant TMD sequences. The data were fit to Eq 2 and gave  $K_d$  values of  $0.94 \pm 0.42$  and  $25.9 \pm 1.1$  nM for WT and TMD mut, respectively. (C) Summary of the  $K_d$  values at indicated lengths of the RodZ nascent chain (sequences in Table S1.1), obtained from the data in Figure S1.1 E and their replicates. Schemes for RNC<sub>RodZ</sub> at each chain length are shown below with ribosome in *grey*, RodZ TMD in *brown*, and sequences upstream of TMD depicted as hexagons. (D) Scheme of sequence elements in wildtype and mutant RodZ nascent chain used for the RNC-SecA binding measurements in parts E and F. MBD (*purple*) denotes the MreB-binding domain (residues 1-103), 6KR (*blue*) denotes the <sup>104</sup>KKRKRR<sup>109</sup> sequence, the RodZ TMD is in *brown*, RodZ peri (*red*) and FtsQ peri (*green*) denote the early periplasmic regions of RodZ (residues 134-160) and FtsQ (residues 50-74), respectively. All the mutant constructs are derived from RodZ160 in Fig 3 C. See Table S1.1 for detailed sequences. (E) Equilibrium titrations to measure the binding of SecA to RNCs bearing the wildtype and mutant RodZ nascent chain depicted in part D. (F) Summary of the  $K_d$  values for RNCs bearing wildtype and mutant RodZ nascent chain, obtained from the data in parts B and E. (G) Scheme for the competition assay to measure the binding of SUMO fusion proteins to SecA. BDP-labeled SecA was allowed to form a complex with RNC<sup>Cm</sup>. This binding equilibrium is perturbed if the inhibitor binds SecA<sup>BDP</sup> and traps it into a SecA•SUMO-variant, generating free RNC<sup>Cm</sup> and resulting in loss of FRET (i.e., increase of Cm fluorescence). (H) Competition reactions to measure the binding of SUMO and SUMO-variants to SecA. SUMO, SMT3 (residues 1-101). SUMO-RodZ(peri), SMT3 fused to the N-terminus of RodZ periplasmic region (residues 134-160). SUMO-FtsQ(peri), SMT3 fused to the N-terminus of FtsQ periplasmic region (residues 50-74). The data with SUMO-RodZ(peri) were fit to Eq 8, and gave a  $K_d$  value of  $1.2 \pm 0.7$   $\mu$ M. In contrast, SUMO and SUMO-FtsQ(peri) did not give robust competition. Values represent mean  $\pm$  S.D., with  $n = 2-3$ .

nascent chain is 160 amino acids (Figure 1.3 C). Collectively, these data strongly suggest that SecA recognizes the TMD of the RodZ nascent chain.

The RodZ TMD is preceded by an extended N-terminal element (NTE), comprised of a helical MreB-binding domain (MBD; residues 1-103) followed by a consecutive sequence of six basic residues (KRRKRR; residues 104-109). Deletion of the MBD did not perturb high affinity binding between SecA and RNC<sub>RodZ</sub>, whereas deletion of the basic residues preceding the TMD weakened binding over 10-fold (Figure 1.3 D-F). These results are consistent with previous findings that positively charged residues N-terminal to the signal sequence enhance preprotein binding and translocation by SecA (Akita et al., 1990; Hikita and Mizushima, 1992; Gelis et al., 2007).

The enhancement in the RNC binding affinity of SecA when the RodZ nascent chain was lengthened from 140 to 160 amino acids suggests the possibility of additional interactions of SecA with the periplasmic sequence of RodZ following the TMD. To test this hypothesis, we replaced the sequences in the N-terminal periplasmic region of RNC<sub>RodZ160</sub> (residues 134-160) with the corresponding sequence from FtsQ (Figure 1.3 D, Peri swap). This mutation weakened the affinity of SecA for RNC<sub>RodZ160</sub> over 40-fold (Figure 1.3 E, F), indicating that the periplasmic sequence of RodZ following its TMD also contributes significantly to high affinity SecA recruitment. Intriguingly, this periplasmic region of RodZ does not belong to the “hydrophobic patch” that binds SecA described by previous studies (Gouridis et al., 2009; Chatzi et al., 2017). Instead, this periplasmic region of RodZ has unusually enriched acidic residues (net charge -4), whereas the corresponding region of FtsQ has net charge of 0 (Table S1.1). To test if the acidic residues contribute to SecA recruitment, we increased the net negative charge of Peri swap by introducing mutations (R54E/K59E/R66E/H67D/R70D; Figure 1.3 D, Peri swap acidic). The binding assay showed the extra acidic residues restored the SecA recruitment (Figure 1.3 E and F, peri swap vs. peri swap acidic), suggesting the acidic residues are critical for SecA recognition.

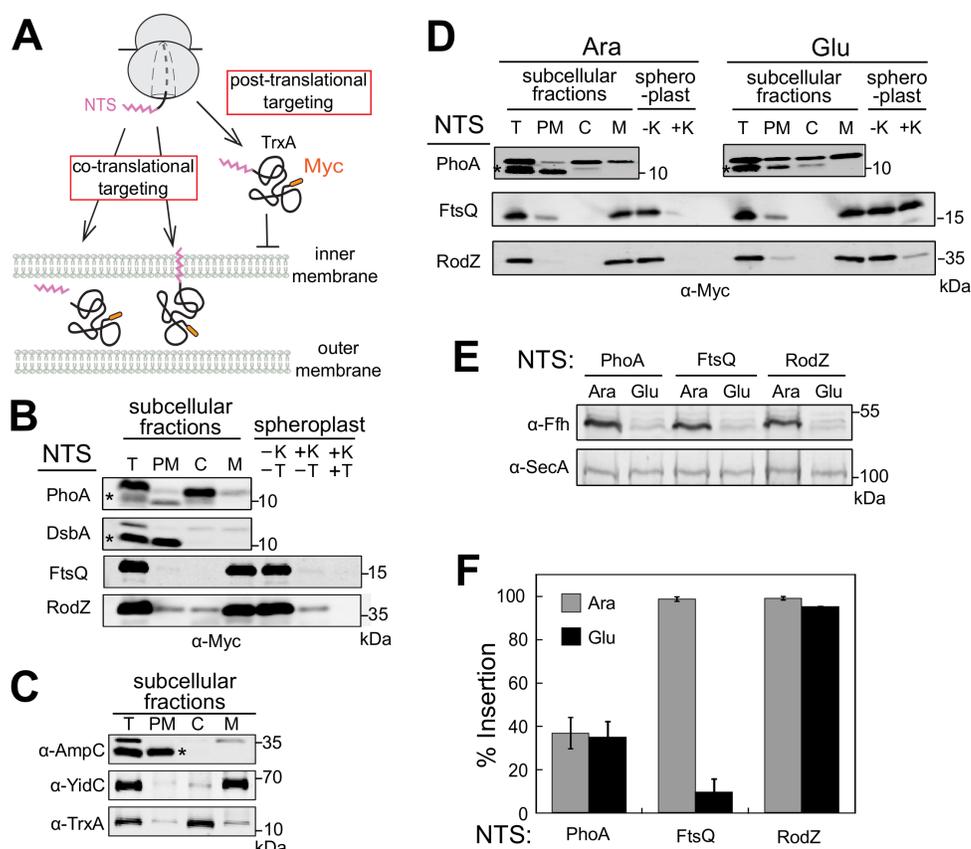
Finally, to distinguish whether the periplasmic sequence of RodZ exerts its effect directly, by interacting with SecA, or indirectly, by altering the conformation of the remainder of the RodZ nascent chain, we fused this sequence (RodZ residues 134-160) or the corresponding periplasmic sequence from FtsQ (residues 50-74) to the well-folded SUMO protein. We tested whether the resulting fusion proteins act as competitive inhibitors of the interaction between SecA and RNC<sub>RodZ</sub>. If the periplasmic sequence of RodZ directly binds SecA, it should be able to compete with RNC<sub>RodZ</sub> for SecA binding and thus restore the fluorescence signal of donor-labeled RNC due to loss of FRET between RNC<sup>Cm</sup> and SecA<sup>BDP</sup> (Figure 1.3 G). Dose-dependent, saturable restoration of the fluorescence of RNC<sup>Cm</sup> was indeed observed with SUMO-RodZ(peri) (Figure 1.3 H). In contrast, SUMO by itself did not compete with RNC<sub>RodZ</sub>, and SUMO-FtsQ(peri) provided significantly less effective competition than SUMO-RodZ(peri) (Figure 1.3 H). Quantitative analysis of this competition reaction yielded an estimated  $K_i$  value of 1.2  $\mu$ M for the interaction between SecA and SUMO-RodZ(peri).

Collectively, the results in this section show that the strong interaction of RNC<sub>RodZ</sub> with SecA are contributed by three sequence elements on the RodZ nascent chain: (i) the consecutive

positively charged residues upstream of the RodZ TMD; (ii) the hydrophobic TMD of RodZ; and (iii) the negatively charged residues in the periplasmic region of RodZ following its TMD. It is likely that each of these elements contributes a modest affinity, but together they enable high-avidity SecA recognition by providing multiple, simultaneous interactions.

### ***RodZ is co-translationally targeted and translocated in vivo independently of SRP***

The co-translational recruitment of SecA to nascent RodZ in *vitro* raised the possibility of SecA-mediated targeting and translocation of RodZ. Previous work showed that the *in vivo* insertion of RodZ is strictly SecA dependent (Rawat et al., 2015). To further test if the targeting and translocation of RodZ occurred co-translationally, we adapted a previously developed *in vivo* assay based on fusion of the N-terminal targeting sequence (NTS; Figure 1.4 A and Table S1.1) of the protein of interest to thioredoxin (Schierle et al., 2003; Huber et al., 2005b;). Thioredoxin (TrxA) folds rapidly and tightly in the cytosol, which would block its translocation across the membrane if targeting and translocation of the fusion protein occurred after the C-terminal TrxA is fully synthesized. Only if the NTS enables a co-translational mode of targeting and translocation would TrxA be successfully translocated across the inner-membrane (Figure 1.4 A). A Myc tag at the C-terminus of NTS-TrxA constructs allowed us to monitor the localization of the fusion protein in cell fractionation experiments. In addition, secretory proteins contain signal sequences that are cleaved by the signal peptidase upon successful translocation across the inner membrane (Figure 1.4A), providing an independent readout for their secretion into periplasm. For proteins that contain a TMD anchored in the bacterial inner membrane, successful translocation of TrxA across the inner-membrane exposes the Myc tag to the periplasm where it is susceptible to proteinase K digestion (Figure 1.4 A); this provides an independent readout for the proper insertion of the fusion protein at the inner membrane.



**Figure 1.4. RodZ is co-translationally targeted and translocated *in vivo*.**

(A) Scheme of the *in vivo* assay to distinguish between co- and post-translational modes of targeting and translocation based on NTS-TrxA fusions. All NTS sequences are provided in Table S1.1. (B) Left, subcellular localization of NTS-TrxA fusion proteins. Asterisks denote mature, translocated secretory proteins whose signal sequences have been cleaved. T, total; PM, periplasm; C, cytosol; M, membrane. Right, assay for translocation of the C-terminus of the NTS-TrxA fusion proteins into periplasm based on protection against proteinase K. K, proteinase K; T, triton X-100. (C) Controls for cell fractionation. Mature AmpC is secreted into the periplasm (asterisk). YidC is an inner-membrane protein. TrxA is cytoplasmic protein. (D) Effects of Ffh depletion on the targeting and translocation of NTS-TrxA fusions. *In vivo* targeting and insertion were measured and analyzed as in part B. Ffh expression is under control of the arabinose promoter. (E) Ffh is depleted in WAM121 cells grown in glucose, without significantly affecting SecA abundance. (F) Translocation efficiency of NTS-TrxA constructs, derived from the data in part D and their replicates. Values represent mean  $\pm$  S.D., with  $n = 2-3$  (biological replicates).

When the PhoA signal sequence (residues 1-21; Table S1.1) was used as the NTS, only a small fraction of the fusion protein was successfully translocated into the periplasm (Figure 1.4 B and C), consistent with previous work showing that PhoA is primarily post-translationally targeted by SecA (Schierle et al., 2003; Gouridis et al., 2009). As previously reported (Schierle et al., 2003), the more hydrophobic signal sequence from DsbA (residues 1-19; Table S1.1) enabled efficient translocation of TrxA into the periplasm (Figure 1.4 B). The N-terminal sequence containing the TMD of FtsQ (residues 1-33; Table S1.1), a substrate of the co-translational SRP pathway, directed efficient targeting of the fusion protein to the inner-membrane (Figure 1.4 B). The C-terminal Myc tag in FtsQ(1-33)-TrxA was susceptible to proteinase K digestion in spheroplasts, indicating that its C-terminus is successfully translocated across the bacterial inner-membrane (Figure 1.4 B). These data validated the robustness of the TrxA-based assay to distinguish co- versus post-translational modes of targeting and insertion *in vivo*. Importantly, when RodZ residues 1-150 encompassing its TMD was fused to TrxA (Table S1.1), the fusion protein was efficiently targeted to and translocated across the bacterial inner membrane analogously to FtsQ (Figure 1.4 B), indicating that the RodZ-TrxA fusion protein was co-translationally targeted and inserted *in vivo*.

To further test the dependence of the targeting reaction on the SRP protein, Ffh, we used the WAM121 strain in which Ffh expression is under control of the *ara* promoter (de Gier et al., 1996). In contrast to FtsQ, which depends on Ffh for proper insertion into the membrane, RodZ was not sensitive to Ffh depletion (Figure 1.4 D and F), consistent with the previous report that RodZ requires SecA, but not SRP or the SRP receptor FtsY, for membrane insertion (Ulbrandt et al., 1997; Rawat et al., 2015). Thus, the N-terminal sequence of RodZ is sufficient to direct the co-translational targeting of the remainder of the protein via an SRP-independent pathway.

Although SecA dependence was observed for the targeting and translocation of all the NTS-TrxA fusion proteins tested in the *in vivo* assay (Figure S1.2 A), these data likely reflect a requirement for SecA during the translocation of the TrxA moiety and cannot be used to conclusively infer the involvement of SecA in their targeting (Schierle et al., 2003). In addition, *in vivo* experiments could only demonstrate the requirement, but not sufficiency, for specific factors. These limitations were addressed by *in vitro* reconstitution experiments described in the next section.

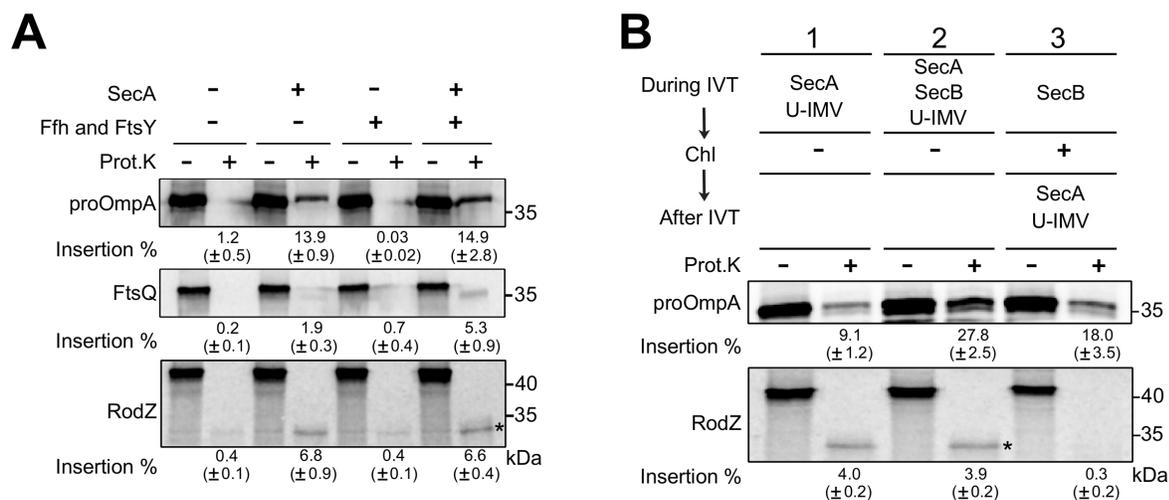
***SecA provides the minimal factor sufficient to drive the co-translational targeting and insertion of RodZ in vitro***

We sought to reconstitute the targeting and insertion of nascent RodZ using the PURE *in vitro* translation (IVT) system (Shimizu et al., 2001) coupled with urea-washed inner-membrane vesicles (U-IMV; Kuruma et al., 2008); successful translocation of substrate proteins across U-IMV leads to their partial or complete protection from proteinase K digestion. This homologous IVT-translocation system contains no endogenous targeting factors, allowing us to probe the contribution of specific factors to the targeting and translocation of protein substrates of interest.

OmpA is a well-studied outer-membrane protein that is post-translationally targeted and translocated by SecA (Hoffschulte et al., 1994; Kuruma et al., 2008). Consistent with these expectations, proOmpA exhibited SecA-dependent but SRP- and FtsY-independent targeting and translocation across U-IMV in the IVT-translocation assay (Figure 1.5 A and replicates in Figure S1.3 A). On the other hand, FtsQ requires SRP and FtsY for targeting to the membrane, and SecA for translocation of its periplasmic loop (Scotti et al., 1999; Kuruma et al., 2008). The coupled IVT-translocation assay recapitulated the dependences of FtsQ on both factors (Figure 1.5 A and replicates in Figure S1.3 A). Importantly, RodZ was inserted in the presence of SecA alone in this assay, and the additional presence of SRP/FtsY did not improve its translocation efficiency (Figure 1.5 A and Figure S 1.3). These data are consistent with the *in vivo* observation that RodZ requires SecA, but not SRP and FtsY, for its proper biogenesis (Figure 1.4; Ulbrandt et al., 1997; Rawat et al., 2015). Moreover, they strongly suggest that SecA provides the minimal factor that can mediate the targeting and insertion of newly synthesized RodZ.

To test the co-translational requirement for RodZ insertion in this assay, we changed the order of addition of targeting/translocation components. Robust insertion of RodZ was only observed if SecA and U-IMV were added during IVT (Figure 1.5 B and Figure S1.3 B, reaction 1). In contrast, if SecA and U-IMV were added after termination of translation by chloramphenicol, no insertion was observed (Figure 1.5 B and Figure S1.3 B, reaction 3). Under this obligatorily post-translational condition, proOmpA was still efficiently inserted, albeit with lower efficiency than if SecA and U-IMV were supplied during IVT (Figure 1.5 B and Figure S1.3 B, reactions 2 vs. 3). Finally, while inclusion of the post-translational chaperone SecB improved the insertion efficiency of proOmpA, as previously reported (Kuruma et al., 2008), SecB did not affect the targeting and insertion of RodZ (Figure 1.5 B and Figure S1.3 B,

reactions 1 vs. 2; Rawat et al., 2015). Together, these results support the model that SecA provides the minimal machinery sufficient for the co-translational targeting and insertion of RodZ.



**Figure 1.5. Reconstitution of RodZ targeting and translocation *in vitro*.**

(A) Effect of SecA and SRP/FtsY on the translocation of indicated substrates into U-IMV during PURE-IVT. Reactions contained 400 nM Ffh, 1  $\mu$ M FtsY, and 0.94  $\mu$ M SecA where indicated. 4.5S RNA was included in the tRNA mix (Kuruma et al., 2008). (B) Targeting and translocation of RodZ is strictly co-translational, whereas that of proOmpA is not. Reactions contained 0.94  $\mu$ M SecA and 2.5  $\mu$ M SecB where indicated. Chl, chloramphenicol. Values under each lane are quantifications of % translocation from these data and their replicates (Figure S1.3) and represent mean  $\pm$  S.D., with  $n = 2-3$ . Asterisks denote the protected fragment after proteinase K digestion.

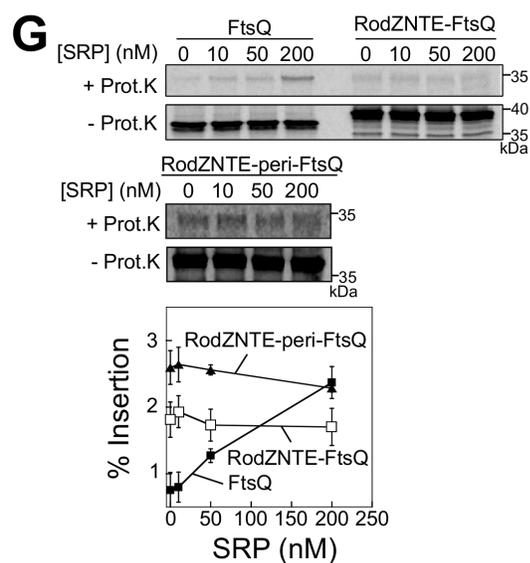
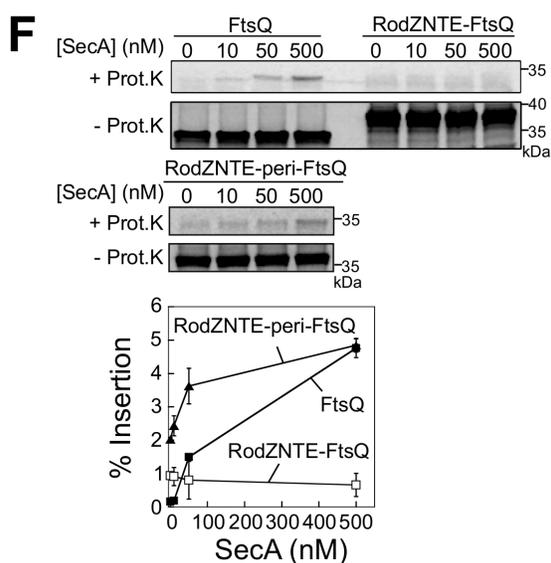
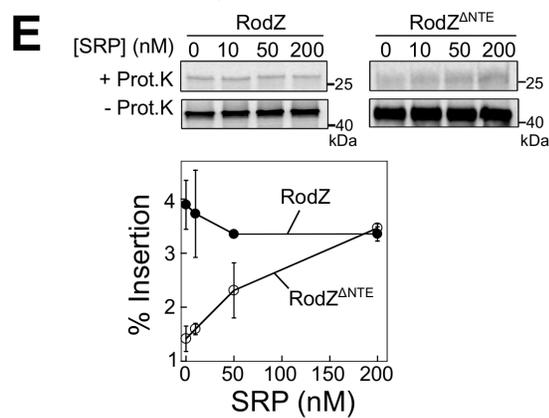
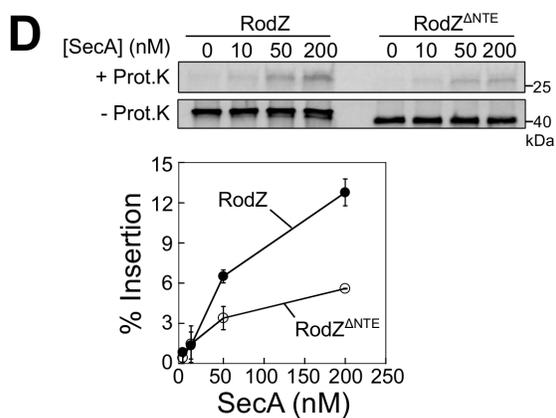
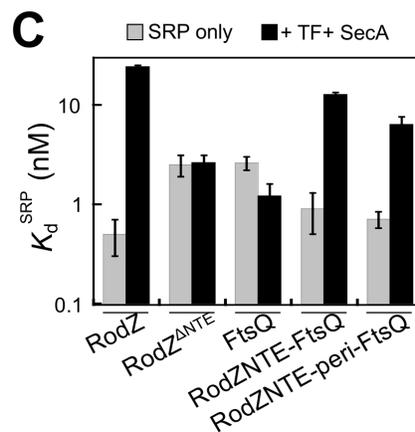
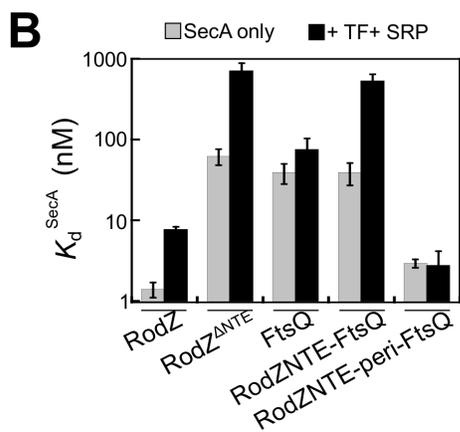
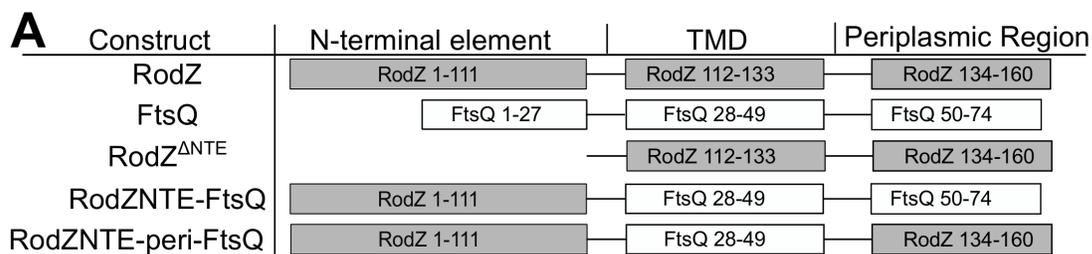
### ***The extended N-terminal extension and early periplasmic region of RodZ dictate its selection by SecA over SRP***

The majority of the bacterial inner-membrane proteome is generally thought to be targeted by SRP, which recognizes hydrophobic TMDs or signal sequences on the nascent polypeptide. The observation that SecA also co-translationally recognizes the RodZ-TMD raises the intriguing question of how nascent membrane proteins are selected between these two factors. Comparison of RodZ with well-studied SRP substrates, such as FtsQ, suggested the 111 residue N-terminal extension (NTE) of RodZ preceding its TMD as a potential distinguishing

feature. Another SecA substrate, EspP, was shown to be excluded from the SRP pathway due to its extended NTE, and deletion of this NTE re-routes EspP to the SRP pathway (Peterson et al., 2003; von Loeffelholz et al., 2013). We therefore hypothesized that, analogous to EspP, the extended NTE of RodZ disfavors its engagement with SRP.

To test this hypothesis, we deleted the NTE of RodZ (RodZ<sup>ΔNTE</sup>) or fused the RodZ NTE to the N-terminus of FtsQ-TMD (RodZNTE-FtsQ) (Figure 1.6 A and Table S1.1). We tested the effects of this mutation on multiple activities: (i) the binding affinity of SecA and SRP for RNCs displaying wildtype and mutant nascent chains (Figure 1.6 B and C); (ii) the SecA- and SRP-dependence of preprotein targeting and translocation across U-IMVs *in vitro* (Figure 1.6 D-G); and (iii) the SRP-dependence of translocation of NTS-TrxA fusion proteins *in vivo* (Figure S1.2 C and D). Deletion of the NTE significantly weakened the binding of SecA to RNC<sub>RodZ</sub>, and the weakened binding was exacerbated in the presence of competing TF and SRP (Figure 1.6 B, RodZ vs. RodZ<sup>ΔNTE</sup>). RodZ<sup>ΔNTE</sup> also exhibited reduced SecA-dependent targeting and translocation across U-IMVs *in vitro* than RodZ (Figure 1.6 D). These results are consistent with our earlier finding that the basic residues in the RodZ NTE are important for high affinity SecA recruitment (Figure 1.3 C).

