# Structural Analysis of Membrane Protein Biogenesis and Ribosome Stalling by Cryo-Electron Microscopy

Lukas Sebastian Bischoff

aus Gräfelfing

2015

**Parts of this thesis have been published:**


Visualization of a polytopic membrane protein during SecY-mediated membrane insertion.
**Bischoff L**, Wickles S, Berninghausen O, van der Sluis EO, Beckmann R.
Nature Communications. 2014 Jun 10;5:4103. doi: 10.1038/ncomms5103.


Molecular Basis for the Ribosome Functioning as an L-Tryptophan Sensor.
**Bischoff L**, Berninghausen O, Beckmann R.
Cell Reports. 2014 Oct 7. pii: S2211-1247(14)00779-7. doi: 10.1016/j.celrep.2014.09.011.


A structural model of the active ribosome-bound membrane protein insertase YidC.
Wickles S, Singharoy A, Andreani J, Seemayer S, **Bischoff L**, Berninghausen O, Soeding J, Schulten K, van der Sluis EO, Beckmann R.
Elife. 2014 Jul 10;3:e03035. doi: 10.7554/eLife.03035.


Dynamic behavior of Trigger Factor on the ribosome
Deeng J, Chan KY, van der Sluis EO, **Bischoff L**, Berninghausen O, Han W, Gumbart J, Schulten K, Beatrix B. and Beckmann R.
*SUBMITTED*


**Parts of this thesis have been presented at international conferences:**

Kuo Symposium on 3D cryo-EM Molecular Imaging,
26.07.14 – 30.07.14, Shanghai, China, poster presentation


Gordon Research Conference: 'Three-Dimensional Electron Microscopy',
22.06.14 – 27.06.14, Girona, Spain, poster presentation


Gordon Research Conference: 'Protein Transport Across Cell Membranes',
09.03.14 – 14.03.14 Galveston, USA, oral and poster presentation


EMBO Conference Series: 'From Structure to Function of Translocation Machines',
13.04.13 – 17.04.13, Dubrovnik, Croatia, poster presentation


Gordon Research Conference: 'Protein Transport Across Cell Membranes',
11.03.12 – 16.03.12 Galveston, USA, oral and poster presentation

# Table of Content

# 1. Introduction

## 1.1 Protein translocation in bacteria

About 25 % to 30 % of all proteins in bacteria have to be translocated from the cytoplasm either across or into the plasma membrane. Bacteria employ several different strategies to conduct this transport; however, the vast majority of the transported proteins use a universal protein conducting channel, the SecYEG translocon to transport proteins over or into the inner membrane. The Sec translocon is conserved in all three domains of life to transport proteins in or across cell membranes.

Nevertheless, the protein-conducting channel is only one part of the translocation machinery in the bacterial membrane (see Figure 1). Depending on the substrate, several other membrane proteins like the insertase and membrane chaperone YidC (that also inserts membrane proteins independently) the SecDF complex or the signal peptide peptidase (LepB) associate with SecY and assist in the translocation, lateral insertion or maturation of proteins. For review see (Driessen and Nouwen, 2008; Park and Rapoport, 2012b)
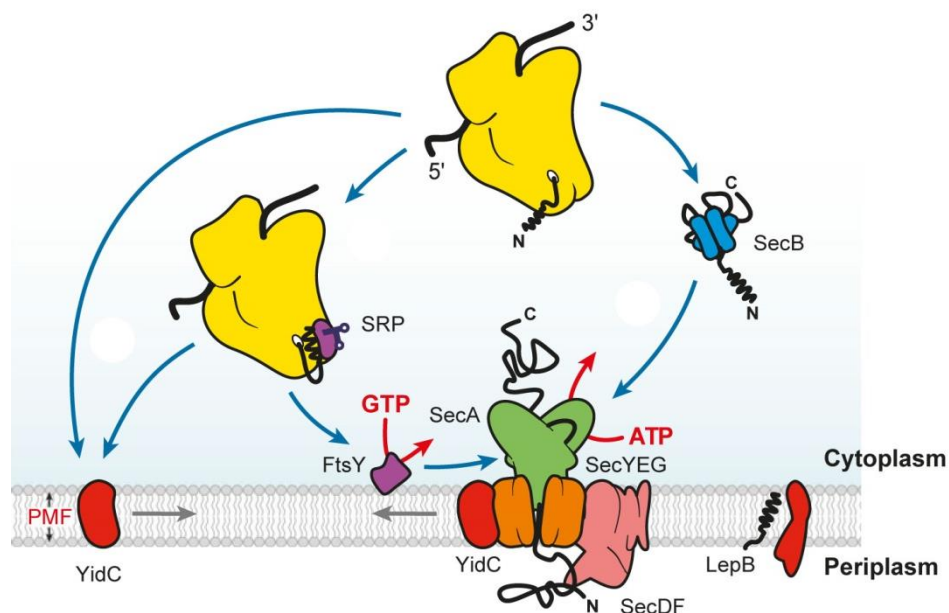


**Figure 1: Schematic overview over protein targeting to the Sec translocon and the protein insertase YidC**
The nascent membrane protein can get targeted by SRP to either the Sec translocon or YidC.
Secretory proteins are translocated post-translationally vie the SecA/B pathway.
Figure modified from (Driessen and Nouwen, 2008).

Since all secretory and inner membrane proteins are synthesized at ribosomes in the cytoplasm they need to be recognized and targeted to the cell membrane and the Sec translocon, respectively.

There are two general modes of protein translocation: the post-translational translocation and the co-translational translocation.

Secretory proteins that are generally transported post-translationally are synthesized in the cytoplasm with an N-terminal signal sequence that is not a part of the mature protein and needs to be cleaved at a yet unknown stage of protein translocation (von Heijne, 1990). Here, the ATPase SecA recognizes and binds the signal sequence of the secretory protein and together with the small chaperone SecB guides the nascent peptide to the Sec translocon. The poly-peptide is then translocated through the protein-conducting channel with assistance of SecA that utilizes the energy derived from ATP hydrolysis (recently reviewed by (Park and Rapoport, 2012b)).

Inner membrane proteins are translocated in a co-translational mode. The targeting signal for inner membrane proteins can either be a cleavable signal sequence as for the secretory proteins, but also a hydrophobic stretch of amino acids somewhere at the N-terminus of the nascent membrane protein that remains as the first transmembrane domain in the mature protein. The latter is referred to as a signal anchor.

As soon as the signal sequence or the signal anchor emerges from the ribosomal exit tunnel, it is recognized by the universally conserved SRP/SR system. The ribonucleoprotein signal recognition particle (SRP) recognizes and binds the signal sequence or the signal anchor of the nascent membrane protein and targets the whole ribosome nascent chain complex to the membrane. Here SRP interacts with its membrane-bound counterpart, the SRP receptor (SR). Via a still unknown mechanism that involves GTP hydrolysis in both, SRP and SR, the RNC is handed over to the Sec translocon. (for review see (Zhang and Shan, 2014)). The ribosome binds now directly to the protein-conducting channel and by continuing translation, the protein gets integrated into the membrane. Thereby, incoming transmembrane segments must be recognized, most likely by the Sec translocon itself, and then laterally released into the membrane. These transmembrane segments can be predicted by computational means as they usually comprise hydrophobic stretches of 15-30 residues in length.

## 1.1.1 The SecYEG translocon

The SecYEG translocon consists of three subunits, SecY (Sec61α in eukaryotes), SecE (Sec61γ) and SecG (Sec61β) with SecY forming the actual pore for protein secretion. This heterotrimeric composition of the translocon complex is conserved in all kingdoms of life (Brundage et al., 1990; Akimaru et al., 1991; Gorlich and Rapoport, 1993). While SecY and SecE are essential and also show sequence conservation among different species, the SecG subunit is not universally conserved and also not essential for cell viability.

In 2004 Van den Berg *et al*. solved the crystal structure of *Methanococcus jannaschii* SecYEβ and thus provided important insights in the mechanisms of proteins translocation by the Sec system (see Figure 2).

SecY consists of 10 transmembrane helices that can be divided into two pseudo-symmetric halves, transmembrane segments (TMS) 1-5 and TMS 6-10 that form a central pore. The loop between TMS 5 and 6 can be seen as a hinge that allows an opening between TMS 2b and 7 on the other side of SecY. This opening between TMS 2b and 7 is known as the "lateral gate" of SecY and is most likely responsible for the lateral release of transmembrane domains into the membrane by opening (Tsukazaki et al., 2008; Zimmer et al., 2008; Egea and Stroud, 2010).

In the plane of the membrane, SecY forms an hourglass-shaped pore with a water-filled cytoplasmic and periplasmic funnel. In an inactive state of SecY the periplasmic funnel is sealed by helix 2a of SecY, the so-called 'plug' domain. This 'plug' domain is thought to move upon protein translocation. The central constriction of the hourglass-shaped SecY is formed by six hydrophobic residues known as the 'pore-ring' (Van den Berg et al., 2004).

The second essential subunit, SecE, links the two halves of SecY together by a long transmembrane helix that extends diagonally across the interface of SecY. Additionally, N-terminal of that long transmembrane helix, SecE possesses an amphipathic helix that lies on the membrane close to TMS 8 and 9 of SecY and might assist ribosome binding. In *E.coli* SecE additionally has two N-terminal transmembrane segments, that are however not essential and whose function is still unclear (Murphy and Beckwith, 1994; Breyton et al., 2002).

The non-essential SecG makes only a few contacts to SecY in the area of TMS 4 and 5, which might explain why it is expendable for complex stability and function. The archeal Secβ has an additional transmembrane segment N-terminal of the conserved helix (Joly et al., 1994).
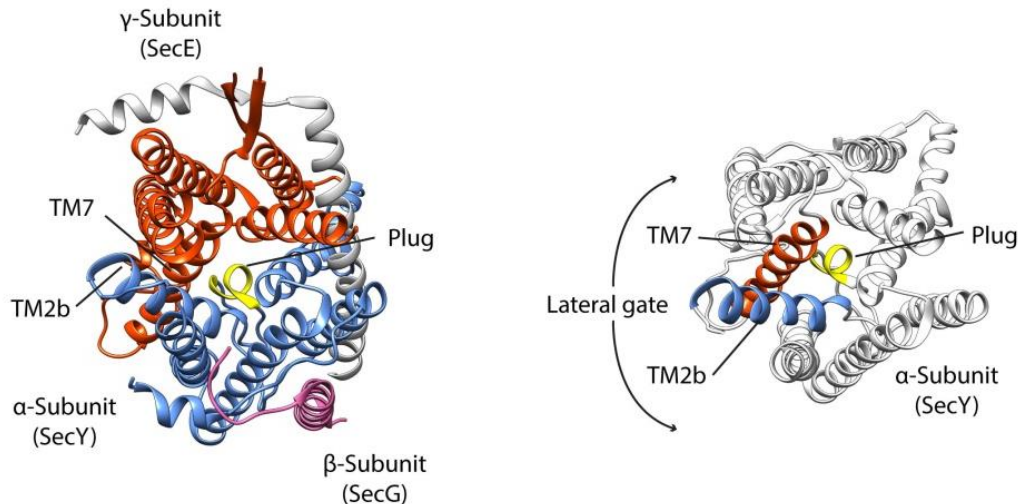
**Figure 2: Crystal structure of the *Methanococcus jannaschii* SecYEβ complex**
Left: View on the translocon from the cytoplasm, the C-terminal half is colored in red, the N-terminal half in blue. The 'plug' domain is highlighted in yellow. SecE is colored in grey and Secβ in purple.
Right: Highlight of the 'lateral gate' between helices 2b and 7.

Interestingly, various crystal structures of bacterial and archaeal SecYEG complexes all show the same overall architecture, however, the conformation of the lateral gate that is triggered by interaction of SecY with a ligand differs and thus already gives hints and insights in the potential function of SecY (Zimmer et al., 2008; Egea and Stroud, 2010). To date no crystal structure of a eukaryotic Sec61 is available; however, several cryo-electron microscopy-derived models suggest structural conservation of the protein-conducting channel (Becker et al., 2009; Gogala et al., 2014; Voorhees et al., 2014).


## 1.1.2 Co-translation protein targeting by SRP and FtsY

In co-translational translocation the signal recognition particle, a ubiquitous ribonucleoprotein conserved in all domains of life, recognizes and binds the first hydrophobic domain of either the signal sequence or the signal anchor as soon as it emerges from the ribosomal exit tunnel (Blobel and Dobberstein, 1975b; Blobel and Dobberstein, 1975a; Walter and Blobel, 1980; Gilmore et al., 1982).
In eukaryotes, binding of SRP to the ribosome induces a stop or slow-down of the translation, however that was not observed in bacteria as the SRP is here of much less complexity. The *E.coli* SRP consists of a 4.5S RNA and the GTPase Ffh (fifty-four homologue) (Grudnik et al., 2009). The Ffh-protein consists of a NG-domain that harbors the GTPase function and a M-domain (Keenan et al., 1998). The methionine-rich M domain contains a deep groove that is responsible for binding of the signal sequence or the signal anchor

(Ramirez et al., 2002). Close to this groove in the M domain binds the conserved domain IV of the SRP RNA suggesting that the binding site consists of both, protein and RNA. For review see (Zhang and Shan, 2014).

SPR bound to a ribosome nascent chain complex then interacts with its receptor FtsY that is bound to the inner membrane (Angelini et al., 2005). FtsY contains a Ffh-homologous NG domain with GTPase activity (Montoya et al., 1997). Additionally, it contains an acidic A-domain that is believed to be important for binding of FtsY to the membrane via anionic phospholipids and for the interaction of FtsY with SecY (Powers and Walter, 1997; Angelini et al., 2006; Braig et al., 2009). The interaction of SRP and FtsY is facilitated mainly by an interaction of the two NG domains (NG-Twin) in the GTP bound state (Shan and Walter, 2005; Reyes et al., 2007). Recent structural and biochemical studies proposed a model in which the two NG-domains assemble in a complex at the tetra-loop on the SRP-RNA and then re-localize to the distal end where GTP hydrolysis is activated (Ataide et al., 2011; Shen et al., 2012). The interaction of the two GTP-bound NG-domains reciprocally stimulates GTP hydrolysis that is followed by a dissociation of SRP from FtsY and the ribosome. During this GTP hydrolysis and SRP dissociation the ribosome nascent chain complex with the emerged signal sequence is handed over to SecY via a still unknown mechanism (Halic et al., 2004; Bradshaw and Walter, 2007).
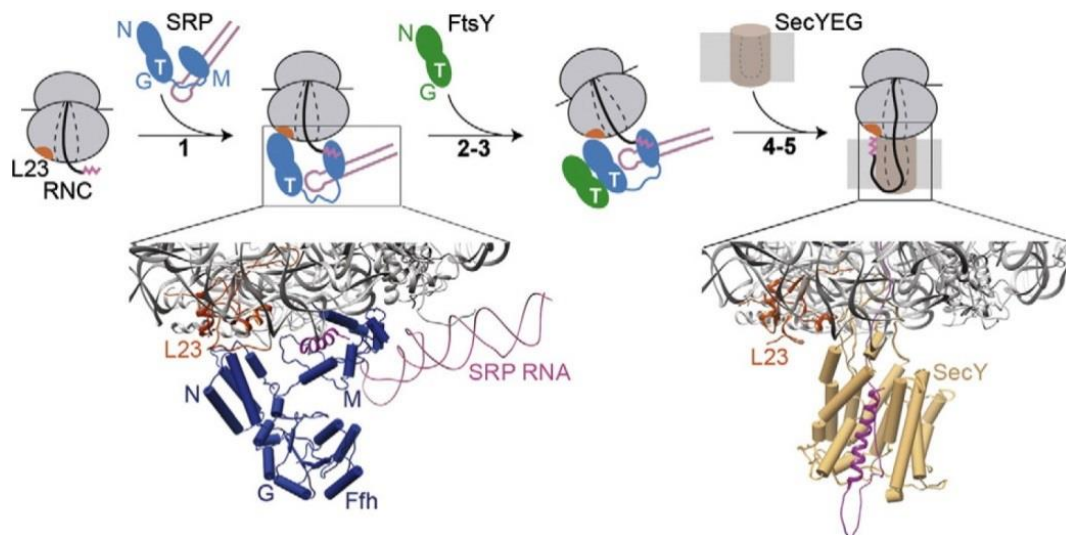


**Figure 3: Overview of the co-translational targeting by *E.coli* SRP**
1. The ribosome nascent chain complex with an emerged signal sequence (magenta) is recognized by SRP. The molecular model shows the signal sequence bound in the M-domain of SRP.
2. The loaded SRP binds its receptor (SR) on the membrane via their homologues NG-domains.
3. The signal sequence and the ribosome nascent chain complex is handed over to the SecY translocon. The SRP/SR complex dissociates upon GTP hydrolysis.
Figure modified from (Saraogi and Shan, 2014).

## 1.1.4 Post-translation protein translocation by the SecA/B pathway

SecA

The cytosolic ATP-dependent motor protein SecA consists of several domains: Two nucleotide-binding domains (NBD1 and 2), a helical wing domain (HWD), a polypeptide cross-linking domain (PPXD), and a helical scaffold domain that is composed of a long helix and two shorter helices, the 'two finger helices' (Hunt et al., 2002; Or et al., 2005; Vassylyev et al., 2006; Zimmer et al., 2008). Several crystal structures of SecA were obtained in different nucleotide-bound states and from different species that differ quite significantly in the relative position of PPXD and the HWD domains (Osborne et al., 2004; Papanikolau et al., 2007). The groove between these two domains is known as the clamp. This suggests that high flexibility in these domains and that the clamp is important for the function of SecA. Additionally, several different architectures for SecA dimers have been reported in solution and in crystal structures (Woodbury et al., 2002; Ding et al., 2003). To date the oligomeric state of SecA in solution and in active translocation is still subject of discussion and will not be further discussed here.

In 2008 Zimmer and co-workers reported the first crystal structure of SecA bound to SecY. In this structure, one copy of SecA is bound to SecY being oriented almost parallel to the membrane and binding is moderated by interactions of residues within the PPXD domain of SecA and loop 8/9 of SecY.

The two-finger loop of SecA is inserted into the cytoplasmic funnel of SecY right above the central pore. It was suggested that ATP hydrolysis of SecA results in an up and down movement of the two finger-loop, thus translocating an unfolded polypeptide (Bauer et al., 2014), however the actual mechanism of how SecA provides the force for protein translocation is still under debate.

SecB

Since SecA-dependent substrates are translocated post-translationally in their translocation-competent, largely unfolded state, they need to be hindered from folding in the cytoplasm.

This unfolded state is stabilized by the molecular chaperone SecB. SecB is organized as a homotetramer that forms a binding groove that might recognize distinct, hydrophobic features of secretory proteins.

It could be shown that a conserved stretch of amino acids at the C-terminus of SecA binds the SecB oligomer, thus SecA and SecB act in concert to translocate secretory proteins. For review see (Castanie-Cornet et al., 2014)

Model for the SecA/B mediated translocation of secretory proteins

Based on the current knowledge, the following model of SecA/B mediated post-translation translocation can be put forward. SecA binds the signal sequence of a fully synthesized secretory protein that is kept unfolded and in solution by several SecB oligomers. The oligomeric state of SecA at this stage is not fully understood. For review see (Park and Rapoport, 2012b). However, this view was challenged in recent publications where it was shown that SecA can bind the ribosome directly and might already bind to the signal sequence of the secretory protein before it is fully translated (Huber et al., 2011).

In the second step, SecA binds to SecY and starts to hydrolyze ATP that triggers translocation of the nascent substrate; SecB is released during this process. It is suggested that repeating cycles of ATP hydrolysis cause up and down movement of the two-finger helix and thus translocation of the nascent substrate. Back-sliding of the substrate would be hindered by the cleft-domain that grabs the substrate between the ATP hydrolysis cycles. Upon final translocation of the substrate, SecA dissociates from SecY (Erlandson et al., 2008; Zimmer et al., 2008; Bauer and Rapoport, 2009).
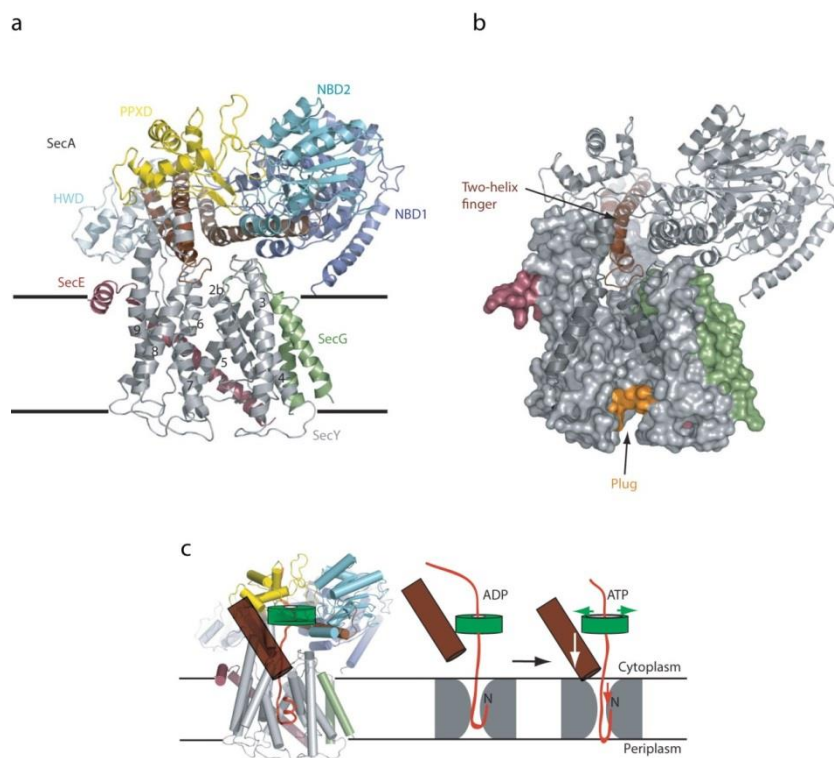


**Figure 4: Crystal structure of the *Thermatoga maritima* SecY-SecA complex and a hypothetical model of SecA-dependent protein translocation**
a-b: Atomic and space filled model of the SecY-SecA complex. The space filled model reveals the alignment of the hydrophilic channel in SecY with the two-helix finger of SecA.
c: The nascent substrate is thought to be pushed through the translocon by the two-helix finger upon cycles of ATP hydrolysis. The clamp prevents back-sliding of the substrate.
Figure modified from (Zimmer et al., 2008).

7

Another aspect of SecA-dependent protein translocation is the role of SecA is the biogenesis of inner membrane proteins with large periplasmic domains. Here, the translocation switches from a co-translational mode to post-translational. It has been shown for various substrates, that inner membrane proteins are firstly targeted and integrated into the membrane by the universal SRP/SR/SecY pathway; however, when a transmembrane segment is followed by a particular large periplasmic loop or domain, translocation of the latter is SecA-dependent (Neumann-Haefelin et al., 2000). That requires that the ribosome dissociates from SecY allowing SecA to bind the nascent substrate as they cannot bind to SecY at the same time (Kuhn et al., 2011; Wu et al., 2012) This alternating between co- and post-translational translocation is still subject of investigation and further addressed in the results section of this thesis.

## 1.1.5 Translocation of secretory proteins through the SecYEG complex

The most characterized reaction conducted by the SecYEG complex is the translocation of secretory substrates, when the signal sequence is targeted and inserted into SecY and the following hydrophilic polypeptide is translocated through the pore.

Interestingly, in a crystal structure of SecA bound to SecY (Zimmer et al., 2008) it could be observed that the proposed lateral gate of SecY between helices 2b and 7 is already more opened than that in the *M. jannaschii* crystal (Van den Berg et al., 2004). Furthermore, the co-crystallization of SecY with a Fab-fragment against the loop 8/9 also showed this 'pre-opened' conformation of SecY (Tsukazaki et al., 2008). Thus, it was proposed that binding of a co-factor to loop 8/9 of SecY already primes the translocon for the actual translocation process. This view was additionally supported by recent cryo-EM reconstructions of Sec complexes bound to non-translating ribosomes (Park et al., 2014; Voorhees et al., 2014).

Hence, the translocation of a secretory protein can be divided in three parts: First, the translocon gets primed for translocation by the binding of a co-factor, like SecA or the ribosome. In a second step, the signal sequence intercalates between SecY helices 2b and 7 in the lateral gate. This intercalation triggers a further opening of the lateral gate and a stable docking of the signal sequence at the SecY-lipid interface (Plath et al., 1998). This opening of the channel will trigger movement of the 'plug' domain, thus the channel is open and ready for the translocation of the hydrophilic polypeptide following the signal sequence. This is the third step in protein translocation. This model for the passage of the nascent secretory protein is supported by various biochemical and structural studies (Derman et al., 1993; Cannon et al., 2005; Smith et al., 2005). When the channel is opened,

the signal sequence remains engaged with the Sec complex until at a thus far unknown moment it is cleaved by the signal peptide peptidase (for review see (Paetzel, 2014)). When the transport of the secretory protein is completed, the 'plug' helix is thought to move back and seal the channel again in its idle state (Park and Rapoport, 2011).

## 1.1.6 Sealing of the channel

Multiple studies investigated how the Sec complex maintains the membrane barrier for small molecules and ions in its idle and active states. It could be shown that in the idle state the sealing is maintained by the 'plug'-domain and the pore ring as it has been already imposed from the first crystal structures. Both, deletion of the 'plug'-domain and mutation of residues in the pore ring lead to leakiness of SecY (Park and Rapoport, 2012b). Since the 'plug' interacts with the pore ring in the inactive state it can explain why both are required for tight sealing and also why intercalation of a signal sequence in the lateral gate can trigger 'plug' movement.

In active translocation the 'plug' is removed, but the pore ring is still in place forming a ring around the secretory substrate, thus maintaining the membrane as a barrier for small molecules. Consequently, whenever a substrate leaves SecY, either into the periplasm or laterally into the membrane, the 'plug' has to move back and re-seal the channel (Park and Rapoport, 2011).

## 1.1.7 Membrane protein integration by SecY

Binding of SecY to the ribosome

Already the first cryo-electron microscopy studies of a complex formed by a ribosome and the protein-conducting channel indicated that the ribosomal exit tunnel aligns well with the central pore of the channel (Beckmann et al., 1997; Beckmann et al., 2001). Cryo-EM studies with higher resolution could further support the idea that the nascent membrane protein is directly transferred from the ribosomal exit tunnel into the channel (Becker et al., 2009; Frauenfeld et al., 2011; Gogala et al., 2014).

However, early structural studies of ribosome-bound Sec channels could not unambiguously identify the oligomeric state of the Sec complex since in low-resolution structures it was hardly possible to distinguish between the protein part and density that attributed to lipids or detergents. Thus, different studies proposed oligomers of SecY to be bound to the

ribosome and to form the functional unit for protein translocation (Beckmann et al., 1997; Beckmann et al., 2001; Mitra et al., 2005).

With increasing resolution it became obvious that only one copy of the Sec complex binds to the ribosome and it could be additionally shown biochemically that one Sec complex is sufficient for both reactions, the translocation of proteins across and integration of proteins into the membrane (Frauenfeld et al., 2011; Kedrov et al., 2011; Park and Rapoport, 2012a; Taufik et al., 2013; Gogala et al., 2014; Voorhees et al., 2014). However, *in vivo* cross-linking experiments indicated the oligomerization of Sec complexes (Meyer et al., 1999; Bessonneau et al., 2002; Duong, 2003); the physiological role of those however is not yet understood.

Models based on high-resolution cryo-EM densities helped to reveal the binding sites of the Sec complex to ribosome. SecY binds to the ribosome mainly via the cytosolic loops 8/9 and 6/7. These loops contact the ribosome at the rRNA helices 50, 53 and 59, as well as the ribosomal protein uL23 at the exit tunnel region (Becker et al., 2009; Frauenfeld et al., 2011; Park et al., 2014). Additionally, it was proposed that the C-terminus of SecY bind uL24, as well as the rRNA helices 24 and 50. The amphipathic helix of SecE exhibits contacts to the ribosomal proteins uL23 and uL29 (Frauenfeld et al., 2011). A recent cryo-EM structure at nearly atomic resolution described the contacts of mammalian Sec61 to the mammalian ribosome for the first time with amino acid precision (Voorhees et al., 2014). Such a highly resolved structure for the bacterial SecY:ribosome complex is still missing. Collectively, however, it could be shown by the various models that the binding site of the Sec translocons to the ribosome is universally conserved in all kingdoms of life.

Biogenesis of polytopic inner membrane proteins

Inner membrane proteins in *E.coli* are usually translocated in a co-translational fashion via the Sec pathway. As soon as the nascent membrane protein emerges from the ribosomal exit tunnel, its N-terminal signal sequence or a signal anchor gets recognized and bound by SRP and targeted to the membrane and the Sec translocon, respectively, as described above.

As the SecY-bound ribosome translates, the polypeptide chain moving through the translocon gets continuously scanned by the channel. The path followed by these nascent transmembrane segments from the ribosomal exit tunnel to the inside of the SecY channel has been extensively studied by cross-linking, cryo-EM and by indirect measurements of 'pulling forces' during the biogenesis of membrane proteins (Cannon et al., 2005; Frauenfeld

et al., 2011; Ismail et al., 2012; Cymer and von Heijne, 2013). The current model of co-translational membrane integration involves a series of translocation and membrane integration steps. The first transmembrane segment would function as a signal anchor to target the translating ribosome to the membrane. Upon opening of the lateral gate, the transmembrane segment will partition into the membrane (Skach, 2009). An interaction of the hydrophobic pore ring of SecY with a nascent transmembrane segment might be an important trigger for the opening of the lateral gate (Demirci et al., 2013). The next transmembrane segment would then act as a stop-transfer sequence to arrange the proper translocation of a loop to the periplasm or to allow cytoplasmic loops to emerge from the ribosome:translocon junction. The third transmembrane segment would re-initiate membrane integration into the lipid bilayer and so forth. Hence, the channel goes through alternating stages of gating and translocation to direct loops to either periplasm or cytoplasm and to establish the correct topology of the protein (Skach, 2009).
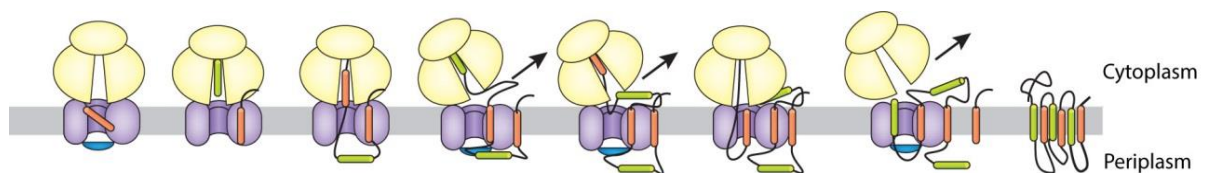


**Figure 5: Schematic of the insertion of a polytopic membrane protein**
Sequential integration of signal anchors (orange) and stop-transfer sequences (green) ensure a topologically correct inserted membrane protein.
Figure modified from (Skach, 2009).