On the other hand, deletion of the NTE from RodZ enabled strong SRP binding to the RNC even in the presence of competing SecA and TF (Figure 1.6 C, RodZ vs. RodZ<sup>ΔNTE</sup>). In agreement with the binding data, deletion of the NTE converted RodZ into an SRP-dependent substrate in the IVT-translocation assay *in vitro* (Figure 1.6 E) and increased the SRP-dependence of the translocation of RodZ-TrxA fusion proteins *in vivo* (Figure S1.2 C and D, RodZ vs. RodZ<sup>ΔNTE</sup>). These data suggest that the NTE of RodZ disfavors SRP binding. As predicted from this hypothesis, fusion of the RodZ NTE to the N-terminus of FtsQ TMD destabilized SRP binding to RNC<sub>FtsQ</sub> in the presence of SecA and TF (Figure 1.6 C, FtsQ vs. RodZNTE-FtsQ, black bars). Consistent with these binding data, fusion to the RodZ NTE also abolished the SRP-dependence of FtsQ targeting to U-IMVs *in vitro* (Figure 1.6 G, FtsQ vs. RodZNTE-FtsQ) and reduced the SRP-dependence of the targeting and insertion of FtsQ-TrxA *in vivo* (Figure S1.2 C and D, FtsQ vs. RodZNTE-FtsQ). Thus, the N-terminal extension of RodZ

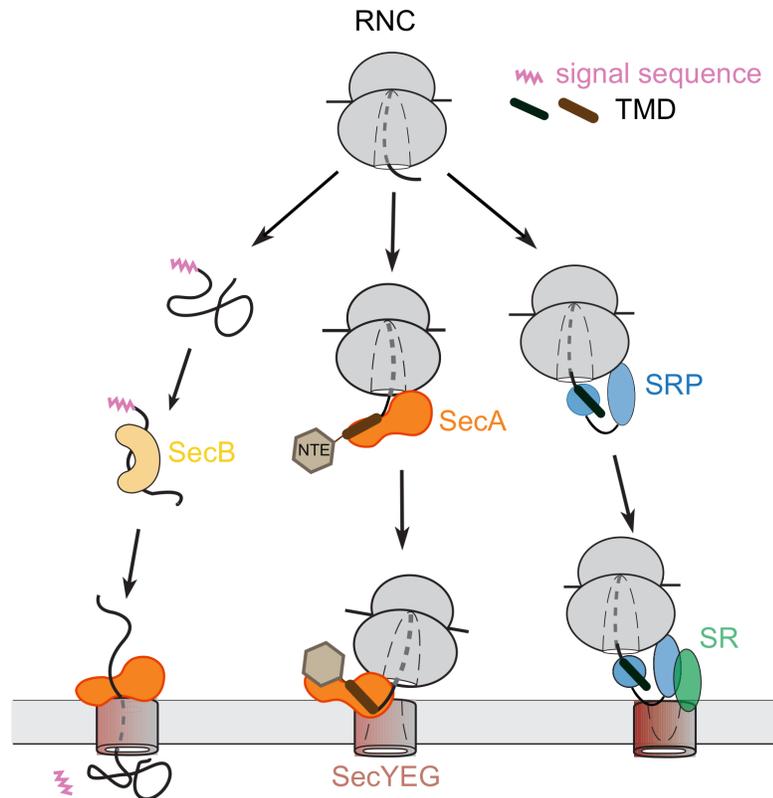


**Figure 1.6. The NTE and early periplasmic region of RodZ together dictate the selection of a membrane protein into the SecA versus SRP pathway.**

(A) Scheme of the sequence elements of the substrate variants tested in this figure. Detailed sequences are in Table S1.1. (B, C) Summary of the  $K_d$  values of RNCs bearing different nascent chains for binding to SecA (B) or SRP (C), derived from the equilibrium titrations in Figure S1.4. All titrations contained 20 nM RNC and 2  $\mu$ M TF, 400 nM SRP or 2  $\mu$ M SecA where indicated. (D, E) *In vitro* translocation assays of wildtype RodZ or mutant RodZ <sup>$\Delta$ NTE</sup>, and their dependence on SecA (D) or SRP (E). (F, G) *In vitro* translocation assays of wildtype FtsQ and mutants RodZNTE-FtsQ and RodZNTE-peri-FtsQ. The dependence of the reaction on SecA was shown in (F), and the dependence on SRP was shown in (G). The reactions in parts D, F contained 3.8  $\mu$ M TF, 400 nM Ffh, 1  $\mu$ M FtsY, and indicated concentrations of SecA. The reactions in parts E, G contained 50 nM SecA, 3.8  $\mu$ M TF, and indicated concentrations of SRP and a five-fold excess of FtsY over SRP. Values represent mean  $\pm$  S.D., with n = 2-3.

is necessary and sufficient to prevent the nascent protein from engaging the SRP targeting machinery.

However, fusion of the RodZ NTE to the N-terminus of FtsQ did not confer tight SecA binding (Figure 1.6 B, RodZNTE-FtsQ) nor efficient SecA-dependent targeting into U-IMVs (Figure 1.6 F, RodZNTE-FtsQ), indicating that the NTE of RodZ is not sufficient to re-route an SRP substrate to a SecA-dependent pathway. Since the periplasmic region of RodZ following its TMD is also important for high affinity SecA recognition (Figure 1.3), we further replaced the sequences in the FtsQ periplasmic domain following its TMD (residues 50-74) with the corresponding sequence from RodZ (Figure 1.6A, RodZ NTE-peri-FtsQ). RNCs bearing the resulting construct bound tightly to SecA (Figure 1.6 B) and displayed SecA-dependent targeting and insertion into U-IMVs *in vitro* (Figure 1.6 F). RodZ NTE-peri-FtsQ did not bound strongly to SRP (Figure 1.6 C), nor was it targeted and inserted into U-IMVs in an SRP-dependent manner (Figure 1.6 G), indicating that it resembles RodZ as a SecA-dependent and SRP-independent substrate. Thus, the extended NTE together with the early periplasmic region of RodZ are sufficient to re-route an SRP-dependent membrane protein into the alternative SecA-mediated co-translational targeting pathway.



**Figure 1.7. Diverse targeting pathways deliver nascent proteins to the SecYEG translocon at the inner-membrane.**

Left path, proteins with weakly hydrophobic signal sequences are maintained soluble by SecB and targeted to membrane via interaction with SecA, which translocates the nascent polypeptide across SecYEG. Right path, proteins containing hydrophobic TMDs or signal sequences are co-translationally recognized by SRP and targeted to SecYEG via the SRP-SRP receptor (SR) interaction. Middle path, proteins harboring internal TMDs are co-translationally recognized and targeted by SecA.

## 1.4 Discussion

Protein targeting to the bacterial cytoplasmic membrane was generally thought to occur via two major pathways (Figure 1.7). The majority of periplasmic, secretory, and outer membrane proteins contain weakly hydrophobic signal sequences and are targeted post-translationally, with or without the aid of the chaperone SecB, to SecA•SecYEG complexes that translocate preproteins across the inner membrane (left path). Proteins containing TMDs or highly hydrophobic signal sequences near the N-terminus are recognized by SRP as soon as they emerge from the ribosome exit tunnel, and are delivered co-translationally to the SecYEG translocation machinery via interaction between SRP and the SRP receptor (right path). This work demonstrated the existence of an alternative targeting route, mediated by SecA, for co-translational targeting to SecYEG sites and integration into membrane (Figure 1.7, middle path). The complete repertoire of substrate proteins utilizing this targeting route remains to be defined. Nevertheless, together with the finding of other substrates that exhibit distinct requirements for alternative translocases (Samuelson et al., 2000; van der Laan et al., 2004), our results add to the diversity of protein targeting mechanisms in bacteria.

SecA is an essential ATPase in bacteria known to drive the post-translational translocation of secretory and outer membrane proteins across the SecYEG translocation machinery. The recent findings that SecA also binds ribosomes near the nascent polypeptide exit site (Huber et al., 2011; Singh et al., 2014) suggested additional roles for this protein, but the function of the SecA-ribosome interaction has been unclear. The previous model, in which nascent proteins contact SecA during translation and then engage SecB for membrane delivery after they are released from the ribosome (Huber et al., 2011), regresses to a largely post-translational mechanism of targeting. The results here demonstrate a new possibility: SecA can specifically recognize and mediate the targeting/translocation of some inner membrane proteins in a strictly co-translational manner. Although the interactions of SecA with nascent periplasmic and outer membrane proteins have been previously characterized and are known to facilitate translocation (Karamyshev and Johnson, 2005; Huber et al., 2011; Huber et al., 2016), the interaction and activity of SecA on RodZ nascent chain observed here is the first example in which the co-translational mode of SecA action is mechanistically obligatory for the proper biogenesis of the substrate protein. Thus, this work provides a potential mechanism by which the SecA-ribosome interaction plays an essential role in nascent protein biogenesis. Additional

mechanistic roles for the SecA-ribosome interaction include providing an early chaperone for nascent polypeptides, or facilitating the translocation of large periplasmic loops for proteins still bound to the ribosome; these possibilities remain to be explored.

Nascent RodZ was shown to bind SRP in the ribosome profiling experiments (Schibich et al., 2016). This is consistent with our observation here that RodZ still binds SRP with a  $K_d$  value of 24 nM in the presence of physiological concentrations of TF and SecA (Figure 1.6 C). Indeed, SRP altered the FRET value of the RNC-SecA complex, and the weakening effect of SRP on RNC-SecA binding saturated above 50 nM (Figure S1.1 C). These observations argue against a model in which the binding of SRP and SecA to RNC<sub>RodZ</sub> is mutually exclusive, and is more consistent with a model in which these two factors allosterically modulate the affinity and conformation of one another at the ribosome exit site (please see Ariosa et al., 2014), for a formulation of the different models and their experimental predictions). However, as we have described extensively (Zhang et al., 2009; 2010), binding of SRP to an RNC does not necessarily turn on downstream steps in the targeting pathway, including efficient assembly of SRP with the SRP receptor, regulated GTP hydrolysis in the SRP•SR complex, and cargo unloading at the membrane translocon. Given the observation that RodZ does not require SRP for insertion *in vitro* and *in vivo*, the observed binding of SRP on RNC<sub>RodZ</sub> likely represents a ‘standby’ interaction mode of SRP that does not lead to SRP-dependent targeting.

The ribosome exit site is a crowded environment at which multiple protein biogenesis factors can bind and access the nascent polypeptide. The ability of SecA to co-translationally interact with nascent proteins further increases the complexity of this environment. This raises intriguing questions as to how nascent proteins are selected by the proper biogenesis factor(s); the preference of both SecA and SRP for hydrophobic TMDs renders this selection particularly challenging. Although the precise mechanism remains to be determined, the results here provided important information. Firstly, the extended NTE of RodZ effectively weakens the interaction of SRP with nascent proteins. This is analogous to the long NTEs preceding the signal sequences of bacterial autotransporters, such as EspP, which also act as self-sufficient SRP avoidance sequences (Peterson et al., 2003; von Loeffelholz et al., 2013). Interestingly, the recently discovered SRP-independent pathway in yeast primarily targets ER-destined membrane proteins harboring internal TMDs (Ast et al., 2013; Aviram et al., 2016), suggesting that a long N-terminal sequence preceding a downstream TMD might be a general feature to disfavor

engagement with the SRP machinery. In addition, the consecutive basic residues immediately upstream of the TMD facilitates SecA recruitment. Enrichment of N-terminal basic residues correlated with enhanced signal sequence binding and preprotein translocation by SecA (Akita et al., 1990; Hikita and Mizushima, 1992; Gelis et al., 2007). The same enrichment of basic residues was also observed in the NTE of EspP (Peterson et al., 2003) and might provide another distinguishing feature that favors the selection for SecA over SRP. Finally, the periplasmic region of RodZ following its TMD is also required for directing the substrate into the SecA-dependent targeting pathway. Acidic residues in early periplasmic region have been shown to be important for the translocation of secretory protein across membrane (Kajava et al., 2000). The results here suggest a function of these acidic residues to directly interact with SecA to facilitate translocation. Given the challenges in recognizing degenerate topogenic signals on nascent proteins amongst a multitude of biogenesis factors, such a 'multiplexed' recognition mechanism might be an effective strategy to ensure accurate nascent protein selection into the appropriate biogenesis pathway.

## 1.5 Methods

**Protein expression and purification.** N-terminally His<sub>6</sub>-tagged SecA (wildtype and mutant) was cloned in pET28a and expressed in BL21 (DE3) cells. Expression was induced by 0.5 mM IPTG at OD<sub>600</sub> = 0.5, 30 °C for 4 hrs. Cells were lysed by French pressing in SecA500 buffer (20 mM KHEPES, pH 7.5, 500 mM KCl, 4 mM MgCl<sub>2</sub>, 4 mM BME) containing 10 mM Imidazole and cOmplete™ protease inhibitor cocktail (Roche). Clarified lysate were loaded onto Ni-NTA resin and washed with SecA500 buffer. Protein was eluted with SecA500 buffer containing 250mM Imidazole. The His<sub>6</sub>-tag was removed by TEV protease digestion in SecA200 buffer (20 mM KHEPES, pH 7.5, 200 mM KCl, 4 mM MgCl<sub>2</sub>, 10 mM Imidazole, 4 mM BME) at 4 °C overnight and reloaded onto Ni-NTA. Flowthrough was collected, exchanged into SecA50 buffer (20 mM Tris-HCl, pH 8.0, 50 mM KCl, 4 mM MgCl<sub>2</sub>, 2 mM DTT), and further purified on MonoQ 10/100 GL (GE healthcare) using a linear gradient of 50 – 1000 mM KCl.

Ffh, FtsY, and TF were expressed and purified as described (Peluso et al., 2000; Jagath et al., 2000; Ariosa, et al., 2014). pHKSB366 encoding SecB was a gift from Andrey Karamyshev (Fekkes et al., 1998). SecB was expressed in BL21(DE3) using 1 mM IPTG at OD<sub>600</sub> = 1.0, 37 °C for 2 hrs. Cell was lysed by sonication in SecB buffer 1 (50 mM potassium phosphate, pH 7.5, 300 mM NaCl) containing 20 mM Imidazole. Clarified lysate was precipitated with 50% ammonium sulphate and centrifuged at 10,000 g for 10 min. The pellet was resuspended in SecB buffer 1 and loaded onto Ni-NTA pre-equilibrated with SecB buffer 1. SecB was eluted with SecB buffer 1 containing 500 mM Imidazole followed by dialysis in 50 mM Tris, pH 7.5. After ultracentrifugation in TLA100.3 (Beckman Coulter) at 60,000 g for 1hr, the supernatant was loaded onto MonoQ equilibrated in SecB buffer 2 (50 mM Tris-HCl, pH 7.5, 30 mM NaCl), and eluted with a linear gradient of 30 – 1000 mM NaCl. The protein was desalted in SecB buffer 2.

SUMO and SUMO fusions to the periplasmic segments of RodZ or FtsQ were expressed using a pET28 vector encoding N-terminal His<sub>6</sub>-tag, full length SUMO family protein SMT3 from *S. cerevisiae*, and RodZ residues 134-160 or FtsQ residues 50-74 where applicable. Proteins were expressed in BL21(DE3) using 0.5 mM IPTG at OD<sub>600</sub> = 0.5, 37 °C for 3 hrs. Cell was lysed by sonication in SUMO buffer 1 (20 mM KHEPES, pH 7.5, 300 mM NaCl, 4 mM BME) containing 20 mM Imidazole and cOmplete™ protease inhibitor cocktail (Roche). Clarified lysate were loaded onto Ni-NTA resin and washed with SUMO buffer 1. Protein was eluted with SUMO buffer 1 containing 250 mM Imidazole. Proteins were dialyzed against

SUMO buffer 2 (20 mM KHEPES, pH 7.5, 300 mM NaCl, 10% glycerol, 2 mM TCEP) at 4 °C and stored at -80°C.

**RNC preparation.** Cm-labeled RNCs were generated by *in vitro* translation in S30 extract supplemented with Cm (Bachem), tRNA<sup>Cm</sup>, and Cm tRNA synthetase, as described previously (Schaffitzel et al., 2006; Saraogi et al., 2011). To prepare RNCs harboring mutant L23(F51A/E52A/E54A/E56A/E89A), S30 extract was prepared from the strain KC623 harboring this L23 mutant (KC6  $\Delta rplW::kan$  pL23(F51A/E52A/E54A/E56A/E89A), see Supplementary Methods for strain construction). All RNCs were purified via N-terminal Strep<sub>3</sub>-tags on the nascent protein using Strep-Tactin Sepharose resin (IBA) as described (Schaffitzel et al., 2006).

**Fluorescent labeling.** The single cysteine mutant, Ffh (C406S/D421C), was purified and labeled with BODIPY-FL maleimide (ThermoFisher scientific) as described previously (Akopian et al., 2013). The single cysteine mutant of SecA, SecA(C98S/S12C), was reduced with 2 mM DTT at 4 °C for 30 min followed by dialysis in Labeling buffer (20 mM KHEPES, pH 7.0, 300 mM KCl, 10% glycerol, 2 mM TCEP) to remove DTT. 40  $\mu$ M SecA (C98S/S12C) was mixed with a 20 fold excess of BODIPY-FL maleimide on a rotary shaker at 4 °C for 4 hr. After quenching with 10 mM DTT, free dye was removed by chromatography on Sephadex G-25 column (Sigma-Aldrich) in SRP buffer (50 mM KHEPES, pH 7.5, 150 mM KOAc, 10 mM Mg(OAc)<sub>2</sub>, 2 mM DTT, 10% glycerol). Labeling efficiencies were 86% and 78% for Ffh and SecA, respectively, determined using the adsorption coefficient of  $\epsilon=73,000 \text{ M}^{-1}\text{cm}^{-1}$  for BODIPY-FL maleimide in aqueous buffer (Stray et al., 2006). The cysteines in the zinc-finger domain of SecA are coordinated by Zn<sup>2+</sup> and were not labeled (data not shown).

**Fluorescence measurements.** All proteins were ultracentrifuged in TLA100 (Beckman Coulter) at 100,000 g for 1hr prior to fluorescence measurements. Fluorescence assays were performed as described previously (Zhang et al., 2010; Ariosa et al., 2014) at room temperature in Assay buffer (50 mM KHEPES, pH 7.5, 150 mM KOAc, 10 mM Mg(OAc)<sub>2</sub>, 2 mM DTT, 0.1 mg/ml BSA). Equilibrium titrations were performed using 20 nM Cm-labeled RNC, indicated concentrations of cytosolic competitors where applicable, and SecA or Ffh as the titrant. The observed FRET value at individual titrant concentrations (FRET<sub>obsd</sub>) was calculated from Eq 1,

$$\text{FRET}_{\text{obsd}} = 1 - D_A/D_0 \quad (1)$$

in which  $D_0$  is the donor fluorescence signal in the absence of the FRET acceptor, and  $D_A$  is the donor fluorescence signal in the presence of the acceptor-labeled titrant.

The concentration dependence of  $\text{FRET}_{\text{obsd}}$  in a titration curve were fit to Eq 2 (Cooper, 2004):

$$\text{FRET}_{\text{obsd}} = \text{FRET}_{\text{max}} \times \frac{[\text{RNC}] + [\text{titrant}] + K_d - \sqrt{([\text{RNC}] + [\text{titrant}] + K_d)^2 - 4 \times [\text{RNC}][\text{titrant}]}}{2 \times [\text{RNC}]} \quad (2)$$

in which  $[\text{RNC}]$ ,  $[\text{titrant}]$ , and  $\text{FRET}$  are input values,  $\text{FRET}_{\text{max}}$  is the FRET value at saturating titrant concentration, and  $K_d$  is the dissociation constant of the complex of interest.

To facilitate comparison of complexes with different  $K_d$  values,  $\text{FRET}_{\text{obsd}}$  were further divided by the  $\text{FRET}_{\text{max}}$  values, obtained from fitting the data to Eq 2, to generate normalized titration curves. These curves are described by Eq 3,

$$\text{Normalized } \Delta F = 1 \times \frac{[\text{RNC}] + [\text{titrant}] + K_d - \sqrt{([\text{RNC}] + [\text{titrant}] + K_d)^2 - 4 \times [\text{RNC}][\text{titrant}]}}{2 \times [\text{RNC}]} \quad (3)$$

To measure the binding of the SUMO-RodZPeri fusion protein to SecA, 50 nM SecA<sup>BDP</sup> was pre-incubated with 20 nM Cm-labeled RNC<sub>RodZ</sub>. Increasing concentrations of SUMO-RodZ(peri) was added as a competitive inhibitor of the FRET between SecA<sup>BDP</sup> and RNC<sup>Cm</sup>, and the observed changes in fluorescence intensity of Cm-labeled RNC<sub>RodZ</sub> ( $F_{\text{obsd}}$ ) was recorded. The data were fit to Eq 8, derived by numerically solving the four relationships (Eq 4-7) according to the reaction scheme in Figure 1.3G,

$$[\text{SecA}^{\text{BDP}}] + [\text{SecA}^{\text{BDP}} \cdot \text{RNC}_{\text{RodZ}}] + [\text{SecA}^{\text{BDP}} \cdot \text{SUMO variant}] = 50 \text{ nM} \quad (4)$$

$$[\text{SecA}^{\text{BDP}} \cdot \text{RNC}_{\text{RodZ}}] + [\text{RNC}_{\text{RodZ}}] = 20 \text{ nM} \quad (5)$$

$$\frac{[\text{SecA}^{\text{BDP}}] \times [\text{RNC}_{\text{RodZ}}]}{[\text{SecA}^{\text{BDP}} \cdot \text{RNC}_{\text{RodZ}}]} = K_d = 1 \text{ nM} \quad (6)$$

$$\frac{[\text{SecA}^{\text{BDP}}] \times [\text{SU}]}{[\text{SecA}^{\text{BDP}} \cdot \text{SUMO variant}]} = K_i \quad (7)$$

$$F_{\text{obsd}} = \frac{-[\text{SU}] - 31 \times K_i + \sqrt{[\text{SU}]^2 + 142 \times [\text{SU}] \times K_i + 1041 \times K_i^2}}{2 \times K_i} \times m + F_0 \quad (8)$$

In Eq 8, [SU] is the concentration of SUMO variant.  $K_i$  is the inhibition constant of the competitors for SecA.  $F_0$  is the initial fluorescence intensity of Cm-labeled RNC<sub>RodZ</sub> in the SecA<sup>BDP</sup>•RNC<sup>Cm</sup> complex, and  $m$  is the contribution to fluorescence intensity per nM of RNC<sup>Cm</sup>.

***In vivo translocation assay of NTS-TrxA fusions.*** pEK20 plasmids coding NTS-TrxA-myc fusion proteins were transformed into *E. coli* strains EO527 and WAM121, in which the expression of SecA and Ffh, respectively, were under control of the arabinose promoter (de Gier et al., 1996; Or et al., 2005). To deplete Ffh, WAM121 cells were grown to OD<sub>600</sub> = 0.5 in LB supplemented with 0.02% (w/v) L-arabinose, washed twice with LB supplemented with 0.4% (w/v) D-glucose, and sub-cultured in LB supplemented with 0.4% (w/v) D-glucose. Ffh level was reduced to < 5% after 3 hrs of media shift. SecA depletion in EO527 was performed similarly to Ffh depletion, except that the subculture was grown for 5 hrs to deplete SecA. At OD<sub>600</sub> = 0.4-0.6, NTS-TrxA-myc expression was induced by addition of IPTG (5  $\mu$ M for RodZ and RodZNTE-FtsQ, 50  $\mu$ M for all other constructs to achieve similar expression levels; Figure S1.2 B) for 30 min at 37 °C. Cells were harvested and resuspended in cold TrxA buffer 1 (0.1 M Tris-HCl, pH 8.0, 20% sucrose). 0.5 mM EDTA, pH 8.0 and 50  $\mu$ g/ml Lysozyme were added and the suspension was incubated at RT for 15 min. 20 mM MgSO<sub>4</sub> was added to stabilize spheroplasts. Spheroplasts were separated from the periplasmic fraction by centrifugation at 3,140 g for 10 min. For the proteinase K protection assay, spheroplasts were resuspended in cold TrxA buffer 2 (0.1 M Tris-HCl, pH 8.0, 20% sucrose, 20 mM MgSO<sub>4</sub>), and incubated with or without 0.5 mg/ml proteinase K on ice for 1hr. Reactions were stopped by addition of 5 mM PMSF. To further separate the cytosol from the membrane fraction, spheroplasts were resuspended in TrxA buffer 3 (50 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 1 mM PMSF), lysed by one freeze-thaw cycle in liquid nitrogen, and clarified in TLA120.1 rotor at 63,000 rpm for 1hr. The supernatant was the cytosolic fraction, and the membrane pellet was resuspended with TrxA buffer 4 (20 mM Tris-HCl, pH 8.0, 5 mM EDTA, 0.5% SDS). The translocation efficiencies for secretory proteins were calculated from the ratio of the secreted fraction (asterisk) to total protein amount. The translocation efficiencies for membrane proteins were calculated from the ratio of protein intensity after/before proteinase K digestion.

**Western blot.** Rabbit anti-SecA antibody was a kind gift from T.A. Rapoport. Rabbit anti-Ffh antibody was a kind gift from P. Walter. Rabbit anti-YidC antibody was a gift from R. E. Dalbey. All other antibodies were purchased from Abcam. Primary antibodies were incubated with IRDye<sup>®</sup> 800CW goat anti-rabbit IgG (LI-COR) for detection. Protein band intensity was quantified by the Odyssey<sup>®</sup> CLx imaging system.

**Preparation of U-IMV.** SecYEG was overexpressed in MRE600 by induction with 0.5 mM IPTG for 2 hr. Cells were harvested in IMV buffer 1 (50 mM TEA-OAc, pH 7.5, 250 mM sucrose, 1 mM EDTA, 1 mM DTT, 0.5 mM PMSF), and lysed at 8,000 psi by FRENCH PRESS (thermo scientific). Unbroken cells were removed by centrifugation at 4,000 g for 10 min. Membranes were further pelleted in Ti70 (Beckman Coulter) at 45,000 rpm for 2 hrs, and resuspended in IMV buffer 1. The membrane suspension was layered onto a 5-step sucrose gradient (0.8, 1.0, 1.2, 1.4, 1.6 M sucrose in IMV buffer 1), ultracentrifuged in SW32 (Beckman Coulter) at 24,000 rpm for 16 hrs. IMV fractions were collected from the lower one third of the gradient as described (Müller and Blobel, 1984a). To make urea-washed IMV, 4 volumes of IMV buffer 2 (50 mM TEA-OAc, pH 7.5, 250 mM sucrose, 1 M KOAc, 7.5 M urea) was added to IMV. The mixture was incubated on ice for 1 hr, after which the urea concentration was adjusted to 3M before pelleting through a sucrose cushion (50 mM TEA-OAc, pH 7.5, 750 mM sucrose, 1 M KOAc, 1 mM DTT) in TLA100.3 at 60,000 rpm for 2 hrs. The pellet was resuspended in IMV buffer 3 (50 mM TEA-OAc, pH 7.5, 250 mM sucrose, 1 mM DTT) (Helde et al., 1997; Müller and Blobel, 1984b).

**In vitro translocation assay in PURE system.** Translation was performed at 30 °C using PURExpress<sup>®</sup> *in vitro* protein synthesis kit (NEB), supplemented with <sup>35</sup>S-Methionine (1.5 mCi/ml, PerkinElmer) and indicated concentrations of cytosolic factors (SecA, SecB, Ffh, FtsY or TF). Unless otherwise indicated, 0.5 mg/ml U-IMV was added 5 min after initiation of translation. The reaction was continued for 85 min at 30 °C, after which it was split equally into two samples, one of which was digested with 0.5 mg/ml proteinase K for 30 min at 25 °C. Digestion was stopped by addition of 5 mM PMSF, after which the sample was incubated on ice for 10 min. Samples with and without proteinase K treatment were analyzed by SDS-PAGE and autoradiography. The insertion efficiency was calculated from the ratio of the intensity of

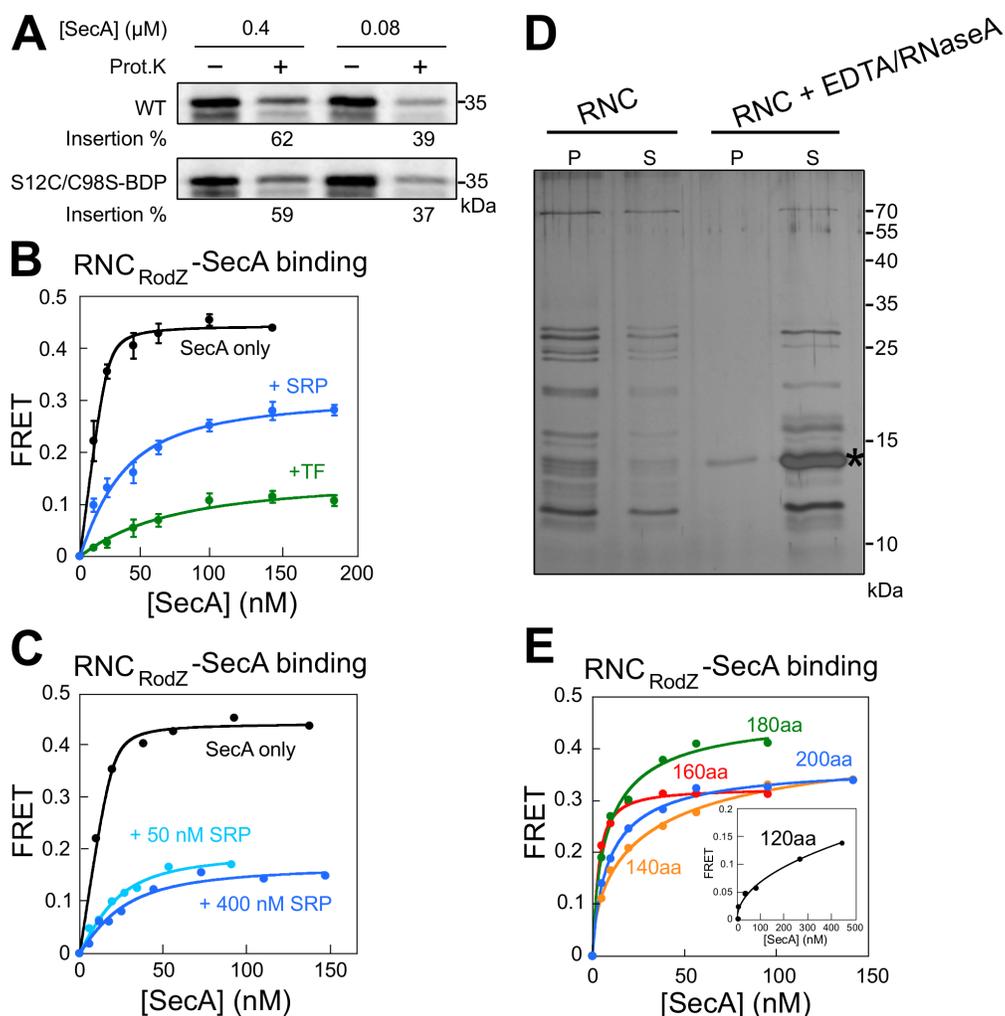
substrate protein bands after and before proteinase K treatment. For secretory proteins containing cleavage signal sequence, the loss of methionine(s) after signal sequence cleavage was corrected before calculation of insertion efficiency.

**Strains.** *E. coli* strains EO527 and WAM121 have been described (de Gier et al., 1996; Or et al., 2005). To construct the strain KC623 harboring mutant L23 (KC6  $\Delta rplW::kan$  pL23<sub>F51A/E52A/E54A/E56A/E89A</sub>), DNA coding L23 mutations was cloned into pEK20 by Gibson assembly (Gibson et al., 2009), and transformed into the *E. coli* strain KC6 (A19  $\Delta endA$   $met^+$ <sub>SEP</sub>  $\Delta tonA$   $\Delta speA$   $\Delta tnaA$   $\Delta sdaA$   $\Delta sdaB$   $\Delta gshA$ ) (Calhoun and Swartz, 2006). The genomic L23 in KC6 harboring pL23<sub>F51A/E52A/E54A/E56A/E89A</sub> was then knocked out by lambda-red recombination (Datsenko and Wanner, 2000).

**RNaseA/EDTA treatment of RNC.** To release nascent chains from the ribosome, RNCs were incubated with 20 mM EDTA (pH 8.0) and 50  $\mu$ g/ml RNaseA at 37 °C for 30 min (Ziehr et al., 2010). To verify the effectiveness of this treatment, RNCs before and after the treatment were sedimented in a TLA100 (Beckman Coulter) rotor at 100,000 g for 2.5 hr. The pellet was resuspended with SDS loading buffer at equal volume as the supernatant; both pellet and supernatant fractions were subject to SDS-PAGE analysis.

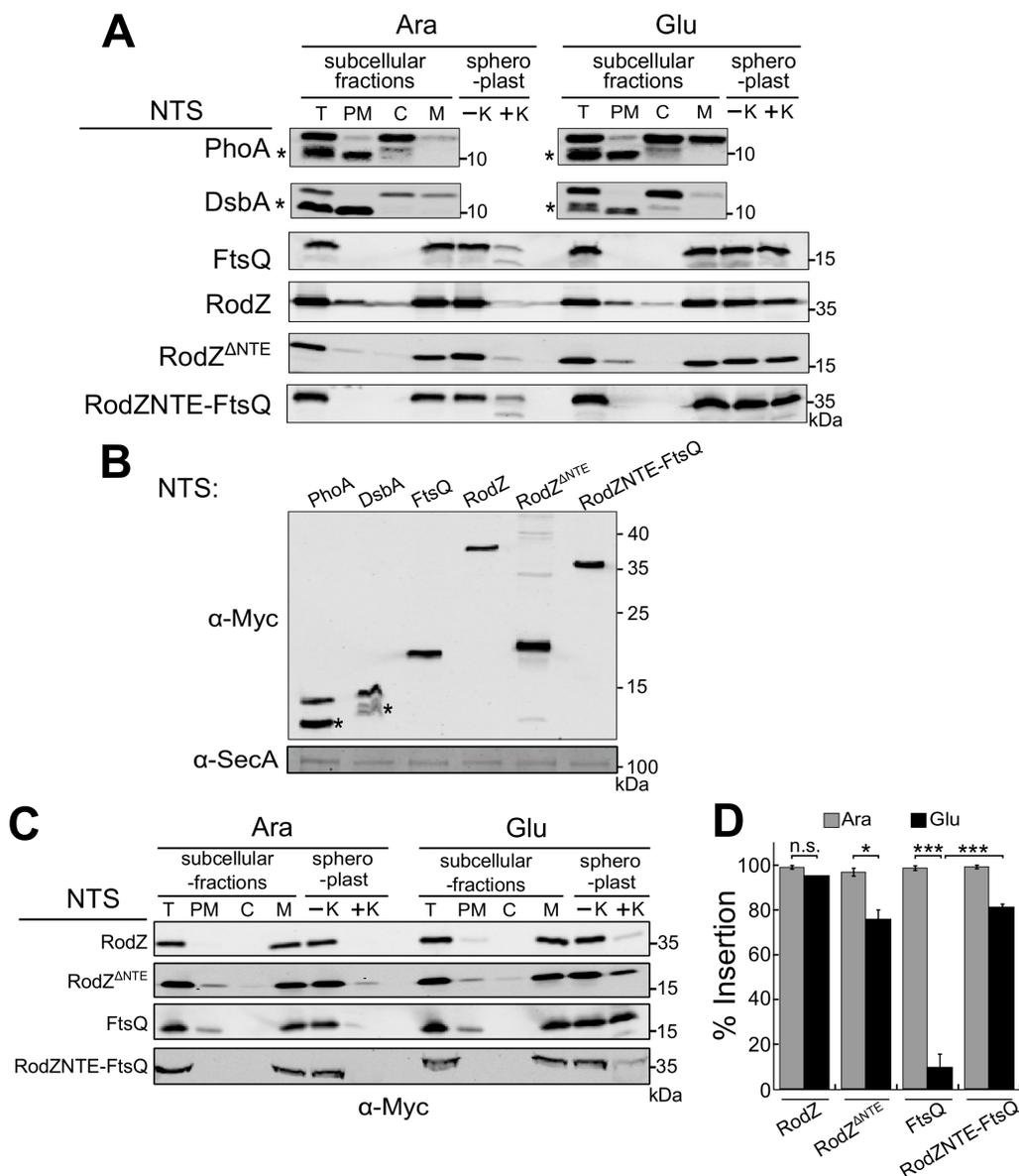
**ProOmpA translocation in wheat germ lysate.** ProOmpA mRNA was *in vitro* transcribed and purified as described (Behrmann et al., 1998). ProOmpA was translated using wheat germ extract (Promega) in the presence of <sup>35</sup>S-Methionine (1.5 mCi/ml, PerkinElmer) at 26 °C for 30 min, followed by incubation with U-IMV at 37 °C for 15 min in the presence of 10 mM phosphocreatine, 0.05 mg/ml creatine kinase, 4 mM Mg(OAc)<sub>2</sub>, 2 mM ATP, 0.5 mg/ml BSA, 10 mM DTT, and indicated concentrations of SecA. Samples were digested with 0.1 mg/ml proteinase K on ice for 15 min. Digestion was stopped by addition of 5 mM PMSF. All samples were precipitated by TCA and analyzed by SDS-PAGE and autoradiography.

## 1.6 Supplementary Figures and Tables



**Figure S1.1. Fluorescence measurements of SecA-RNC<sub>RodZ</sub> affinity.**

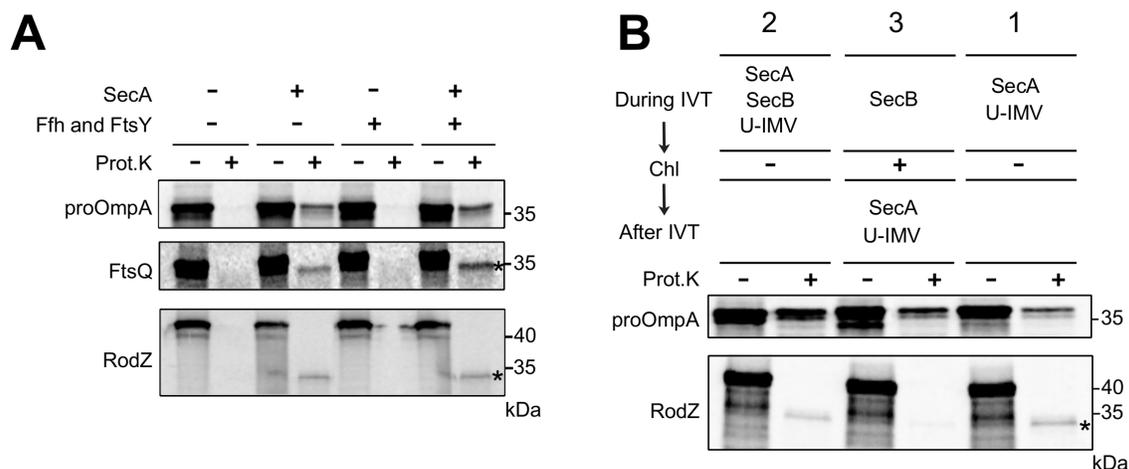
(A) Activities of wildtype SecA (upper) and mutant SecA(S12C/C98S)<sup>BDP</sup> (lower) in mediating proOmpA translocation across the U-IMV, carried out as described in Supplementary Methods. The bands indicate proOmpA. Translocation efficiency is the ratio of protected fragment after proteinase K digestion to the undigested amount. The values underneath each lane denote % insertion. (B, C, E) Equilibrium titration curves of SecA-RNC<sub>RodZ</sub> binding shown in Figure 1.1 C, D, and Figure 1.3 C, respectively, showing FRET signals before normalization of the fluorescence change. (D) Disassembly of the ribosomes in RNC by EDTA and RNaseA treatment, assayed after ultracentrifugation as described in Supplementary Methods. P, pellet; S, supernatant. The asterisk denotes RNaseA. (E) Titration of SecA to RNC<sub>RodZ</sub> harboring different lengths of RodZ nascent chain. Curves were fit to Eq 2. Values represent mean  $\pm$  S.D., with  $n = 2-3$ .



**Figure S1.2. Translocation of NTS-TrxA fusion proteins and their dependence on SecA or SRP.**

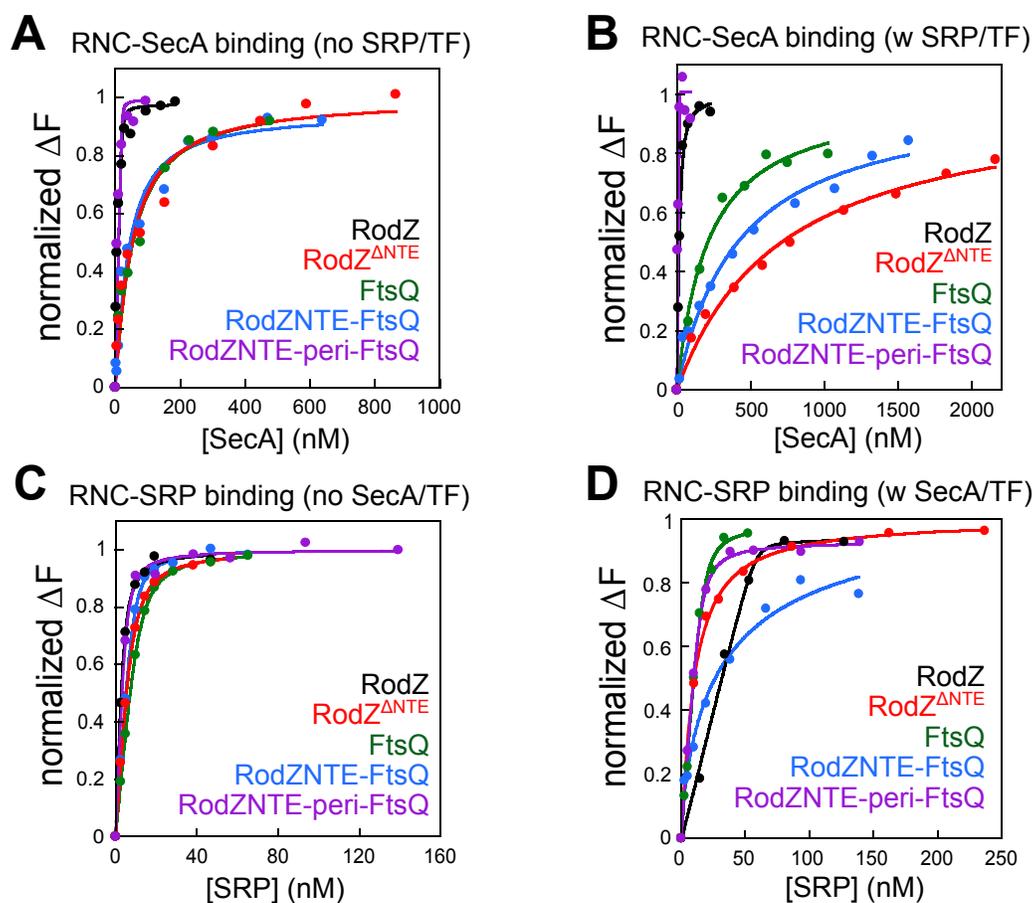
(A) Translocation of NTS-TrxA constructs across the inner membrane in wildtype (Ara) and SecA-depleted (Glu) cells. Translocated secretory proteins are indicated by asterisks. T, total; PM, periplasm; C, cytosol; M, membrane. K, proteinase K; T, triton X-100. (B) Western blot analysis of the expression levels of tested NTS-TrxA fusion proteins (upper gel), and the levels of SecA in cells expressing each substrate (lower gel). SecA was used as loading control. Signal sequence-cleaved secretory proteins are indicated by asterisks. (C) Translocation of NTS-TrxA fusion proteins across the bacterial inner membrane in wildtype (Ara) and Ffh-depleted (Glu) cells. Ffh is under the control of arabinose promoter. (D) Quantification of the translocation efficiency of NTS-TrxA constructs, from the data in (C) and their replicates. Values represent mean  $\pm$  S.D., with

n = 2-3. P-values in panel D were calculated using *Welch's* t-test. n.s., not significant; \*, P < 0.05; \*\*\*, P < 0.001.



**Figure S1.3. Replicates for reconstitution of RodZ targeting and translocation *in vitro*.**

(A) Effect of SecA and SRP/FtsY on the translocation of indicated substrates into U-IMV during PURE-IVT. Reactions contained 400 nM Ffh, 1  $\mu$ M FtsY, and 0.94  $\mu$ M SecA where indicated. 4.5S RNA was included in the tRNA mix (Kuruma et al., 2008). (B) Targeting and translocation of RodZ is strictly co-translational, whereas that of proOmpA is not. Reactions contained 0.94  $\mu$ M SecA and 2.5  $\mu$ M SecB when indicated. Chl, chloramphenicol. Asterisks denote the protected fragment after proteinase K digestion.



**Figure S1.4. Equilibrium titrations for RNC binding to SecA or SRP.**

(A, B) Representative equilibrium titrations to measure the binding of SecA to indicated RNCs without (A) or with (B) 2  $\mu$ M TF and 400 nM SRP present. (C, D) Representative equilibrium titrations to measure the binding of SRP to indicated RNCs without (C) and with (D) 20  $\mu$ M TF and 2  $\mu$ M SecA present.



Nascent chain	Amino acid sequence
PhoA	MKQSTIALALLPLLFTAmbVTKARTPEMPVLENRAAQGDITAPGGARRLTGDQTAALRDSLSDKPAKNII LLIGDGMGDSEITAARNYAEGAGGFFKGIDALDI
FtsQ	MGQAALNTRNSEEVEVSSRRNNGTRLGILFLLTVLTVLVSGWVVLGAmbMEDAQRLPLSKLVLTGER HYTRNDDI
RodZ	MNTEATHDQNEALTTGARLRNAREQLGLSQQAVAERLCLKVSTVRDIEEDKAPADLASTFLRGYIRSYA RLVHIPEEELLPGLEKQAPLRAAKVAPMQSFSLGKRRKKRDAmbWLMFTTWLVLFVIGLSGAWWWQ DRKAQEEITTMADQSSAELSSNSEQQQSVPLNTSTTTDPATTSTP
RodZ <sup>ANTE</sup>	MRDAmbWLMFTTWLVLFVIGLSGAWWWQDRKAQEEITTMADQSSAELSSNSEQQQSVPLNTSTT TDPATTSTP
RodZNTE-FtsQ	MNTEATHDQNEALTTGARLRNAREQLGLSQQAVAERLCLKVSTVRDIEEDKAPADLASTFLRGYIRSYA RLVHIPEEELLPGLEKQAPLRAAKVAPMQSFSLGKRRKKRDAmbILFLLTVLTVLVSGWVVLGWMEDA QRLPLSKLVLTGERHYTRNDDI
RodZNTE-peri-FtsQ	MNTEATHDQNEALTTGARLRNAREQLGLSQQAVAERLCLKVSTVRDIEEDKAPADLASTFLRGYIRSYA RLVHIPEEELLPGLEKQAPLRAAKVAPMQSFSLGKRRKKRDAmbILFLLTVLTVLVSGWVVLGWMQDR KAQEEITTMADQSSAELSSNSEQ
peri swap	MNTEATHDQNEALTTGARLRNAREQLGLSQQAVAERLCLKVSTVRDIEEDKAPADLASTFLRGYIRSYA RLVHIPEEELLPGLEKQAPLRAAKVAPMQSFSLGKRRKKRDAmbWLMFTTWLVLFVIGLSGAWWWE DAQRLPLSKLVLTGERHYTRNDDI
peri swap acidic	MNTEATHDQNEALTTGARLRNAREQLGLSQQAVAERLCLKVSTVRDIEEDKAPADLASTFLRGYIRSYA RLVHIPEEELLPGLEKQAPLRAAKVAPMQSFSLGKRRKKRDAmbWLMFTTWLVLFVIGLSGAWWWE DAQELPLSELVLTGEEDYTDNDDI
RodZ 120	MNTEATHDQNEALTTGARLRNAREQLGLSQQAVAERLCLKVSTVRDIEEDKAPADLASTFLRGYIRSYA RLVHIPEEELLPGLEKQAPLRAAKVAPMQSFSLGKRRKKRDAmbWLMFTTWLV
RodZ 140	MNTEATHDQNEALTTGARLRNAREQLGLSQQAVAERLCLKVSTVRDIEEDKAPADLASTFLRGYIRSYA RLVHIPEEELLPGLEKQAPLRAAKVAPMQSFSLGKRRKKRDAmbWLMFTTWLVLFVIGLSGAWWWQ DRKAQQ
RodZ 160	MNTEATHDQNEALTTGARLRNAREQLGLSQQAVAERLCLKVSTVRDIEEDKAPADLASTFLRGYIRSYA RLVHIPEEELLPGLEKQAPLRAAKVAPMQSFSLGKRRKKRDAmbWLMFTTWLVLFVIGLSGAWWWQ DRKAQEEITTMADQSSAELSSNSEQ
RodZ 180	same as RodZ
RodZ 200	MNTEATHDQNEALTTGARLRNAREQLGLSQQAVAERLCLKVSTVRDIEEDKAPADLASTFLRGYIRSYA RLVHIPEEELLPGLEKQAPLRAAKVAPMQSFSLGKRRKKRDAmbWLMFTTWLVLFVIGLSGAWWWQ DRKAQEEITTMADQSSAELSSNSEQQQSVPLNTSTTTDPATTSTPPASVDTTANTQTPAVTAPA



NTS	Amino acid sequence
PhoA	MKQSTIALALLPLLFTPVTKA
DsbA	MKKIWLALAGLVLAFSASA
FtsQ	MSQAALNTRNSEEVEVSSRRNNGTRLGILFLLTVLTVLVSGWVVLGWMEDAQR
RodZ	MNTEATHDQNEALTTGARLRNAREQLGLSQQAVAERLCLKVSTVRDIEEDKAPADLASTFLRGYIRSYA RLVHIPEEELLPGLEKQAPLRAAKVAPMQSFSLGKRRKKRDGWLMTFTWLVLFVIGLSGAWWWQDR KAQEEITTMADQS
RodZ <sup>ANTE</sup>	MRDGWLMTFTWLVLFVIGLSGAWWWQDRKAQEEITTMADQS
RodZNTE-FtsQ	MNTEATHDQNEALTTGARLRNAREQLGLSQQAVAERLCLKVSTVRDIEEDKAPADLASTFLRGYIRSYA RLVHIPEEELLPGLEKQAPLRAAKVAPMQSFSLGKRRKKRDGILFLLTVLTVLVSGWVVLGWMEDAQR

**Table S1.1. Constructs used in RNC preparation (top) and *in vivo* translocation assays (bottom).**

## Chapter 2 Molecular mechanism of co-translational membrane protein recognition and targeting by SecA

A modified version of this section is in press:

*Molecular mechanism of co-translational membrane protein recognition and targeting by SecA.* Shuai Wang\*, Ahmad Jomaa\*, Mateusz Jaskolowski, Chien-I Yang, Nenad Ban and Shouou Shan. *Nat. Struct. Mol. Biol.* **2019** (\*equal contribution)

### 2.1 Abstract

Co-translational protein targeting is a conserved process for the biogenesis of membrane proteins. In a recently described pathway, the essential ATPase SecA is necessary and sufficient to co-translationally recognize and deliver some nascent membrane proteins to the SecYEG translocase at the bacterial inner membrane; however, the molecular mechanism of this pathway is unclear. In this work, biochemical and cryoEM analyses show that the N-terminal amphipathic helix on SecA and the ribosomal protein uL23 together form a composite binding site for the transmembrane domain (TMD) on the nascent chain. This interaction positions additional sites on the ribosome and SecA for recognition of the charged residues on both sides of the TMD, explaining the substrate specificity of SecA recognition. Quantitative kinetic analyses demonstrate that membrane-embedded SecYEG can associate with and remodel the SecA-bound ribosome-nascent chain complex, which together with elongation of the nascent polypeptide facilitates handover of the translating ribosome to the translocase. Our work shows how the ribosome induces a distinct mode of nascent protein recognition and delivery by SecA.

### 2.2 Introduction

Membrane protein biogenesis is crucial for cell viability, due to the abundance (~30% of proteome) of membrane proteins and their participation in numerous essential cellular functions such as energy generation, molecular transport, and cell-cell communication (Brandman and Hegde, 2016; Shao and Hegde, 2015; Zhang and Shan, 2014). The localization, insertion, and folding of transmembrane domains (TMDs) are energetically costly (Cymer et al., 2014) and kinetically demanding (Elvekrog and Walter, 2015). To overcome these challenges, a strategy

widely used by cells is to recruit molecular chaperones when the TMD on a nascent polypeptide emerges from the ribosomal exit tunnel. These chaperones protect the nascent TMD from aggregation and also act as or in collaboration with dedicated targeting machinery to co-translationally deliver the nascent membrane protein to the translocation machinery on the target membrane (Nyathi et al., 2013).

Diverse membrane protein targeting pathways have been discovered. The most well-studied route is mediated by the signal recognition particle (SRP) (Zhang et al., 2010; Zhang and Shan, 2014), which is co-translationally recruited to ribosomes (Chartron et al., 2016; Schibich et al., 2016) and shields TMDs or hydrophobic signal sequences adjacent to the N-terminus of the nascent protein (Jomaa et al., 2016). The interaction of SRP with its membrane receptor SR delivers the ribosome•nascent chain complex (RNC) to the Sec61p translocase at the eukaryotic endoplasmic reticulum (ER), or the SecYEG translocase at the bacterial plasma membrane (Jomaa et al., 2016; Jomaa et al., 2017). In eukaryotic cells, the SRP-independent targeting (SND) components help in the delivery and insertion of a subclass of membrane proteins harboring internal TMDs to the ER, possibly before the nascent protein finishes its synthesis (Ast et al., 2013; Aviram et al., 2016). The ER membrane protein complex (EMC) could insert a subset of nascent TMDs into the membrane both co- and post-translationally (Chitwood et al., 2018; Guna et al., 2018; Shurtleff et al., 2018). The diversity of membrane protein targeting and translocation machineries are suggested to accommodate different properties of the nascent membrane proteome, such as hydrophobicity (Ast et al., 2013; Guna et al., 2018; Shurtleff et al., 2018), location (Aviram et al., 2016), or topology of the TMDs (Ast et al., 2013; Chitwood et al., 2018; Shurtleff et al., 2018).

SecA is another emerging bacterial protein biogenesis factor that can mediate the co-translational targeting and translocation of some of the membrane proteins (Huber et al., 2011; Singh et al., 2014; Wang et al., 2017). SecA binds to the ribosome near uL23 in proximity to the exit tunnel (Huber et al., 2011; Singh et al., 2014), and could be recruited to many membrane proteins during translation (Huber et al., 2016). The most well-characterized membrane protein substrate for co-translational delivery by SecA is RodZ (Rawat et al., 2015; Wang et al., 2017), a single pass type II membrane protein essential for cell division. SecA is necessary and sufficient for the

targeting of RodZ to the SecYEG translocon in a strictly co-translational mechanism *in vitro* and *in vivo* (Wang et al., 2017). SecA binds to RNCs bearing the RodZ nascent chain with high affinity ( $K_d \leq 1$  nM), and this binding survives the competition from other ribosome-associated protein biogenesis factors such as SRP and trigger factor (TF). The RodZ TMD is flanked by basic residues at the N-terminus and acidic residues at the C-terminus (net charge of -4), both of which are important for high affinity binding of SecA in preference over SRP (Wang et al., 2017). However, little is known about how SecA protects hydrophobic TMDs emerging from the ribosome exit tunnel, nor the molecular basis of its charge preferences during this recognition.

SecA was known to be an essential ATPase that drives the post-translational translocation of secretory proteins harboring less hydrophobic signal sequences across SecYEG (Hartl et al., 1990), often in collaboration with the chaperone SecB. In this post-translational mode, SecA binds the signal sequence via a hydrophobic groove in the pre-protein crosslinking (PPXD) domain (Gelís et al., 2007; Kimura et al., 1991). Another surface on SecA, Patch A, provides additional contact sites for hydrophobic segments in the mature regions of secretory proteins (Chatzi et al., 2017). In its recently described co-translational mode of targeting, it is unclear whether SecA uses the same preprotein binding sites to recognize nascent TMDs emerging from the ribosome. Furthermore, SecA binds with high affinity to anionic phospholipids and to SecYEG (Bauer and Rapoport, 2009; Zimmer et al., 2008), and previous structures (Frauenfeld et al., 2011; Park et al., 2014) suggested that SecA and the ribosome share partially overlapping binding sites on SecYEG. Biochemical data also indicate that SecA and the 70S ribosome compete for binding to SecYEG (Wu et al., 2012). How the co-translational recognition by SecA leads to the efficient delivery of nascent membrane proteins to SecYEG (Rawat et al., 2015; Wang et al., 2017) remains an outstanding puzzle.

To address these questions, we combined biochemical and structural analyses to study the molecular mechanism of this pathway. Site-specific crosslinking showed that the ribosome induces a distinct mode of nascent protein recognition by SecA. A cryoEM structure of SecA bound to RNC<sub>RodZ</sub> showed that the nascent TMD is sandwiched in a composite binding site formed by the N-terminal amphipathic helix of SecA and a hydrophobic groove on uL23 of the ribosome, and revealed the molecular basis for the charge preference during nascent protein

recognition by SecA. Finally, quantitative kinetic analyses demonstrate that SecYEG remodels the RNC-bound SecA to facilitate nascent protein transfer to SecYEG, and the transfer process is further facilitated by the elongation of the nascent polypeptide.

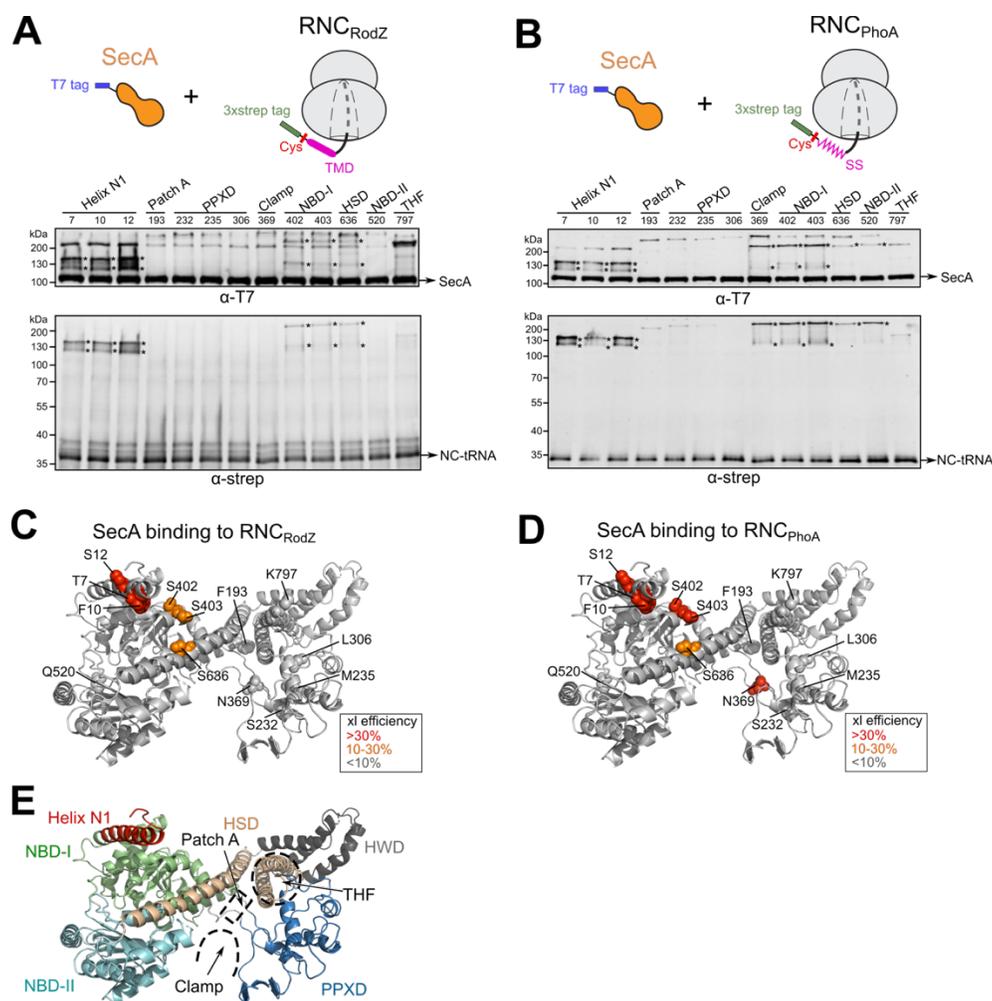
## 2.3 Results

### The ribosome promotes nascent protein interaction with the N-terminal amphipathic helix of SecA

As a model co-translational SecA substrate, we used the inner membrane protein RodZ (Rawat et al., 2015), which is co-translationally targeted and translocated by SecA *in vitro* and *in vivo* (Wang et al., 2017). To systematically probe how SecA interacts with the nascent RodZ TMD, we used thio-specific crosslinking with bismaleimido-hexane (BMH) to test the proximity between a single cysteine (C111) one residue upstream of the RodZ TMD and individual cysteines engineered at various positions on the SecA surface. All the engineered cysteine variants of SecA are functional in mediating the post-translational translocation of proOmpA (Figure S2.1A). We probed the co-translational nascent chain interactions of SecA using purified RNCs (Nakatogawa and Ito, 2002; Wang et al., 2017) bearing residues 104-160 of the RodZ nascent chain. The RNCs are stalled using the SecM arrest peptide (SecM residues 133-170) fused to the C-terminus of the RodZ nascent chain, so that the RodZ TMD (residues 112-132) is exposed outside the ribosome exit tunnel. The MreB-binding domain (MBD; residues 1-103) of RodZ is not essential for SecA recruitment (Wang et al., 2017) nor for the SecA-dependent membrane targeting and integration of RodZ (Figures S2.1B-D), and was therefore removed in this work.