In order to shield the unfavorable polar groups at the protein backbone, the transmembrane segment needs to form an α-helix latest when actually partitioning into the membrane (Sadlish et al., 2005). The distinct time-point for the helix formation is still under debate, as it was shown that helix formation can be induced inside the channel, but also already inside the ribosomal exit tunnel (Lu and Deutsch, 2005; Mackinnon et al., 2014).
The exact time-point for the formation of helices is of particular importance in the biogenesis of complex polytopic membrane proteins. Here, the formation of helix-helix interactions is crucial for the correct folding and assembly of the protein in the membrane. The formations of these interactions is especially important for helices, whose hydrophobicity alone is not high enough for partition into the membrane (Elofsson and von Heijne, 2007). These helices might already form interactions with other transmembrane segments of the same protein, thus integration of a helix bundle is energetically favorable. Additionally, positively and negatively charged amino acids have a high energy barrier when integrating into the membrane (Hessa et al., 2005; Xie et al., 2007; Ojemalm et al., 2013);

however, reciprocal saturation of the charges favors the integration of the helices. Whether these interactions are established in the translocon or at the cytoplasm:lipid interface is still subject of investigation. The current model suggests that the co-translational formation of helix-helix interactions and the sequential release of individual helices or helix bundles into the membrane occur via the lateral gate of SecY.
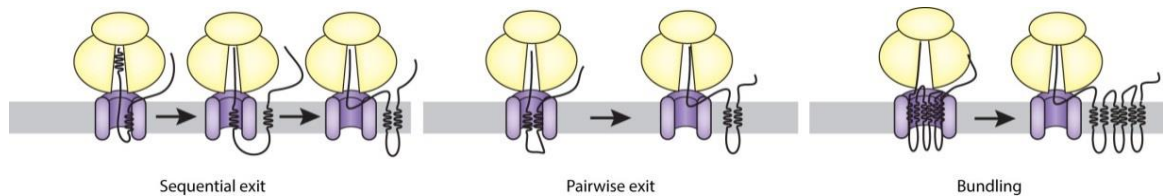


**Figure 6: Different mechanisms of transmembrane helix integration**
Helices could be integrated one-by-one in a sequential manor. Another model suggests the pair-wise exit of helices, while a third model postulates the exit of groups of transmembrane helices into the membrane.
Figure modified from (Skach, 2009).

It has been shown that the nascent transmembrane segments still stay in contact with the Sec translocon even after emerging into the lipid bilayer and that they assist the folding of downstream C-terminal helices. That suggests that the environment outside of the translocon itself might act as an inner membrane chaperone assisting in the folding of polytopic membrane proteins (Sadlish et al., 2005). Additionally, the role of the membrane insertase YidC in the biogenesis of polytopic membrane proteins is still not completely clear.

## 1.1.8 Membrane protein topology and the positive inside rule

The inner membrane proteins in bacteria usually consist of α-helical bundles that are integrated by the Sec translocon. However, one decisive point in the biogenesis of inner membrane proteins that is still not understood is the molecular mechanism by which the topology of the membrane protein is determined. Proteome-wide analysis revealed that positively charged amino acids are enriched in cytosolic loops, flanking the transmembrane segments, which is known as the 'positive-inside rule' (von Heijne, 1986; Boyd et al., 1987; Boyd and Beckwith, 1990; von Heijne, 1992; Hessa et al., 2005; Hessa et al., 2007; Hessa et al., 2009). Interestingly, this positive-inside rule also applies for membrane proteins in the ER, mitochondria or chloroplasts (Gavel et al., 1991; Gavel and von Heijne, 1992). Additionally, negatively charged amino acids in the periplasm can also influence the topology of a membrane protein (Kiefer et al., 1997; Rutz et al., 1999).
In polytopic membrane proteins, the orientation of each transmembrane segment has to be controlled carefully to ensure the correct tertiary structure of the protein. However, it was

shown that a single residue in one cytoplasmic loop can alter the topology of a whole protein, even when that residue is located at the very C-terminal end of the protein (Seppala et al., 2010). That would presume that the whole nascent membrane protein is able to flip its topology after almost complete synthesis, which was also observed for even larger membrane protein (Vitrac et al., 2013). The role of SecY or SecY-associated proteins in this 'flipping' is not understood thus far. To date the underlying molecular mechanism for the 'positive-inside rule' is still obscure. In very early studies it has been suggested that the membrane potential align the membrane proteins according to the 'positive-inside rule' (Andersson and von Heijne, 1994). However, that view has been challenged since acidophilic bacteria have an inverted membrane potential across the inner membrane, while still following the 'positive-inside rule' (van de Vossenberg et al., 1998). Additionally, it was shown that mutations within the Sec translocon also affect the correct topology of membrane proteins (Goder et al., 2004). Moreover, the lipid composition of the membrane was shown to play a role in topology determination and can additionally cause 'flipping' of a membrane protein (Prinz et al., 1998; Junne et al., 2007; Dowhan and Bogdanov, 2009).

## 1.1.9 The bacterial membrane insertase YidC

YidC is a bacterial membrane insertase (Samuelson et al., 2000) that catalyzes the insertion of inner membrane proteins, either alone or in concert with SecYEG. As well as SecY, YidC can bind to the ribosome close to the ribosomal exit tunnel (for review see (Dalbey et al., 2014)). *E.coli* YidC is composed of 6 transmembrane segments and a large periplasmic domain (P1) following transmembrane segment 1 (Saaf et al., 1998). Biochemical analysis revealed that neither transmembrane segment 1 nor the periplasmic domain are essential for YidC functioning and only the conserved transmembrane segments 2-6 are essential for membrane protein integration (Jiang et al., 2003). This conserved core has homologues in mitochondria (Oxa1) and in chloroplasts (Alb3) and also in some bacteria. Additionally, the oligomeric state of YidC has been intensively debated. A first model based on a low-resolution cryo-EM structure postulated a dimer of YidC as the functional unit bound to the ribosome (Kohler et al., 2009); however biophysical analysis suggested that only one copy of YidC binds to the ribosome (Kedrov et al., 2013; Wickles et al., 2014).

Insertion of membrane proteins via the YidC-only pathway
Interestingly, some YidC substrates e.g. MscL are targeted to YidC via the conserved SRP/SR pathway (Facey et al., 2007). However, other substrates e.g. $F_o$c do not require any particular targeting to the membrane (Yi et al., 2004).

To date there have been a few YidC-only substrates identified (Dalbey et al., 2014). Those have in common that they possess only very small hydrophilic domains that need to be translocated (Serek et al., 2004; van der Laan et al., 2004). The molecular mechanism how YidC integrates membrane proteins was subject of extensive biochemical and biophysical investigations; however, the full mechanism could not be revealed by these experiments (Saller et al., 2012; Dalbey et al., 2014). The breakthrough was the recently solved, first crystal structure of *Bacillus halodurans* YidC2 (see Figure 7 right) (Kumazaki et al., 2014a; Kumazaki et al., 2014c), as well as an accurate model of ribosome-bound YidC form *E.coli* engaged with a substrate based on cryo-EM, sequence-based co-variation analysis an molecular dynamics simulations (Wickles et al., 2014) (see Figure 7 left). A recent crystal structure of *E.coli* YidC did not reveal substantial differences to the earlier published structures (Kumazaki et al., 2014b). These two models gave first insights on the molecular mechanism of YidC-mediated membrane protein insertion. Remarkably, both structures reveal a monomer of YidC as the functional unit. Furthermore, both structures revealed that the conserved core of YidC forms a positively charged groove, which is opened towards the lipid environment and the cytoplasm and closed to the periplasm. In the YidC:ribosome complex this groove is well aligned with the ribosomal exit tunnel (Wickles et al., 2014).
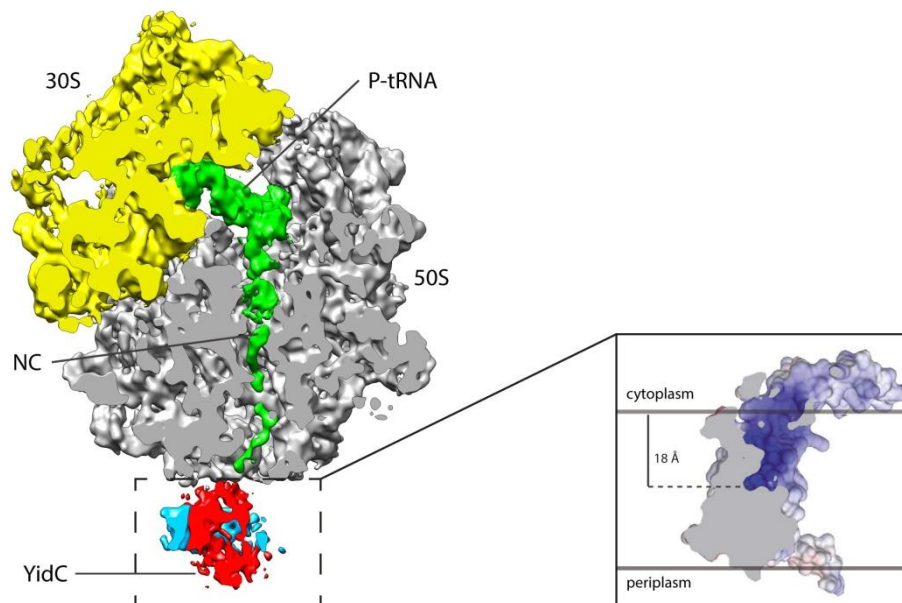


**Figure 7: Cryo-EM structure of *E.coli* YidC bound to a translating ribosome and crystal-structure of *Bacillus halodurans* YidC2**

Left: A monomeric YidC is (red) in a detergent micelle (blue) is bound to a translating ribosome (30S subunit in yellow, 50S subunit in grey, the P-site tRNA and the nascent peptide in green). The structure reveals alignment of the hydrophilic core of YidC with the ribosomal exit tunnel. Membrane integration occurs at the interface of the protein and lipids.

Right: space filled model of *B.halodurans* YidC 2 shows the hydrophilic groove of YidC (shown in blue) at the protein-lipid interface.

Figure modified from (Wickles et al., 2014) left and (Kumazaki et al., 2014a) right.

Moreover, the molecular dynamic simulations revealed that the proximate lipid membrane is significantly thinned by YidC (Wickles et al., 2014). Based on these models it was suggested that the nascent membrane protein inserts on the interface of the hydrophilic groove and the lipid environment, while the hydrophilic groove reduces the required free energy to translocate flanking short hydrophilic loops over the membrane. Furthermore, the free energy could be further decrease by tinning of the membrane that in induced by YidC. Taken together, the mechanism of membrane insertion by YidC might be explained by YidC providing a 'slide' that reduces the free energy of translocating small hydrophilic loops over the membrane and favors the integration of transmembrane segments.

Membrane protein insertion by SecY in cooperation with YidC
Co-purification experiments suggested that SecYEG and YidC form a functional complex in the membrane together with other accessory factors like SecDFyajC (Scotti et al., 2000; Nouwen and Driessen, 2002). However, the presence of SecDFyajC is not essential for the function of a SecYEG:YidC complex (Xie et al., 2006). It was shown by cross-linking experiments that YidC occupies the area in front of the lateral gate of SecY and that polytopic substrates can get cross-linked to YidC in an early stage of their biogenesis. However, YidC gets displaced from the lateral gate of SecY when the synthesis of the polytopic membrane protein continues (Sachelaru et al., 2013). That suggests that the integration of a SecY-YidC substrate e.g. CyoA (Facey et al., 2007) in *E.coli* occurs at the SecY:YidC interface and that SecY and YidC together might form a space in the membrane with special biophysical properties to favor the correct folding and assembly of membrane proteins. Thus far, structural data on a functional YidC:SecY assembly that could shed further light on the underlying mechanism of membrane protein insertion are missing.

## 1.1.10 A thermodynamic view on membrane protein biogenesis

In the last years, the process of membrane protein insertion was also addressed from a totally different point of view, by examining the thermodynamics that underlie and drive the different stages of membrane protein insertion.

The researchers describe four different steps in this process: (1) the partitioning, (2) the folding, (3) the insertion and (4) the association (White and Wimley, 1999; Cymer et al., 2014).

1: An unfolded transmembrane segment in the cytoplasm is energetically highly unfavorable, since exposed hydrophobic residues decrease the entropy in the solvent dramatically. This phenomenon is generally known as the hydrophobic effect. Instead, the unfolded transmembrane segment either aggregates in solution or associates extremely fast with the membrane interface, where both the hydrophobic side-chains and the polar backbone can form contacts with the hydrophobic lipid tails and the polar lipid head groups, respectively. This process is thermodynamically so highly favored, that it could hardly be described in molecular simulations (Wimley and White, 1996; White and Wimley, 1999; Popot and Engelman, 2000; Mackenzie, 2006; Ulmschneider et al., 2014).

2: The folding of the peptide to an α-helix at the interface of the membrane is per se not favored, however, when the peptide bonds of the backbone can engage in H-bonds, as found in transmembrane helices, the process becomes highly favored (White and Wimley, 1998; Wimley et al., 1998; Ladokhin and White, 1999).

3: For insertion of the interface-bound α-helical membrane segment two components of the free energy have to be considered. On one hand, the thermodynamic cost of the dehydration of the peptide bonds and the forming H-bonds in the hydrophobic environment, and on the other hand, the gain of the free energy when integrating hydrophobic side-chains into the hydrophobic lipid environment in the membrane. Only if the latter free energy compensates for the loss of the peptide backbone dehydration, the transmembrane segment gets inserted into the membrane. That explains very clearly why transmembrane segments have to be fairly hydrophobic, at least considering single isolated transmembrane segments (Ben-Shaul et al., 1996; Ben-Tal et al., 1996; Almeida et al., 2012).

4: The model presented above highlights the basic principle of a thermodynamic view on the insertion of a single transmembrane helix; however, membrane proteins consist very often of several transmembrane segments. Here, the thermodynamics of helix-helix interactions have to be additionally considered. These tertiary interactions can favor the insertion of two transmembrane segments that alone would not be that stable in the membrane. For

example, by reciprocal saturation of charged residues or formation of H-bonds the energy landscape can be changed dramatically, thus favoring the insertion of these two transmembrane segments. In a recent study it has been shown that helices of polytopic membrane proteins already form interactions in very early stages of their synthesis and integration into the membrane (Cymer and von Heijne, 2013). These considerations can be made arbitrary complex for membrane proteins containing up to a dozen transmembrane segments (Zhou et al., 2001; Doura and Fleming, 2004; Hong and Bowie, 2011; Cymer and von Heijne, 2013).
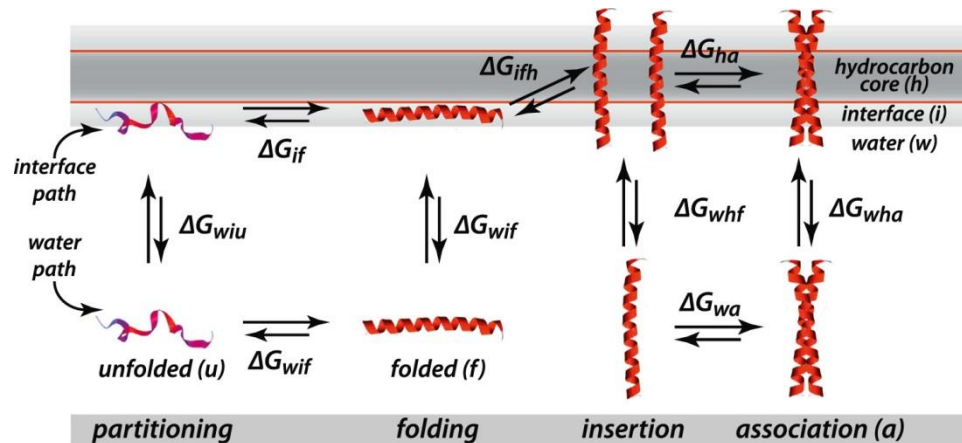


**Figure 8: Model of thermodynamic integration of a transmembrane helix**
Schematic overview of the four stages, partitioning, folding, insertion and association in a thermodynamic model for the integration of a single transmembrane helix and the considered free-energies for each step. Figure modified from (Cymer et al., 2014).

All these thermodynamic considerations made above describe translocon-free insertion of a transmembrane helix. However, it could be shown that the general free energies of transmembrane segment insertion are parallel shifted to a more favorable free energy when the translocon is included in the calculations. That indicated that the translocon only catalyzes the insertion, but does not alter the thermodynamics of insertion completely.

That leads to a model in which the transmembrane helices do not fully insert into the translocon to then egress laterally through the gate into the membrane. They rather use the area of the lateral gate as a 'slide' to lower the free energy of the insertion step described above. The hydrophilic interior of the channel would then also assist the translocation of small periplasmic loops. As introduces above, larger periplasmic loops or domains even need the assistance of the motor protein SecA that might indicated a change in the translocation mode (Cymer et al., 2014).

Interestingly, the model of the translocon acting as a slide to lower the energetic costs of membrane insertion is very similar to the proposed function of the membrane protein insertase YidC, as described above (see chapter 1.1.9).

## 1.2 Translation regulation by arrest peptides

### 1.2.1 General mechanisms of translation regulation by arrest peptides

When the ribosome synthesizes a protein, the nascent polypeptide emerges through the ribosomal exit tunnel and leaves the ribosome after complete synthesis and translation termination. The ribosomal exit tunnel is about 100 Å long and can accommodate 30-40 residues of a nascent peptide (Nissen et al., 2000; Voss et al., 2006; Wilson and Beckmann, 2011).

The average diameter of the tunnel is around 15 Å and it was shown that the formation of α-helical secondary structure is already possible inside the tunnel (Bhushan et al., 2010a). The wall of the tunnel is mainly composed of ribosomal RNA. However, at about one third of the distance from the Peptidyl-transferase center (PTC) to the exit site two ribosomal proteins, uL22 and uL4 (in bacteria) have extended loops reaching into the tunnel and forming a central constriction. In the beginning of ribosomal research the tunnel was thought to be of 'Teflon'-like character, not interacting at all with the emerging substrate (Nissen et al., 2000). This view, however, has been challenged by various studies showing extensive interactions of components of the tunnel wall with nascent substrates or small molecules (Seidelt et al., 2009; Bhushan et al., 2010b; Bhushan et al., 2011; Wilson and Beckmann, 2011; Arenz et al., 2014a; Arenz et al., 2014b).

A certain class of proteins called ribosome-arrest peptides contain specific amino acid sequences that interact with distinct features of the ribosomal exit tunnel and induce an arrest of their own translation. These arrest peptides represent a *cis*-specific regulation of the ribosome, as the information to stop translation is encoded in the peptide itself and does not come from an external factor acting *in trans* as, for example, translational stop by SRP. Most of the known arrest peptides are up to 20 amino acids in length; however, they often contain only a few conserved residues crucial for the translation arrest. There are two classes of arrest peptides: one class that can arrest translation alone, like the SecM or the MifM arrest peptides, and a second class that arrests translation only in response to small molecules, like an amino acid (TnaC, AAP) or an antibiotic (ErmA,B,C,D). For review see (Ito and Chiba, 2013).
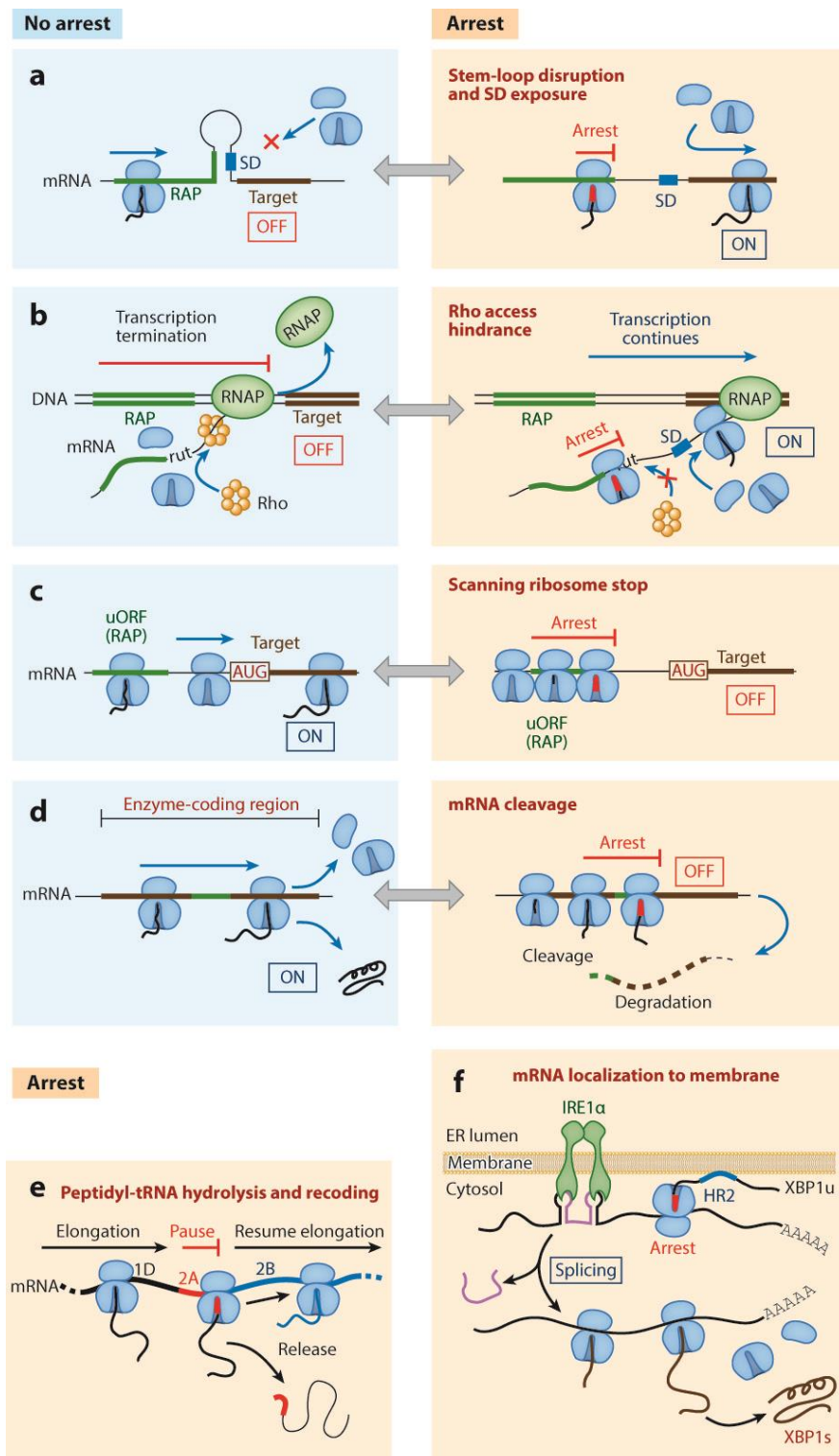
**Figure 9: Overview of several known mechanisms of translational control by regulatory arrest peptides**
a: effect of ribosome stalled on an arrest peptide on the secondary structure of the mRNA.
b: ribosome stalling affects binding of transcription factors.
c: stalling on arrest peptides inhibits scanning and translation of down-stream ORFs.
d: ribosomal stalling triggers mRNA cleveage.
e: small arrest peptides trigger unusual termination and re-initiation.
f: ribosomal stalling can affect the localization of a mRNA.
Modified from (Ito and Chiba, 2013).

The actual stalling of the arrest peptide in the ribosome relies on different molecular mechanisms. A common mechanism is the inhibition of peptidyl transfer, so that the amino acid on the A-site tRNA cannot be transferred to the nascent peptide (e.g. SecM, MifM). Secondly, the termination of the arrest peptide is inhibited (e.g. TnaC, AAP) and, thirdly, the ribosome gets translocated (e.g. CGS1) (Ito and Chiba, 2013).

However, all arrest peptides have in common, that the translational arrest alters the expression of downstream genes. This can be achieved by: changing the folding state of the mRNA (ErmA,B,C,D, MifM) (Arenz et al., 2014a; Arenz et al., 2014b); by hindering the access of the transcription termination factor Rho (TnaC) (Seidelt et al., 2009; Bischoff et al., 2014a); by induction of mRNA cleavage (AdoMet); by peptidyl hydrolysis and recoding (2A) or by localization of the mRNA to the membrane (XPB1) (Ito and Chiba, 2013).

## 1.2.2 Ribosomal stalling on the TnaC peptide in *E.coli*

In *E.coli* the homeostasis of tryptophan is tightly controlled by the TnaC arrest peptide. The tryptophan metabolism operon *tna* consist of three genes. *TnaC*, the controlling arrest peptide, followed by *tnaA* encoding for the tryptophan degrading tryptophanase and *tnaB* encoding for a tryptophan specific permease. The area between *tnaC* and *tnaA* comprises several Rho dependent transcription termination sites.
Intriguingly, the transcription and thus the expression of TnaA and TnaB is regulated by TnaC in response to free tryptophan.
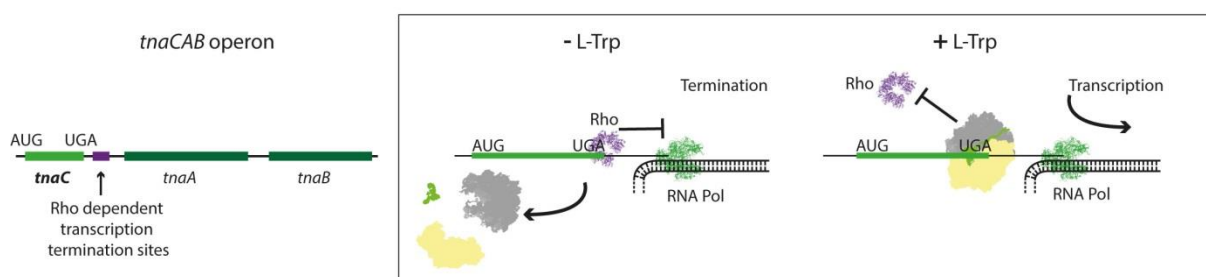


**Figure 10: Schematic for the tryptophan dependent stalling on the TnaC peptide**
Left: general organization of the *tnaCAB* operon.
Right: schematic of the transcription termination by Rho in the absence of L-Trp and inhibition of transcription termination in the presence of Rho.
Figure modified from (Bischoff et al., 2014a).

At low intracellular tryptophan concentration, the ribosome translates the 24 amino acid comprising TnaC leader peptide. Translation gets terminated at the UGA stop codon of TnaC

and the ribosome gets released from the mRNA. Thus, transcription termination Rho is able to bind the mRNA and to remove the RNA polymerase from the mRNA before the transcription of *tnaA* and *tnaB* is complete (Gong and Yanofsky, 2002b; Gong and Yanofsky, 2002a).

At high intracellular levels of tryptophan, however, the free L-Trp induces stalling of the ribosome on the UGA stop codon of the TnaC leader peptide by inhibiting the release of the peptide by release factor 2 (Gong et al., 2001; Gong and Yanofsky, 2001). Consequently, the binding sites for transcription termination factor are occupied by the ribosome and the transcription of *tnaA* and *tnaB* can be completed. That leads finally to synthesis of TnaA and TnaB which then degrade the excess of free L-Trp in the cell.

To sum it up, the ribosome gets stalled on the TnaC leader peptide at the P24 residue of TnaC in the P-site and the UGA stop codon in the A-site in the presence of elevated levels of L-Trp. The exact concentration of L-Trp that induces stalling *in vivo* is not yet known (Ito and Chiba, 2013).

Extensive mutational studies and analyses of sequence conservation could identify P24, D16 and W12 of TnaC to be the invariant residues for effective stalling. Additionally, the distance between these three residues is also of particular importance (Gong and Yanofsky, 2002b; Gong and Yanofsky, 2002a; Cruz-Vera et al., 2005; Cruz-Vera and Yanofsky, 2008). Furthermore, several rRNA bases in the ribosome, as well as amino acids in the proteins of the central constriction uL22 and uL4 have been identified as important components for induction of stalling (Cruz-Vera et al., 2005; Cruz-Vera et al., 2007; Yang et al., 2009; Martinez et al., 2014). Interestingly, it could be shown that TnaC inhibits specifically the action of RF2, replacement of the UGA stop codon to UAA or UAG that are recognized by RF1 alleviate stalling (Konan and Yanofsky, 1999). Crucial results of the mutational studies are summarized in Table 1.