The results of this cysteine scan revealed strong crosslinks between the RodZ nascent chain and the cysteines engineered at SecA residues 7, 10 and 12 (Figures 2.1A, 2.1C, and 2.1E), located in the conserved amphipathic helix (Koch et al., 2016) at the N-terminus of SecA (termed “helix N1”). Weaker but detectable crosslinks were also observed with the cysteines in nucleotide-binding domain-I (NBD-I, residues 34, 56, 402 and 403) and helical scaffold domain (HSD, residue 636) of SecA, all of which are within  $\sim 30$  Å of helix N1 (Figures 2.1A and S2.1G). These crosslinks are dependent on the presence of crosslinker, RNC<sub>RodZ</sub> and SecA (Figure

S2.1E). In contrast, no crosslinks were detectable with the cysteines engineered in the SecA PPXD domain (residues 232, 235, and 306) or in Patch A (residue 193), which are known to bind



**Figure 2.1. Thio-specific crosslinking to map SecA surface residues that contact hydrophobic sequences of nascent protein on the ribosome**

(A, B) Engineered single cysteines at indicated positions of SecA were tested for crosslinking by BMH to the TMD of RNC<sub>RodZ</sub> (A) or the signal sequence of RNC<sub>phoA</sub>. (B) SecA is T7-tagged at the flexible C-terminus. The RodZ (residues 104-160) or phoA (residues 1-52) nascent chain was fused to a C-terminal SecM stall sequence and an N-terminal 3xstrep tag, and contains a single cysteine upstream of the RodZ TMD (C111) or the phoA signal sequence (C4). Crosslinking reactions contained 1  $\mu$ M SecA and 500 nM RNC. Asterisks indicate crosslinking products detected by both the anti-T7 and anti-strep antibodies. ‘NC’, nascent chain. (C, D) Crosslinking efficiency from the data in parts A and B are mapped onto the structure of SecA generated in this work. Crosslinking efficiencies were calculated by dividing the amount of crosslinked nascent chain over the total amount of nascent chain, based on western-blot against the strep tag. Residues are shown in

spheres and colored based on crosslinking efficiency as indicated. (E) The individual domains are highlighted in SecA structure. The following color scheme was used: helix N1, red; NBD-I, green; NBD-II, cyan; PPXD, blue; HSD, wheat; HWD (helical wing domain), gray. Patch A, clamp and THF regions are highlighted by dashed lines.

hydrophobic segments on preproteins post-translationally (Gelís et al., 2007; Kimura et al., 1991). The SecA clamp region (residue 369) (Bauer and Rapoport, 2009) and the two-helix finger loop (THF, residue 797) (Erlandson et al., 2008) were previously reported to contact the translocating polypeptide, but they also failed to crosslink to the RodZ nascent chain.

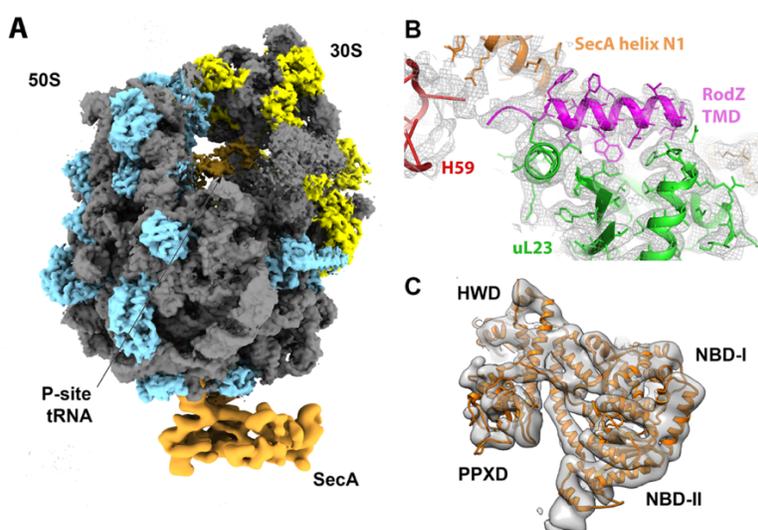
The following observations corroborated these interaction patterns of SecA during its co-translational recognition of nascent proteins. First, another thio-specific crosslinker bismaleimidoethane (BMOE), which has a shorter spacer length than BMH (8.0 Å vs. 13.0 Å, respectively), generated a similar crosslinking pattern between the RodZ nascent chain and SecA, albeit with slightly lower efficiency (Figure S2.1G). Moreover, RNC exposing a less hydrophobic signal sequence (SS) from the secretory protein, PhoA, exhibited a SecA crosslinking pattern similar to that of RNC<sub>RodZ</sub> (Figures 2.1B, 2.1D, and 2.1E), arguing against the notion that the observed crosslinking pattern is specific to the RodZ TMD. Finally, as an orthogonal approach to detect the distance between SecA and nascent chain on the ribosome, we incorporated a fluorescent amino acid, 7-hydroxycoumarin ethylglycine (Cm), immediately upstream of the RodZ TMD (residue 111) or the PhoA SS (residue 4) on the RNC. We monitored the Förster resonance energy transfer (FRET) between RNC<sub>RodZ</sub><sup>Cm</sup> or RNC<sub>PhoA</sub><sup>Cm</sup> and an acceptor dye, BODIPY-FL (BDP), incorporated at various positions on SecA (Figure S2.2A). The highest FRET efficiency was observed with BDP labeled at helix N1 of SecA, whereas BDP labeled at sites away from the N-terminus of SecA, such as PPXD and Patch A, exhibited low FRET efficiency with the Cm dye on the nascent chains (Figures S2.2A-C). Thus, helix N1 of SecA is the primary binding site for TMDs or signal sequences as the nascent polypeptide emerges from the ribosome.

These results are surprising, because previous work identified PPXD or Patch A as the sites used by SecA to interact with hydrophobic sequences on preprotein substrates. To test if the ribosome is responsible for this difference, we purified SUMO-RodZ and SUMO-PhoA fusion proteins, in

which the RodZ TMD or the PhoA SS was C-terminally fused to the SUMO protein, and probed the post-translational interactions of these proteins with SecA using thio-specific crosslinking. Both substrates crosslinked efficiently to the cysteines in SecA PPXD and/or Patch A, as well as multiple sites across all the domains of SecA (Figures S2.2D-G), consistent with previous observations using PhoA as the model post-translational substrate (Chatzi et al., 2017; Kimura et al., 1991; Or et al., 2002). Compared to the co-translational recognition mode of SecA, helix N1 was a less dominant binding site in the absence of ribosome. These results strongly suggest that the ribosome induces a distinct mode of substrate recognition by SecA and confines nascent protein interactions to its N-terminus.

### Structure of SecA bound to nascent RodZ on the ribosome

High-resolution structural information on a complex formed between the translating ribosome and SecA is still lacking. To better understand the interaction between the nascent chain and SecA during the co-translational targeting pathway, we set out to determine the cryo-EM structure of the RNC<sub>RodZ</sub>•SecA complex. Initial efforts to obtain a stable complex for cryo-EM did not yield high resolution information, suggesting that the binding of SecA on the ribosome is flexible. To increase the stability of this complex, BMH crosslinking was used to maintain the contact between SecA (C12) and a specific cysteine engineered upstream (C111) or downstream (C146) of the RodZ TMD (Figure S2.2H). RNC<sub>RodZ</sub> (C146), which gave the most efficient (45%) crosslink to SecA, was used for the cryo-EM studies.

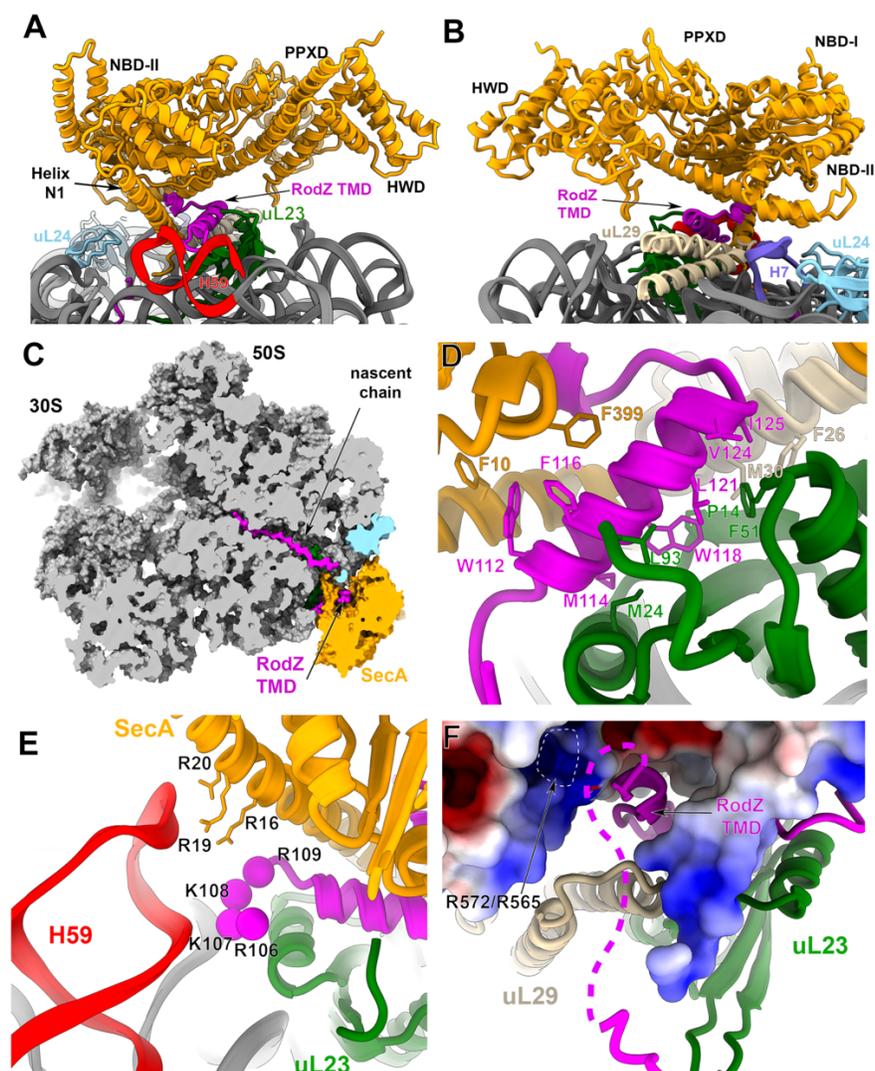


## Figure 2.2. Cryo-EM structure of the SecA•RNC<sub>RodZ</sub>

(A) Overview of the cryo-EM model of the SecA•RNC<sub>RodZ</sub> complex. Ribosomal RNA is in grey, ribosomal proteins in the large and small subunits are in blue and yellow, respectively, the P-site tRNA is in beige, and EM-densities of the ribosome and SecA are from the global refinement and focused refinement, respectively, and are filtered to the respective resolutions of 3.1 and 5.7 Å. (B) Representative fit of the RodZ TMD and the contact points between the ribosome and SecA with the overlaid EM-density (grey mesh). The ribosomal protein uL23 and the RodZ TMD are colored in green and magenta, respectively. (C) Representative fit of the SecA model with the overlaid EM-density obtained from the focused refinement approach.

Despite the high crosslinking efficiency, our initial structural analysis showed a low occupancy of SecA on the ribosome, which underscores the sensitivity of this complex under cryo-EM freezing conditions. We therefore collected a large dataset and employed an extensive focused 3D classification and refinement scheme, which resulted in a cryo-EM structure of the RNC<sub>RodZ</sub>•SecA complex at an overall resolution of 3.1 Å (Figures 2.2A, S2.3, and methods). The contact points between SecA and the ribosome as well as the RodZ TMD were resolved to side chain resolution (3.1 – 3.5 Å), which allowed us to build these regions *de novo* and assign the registry and directionality of the RodZ TMD (Figures 2.2B and S2.4). The local resolution of SecA was further improved using a 3D refinement scheme that focused on the SecA region and masked out the rest of the ribosome. This strategy yielded a SecA structure at a local resolution of 5.7 Å, where secondary structural elements can be clearly resolved (Figures 2.2C and S2.4). This also allowed us to manually adjust the  $\alpha$ -helices and place the PPXD of SecA, which is known from previous structural and biochemical studies to adopt multiple conformations (Osborne et al., 2004; Sharma et al., 2003; Zimmer et al., 2008), as a rigid body into the EM density.

The density of SecA covers the ribosome exit tunnel and lies parallel to the ribosome surface, with NBD-I most proximal to the exit tunnel, whereas PPXD and HWD points away (Figures 2.3A and 2.3B). Interactions between SecA and the ribosome are mediated exclusively through contacts with regions of 23S rRNAs in the vicinity of ribosomal proteins uL23, uL24, and uL29, consistent with previous observations (Huber et al., 2011; Singh et al., 2014). Specifically, rRNA H59 contacts basic amino acids (R16, R19, R20) on the positively charged face of the amphipathic helix N1 of SecA, which extends down from NBD-I towards the ribosomal exit



**Figure 2.3. Snapshots of the SecA•RNC<sub>RodZ</sub> model**

(A) and (B) Close-ups of the contact points between SecA and the surface of the ribosome. The two views are rotated relative to each other by 180° along the vertical axis of the model. (C) Surface representation of the SecA•RNC<sub>RodZ</sub> model with a cross-section of the ribosome tunnel region where the nascent chain and the TMD of RodZ can be visualized. (D) Close-up view of the composite TMD binding pocket formed by uL23 and SecA. uL29 may also contribute to this binding pocket. Hydrophobic residues that contact the RodZ TMD are highlighted in sticks. (E) 23S rRNA H59 contacts basic residues (spheres) preceding the RodZ TMD and basic residues (sticks) on the hydrophilic side of SecA helix N1. (F) Sequence downstream of the RodZ TMD (dashed line) may be positioned in the vicinity of a basic patch on SecA, which is shown as electrostatic surface. The basic surface provided by R572 and R565 is highlighted. The following color scheme was used: uL23, green; uL24, cyan; uL29, wheat; SecA, orange; RodZ TMD, magenta; H59, red; H7, deep blue.

tunnel (Figures 2.3A, 2.3E, 2.4A and S2.5A). In addition, 23S rRNA H7 contacts basic residues (R602, K609) on NBD-II of SecA (Figure 2.4A, blue).

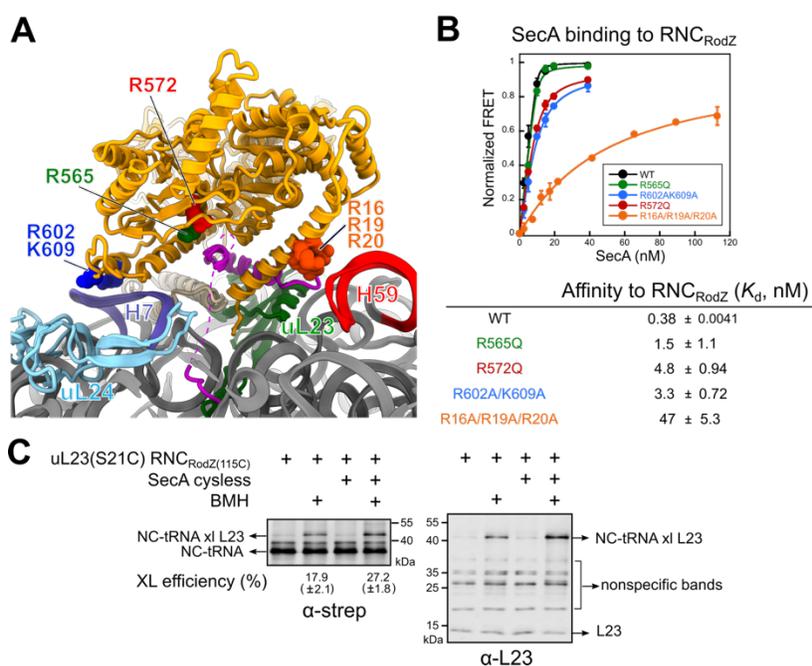
The RodZ nascent chain is resolved within the ribosome tunnel from the CAA end of the P-site tRNA to the tunnel exit (Figures 2.3C and S2.5C). The RodZ TMD is bound within a composite pocket contributed by: (1) residues (P14, M24, F51, L93) from ribosomal protein uL23, (2) the hydrophobic face (L2, I3, L5-6, F10) of the amphipathic helix N1 of SecA, (3) residue F399 in NBD-I of SecA, and (4) residue F639 in HSD of SecA (Figures 2.3D, S2.5A, and S2.5B). Residues M30 and F26 of uL29 may also contribute to part of the binding pocket. These interactions shield the nascent TMD from the aqueous cytosolic environment prior to membrane insertion. The RodZ TMD is preceded by six consecutive positively charged residues, four of which are resolved and contact the 23S rRNA at H59 (Figure 2.3E) in a mode similar to previous observations with an SRP-bound signal sequence on the RNC (Jomaa et al., 2016). This observation implicates a potential role of H59 in multiple co-translational targeting pathways.

To test the role of the SecA-ribosome contacts observed in the structure, we measured the equilibrium dissociation constant ( $K_d$ ) of RNC<sub>RodZ</sub> bound to various SecA mutants based on FRET between Cm-labeled RNC<sub>RodZ</sub> and BDP-labeled SecA described above (Figures S2.2A). Mutation of the basic residues (R16, R19 and R20) on SecA that contact 23S rRNA H59 reduced its binding affinity for RNC<sub>RodZ</sub> over 100-fold (Figures 2.4A and 4B, orange), indicating the essential role of this contact in stabilizing the SecA-RNC interaction. Mutation of R602 and K609 in SecA that contact 23S rRNA H7 had a more modest effect, ~10-fold (Figures 2.4A and 2.4B, blue), suggesting this to be an ancillary ribosome contact site.

We previously showed that the hydrophobic TMD of RodZ as well as enrichment of basic and acidic residues N- and C-terminal to the TMD, respectively, are important for high affinity binding between SecA and RNC<sub>RodZ</sub> (Wang et al., 2017). While the basic residues N-terminal to the RodZ TMD were resolved in the structure and contact rRNA H59, the region C-terminal to

the RodZ TMD was not resolved. Nevertheless, the C-terminus of the RodZ TMD points towards a SecA surface rich in positively charged residues that could provide a contact site for the acidic sequence C-terminal to the RodZ TMD (Figure 2.3F). In support of this model, conservative mutation of two Arg residues in this surface (565 or 572) to Gln each caused a 4-12 fold weakened binding of SecA to RNC<sub>RodZ</sub> (Figures 2.4A and 2.4B, green and red). These results provide a structural basis to explain the charge preferences of SecA during its co-translational nascent chain recognition.

The composite TMD binding pocket formed by both uL23 and SecA observed in the structure is in good agreement with our crosslinking data that SecA residues 1, 7, 12, 402, 403, and 636 are in close proximity to the RodZ TMD on the ribosome. To test whether the hydrophobic cleft on uL23 provides a potential docking site for the nascent TMD, we deleted the genomic uL23 and supplied a single cysteine mutant of uL23 (C21) on a plasmid. We purified RNC<sub>RodZ</sub> harboring uL23 (C21) and a cysteine at residue 115 in the RodZ TMD. Addition of BMH induced a significant crosslink between uL23 and the RodZ nascent chain, and the presence of SecA increased the crosslinking efficiency by ~1.5 fold (Figure 2.4C). These results suggest that nascent TMD has an intrinsic preference to dock at uL23, and this interaction is further stabilized by SecA.



### Figure 2.4. Validation of the interaction sites between RNC<sub>RodZ</sub> and SecA

(A) SecA mutations tested biochemically are highlighted in spheres in the SecA•RNC<sub>RodZ</sub> model. Residues are shown in spheres and colored as in (B). The remainder of the structure is colored as in Figure 2.3.

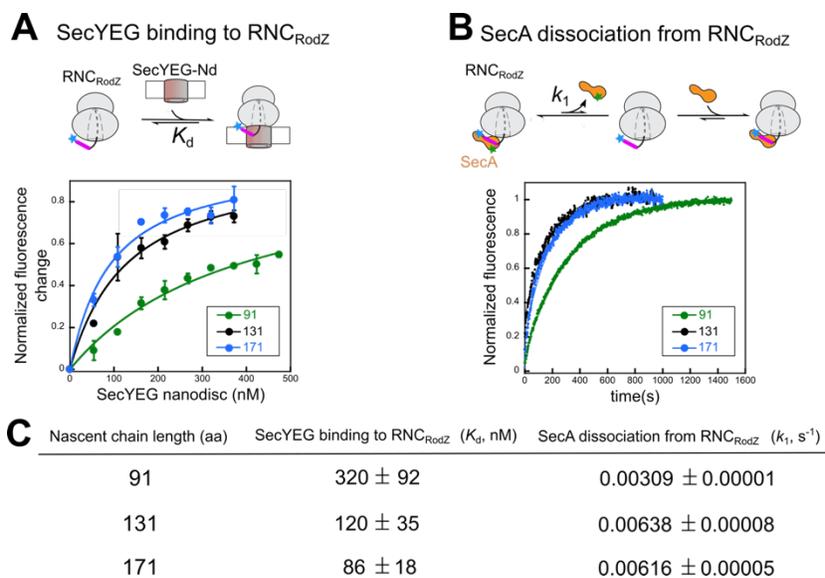
(B) Equilibrium titrations to measure the  $K_d$  values for the binding of RNC<sub>RodZ</sub> to wild type (WT) SecA and indicated SecA mutants. Cm was incorporated at residue 111 upstream of the RodZ TMD on RNC. All SecA variants were labeled by BDP at residue 12. Lines are fits of the data to Eq. 2, and the obtained  $K_d$  values are summarized in the lower panel. All values are reported as mean  $\pm$  SD, with  $n = 2-3$ .

(C) Crosslinking between a pair of engineered cysteines at residue 115 in the RodZ TMD and residue 21 of uL23 in RNC<sub>RodZ</sub>. Crosslinking was induced by BMH, and the crosslinked product was detected by western blot with anti-strep and anti-L23 antibodies. The numbers underneath the  $\square$ -strep blot indicate crosslinking efficiency, calculated from the ratio of the intensity of crosslinked bands relative to the total intensity of tRNA-linked nascent chain (NC-tRNA). XL, crosslinking. All values are reported as mean  $\pm$  SD, with  $n = 2$ .

### SecYEG and nascent chain elongation facilitate RNC handover from SecA to SecYEG

The results above provided the molecular basis for the initial recognition of RNC<sub>RodZ</sub> by SecA, which subsequently targets the RodZ nascent chain to SecYEG for membrane integration in a strictly co-translational pathway (Rawat et al., 2015; Wang et al., 2017). Intriguingly, our structure suggested that the RNC-binding surface of SecA heavily overlaps with its anionic phospholipid interaction surface (Koch et al., 2016) and was on the same face as its SecYEG docking site (Li et al., 2016) (Figure S2.6A). This raises questions as to how the SecA-bound RNC<sub>RodZ</sub> is delivered to SecYEG.

To address this question, we first asked whether the nascent chain length affects the interaction of RNC<sub>RodZ</sub> with SecYEG embedded in the phospholipid bilayer. We prepared fluorescently labeled RNC<sub>RodZ</sub><sup>Cm</sup> with three nascent chain lengths: RNC<sub>RodZ91</sub> (the same construct characterized above), RNC<sub>RodZ131</sub>, and RNC<sub>RodZ171</sub>. We reconstituted purified SecYEG complex in large lipid nanodiscs (Nd) formed by the ApoE422K scaffold protein (Figures S2.6B-E) (Koch et al., 2016). Interaction with SecYEG-Nd leads to quenching of the Cm fluorescence on RNC<sub>RodZ</sub><sup>Cm</sup> (Figure S2.6F, lanes 1 and 4), which was used to measure RNC-SecYEG binding. SecYEG-Nd binds RNC<sub>RodZ91</sub> with modest affinity, and this binding affinity increased 3-4 fold with RNC<sub>RodZ131</sub> and RNC<sub>RodZ171</sub> (Figures 2.5A and 5C). To further test if elongation of the nascent chain impacts the interaction of SecA with RNC<sub>RodZ</sub>, we measured the rate constant of SecA dissociation from

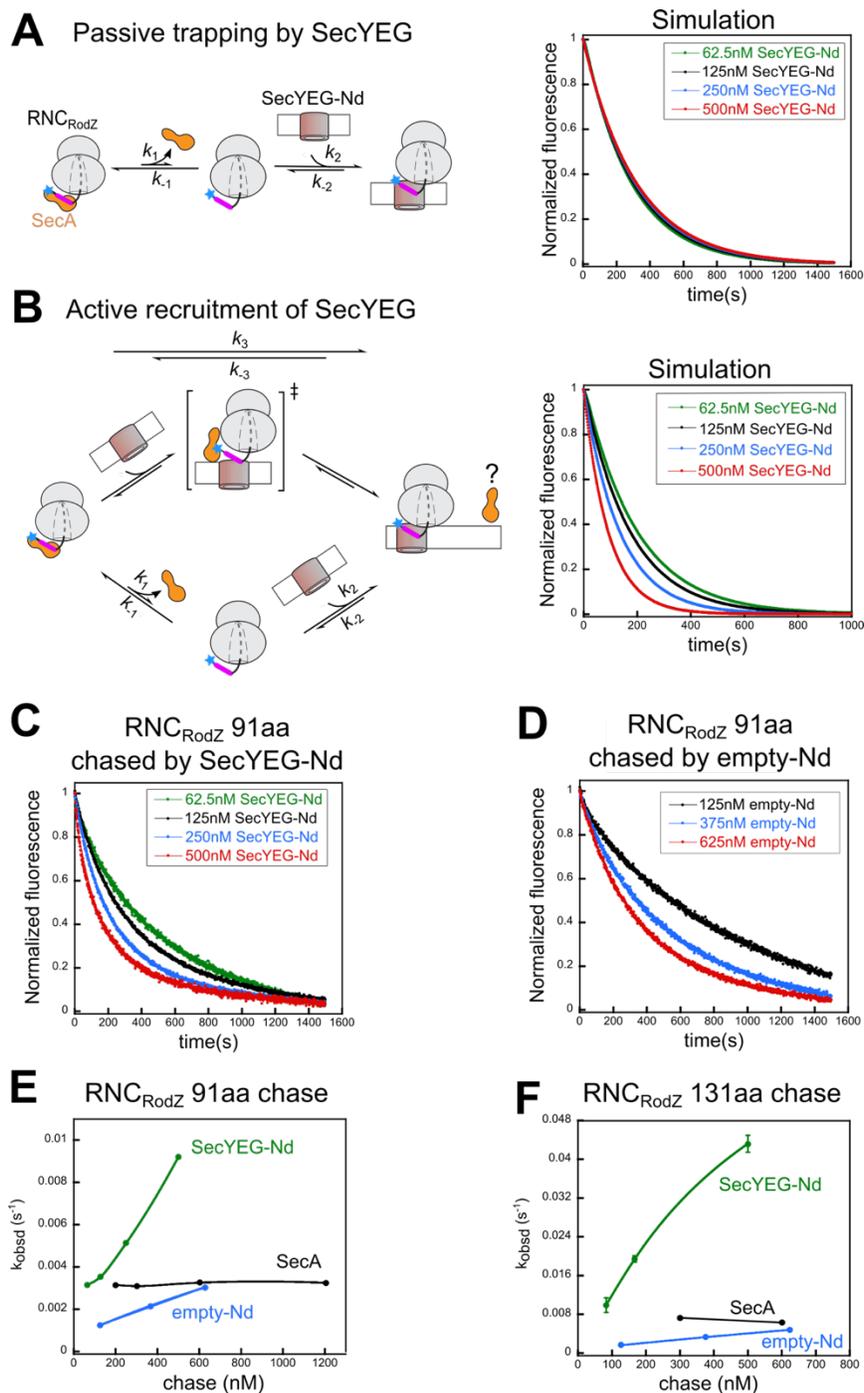


**Figure 2.5. Effects of nascent chain length on SecYEG binding to and SecA dissociation from RNC<sub>RodZ</sub>**

(A) Equilibrium titrations to measure the  $K_d$  value of the SecYEG•RNC<sub>RodZ</sub> complex. Top panel, scheme of the assay: RNC<sub>RodZ</sub> labeled with Cm (blue star) at RodZ residue 111 upstream of its TMD (magenta) was incubated with indicated concentrations of SecYEG-Nd, prepared and quantified as described in the methods. Bottom panel, representative titration curves of SecYEG binding to RNC<sub>RodZ</sub> at different nascent chain lengths. Complex formation was monitored by quenching of the Cm fluorescence by SecYEG-Nd (Figure S2.6F, lanes 1 and 4). The lines are fits of the data to Eq. 2, and the obtained  $K_d$  values are summarized in part C. (B) Measurement of the dissociation rate constant ( $k_1$ ) of SecA from RNC<sub>RodZ</sub>. Top panel, scheme of the assay: RNC<sub>RodZ</sub> labeled as in (A) was pre-incubated with SecA labeled with BDP (green star) at residue 12, and excess (300 nM) unlabeled SecA was added to chase the FRET signal (Figure S2.6F, lanes 2 and 3). Bottom panel, representative time courses of the chase reactions for RNC<sub>RodZ</sub> with indicated nascent chain lengths. The data were fit to Eq. 3, and the obtained  $k_1$  values are summarized in part C. (C) Summary of the  $K_d$  and  $k_1$  values measured from the experiments in parts A, B and their replicates. All values represent mean ± SD, with  $n = 2-3$ .

RNC<sub>RodZ</sub> ( $k_1$ ) based on the loss of FRET from a preformed RNC<sub>RodZ</sub><sup>Cm</sup>•SecA<sup>BDP</sup> complex upon chase with excess unlabeled SecA (Figures 2.5B and S2.6F, lanes 2 and 3). SecA dissociation from RNC<sub>RodZ91</sub> was slow, with a  $k_1$  of 0.0031 s<sup>-1</sup>, and was accelerated ~2-fold with longer nascent chain (Figures 2.5B and 2.5C). Thus, elongation of the nascent chain modestly enhances the binding of RNC to membrane-embedded SecYEG and reduces the kinetic stability of the

RNC<sub>RodZ</sub>•SecA complex, which could potentiate the nascent chain for handover from SecA to SecYEG.



**Figure 2.6. Measurements to distinguish different mechanisms of RNC<sub>RodZ</sub> delivery to SecYEG by SecA**

(A) Scheme (left panel) and kinetic simulations (right panel) for the model in which SecYEG passively binds RNC<sub>RodZ</sub> that has dissociated from SecA. RodZ TMD is colored in magenta. (B) Scheme (left panel) and

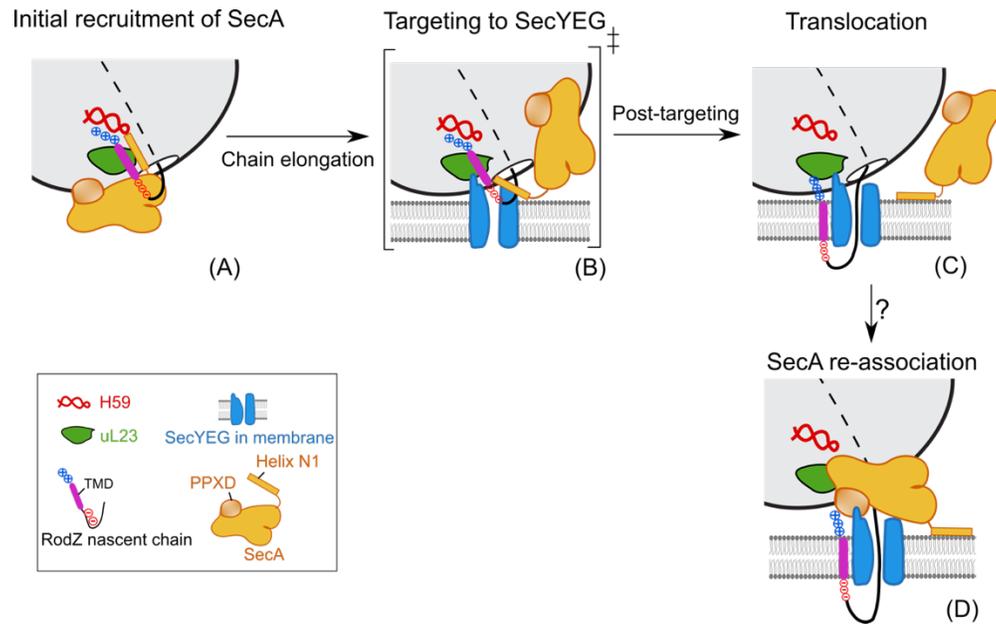
kinetic simulations (right panel) for the model in which SecYEG forms a ternary complex with SecA-bound RNC<sub>RodZ</sub> to actively displace SecA (upper route). The passive pathway was included in the simulation for completeness (lower route). The ‘?’ denotes that it is unclear whether SecA is completely displaced after the transfer, or significantly repositioned relative to its initial mode of interaction with the RNC. (C) Representative fluorescence time traces to measure the transfer of RNC<sub>RodZ</sub> at a chain length of 91aa. RNC<sub>RodZ</sub> labeled with Cm (blue star) at RodZ residue 111 was pre-incubated with unlabeled SecA, and challenged with indicated concentrations of SecYEG-Nd. Docking of RNC<sub>RodZ</sub> onto SecYEG-Nd was monitored by quenching of Cm fluorescence (Figure S2.6F, lanes 5 and 7). Each trace is the average of 6-8 measurements. The data were fit to Eq. 3 to obtain the apparent rate constant ( $k_{\text{obsd}}$ ) of transfer. Note that the time traces are biphasic, and control experiment indicated that the slow phase was due to dye bleaching (see Methods). (D) Same as in (C), except that empty nanodisc was used instead of SecYEG-Nd. (E, F) Summary of the observed rate constants for transfer of SecA-bound RNC<sub>RodZ</sub> to SecYEG-Nd (green) or to empty nanodiscs (blue), and for spontaneous dissociation of SecA from RNC<sub>RodZ</sub> (black) at RodZ nascent chain lengths of 91 aa (E) and 131 aa (F). Error bar may not be visible. All values represent mean  $\pm$  SD, with  $n = 2-3$ .

The slow SecA dissociation from RNC<sub>RodZ</sub> ( $t_{1/2} = 108-220$  s) raised questions as to how RodZ targeting occurs with kinetic competence; for comparison, co-translational targeting mediated by SRP occurs in less than 5 s (Zhang and Shan, 2014). We therefore tested alternative models for how RNC pre-bound by SecA could be delivered to SecYEG. If RNC<sub>RodZ</sub> must dissociate from SecA before it binds SecYEG (Figure 2.6A, “passive” model on the left), the rate of RNC<sub>RodZ</sub> engagement with SecYEG-Nd would be limited by the slow dissociation of SecA from RNC<sub>RodZ</sub> and independent of the concentration of SecYEG-Nd (Figure 2.6A, simulation results on the right and Figures S2.7C-D). In contrast, if SecYEG directly associates with SecA•RNC<sub>RodZ</sub> and alters its conformation to facilitate the handover (Figure 2.6B, “active recruitment” model on the left), the transfer reaction will be accelerated by increasing concentrations of SecYEG-Nd and significantly faster than spontaneous SecA dissociation from RNC<sub>RodZ</sub> (Figure 2.6B, simulations on the right and Figures S2.7E-F).

To distinguish between these models, we pre-formed a complex of RNC<sub>RodZ</sub><sup>Cm</sup> with unlabeled SecA and challenged the complex with varying concentrations of SecYEG-Nd. As the fluorescence intensity of RNC<sub>RodZ</sub><sup>Cm</sup> was unaffected by unlabeled SecA but was quenched when bound to SecYEG-Nd (Figure S2.6F, lanes 1, 5 and 7), transfer of RNC<sub>RodZ</sub><sup>Cm</sup> to SecYEG was monitored by quenching of Cm fluorescence (Figure 2.6C). At the end of the transfer reaction,

quenching of the Cm fluorescence on  $\text{RNC}_{\text{RodZ}}^{\text{Cm}}$  was the same, within error, as that obtained after direct binding of  $\text{RNC}_{\text{RodZ}}^{\text{Cm}}$  to SecYEG-Nd (Figure S2.6F, lanes 4 vs. 7). In addition, the Cm fluorescence intensity was similar regardless of whether  $\text{RNC}_{\text{RodZ}}^{\text{Cm}}$  was prebound to  $\text{SecA}^{\text{BDP}}$  or unlabeled SecA (Figure S2.6F, lanes 6 vs. 7), indicating that FRET between  $\text{RNC}_{\text{RodZ}}^{\text{Cm}}$  and  $\text{SecA}^{\text{BDP}}$  was lost. These observations strongly suggest that at the end of the transfer reaction, the Cm dye on  $\text{RNC}_{\text{RodZ}}^{\text{Cm}}$  was embedded in an environment dominated by SecYEG-Nd, whereas SecA was displaced from its initial position on  $\text{RNC}_{\text{RodZ}}$ .

With  $\text{RNC}_{\text{RodZ}91}$ , the observed transfer rate constant ( $k_{\text{obsd}}$ ) was accelerated by increasing concentrations of SecYEG-Nd (Figures 2.6C and 2.6E) and became significantly faster than spontaneous SecA dissociation at SecYEG-Nd concentrations above 300 nM (Figure 2.6E, green vs black), indicating that the active model became the dominant pathway for the delivery and transfer of  $\text{RNC}_{\text{RodZ}}$  at moderate SecYEG concentrations. Control reactions using empty nanodiscs were significantly slower than those using SecYEG-Nd, indicating a role of SecYEG in the accelerated transfer (Figures 2.6D, 2.6E and S2.6F, lanes 5 and 8). As another negative control, spontaneous dissociation of SecA from  $\text{RNC}_{\text{RodZ}}$  was independent of chase concentration (Figures 2.6E and 2.6F, black; Figure S2.7H), as would be expected for a unimolecular reaction (Figure S2.7G). Analogous results were observed for transfer of  $\text{RNC}_{\text{RodZ}131}$  to SecYEG-Nd (Figures 2.6F and S2.7, I-K), except that the longer nascent chain length increase the rate of SecYEG-mediated transfer reaction by 3-4 fold (Figure 2.6, E vs F). As the reconstituted SecYEG-Nd contains ~0.55 copy of SecYEG per copy of nanodisc (Figure S2.6C), the observed transfer kinetics were a lower estimate of the efficiency at which SecYEG stimulates RNC transfer. Finally, consideration of the reaction equilibrium indicated that the  $\text{RNC}_{\text{RodZ}} \cdot \text{SecYEG}$  complexes generated during the transfer reaction were much more stable than those obtained from binding of free  $\text{RNC}_{\text{RodZ}}$  to SecYEG (supplementary discussion), providing additional evidence that transfer occurred via an active mechanism that bypasses the formation of free  $\text{RNC}_{\text{RodZ}}$ . Together, these results showed that SecYEG is recruited to SecA-bound  $\text{RNC}_{\text{RodZ}}$  and actively promotes transfer of the nascent protein from SecA to this translocase, and this process is further facilitated by elongation of the nascent polypeptide.



**Figure 2.7 Model of SecA-mediated co-translational recognition and targeting of membrane proteins.**

(A) Nascent TMD emerges from exit tunnel and recruits SecA. (B) SecYEG invades the SecA-bound RNC, generating a transient ternary complex in which SecA is repositioned and its contacts with the RNC are weakened. (C) The nascent TMD is released from SecA and docks onto SecYEG, which initiates translocation. SecA could remain bound to the membrane surface via helix N1. (D) As the nascent chain further elongates, SecA could also re-associate with the translocation complex and use its ATPase cycle to drive translocation of the periplasmic domain. The “?” denotes that it is unclear whether step (D) occurs, nor what the molecular signals are that trigger SecA re-association.

## 2.4 Discussion

Emerging data indicate that nascent membrane proteins are co-translationally delivered to and inserted into their membrane destinations via diverse pathways in both prokaryotic and eukaryotic organisms (Aviram et al., 2016; Guna et al., 2018; Zhang et al., 2010). Besides the well-studied co-translational SRP pathway, little is known about the molecular mechanism of the alternative co-translational targeting pathways. This work elucidates the molecular basis of co-translational nascent membrane protein recognition and delivery by SecA in bacteria. We show that the hydrophobic TMD on the nascent polypeptide emerging from the ribosome is sandwiched in a composite binding site formed by both SecA and the ribosomal protein uL23, and explain the structural basis for the charge preference of SecA during its co-translational recruitment. Furthermore, the SecYEG complex in the membrane can associate with and actively remodel the SecA-bound RNC, which together with elongation of the nascent polypeptide facilitates handover of the RNC to the membrane translocon.

The structure here provides a precedent for active participation of the ribosome exit site in forming a shared nascent TMD docking site with a ribosome-associated protein biogenesis factor (RPB). This is distinct from previous observations where the ribosome simply provides a docking site for an RPB (SRP or TF for example) which is responsible for interacting with the nascent polypeptide (Ferbitz et al., 2004; Jomaa et al., 2016; Knorr et al., 2019). The position of the nascent TMD was also distinct from that in the RNC•SRP structure (Figure S2.8). The observation of crosslink between the RodZ nascent chain and the hydrophobic groove of uL23 suggests that uL23 provides a transient early binding site to shield hydrophobic sequences on the nascent polypeptide before the latter engages with an RPB. This interaction could be further regulated by the RPBs, as was observed here for SecA. In addition, H59 of the 23S rRNA provides a contact site for basic residues upstream of a hydrophobic TMD (this work) or signal sequence (Jomaa et al., 2016; Jomaa et al., 2017), raising the possibility that H59 acts as a hub to select for enrichment of basic residues in the targeting sequences of membrane and secretory proteins.

The biochemical and structural work here also demonstrate that the ribosome induces a distinct mode of nascent protein recognition by SecA. On SecA-bound RNC, the substrate recognition

site is confined to the N-terminal region of SecA, whereas the previously identified preprotein binding sites on SecA, PPXD and Patch A, are not involved in recognition. In contrast, the same TMD or signal sequence can crosslink to PPXD, Patch A and multiple other sites on SecA in the absence of the ribosome (Figures 2.1 and S2.2). These observations strongly suggest that SecA can alternate between two modes of substrate recognition: the post-translational mode defined previously, and the co-translational mode described in this work. In addition to the ribosome, multiple determinants, such as hydrophobicity of the TMD and enrichment of charges flanking the TMD, could bias the relative energetics and hence the selection of SecA's recognition modes.

The ribosome exit tunnel is a crowded environment where multiple RPBs in addition to SecA can dock and access the nascent polypeptide (Akopian et al., 2013b). Although SecA can provide a docking site for both TMDs and signal sequences emerging from the ribosome, a variety of factors, such as the affinity of SecA for the RNC and regulation by other RPBs, likely dictate which substrates enter the SecA-mediated co-translational targeting pathway *in vivo*. We previously showed that, although both SRP and SecA can recognize the RodZ TMD, the RodZ N-terminal element preceding its TMD weakens SRP binding (Wang et al., 2017). Likewise, although SecA by itself binds with reasonable affinity to RNCs exposing the TMD of FtsQ, an SRP substrate, this interaction did not withstand competition from SRP. In addition, RNC bearing the nascent chain of PhoA, a post-translational SecA substrate, binds 200-fold more weakly to SecA than RNC<sub>RodZ</sub>, likely due to the lower hydrophobicity of its signal sequence and the lack of C-terminal acidic residues. These and other observations (Ariosa et al., 2014; Ranjan et al., 2017; Zhang et al., 2010) suggest that multiple sequence and structural elements on the nascent polypeptide, the preferential recognition of each element by individual RPBs, and competition or regulation by other RPBs together dictate the selection of nascent proteins into distinct biogenesis pathways. Reciprocally, the diversity of protein targeting factors and the dual mode of substrate recognition by SecA could accommodate the targeting needs of diverse nascent proteins with different hydrophobicity, charge distribution, and TMD location or topology.

The available structures show that SecA- and SecYEG-bound RNCs share multiple overlapping binding sites. Helix N1 of SecA, which provides an important TMD- and ribosome binding site,

is also vital for its interaction with anionic phospholipids. There are also extensive overlaps in the SecA and ribosome docking sites on SecYEG as well as the SecA and SecYEG binding sites on the ribosome (Figure S2.8). Given these overlaps, it is puzzling how SecA-bound RNCs are targeted to the SecYEG translocon. Our data here provide kinetic evidence that SecYEG can directly associate with SecA-bound RNCs to generate a transient ternary intermediate in which the SecA-RNC interaction is weakened, allowing facilitated transfer of the RNC to SecYEG in a concerted pathway. Such a mechanism is possible, in part due to the multiple interaction sites of SecA on the ribosome. We showed here that basic residues near the N-terminus (R16/R19/R20) and in the HSD (R602/K609) of SecA both contribute to ribosome binding; a previous work suggested that additional basic residues in the SecA HSD (K625/R633) may also provide a ribosome contact site (Huber et al., 2011). The multi-dentate, electrostatically driven interaction could allow SecYEG to ‘invade’ part of the SecA-ribosome interaction surface without waiting for complete SecA dissociation, thus generating an accelerated path for RNC handover. These observations resonate with those during SRP-dependent co-translational protein targeting, in which RNCs pre-bound to SRP and SR are transferred to SecYEG in a concerted mechanism involving a major rearrangement of the SRP•SR complex on the ribosome (Jomaa et al., 2016; Saraogi et al., 2014). Such a concerted mechanism minimizes the loss of RNC during the handover while allowing a kinetically more facile path for the transfer, and could represent a general mechanism in membrane protein biogenesis pathways.

We propose the following model for SecA-mediated co-translational protein targeting (Figure 2.7). SecA is recruited co-translationally to nascent proteins emerging from the ribosome via multiple interactions, including recognition of the TMD via the composite binding site formed by SecA helix N1 and the hydrophobic groove on uL23, and recognition of charged residues flanking the TMD via H59 of 23S rRNA and the basic SecA surface near helix N1 (Figure 2.7A). SecYEG associates with and actively remodels the RNC•SecA complex, generating a transient intermediate in which SecA is repositioned on the RNC with weakened contacts (Figure 2.7B), thus facilitating handover of the TMD from SecA to SecYEG (Figure 2.7C). The hydrophobic groove of uL23 and H59 could help stabilize the nascent TMD during the handover, and the interaction of SecA helix N1 with anionic phospholipids could also promote its repositioning in this intermediate (Figure 2.7B). After the nascent polypeptide docks onto

SecYEG and initiates translocation, SecA may remain bound at the membrane via helix N1 (Koch et al., 2016) (Figure 2.7C). As the ribosome-translocon junction can be transiently disrupted during membrane protein integration (Devaraneni et al., 2011), and as SecA is required for the translocation of membrane proteins containing large periplasmic loops (Gebert et al., 1988; Sääf et al., 1995; Traxler and Murphy, 1996), it is plausible that SecA could reassociate with the translocation complex as periplasmic domains emerge from the ribosome and use its ATPase cycle to drive translocation (Figure 2.7D).

## 2.5 Methods

### Protein expression and purification

N-terminally His<sub>6</sub>-tagged *E. coli* SecA contains a mutation (C98S) to remove the surface exposed cysteine, as described before (Wang et al., 2017). For SecA used in crosslinking and cryoEM studies, the three cysteines ( $\Delta$ 885-896) at the non-essential (Zimmer et al., 2008) C-terminus of SecA was also removed. SecA variants used for crosslinking and fluorescent analyses were expressed and purified as described before (Wang et al., 2017). SecA used for cryoEM was further purified by size exclusion chromatography on Superdex 200 10/300 GL (GE healthcare) in buffer containing 50 mM KHEPES, pH 7.5, 150 mM KOAc, 10 mM Mg(OAc)<sub>2</sub>, 2 mM DTT.

SecYEG containing N-terminally His<sub>6</sub>-tagged SecY was expressed and purified as described before (Akopian et al., 2013a) with slight modifications. Cells were induced at log phase by 0.5 mM IPTG for 3 hrs at 37 °C. Harvested cells were resuspended in KC300G buffer (50 mM KHEPES pH 7.5, 300 mM NaCl, 10% glycerol) and lysed by sonication. Lysate was clarified by centrifugation at 12,000 g, 4 °C for 20 min in JA 20 rotor (Beckman Coulter). The supernatant was ultracentrifuged at 42,000 rpm, 4 °C in Ti70 rotor (Beckman Coulter) for 50 min. The membrane fraction was resuspended in KS200G buffer (50 mM KHEPES pH 7.5, 200 mM NaCl, 10% glycerol) by dounce homogenizer (Wheaton). N-Dodecyl- $\beta$ -D-Maltopyranoside (DDM) and KS200G buffer were added to the membrane suspension to adjust the total protein concentration to 10 mg/ml and DDM concentration to 10% (w/w). The suspension was clarified by ultracentrifugation at 42,000 rpm for 50 min in Ti 70 rotor, and purified by Ni-NTA agarose. Protein was loaded and washed in SecYEG buffer 1 (50 mM KHEPES pH 7.5, 200 mM NaCl, 10% glycerol, 10 mM imidazole, 0.02% DDM), and eluted in SecYEG buffer 2 (50 mM KHEPES pH 7.5, 200 mM NaCl, 10% glycerol, 300 mM imidazole, 0.02% DDM). Eluted protein was dialyzed into KS50G (50 mM KHEPES pH 7.5, 50 mM NaCl, 10% glycerol, 0.02% DDM), loaded onto SP sepharose Fast Flow (GE healthcare) in KS50G, and eluted using a gradient of 50-1000 mM NaCl. Protein fractions were pooled and concentrated to ~75  $\mu$ M using Amicon, 50K MWCO centrifugal filter unit (MilliporeSigma).

SecB, SRP, FtsY, trigger factor, and ApoE422K were expressed and purified as described (Koch et al., 2016; Wang et al., 2017). SUMO fusions to the TMD of RodZ (residue 104-133) or signal sequence of PhoA (residues 1-21) were expressed and purified as described previously (Wang et al., 2017).

### **RNC preparation**

Stalled RNCs were generated by in vitro translation in S30 extract as described previously (Wang et al., 2017). 7-hydroxycoumaryl ethylglycine (Cm)-labeled RNCs were prepared similarly, except that the translation reaction was supplemented with Cm (Bachem), tRNA<sup>Cm</sup>, and Cm tRNA synthetase, as described before (Saraogi et al., 2011; Wang et al., 2017). RNC containing an engineered cysteine at uL23 was prepared using S30 extracts from strain KC624 harboring uL23(S21C) (KC6  $\Delta rplW::kan$  pL23<sub>S21C</sub>). KC624 was made from *E. coli* strain KC6 (Calhoun and Swartz, 2006), by transforming with the plasmid pEK20 containing single cysteine mutant uL23, and subsequent knocking out genomic uL23 via lambda-red recombination (Datsenko and Wanner, 2000). RNC for cryoEM study was further purified by sucrose gradient (10-50 %) to enrich monosome.

### **Crosslinking**

SecA, RNCs, and SUMO fusion proteins were buffer exchanged into labeling buffer A (50 mM KHEPES, pH 7.0, 150 mM KOAc, 10 mM Mg(OAc)<sub>2</sub>, 5 mM TCEP). Crosslinking reactions between SecA and RNC were performed in labeling buffer A and contained 1  $\mu$ M SecA, 0.5  $\mu$ M RNC, and 0.2 mM BMH (or BMOE). Crosslinking reaction between SecA and SUMO fusion proteins were performed in labeling buffer B (20 mM KHEPES, pH 7.0, 50 mM KOAc, 2 mM Mg(OAc)<sub>2</sub>, 5 mM TCEP) and contained 1  $\mu$ M SecA, 8.3  $\mu$ M SUMO fusions, and 2 mM BMH. All crosslinking reactions were carried out at room temperature for 45 min, and quenched by addition of DTT to 100 mM. Samples were analyzed by SDS-PAGE and western blotting using the indicated antibodies.

### **Cryo-EM data collection**

RNCs (500 nM) were mixed with SecA (1  $\mu$ M) in the presence of 0.2 mM bismaleimido-hexane (BMH) for 50 minutes protected from light and at room temperature. The final reaction size was

10  $\mu\text{l}$  in buffer C (50 mM HEPES-KOH pH 7.4, 100 mM KOAc, 10 mM  $\text{Mg}(\text{OAc})_2$ ). After quenching the reaction with 100 mM Dithiothreitol (DTT) for 5 minutes, reaction was diluted to 60  $\mu\text{l}$  in buffer C to a final concentration of 80 nM of RNCs and cooled on ice for 15 minutes. The sample (5  $\mu\text{l}$ ) was then applied to Quantifoil grids (R2/2) freshly coated with a thin layer of carbon, incubated for one minute before plunge frozen into a liquid ethane/propane mix cooled to liquid nitrogen temperature using a Vitrobot Mark IV at 95% relative humidity and previously cooled to 4  $^{\circ}\text{C}$ . Cryo-EM data collection was performed using a Titan Krios electron microscope (ThermoFisher Scientific) operated at 300 KV and equipped with Falcon 3EC direct electron detector. Micrographs were recorded in integrating mode as movie stacks with exposure time of 1.66 seconds and a total of 33 frames were stored per movie stack. The defocus range applied was -1.5 to -2.8  $\mu\text{m}$ . The calibrated magnification of the data acquisition was 100,719 $\times$ , which resulted in a pixel size of 1.39  $\text{\AA}$  per pixel and an electron dose of 40  $\text{e}^-/\text{\AA}^2$  was applied. The EPU software was used as a setup for the automatic data collection and a total of 15,162 movie stacks were collected at a rate of 100 images per hour.

### **Data processing and map calculation**

Motioncorr2 (Zheng et al., 2017) was used for performing drift collection and dose weighting on the movie stacks. Contrast transfer function (CTF) was first calculated using GCTF (Zhang, 2016) for aligned and non-dose weighted frames. CTF was then carefully inspected for drift and only images that extend beyond 5  $\text{\AA}$  were retained. A total of 2,613,025 particle-images were picked from the dose-weighted frames with BATCHBOXER implemented in EMAN (Ludtke et al., 1999) and using projections of an empty 70S ribosome filtered to 40  $\text{\AA}$  resolution as a reference. After 25 iterations of two dimensional (2D) image classification in RELION3 (Zivanov et al., 2018) on binned images (5.56  $\text{\AA}$  per pixel), a total of 2,477,544 particle-images were selected and further refined following the 3D refine approach in RELION3 and using a bacterial ribosome a reference filtered to 60  $\text{\AA}$  resolution. Images were then subjected to 3D focused classification without alignments by applying a circular mask onto the ribosome tunnel exit site. The 3D classification yielded two classes, one of which contained a density of the SecA protein bound to the ribosome (329,048 particle-images). The remaining classes were either of an empty ribosome or contained weak density on the exit tunnel and thus were discarded. To further improve the density of SecA, a second round of 3D classification was performed using 2-

fold binned images (2.78 Å per pixel), and the 3D class displaying high resolution features of SecA was retained (140,665 particle-images). A final round of 3D classification was performed by adjusting the tau value (T=10) in RELION and this yielded a 3D class where secondary structural elements of SecA can be resolved. The selected particles (37,334 particles) were refined at a full pixel size without binning (1.39 Å per pixel; box size is 320 x 320 pixels) resulting an overall average reconstruction of 3.1 Å resolution (Map1) using the 3D refinement approach in cryoSPARC (Punjani et al., 2017). Although the contact points between the ribosome and SecA were resolved to side-chain resolution (< 3.5 Å), the outer shell of SecA was around 8 Å resolution. To improve the local resolution of SecA, a focused refinement approach was performed by masking out the ribosome of the refined map, then re-centering the picked particle images around SecA density. Local angular searches were then applied in addition to small angular increments (1.8 degrees), which yielded a final reconstruction of the SecA protein at a local resolution of 5.7 Å (Map2). Local resolution and gold standard FSC plots using FSC=0.143 criterion were calculated as implemented in RELION3. Final maps were sharpened either in RELION3 or with the auto-sharpen option implemented in PHENIX (Afonine et al., 2018a).

### **Model building**

For the model building of the RNC•SecA complex, coordinates of the 50S (PDB ID:5GAG) and of SecA (PDB ID: 2FSF) and the coordinates of SecA PPXD (PDB ID: 2VDA) were docked as rigid body elements into the cryo-EM map using USCF CHIMERA (Pettersen et al., 2004). Manual adjustments of the protein  $\alpha$ -helices of SecA were done using COOT (Emsley and Cowtan, 2004) first into Map2 as the secondary structural elements of SecA were resolved. The contact sites between the ribosome and SecA in Map1 were resolved to 3.1- 3.5 Å and allowed us to build these regions *de novo*, in particular, the amphipathic helix at the N-terminal end of SecA and the nascent chain within the ribosome tunnel. H59 of the ribosomal RNA and uL23 were manually adjusted to better fit the EM density. The density of the transmembrane domain (TMD) of RodZ was resolved and assigned based on the side chain density of this region, which allowed to establish the registry and directionality of the TMD. The rest of the nascent chain region in the ribosome tunnel was traced as poly-alanine chains. The resulting model was then refined into the corresponding EM densities and subjected to five cycles of real space

refinements using phenix.real\_space\_refine (Afonine et al., 2018b), during which protein secondary structure, Ramachandran and side chain rotamer restraints, RNA base pair restraints were applied. The fit of the EM map was validated using the real space correlation coefficients ( $CC_{\text{mask}}$ ) between the model and the versus map Fourier Shell Correlation (FSC) at FSC=0.5 as a cut-off criterion and resulted in similar resolution as the half-set map FSC using FSC=0.143 criterion. Numbering of the TMD residues of RodZ corresponds to the numbering of residues in the full protein sequence from *E. coli*. In the deposited model, residues numbering of the nascent chain construct starts from the first methionine as residue number 1. Images were prepared in either Chimera, ChimeraX, or PyMOL.

### **Western blot**

Rabbit anti-uL23 antibody was customized from GenScript using CGKVKRHGQRIGRRS as the epitope. Anti-T7 antibody was purchased from Abcam. Anti-strep, anti-HA and anti-SUMO antibody were purchased from GenScript. Primary antibodies were incubated with IRDye<sup>®</sup> 800CW secondary antibodies (LI-COR) for detection. Protein band intensity was quantified by the Odyssey<sup>®</sup> CLx imaging system.