**Table 1: Summary of important mutational studies involving the TnaC leader peptide and components of the ribosomal tunnel wall**

| Molecule | Residue | Experimental Data | Reference |
|---|---|---|---|
| TnaC | M1-W12 | Mutations have little or no effect on TnaC stalling | (Gong and Yanofsky, 2002b) |
| | K11 | UV crosslinks with A750 | (Cruz-Vera et al., 2005) |
| | W12 | W12R, W12L, substitution with 33 unique codons alleviate stalling | (Gollnick and Yanofsky, 1990; Gish and Yanofsky, 1995; Kamath and Yanofsky, 1997; Konan and Yanofsky, 1997; Gong and Yanofsky, 2002b; Yang et al., 2009) |
| | F13 | F13C, F13L, F13I reduce stalling | (Gollnick and Yanofsky, 1990; Gish and Yanofsky, 1995; Gong and Yanofsky, 2002b) |
| | N14 | N14I, N14S, N14K eliminate stalling | (Gish and Yanofsky, 1995) |
| | I15 | I15F and I15N alleviate stalling | (Gish and Yanofsky, 1995; Cruz-Vera and Yanofsky, 2008) |
| | D16 | D16A, D16N, D16E, D16W, D16S, D16L, D16C, D16V, D16K Eliminate stalling | (Gish and Yanofsky, 1995; Cruz-Vera and Yanofsky, 2008; Cruz-Vera et al., 2009) |
| | K18 | K18R has no effect on stalling, K18Q reduces stalling | (Gish and Yanofsky, 1995; Cruz-Vera et al., 2005; Cruz-Vera and Yanofsky, 2008) |
| | I19 | I19M, I19W reduce stalling | (Martinez et al., 2014) |
| | P24 | P24S and P24A eliminate stalling | (Gong and Yanofsky, 2002b; Cruz-Vera et al., 2009) |
| 23S rRNA | A751 | Insertions at A751 eliminate stalling | (Cruz-Vera et al., 2005; Cruz-Vera et al., 2007) |
| | A752 | A752C and A752T reduce stalling | (Cruz-Vera et al., 2007) |
| | A2058 | A2058U reduces stalling | (Martinez et al., 2014) |
| | A2059 | A2058U/G reduces stalling | (Martinez et al., 2014) |
| | U2585 | Base was supposed to inhibit RF2 accommodation | (Seidelt et al., 2009) |
| | A2602 | Base was supposed to inhibit RF2 accommodation | (Seidelt et al., 2009) |
| | U2609 | U2609A and U2609C eliminate stalling | (Cruz-Vera et al., 2005; Cruz-Vera et al., 2007) |
| uL22 | K90 | K90W, K90H, K90A, K90W eliminate stalling | Cruz-Vera et al., 2007; Cruz-Vera et al., 2005) |
| uL4 | K57-K58 | K57E-K58Q, K57Q-K58Q, K57E-K58E, K57Q-K58E slightly reduce stalling | Cruz-Vera et al., 2005) |

Table inspired by (Trabuco et al., 2010)

Since mutational studies also revealed the importance of several residues in the PTC for effective TnaC stalling and that TnaC competes with the antibiotic sparcomycin, that binds in the PTC behind A2602, it was suggested that the free L-Trp molecule(s) bind in that area in the PTC close to the A-site (Cruz-Vera et al., 2006; Cruz-Vera et al., 2007; Cruz-Vera and Yanofsky, 2008; Yang et al., 2009). This view was supported by observations that TnaC is partially resistant against the antibiotic puromycin that also acts directly in the PTC (Hansen et al., 2003; Cruz-Vera et al., 2006).
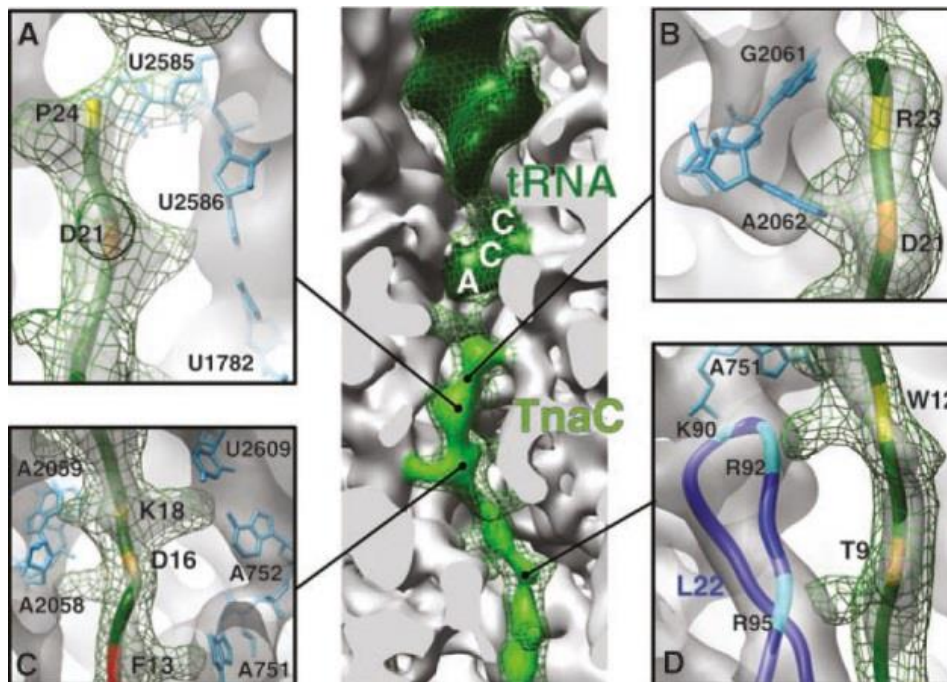


**Figure 11: Contacts on the TnaC peptide to the ribosomal exit tunnel observed by (Seidelt et al., 2009)**
Seidelt et al., 2009 identified strong contacts of D21, K18 of TnaC and K90 of uL22 between the nascent peptide and the ribosome.
Figure modified from (Seidelt et al., 2009).

The molecular model of TnaC based on a cryo-EM structure at 5.8 Å could reveal that TnaC stalls indeed with the P24 tRNA in the P-site and several contacts of the nascent chain to the ribosomal exit tunnel could be identified. Additionally, two rRNA bases in the PTC, U2585 and A2602 were found in unusual, stable conformations. These conformations preclude the accommodation of RF2 and thus explain the inability to terminate. However, the limited resolution of the cryo-EM based model did not allow for the identification of the position the free L-Trp molecule(s) (Seidelt et al., 2009). Collectively, the so far performed mutational and structural studies gave a comprehensive overview of decisive residues in TnaC and the ribosomal exit tunnel and how RF2 termination is inhibited; however, the underlying molecular mechanism of how free L-Trp induces this stalling and release inhibition is not known so far.

## 1.3 Cryo-EM and the resolution revolution

For the determination of the three dimensional structure of macromolecular complexes like the ribosome, the proteasome or viruses, cryo-EM became the method of choice. While other techniques to determine molecular structures, like X-ray crystallography or Nuclear magnetic resonance (NMR), promise higher resolutions in the final structures, they have significant drawbacks. X-ray crystallography requires diffracting crystals, and relatively large sample amounts are necessary for NMR studies. Cryo-EM, however, can deal with sample concentrations in the low nano-molar or even pico-molar range in volumes less than 10 µL. Furthermore, cryo-EM does not require the sample to be ordered in any form.

Already in the first cryo-EM studies, the ribosome has been identified as the prominent sample, because of its relatively high molecular mass of 2-3 MDa and its high RNA content that gives high contrast in the micrographs (Frank et al., 1991).

In cryo-EM the sample is flash frozen by quickly plunging the EM grid in liquid ethane. This extreme temperature gradient of 100,000 °C/s prevents the solvent from forming ice-crystals, resulting in a specimen embedded in vitreous ice. Vitreous ice does not interact with the electron beam of the microscope, like crystalline ice would. Thus it is possible to detect a signal that is the result from inelastic scattering events of the electron beam with the embedded specimen.

However, due to the low electron density of biological macromolecules (main components: Carbon, Oxygen, Hydrogen, Sulfur, Nitrogen, Phosphor) the signal-to-noise ratio is rather low. Additionally, the biological specimen suffers from radiation damage that limits the electron dose to about 20 e$^-$/Å$^2$. Unfortunately, low electron-dose additionally lowers the signal-to-noise ratio. So, the microscope needs to be operated with compromising between a high signal-to-noise-ratio while avoiding radiation damage. Furthermore, the incoming electron beam induces charging of the sample and results in beam-dependent motion. That induced motion of the time of the exposure results in a blur of the image (Leapman and Sun, 1995).

In the ideal case, the individual molecules of the sample get randomly oriented on the EM grid so that all possible orientations are represented. The EM grid is then mounted into the transmission electron microscope and images at a magnification of 75,000 x to 150,000 x are acquired. Since the microscope is operated in the transmission mode, the resulting images of the individual molecules represent 2D projections of the 3D molecule. These 2D projections are then windowed out of the EM-micrograph by computational means and treated as individual particles. A cryo-EM dataset can contain up to or even more than 1,500,000 particles.

To reconstruct the three dimensional structure, the relative orientation of all 2D projections needs to be precisely determined and subsequently the 2D projections can computationally be back-projected to a three dimensional volume. By classification of the particles and iterating orientation determination and back-projection the resolution of the final 3D volume can be increased.

The final resolution depends on several factors: the homogeneity of the dataset, the number of individual particles and intrinsic flexibility in the molecule itself and, most importantly, on the electron detection device that is used.

Since the precise orientation determination requires a high signal-to-noise ratio, the detection device that delivers the highest signal-to-noise ratios over the whole frequency spectrum will result in the best-resolved volume (Henderson, 1992; Henderson, 1995).

The first electron micrographs were recorded on film, which has reasonable good detective quantum efficiency (DQE: frequency dependent measure of the signal-to-noise performance of the detection device). However, film is difficult to handle and is not suitable for automation and high-throughput. Nevertheless, structures of the ribosome recorded on film have been published to resolution up to 5.8 Å (Seidelt et al., 2009).

The film was then replaced with 'charge-coupled device' (CCDs) cameras. Here, the incoming electron hit a scintillator plate and is converted to a light signal that is recorded by the CCD camera. This made it possible to collect huge datasets automatically, however, with a poor DQE, especially at higher acceleration voltage at which most modern electron microscopes operate best. Ribosomal structures have been reported up to 5.5 Å resolution with datasets collected on CCD cameras (Anger et al., 2013).

## 'The resolution revolution' (Kuhlbrandt, 2014)

In the beginning of the 2010s the first commercially produced direct electron detectors or 'direct-detection devices' (DDDs) entered the market and revolutionized the field of cryo-EM (Bai et al., 2013). These DDDs are equipped with chips that directly recognize the incoming electrons, resulting in an unprecedented signal-to-noise ratio. The extremely thin chips that prevent electrons from backscattering additionally enhance the measured signal over the background noise. Furthermore, their DQE at higher frequencies is superior to both CCD and film. However, the perhaps most important advantage is the much faster readout of the DDDs. Modern DDDs routinely read out up to 35 frames per second without a loss in the signal. This mode of data collection is referred to as the 'movie-mode'. This movie-mode now enables to correct for beam-induced drift of the particles on both 2D and 3D level. Moreover, the impact of radiation damage can be controlled, since frames that suffered

from an excessive electron dose can be excluded from the final back-projections (McMullan et al., 2009a; McMullan et al., 2009b; McMullan et al., 2009c).

This remarkable improvement in the detection devices resulted in a flow of cryo-EM structure at atomic resolution. Ribosomal structures up to 3 Å are now reported routinely (Amunts et al., 2014; Brown et al., 2014; Fernandez et al., 2014; Voorhees et al., 2014; Wong et al., 2014), but also small protein complexes as ion-channels (Liao et al., 2013) or the γ-secretase (Lu et al., 2014) have been solved to almost atomic resolution.

In this thesis we present two cryo-EM structures that also illustrate that dramatic advancement in the field. The structure of the tightly coupled PR2Q-SecYE complex was solved from EM data recorded on CCD to a resolution up to 7.3 Å which was, at that time, close to the possible limit for a programmed ribosome engaged with a ligand. However, in the second part we present the structure of a TnaC-stalled ribosome reconstructed from EM data collected on a DDD to a resolution up to 3.5 Å, which allows us building an atomic model and elucidating the underlying biochemical reaction on an atomic level.

# 2. Aims of this thesis

The insertion of polytopic membrane proteins into the bacterial membrane occurs co-translationally on ribosomes that are bound to a protein-conducting channel, the SecYEG complex. Although the path followed by the nascent membrane protein inside the ribosome and the Sec complex in relatively well-established, it remains unclear when and how the N-terminal transmembrane domains insert into the membrane, while the C-terminus of the protein is not yet fully synthesized. As a first major aim of this thesis we set out to solve the three dimensional structure of a membrane protein insertion intermediate by cryo-EM and single-particle reconstruction. Since the *in vitro* reconstitution of a ribosome nascent chain complex carrying a nascent polytopic membrane protein with the separately purified SecYEG does most likely not lead to the formation of a physiologically relevant complex, we aim for a protocol to assemble these insertion intermediates *in vivo* in the intact cell. To this end TnaC-stalled membrane protein constructs based on Proteorhodopsin have been expressed in *E.coli* and purification of these complexes followed by cryo-EM and single particle reconstruction has provided unprecedented insights into the biogenesis of polytopic membrane proteins in bacteria.

In *E.coli* the expression of tryptophan-catabolizing enzymes in controlled by the TnaC leader peptide. Thereby, the 12 most C-terminal amino acids of TnaC in concert with free L-tryptophan molecules engage in specific interaction with the ribosomal exit tunnel and lead to stalling of the peptide by inhibiting peptide release by release factor 2.

In the first part of this thesis we used the TnaC stalling sequence to create uniformly stalled membrane protein intermediates, although the underlying molecular mechanism how the TnaC peptide together with the free L-Trp induces the stalling of the ribosome remained unknown.

Thus, we set out to solve the molecular mechanism of TnaC stalling by cryo-EM and single particle reconstruction. We took advantage of the recent development of direct electron detection devices that allowed the reconstruction of cryo-EM data up to resolutions of 3 Å.

These high resolution reconstructions allowed the unambiguous building of an atomic model that eventually revealed the mechanism of the L-Trp induced stalling of a bacterial ribosome on the TnaC leader peptide.

# 3. Materials and Methods

## 3.1 General

Sterile laboratory material and autoclaved glassware was used. Water for buffer preparation was deionized and buffers were sterile filtered before being used in reaction mixtures. Nuclease-free water was used in reactions containing DNA or single stranded RNA.

## 3.2 Vectors

All TnaC-stalled nascent chain constructs were cloned in the pBAD vector (Invitrogen) between the NcoI and the HindIII restriction site. If necessary, an additional glycine was cloned in after the initiator methionine to create the NcoI restriction site. In this vector, protein expression is under the control of the araBAD promotor. In the presence of L-arabinose, protein expression is induced, while the absence of L-arabinose only results in a very low basal expression. By varying the L-arabinose concentration, protein expression level can be optimized.

SecYEG and FtsY were expressed from the pRSF DUET plasmid, the protein SecA was cloned into the pRSF-ek-lic plasmid. Both plasmids are under the control of a T7 promotor-based system.

## 3.3 PCR Primers

**Table 2: Primers with Sequences used in this study**

| Description | Name | Enzyme | Sequence |
|---|---|---|---|
| Fwd 1 for FtsQ$_{85}$ | es236 | NcoI | CATGCCATGGGTCATCACCATCACCATCACCATCACGATTACGATATC |
| Rev 1 for FtsQ$_{85}$ | es237 | ApaI | TATAGGGCCCAGCGTAATCTGGAACATCGTATGGGTA |
| Fwd 2 for FtsQ$_{85}$ | es238 | ApaI | TGTGGGGCCCAATATCTTACATATAAGTGTGACCTCAAAATGG |
| Rev 2 for FtsQ$_{85}$ | T7 term | none | CTAGTTATTGCTCAGCGGT |
| Fwd for FtsQ$_{119}$ | es245 | NcoI | CATGCCATGGGTCACCACCATCATCATCATCATCATTACCCATAC |
| Rev for FtsQ$_{119}$ | es246 | ApaI | TTAAGGGCCCGATGTTGACATCCTGGGTCATAAAGG |
| Fwd for strep-PR | es146 | PstI | CCTCTGCAGGAGGTTCCGGAGGTGGATCGGGAGGTGGATCGTGGAGCCACCCGCAGTTCGAAAAGAGCTCCGAAAACCTGTATTTTCAGGGGAGCTCTGCGGGTGGTGGTGACCTTGATGC |
| Rev for strep-PR | es131 | PstI | TGCACTGCAGATTTTTCGAACTGCGGGTGGCTCCAACCACCACCAGCAGCAAATGTAGGAAGTGC |
| Rev for PR0 | es211 | SacII | GTCACCACCACCCGCGGAACTCTTTTCGAACTGCGG |

| | | | |
|---|---|---|---|
| Rev for PR1 | es203 | SacII | GATGTTTTCCATTCCGCGGAAACTCTATCTCTTTCAAC |
| Rev for PR2 | es206 | SacII | GGCGAATCACCAGCCGCGGTCCATACCCCTCTCATGTAC |
| Rev for PR3 | es202 | SacII | CTAGTAATTTCTTAACCGCGGACCCAGCAACATTAGTTGC |
| Fwd for PRQ2/4/6 | es289 | SacII | GAAGCCGCGGGTCGCCTGCCGCTCTCAAAGCTGGTG |
| Fwd on CyoA | LB1 | SacII | TCCCCGCGGTGTAATTCTGCGCTGTTAGATCCC |
| Rev for CyoA$_{27-151}$ | LB2 | SacII | GTCACCACCACCCGCGGAACTCTTTTCGAACTGCGG |
| Fwd for proOmpA-(FtsQ)-TnaC | LB32 | NcoI | CGCCATGGGAAAAAAGACAGCTATCGCGATTGCAGTG |
| Rev for proOmpA-TnaC | LB17 | ApaI | GCGGGCCCAGCGTAATCTGGAACATCGTATGGGTAAGCCTGCGGCTGAGTTACAACGTC |
| Fwd put PR3 in FtsQ | LB30 | blunt end ligation | TGGTTACTAACAGTTCCTCTATTAATATGTGAATTCTACTTAATTCTTGCTGCTGCATACCCATACGATGTTCCAGATTACGCTGGG |
| Rev to put PR3 in FtsQ | LB31 | blunt end ligation | GAGCGGCTGGGTCGTGTTGGGCTGGATGGAAATGAGAGGGGTATGGATTGAAACTGGTGATTCGCCAACTGTATTTAGATACATTGAT |
| Fwd put FtsQ loop in OmpA | LB32 | blunt end ligation | TACACACGTAATGACGATATCCGGCAGTCGATCCTGGCATTGGGTGAGCCGGGTTACCCATACGATGTTCCAGATTACGCTGGG |
| Rev put FtsQ loop in OmpA | LB33 | blunt end ligation | ATGGCGTTCACCGGTCAACACCAGCTTTGAGAGCGGCAGGCGTTGCGCATCTTCGTCACAGGTGTTGCCAGTAACCGGGTTGGATTC |
| Fwd removal of TnaC stalling | LB39 | ApaI | TATAGGGCCCTGAAAGCTTACGGTCTCCAGCTTGGCTGTTTTG |
| Fwd cloning of SecA in pET lic | LB40 | Lic cloning | GACGACGACAAGATGCTAATCAAATTGTTAACTAAAGTTTTCG |
| Rev cloning of SecA in pET lic | LB41 | Lic cloning | GAGGAGAAGCCCGGTTATTGCAGGCGGCCATGGCACTGCTTGTATTTTTTACC |
| Rev elongating FtsQ to K$_{120}$ | LB49 | ApaI | TTGGGCCCAGCGTAATCTGGAACATCGTATGGGTACAATTCATCAGGCCACTGCTTTCTGACGCT |
| Rev elongating FtsQ to I$_{130}$ | LB50 | ApaI | TTGGGCCCAGCGTAATCTGGAACATCGTATGGGTAAATCGGCACATATTCAACCAGATGAATCTT |
| Fwd PR2Q V64amb | LB76 | blunt end ligation | GAAAACATCATTAACT**TAG**TCTGGTCTTGTTACTGGTATTGCTTTCTGG |
| Rev PR2Q V64amb | LB77 | blunt end ligation | CATTTTGCAGAAACTCTATCTCTTTCAACAAAGAAAAATAC |
| Fwd removal of 4 positive charges from loop1&2 of PR2Q | LB62 | blunt end ligation | GTATTTTTCTTTGTTGAAGCGGATGCGGTTTCCGCAGCGTGGGCGACATCATTAACTGTATCTGGT |
| Fwd removal of 2 positive charges from loop1&2 of PR2Q | LB64 | blunt end ligation | GTATTTTTCTTTGTTGAAAGAGATGCGGTTTCCGCAGCGTGGAAAACATCATTAACTGTATCTGGT |
| Rev removal of 2 positive charges from loop1&2 of PR2Q | LB65 | blunt end ligation | ACCAGATACAGTTAATGATGTTTTCCACGCTGCGGAAACCGCATCTCTTTCAACAAAGAAAAATAC |

## 3.4 Bacteria and Media

The following *E.coli* strains were used for plasmid amplification, protein expression and expression of ribosome nascent chain complexes.

**Table 3: Bacterial strains with the respective genotype used in this study**

| Strain | Usage | Genotype | Reference |
|---|---|---|---|
| DH5α | Plasmid amplification | F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK- mK+), λ– | Stratagene |
| XL1Blue | Plasmid amplification | endA1 gyrA96(nalR) thi-1 recA1 relA1 lac glnV44 F'[ ::Tn10 proAB+ lacIq Δ(lacZ)M15] hsdR17(rK- mK+) | Stratagene |
| BL21(DE3) | Protein expression | F– ompT gal dcm lon hsdSB(rB- mB-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5]) | Stratagene |
| KC6 ΔssrAΔsmpB | Expression of TnaC-stalled nascent chains | Rna-19 gdhA2 his-95arelA1 spoT1 metB1 ΔendA met+ ΔtonA ΔspeA ΔtnaA ΔsdaA ΔsdaB ΔgshA ΔsmpB ΔssrA | (Calhoun and Swartz, 2006; Seidelt et al., 2009) |

## 3.5 Molecular Cloning

### 3.5.1 Polymerase chain reaction (PCR)

For the amplification and the mutagenesis of double-stranded DNA fragments, polymerase chain reaction was used. The DNA was amplified using the Phusion Flash High-Fidelity Master Mix (Finnzymes) or the KOD Hot Start DNA Polymerase (Merck/Millipore), both according to the manufacturer's protocols. The annealing temperature was not set to a fixed value. Instead it was set around 5 °C – 10 °C above the higher melting points of the two primers that is provided by company synthesizing the DNA oligo (Metabion or MWG). Subsequently, the annealing temperature is decreased by 0.5 °C in each cycle of the PCR reaction.

Typically 20-25 cycles were run per PCR with an elongation time of 1min/kbp. A typical PCR program is shown in Table 4.

**Table 4: Program for 'Touchdown' PCR**

| 'Touchdown' PCR Program | | | | |
|---|---|---|---|---|
| Step 1 | Initial DNA melting | 98 °C | 120 s | |
| Step 2 | DNA melting | 98 °C | 10 s | Steps 2 – 4 |
| Step 3 | Primer annealing | 65 °C − 0.5 °C per cycle | 10 s | 20 – 25 cycles |
| Step 4 | Elongation | 68 °C | 1 min per kbp | |
| Step 5 | Final elongation | 68 °C | 2.5 min per kbp | |

Preparation and mutagenesis of plasmids

For the preparation and mutagenesis of expression vectors, two major methods were used.
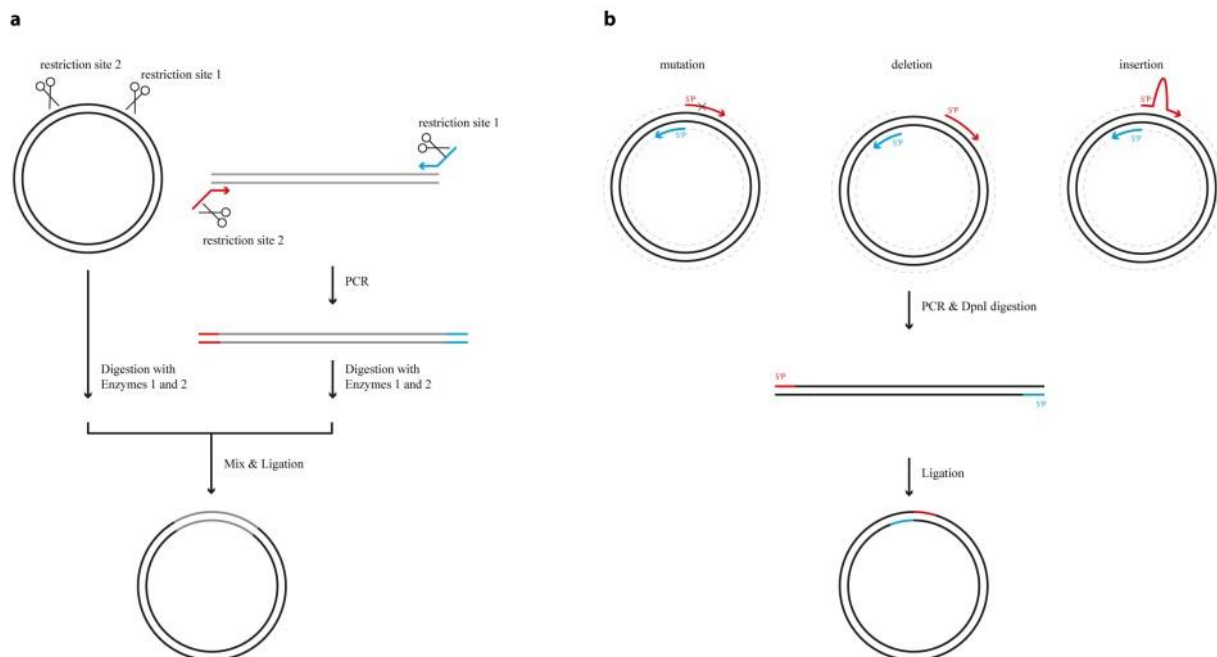


**Figure 12: Schematic of molecular cloning strategies**
a: molecular cloning using restriction enzymes.
b: restriction free, blunt-end mutagenesis.

a: Molecular cloning using restriction enzymes

For cloning using restriction enzymes, a primer pair containing recognition sites (if necessary in an overhang) for two orthogonal restriction enzymes were used to amplify the desired DNA fragment. Subsequently, the amplified DNA fragments, as well as the target vector,

31

were digested using these restriction enzymes. After purification of the digested vector and insert from an agarose gel, the insert was subsequently ligated into the vector (see Figure 12a).

b: Restriction-free blunt end ligation

For short insertions, deletions or point mutants, the method of blunt end mutagenesis was used. Thereby, the whole plasmid is amplified by PCR using a 5' phosphorylated primer pair containing the mutation or an overhang to create an insertion. After digesting the parental DNA with DpnI, the purified DNA fragment containing phosphorylated 5' ends can be set in a ligation reaction (see Figure 12b).

## 3.5.2 Purification of PCR products

PCR products were purified using the QIAquick PCR purification kit (Qiagen) according to the manufacturer's protocol. Optional steps were omitted and the final DNA was eluted in 40 µL ddH$_2$O.

## 3.5.3 Restriction

The purified PCR products were digested with the two chosen restriction enzymes simultaneously following suggested DoubleDigest protocols (NEB). The reaction conditions and buffers were chosen according to the NEB catalogue.

## 3.5.4 Degradation of parental DNA with DpnI

DpnI is a restriction endonuclease that digests exclusively methylated DNA, hence DNA amplified by PCR remains unaffected. For the digestion, 1 µL of DpnI (NEB) was added to 40 µL of purified PCR product in the buffer provided by the manufacturer and incubated for at least 1 h at 37 °C.

## 3.5.5 Analysis of DNA fragments by agarose gel electrophoresis

Agarose gel electrophoresis separates DNA strands according to their size. To this end a 1 % agarose gel in TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) was prepared by

heating the buffer agarose suspension until all agarose was dissolved. After pouring and cooling/polymerization of the agarose gel, it was covered with TAE buffer. The DNA samples were mixed 1:10 with 10x sample buffer and loaded next to a DNA standard marker. Electrophoresis was carried out with a current of 120 V for 30 min – 45 min and detection was performed on a Multi Gel Jet (INTAS) system.

### 3.5.6 Purification of DNA fragments from agarose gels

DNA fragments were purified by agarose gel electrophoresis and subsequently re-isolated using the QIAquick Gel extraction kit (Qiagen) according to the manufacturer's protocol. The DNA was finally eluted in 40 µL ddH$_2$O.

### 3.5.7 Ligation

To ligate an insert into an expression plasmid the restrictase-treated insert was added to the treated vector in 5-10 fold molar excess and subsequently ligated using T4 ligase (NEB) for at least 1 h at 16 °C. In a more straightforward protocol leading to a higher yield of positive clones, the insert and the vector were isolated from the agarose gel using the same gel extraction column. Finally the vector/insert mix was eluted in 40 µL ddH$_2$O and 17 µL were set in the ligation reaction.
For the self-ligation of 5'-phosphorylated DNA fragments 17 µL of the elution after the agarose gel extraction were set in the ligation reaction.

**Table 5: Reaction mixture for DNA ligation**

| Ligation reaction | |
|---|---|
| Mixed vector- and insert-DNA | 17 µL |
| 10x T4-DNA-ligase buffer | 2 µL |
| T4 DNA-ligase | 1 µL |

The ligation products were directly transformed in competent *E.coli* DH5α or XL1blue cells.

### 3.5.8 Plasmid isolation

For the amplification of plasmid DNA, 5 mL of *E.coli* DH5α or XL1blue cells harboring the respective plasmid were grown over-night in 5 mL LB medium. Cells were harvested by centrifugation and plasmid DNA was isolated with the QIAprep Miniprep Kit (Qiagen) following the manufacturer's instructions.

### 3.5.9 DNA sequencing

For sequencing 500 ng – 1000 ng of DNA in 15 µL ddH$_2$O were mixed with 0.3 µL of sequencing primer (100 µM). Standard sequencing primers could optionally be added by the sequencing company (MWG).

## 3.6 Analysis of proteins

### 3.6.1 Protein concentration

To concentrate protein solutions an Amicon Ultra-15 Centrifugal Filter Unit / Ultracel-10 membrane (Millipore) with the proper molecular weight cut-off was used. Before concentrating the protein solution, the membrane was equilibrated with 2 mL of the corresponding buffer. Thereafter, the protein solution was concentrated by centrifuging at 3000 g and 4 °C until the desired volume or concentration was reached.

### 3.6.2 Protein concentration determination

To determine the concentration of protein solutions, the absorption at 280 nm of a small sample (usually 1 µL – 2 µL) was measured on a Nanodrop ND-1000 (Peqlab). From the measured absorption and the specific extinction factor (calculated on the webserver: www.scripps.edu/~cdputnam/protcalc.html) the protein concentration could be determined.

### 3.6.3 Protein precipitation

Small amount of proteins were precipitated with trichloracetic acid (TCA) before being analyzed by SDS-PAGE. To this end, the sample was filled up to 1 mL with cold ddH$_2$O before 100 µL of 72 % TCA and 100 µL of 15 % Na-deoxycholate were added. The samples were mixed and incubated on ice for at least 1 h or at -20 °C over-night. Subsequently, the precipitated proteins were pelleted by centrifugation in a pre-cooled tabletop centrifuge (Eppendorf) for 30 min at 14.000 rpm at 4 °C. The supernatant was discarded and the protein pellet was washed with 500 µL ice-cold acetone. After re-pelleting the precipitated proteins were resuspended in an appropriate volume of SDS sample buffer.

### 3.6.4 SDS-polyacrylamid gel electrophoresis (PAGE)

For the analysis of standard protein samples, they were separated based on their molecular weight by performing an SDS-PAGE. Protein samples were mixed with the proper amount of 4x SDS sample buffer (200 mM Tris 6.8, 8 % (w/v) SDS, 0.4 % (w/v) bromphenol blue, 40 % (v/v) glycerol, 400 mM DTT).

The anionic detergent sodium dodecylsulfate (SDS) denatures the protein and superposes the intrinsic charge of the protein with a constant negative charge; thus the proteins are only separated according to their molecular weight. The commonly used gels contained 15 % acrylamide in the separation gel (Laemmli) and were run in a Mini-Protean II electrophoresis cell (Bio-Rad) at 200 V in standard SDS running buffer (25 mM Tris, 192 mM glycine and 0.1 % SDS).

For the analysis of some special protein samples, like cross-linking products, pre-casted NuPAGE 4-12 % gradient gels (Invitrogen) were used, following the provider's instructions. These gels were run in MES running buffer (50 mM MES, 50 mM Tris base, 3.5 mM SDS, 1.0 mM EDTA free acid).

To visualize the proteins separated by SDS-PAGE, the gels were stained with Coomassie staining solution (0.25 % (w/v) Coomassie Blue R 250, 50 % (v/v) ethanol, 10 % (v/v) acetic acid) or SimplyBlueTM SafeStain (Invitrogen). The gel was subsequently destained applying several washing steps in destaining solution (40 % (v/v) ethanol, 10 % (v/v) acetic acid).

Small protein amounts, not detectable by Coomassie staining, could be detected by SYPRO Orange staining. Therefore, the gel was rinsed with water and subsequently incubated for 45 min in 10 mL 7.5 % acetic acid and 1:5000 SYPRO Orange. Afterwards, the gel was rinsed with 7.5 % acetic acid for 30 s and scanned on the Typhoon 9400 Variable Mode Imager (Amersham Biosciences).

### 3.6.5 Western blotting

<u>Semi-dry western blotting</u>

Western blotting was used for the qualitative and semi-quantitative detection of tagged proteins. For HA-tag or SecY detection, SDS-PAGE gels were blotted for 60 min at constant current of 75 mA per gel, in the blotting buffer (48 mM Tris, 39 mM glycine, 0.037 % (w/v) SDS, 20 % methanol (v/v)) on PVDF or nitrocellulose membranes, using a standard semi-dry blotting apparatus (Bio-Rad).

<u>Tank western blot</u>

The pre-casted NuPAGE 4-12 % gradient gels (Invitrogen) were blotted on PVDF membrane in a tank using NuPAGE blotting buffer (25 mM Bicine, 25 mM Bis-Tris, 1 mM EDTA free acid, 20 % methanol (v/v).

### 3.6.6 Antibody detection

<u>αHA antibody detection</u>

For αHA antibody detection, the membrane was incubated 30 min in 5 % milk/TBS (50 mM Tris 7.4, 150 mM NaCl) at room temperature to minimize unspecific binding. Subsequently, the membrane was incubated with the first antibody: HA-F/7 1:500 (Santa Cruz) in 2 % BSA/TBS at 4 °C over-night. After washing the membrane with 2x TBS-T (50 mM Tris 7.4, 150 mM NaCl, 0.1 Tween 20 (v/v)) and 1x with TBS, for 5 min each, the second HRP-conjugated antibody (goat anti-mouse) (Santa Cruz) in 5 % milk TBS was applied to the membrane for 1 h at room temperature. Afterwards, the membrane was again washed with 2x TBS-T buffer and 1x TBS buffer, 5 min each step.

<u>αStrep antibody detection</u>

For αStrep antibody detection, the membrane was incubated 1 h in 5 % milk/TBS (50 mM Tris 7.4, 150 mM NaCl) at room temperature to minimize unspecific binding. Subsequently, the membrane was incubated with the anti-Strep antibody (IBA Lifescience) in TBS at room-temperature for at least 1 h. The membrane was washed twice in TBS-T for 1 min each and additionally 1 min with TBS.

<u>αSecY antibody detection</u>

In order to detect SecY with an antibody, the membrane was incubated over-night in 5 % milk/TBS-T at 4 °C to minimize unspecific binding. Subsequently, the membrane was

incubated with the anti-SecY antibody: 1:3000 (generous gift from A. Driessen, Groningen, The Netherlands) in 5 % milk/TBS-T for 1 h at room temperature. After washing the membrane with 3x TBS-T (for 10 min each) the second antibody (HRP-anti rabbit, Santa Cruz) was applied to the membrane for 1 h at room temperature in 5 % milk/TBS-T. Afterwards the membrane was again washed with 2x TBS-T buffer and 1x TBS buffer, 5 min each step.

<u>Detection</u>

The detection was carried out on film (GE Healthcare) with the Enhanced Chemi Luminescence Detection Kit for Western Blot (ApplieChem) as prescribed by the provider. The film was exposed for 10 s to 20 min until descent signal intensity was reached.

## 3.7 Protein expression and purification from *E.coli*

### 3.7.1 Pre-culture, induction and expression in general

A single colony was picked from the selection plate and transferred to 2 mL of LB medium containing the specific antibiotic. This small pre-culture was shaken at 37 °C, 140 rpm. After several hours 100 mL – 200 mL LB containing the specific antibiotic were inoculated with the small pre-culture and shaken over night at 37 °C, 140 rpm.

Finally, the expression was carried out in pre-warmed LB medium containing the specific antibiotic inoculated 1:50 with the over-night pre-culture. The expression culture was further grown at 37 °C and shaken at 110 rpm until the optical density at 600 nm ($OD_{600}$) reached ~ 0.5. To induce the protein expression of proteins under the control of the lac promotor, 0.5 mM IPTG (isopropyl-D-thiogalacto-pyranoside) was added. Expression of proteins und the control of an arabinose promotor was induced by adding 0.2 % arabinose.

### 3.7.2 Purification of *E.coli* SecY(His)EG

Overexpression of SecYEG was carried out in BL21(D3) cells, which were grown to an $OD_{600}$ of 0.5 before expression was induced for 3 h at 37 °C, as described above. The cells were harvested by centrifugation in a pre-cooled SLC6000 rotor (Beckman Coulter) for 10 min at 4.000 rpm and 4 °C. Cell pellets were resuspended in the resuspension buffer (20 mM Tris 7.5, 200 mM NaCl, 5 mM $MgCl_2$, 10 % glycerol 1:500 EDTA-free Complete proteinase inhibitors (Roche Applied Science)). The whole purifications procedure was carried out at

4 °C or on ice. Cells were lysed in a microfluidizer (M-110L, Microfluidics) by passing them through at 18 kPsi for 3 times. Subsequently, cell debris was removed by centrifugation at 16,000 rpm in a SS34-rotor (Sorvall) at 4 °C. In the next step, the *E.coli* membranes were isolated by spinning the SS34 supernatant in a Ti45 rotor (Beckman Coulter) for 1 h at 40,000 rpm and 4 °C. Following, the brownish membrane pellet was resuspended in the solubilization buffer (20 mM Tris 7.5, 1 M NaCl, 5 mM $MgCl_2$, 10 % glycerol, 1 % DDM, 1:500 EDTA-free Complete proteinase inhibitors) and the membranes were solubilized by incubation in solubilization buffer for 1 h on a rotating wheel. Non-solubilized material was removed by centrifugation in a Ti45 rotor for 30 min at 40,000 rpm and 4 °C. Meanwhile, 1 mL Talon slurry (Clontech) per liter of cell culture was pre-washed twice with the solubilization buffer and transferred to a fresh Eppendorf column, and the solubilized membrane proteins were incubated on the Talon beads for 1 h at 4 °C. The flow-through was discarded and the beads were washed with 10 column volumes (CV) of the washing buffer (20 mM Tris 7.5, 200 mM NaCl, 5 mM $MgCl_2$, 10 % glycerol, 20 mM imidazole 0.1 % DDM. 1:500 EDTA-free Complete proteinase inhibitors) and then eluted in 2 CV elution buffer (20 mM Tris 7.5, 200 mM NaCl, 5 mM $MgCl_2$, 10 % glycerol, 200 mM imidazole 0.1 % DDM, 1:500 EDTA-free Complete proteinase inhibitors). Elution was analyzed by SDS-PAGE and SecYEG-containing fractions were pooled and concentrated to 0.5 mL as described above. Finally, SecYEG was further purified on an S200 gel-filtration column (Äkta System, GE Healthcare). Therefore the column was equilibrated in the gel-filtration buffer (20 mM Tris 7.5, 100 mM NaCl, 5 mM $MgCl_2$, 10 % glycerol, 0.1 % DDM, 1:500 EDTA-free Complete proteinase inhibitors) prior to loading of the protein sample. Elution fractions were analyzed by SDS-PAGE and SecYEG-containing fractions were pooled, concentrated and flash frozen in liquid nitrogen. Samples were stored at -80 °C until further usage.

### 3.7.3 Purification of *E.coli* SecA(His)

Overexpression of SecA was carried out in BL21(D3) cells, which were grown to an $OD_{600}$ of 0.5 and expression was induced for 1.5 h at 37 °C as described above. The cells were harvested by centrifugation in a pre-cooled SLC6000 rotor for 10 min at 4,000 rpm at 4 °C. Cell-pellets were resuspended in resuspension buffer (20 mM Tris 7.5, 50 mM KOAc, 5 mM $MgOAc_2$, 10 % glycerol 1:500 EDTA-free Complete proteinase inhibitors, 1 mM DTT). The whole purifications procedure was carried out at 4 °C or on ice. Cells were lysed in a microfluidizer (M-110L, Microfluidics) by passing them through at 18 kPsi for 3 times. Subsequently, cell debris was removed by centrifugation at 16,000 rpm in a SS34-rotor (Sorvall) at 4 °C. The SS-34 supernatant was subsequently incubated for 1 h at 4 °C on Ni-

NTA beads (Invitrogen, 1 mL per liter cell culture) that were pre-washed in resuspension buffer. Subsequently, the beads were washed with 10 CV of washing buffer (20 mM Tris 7.5, 500 mM KOAc, 5 mM MgOAc$_2$, 20 mM imidazole, 10 % glycerol 1:500 EDTA-free Complete proteinase inhibitors, 1 mM DTT) and eluted in 5 CV elution buffer (20 mM Tris 7.5, 50 mM KOAc, 5 mM MgOAc$_2$, 200 mM imidazole, 10 % glycerol 1:500 EDTA-free Complete proteinase inhibitors, 1 mM DTT).

The elution fractions were analyzed by SDS-PAGE and SecA containing fractions were pooled and concentrated to 0.5 mL as described above. Finally SecA was further purified on a Superose6 gel-filtration column (Äkta System, GE Healthcare). Therefore the column was equilibrated in gel-filtration buffer buffer (20 mM Tris 7.5, 50 mM KOAc, 5 mM MgOAc$_2$, 10 % glycerol, 1 mM DTT) prior to loading of the protein sample. Elution fractions were again analyzed by SDS-PAGE and SecYEG containing fractions were pooled, concentrated and flash frozen in liquid Nitrogen. Samples were stored at -80 °C until further usage.

### 3.7.4 Purification of *E.coli* FtsY(His)

Overexpression of FtsY was carried out in BL21(D3) cells, which were grown to an OD$_{600}$ of 0.5 and expression was induced for 3 h at 37 °C as described above. The cells were harvested by centrifugation in a pre-cooled SLC6000 rotor (Beckman Coulter) for 10 min at 4,000 rpm at 4 °C. Cell-pellets were resuspended in resuspension Buffer (50 mM Tris 7.5, 300 mM KCl, 0.1 % DDM, 1:500 EDTA-free Complete proteinase inhibitors (Roche Applied Science), 20 mM Imidazole). The whole purifications procedure was carried out at 4 °C or on ice. Cells were lysed in a microfluidizer (M-110L, Microfluidics) by passing them through at 18 kPsi for 3 times. Subsequently, cell debris was removed by centrifugation at 16,000 rpm in a SS34-rotor (Sorvall) at 4 °C. The supernatant was loaded on Ni-NTA beads (Invitrogen), pre-washed in resuspension buffer (1 mL Ni-NTA slurry per liter of cell culture). After incubation for 1 h at 4 °C, the beads were washed with 10 CV washing buffer (50 mM Tris 7.5, 500 mM KCl, 0.1 % DDM, 1:500 EDTA-free Complete proteinase inhibitors, 20 mM Imidazole) and subsequently eluted in 2 CV elution buffer (20 mM Tris 7.5, 200 mM KCl, 0.1 % DDM, 1:500 EDTA-free Complete proteinase inhibitors, 200 mM Imidazole, 10 % glycerol). The elution fractions were analyzed by SDS-PAGE and SecA containing fractions were pooled and concentrated to 0.5 mL as described above. Finally, FtsY was further purified on a S200 gel-filtration column (Äkta System, GE Healthcare). Therefore the column was equilibrated in gel-filtration buffer (20 mM Tris 7.5, 200 mM KCl, 0.01 % DDM, 1:500 EDTA-free Complete proteinase inhibitors, 200 mM Imidazole, 10 % glycerol) prior to loading of the protein sample. Elution fractions were again analyzed by SDS-PAGE and FtsY

containing fractions were pooled, concentrated and flash frozen in liquid Nitrogen. Samples were stored at -80 °C until further usage.

# 3.8 Purification of TnaC-stalled Ribosome nascent chain complexes (RNCs)

## 3.8.1 Purification of TnaC-stalled ribosome nascent chain complexes from the whole cell lysate

The TnaC-stalled ribosome nascent chain complexes were purified from *E. coli* strain KC6 Δ*smpB*Δ*ssrA*. All constructs encoding for nascent chains used N-terminal of the TnaC stalling sequence were subcloned in the pBAD vector (Invitrogen) under the control of an arabinose promotor. *E.coli* cells were grown at 37 °C in LB medium to an $OD_{600}$ of 0.5. Expression of the nascent chain was induced for 1h by adding 0.2 % arabinose.

Cells were harvested and resuspended in buffer A (50 mM HEPES 7.2, 250 mM KOAc, 25 mM $MgOAc_2$, 2 mM tryptophan, 0.1 % DDM and 0.1 % EDTA-free Complete proteinase inhibitors (Roche Applied Science)). Cells were lysed by passing two times through a microfluidizer (M-110L, Microfluidics) and cell debris was removed by centrifugation for 20 min at 16,000 rpm in a SS34 rotor (Sorvall). The cleared lysate was centrifuged through a sucrose cushion (750 mM sucrose) in buffer A at 25,000 rpm for 20 h in a Ti45 rotor (Beckman Coulter).

The crude ribosomal pellet was resuspended in a small volume of buffer A.

Ribosomes carrying the nascent chain were separated by affinity chromatography using Talon beads (Clontech), that were additionally pre-incubated with 10 µg/mL *E. coli* tRNAs to minimize unspecific binding of ribosomes. After incubating for 1 h at 4°C the beads were washed with at least 10 column volumes (CV) of buffer B (50 mM HEPES pH 7.2 500 mM KOAc, 25 mM $MgCl_2$, 0.1 % DDM, 2 mM tryptophan). RNCs were eluted in Buffer B + 150 mM imidazole and loaded on a linear sucrose gradient (10% - 40% sucrose in buffer B). After spinning for 3 h at 40,000 rpm in a SW40 rotor (Beckman Coulter) the 70S peak was collected and diluted three times with buffer B. RNCs were finally concentrated by spinning for 4 h at 40,000 rpm in a Ti70 rotor (Beckman Coulter) and resuspended in an appropriate volume of Grid buffer (20 mM HEPES pH 7.2, 50 mM KOAc, 5 mM $MgOAc_2$, 125 mM sucrose, 2 mM tryptophan, 0.03 % DDM). (protocol adapted from(Bischoff et al., 2014b)).

## 3.8.2 Purification of TnaC-stalled ribosome nascent chain-ligand complexes

In order to analyze ligands interacting with the expressed nascent chains, the protocol was essentially conducted as described above, however, the first centrifugation through the sucrose cushion was replaced by a simple centrifugation step in the Ti45 rotor (Beckman Coulter) for 1 h at 40,000 rpm and 4 °C. The resulting membrane/ribosome pellet was the solubilized for 1 h in buffer A + DDM. Additionally, buffer conditions were changed:

Buffer A: 50 mM HEPES pH 7.2, 150 mM $NH_4Cl$, 10 mM $MgCl_2$, 2 mM tryptophan, 0.1 % DDM and 0.1 % EDTA-free Complete proteinase inhibitors (Roche Applied Science)

Buffer A + DDM: 50 mM HEPES pH 7.2, 150 mM $NH_4Cl$, 10 mM $MgCl_2$, 2 mM tryptophan, 1 % DDM and 0.1 % EDTA-free Complete proteinase inhibitors

Buffer B: 50 mM HEPES pH 7.2, 150 mM $NH_4Cl$, 10 mM $MgCl_2$, 2 mM tryptophan, 10 mM Imidazole 0.1 % DDM and 0.1 % EDTA-free Complete proteinase inhibitors

Grid buffer: 20 mM HEPES pH 7.2, 50 mM $NH_4Cl$, 5 mM $MgCl_2$, 125 mM sucrose, 2 mM tryptophan, 0.03 % DDM

## 3.8.3 Purification of membrane bound Ribosome nascent chain complexes

To analyze the interaction of nascent membrane proteins with SecYEG and other co-factors involved in membrane protein biogenesis, only RNCs bound to the inner *E.coli* membrane were purified.

Cells were harvested and resuspended in Buffer A (20 mM Tris pH 7.5, 150 mM $NH_4Cl$, 10 mM $MgCl_2$, 1 mM tryptophan, 250 µg/mL Chloramphenicol and 0.1 % EDTA-free Complete proteinase inhibitors (Roche Applied Science). Cells were lysed by passing two times through a microfluidizer (M-110L, Microfluidics) and debris were removed by centrifugation for 20 min at 16,000 rpm in a SS34-rotor (Sorvall). The cleared lysate was centrifuged for 1 h at 40,000 rpm in a Ti45 rotor (Beckman Coulter), the resulting crude membrane pellet was resuspended in a small volume of buffer A and then equally divided over several 5 step sucrose density gradients consisting of 0.5 mL 55 %, 50 %, 45 %, 40 % and 35 % sucrose in buffer A in a TLA110 (Beckman Coulter) tube. After spinning for 30 min at 100,000 rpm, the brownish layer of inner membranes was taken out from the lower third of the tube and subsequently solubilized in Buffer A + 1 % DDM.

Ribosomes carrying the nascent chain were separated by affinity chromatography using StrepTactin Beads (IBA) or Talon (Clontech). Talon beads were additionally pre-incubated with 10 µg/mL *E.coli* tRNAs to minimize unspecific binding of ribosomes. After incubating for 1 h at 4°C the beads were washed with at least 10 column volumes (CV) of Buffer B (20 mM Tris pH 7.5 150 mM $NH_4Cl$, 10 mM $MgCl_2$ 250 mM sucrose, 0.1 % DDM). RNCs were eluted in Buffer B + 2.5 mM Desthiobiotin or 150 mM imidazole and loaded on a linear sucrose gradient, 10 % - 40 % sucrose in buffer B. After spinning for 3 h at 40,000 rpm in a SW40 rotor (Beckman Coulter) the 70S peak was collected and diluted three times with buffer B. RNCs were finally concentrated by spinning for 4 h at 40,000 rpm in a Ti70 rotor (Beckman Coulter) and resuspended in an appropriate volume of buffer B. RNCs were flash frozen in liquid $N_2$ and stored at -80 °C. (protocol adapted from (Bischoff et al., 2014b)).

## 3.9 RNC-SecYEG complex purification for cryo-EM data collection

*E. coli* strain KC6 Δ*smpB*Δ*ssrA* cells were transformed with a plasmid encoding for the respective nascent chain and SecY(His)EG and grown at 37 °C to an $OD_{600}$ of 0.5. Expression of SecY(His)EG was generally not induced, in some cases SecYEG was induced by 10 µM IPTG 30 min prior to induction of the nascent chain. The nascent chain was induced for 1 h by adding 0.2 % arabinose.

Inner membranes were isolated and solubilized as described above for the purification of membrane-bound RNCs. SecY-RNC complexes were purified on Talon beads (Clontech), pre-incubated in buffer B + 20mM imidazole and 10 µg/ml E. coli tRNAs to minimize unspecific binding of ribosomes. After washing the beads with 50 column volumes buffer B + 20 mM imidazole, SecY-RNC complexes were eluted in buffer B + 150 mM imidazole. Isolation of monosomal complexes over a linear sucrose gradient and final concentration was performed as described above. (Protocol adapted from(Bischoff et al., 2014b))

## 3.10 Site-directed cross-linking with an unnatural amino acid

For *in vivo* and *in vitro* site-directed cross-linking, cells need to express a protein containing the UV-activatable, unnatural amino acid p-benzoyl-L-phenlylalanine (pBPA). The following protocol was adapted from (Farrell et al., 2005).

*E.coli* KC6 were co-transformed with the pEVOL-pBPF (Chin et al., 2002) plasmid encoding for a mutated amino acyl tRNA synthetase as well as the amber suppressor tRNA and the plasmid encoding for the protein or nascent chain, containing an amber codon (TAG) at the desired position. Cells are then grown and expression is induced as described in chapter 3.7.1. in LB medium containing 1 mM pBPA. pBPA is loaded on the amber suppressor tRNAs by the mutated amino acyl tRNA synthetase and thus incorporated in a protein or a nascent chain of interest.

For *in vivo* cross-linking, cells were harvested by centrifugation in a Rotanta 460R centrifuge (Hettich) and resuspended in 10 mL resuspension buffer. Subsequently, 5 mL of the resuspended cells were poured into a sterile Petri dish and irradiated with a pre-warmed (15 min) Black Ray B100AP UV lamp from 4 – 5 cm distance at 4 °C for 20 min. The other 5 mL of cell suspension was not treated with UV-light as a negative control. Subsequently, RNCs were purified via an affinity column as described above and analyzed by SDS-PAGE and western blotting.

For *in vitro* cross-linking the pBPA containing RNCs in complex with SecY were purified by affinity purification as described above. When indicated, separately purified SecYEG was added to the sample prior to cross-linking. The sample was irradiated with a pre-warmed (15 min) Black Ray B100AP UV lamp from 4 -5 cm distance at 4 °C for 20 min and samples were analyzed by SDS-PAGE and western blotting.

## 3.11 *In vitro* translocation assay

### 3.11.1 Preparation of inverted membrane vesicles

In order to prepare inverted membrane vesicles (IMVs) *E.coli* BL21(DE3) cells were grown to an $OD_{600}$ of 1.0 and inner membranes were isolated over a 5-step sucrose gradient as described above. The layer of inner membranes was taken from the upper half of the tube and incubated with 6 M Urea for 30 min to remove any membrane-bound or membrane-associated protein. Subsequently, the sample was diluted 1:10 and the urea-treated

membranes were concentrated by centrifugation through a 750 mM sucrose cushion (17 h, 24,000 rpm in a Ti70 rotor (Beckman Coulter)).

### 3.11.2 Preparation of SecYEG proteoliposomes

To prepare SecY-containing proteoliposomes, the synthetic lipids DOPC, DOPE, DOPG and DAG (all purchased from Avanti lipids) were mixed in a 4:3:3:1 molar ratio and resuspended in water at a concentration of 20 mg/mL. Then, 80 µL of this suspension was diluted with 300 µL buffer A (20 mM Tris pH 7.5, 150 mM $NH_4Cl$, 10 mM $MgCl_2$ ), 20 µL 10 % Triton-X100 and 20 µL of SecYEG (3.75 mg/mL). The mixture was then incubated with 80 mg Bio-beads SM2 sorbent (Bio-Rad) at 4 °C. The Bio-beads were replaced after 2 h and the sample was incubated with the fresh Bio-beads for another hour. Subsequently, the proteoliposomes were collected by centrifugation and resuspended in buffer A.

### 3.11.3 Translocation assay

The *in vitro* translocation reaction was carried out using the PURE translation system (NEB). For the initial tests described in this study, samples of 50 µL reaction volume were set up containing 2 µL of either proteoliposomes or IMVs and always 1 µL of the protein co-factors, like SecA (described in 3.7.3 Purification of *E.coli* SecA(His), Page 38), SecB (gift from P. Flagmeier), SRP (gift from E. van der Sluis) and FtsY (described in 3.7.4 Purification of *E.coli* FtsY(His). Translation was started by adding 150 ng of plasmid DNA and carried out for 2 h at 37 °C. Subsequently, the reaction was stopped by adding CAM and the insertion efficiency of the membrane protein was estimated by the protease protection assay. Therefore, the sample was split in 3 aliquots: One was treated directly with 4x SDS sample buffer, one was incubated with 0.2 mg/mL proteinase K (NEB) for 20 min on ice directly and the third was incubated with 1 % Triton-X100 for 5 min at room temperature before proteinase K treatment. The protease was inhibited by adding PMSF before all samples were analyzed by SDS-PAGE and western blotting.

## 3.12 Ribosome binding assay

To test the binding of a potential ligand to the ribosome or the ribosome nascent chain complex, 10 pmol – 20 pmol of ribosomes (or RNCs) were mixed with a 10x excess of the

44

potential ribosome interactor in a total volume of 20 µL. The mixture was incubated at 37 °C for 15 min and then loaded on 150 µL of a 750 mM sucrose cushion in a TLA100 tube (Beckman Coulter). After the sample was centrifuged for 80 min at 50.000 rpm, the supernatant was quickly removed from the tube and TCA precipitated. The pellet was directly resuspended in 1x SDS sample buffer. Both samples were analyzed by SDS-PAGE.

## 3.13 Cryo-EM data collection

### 3.13.1 Data collection of the PR2Q-SecYE complex

The freshly prepared PR2Q-SecYE complex was applied to 2 nm pre-coated Quantifoil R3/3 holey carbon supported grids in a concentration of 3 $A_{260mn}$/mL. The grids were vitrified using a Vitrobot mark IV (FEI Company). Cryo-EM data was collected on a Titan Krios TEM (FEI Company) under low-dose conditions (about 20 $e^-/Å^2$) at an acceleration voltage of 300 kV. Data was recorded on a CCD camera (TemCam-F416 CMOS, TVIPS GmbH; 4096 x 4096 pixels, 15.6 µm/pixel, 1 s per full frame) at a nominal defocus between -1 µm and 3.5 µm. Magnification of 148,721 at the plane of the CCD resulted in a pixel size of 1.049 Å. Vitrification and data collection was performed by Dr. Otto Berninghausen and Charlotte Ungewickell.

### 3.13.2 Data collection of the TnaC-stalled ribosome

The freshly prepared TnaC-stalled ribosome nascent chain complex at a concentration of 3 $A_{260nm}$ was mixed with a 5x excess of *E.coli* SRP, incubated for 15 min at 30 °C and then applied to 2 nm pre-coated Quantifoil R3/3 holey carbon supported. The grids were vitrified using a Vitrobot mark IV (FEI Company). Cryo-EM data was collected on a Titan Krios (FEI Company) equipped with a Cs-corrector and operated at an acceleration voltage of 300 kV. Data was recorded on a back-thinned Falcon II DDD (FEI Company). The detector was calibrated for a nominal magnification of 125,085x resulting in a pixel size of 1.10 Å at the plane of the specimen. 8 frames per second were recorded with an electron dose of 4 $e^-$ per frame and a defocus between 0.8 µm and 2.2 µm. The individual frames were aligned using the 'Gatan Microscopy Suite 2.30.463.1' to correct for drift during exposure and were subsequently summed up using the SPIDER command AD S.

The sample was vitrified by Charlotte Ungewickell (LMU) and data was collected by Dr. Sasha De Carlo at NeCEN (Leiden, Netherlands).

## 3.14 Cryo-EM data processing

### 3.14.1 Preprocessing of cryo-EM data

Cryo-EM data was pre-processed using the SPIDER and WEB software package (Frank et al., 1996).

The collected 16 bit *.tiff imaged collected on the CCD camera for the PR2Q-SecYE complex and the 32 bit *.raw images collected on the Direct Detection Device were first converted into SPIDER-compatible files using the SPIDER command 'CP FROM RAW'. The Contrast Transfer Function (CTF) and the correct defocus were determined with the command 'TF ED' for each micrograph. Visual inspection of all micrographs and power-spectra was carried using the program JWEB. All micrographs showing drift, contamination or astigmatism were excluded from the dataset. The coordinates of the individual particles were extracted using the program SIGNATURE (Chen and Grigorieff, 2007) with a 70S ribosome as reference. For the PR2Q-SecYE sample the dataset of individual particles was further classified applying the machine learning algorithm MAPPOS (Norousi et al., 2013) for the detection of false positive particles.

For faster parallel computing, the dataset was finally reorganized by dividing the dataset into groups of particles with similar defocus; up to 2,500 particles were merged into one defocus group.

The initial alignment of the dataset was performed at 3x decimation using a cross-correlation based projection matching algorithm, where each particle was aligned to 84 reference projections of an empty *E.coli* ribosome as initial reference executing the SPIDER command 'AP MQ'.