### **Nanodisc reconstitution**

Reconstitution was carried out as described previously (Koch et al., 2016) with slight modifications. A lipid mixture containing DOPC:DOPG:DOPE at a molar ratio of 4:3:3 (Avanti) was dried under nitrogen gas and then in a vacuum desiccator overnight, and resuspended in lipid buffer (20 mM HEPES pH 7.5, 112 mM KCl, 0.4 mM TCEP, 46 mM Na-cholate) at a total lipid concentration of 22 mM. Reconstitution reactions were performed in nanodisc buffer (25 mM HEPES pH 7.5, 140 mM KCl, 0.5 mM TCEP, 46 mM Na-cholate) and contained 13.4  $\mu\text{M}$  SecYEG, 133.8  $\mu\text{M}$  ApoE422K and 12.2 mM lipid mixture for SecYEG nanodisc; and 133.8  $\mu\text{M}$  ApoE422K and 12.2 mM lipid mixture for empty nanodisc. The mixtures were incubated at 4 °C for 1 hr on a rotary shaker. Bio-bead SM-2 resin (Bio-rad) was washed by methanol, ddH<sub>2</sub>O and bead buffer (25 mM HEPES pH 7.5, 140 mM KCl, 0.5 mM TCEP). Remove extra washed bio-bead so that the remaining bead volume is equal to the volume of the reconstitution reaction mixture. The reconstitution reaction mixture was added to the remaining bead and incubated at 4 °C for overnight on rotary shaker. The mixture was filtered to remove bio-bead and pelleted at

77,000 rpm for 25 min in TLA 120.2 (Beckman Coulter) to remove aggregates. The supernatant was concentrated in Amicon, 30K MWCO centrifugal filter unit (MilliporeSigma). The concentration of nanodisc was calculated as follows: the concentration of ApoE422K in nanodisc was quantified by SDS-PAGE using known concentrations of purified ApoE422K as standards. As each large nanodisc (~40 nm) contains ~8 copies of ApoE422K on average (Blanchette et al., 2008; Koch et al., 2016), the concentration of ApoE422K was divided by 8 to obtain the concentration of nanodisc. The concentration of SecYEG in nanodisc was determined by SDS-PAGE using known concentrations of purified SecYEG as standards.

### **Negative stain electron microscopy**

10 nM nanodiscs were applied onto a glow discharged ultrathin C film on holey carbon support film, 400 mesh, Cu grids (Ted Pella, Inc.). Samples were stained with 3% uranyl acetate. Data were collected using a FEI Tecnai T12 transmission electron microscope at 120 keV on a Gatan Ultrascan 2k x 2k CCD detector. Images were acquired using a 1 s exposure time at a nominal magnification of 42,000x at 2-3 $\mu$ m defocus, resulting in 2.5 Å per pixel.

### **Fluorescence labeling of SecA**

The single cysteine mutant of SecA (C98S/S12C) and its derivatives were reduced with 2 mM DTT at 4 °C for 30 min followed by dialysis in Labeling buffer (20 mM KHEPES, pH 7.0, 300 mM KCl, 10% glycerol, 2 mM TCEP) to remove DTT. 40  $\mu$ M SecA (C98S/S12C) was mixed with a 20-fold excess of BODIPY-FL maleimide on a rotary shaker at 4 °C for 4 hr. After quenching with 10 mM DTT, free dye was removed by chromatography on Sephadex G-25 column (Sigma-Aldrich) in SRP buffer (50 mM KHEPES, pH 7.5, 150 mM KOAc, 10 mM Mg(OAc)<sub>2</sub>, 2 mM DTT, 10% glycerol). Labeling efficiencies was ~90 %, determined using the adsorption coefficient of  $\epsilon = 73,000 \text{ M}^{-1}\text{cm}^{-1}$  for BODIPY-FL maleimide in aqueous buffer (Wang et al., 2017). The cysteines in the zinc-finger domain of SecA are coordinated by Zn<sup>2+</sup> and were not labeled (data not shown).

### **Fluorescence measurements**

Equilibrium titrations were performed using a Fluorolog-3-22 spectrofluorometer (Jobin Yvon) at room temperature in Assay buffer (50 mM KHEPES, pH 7.5, 150 mM KOAc, 10 mM

Mg(OAc)<sub>2</sub>, 2 mM DTT, 0.1 mg/ml BSA). Unless otherwise specified, experiments used an excitation wavelength of 360 nm and an emission wavelength of 455 nm. The FRET efficiency was calculated based on equation 1:

$$\text{FRET} = 1 - \frac{F_e}{F_0} \quad (1)$$

in which  $F_0$  is the fluorescence intensity at 455 nm for Cm-labeled RNC alone.  $F_e$  is the fluorescence intensity at 455 nm when the Cm-labeled RNC is incubated with saturating amount of BDP-labeled SecA.

Equilibrium titrations used 10 nM Cm-labeled RNC, and indicated concentrations of SecA or SecYEG nanodisc as the titrant. The data were fit to equation 2:

$$\Delta F_{\text{norm}} = 1 \times \frac{[\text{RNC}] + [\text{titrant}] + K_d - \sqrt{([\text{RNC}] + [\text{titrant}] + K_d)^2 - 4 \times [\text{RNC}][\text{titrant}]}}{2 \times [\text{RNC}]} \quad (2)$$

in which “Normalized fluorescence change” ( $\Delta F_{\text{norm}}$ ) was calculated by dividing the observed fluorescence change at each titrant concentration over the fluorescence change at saturating titrant concentration, so that all titration curves start at 0 and plateau at 1, and the curvature of the titration curves directly reflect the  $K_d$  value.

Dissociation rate constants of SecA from RNC were measured using a Kintek stopped flow apparatus at room temperature as described previously (Rome et al., 2014). 10 nM Cm-labeled RNC and 30 nM BODIPY-FL-labeled SecA were preincubated in Assay buffer, followed by addition of unlabeled SecA at indicated concentrations as the chase to initiate dissociation of the preformed complex. The time course of observed fluorescence ( $F$ ) was fit to a double exponential function (equation 3):

$$F = F_e + \Delta F_a \times e^{-k_a t} + \Delta F_b \times e^{-k_b t} \quad (3)$$

in which  $F_e$  is the fluorescence when the reaction reaches equilibrium,  $\Delta F_a$  and  $k_a$  are the magnitude and rate constant of the fast phase, and  $\Delta F_b$  and  $k_b$  are the magnitude and rate constant of the slow phase. The magnitude and rate constants of the slow phase are consistent with fluorescence bleaching of the Cm dye determined in parallel measurements. Hence, the first phase was assigned to SecA dissociation from RNC, and  $k_a$  represents the dissociation rate

constant ( $k_1$ ). Normalized fluorescence was calculated by dividing the observed fluorescence change at each time point over the fluorescence change when the reaction is complete, so that all the traces start at 0 and plateau at 1.

Measurements of RNC<sub>RodZ</sub> transfer from SecA to SecYEG-Nd or empty nanodisc were performed using a Kintek stopped flow apparatus at room temperature in Assay buffer supplemented with 0.5 mM AMP-PNP. 10 nM Cm-labeled RNC was preincubated with 30 nM unlabeled SecA followed by addition of SecYEG nanodisc or empty nanodisc at indicated concentrations. The time course of observed fluorescence ( $F$ ) was fit to Eq. 3, in which  $k_a$  represents the apparent rate constant of RNC transfer. Normalized fluorescence was calculated by subtracting the observed fluorescence at each time point by  $F_e$ , and is then divided by  $\Delta F_a$  so that all the traces start at 1 and plateau at 0.

### Kinetic simulations

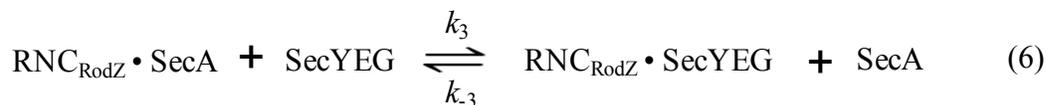
Simulations in Figures 2.6A-B, Figures S2.7C-F were performed using the Berkeley Madonna software.

For the passive model in Figure 2.6A and Figures S2.7C-D, the following reactions were modeled:



For the active model in Figure 2.6B and Figures S2.7E-F, the passive pathway (equations 4 and 5) was included in the simulation of the active model for completeness. The following reactions were modeled:





The initial concentrations of all species were set based on experiment conditions as described under “Fluorescence measurements”:

$$[\text{RNC}_{\text{RodZ}} \cdot \text{SecA}]_0 = 10 \text{ nM}$$

$$[\text{SecA}]_0 = 20 \text{ nM}$$

$$[\text{RNC}_{\text{RodZ}}]_0 = 0 \text{ nM}$$

$$[\text{RNC}_{\text{RodZ}} \cdot \text{SecYEG}]_0 = 0 \text{ nM}$$

$$[\text{SecYEG}]_0 : \text{varied concentrations as indicated}$$

The SecA dissociation rate constant,  $k_1$ , was experimentally determined for  $\text{RNC}_{\text{RodZ91}}$  ( $0.00309 \text{ s}^{-1}$ , Figure 2.5C). To measure the SecA association rate constant ( $k_{-1}$ ), we monitored the association of  $\text{RNC}_{\text{RodZ}}$  with varying concentrations of SecA using the FRET assay (Figure S2.7A). The observed rate constant of SecA binding to  $\text{RNC}_{\text{RodZ}}$  ( $k_{\text{obsd}}$ ) was plotted as a function of SecA concentration and fit to equation 7:

$$k_{\text{obsd}} = k_{\text{on}}[\text{titrant}] + k_{\text{off}} \quad (7)$$

in which SecA is the titrant, and  $k_{\text{on}}$  is the SecA association rate constant ( $k_{-1}$ ) and was determined to be  $1.48 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ . The dissociation constant ( $K_{\text{d1}}$ ) of the  $\text{RNC}_{\text{RodZ91}} \cdot \text{SecA}$  complex was calculated to be 2.2 nM ( $K_{\text{d1}} = k_1/k_{-1}$ ).

To obtain the dissociation constant ( $K_{\text{d2}}$ ) of the  $\text{RNC}_{\text{RodZ91}} \cdot \text{SecYEG}$  formed in the transfer reaction, we titrated SecYEG-Nd during the transfer. We preformed a complex of 10 nM  $\text{RNC}_{\text{RodZ}}^{\text{Cm}}$  with 30 nM BDP-labeled SecA, and added increasing amounts of SecYEG-Nd; the increase in Cm fluorescence due to the loss of FRET was used to monitor the transfer reaction (Figure S2.7B). The data were fit to equation 8:

$$F = F_0 + (F_e - F_0) \times \frac{[\text{SecYEG}]}{[\text{SecYEG}] + K_{1/2}} \quad (8)$$

in which  $F$  is the observed Cm fluorescence,  $F_0$  and  $F_e$  are the Cm fluorescence at the beginning and end of the titration, respectively, and  $K_{1/2}$  is the concentration of SecYEG-Nd required for 50% complete transfer.  $K_{1/2}$  was determined to be 45 nM for RNC<sub>RodZ91</sub>. At this concentration, we have:

$$[\text{RNC} \cdot \text{SecA}] = [\text{RNC} \cdot \text{SecYEG}] \quad (9)$$

$$[\text{RNC} \cdot \text{SecYEG}] + [\text{SecYEG}] = 45.4 \text{ nM} \quad (10)$$

$$[\text{RNC} \cdot \text{SecA}] + [\text{SecA}] = 30 \text{ nM} \quad (11)$$

$$[\text{RNC} \cdot \text{SecA}] + [\text{RNC} \cdot \text{SecYEG}] + [\text{RNC}] = 10 \text{ nM} \quad (12)$$

$$K_{d1} = \frac{[\text{RNC}] \times [\text{SecA}]}{[\text{RNC} \cdot \text{SecA}]} = 2.2 \text{ nM} \quad (13)$$

$$K_{d2} = \frac{[\text{RNC}] \times [\text{SecYEG}]}{[\text{RNC} \cdot \text{SecYEG}]} \quad (14)$$

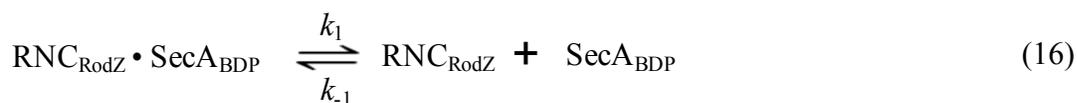
$$K_{\text{trans}} = \frac{[\text{RNC} \cdot \text{SecYEG}] \times [\text{SecA}]}{[\text{RNC} \cdot \text{SecA}] \times [\text{SecYEG}]} = \frac{K_{d1}}{K_{d2}} = \frac{k_3}{k_{-3}} \quad (15)$$

Solving equations 9-15 gave  $K_{d2} = 3.4 \text{ nM}$  and  $K_{\text{trans}} = 0.62$ . The association rate constant of RNC<sub>RodZ</sub> binding to SecYEG ( $k_2$ ) was assumed to be  $1 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ , which is typical for bimolecular association. This results in a dissociation rate constant ( $k_{-2} = K_{d2} \times k_2$ ) for the RNC<sub>RodZ</sub>•SecYEG complex of  $0.00338 \text{ s}^{-1}$ , which is consistent with the previous observation that the half-life of the RNC•SecYEG complex is  $\sim 250 \text{ s}$  (Wu et al., 2012). Varying the values of  $k_2$  and  $k_{-2}$  while maintaining the value of  $K_{d2}$  did not affect the outcome of the simulation (Figure 2.7C-F).

To measure the rate constant  $k_3$ , the observed transfer rate of RNC<sub>RodZ91</sub> in Figure 2.6E (green) was fit to equation 7, where  $k_3 = k_{\text{on}}$  and was determined to be  $1.4 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$ .  $k_{-3}$  was calculated to be  $2.3 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$  based on equation 15.

Because the formation of  $\text{RNC}_{\text{RodZ}} \cdot \text{SecYEG}$  complex causes fluorescence quenching of  $\text{RNC}_{\text{RodZ}}$ , the simulated fluorescence ( $F_{\text{sim}}$ ) starts at 1 at time = 0s, and is proportional to the sum of the fraction of  $\text{RNC}_{\text{RodZ}}$  and  $\text{RNC}_{\text{RodZ}} \cdot \text{SecA}$  complex. Normalized fluorescence was simulated as  $(F_{\text{sim}} - F_{\text{sim,e}})/(1 - F_{\text{sim,e}})$  in which  $F_{\text{sim,e}}$  is the fraction of  $F_{\text{sim}}$  when the reaction is complete, so that the traces start at 1 and plateau at 0.

To simulate the chase experiments to measure SecA dissociation from  $\text{RNC}_{\text{RodZ}}$  in Figure S2.7G, the following reactions were used:



In which  $\text{SecA}_{\text{BDP}}$  and  $\text{SecA}$  denote BODIPY-FL labeled and unlabeled SecA, respectively. The initial concentrations of all species were set based on experiment conditions as described above under ‘‘Fluorescence measurements’’:

$$[\text{RNC}_{\text{RodZ}} \cdot \text{SecA}_{\text{BDP}}]_0 = 10 \text{ nM}$$

$$[\text{SecA}_{\text{BDP}}]_0 = 20 \text{ nM}$$

$$[\text{RNC}_{\text{RodZ}}]_0 = 0 \text{ nM}$$

$[\text{SecA}]_0$ : varied concentrations as indicated

As described above,  $k_1$  and  $k_{-1}$  were set to  $0.00309 \text{ s}^{-1}$  and  $1.48 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ , respectively.

Normalized fluorescence change during the chase is proportional to the fraction of the  $\text{RNC}_{\text{RodZ}} \cdot \text{SecA}$  complex and was simulated as  $[\text{RNC}_{\text{RodZ}} \cdot \text{SecA}]/[\text{RNC}_{\text{RodZ}} \cdot \text{SecA}]_e$ , in which  $[\text{RNC}_{\text{RodZ}} \cdot \text{SecA}]_e$  is the  $\text{RNC}_{\text{RodZ}} \cdot \text{SecA}$  concentration when the reaction is complete.

### **In vitro translation-translocation in PURE system**

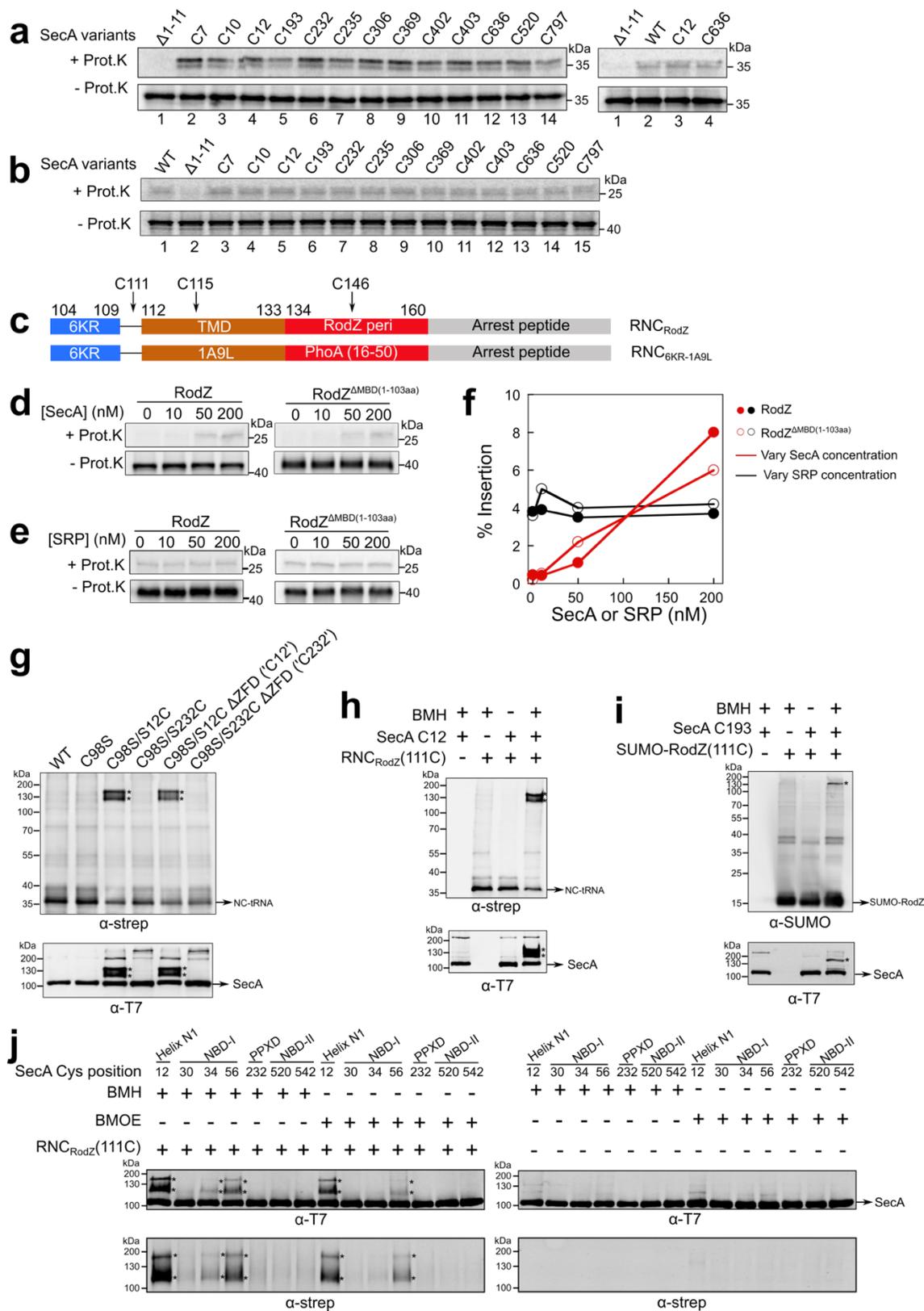
Translation was performed at 30 °C using PURExpress<sup>®</sup> *in vitro* protein synthesis kit (NEB) supplemented with <sup>35</sup>S-Methionine (1.5 mCi/ml, PerkinElmer) and the indicated concentrations of cytosolic factors (SecA, SecB, Ffh, FtsY or TF). Unless otherwise indicated, 0.5 mg/ml Urea-washed IMV (Wang et al., 2017) was added 5 min after initiation of translation. The reaction

was continued for 85 min at 30 °C, after which it was split equally into two samples, one of which was digested with 0.5 mg/ml proteinase K for 30 min at 25 °C. Digestion was stopped by addition of 5 mM PMSF, after which the sample was incubated on ice for 10 min. Samples with and without proteinase K treatment were analyzed by SDS-PAGE and autoradiography.

### **Additional information**

Cryo-EM maps and model coordinates are deposited in the EMDB as EMD-10073 and EMD-10074, and in the PDB as PDB ID 6S0K. Other data are available from corresponding authors upon reasonable request.

## 2.6 Supplementary Figures and Legends



**Figure S2.1. Controls and additional data to map the interaction of SecA with hydrophobic sequences on nascent protein in the co- versus post-translational mode.**

**a-b,** The activity of SecA variants was tested by assaying the translocation of proOmpA, a model post-translational SecA substrate (a), and RodZ, a model co-translational SecA substrate (b).

ProOmpA and RodZ was *in vitro* translated using the PURE system supplemented with 1  $\mu$ M SecA variants and urea-washed inverted membrane vesicles as described previously. Successful insertion into the membrane was detected by protection against proteinase K (Prot.K) digestion. In the left panel of part (a), lane 1 shows the reaction of mutant SecA with deletion on residues 1-11, and lanes 2-14 show the data with single cysteine variants of SecA. The right panel of part (a) shows the comparison of representative single cysteine mutants of SecA relative to wild type (WT) SecA. The data in each gel are from side-by-side experiments.

**c,** Scheme of the composition of the nascent chains on stalled RNCs used in this study. MreB-binding domain (MBD, residues 1-103)<sup>18</sup> was removed from the RodZ nascent chain. Arrest peptide is from SecM residues 133-170. The positions of engineered cysteines at residues 111, 115 and 146 are indicated. 1A9L nascent chain was constructed by replacing the TMD of RodZ with 1A9L<sup>10</sup> followed by the mature region (residues 16-50) of PhoA. RNC, ribosome-nascent chain complex.

**d-f,** The MreB-binding domain (MBD; residues 1-103) of RodZ is not essential for SecA-dependent co-translational translocation in a coupled *in vitro* translation-translocation assay. As previously described<sup>17,47,48</sup>, RodZ or RodZ $\Delta$ MBD was translated using the PURE system supplemented with the indicated concentrations of SecA (d) or SRP (e) and urea-washed, inverted membrane vesicles. Successful insertion into the membrane was detected by protection against proteinase K (Prot.K) digestion. The reactions in (d) also contained 3.8  $\mu$ M trigger factor, 400 nM SRP, 1  $\mu$ M FtsY. The reactions in (e) also contained 3.8  $\mu$ M trigger factor, 50 nM SecA, and a fivefold excess of FtsY over SRP. (f) Summary of the insertion efficiency of RodZ or RodZ $\Delta$ MBD from the data in (d) and (e). Insertion efficiency was calculated by dividing the amount of proteinase K-resistant protein by the total amount of protein, normalized by the number of methionines before and after Prot.K digestion.

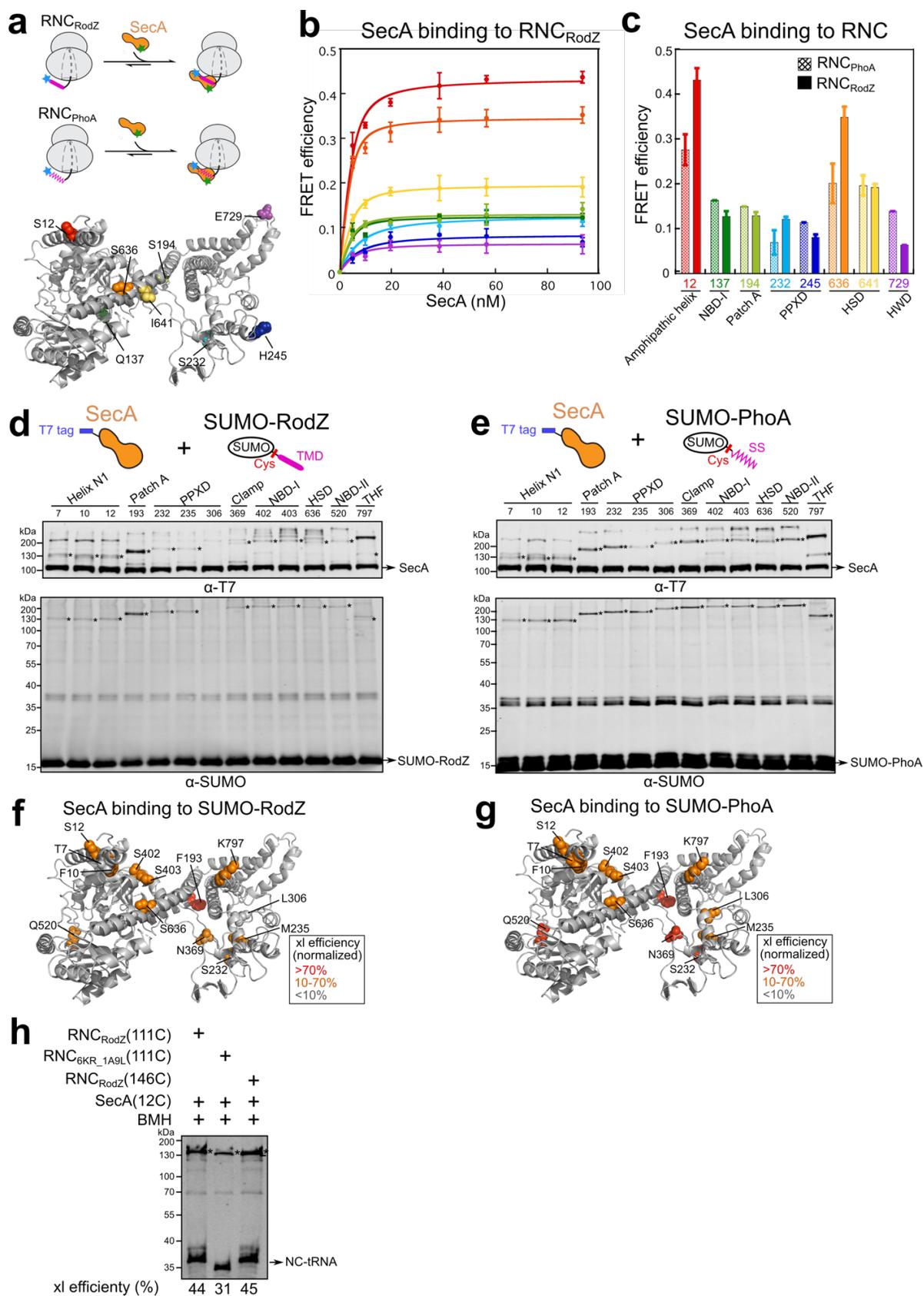
**g,** Crosslinking of SecA to C111 in the RodZ nascent chain depends on engineered cysteine on SecA. All lanes contain the BMH crosslinker. Wild type (WT) SecA contains four cysteines (residue 98, 885, 887, 896), none of which crosslinked to RodZ nascent chain. In C98S, the cysteine at residue 98 was mutated to serine. In  $\Delta$ ZFD, the non-essential C-terminus of SecA

containing three cysteines was removed. All the other crosslinking experiments in this work contained the C98S mutation and  $\Delta$ ZFD deletion for clean interpretation of results. Asterisks indicate crosslinked products detected by the anti-strep and anti-T7 antibodies.

**h**, Crosslinking of SecA(C12) to C111 in the RodZ nascent chain depends on the crosslinker, SecA and RNC<sub>RodZ</sub>. Asterisks indicate crosslinked products detected by the anti-strep and anti-T7 antibodies.

**i**, Crosslinking of SecA (C193) to C111 in SUMO-RodZ depends on the crosslinker, SecA and SUMO-RodZ. Asterisks indicate crosslinked products detected by the anti-SUMO and anti-T7 antibodies.

**j**, Engineered single cysteines at the indicated positions of SecA were tested for crosslinking to RodZ (C111) on RNC by BMH or BMOE. Asterisks indicate crosslinked products that are detected by the anti-strep and anti-T7 antibodies.



**Figure S2.2. Additional data to map the interaction surface of SecA with nascent chains on RNC and with post-translational substrates.**

**a**, FRET experiments to monitor the proximity between Cm (blue star)-labeled RNC<sub>RodZ</sub> or RNC<sub>phoA</sub> and BDP (green star) labeled at indicated positions on SecA. Cm was incorporated at residue 111 immediately upstream of the RodZ TMD (magenta) or residue 4 upstream of the phoA signal sequence (magenta). Top left panel, scheme of the FRET-based binding assay. Lower left panel, SecA residues for acceptor labeling are mapped onto the structure of SecA from this work.

**b**, Representative equilibrium titrations showing the binding of SecA<sup>BDP</sup> to Cm-labeled RNC<sub>RodZ</sub>. Reactions used 20 nM RNC<sup>Cm</sup> (donor) and indicated concentrations of SecA<sup>BDP</sup> (acceptor). All titrations saturated above 20 nM SecA, indicating tight binding of all the fluorescently labeled SecA variants. The data for individual SecA variants are colored as in the lower panel of (a).