### 3.14.2 Refinement and Sorting

The quality and resolution of the reconstructed 3D volume was improved by an iterative projection matching process using the SPIDER software package (Frank et al., 1996). In each round of refinement, all particles are aligned to a masked and filtered reference 3D volume that was created in the previous refinement round; the initial reference was used as reference in the first refinement round. For each round of refinement, the angular increment of the reference projection was gradually decreased resulting in a finer sampling rate and eventually in a higher resolved 3D reconstruction after each refinement round. A Butterworth low-pass filter (SPIDER command FQ) was set anew in every refinement round

to remove frequencies that were higher than the noise level in a calculated Fourier Shell Correlation (FSC).

After the electron density is fully refined, e.g. no resolution or feature improvement is achieved over a couple of refinement rounds, wrongly aligned particles are removed from the dataset. Therefore, particles showing a significantly lower cross-correlation value than the average particle in the same defocus group were considered to be wrongly aligned and thus discarded from the dataset.

For the TnaC-stalled ribosome sample the filter was not set below 8 Å for the whole refinement procedure to prevent potential over-fitting of the data (Scheres and Chen, 2012).

The initially heterogeneous dataset was sorted into more homogeneous subsets using cross-correlation-based competitive projection matching, offering more than one reference ('semi-supervised' sorting). Hence, every particle is aligned to several distinct references and assigned to the class of the reference with the highest cross-correlation value during alignment. Every subset of particles is back-projected individually to a new 3D volume and subsequently used as a new reference in the next round of sorting. This iterative process is continued until the number of particles in each group remains stable over a couple of sorting rounds. Sorting is generally carried out going from global to minor differences in the offered references. A 3D volume reconstructed from contamination and carbon edges ('edge-volume') was usually used in the first round of sorting to clean the dataset from non-ribosomal particles.

In order to sort for only minor differences, e.g. presence of a small ligand in otherwise identical 3D volumes, a focused-sorting strategy was applied. Particles were aligned to several distinct references as describe above, however, the cross-correlation, determining the subset to which particle is finally assigned to, is only calculated in a small area. This area is defined by a mask in the area of the expected difference between the two 3D volumes.

### 3.14.3 Resolution determination

Lacking a clear physical resolution criterion as in X-ray crystallography, the FSC is used to estimate the resolution in cryo-EM. For the PR2Q-SecYE complex the FSC = 0.5 criterion was used to estimate the resolution of the reconstruction. Therefore, the FSC was calculated as described above using the SPIDER command 'RF3' and the spatial frequency, where the FSC drops below the threshold of 0.5 was defined as the estimated resolution.

47

For the TnaC-stalled RNC 8 Å frequency cut-off during refinement as well as the so-called 'gold-standard' refinement (refining two half datasets totally independent and calculation of the FSC between these independent half volumes) was applied in order to prevent the potential over-fitting of the data (Scheres and Chen, 2012). Following these strategies, the resolution is estimated by the FSC = 0.143 criterion. The obtained 3D volume of the TnaC-stalled ribosome was B-factor sharpened using automated methods in the program EM-BFACTOR (Fernandez et al., 2008).

### 3.14.4 Interpretation of the electron densities and model building

All electron densities were visualized using the programs COOT (Emsley and Cowtan, 2004) and UCSF Chimera (Pettersen et al., 2004). For the interpretation of the electron densities, existing protein and RNA models were docked in the density as a rigid body. If necessary, *E.coli* homology models were created on existing structures using HHPred (Soding et al., 2005; Hildebrand et al., 2009) and MODELLER (Eswar et al., 2006; Eswar et al., 2007; Eswar et al., 2008). The models were then manually adjusted to fit the obtained electron density. However, the modeling was restricted by the resolution of the obtained 3D volume, thus only secondary structure elements, as α-helices, were modelled as rigid body in the electron density of the PR2Q-SecYE. The higher resolution of the TnaC-stalled RNC, however, allowed the modelling of individual RNA bases as well as bulky amino acid side-chains. All newly modeled parts were refined using COOT (Emsley and Cowtan, 2004).

### 3.14.5 Model validation

To test for over-fitting and to validate the model of the TnaC-stalled RNC, the FSC between the built model and both of the half-maps from the 'gold standard' refinement were calculated.

### 3.14.6 Figure preparation

All figures showing molecular models and electron density were created in UCSF Chimera (Pettersen et al., 2004). Densities were separated using the 'Color Zone' and 'Split Map' function of UCSF Chimera.

# 4. Results

## 4.1 Visualization of a polytopic membrane protein during SecY-mediated membrane insertion

### 4.1.1 Purification of *in vivo* assembled ribosome nascent chain complexes

To study the mechanisms of membrane protein insertion *in vivo* on a structural level, we established a protocol that allows the isolation of *in vivo* assembled RNCs from *E.coli* in high yield and quality. The efficient tryptophan dependent TnaC stalling sequence (Seidelt et al., 2009; Ito and Chiba, 2013; Bischoff et al., 2014a) was used to create RNCs in a special *E.coli* strain. This strain, *E.coli* KC6ΔsmpBΔssrA is tryptophanase-deficient, leading to artificially high levels of tryptophan in the cell that trigger efficient TnaC stalling. Additionally, the tmRNA system that normally rescues ribosomes that are stalled on a nascent peptide is also knocked out.
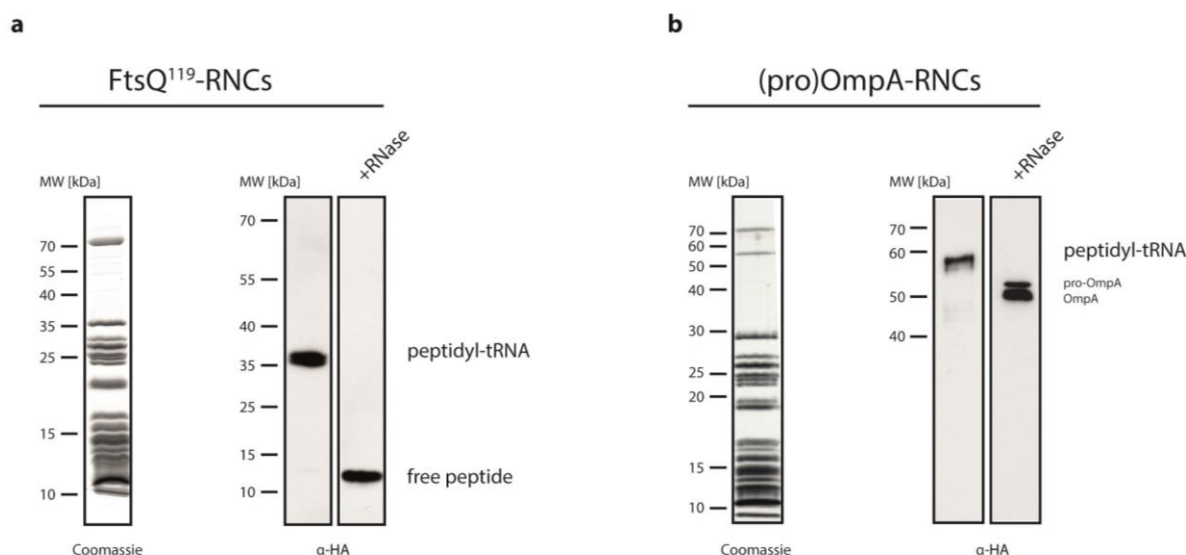


**Figure 13: Purification of *in vivo* assembled ribosome nascent chain complexes**
Coomassie stained SDS-PAGE gel and western blot decorated with an anti-HA antibody of FtsQ$_{119}$ (a) and (pro)OmpA (b) RNCs. The presence of the peptidyl-tRNA was verified by incubation with RNase. The peptidyl-tRNA of (pro)OmpA is visible as an extra band among the ribosomal proteins. The FtsQ$_{119}$ peptidyl-tRNA is not visible as it is hidden behind the ribosomal proteins.
Figure modified from (Bischoff et al., 2014b).

As a model membrane protein, Proteorhodopsin (PR) from the Gamma-proteobacterium EBAC31A08 (Beja et al., 2000; Beja et al., 2001) was chosen because:

(1) a NMR structure of PR in solution (Reckel et al., 2011) revealed only short loops connecting the transmembrane helices on the cytoplasmic site, making it unlikely to hamper a tight engagement of the RNC with the protein conducting channel.

(2) Placement of an affinity tag between the cleavable signal sequence and the first transmembrane segment does not alter the N-terminus of the protein and thus should not impair the proper targeting of PR to the membrane. Additionally, after signal sequence cleavage, an affinity tag on the periplasmic site should be freely accessible for affinity purification.

Several versions of TnaC PR, containing only the signal sequence (PR0TM) and increasing numbers of transmembrane segments were constructed (PR1TM to PR7TM (wildtype)). A StrepII tag was introduced after the predicted signal sequence cleavage site, C-terminal the membrane protein followed by a hemaglutinine (HA) tag for detection and the TnaC stalling sequence (see Figure 14).

## 4.1.2 Tightly vs. loosely coupled RNC-SecY complexes

The different nascent membrane proteins were expressed in *E.coli*KC6ΔsmpBΔssrA and *in vivo* assembled ribosome nascent chain complexes were purified under mild conditions to allow the co-purification of the SecYEG complex.

As a control RNCs carrying the functional (Urbanus et al., 2001) and structural (Frauenfeld et al., 2011) well-characterized type II membrane protein $FtsQ_{119}$ were purified as well as a shorter version, $FtsQ_{85}$.
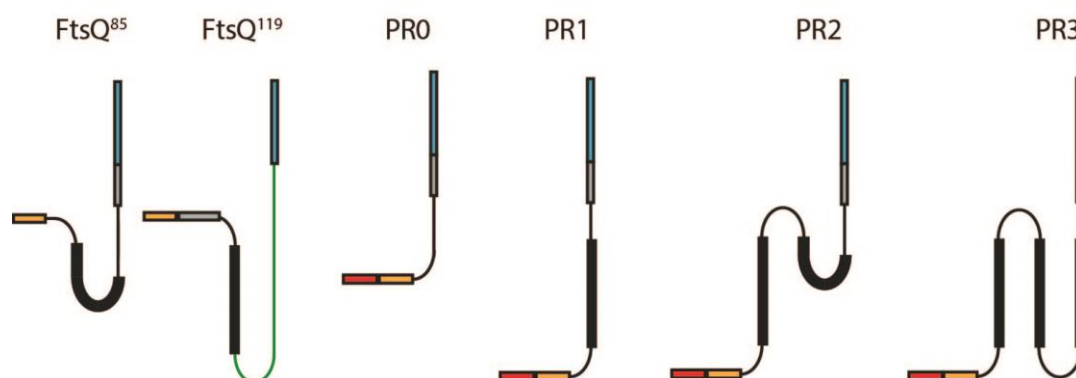


**Figure 14: Schematic illustration of the nascent membrane protein constructs**
Turquois: TnaC stalling sequence, Grey: HA tag, Black bold: Transmembrane segments, orange: affinity tag for purification, red: signal sequence, green: hydrophilic stretch of FtsQ.
Figure modified from (Bischoff et al., 2014b).

As expected, a stable engagement of FtsQ$_{119}$ with SecY could be established and considerable amounts of SecY could be co-purified with the respective RNCs, the shorter FtsQ$_{85}$ co-purify only very small amounts of SecY (see Figure 15).

Unexpectedly, no detectable amounts of SecY could be purified at all together with the PR0 to PR4.



**Figure 15: Purification of RNC-SecY complexes, I**
Western blot analysis to detect co-purified SecY with RNCs derived from PR reveals that SecY only co-purifies with FtsQ RNCs.
Figure modified from (Bischoff et al., 2014b).

Assuming that the hydrophilic loop of FtsQ results in a different affinity of SecY to the RNC, we designed hybrid substrates of PR and FtsQ by fusing the hydrophilic loop of FtsQ (amino acids 54-85) at C-termini of the transmembrane segments 2, 4 and 6 of the PR constructs resulting in the PR2Q, PR4Q and PR6Q hybrid proteins (see Figure 16). To exclude that these unnatural chimeric proteins induce an 'off-pathway' and artificially induce tight coupling, amino acids 27-151 of cytochrome b$_o$3 oxidase (CyoA27-151), a well-studied SecY substrate (du Plessis et al., 2006), were inserted between the PR signal sequence and the HA tag, resulting in an endogenous *E.coli* protein that has a similar pattern of hydrophobic and hydrophilic stretches as the PRxQ hybrid constructs.
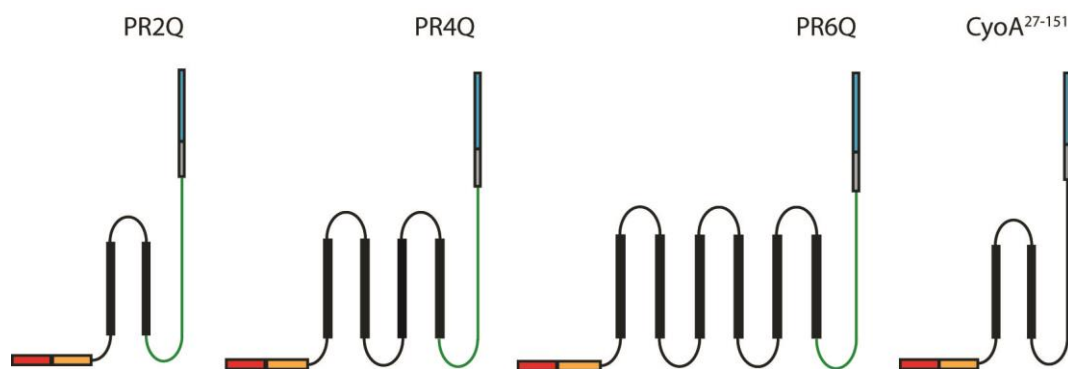


**Figure 16: Schematic illustration of the chimeric PR-FtsQ (PRxQ) nascent membrane protein constructs and CyoA(27-151)**
Turquois: TnaC stalling sequence, Grey: HA tag, Black bold: Transmembrane segments, orange: affinity tag for purification, red: signal sequence, green: hydrophilic stretch of FtsQ.
Figure modified from (Bischoff et al., 2014b).

Analysis of the PRxQ hybrid constructs as well as CyoA$_{(27-151)}$ revealed tight coupling of the RNCs with SecY and thus the co-purification of considerable amounts of SecY.
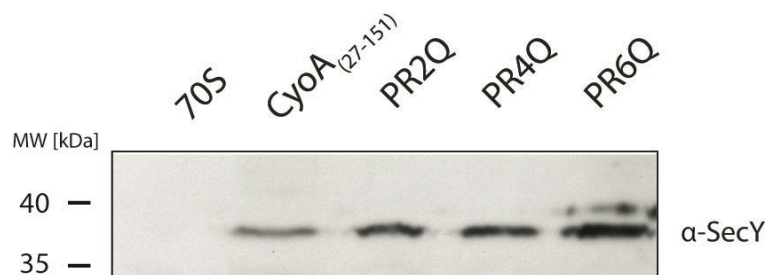


**Figure 17: Purification of RNC-SecY complexes, II**
Western blot analysis to detect co-purified SecY with RNCs derived from PRxQ and CyoA reveals that SecY co-purifies with the chimeric nascent chains and the endogenous CyoA.
Figure modified from (Bischoff et al., 2014b).

To further characterize the interaction of RNCs with SecY *in vivo*, we additionally purified RNCs from a well-characterized, post-translationally translocated protein, proOmpA. Detection of the nascent chain revealed, that most of the nascent chain was indeed inserted into the translocon, as most of the signal sequence has been cleaved, however, no SecY could be co-purified with these RNCs. In order to assure that co-purification of SecY with FtsQ$_{119}$ and the PRxQ hybrid construct is not only due to the hydrophilic part of FtsQ, it was fused C-terminally to proOmpA (OmpA-FtsQ). For both wild-type proOmpA and OmpA-FtsQ a stable RNC:SecY complex could not be observed (see Figure 18). Moreover, inserting a hydrophobic transmembrane segment (TM3 of PR) in the hydrophilic loop of the PRxQ hybrid constructs (FtsQ-PR3TM) leads to loose coupling of SecY to the RNC again.
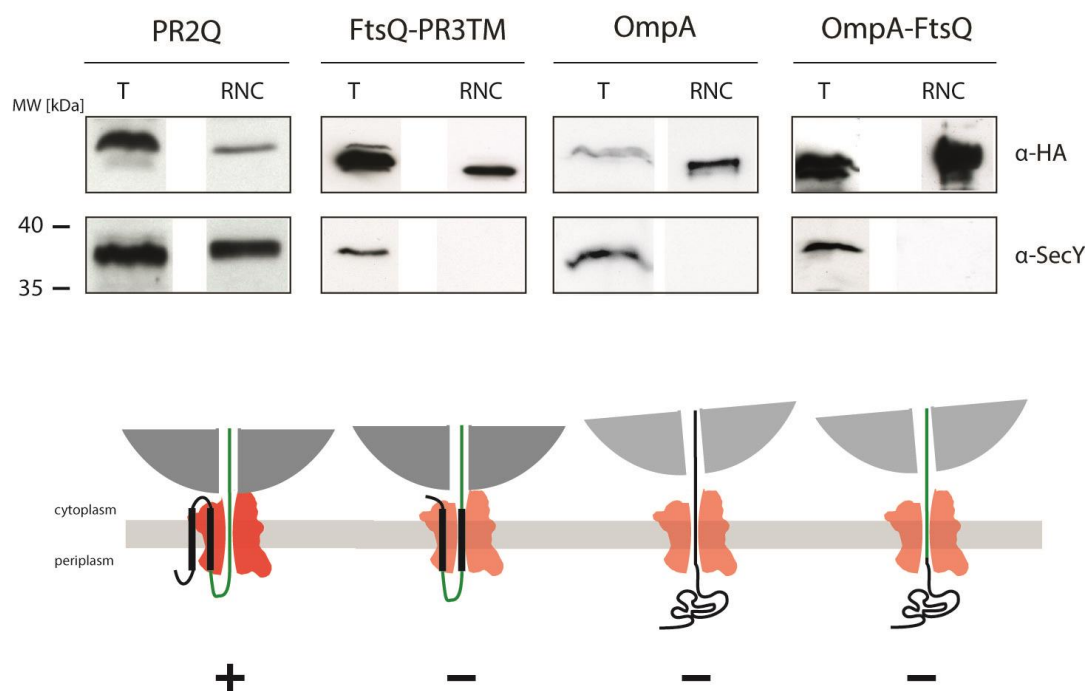
**Figure 18: Requirements for tightly coupled RNC-SecY complexes**
Tight coupling of SecY to the translating ribosomes requires a co-translational mode of translocation and a hydrophilic stretch following the hydrophobic transmembrane segment.
T: sample taken from the total cell lysate, RNC: final purified RNC construct, Western blots are decorated with antibodies against SecY and the nascent chain (HA). Figure modified from (Bischoff et al., 2014b).

These results demonstrate that tight coupling of SecY to the translating ribosome *in vivo* is only established during co-translational translocation, when inserting a transmembrane segment with type II topology that is followed by a hydrophilic loop.

Interestingly, elongation of the FtsQ hydrophilic loop leads to loose coupling and to a complete loss of SecY, as soon as more than 55 to 60 amino acids are outside the ribosomal exit tunnel.
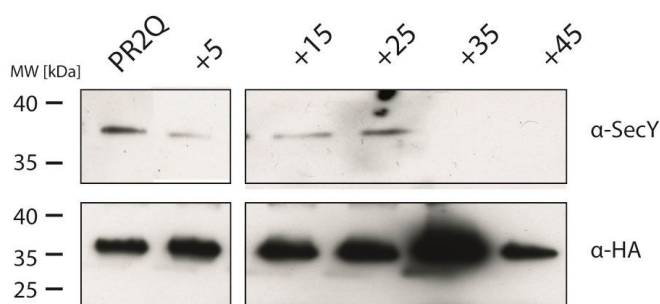


**Figure 19: Coupling of SecY to the RNC depends on the length of the nascent chain**
Elongation of the hydrophilic stretch of PR2Q by 25 amino acids converts a tightly coupled RNC-SecY complex to a loosely coupled.
Figure modified from (Bischoff et al., 2014b)

53

Taken together, it was possible to generate tightly coupled RNC-SecY complexes under totally native conditions and to purify them for structural studies by cryo-EM and single particle reconstruction.

### 4.1.3 Cryo-EM structure of a tightly coupled RNC-SecY complex

Since the yields and the quality of PR2Q were best amongst the model substrates, this construct was expressed to assemble an RNC-SecY complex *in vivo*. To increase the stoichiometry of SecYEG over the RNCs, cells were additionally transformed with a plasmid encoding for a C-terminally His-tagged SecYEG via which the RNC-translocon complexes were purified.
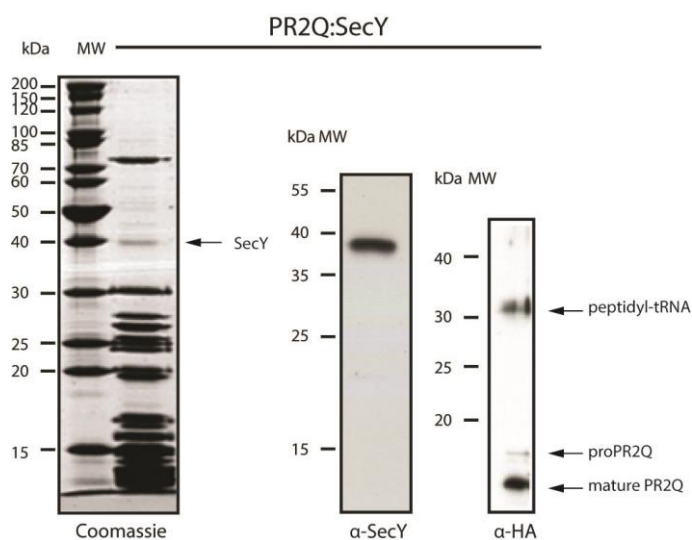


**Figure 20: Evaluation of the purified PR2Q-SecY complex**
Coomassie stained SDS-PAGE gel of the PR2Q-SecY complex; the presence of SecY and a nascent PR2Q chain was proved by western blotting. The nascent chain appears almost exclusively with a cleaved signal sequence. Figure modified from (Bischoff et al., 2014b).

Cryo-EM analysis revealed heterogeneity of the sample, however, applying several global (Becker et al., 2009; Frauenfeld et al., 2011) and focused (Bradatsch et al., 2012) sorting steps on the dataset resulted in a stable sub-volume of 47,471 particles that could be refined to an overall resolution of 7.3 Å according to the FSC = 0.5 criteria.
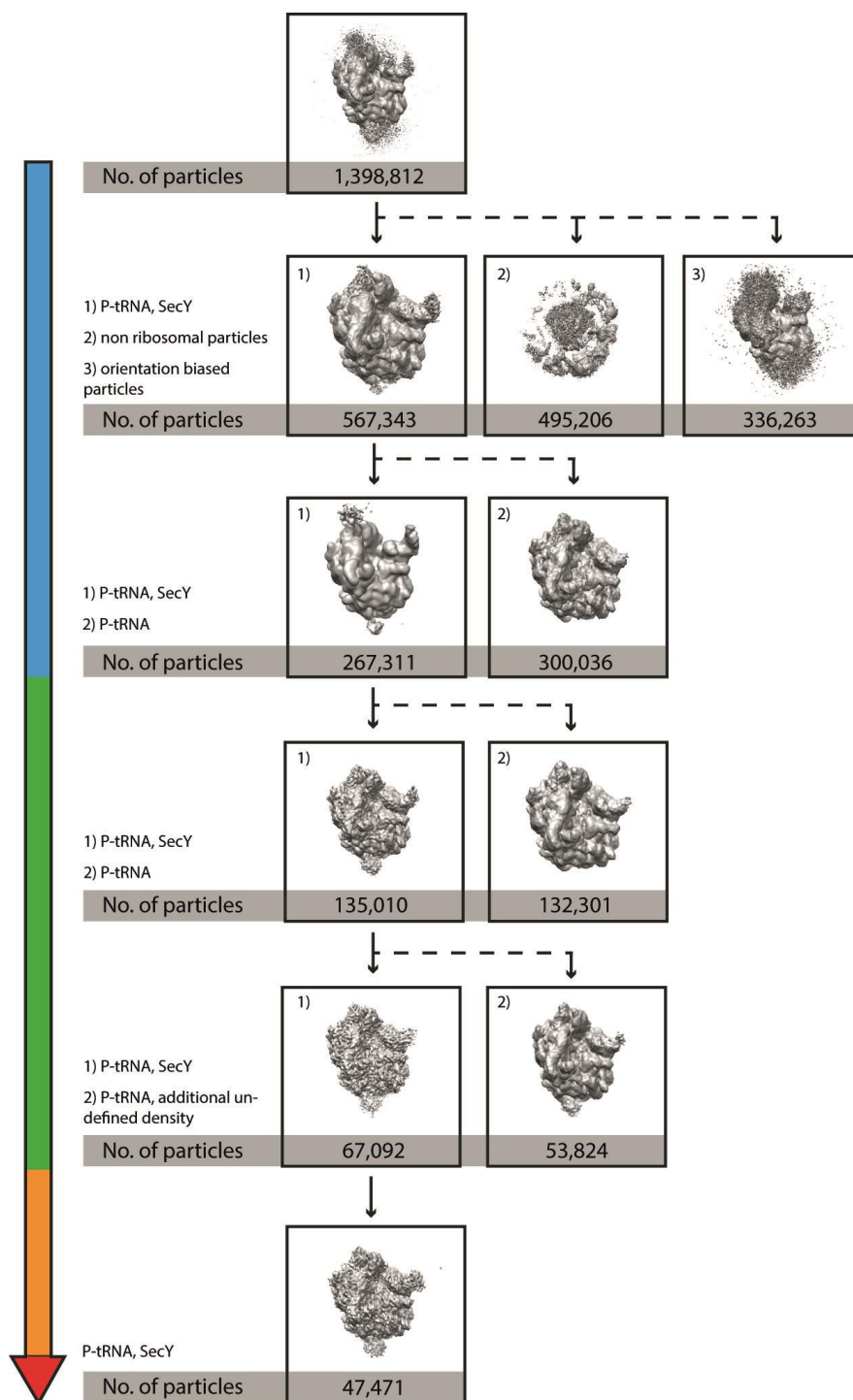
**Figure 21:** *In silico* **sorting scheme of the PR2Q-SecY dataset**
In a first global sorting step (blue) non ribosomal particles and particles of preferred orientations could be sorted out.
In two rounds of focused sorting (green) ribosomes without bound SecY and with a displaced density at the ribosomal exit tunnel could be sorted out.
Removal of particles with an extremely low cross-correlation coefficient (orange) leads to a higher resolved volume.
A final sub-dataset of 47,471 particles was refined to 7.3Å (red).

The cryo-EM reconstruction of this complex shows the appearance of a programmed 70S ribosome, with an expected additional density at the ribosomal exit tunnel accounting for the co-purified translocon (see Figure 23). Comparing the size of the additional density with available crystal-structures of SecYEG clearly shows that only a monomer of SecYE is bound to the ribosome. Calculation of the local resolution (Kucukelbir et al., 2014) of the volume revealed a homogeneously resolved ribosome with a resolution of about 7 Å at the periphery and even higher resolution in the core of the ribosome.



**Figure 22: Overall and local resolution of the refined electron density**
a: The PR2Q-SecY complex could be refined to an overall resolution of 7.3 Å according to the FSC = 0.5 criterion.
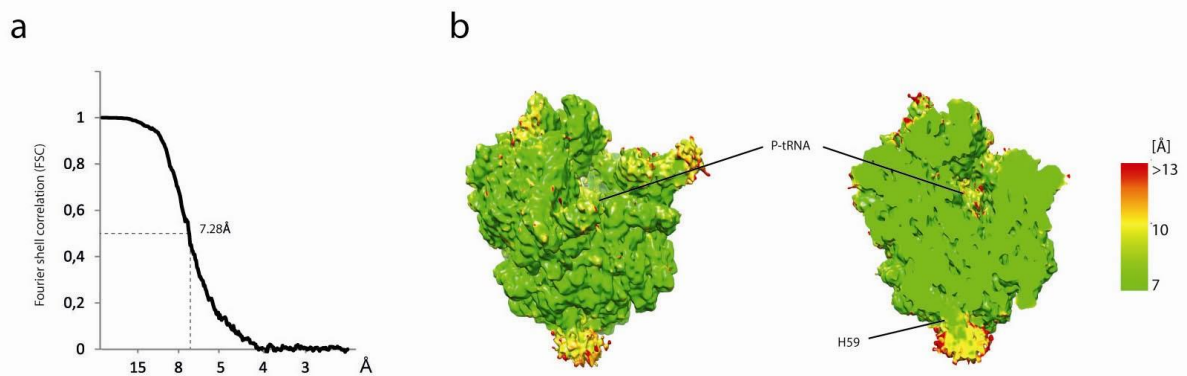b: Calculation of the local resolution revealed a homogenously resolved ribosome and a less resolved ligand (full view and cut-through). Areas of the ligand that are connected to the ribosome are higher resolved, while flexible parts are resolved between 7Å and 10Å. The ligand is surrounded by a less resolved detergent micelle. Figure modified from (Bischoff et al., 2014b).