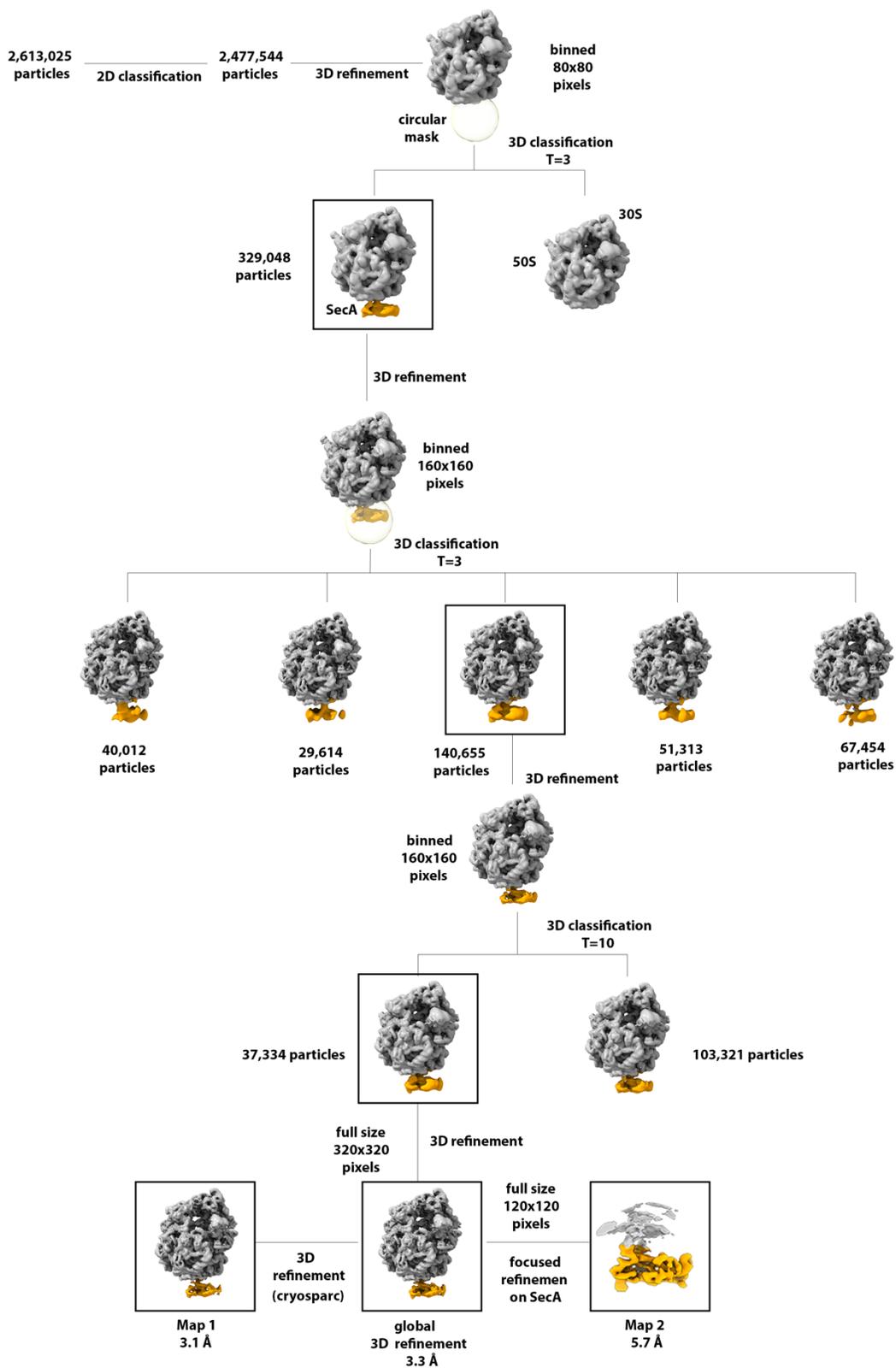
**c**, Summary of FRET efficiency in the complexes formed between the indicated SecA variants and RNC<sub>RodZ</sub> or RNC<sub>phoA</sub>. FRET efficiency was calculated at 500 nM SecA<sup>BDP</sup> according to Eq. 1. The data for individual SecA positions are colored as in the lower panel of (a). All values represent mean  $\pm$  SD, with  $n = 2-3$ .

**d-e**, Engineered single cysteines at indicated positions on SecA were tested for crosslinking by BMH to the RodZ TMD (residues 104-133) or the phoA signal sequence (residues 1-21) fused to the C-terminus of SUMO (SMT3 residues 1-101). The cysteines on RodZ and PhoA are at the same locations as in Figure 2.1a and b, respectively. Crosslinking reactions used 8.3  $\mu$ M SUMO fusion proteins and 1  $\mu$ M SecA. Asterisks denote crosslinked products detected by both the anti-SUMO and anti-T7 antibodies.

**f-g**, Crosslinking efficiency from the data in parts (d) and (e), respectively, are summarized in the structural model of SecA from this work. Crosslinking efficiencies (normalized) were relative to the crosslinked product formed by SecA(C193), based on western-blot against SUMO and strep-tag. Residues are colored based on crosslinking efficiency as indicated.

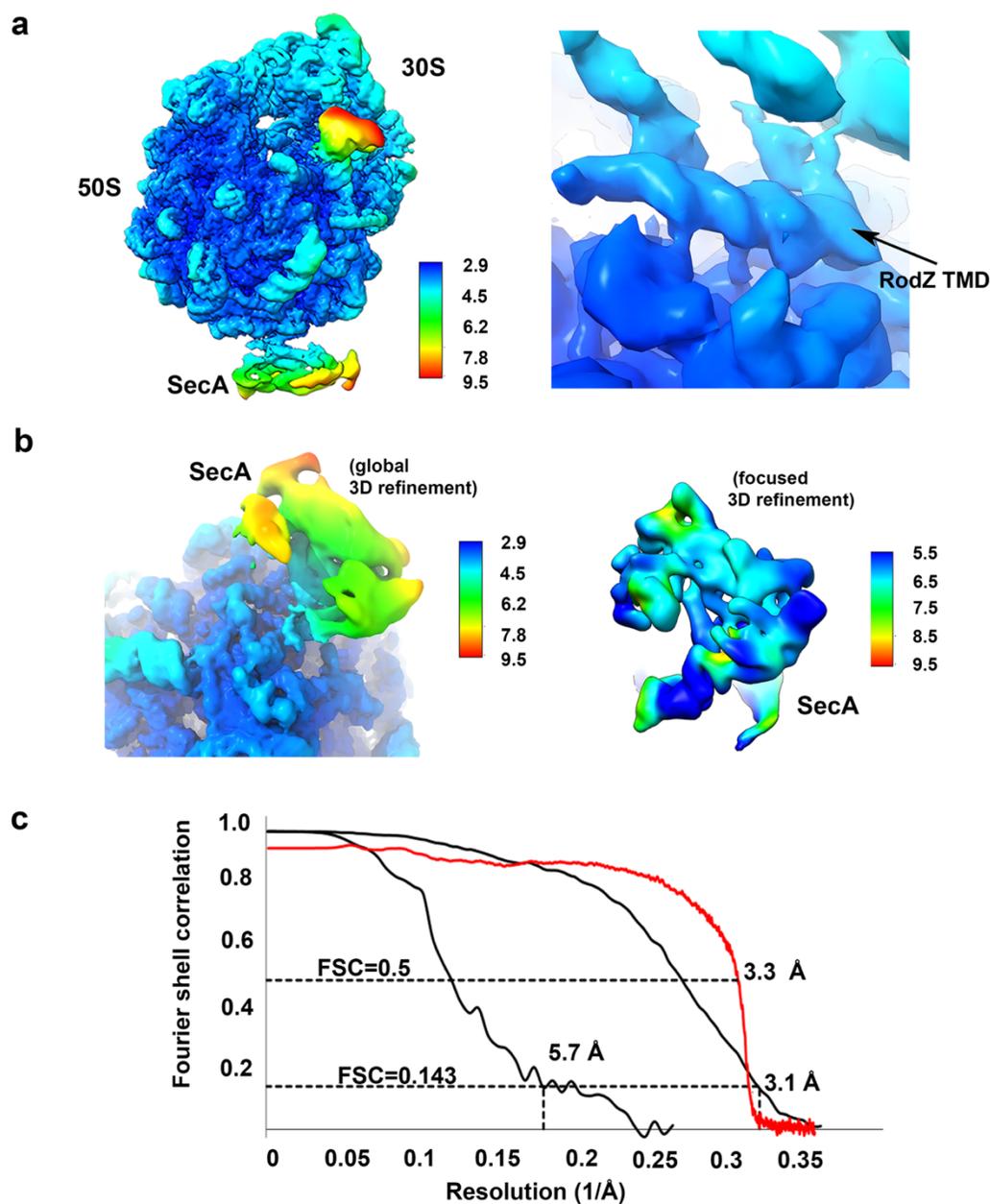
**h**, Characterization of samples for cryoEM. RNCs were tested for crosslinking between the indicated cysteines on the nascent chain and SecA (C12). RNC<sub>6KR\_1A9L</sub> contained a model signal sequence 1A9L in place of the RodZ-TMD preceded by six consecutive basic residues derived from residues 104-109 of RodZ. Single bands were observed for both the tRNA-linked nascent chain and crosslinked products with SecA, probably due to the removal of polysomes during preparation of the samples for cryoEM. Asterisks denote major crosslinked products detected by

anti-strep antibody. Crosslinking efficiency was quantified from the ratio of the intensity of crosslinked nascent chain relative to the total intensity of bands containing the nascent chain.



**Figure S2.3. Image classification and refinement of the structure of the RNC<sub>RodZ</sub>•SecA complex.**

An initial 2D classification was performed on 4-fold binned particles with pixel size of 5.56 Å per pixel (box size 80 x 80 pixels) to remove bad particles. The selected particle images were then subjected to 3D refinement in RELION3 to obtain an initial map of the ribosome. Using a circular mask applied at the ribosome tunnel region, a 3D focused classification without alignment was performed. This approach yielded a 3D class with a density corresponding to SecA. The remaining classes contained either no or weak density at the exit tunnel region and were discarded. A second round of focused 3D classification on two-fold binned images (160 x 160 pixels) yielded a class with an improved EM density of SecA. A final round of focused 3D classification by adjusting tau values in RELION3 (T=10), which yielded a 3D class with resolved secondary structure elements in the EM density of SecA. The selected particle images in this 3D class were subjected to a 3D refinement using full size images without binning (320 x 320 pixels) in RELION3, which yielded a map with an overall resolution of 3.3 Å, and was further improved to 3.1 Å when refined in cryoSPARC (Map1). To improve the local resolution of SecA, a focused 3D refinement scheme was used by first shifting the center of the box from the ribosome to SecA and re-extracting the new particle coordinates using the re-centering option in RELION3 (box size 120 x 120 pixels). Local searches along with a mask around the SecA density were then applied, which resulted in a map of SecA resolved to 5.7 Å resolution (Map2).



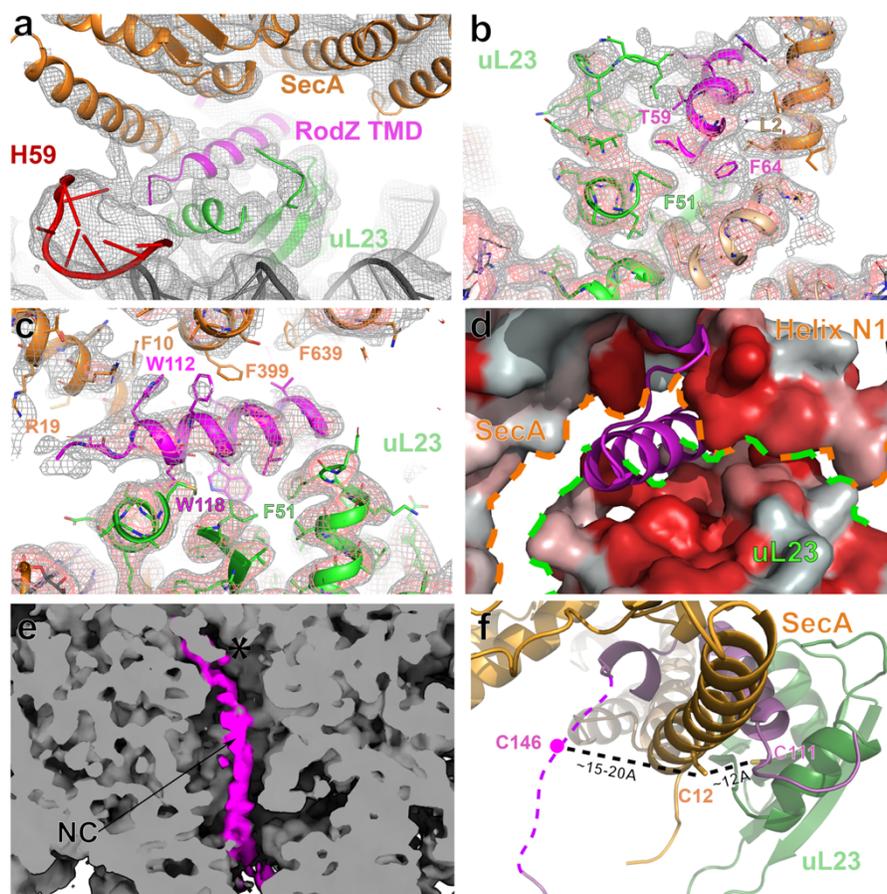
**Figure S2.4. Local resolution and validation of the cryoEM maps.**

**a**, Local resolution of Map1 obtained from the 3D refinement (left panel), and close-up of the contact points and the resolved density for the RodZ TMD at a similar resolution as the overall resolution (right panel).

**b**, Comparison of the local resolution plot for the EM density corresponding to SecA using the global refinement approach (left panel) and the focused refinement approach (right panel).

Corresponding color keys are shown on the right side of each map.

c, Fourier Shell Correlation (FSC) plots for Map1 (3.1 Å) and Map2 (5.7 Å) using the gold standard FSC criteria cutoff (FSC=0.143) using independent two half maps as implemented in RELION3 and cryoSPARC. Map1 versus obtained model plot shown in red and depicts a similar resolution 3.3 Å using cutoff (FSC=0.5) as that of the cryoEM Map1.



**Figure S2.5. Overview of the RodZ TMD binding pocket and the ribosome tunnel region.**

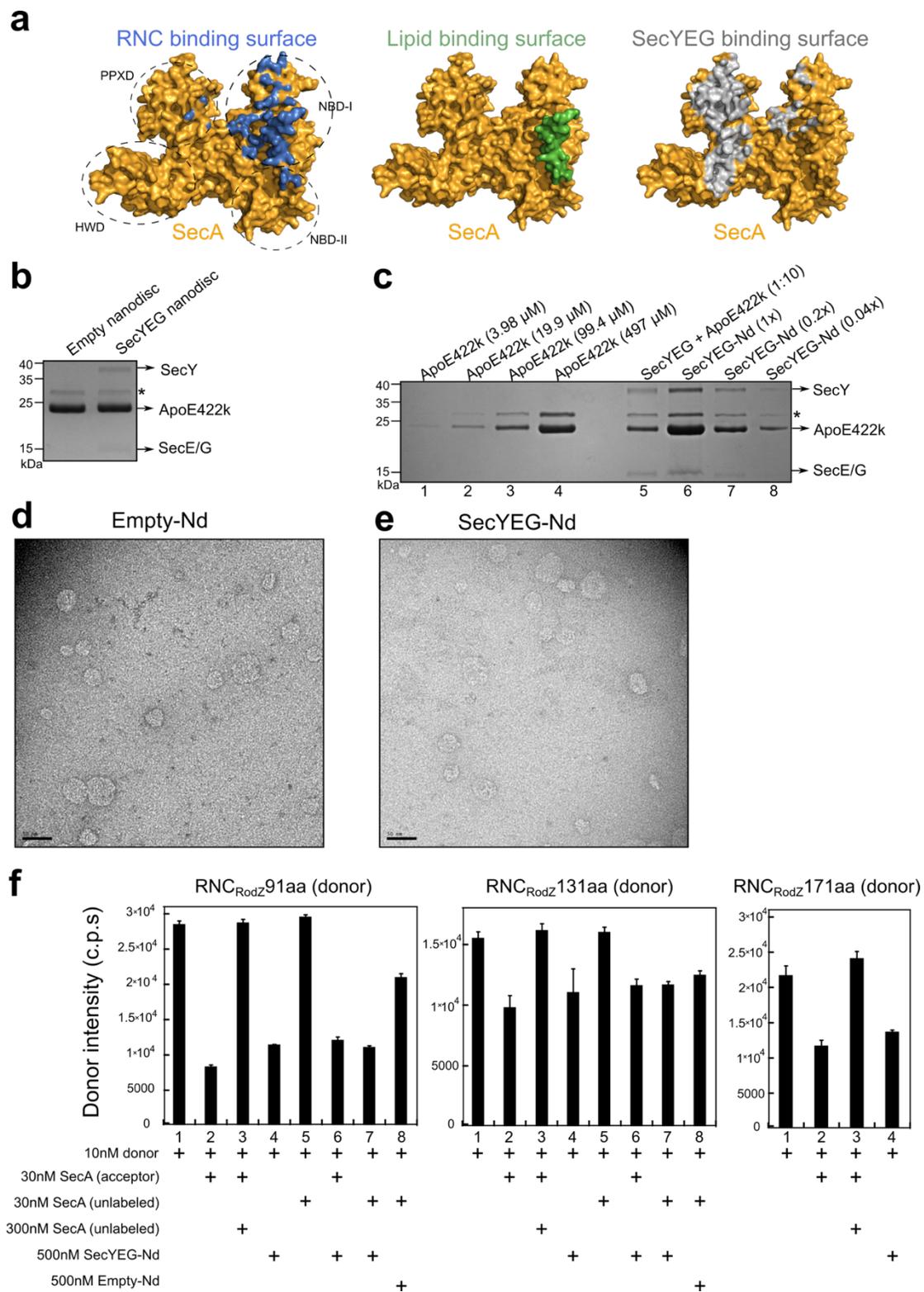
**a**, Close-up of the N-terminal amphipathic helix of SecA in the  $\text{RNC}_{\text{RodZ}} \cdot \text{SecA}$  complex with an overlay of the EM-density. Color scheme is the same as in Figure 2.3. The EM density was filtered based on the local resolution.

**b-c**, Local EM-densities outlining regions of RodZ TMD interactions with SecA, uL23, and uL29 are shown with fitted atomic models. EM-densities are low-pass filtered to 3.5 Å resolution for clarity.

**d**, Hydrophobicity gradient of the TMD binding pocket formed by SecA and uL23. Hydrophobic gradient shows hydrophobic amino acids in red and non-hydrophobic residues to white, applied from script “color\_h” in pymol. Orange and green dashed lines outline the surfaces from SecA and uL23, respectively.

**e**, A cross-section of the ribosome tunnel region with the EM-density of the RodZ nascent chain colored in magenta. The asterisk indicates the position of the CAA end of the P-site tRNA. EM-densities of the RodZ nascent chain were filtered to 4.5 Å resolution for clarity.

**f**, Distance between C12 of SecA and C111/C146 on the RodZ nascent chain. C12 and C111 are shown in sticks, and the hypothetical location of C146 is shown in sphere. Residues 134-160 of nascent chain are not resolved and are shown as a dashed line.



**Figure S2.6. Characterization of the samples for fluorescence measurements of RNC<sub>RodZ</sub> transfer from SecA to SecYEG.**

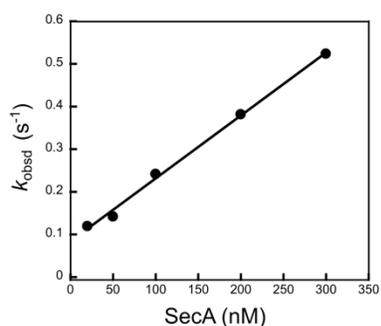
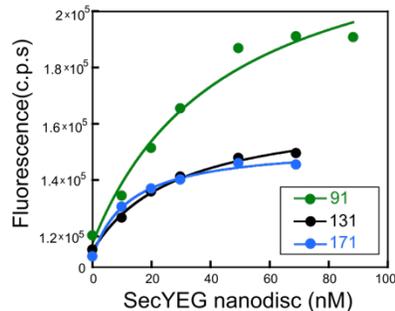
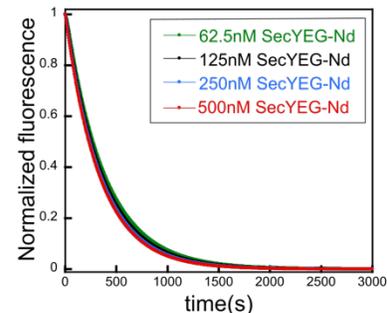
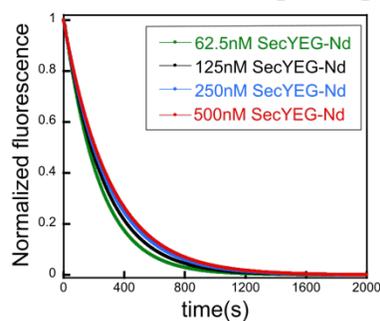
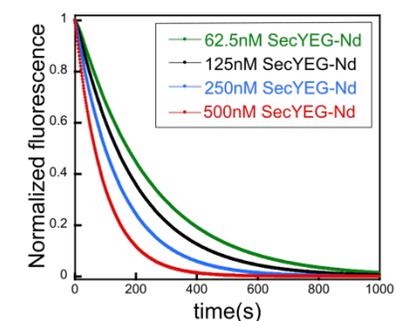
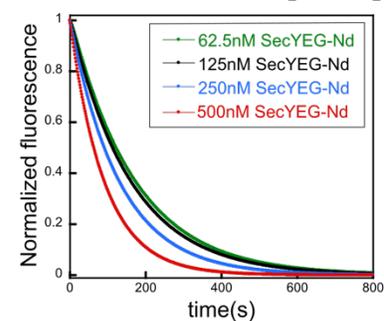
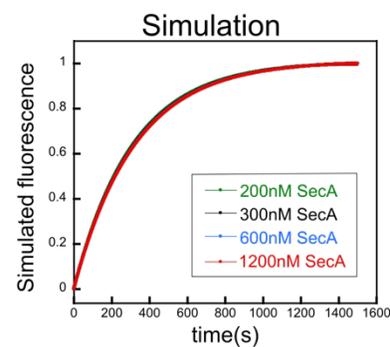
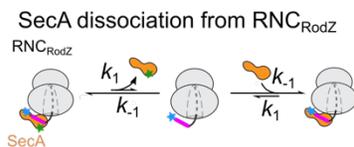
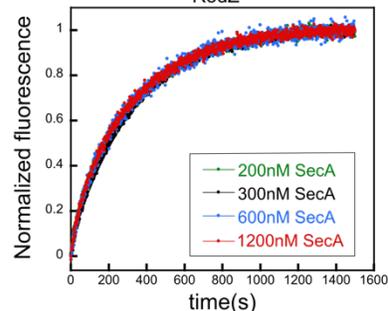
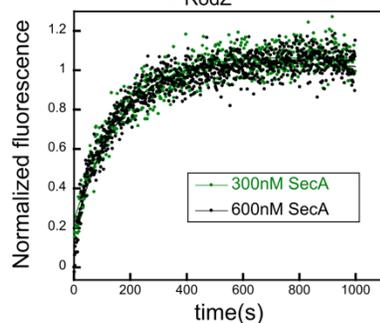
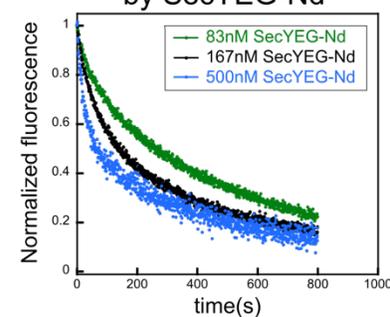
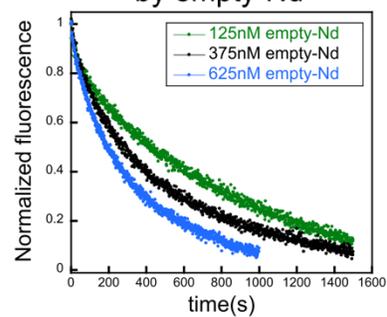
**a**, SecA surfaces contacting RNC (blue), anionic phospholipid (green), and SecYEG (grey). Blue and grey highlight SecA residues within 6 Å of RNC (this work) and SecYEG (PDB ID 5EUL), respectively. Green highlights residues 1-20 of SecA that mediate its lipid binding<sup>33</sup>.

**b**, Coomassie-blue stained gel showing reconstituted ApoE422k nanodiscs with and without SecYEG. ApoE422k contains two thrombin cleavage sites at the N-terminus, and the observed minor band (asterisk, ~5% of total) may represent incompletely cleaved ApoE422k.

**c**, Quantification of SecYEG-Nd by Coomassie-blue stained gel. Lanes 1-4 are purified ApoE422k at known concentrations. Lanes 6-8 are reconstituted SecYEG-Nd at different dilutions. Lane 5 shows the mixture of SecYEG:ApoE422k:lipid (molar ratio = 0.1:1:91) before removing detergent (see methods). The intensity of bands with purified ApoE422k was used to generate a standard curve from which we calculated the concentration of ApoE422k in the nanodisc. The concentration of nanodisc was determined to be 136.5 μM by dividing the concentration of ApoE422k in nanodisc by 8 (see method). The concentration of SecYEG in nanodisc was 75.4 μM, determined as with ApoE422k using purified SecYEG to construct a standard curve. These values indicate that there is 0.55 copy of SecYEG per copy of nanodisc on average. The observed minor band (asterisk, ~5% of total) may represent incompletely cleaved ApoE422k.

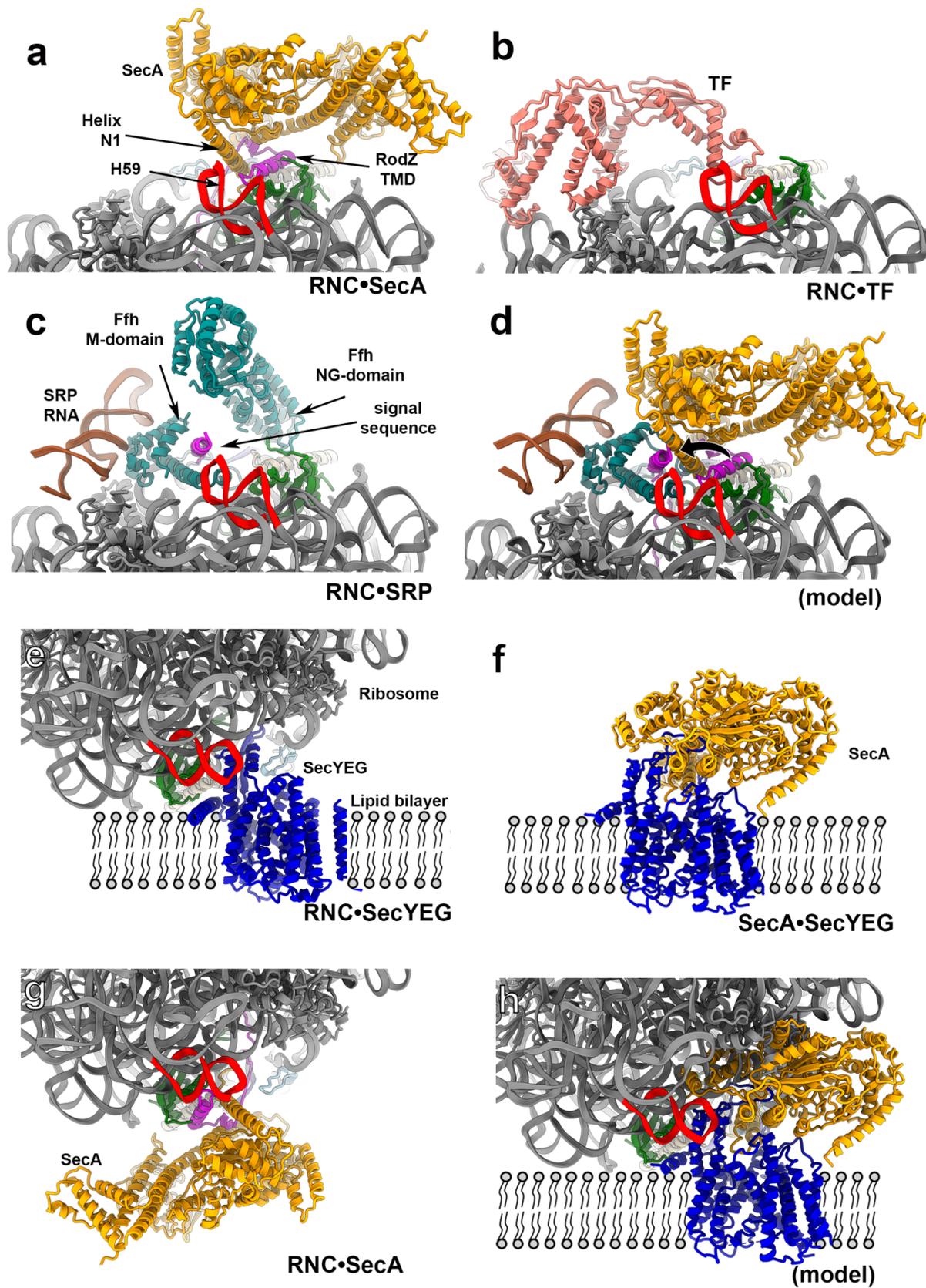
**d-e**, Negative stain electron microscopy images of the empty (d) and SecYEG (e) nanodiscs. Scale bar: 50 nm.