Notably, the density accounting for SecY is resolved to 7 - 8 Å in areas connected to the ribosome and to 8 - 10 Å in areas not connected to the ribosome and thus more flexible regions. In detergent-covered, solvent-exposed regions the resolution decreases above 1 nm. However, the sub-nanometer resolution in most parts of the ligands allows the identification of secondary structure element such as α-helices.
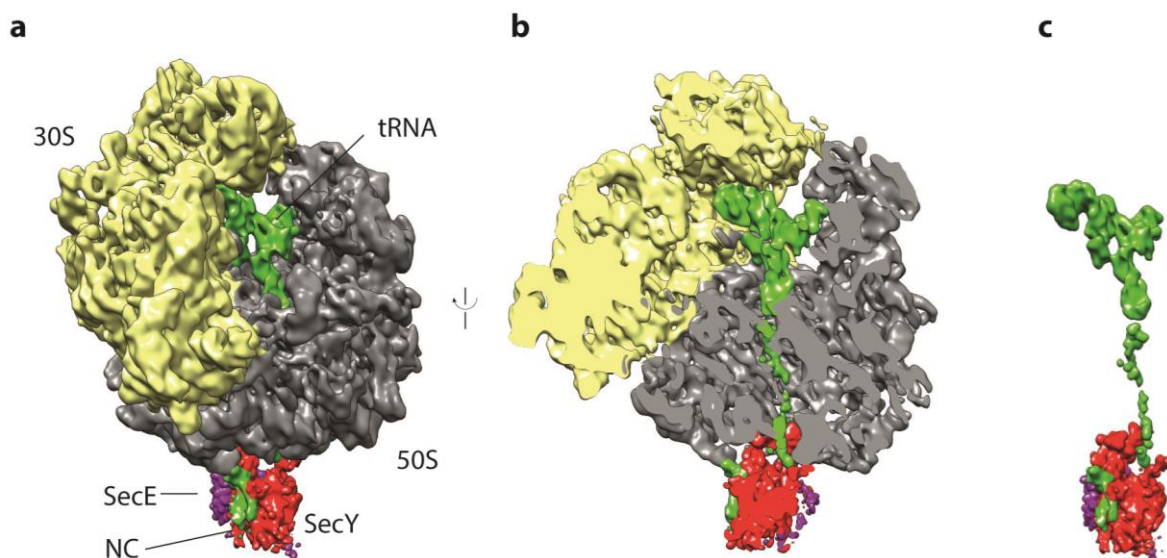
**Figure 23: Cryo-EM reconstruction of a tightly coupled RNC-SecY complex**
a: Overall appearance of the tightly coupled PR2Q-RNC-SecY complex. The small subunit (30S) is depicted in yellow, the large subunit (50S) in grey. The nascent chain is colored in green. SecY is shown in red and SecE in purple.
b: The same view as in a but cut-through perpendicular to the membrane to visualize the nascent chain in the ribosomal exit tunnel.
c: Isolated density for the P-tRNA, the nascent chain and SecYE.
Figure modified from (Bischoff et al., 2014b).

## 4.1.4 Molecular model of the tightly coupled RNC-SecY complex

In order to interpret the obtained electron density on a molecular level, a homology model of *E.coli* SecY based on a crystal structure of the SecY-SecA complex from *Aquifex aeolicus* (Zimmer et al., 2008) was docked into the density as a rigid body. The initial positioning was based on previously observed contacts of the cytosolic loops 8/9 and 6/7 of SecY to the large ribosomal subunit (Becker et al., 2009; Frauenfeld et al., 2011).

The α helices were then manually fitted into the cryo-EM density using COOT (Emsley and Cowtan, 2004). The C-terminal half of SecY fitted remarkably well with only miniscule changes, whereas the N-terminal half showed more flexibility and thus had to be adjusted by shifting the helices to a larger extent.

The two N-terminal helices of SecE could be fitted into two α-helical densities. Although the putative position of SecG is known (Breyton et al., 2002; Hizlan et al., 2012), we could not assign any density in our structure to SecG. Thus, it is not included in the final model. At the local resolution map already indicated, SecY helices 8 and 9 as well as 6 that have direct connections to the ribosome could be fitted remarkably well into the density, since they are well resolved. However, the more flexible helices in the N-terminal half, especially helix 1

and 4, did not show continuous rod like densities for the whole helix. This could be either due to the overall more flexible N-terminal half of SecY, but also the absence of the stabilizing SecG that allows especially helices 1 and 4 to be more flexible.

Notably, the area around the proposed lateral gate, helices 2b, 3 and 7, is well resolved and allows unambiguous placement of the respective helices.
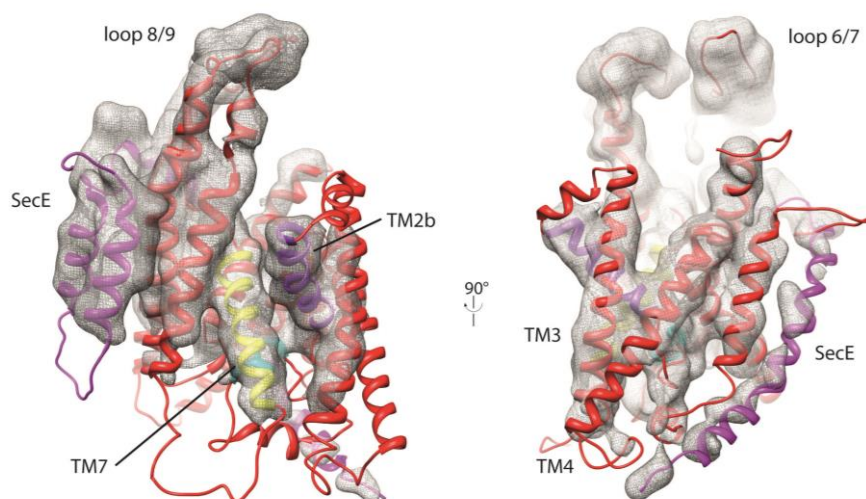


**Figure 24: Homology model of *E.coli* SecYE manually fitted in the experimental density**
Helices TM7 and TM2b of the lateral gate of SecY are colored in yellow and purple, respectively. SecE is colored in magenta.
Figure modified from (Bischoff et al., 2014b).

Identification of the contacts between SecYE and the translating ribosome revealed high similarities to the previously published structure of SecYE in the membrane environment (Frauenfeld et al., 2011) indicating that the solubilization of SecYE in detergent does not alter the binding position of the translocon to the ribosome and maintains the geometry that is set up in the lipid bilayer. Moreover, it is in good agreement with earlier published structures of solubilized SecY (Menetret et al., 2007; Park et al., 2014) and Sec61 (Becker et al., 2009).

Particularly, we find contacts of SecY loop 8/9 that reaches towards the ribosomal exit tunnel with the ribosomal rRNA helices H50, H53 and H59 and the ribosomal protein L23, whereas loop 6/7 contacts rRNA helices H6, H24 and H50. Additionally, both loops contact ribosomal protein L23. As reported before (Frauenfeld et al., 2011), the very C-terminus of SecY shows contacts to rRNA helix 24 as well as ribosomal protein L24. The amphipathic helix of SecE contributes to the binding by contacting ribosomal proteins L23 and L29, which was also shown previously (Frauenfeld et al., 2011).

<u>Conformational state of the tightly coupled SecY complex</u>

The overall conformation of the manually adjusted SecY model is distinct from the template structure (cross-correlation coefficient (ccc) = 0.88).

Comparison of our model to the 'closed' crystal structure of *M. jannaschii* (Van den Berg et al., 2004), shows an overall less compact conformation (ccc = 0.82), while the plug helix is only slightly shifted, to provide space for the translocating peptide. Concerning the lateral gate, our model exhibits a more open conformation, compared to the idle SecY, which is mainly caused by a shift of helix 2b towards helix 3, while helix 7 remains in a similar position (see Figure 25). In the crystal structure of the SecY-SecA complex (Zimmer et al., 2008) SecY also adopts a more compact conformation (ccc = 0.79), possibly due to the absence of a substrate. While the stable C-terminal moiety is very similar, the lateral gate of our model clearly adapts a more closed state, by movement of both, helices 2b and 7.

Interestingly, our model resembles to a high degree (ccc = 0.90) the 'pre-opened' SecY found in the crystal structure form *Thermus thermophilus* SecY that was co-crystallized with a FAB fragment against loop 8/9 (Tsukazaki et al., 2008). The lateral gate shows the same, pre-opened conformation, and the plug helix is also in a very similar position.

Comparing our model to a recently published cryo-EM based model of *E.coli* SecY (Park et al., 2014), on the one hand our model is more opened, when compare to an *E.coli* SecY reconstituted with non-translating ribosomes. On the other hand, it is less opened, when compared to the model of a cross-linked RNC-SecY complex engaging a signal sequence.

However, the structure of the tightly coupled PR2Q-SecY is very similar to the model of a RNC-Sec61 complex translocating a hydrophilic substrate (Gogala et al., 2014).
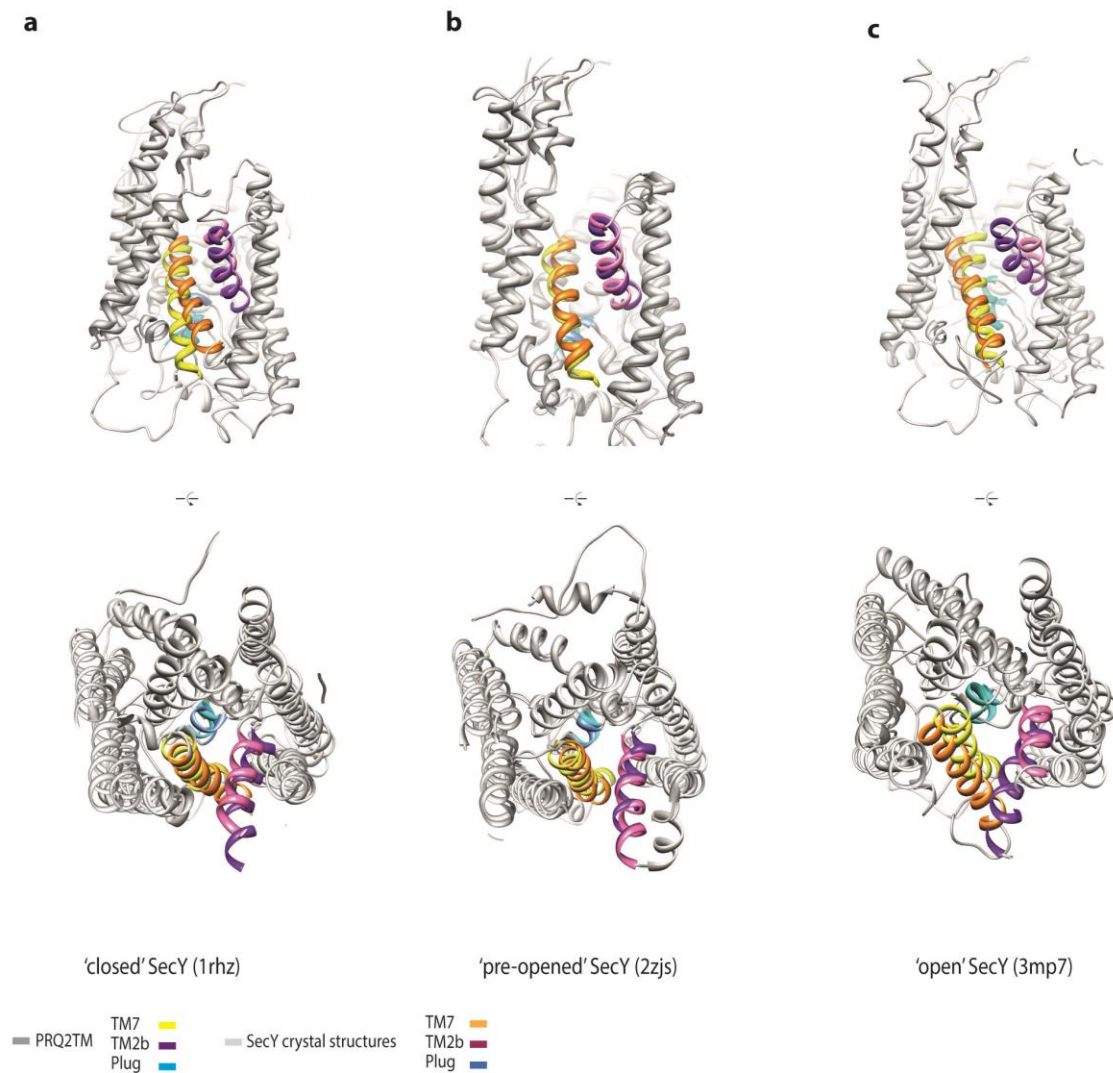
**Figure 25: Molecular model of SecY compared to other published structures**
Comparison of the overall conformation of our model with crystal-structures of SecY in the 'closed', 'pre-opened' and 'open' state.
Figure modified from (Bischoff et al., 2014b).

To sum it up, we provide structural evidence that the lateral gate of SecY is in a 'pre-opened' conformation in the moment of translocation, very similar to the conformation observed by Tsukazaki and colleagues (Tsukazaki et al., 2008).

Moreover, we see only a slight shift of the plug helix compared to the idle (Van den Berg et al., 2004), nevertheless, providing enough space for an unstructured stretch of amino acids to be translocated.

## 4.1.5 Following the nascent chain

Most remarkable, however, analysis of the electron density accounting for SecYE revealed two additional rod-like densities right in front of helices 2b and 7, the lateral gate of SecY that we interpreted as the first two transmembrane segments of PR2Q. The corresponding density was extracted and compared to the NMR structure of Proteorhodopsin (Reckel et al., 2011) filtered to 9 Å. A molecular model based on the latter could indeed be fitted as a rigid body into the density. Notably, the similarity in the two densities allows building of an accurate model for the nascent membrane protein. The TMS that was closer to the lateral gate was assigned to be PR helix 2, however, final evidence for this assignment is lacking.

The position of the two helices relative to SecY provides very strong evidence that transmembrane segments are indeed released from SecY by an opening of helices 2b and 7 that were proposed to form the lateral gate.

The TMS that was closer to the lateral gate was assigned to be PR helix 2, however, final evidence for this assignment is lacking.

Since the resolution in the area of SecY is limited, it was not possible to observe a clear density for the unstructured hydrophilic stretch (FtsQ part) outside the ribosomal exit tunnel and thus no model was built. However, it is most likely still inside the hydrophilic surrounding in the center of SecY since the properties (hydrophilic vs. hydrophobic) of this stretch determine whether SecY remains tightly coupled to the RNC or not. Additionally, cross-links from nascent substrates to the central constriction of SecY have been observed before (Cannon et al., 2005). In a similar study of Sec61 bound to RNCs it was shown that an unstructured stretch of amino acids is able to pass through the central constriction.
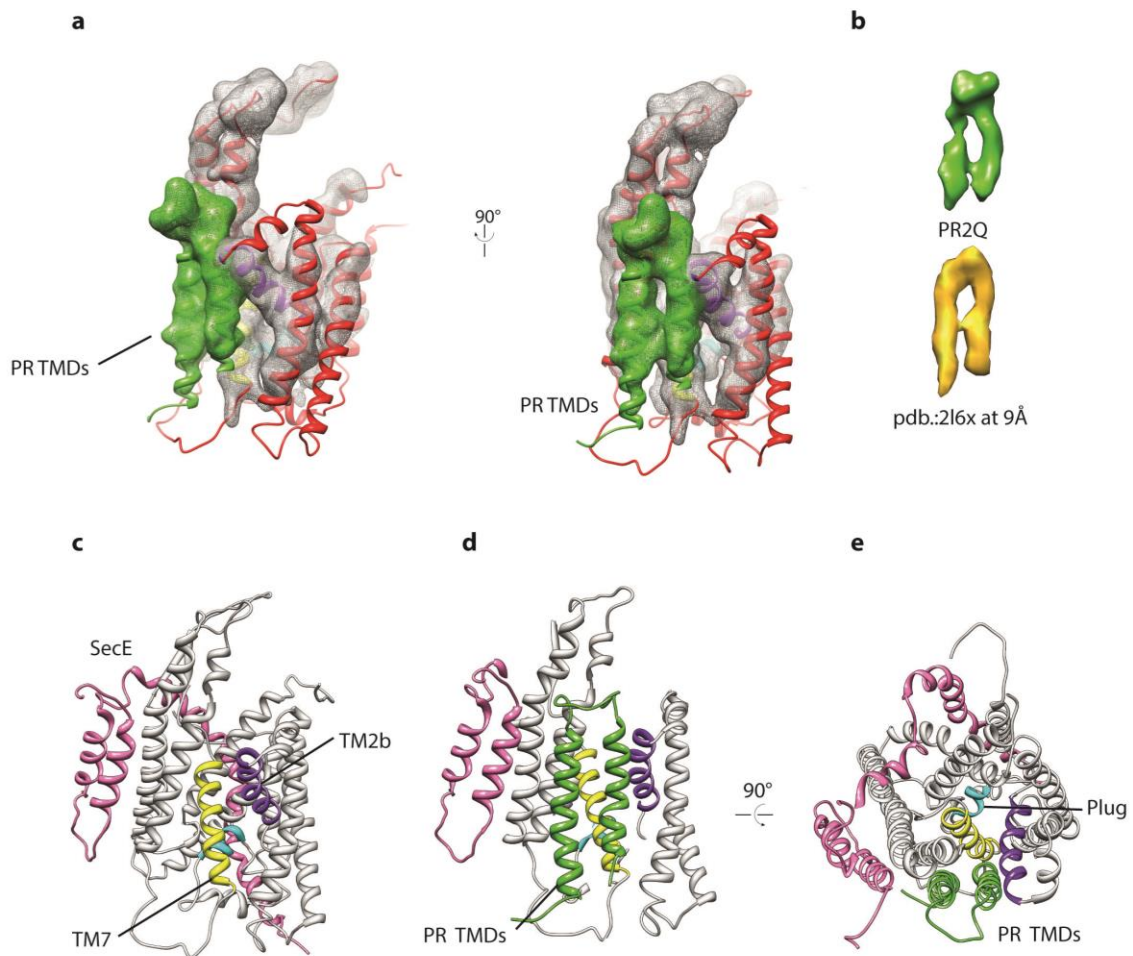
**Figure 26: Visualization of two transmembrane segments of a polytopic membrane protein during SecY-mediated insertion**
a: Positioning of TMS 1 and 2 of PR directly in front of the lateral gate of SecY.
b: Isolated experimental densities for TMSs 1 and 2 of PR (up) reveal high similarity to the corresponding helices in the PR NMR structure filtered to 9Å.
c: Front view on the lateral gate of SecY with the helices of the lateral gate TM7 and TM2b colored in yellow and purple in a conformation similar to the #pre-activated conformation).
d: Front view on SecY in complex with the nascent polytopic membrane protein (green). The nascent chain could not be traced inside the channel and is consequently not part of the model.
e: the same as in (d) seen in a top view from the cytoplasm.
Figure modified from (Bischoff et al., 2014b).

Further analysis of the electron density did not reveal only an interaction of the nascent transmembrane segment back to the lateral gate of SecY, but we also observe a continuous density between ribosomal rRNA helix 59 and the cytoplasmic loop connecting PR transmembrane segment 1 and 2, strongly suggesting a tight contact between this loop and the rRNA helix. Considering the negatively charged phosphate backbone of the rRNA and four positively charged residues within loop1/2 of PR2Q, this could lead to the conclusion

that an ionic interaction between the nascent substrate and rRNA helix H59 exists. The resolution of our map, however, does not allow to unambiguously identify the particular positive residue(s) in loop 1/2 of PR2Q that are responsible for the interaction. Importantly, mutation of two (RAAK) or all four (AAAA) positively charged amino acids to alanine strongly influences the stability of the RNC-SecY complex.
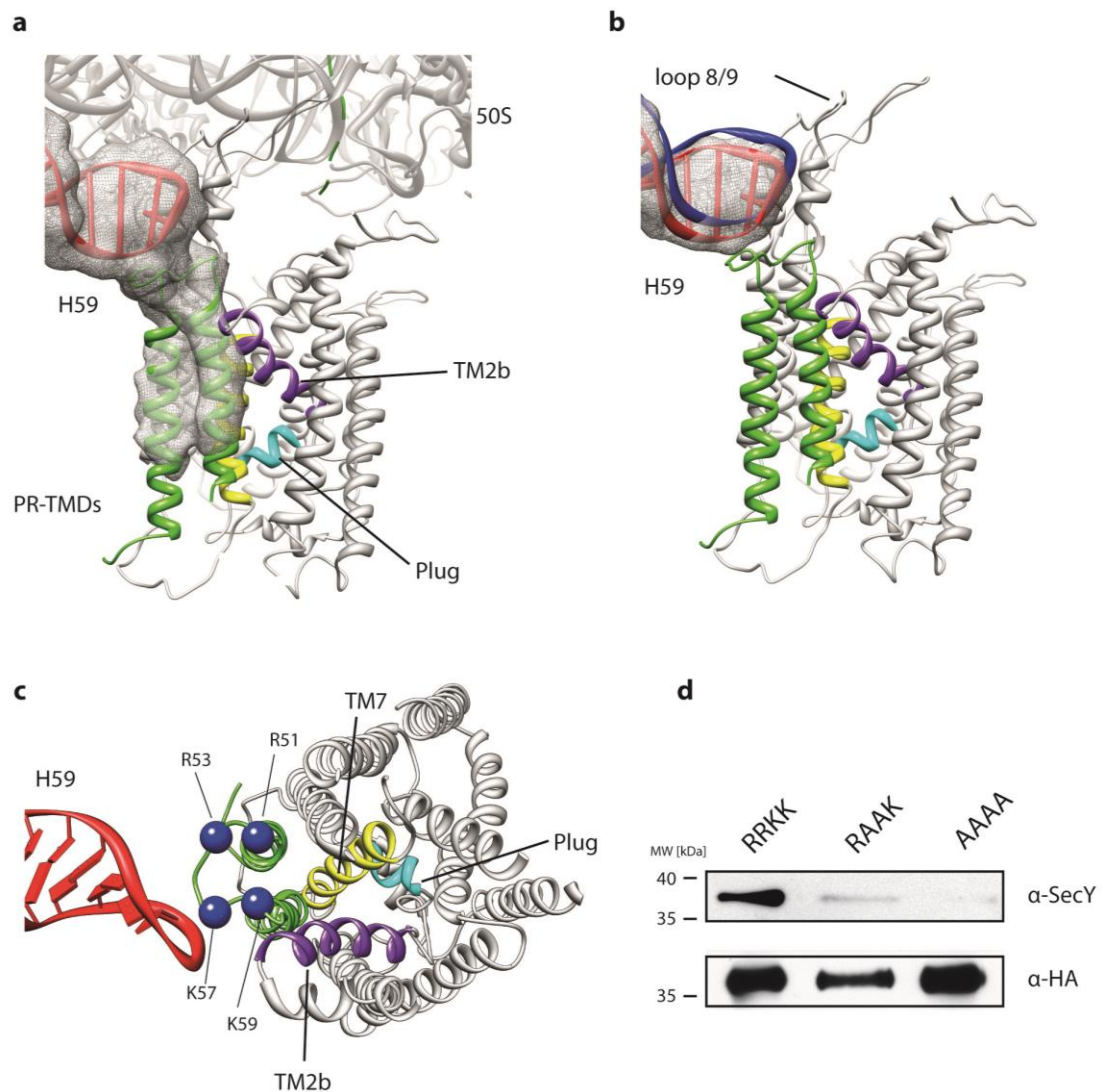


**Figure 27: Interactions between the nascent membrane protein and rRNA helix 59 of the large ribosomal subunit**
a: Continuous electron density between the cytosolic loop of the nascent membrane protein and the rRNA helix 59.
b: The rRNA helix 59 (red) moves down towards the nascent membrane proteins relative to its canonical position (blue).
c: Top view from the cytoplasm with four positively charged amino acids (blue balls) in the cytoplasmic loop 1/2 of PR2Q.
d: Mutation of the positively charged amino acids in the cytosolic loop of PR2Q to alanine converts a tightly coupled RNC-SecY complex to a loosely coupled (RRKK: wt, RAAK: R53A/K57A, AAAA: R51A/R53A/K57A/K59A). Figure modified from (Bischoff et al., 2014b).

Intriguingly, comparing our electron density of the 50S ribosomal subunit to a TnaC-stalled ribosome (Seidelt et al., 2009) does not show any difference apart from the density accounting for rRNA helix H59, which is bend down towards SecY in our structure.

The different location of rRNA helix H59 together with the accumulation of positive charges in the cytosolic loops of membrane protein leads us to speculate that this interaction is of importance for the correct biogenesis of the membrane protein, as it may determine the topology of transmembrane domains.

## 4.2 Site-directed cross-linking

PR2Q mutants $L41_{pBPA}$, $V45_{pBPA}$, $T63_{pBPA}$, $V64_{pBPA}$ and $L67_{pBPA}$ were expressed in *E.coli* together with $SecY_{wt}$, respectively, and purified as described in chapter 3.10. Cross-linking after lysis was followed by concentration of proteins by TCA precipitation and analysis by SDS-PAGE and western blotting against SecY and HA.

Whilst PR2Q mutants $L41_{p\text{-}BPA}$, $V45_{p\text{-}BPA}$, $T63_{p\text{-}BPA}$, and $L67_{p\text{-}BPA}$ did not show any cross-linking products (data not shown), PR2Q $V64_{p\text{-}BPA}$ showed a cross-linking product of approx. 65 kDa in the anti-HA and the anti-SecY blot. Surprisingly, no signal for the P-tRNA in the sample that was not treated with UV could be observed, however, the same amount of RNCs was loaded and the anti SecY signal indicates similar levels of SecY. Much stronger western-blot signals in samples that were exposed to UV light were observed in all tests performed with this cross-linking system (data not shown). Hence, further experiments using another cross-linking system would be necessary for confirmation.
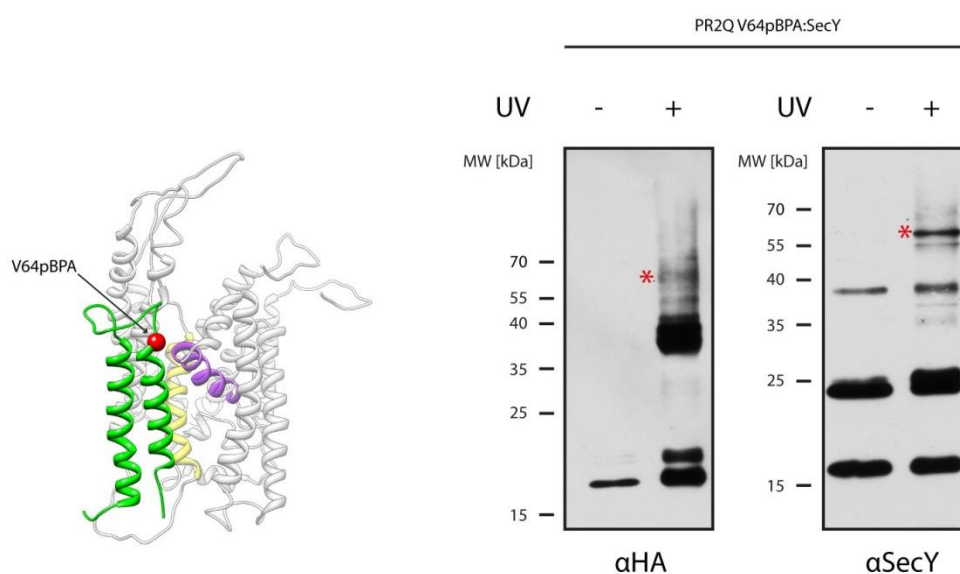


**Figure 28: Crosslinking of PR2Q(V64pBPA) to SecY**
Left: the point mutation in PR2Q (green) is indicated by a red sphere, helix 2b of SecY (light grey) is colored in purple and helix 7 is yellow.
Right: Constructs were expressed in *E.coli* and enriched by affinity purification via the SecY His-tag. The RNC samples were then exposed to UV light for 20 min before TCA precipitation of the proteins. Both blots of the PR2Q $V64_{p\text{-}BPA}$ sample exhibit a UV-dependent signal at ~ 65 kDa marked with an asterisk.

# 4.3 *In vitro* translocation assays to elucidate the translocation mode of PR2Q

Cleavable N-terminal signal sequences are generally found in secretory proteins that are translocated by the SecA-SecY system in bacteria. However, data on the translocation mode of bacterial inner membrane proteins possessing a cleavable N-terminal signal sequence is scarce. Proteorhodopsin that was used as model substrate for the structure of a polytopic inner-membrane protein in complex with SecY is predicted to have such an N-terminal signal sequence.

## 4.3.1 Purification of SecYEG

The protein conducting channel SecYEG was purified from *E.coli* BL21(DE3) cells as described in the methods section. Purification yielded a sample of 500 µL SecYEG in a concentration of 3.75 mg/mL.
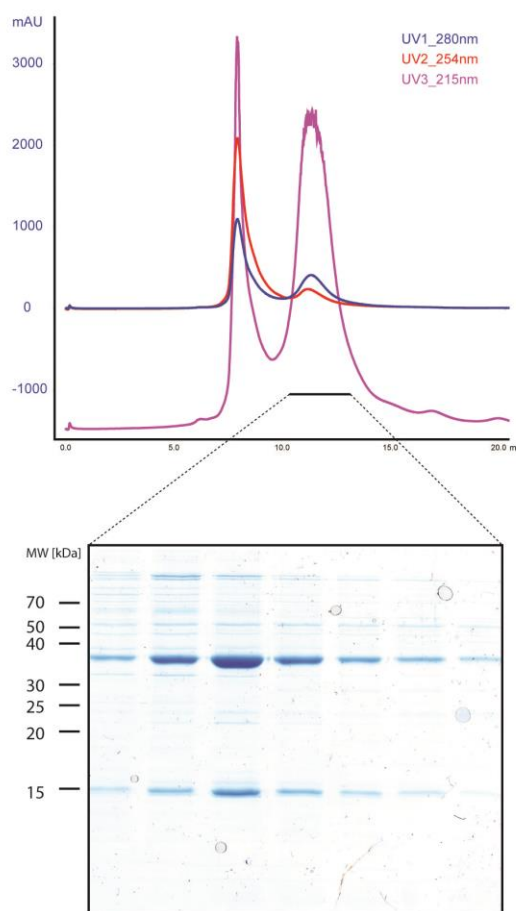


**Figure 29: Gel filtration chromatography of SecYEG**
Chromatogram of the gel filtration chromatography using a Superdex 200 column. The absorptions at 260 nm, 280 nm and 215 nm are shown in red, blue and pink, respectively.
SDS-PAGE analysis (Coomassie staining) of the peak fractions of the gel filtration chromatography. (MW): molecular weight marker [kDa].

## 4.3.2 Purification of SecA

The ATPase SecA was purified from *E.coli* BL21(DE3) cells as described in the methods section. Purification yielded a sample of 500 µL SecA in a concentration of 4 mg/mL.