**f**, Summary of the steady-state fluorescence intensity of the Cm dye on RNC<sub>RodZ</sub> under the indicated reactions. RNC<sub>RodZ</sub> 91aa, 131aa, and 171aa contain RodZ residues 104-160, 104-200, and 104-240, respectively. Nd, nanodisc. c.p.s, counts per second.

**a** RNC<sub>RodZ</sub>•SecA association**b** RNC<sub>RodZ</sub> chase by SecYEG-Nd**c** Passive binding of SecYEG  
2-fold decrease in  $k_2$  and  $k_{-2}$ **d** Passive binding of SecYEG  
2-fold increase in  $k_2$  and  $k_{-2}$ **e** Active recruitment of SecYEG  
2-fold decrease in  $k_2$  and  $k_{-2}$ **f** Active recruitment of SecYEG  
2-fold increase in  $k_2$  and  $k_{-2}$ **g****h** SecA dissociation from  
RNC<sub>RodZ</sub> 91aa**i** SecA dissociation from  
RNC<sub>RodZ</sub> 131aa**j** RNC<sub>RodZ</sub> 131aa chase  
by SecYEG-Nd**k** RNC<sub>RodZ</sub> 131aa chase  
by empty-Nd

**Figure S2.7. Supporting information for the kinetic simulations, and additional data for chase experiments of the RNC<sub>RodZ</sub>•SecA complex.**

- a**, Measurement of the association rate constant of the SecA•RNC<sub>RodZ</sub> complex. 10 nM Cm-labeled RNC<sub>RodZ</sub> was mixed with indicated concentrations of BDP-labeled SecA, and the fluorescence change was fit to Eq. 3 to extract the observed association rate constant ( $k_{\text{obsd}}$ ). The value of  $k_{\text{obsd}}$  was plotted against SecA concentration and fit to Eq. 7 to determine  $k_{-1}$ .
- b**, Equilibrium measurement of the transfer reaction. 10 nM RNC<sub>RodZ</sub><sup>Cm</sup> at indicated chain lengths was pre-incubated with 30 nM BDP-labeled SecA. Increasing amounts of SecYEG-Nd were then added to the preformed complex, and the increase in Cm fluorescence due to the loss of FRET was monitored. The data were fit to Eq. 8 and gave a  $K_{1/2}$  value of  $45 \pm 18$  nM,  $29 \pm 5.4$  nM, and  $13 \pm 2.1$  nM for RNC<sub>RodZ</sub> at nascent chain lengths of 91, 131, and 171aa, respectively.
- c-f**, Changes in the rate constants of RNC<sub>RodZ</sub>•SecYEG association ( $k_2$ ) and dissociation ( $k_{-2}$ ) (the  $K_d$  value for RNC<sub>RodZ</sub>•SecYEG was held constant) do not affect the kinetics behavior for both the passive (c,d) and active (e,f) models.
- g**, Reaction scheme (left) and simulation (right) of the experiments to measure the dissociation rate constant ( $k_1$ ) of SecA from RNC<sub>RodZ</sub>. A preformed complex of Cm (blue star)-labeled RNC<sub>RodZ</sub> with BDP-labeled SecA was chased with excess unlabeled SecA to initiate complex dissociation, and the loss of FRET was monitored in real time.
- h-i**, Representative time courses for measurement of  $k_1$  at nascent chain lengths of 91aa (h) and 131aa (i). The data were fit to Eq. 3. All traces are the average of 6-8 measurements.
- j-k**, Representative fluorescence time traces for chase of SecA-bound RNC<sub>RodZ</sub> complex with SecYEG-Nd (j) or empty nanodisc (k) at a nascent chain length of 131aa. Reactions were carried out and analyzed as in Figure 2.6c,d, and the obtained rate constants are summarized in Figure 2.6f. Note that the time traces are biphasic, and control experiment indicated that the slow phase was due to dye bleaching (see Methods).



**Figure S2.8. Comparison of the structure from this work with previous structures.**

**a-c**, Comparisons of SecA (a; this work), TF (b; PDB ID: 1W26 & 1W28) and SRP (c; PDB ID: 5GAF) bound to RNC.

**d**, Overlay of SecA and SRP on the ribosome. The NG domain of the SRP protein Ffh was removed due to steric clash with SecA. The arrow indicates the difference in the position of the signal sequence versus TMD on the ribosome in the presence of SRP versus SecA. The following coloring scheme is used. TF, salmon; Ffh, cyan; SRP RNA, dark orange; signal sequence and TMD, magenta.

**e-g**, Comparison of the structure of the RNC•SecA complex (g; this work) with the RNC•SecYEG (e; PDB ID: 3J46) and SecA•SecYEG (f; PDB ID: 5EUL) structures. The color scheme is: SecA, orange; SecYEG, blue; uL23, green; H59, red; RodZ TMD, magenta.

**h**, The structures of SecA- (PDB ID 5EUL) and RNC-bound SecYEG (PDB ID: 3J46) were overlaid to show the steric clash between SecA and RNC on SecYEG.

**Table S2.1. Data collection, structure model refinement, and validation statistics.**

<b>Data Collection</b>	<b>RNC<sub>RodZ</sub>•SecA</b>
Total number of particles	2,613,025
Voltage (kV)	300
Electron dose (e <sup>-</sup> /Å <sup>2</sup> )	40
Pixel size (Å)	1.39
Defocus range (μm)	1.5 - 2.8
<b>Data Refinement</b>	
Final number of particles	37,334
Sharpening B-factor (Å <sup>2</sup> )	-130.53
Resolution at FSC=0.143 (Å)	3.1
Resolution (model vs. map) FSC=0.5 (Å)	3.3
<b>r.m.s. deviations</b>	
Bond length (Å)	0.005
Angles (°)	0.776
<b>Average B-factors (min/max/mean)</b>	
Protein	55.69/522.79/148.63
RNA	57.83/405.65/95.06
Ligand	44.14/95.51/60.40
<b>Validation statistics</b>	
Molprobability score	1.52
Clashscore, all atoms	4.21
<b>Protein</b>	
Favored rotamers (%)	92.98
Ramachandran plot	
Favored (%)	95.44
Allowed (%)	4.28
Outliers (%)	0.28
<b>RNA</b>	
Correct sugar puckers (%)	99.73
Good backbone conformation (%)	84.66

## References

- Afonine, P.V., Klaholz, B.P., Moriarty, N.W., Poon, B.K., Sobolev, O.V., Terwilliger, T.C., Adams, P.D., and Urzhumtsev, A. (2018a). New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr D Struct Biol* 74, 814-840.
- Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018b). Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr D Struct Biol* 74, 531-544.
- Akita, M., S. Sasaki, S. Matsuyama, and S. Mizushima. (1990). SecA interacts with secretory proteins by recognizing the positive charge at the amino terminus of the signal peptide in *Escherichia coli*. *J Biol Chem.* 265:8164-8169.
- Akopian, D., Dalal, K., Shen, K., Duong, F., and Shan, S.O. (2013a). SecYEG activates GTPases to drive the completion of cotranslational protein targeting. *J Cell Biol* 200, 397-405.
- Akopian, D., Shen, K., Zhang, X., and Shan, S.O. (2013b). Signal recognition particle: an essential protein-targeting machine. *Annual review of biochemistry* 82, 693-721.
- Ariosa, A., J.H. Lee, S. Wang, I. Saraogi, and Shan, S.O. (2014). Regulation by a chaperone improves substrate selectivity during cotranslational protein targeting. *Proc Natl Acad Sci U S A.* 112:E3169-3178.
- Ast, T., Cohen, G., and Schuldiner, M. (2013). A network of cytosolic factors targets SRP-independent proteins to the endoplasmic reticulum. *Cell* 152, 1134-1145.
- Aviram, N., Ast, T., Costa, E.A., Arakel, E.C., Chuartzman, S.G., Jan, C.H., Hassdenteufel, S., Dudek, J., Jung, M., Schorr, S., et al. (2016). The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. *Nature* 540, 134-138.
- Bauer, B., T. Shemesh, Y. Chen, and T. Rapoport. (2014). A "push and slide" mechanism allows sequence-insensitive translocation of secretory proteins by the SecA ATPase. *Cell.* 157:1416-1429.
- Bauer, B.W., and Rapoport, T.A. (2009). Mapping polypeptide interactions of the SecA ATPase during translocation. *Proc Natl Acad Sci U S A* 106, 20800-20805.
- Behrmann, M., H. Koch, T. Hengelage, B. Wieseler, H. Hoffschulte, and M. Müller. (1998). Requirements for the translocation of elongation-arrested, ribosome-associated OmpA across the plasma membrane of *Escherichia coli*. *J. Biol. Chem.* 273:13898-13904.
- Blanchette, C.D., Law, R., Benner, W.H., Pesavento, J.B., Cappuccio, J.A., Walsworth, V., Kuhn, E.A., Corzett, M., Chromy, B.A., Segelke, B.W., et al. (2008). Quantifying size distributions of nanolipoprotein particles with single-particle analysis and molecular dynamic simulations. *J Lipid Res* 49, 1420-1430.

- Brandman, O., and Hegde, R.S. (2016). Ribosome-associated protein quality control. *Nat Struct Mol Biol* 23, 7-15.
- Calhoun, K.A., and J.R. Swartz. (2006). Total amino acid stabilization during cell-free protein synthesis reactions. *J. Biotechnol.* 123:193-203.
- Chartron, J.W., Hunt, K.C., and Frydman, J. (2016). Cotranslational signal-independent SRP preloading during membrane targeting. *Nature* 536, 224-228.
- Chatzi, K.E., Sardis, M.F., Tsirigotaki, A., Koukaki, M., Sostaric, N., Konijnenberg, A., Sobott, F., Kalodimos, C.G., Karamanou, S., and Economou, A. (2017). Preprotein mature domains contain translocase targeting signals that are essential for secretion. *J Cell Biol* 216, 1357-1369.
- Chitwood, P.J., Juszkiwicz, S., Guna, A., Shao, S., and Hegde, R.S. (2018). EMC Is Required to Initiate Accurate Membrane Protein Topogenesis. *Cell* 175, 1507-1519.e1516.
- Cooper, A. (2004). *Biophysical Chemistry*. Royal Society of Chemistry.
- Cymer, F., von Heijne, G., and White, S.H. (2014). Mechanisms of integral membrane protein insertion and folding. *J Mol Biol* 427, 999-1022.
- Datsenko, K.A., and B.L. Wanner. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* 97:6040-6045.
- de Gier, J., P. Mansournia, Q. Valent, G. Phillips, J. Luirink, and G. von Heijne. (1996). Assembly of a cytoplasmic membrane protein in *Escherichia coli* is dependent on the signal recognition particle. *FEBS. Lett.* 399:307-309.
- Devaraneni, P.K., Conti, B., Matsumura, Y., Yang, Z., Johnson, A.E., and Skach, W.R. (2011). Stepwise insertion and inversion of a type II signal anchor sequence in the ribosome-Sec61 translocon complex. *Cell* 146, 134-147.
- Elvekrog, M.M., and Walter, P. (2015). Dynamics of co-translational protein targeting. *Curr Opin Chem Biol* 29, 79-86.
- Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60, 2126-2132.
- Erlanson, K.J., Miller, S.B., Nam, Y., Osborne, A.R., Zimmer, J., and Rapoport, T.A. (2008). A role for the two-helix finger of the SecA ATPase in protein translocation. *Nature* 455, 984-987.
- Estrozi, L., D. Boehringer, S. Shan, N. Ban, and C. Schaffitzel. (2011). Cryo-EM structure of the *E. coli* translating ribosome in complex with SRP and its receptor. *Nat Struct Mol Biol.* 18:88-90.

- Fekkes, P., J. de Wit, J. van der Wolk, H. Kimsey, C. Kumamoto, and A. Driessen. (1998). Preprotein transfer to the Escherichia coli translocase requires the co-operative binding of SecB and the signal sequence to SecA. *Mol Microbiol.* 29:1179-1190.
- Ferbitz, L., Maier, T., Patzelt, H., Bukau, B., Deuerling, E., and Ban, N. (2004). Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* 431, 590-596.
- Frauenfeld, J., Gumbart, J., Sluis, E.O., 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.
- Gamerding, M., M. Hanebuth, T. Frickey, and E. Deuerling. 2015. The principle of antagonism ensures protein targeting specificity at the endoplasmic reticulum. *Science.* 348:201-207.
- Gebert, J.F., Overhoff, B., Manson, M.D., and Boos, W. (1988). The Tsr chemosensory transducer of Escherichia coli assembles into the cytoplasmic membrane via a SecA-dependent process. *J Biol Chem* 263, 16652-16660.
- Gelis, I., Bonvin, A.M., Keramisanou, D., Koukaki, M., Gouridis, G., Karamanou, S., Economou, A., and Kalodimos, C.G. (2007). Structural basis for signal-sequence recognition by the translocase motor SecA as determined by NMR. *Cell* 131, 756-769.
- Gibson, D.G., L. Young, R.Y. Chuang, J.C. Venter, C.A. Hutchison, 3rd, and H.O. Smith. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods.* 6:343-345.
- Gouridis, G., S. Karamanou, I. Gelis, C.G. Kalodimos, and A. Economou. (2009). Signal peptides are allosteric activators of the protein translocase. *Nature.* 462:363-367.
- Guna, A., Volkmar, N., Christianson, J.C., and Hegde, R.S. (2018). The ER membrane protein complex is a transmembrane domain insertase. *Science* 359, 470-473.
- Hartl, F.U., Lecker, S., Schiebel, E., Hundrick, J.P., and 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.
- Helde, R., B. Wiesler, E. Wachter, A. Neubüser, H. Hoffschulte, T. Hengelage, K. Schimz, R. Stuart, and M. Müller. (1997). Comparative characterization of SecA from the alpha-subclass purple bacterium *Rhodobacter capsulatus* and *Escherichia coli* reveals differences in membrane and precursor specificity. *J Bacteriol.* 179:4003-4012.
- Hikita, C., and S. Mizushima. (1992). The Requirement of a Positive Charge at the Amino Terminus Can be Compensated for by a Longer Central Hydrophobic Stretch in the Functioning of Signal Peptides. *J Biol Chem.* 267:12375-12379.

Hoffschulte, H., B. Drees, and M. Müller. (1994). Identification of a soluble SecA/SecB complex by means of a subfractionated cell-free export system. *J Biol Chem.* 269:12833-12839.

Huber, D., D. Boyd, Y. Xia, M.H. Olma, M. Gerstein, and J. Beckwith. (2005a). Use of thioredoxin as a reporter to identify a subset of *Escherichia coli* signal sequences that promote signal recognition particle-dependent translocation. *J Bacteriol.* 187:2983-2991.

Huber, D., Jamshad, M., Hanmer, R., Schibich, D., Döring, K., Marcomini, I., Kramer, G., and Bukau, B. (2016). SecA cotranslationally interacts with nascent substrate proteins in vivo. *J Bacteriol* 199, e00622-00616.

Huber, D., M.I. Cha, L. Debarbieux, A.G. Planson, N. Cruz, G. Lopez, M.L. Tasayco, A. Chaffotte, and J. Beckwith. (2005b). A selection for mutants that interfere with folding of *Escherichia coli* thioredoxin-1 in vivo. *Proc Natl Acad Sci U S A.* 102:18872-18877.

Huber, D., Rajagopalan, N., Preissler, S., Rocco, M.A., Merz, F., Kramer, G., and Bukau, B. (2011). SecA Interacts with Ribosomes in Order to Facilitate Posttranslational Translocation in Bacteria. *Mol Cell* 41, 343-353.

Jagath, J., M. Rodnina, and W. Wintermeyer. (2000). Conformational changes in the bacterial SRP receptor FtsY upon binding of guanine nucleotides and SRP. *J Mol Biol.* 295:745-753.

Jomaa, A., Boehringer, D., Leibundgut, M., and Ban, N. (2016). Structures of the *E. coli* translating ribosome with SRP and its receptor and with the translocon. *Nat Commun* 7, 10471.

Jomaa, A., Fu, Y.H., Boehringer, D., Leibundgut, M., Shan, S.O., and Ban, N. (2017). Structure of the quaternary complex between SRP, SR, and translocon bound to the translating ribosome. *Nat Commun* 8, 15470.

Kajava, A., S. Zolov, A. Kalinin, and M. Nesmeyanova. (2000). The net charge of the first 18 residues of the mature sequence affects protein translocation across the cytoplasmic membrane of gram-negative bacteria. *J Bacteriol.* 182:2163-2169.

Karamyshev, A.L., and A.E. Johnson. (2005). Selective SecA association with signal sequences in ribosome-bound nascent chains: a potential role for SecA in ribosome targeting to the bacterial membrane. *J Biol Chem.* 280:37930-37940.

Kihara, A., and K. Ito. (1998). Translocation, folding, and stability of the HflKC complex with signal anchor topogenic sequences. *J Biol Chem.* 273:29770-29775.

Kimura, E., Akita, M., Matsuyaman, S., and Mizushima, S. (1991). Determination of a region in SecA that interacts with presecretory proteins in *Escherichia coli*. *J Biol Chem* 266, 6600-6606.

Knorr, A.G., Schmidt, C., Tesina, P., Berninghausen, O., Becker, T., Beatrix, B., and Beckmann, R. (2019). Ribosome-NatA architecture reveals that rRNA expansion segments coordinate N-terminal acetylation. *Nat Struct Mol Biol* 26, 35-39.

- Koch, S., De Wit, J.G., Vos, I., Birkner, J.P., Gordiichuk, P., Herrmann, A., Van Oijen, A.M., and Driessen, A.J.M. (2016). Lipids activate SecA for high affinity binding to the SecYEG complex. *J Biol Chem* 291, 22534-22543.
- Kuruma, Y., K. Nishiyama, Y. Shimizu, M. Müller, and T. Ueda. (2008). Development of a minimal cell-free translation system for the synthesis of presecretory and integral membrane proteins. *Biotechnol. Prog.* 21:1243-1251.
- Li, L., Park, E., Ling, J., Ingram, J., Ploegh, H., and Rapoport, T.A. (2016). Crystal structure of a substrate-engaged SecY protein-translocation channel. *Nature* 531, 395-399.
- Lindner, E., and S.H. White. (2014). Topology, dimerization, and stability of the single-span membrane protein CadC. *J Mol Biol.* 426:2942-2957.
- Ludtke, S.J., Baldwin, P.R., and Chiu, W. (1999). EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* 128, 82-97.
- Luirink, and S. Bhushan. (2014). Cryo-electron microscopic structure of SecA bound to the 70S ribosome. *J Biol Chem.* 289:7190-7199.
- Luirink, J., and I. Sinning. (2004). SRP-mediated protein targeting: structure and function revisited. *Biochim Biophys Acta.* 1694:17-35.
- Müller, M., and G. Blobel. (1984a). In vitro translocation of bacterial proteins across the plasma membrane of *Escherichia coli*. *Proc Natl Acad Sci U S A.* 81:7421-7425.
- Müller, M., and G. Blobel. (1984b). Protein export in *Escherichia coli* requires a soluble activity. *Proc Natl Acad Sci U S A.* 81:7737-7741.
- Nakatogawa, H., and Ito, K. (2002). The ribosomal exit tunnel functions as a discriminating gate. *Cell* 108, 629-636.
- Neumann-Haefelin, C., U. Schäfer, M. Müller, and H. Koch. (2000). SRP-dependent co-translational targeting and SecA-dependent translocation analyzed as individual steps in the export of a bacterial protein. *EMBO J.* 19:6419-6426.
- Nyathi, Y., Wilkinson, B.M., and Pool, M.R. (2013). Co-translational targeting and translocation of proteins to the endoplasmic reticulum. *Biochim Biophys Acta* 1833, 2392-2402.
- Or, E., D. Boyd, S. Gon, J. Beckwith, and T. Rapoport. (2005). The bacterial ATPase SecA functions as a monomer in protein translocation. *J. Biol. Chem.* 280:9097-9105.
- Or, E., D. Boyd, S. Gon, J. Beckwith, and T. Rapoport. (2005). The bacterial ATPase SecA functions as a monomer in protein translocation. *J Biol Chem.* 280:9097-9105.

- Or, E., Navon, A., and Rapoport, T. (2002). Dissociation of the dimeric SecA ATPase during protein translocation across the bacterial membrane. *EMBO J* 21, 4470-4479.
- Osborne, A.R., Clemons Jr., W.M., and Rapoport, T.A. (2004). Monomer closed PPXD; dimer open PPXD. *Proc Natl Acad Sci U S A* 101, 10937-10942.
- Park, E., Menetret, J.F., Gumbart, J.C., Ludtke, S.J., Li, W., Whynot, A., Rapoport, T.A., and Akey, C.W. (2014). Structure of the SecY channel during initiation of protein translocation. *Nature* 506, 102-106.
- Peluso, P., D. Herschlag, S. Nock, D. Freymann, A. Johnson, and P. Walter. (2000). Role of 4.5S RNA in assembly of the bacterial signal recognition particle with its receptor. *Science*. 288:1640-1643.
- Peterson, J.H., C.A. Woolhead, and H.D. Bernstein. (2003). Basic amino acids in a distinct subset of signal peptides promote interaction with the signal recognition particle. *J Biol Chem*. 278:46155-46162.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25, 1605-1612.
- Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods* 14, 290-296.
- Qi, H., and H. Bernstein. (1999). SecA is required for the insertion of inner membrane proteins targeted by the Escherichia coli signal recognition particle. *J. Biol. Chem.* 274:8893-8897.
- Randall, L. (1983). Translocation of domains of nascent periplasmic proteins across the cytoplasmic membrane is independent of elongation. *Cell*. 33:231-240.
- Ranjan, A., Mercier, E., Bhatt, A., and Wintermeyer, W. (2017). Signal recognition particle prevents N-terminal processing of bacterial membrane proteins. *Nat Commun* 8, 15562.
- Rawat, S., Zhu, L., Lindner, E., Dalbey, R.E., and White, S.H. (2015). SecA drives transmembrane insertion of RodZ, an unusual single-span membrane protein. *J Mol Biol* 427, 1023-1037.
- Rome, M.E., Chio, U.S., Rao, M., Gristick, H., and Shan, S.O. (2014). Differential gradients of interaction affinities drive efficient targeting and recycling in the GET pathway. *Proc Natl Acad Sci U S A* 111, E4929-4935.
- Sääf, A., Andersson, H., Gafvelin, G., and von Heijne, G. (1995). SecA-dependence of the translocation of a large periplasmic loop in the Escherichia coli MalF inner membrane protein is a function of sequence context. *Mol Membr Biol* 12, 209-215.

Samuelson, J., M. Chen, F. Jiang, I. Möller, M. Wiedmann, A. Kuhn, G. Phillips, and R. Dalbey. (2000). YidC mediates membrane protein insertion in bacteria. *Nature*. 406:637-641.

Saraogi, I., Akopian, D., and Shan, S.O. (2014). Regulation of cargo recognition, commitment, and unloading drives cotranslational protein targeting. *J Cell Biol* 205, 693-706.

Saraogi, I., Zhang, D., Chandrasekaran, S., and Shan, S.O. (2011). Site-specific fluorescent labeling of nascent proteins on the translating ribosome. *Journal of the American Chemical Society* 133, 14936-14939.

Schaffitzel, C., M. Oswald, I. Berger, T. Ishikawa, J. Abrahams, H. Koerten, R. Koning, and N. Ban. (2006). Structure of the *E. coli* signal recognition particle bound to a translating ribosome. *Nature*. 444:503-506.

Schibich, D., Gloge, F., Pohner, I., Bjorkholm, P., Wade, R.C., von Heijne, G., Bukau, B., and Kramer, G. (2016). Global profiling of SRP interaction with nascent polypeptides. *Nature* 536, 219-223.

Schierle, C.F., M. Berkmen, D. Huber, C. Kumamoto, D. Boyd, and J. Beckwith. (2003). The DsbA Signal Sequence Directs Efficient, Cotranslational Export of Passenger Proteins to the *Escherichia coli* Periplasm via the Signal Recognition Particle Pathway. *J Bacteriol*. 185:5706-5713.

Scotti, P., Q. Valent, E. Manting, M. Urbanus, A. Driessen, B. Oudega, and J. Luirink. (1999). *J Biol Chem*. *J Biol Chem*. 274:29883-29888.

Shao, S., and Hegde, R.S. (2015). Target Selection during Protein Quality Control. *Trends in biochemical sciences* 41, 124-137.

Sharma, V., Arockiasamy, A., Ronning, D.R., Savva, C.G., Holzenburg, A., Braunstein, M., Jacobs, W.R., Jr., and Sacchettini, J.C. (2003). Crystal structure of *Mycobacterium tuberculosis* SecA, a preprotein translocating ATPase. *Proc Natl Acad Sci U S A* 100, 2243-2248.

Shimizu, Y., A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, and U. T. (2001). Cell-free translation reconstituted with purified components. *Nat Biotechnol*. 19:751-755.

Shurtleff, M.J., Itzhak, D.N., Hussmann, J.A., Schirle Oakdale, N.T., Costa, E.A., Jonikas, M., Weibezahn, J., Popova, K.D., Jan, C.H., Sinitcyn, P., et al. (2018). The ER membrane protein complex interacts cotranslationally to enable biogenesis of multipass membrane proteins. *eLife* 7, 1-23.

Siegel, V., and P. Walter. (1988). The affinity of signal recognition particle for presecretory proteins is dependent on nascent chain length. *EMBO J*. 7:1769-1775.

- Singh, R., Kraft, C., Jaiswal, R., Sejwal, K., Kasaragod, V.B., Kuper, J., Bürger, J., Mielke, T., Luirink, J., and Bhushan, S. (2014). Cryo-electron microscopic structure of SecA bound to the 70S ribosome. *J Biol Chem* 289, 7190-7199.
- Stray, S., J. Johnson, B. Kopek, and A. Zlotnick. (2006). An in vitro fluorescence screen to identify antivirals that disrupt hepatitis B virus capsid assembly. *Nat Biotechnol.* 24:358-362.
- Traxler, B., and Murphy, C. (1996). Insertion of the polytopic membrane protein MalF is dependent on the bacterial secretion machinery. *J Biol Chem* 271, 12394-12400.
- Ulbrandt, N., J. Newitt, and H. Bernstein. (1997). The E. coli signal recognition particle is required for the insertion of a subset of inner membrane proteins. *Cell.* 88:187-196.
- van der Laan, M., P. Bechtluft, S. Kol, N. Nouwen, and A.J. Driessen. (2004). F1F0 ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis. *J Cell Biol.* 165:213-222.
- von Loeffelholz, O., K. Knoops, A. Ariosa, X. Zhang, M. Karuppasamy, K. Huard, G. Schoehn, I. Berger, S.O. Shan, and C. Schaffitzel. (2013). Structural basis of signal sequence surveillance and selection by the SRP-FtsY complex. *Nat Struct Mol Biol.* 20:604-610.
- Walter, P., I. Ibrahim, and G. Blobel. (1981). Translocation of proteins across the endoplasmic reticulum. I. Signal recognition protein (SRP) binds to in-vitro-assembled polysomes synthesizing secretory protein. *J Cell Biol.* 91:545-550.
- Wang, S., Yang, C.I., and Shan, S.O. (2017). SecA mediates cotranslational targeting and translocation of an inner membrane protein. *J Cell Biol* 216, 3639-3653.
- Weiss, J., P. Ray, and P.J. Bassford. (1988). Purified secB protein of Escherichia coli retards folding and promotes membrane translocation of the maltose-binding protein in vitro. *Proc Natl Acad Sci U S A.* 85:8978-8982.
- Wickstrom, D., S. Wagner, L. Baars, A.J. Ytterberg, M. Klepsch, K.J. van Wijk, J. Luirink, and J.W. de Gier. (2011). Consequences of depletion of the signal recognition particle in Escherichia coli. *J Biol Chem.* 286:4598-4609.
- Wolfe, P., M. Rice, and W. Wickner. (1985). Effects of two sec genes on protein assembly into the plasma membrane of Escherichia coli. *J Biol Chem.* 260:1836-1841.
- Wu, Z.C., De Keyser, J., Kedrov, A., and Driessen, A.J.M. (2012). Competitive binding of the SecA ATPase and ribosomes to the SecYEG translocon. *J Biol Chem* 287, 7885-7895.
- Zhang, D., M. Sweredoski, R. Graham, S. Hess, and S. Shan. (2012). Novel proteomic tools reveal essential roles of SRP and importance of proper membrane protein biogenesis. *Mol Cell Proteomics.* 11:M111.011585.

- Zhang, J., X. Pan, K. Yan, S. Sun, N. Gao, and S.F. Sui. 2015. Mechanisms of ribosome stalling by SecM at multiple elongation steps. *Elife*. 4:pii: e09684.
- Zhang, K. (2016). Gctf: Real-time CTF determination and correction. *J Struct Biol* 193, 1-12.
- Zhang, X., and Shan, S.O. (2014). Fidelity of cotranslational protein targeting by the signal recognition particle. *Annual review of biophysics* 43, 381-408.
- Zhang, X., Rashid, R., Wang, K., and Shan, S.O. (2010). Sequential checkpoints govern substrate selection during cotranslational protein targeting. *Science* 328, 757-760.
- Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y., and Agard, D.A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* 14, 331-332.
- Ziehr, D.R., J.P. Ellis, P.H. Culviner, and S. Cavagnero. (2010). Production of ribosome-released nascent proteins with optimal physical properties. *Anal. Chem.* 82:4637-4643.
- Zimmer, J., Nam, Y., and Rapoport, T.A. (2008). Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature* 455, 936-943.
- Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* 7.
- Zoschke, R., and A. Barkan. (2015). Genome-wide analysis of thylakoid-bound ribosomes in maize reveals principles of cotranslational targeting to the thylakoid membrane. *Proc Natl Acad Sci U S A.* 112:E1678-1687.