**Figure 30: Gel filtration chromatography of SecA**
Chromatogram of the gel filtration chromatography using a Superose 6 10/30 column. The absorptions at 260 nm, 280 nm and 215 nm are shown in red, blue and pink, respectively.
SDS-PAGE analysis (Coomassie staining) of the peak fractions of the gel filtration chromatography. (MW): molecular weight marker [kDa].

## 4.3.3 Purification of FtsY

The SRP receptor FtsY was purified from *E.coli* BL21(DE3) cells as described in the methods section. Purification yielded a sample of 500 µL FtsY in a concentration of 37 mg/mL.
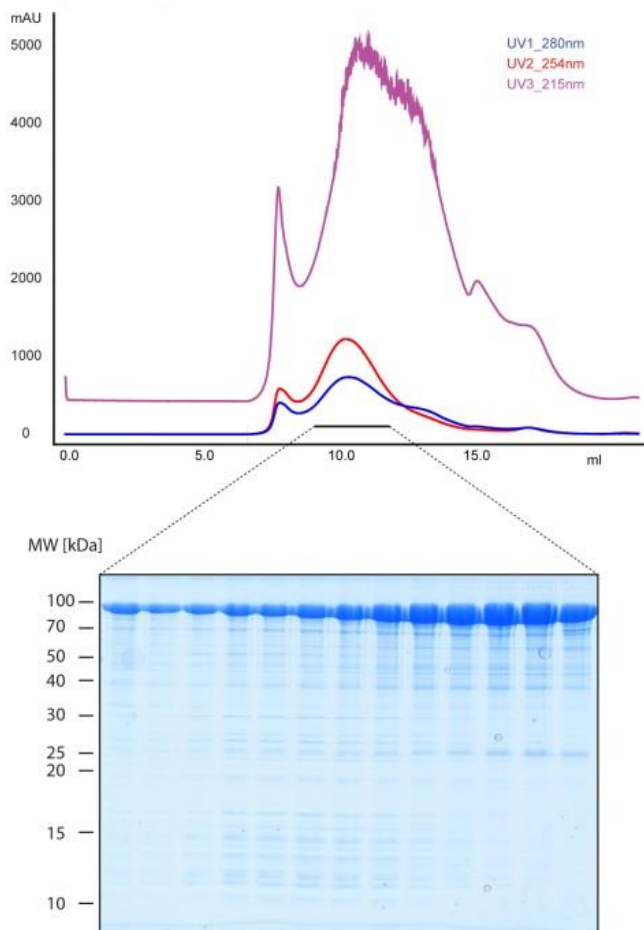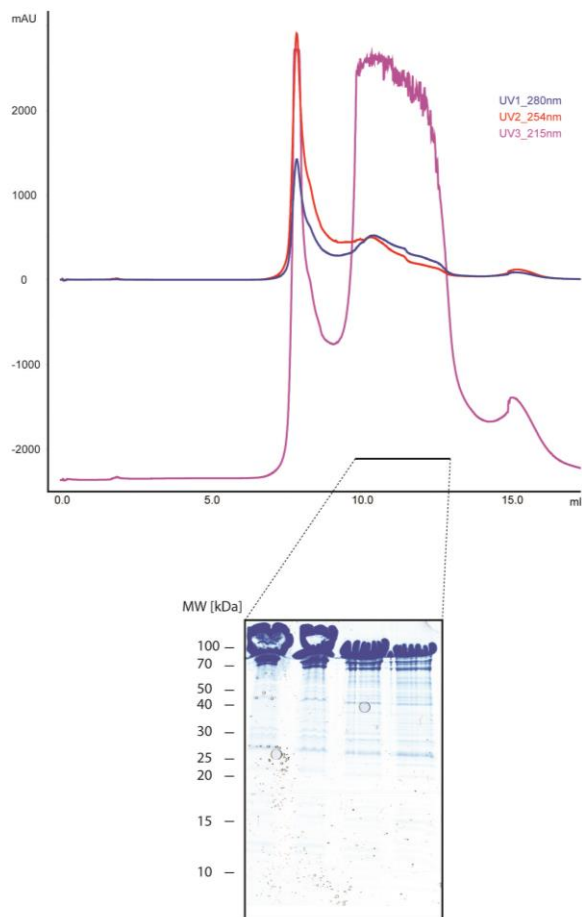


**Figure 31: Gel filtration chromatography of FtsY**
Chromatogram of the gel filtration chromatography using a Superose 6 10/30 column. The absorptions at 260 nm, 280 nm and 215 nm are shown in red, blue and pink, respectively.
SDS-PAGE analysis (Coomassie staining) of the peak fractions of the gel filtration chromatography. (MW): molecular weight marker [kDa].

*E.coli* signal recognition particle and the small chaperone SecB were purified by Eli van der Sluis and Patrick Flagmeier (AG Beckmann, Gene Center Munich, Germany).

### 4.3.4 *In vitro* translocation assays

In order to elucidate the mode by which Proteorhodopsin is targeted to and translocated into the membrane, *in vitro* translocation assays were performed. To this end, PR2Q was translated in a reconstituted *E.coli* translation system (PURE Express, NEB) consisting only of the essential translation factors and an energy recovering system. The translation was carried out in the presence of urea-treated inverted membrane vesicles (U-IMVs) or proteoliposomes reconstituted with SecYEG (PL) in the presence and absences of crucial targeting and translocation factors.

While in the assay containing U-IMVs translocation can be monitored by cleavage of the PR signal sequence by endogenous leader peptidase, translocation into proteoliposomes was tested by protease protection. In case of proper membrane integration, PR is resistant towards digestion by proteinase K. Additionally the proteoliposomes were disrupted by adding the detergent Triton X-100 prior to proteinase K treatment. Samples of the reaction were separated by SDS-PAGE and the products were detected by western blotting against the N-terminal Strep-tag of PR2Q.
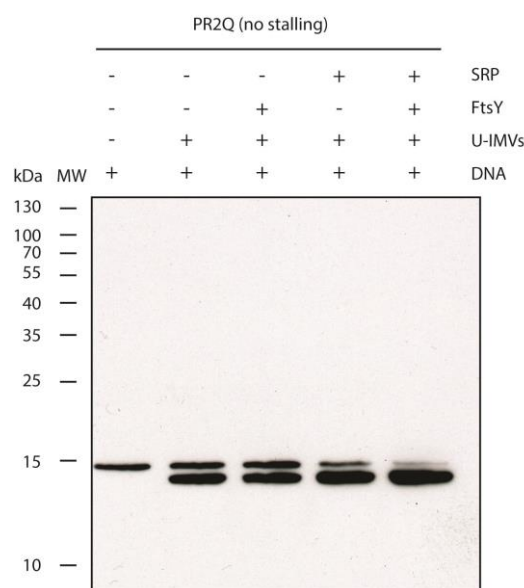


**Figure 32: Translation of PR2Q without TnaC stalling sequence in the presence of urea-treated inverted membrane vesicles**
Samples of the reactions were separated by SDS-PGAE and PR2Q was detected by western blotting against the N-terminal affinity tag.

The translation in the PURE expression system in the absence of any other added factor results, as expected, a single band for the product (lane 1). Translation in the presence of U-IMVs already shows two bands, one for the full-length protein and one with a cleaved signal sequence (lane 2). The addition of either FtsY (lane 3) or SRP (lane 4) results in a

similar pattern of cleaved and full length PR2Q. Interestingly, the addition of SRP and FtsY to the translation system leads exclusively to a processed membrane protein, strongly suggesting that PR is targeted to the membrane by the universal SRP pathway.
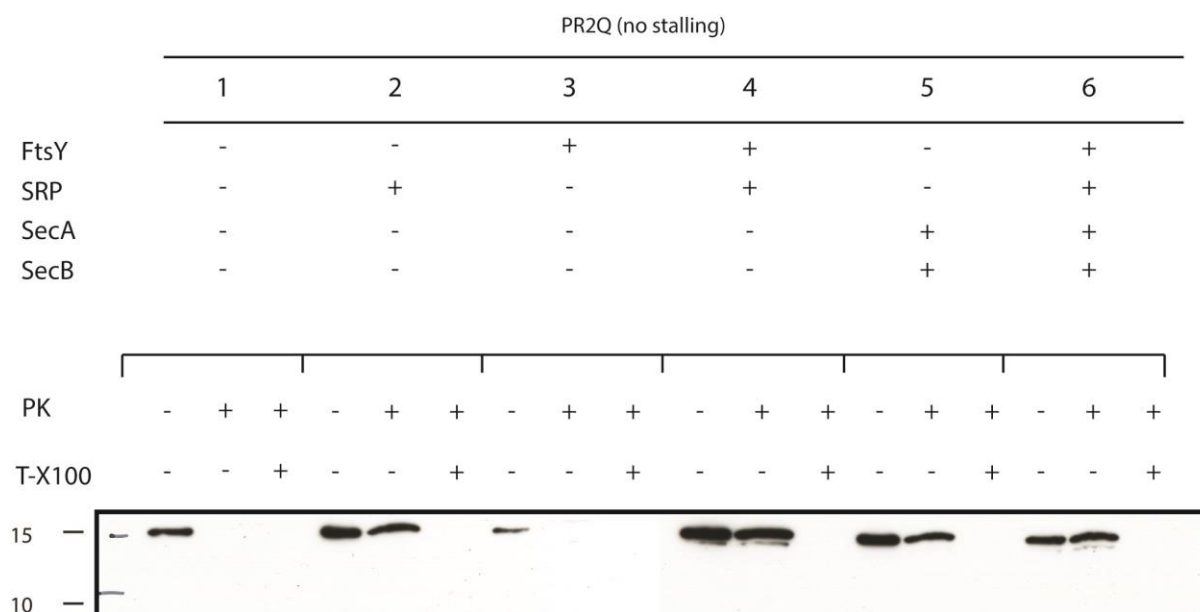


**Figure 33: Translation of PR2Q without TnaC stalling sequence in the presence of proteoliposomes reconstituted with SecYEG**
Samples of the reactions were separated by SDS-PGAE and PR2Q was detected by western blotting against the N-terminal StrepII-affinity tag.

To exclude the influence of possible impurities in the U-IMV preparation the translocation assay was repeated in the presence of proteoliposomes reconstituted with SecYEG. Translocation was monitored by protease protection. Surprisingly, the presence of SRP alone (sample 2) already yielded a significant protease protected protein, while the addition of FtsY alone (sample 3) could not target PR2Q to the lipid bilayer. The addition of both, SRP and FtsY (sample 4) resulted in a strong western blot signal for a protected fragment, indicating proper membrane insertion. Intriguingly, also the addition of SecA and SecB (sample 5) lead to a similar efficient protease protected fragment, as well as the addition of all four targeting and translocation components (sample 6).

Taken together, these translocation experiments suggest that PR2Q can be targeted and integrated into the membrane by the SRP/FtsY pathway, but also by the SecA/SecB system.

# 4.4 Molecular basis for the ribosome functioning as an L-tryptophan sensor

### 4.4.1 Purification of a TnaC-stalled ribosome nascent chain complex

In order to study the molecular mechanisms triggering ribosomal stalling on the TnaC leader peptide, TnaC-stalled RNCs were purified as described above.
FtsQ119-TnaC was expressed in *E.coli* KC6ΔsmpBΔssrA and affinity purification followed by a sucrose density gradient, both in the presence of elevated L-tryptophan levels yielded uniformly stalled ribosome nascent chain complexes.
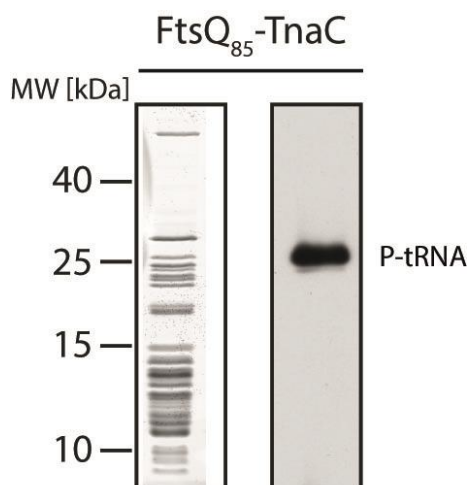


**Figure 34: Evaluation of the *in vivo* assembled ribosome nascent chain complexes**
Coomassie stained SDS-PAGE gel (left) and western blot against the HA tag in the nascent chain reveal a uniformly stalled population of ribosome nascent chain complexes.
Figure modified from (Bischoff et al., 2014a).

### 4.4.2 Cryo-EM structure and molecular model of a TnaC-stalled ribosome

Cryo-EM and single particle reconstruction coupled with *in silico* sorting of the dataset resulted in a subpopulation of 72,468 particles with a stable nascent chain attached to the P-site tRNA that could be refined to 3.8 Å.
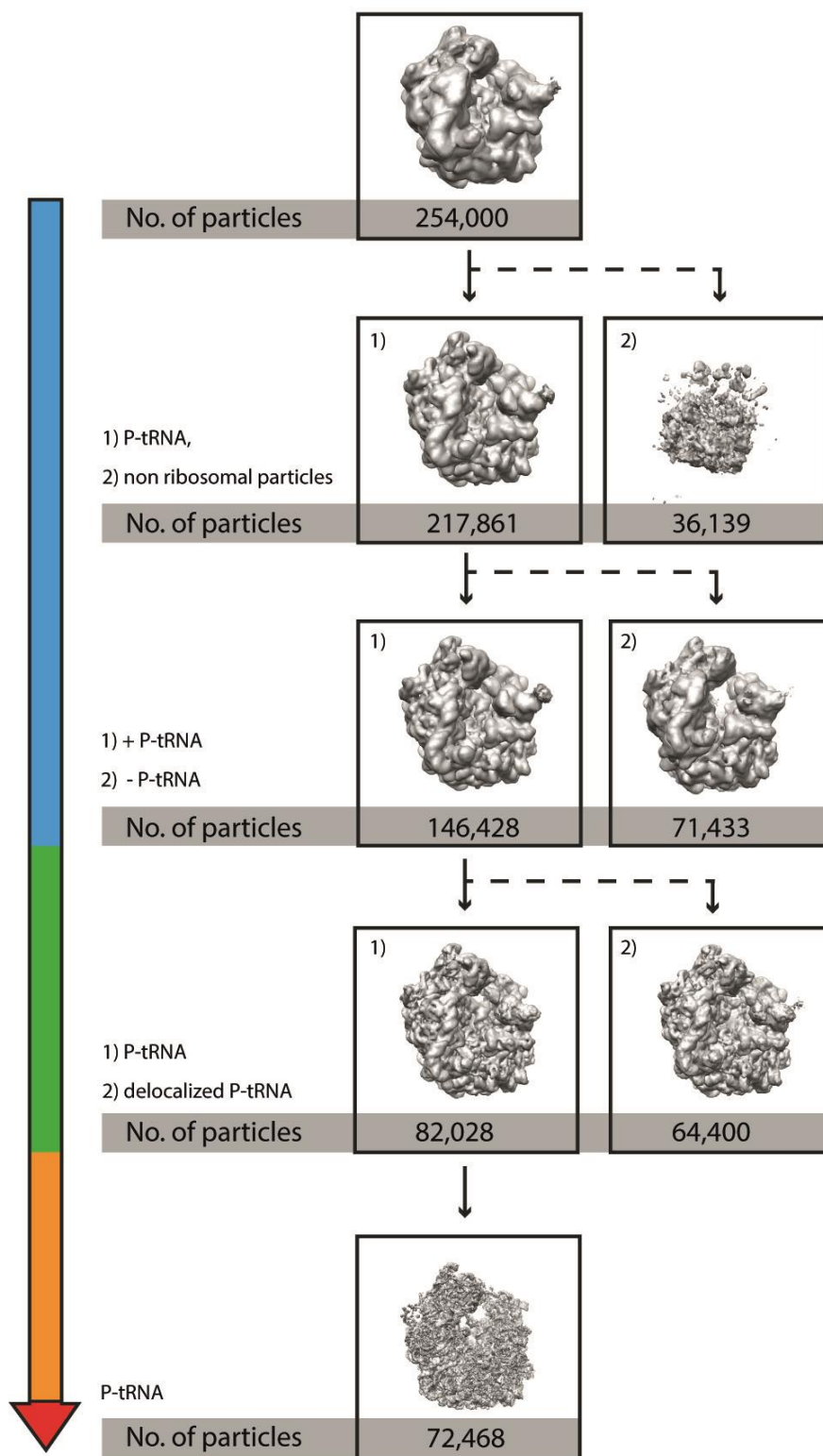
**Figure 35: *In silico* sorting scheme of the TnaC-stalled ribosome dataset**
In a first global sorting step (blue) non-ribosomal particles were sorted out.
In a second global (blue) sorting step ribosome without a P-site could be sorted out.
A third focused sorting step (green) could further sort out particles with a displaced P-site tRNA.
Removal of particle with an extremely low cross-correlation coefficient (orange) leads to a higher resolved volume.
A final sub-dataset of 72,468 particles was refined to 3.8Å (red).

To avoid the major risk of over-fitting the data and thus the accumulation of noise in the volume that could be misinterpreted as high-resolution features, the data was permanently low-pass filtered at 8 Å during the whole sorting and refinement process. Additionally, a second refinement approach following the 'gold standard' protocol was performed (Scheres and Chen, 2012). To this end, the dataset was split in two independent half datasets, that were refined individually and independent and the 'gold standard' resolution was estimated by a FSC of the two half volumes. As expected, both refinement approaches lead to basically the same resolution (see Figure 36). Calculation of the local resolution, however, revealed that the core of the ribosome is significantly better resolved than the average resolution of the ribosome.
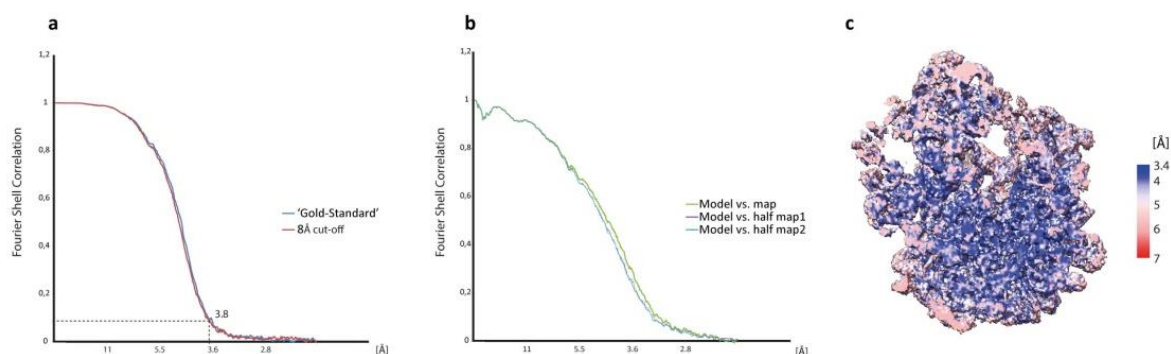


**Figure 36: Global and local resolution of the TnaC-stalled ribosome nascent chain complex**
a: FSC-curves derived following the 'gold standard' (blue) and the 8 Å cut-off (red) protocol.
b: FSC-curves calculated from the model vs. the experimental cryo-EM map (green) and the 'gold-standard' half volumes (purple, blue).
c: calculation of the local resolution reveals that the core of the RNC is resolved to less than 4 Å.
Figure modified from (Bischoff et al., 2014a).

Rigid body docking of an X-ray crystallography base model of the *E.coli* ribosome (Dunkle et al., 2010) revealed excellent agreement with our experimental electron density.

Consistent with the estimated resolution we find distinct features like bulky side-chains, single rRNA bases, separated beta strands and metal ions in the ribosome unambiguously resolved.
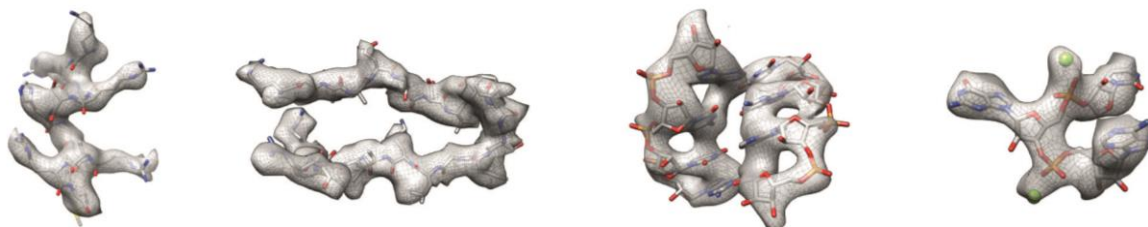


**Figure 37: Representative snap-shots of the experimental density**
From left: α helix of a ribosomal protein, beta turn of a ribosomal protein, rRNA helix, bulged out single rRNA bases and a $Mg^{2+}$ ion.
Figure modified from (Bischoff et al., 2014a).

Moreover, we find the TnaC leader peptide and its contacts with distinct ribosomal rRNA bases and amino acids of ribosomal proteins extremely well resolved in the upper part of the ribosomal exit tunnel until the central constriction formed by the ribosomal proteins uL22 and uL4. The continuous and well-resolved density for the nascent peptide enabled to build a *de novo* model for the 12 C-terminal amino acids of the TnaC leader peptide that are critical for stalling. Additionally, all rRNA bases at and close to the wall of the ribosomal exit tunnel were inspected and their position was adjusted according to the electron density.

Early biochemical assays suggested that the binding pocket for the free L-tryptophan molecules might be in or close to the A-site of the ribosome. Precisely, it has been proposed to overlap with the binding site of the antibiotic sparcomycin. However, that has been contradicted in a recent study and is also clearly not supported by our cryo-EM reconstruction.

Remarkably, two unassigned extra densities were identified in the upper half from the PTC to the central constriction. These extra densities were interpreted as two free L-tryptophan molecules bound in composite binding sites between the nascent TnaC peptide and components of the ribosomal exit tunnel in about 15 Å to 20 Å distance from the PTC.
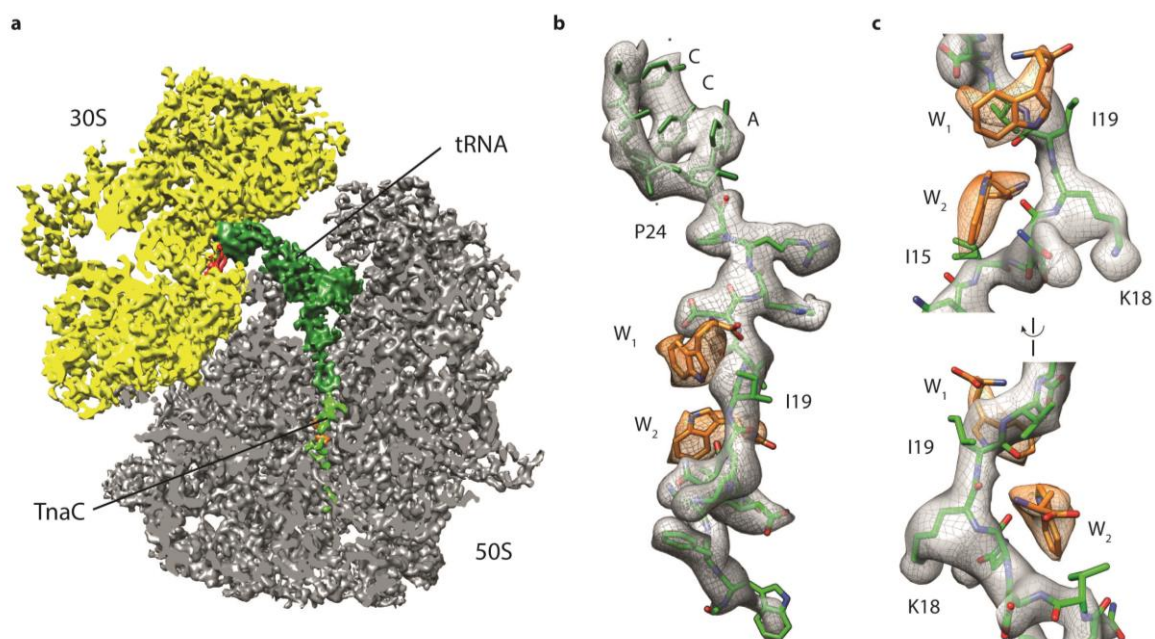


**Figure 38: Cryo-EM structure of a TnaC-stalled ribosome nascent chain complex**
a: cross-section through the cryo-EM density of the TnaC-stalled ribosome nascent chain complex with the small ribosomal subunit (30S) in yellow, the large ribosomal subunit (50S) in grey, the P-site tRNA in green and the nascent chain in light green. The mRNA anticodon in colored in red and the two free L-tryptophan molecules in orange.
b: Density with the *de novo* built model for the CCA-end of the tRNA, residues 24-12 of the TnaC peptide and the free L-tryptophan molecules.
c: as in (b) but close-up on the two free L-tryptophan molecules. Figure modified from (Bischoff et al., 2014a).

## 4.4.3 Interactions of the TnaC leader peptide with components of the ribosomal tunnel

The two free L-tryptophan molecules are bound in two hydrophobic pockets formed by the nascent chain and rRNA bases of the 23S rRNA. The upper L-tryptophan ($W_1$) is bound in a hydrophobic cavity formed by V20 and I19 of TnaC and on the other side engaged in a stacking interaction with rRNA base U2586. Additionally, $W_1$ is stabilized by interactions with D21 and the peptide backbone of TnaC. Interestingly, the nascent chain forms a remarkable kink in the area of N17 and K18, stabilized by interactions of the latter amino acids with the rRNA base-pair U2609 and A752. That kink forms a second hydrophobic pocket, in which the second free L-tryptophan molecule is bound.
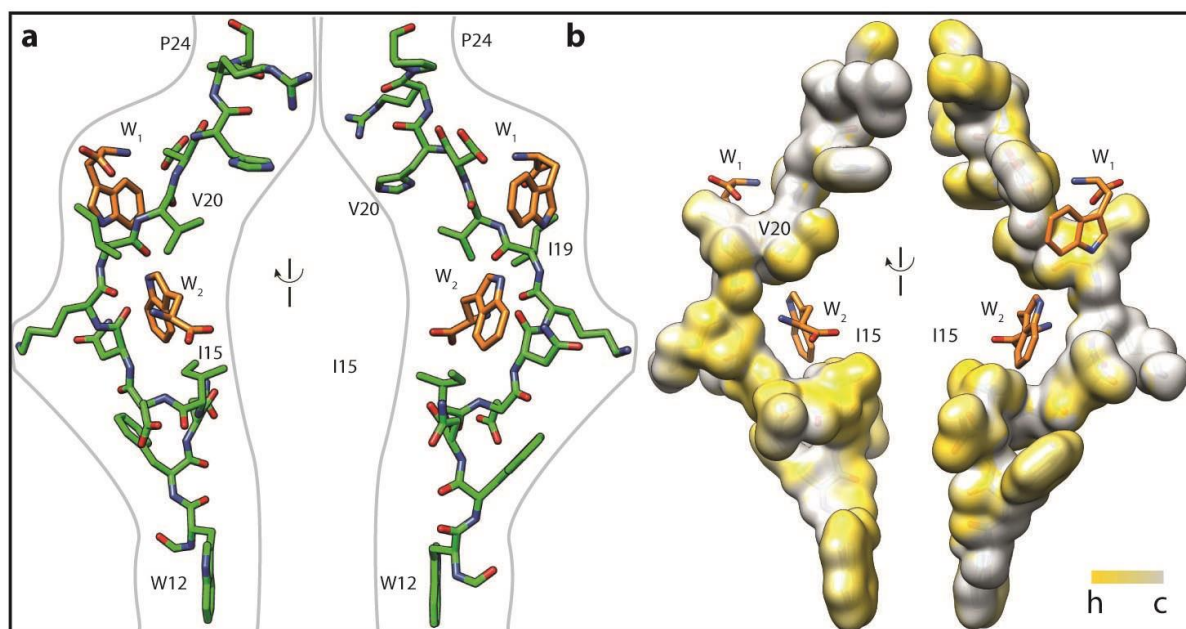


**Figure 39: The two free L-tryptophan molecules in the ribosomal exit tunnel**
a: Schematic of the TnaC leader peptide with the two L-tryptophan molecules in the ribosomal exit tunnel.
b: Surface hydrophobicity plot of the nascent chain (h=hydrophobic, c=charged) reveals two hydrophobic binding pockets for L-tryptophan formed by V20 and I19 as well as I19 and I15 of TnaC.
Figure modified from (Bischoff et al., 2014a).

$W_2$ is additionally interacting with the rRNA bases A2058 and A2059 that form a crevice in the ribosomal exit tunnel, which is also the binding platform for macrolide antibiotics.
Noteworthy, mutations of U2609 and A752 as well as A2058 have been shown to severely affect the stalling efficiency of TnaC. Moreover, mutating the residues I19 and I15 to less hydrophobic or charged side chains also affects the stalling on the TnaC peptide. While distinct density for all side-chains in the TnaC peptide can be assigned, I19 and I15 lack a clear electron density.
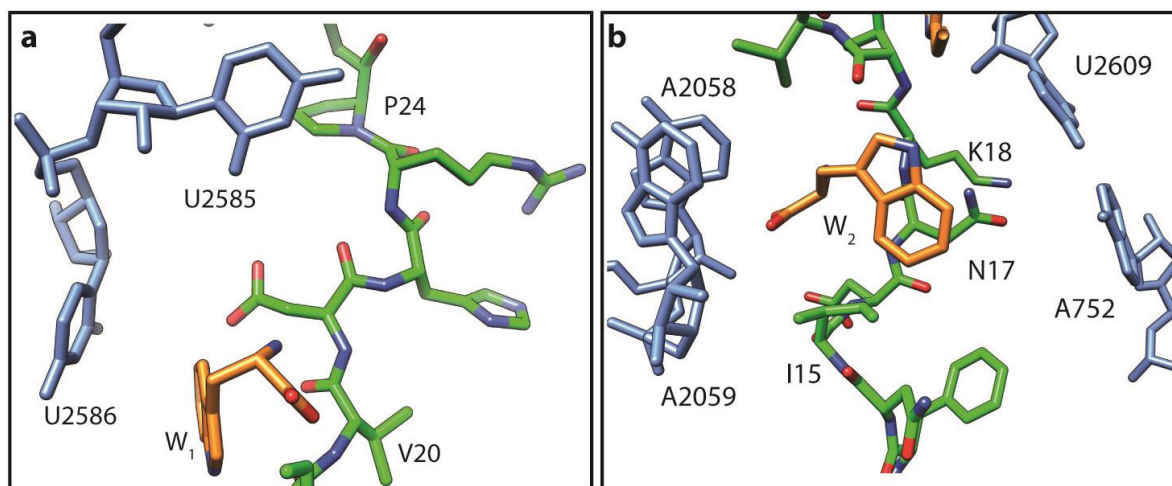
**Figure 40: Interactions of the free L-tryptophan with TnaC and the ribosomal exit tunnel**
a: $W_1$ interacts with V20 and I19 of TnaC and U2596 of the 23S rRNA.
b: $W_2$ interacts with I19 and I15 of TnaC as well as the rRNA nucleotides A2058 and A2059.
Figure modified from (Bischoff et al., 2014a).

Further down in the ribosomal exit tunnel we find additional interactions between the nascent chain and the ribosomal proteins of the central constriction uL22 and uL4. The invariant TnaC residues D16 and W12 engage with the residues K90 and R92 of uL22. This motif seems of particular importance for TnaC induced stalling, since mutations of K90 in uL22 have been shown to reduce the efficiency of TnaC stalling. Additionally, the distance of D16 and W12 from the PTC is also critical for inducing stalling on the TnaC leader peptide. The 'zipper' is furthermore stabilized by an interaction between N14 of TnaC and R61 of the ribosomal protein uL4. The importance of this interaction is supported by mutational data.
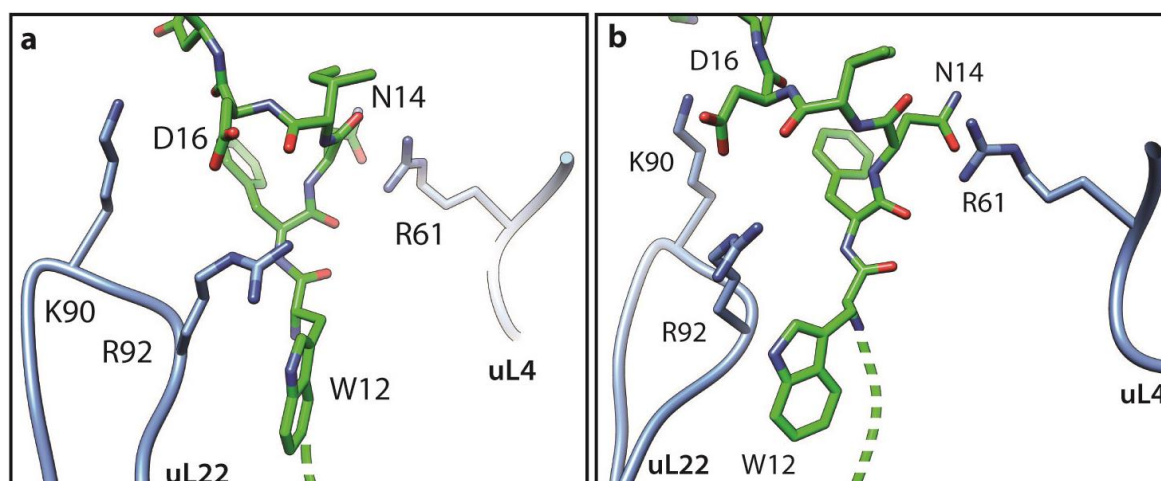


**Figure 41: Interactions of TnaC with the ribosomal proteins of the central constriction**
a: The invariant TnaC residues D16 and W12 engage in a 'zipper' -ike interaction with the residues K90 and R92 of uL22.
b: N14 of TnaC interacts with the residue R62 of uL4. Figure modified from (Bischoff et al., 2014a).

The highly conserved residue P24 of TnaC shows an interaction with the important PTC nucleotide U2585 and most likely contributes to the TnaC stalling by its poor reactivity in peptide bond formation.

### 4.4.4 Silencing of the PTC

The peptide bond formation in the ribosome as well as peptide release by RF2, strictly requires precise positioning of the P- and A-site tRNA, the GGQ motive of the release factor and importantly distinct conformations of highly conserved rRNA nucleotides in the PTC.

We find $W_1$ in a tight stacking interaction with the rRNA base 2586, that might reduce the flexibility of the neighboring U2585 that on its side shows an interaction with P24 of TnaC in the P-site of the ribosome. Furthermore the interactions of the TnaC peptide and $W_2$ in the area of A2058, A2059 and the base pair U2609/A752 result in a shift of U2609 and G2608 towards the nascent chain. This shift enables a new interaction between G2608 and the phosphate backbone between U2586 and U2585 that additionally stabilizes U2585 in its particular conformation.
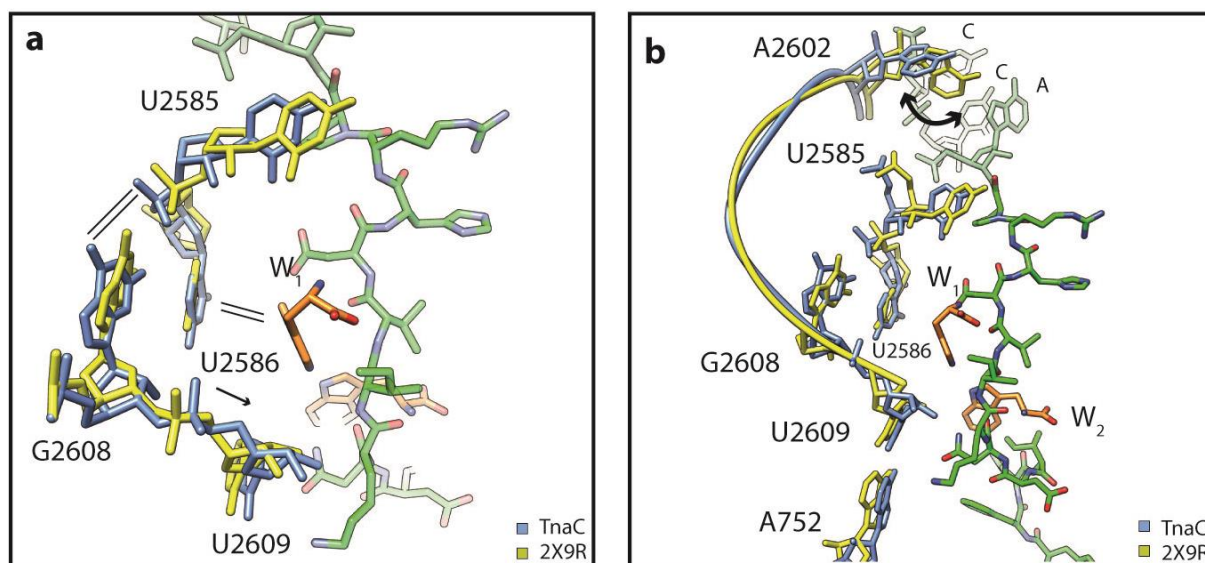


**Figure 42: Relay to PTC and schematic model of the TnaC-stalled ribosome**
A: The interaction of $W_1$ with 2586 and the interaction of K18 and N17 of TnaC with U2609 lead to the formation of a new interaction between G2608 and the phosphate connecting U2586 and U2585 eventually stabilizing U2585 in the observed conformation. Comparison with (Jin et al., 2010).
B: The interactions shown in (A) might decrease the flexibility of the rRNA stretch up to the PTC and could contribute to the stabilization of A2602. Comparison with (Jin et al., 2010).
Figure modified from (Bischoff et al., 2014a).

The observed stabilized conformation of U2585 eventually disables the GGQ motive of RF2 to correctly deliver the water molecule necessary to release the stalled peptide.

Moreover, our structure clearly shows A2602 in a stable conformation that resembles the conformation when the translation inhibitor sparcomycin is bound. It differs severely when compared to the crystal structure of RF2 bound to the ribosome and is inconsistent with the accommodation of RF2 for peptide release (see Figure 43d,e).
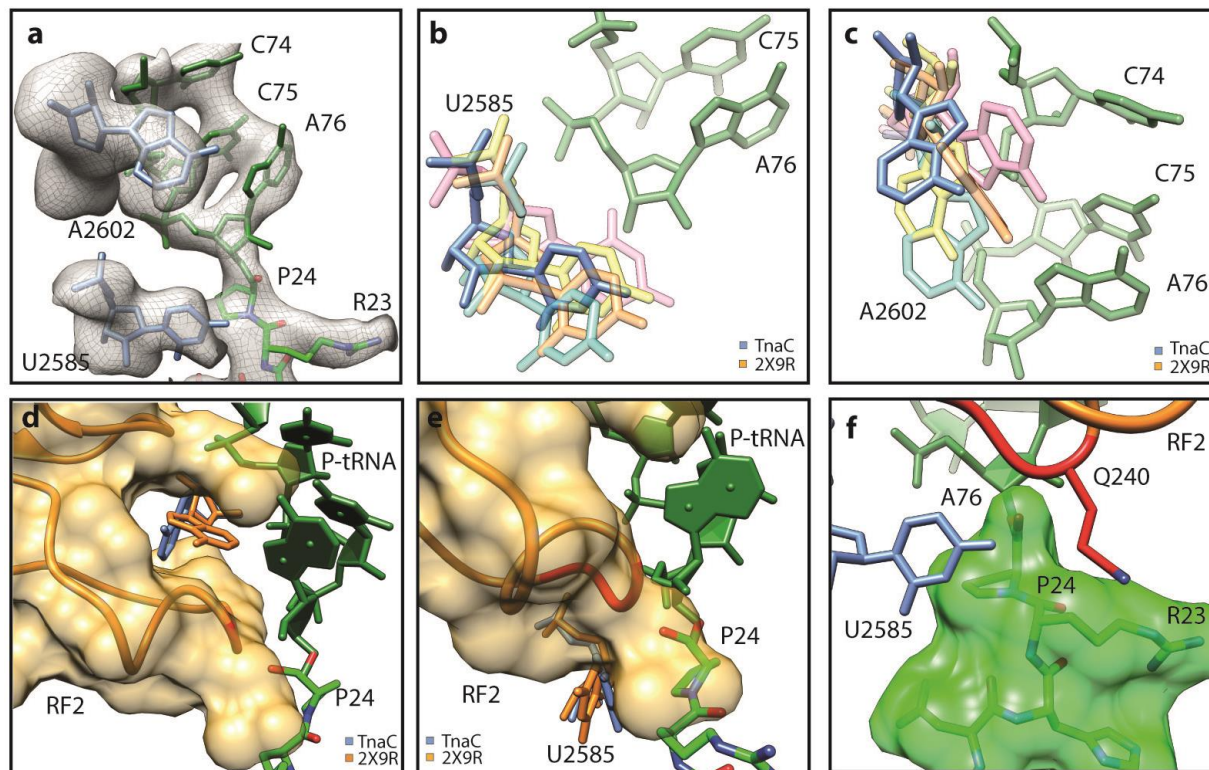


**Figure 43: Silencing of the PTC in the TnaC-stalled ribosome**
a: The cryo-EM density for the 23S rRNA nucleotides U2585 and A2602 adopting distinct conformations in the PTC of the TnaC-stalled ribosome.
b: Conformation of 23S nucleotide U2585 in TnaC (blue) in comparison to a ribosome bound to the antibiotic chloramphenicol ((Dunkle et al., 2010) pdb 3OFC, yellow), the uninduced ((Schmeing et al., 2005b)pdb 1VQ6, pink) and the induced ((Schmeing et al., 2005b)pdb 1VQN, light blue) state of the PTC.
c: Conformation of 23S rRNA nucleotide A2602 in TnaC (blue) in comparison to a ribosome bound to the antibiotic chloramphenicol (pdb 3OFC, orange), the uninduced (pdb 1VQ6, pink) and the induced (pdb 1VQN, light blue) state of the PTC .
d: The conformation of A2602 in TnaC-SRC is inconsistent with the accommodation of release factor 2 (space filled model, orange, GGQ motive colored in red). The rotation of TnaC A2602 in comparison to A2602 in the 70S-RF2 complex (pdb 2X9R, orange) leads to a clash of the nucleotide with RF2.
e: U2585 in the conformation observed in TnaC-RNC would clash with RF2 (orange space filled, GGQ motive colored in red).
f: The contact of U2585 to P24 of TnaC leads to a stabilization of R23 of TnaC (green space filled) in a position that would clash with the correct accommodation of the GGQ motive of RF2.
Figure modified from (Bischoff et al., 2014a).

Furthermore, R23 of TnaC is stably positioned between the bases U2506 and C2452 in the PTC and would clash with Q240 of RF2 thus hampering the correct placement of the GGQ motive necessary for peptide release. The stable position of R23 may also explain the poor reactivity of TnaC towards puromycin through a sterical clash.
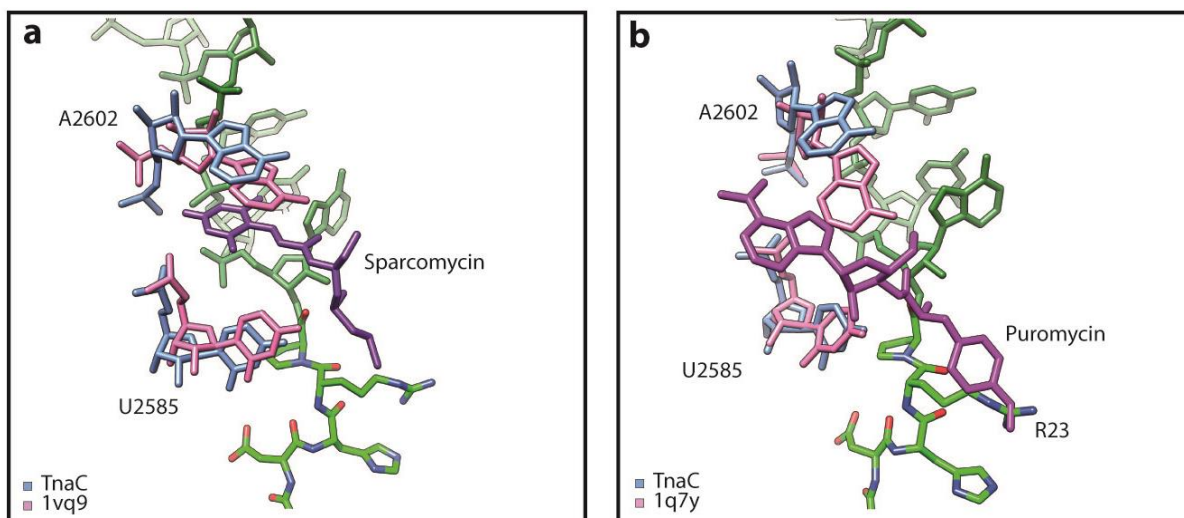
**Figure 44: Comparison of the PTC in the TnaC-stalled ribosomes to the PTC of ribosomes with bound antibiotics**

a: The conformation of A2602 and U2585 of the TnaC-stalled ribosomes are very similar to the structure of sparcomycin bound in the PTC (Dunkle et al., 2010).

b: The poor reactivity of TnaC toward puromycin can be explained by a sterical clash between the drug and R23 of TnaC preventing the correct accommodation of puromycin (Hansen et al., 2002).

Figure modified from (Bischoff et al., 2014a).

Comparing the conformation of U2585 and A2602 in the PTC of our structure with the crystal structure of a ribosome with bound A- and P-site before peptide bond formation reveals that the accommodation of the A-site tRNA and the incorporation of an additional amino acid is not hampered by the conformation of U2585 and A2602 in the PTC of the TnaC-stalled ribosome (data not shown).

# 5. Discussion

## 5.1 Purification of *in vivo* assembled ribosome nascent chain complexes

The protocol to purify ribosome nascent chain complexes from the living cell presented in this thesis offers great new possibilities in studying co-translational events.

The advantages of *in vivo* assembled RNCs compared to *in vitro* translation and purification are manifold: The RNCs are generated under entirely physiological conditions in the living cell. The purification procedure is faster and cheaper compared to *in vitro* systems and can optionally be scaled up to gain large quantities. Additionally, isolation of the RNCs under mild conditions allows the purification of co-factors that can be further used in the unbiased investigation of physiologically relevant complexes.

Although others have reported protocols to isolate *in vivo* assembled RNCs (Jha and Komar, 2012) in earlier publications, the yields and quality that can be achieved by the protocol presented here are still unprecedented.

While most reported protocols for the isolation of RNCs are based on the well-known SecM stalling sequence, we used the TnaC stalling sequence, whose molecular mechanism is also presented in this study. The great advantage of TnaC over SecM is that TnaC induces inhibition of peptide release, while SecM only induces an extreme slowdown of the incorporation of the A-site amino acid. Hence, the SecM stalling gets released over the time of the purification resulting in an inhomogeneous sample. Additionally, SecM is *in vivo* released by a pulling force on the nascent chain created during membrane insertion (Ito and Chiba, 2013), thus it is tempting to speculate that also forces built by the liquid flow during affinity purification may cause the nascent chain release and lead to the low yields of SecM-stalled ribosome nascent chain complexes.

In contrast, however, TnaC is not sensitive to force on the nascent peptide but only to the concentration of free tryptophan. Consequently, a high tryptophan concentration during the whole purification procedure leads to highly homogenous RNCs. Furthermore, the usage of a specialized *E.coli* strain is additionally improving the yield and quality of the RNCs. This *E.coli* strain, KC6 ΔssrAΔsmpB, has a knockout of the enzyme tryptophanase, hence having an unnaturally high tryptophan concentration in the cytoplasm.

Since TnaC stalling is dependent on elevated levels of free tryptophan, this *E.coli* strain is especially suitable for the production of stable RNCs. Additionally, a knockout in the tmRNA ribosome rescue system enhances the yield of stalled ribosomes.

## 5.2 Visualization of a polytopic membrane protein during SecY-mediated membrane insertion

### 5.2.1 A dual binding mode for SecY to the ribosome

The results presented here show that the initial attempt to visualize a polytopic membrane protein during SecY mediated membrane insertion by expressing a simple polytopic membrane protein, where SecY gets trapped on the nascent peptide did not lead to a stable complex. It could be shown by western blotting and mass-spectrometry that the ribosome expressing a polytopic membrane protein, like proteorhodopsin that mainly consists of hydrophobic transmembrane segments only separated by short loop, cannot engage in a stable complex with SecY. Since SecY obviously evolved to efficiently integrate transmembrane segments into the membrane it was in retrospection not too surprising that the simple expression of a polytopic membrane protein - although different lengths were testes - followed by a purification over several days does not result in a stable RNC-SecY complex.

However, expressing the well-known type II membrane protein FtsQ in two different lengths leads to the formation of a stable RNC-SecY complex. The difference here is that SecY encounters a different pattern of hydrophobic and hydrophilic stretches. Assuming that a hydrophobic transmembrane segment as present in the PR constructs rapidly leaves SecY to partition into the membrane it is possible that a hydrophilic stretch following the hydrophobic segment as in $FtsQ_{119}$ might lead to a tightly coupled PR-RNC-SecY complex.

Hence, we designed hybrid membrane protein constructs that consist of the polytopic membrane protein PR on the N-terminus followed by the hydrophilic loop of FtsQ. Interestingly, the hybrid constructs engaged in an equally stable complex with SecY like the $FtsQ_{119}$.

Since the endogenous *E.coli* protein CyoA that has a similar distribution of hydrophobic transmembrane segments and hydrophilic stretches also forms a stable complex with SecY, it can be excluded that the hybrid proteins PRxQ are forced into an off-pathway that artificially induces the tight binding of the RNC to SecY.

These results together with the controls performed (see Figure 18) suggests that the interaction of SecY with the translating ribosome can alter between a stable (tight) and a less stable (loose) coupling, dependent on the nascent chain that is translated and translocated in that moment. Taken together tight coupling of SecY to the translation ribosome *in vivo* is only established during co-translational translocation, when inserting a transmembrane segment with type II topology that is followed by a hydrophilic loop.

This bimodal binding mode of SecY to the translating ribosome in *E.coli* is in stark contrast to the ribosome binding to Sec61 in eukaryotic cells. Here, the ribosome binds Sec61 with very high affinity and stays bound during the whole membrane protein synthesis (Shao and Hegde, 2011; Pfeffer et al., 2012; Dudek et al., 2014; Pfeffer et al., 2014). Moreover, it was suggested that the large subunit of the eukaryotic ribosome stays bound to the membrane even during and after ribosome recycling. A similar behavior in *E.coli* can be excluded from the data presented here.

These findings raise now the question, why SecY exhibits this dual binding mode to the ribosome in bacteria. One major difference in the translocation of (membrane) proteins in *E.coli* and eukaryotic cells is the presence of the cytosolic ATPase SecA that is responsible for the post-translational translocation of periplasmic and outer-membrane proteins, but also for the translocation of large periplasmic loops and domains of inner membrane proteins in bacteria (Chatzi et al., 2014). Since the ribosome and SecA bind to the same region of SecY it can be excluded that they can bind simultaneously (Huber et al., 2011; Kuhn et al., 2011; Wu et al., 2012). Therefore, the ribosome has to dissociate from SecY at the point, where SecA is needed to translocate larger periplasmic domains over the membrane. This is further supported by the findings presented in Figure 19 showing that the coupling of SecY to the ribosome switches from tight to loose upon elongation of the hydrophilic stretch succeeding the last transmembrane segment of PR2Q. Whether SecA plays an active role in dissociating the tightly coupled RNC-SecY complex or whether the elongation of the hydrophilic loop leads to a lower affinity of SecY to the ribosome still needs to be elucidated.

Based on these results, the following model (see Figure 45) can be put forward: (1) the ribosome is tightly coupled to SecY and in the beginning of synthesizing a hydrophilic loop following a hydrophobic transmembrane segment. (2) Upon elongation of this loop the coupling of SecY to the ribosome gets more and more loosely until it finally dissociates. This dissociation might be actively supported by SecA. (3) After final dissociation of the ribosome SecA can bind to SecY and translocate the hydrophilic loop over the membrane. (4) After successful translocation SecA dissociated from SecY.
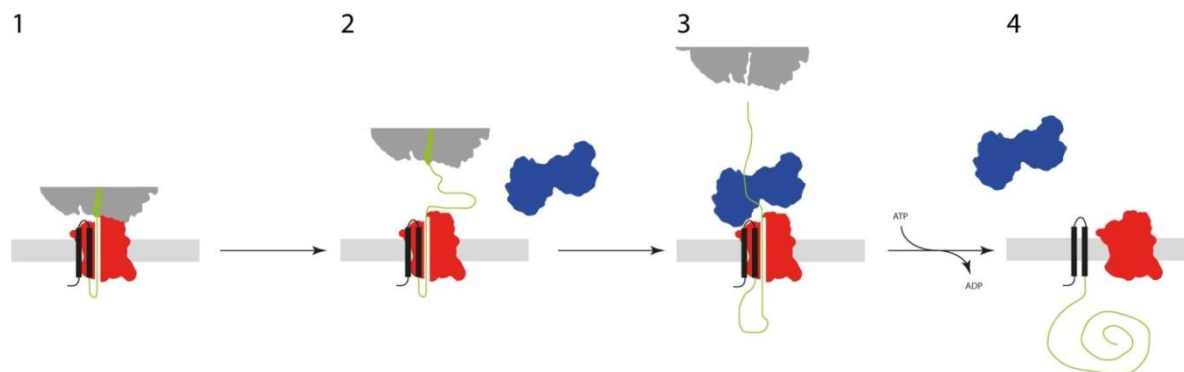
**Figure 45: Schematic model of the transition from tight to loose coupling of the translating ribosome to SecY**

On interesting aspect that still needs to be investigated is however, when the hydrophilic loop that is translocated by SecA is followed by a new transmembrane segment. In this case a new round of SRP binding and re-targeting the whole translating ribosome back to the membrane where it engages again in a tight coupling to the translocon would be required. This is for example the case in the biogenesis of YidC, where five more transmembrane domains are following the large periplasmic P1 domain whose translocation is SecA-dependent.

## 5.2.2 Cryo-EM structure of a tightly coupled RNC-SecY complex

One general concern in all structural studies of membrane proteins is the use of detergent to solubilize the protein out of the membrane and to keep it in solution afterwards and to what extent the isolated membrane protein is then still of physiological relevance.
However, we argue that the use of detergent is not an argument to question the physiological relevance of the structure presented here. As a major point, the complex presented here was assembled under totally physiological conditions in intact membranes and detergent was only applied after translation and membrane insertion. Solubilization of assembled complexes out of the membrane is common practice in structural studies on membrane proteins.

After deciphering conditions, in which a tightly coupled RNC-SecY complex is formed, we set out to purify the latter in quantitative amounts for cryo-EM and single particle reconstruction. The substrate PR2Q was chosen as it is the shortest and thus the most stable of the PRxQ constructs. Single particle reconstruction and *in silico* sorting of the dataset (Figure 21) resulted in a final 3D volume at an average resolution of 7.3 Å as described in the

methods section. The structure revealed as expected additional density at the ribosomal exit tunnel, that could be explained by a single copy of SecY bound to the ribosome. The data clearly excluded a possible oligomer of SecY as it was also put forward in the most recent publications. Since there is no crystal structure of the *E.coli* SecY to date, a homology model was built as a starting model for the interpretation of the additional density on the ribosome.

As a template for the homology model, the crystal structure of *Aquifex aeolicus* SecY from a SecY-SecA (Zimmer et al., 2008) complex was chosen since sequence alignment (Soding et al., 2005; Hildebrand et al., 2009) revealed the highest sequence similarity (86.8 %) to *E.coli* SecY. Thus it is the most reliable structure to serve as a template. The starting model of SecE was taken from a former cryo-EM-model of *E.coli* SecY because none of the existing crystal structures contains a molecular model of three SecE helices as they are present in *E.coli*.

The homology model of SecY was then fitted in the EM density as rigid body based on the position of the loops 8/9 and 6/7 that are the universal binding anchors on SecY for ribosomal binding and that are very well conserved from bacteria to eukaryotes (Menetret et al., 2007; Becker et al., 2009; Frauenfeld et al., 2011; Gogala et al., 2014). As previously reported, the C-terminal part of SecY which binds the ribosome is rather rigid compared to the more flexible N-terminus. In contrast to the C-terminal part of SecY that could be fitted as a rigid body with only miniscule changes, the N-terminal helices were remodeled manually to fit the experimental density best.

Additionally, SecE was modelled into the cryo-EM density manually in agreement with 2D crystal structures of SecYE (Breyton et al., 2002; Hizlan et al., 2012) and a previously published cryo-EM structure of SecYE in the membrane environment (Frauenfeld et al., 2011). Although the putative position of SecG is known, we could not assign any electron density to SecG. This, however, should not be of major concern for several reasons: (i) SecG is known to be loosely bound to SecYE and is missing in many structural studies (Tsukazaki et al., 2008); (ii) SecG is important for post- rather than co-translational protein translocation (Belin et al., 2015); (iii) a comparison of SecYE and SecYEG-SecA structures does not reveal any major conformational differences (Tsukazaki et al., 2008; Zimmer et al., 2008); (iv) electron density for SecG was also not observed for Nanodisc-reconstituted SecY (Frauenfeld et al., 2011); and (v) SecG is not essential in *E.coli* (Flower et al., 2000). After building a model of SecY that fully satisfied the experimental electron density we compared the model to existing (crystal) structures of SecY in different states. Our model is clearly distinct from the model of the archeal *M.jannaschii* SecY that was crystallized in the 'closed' idle conformation (Van den Berg et al., 2004). Furthermore, it can be excluded that the tightly coupled SecY is in a state with the lateral gate in an 'open' conformation by

comparing it with the structure of the 'open' SecY from *Pyrococcus furiosus (*Egea *and Stroud, 2010)*. However, in this crystal two SecY molecules were forming an artificial dimer with helix 10 of one SecY inserting in the hydrophilic funnel of the second copy inducing an opening of the lateral gate. To what extent this opening is physiologically relevant remains to be shown.

However, our model is in best agreement with the 'pre-opened' structure of SecY published by (Tsukazaki et al., 2008). This SecY was crystallized with a FAB fragment bound to SecY loop 8/9. Thus, it is tempting to speculate that binding to the loop 8/9 of SecY either by the ribosome or other factors like SecA might already prime the translocon for effective membrane protein insertion. This hypothesis was also recently put forward when interpreting the first high-resolution structure of eukaryotic Sec61 bound to the ribosome (Voorhees et al., 2014).

To conclude, our cryo-EM density of a tightly coupled RNC-SecY complex leads to a pseudo-atomic model of SecY in a 'pre-opened' conformation. Since careful inspection of the electron density 'inside of SecY' did not reveal any additional density that could be attributed to the hydrophilic loop of the substrate, it cannot be clearly resolved whether this 'pre-opening' is triggered by the substrate or the binding to the ribosome.

### 5.2.3 Following the path of the nascent chain

Since the biochemical assays performed with the PRxQ constructs could not reveal the state of membrane insertion, the experimental density was carefully examined to identify the nascent substrate. As expected for an unstructured peptide, we could not identify any electron density corresponding to the hydrophilic loop of PR2Q, thus it was excluded from a pseudo-atomic model of the nascent substrate. However we assume that it is still in the hydrophilic funnel inside SecY based on three arguments: (1) the biochemical findings presented here; (2) previously published cross-linking studies that indicated that hydrophilic substrates pass through the central constriction of SecY (Cannon et al., 2005); and (3) a recently published ribosomes-Sec61 structure containing a hydrophilic stretch inside Sec61 (Gogala et al., 2014). Comparison of our model to the latter reveals high similarity in the conformations of the lateral gate and the plug domain, suggesting that the hydrophilic loop of SecY follows the same path as a secretory protein.

In contrast to the unstructured hydrophilic loop, we identified two rod-like densities right in front of the lateral gate of SecY that were interpreted as the two helices of the substrate PR

since they showed high similarity to the first two helices in the NMR structure of proteorhodopsin (Reckel et al., 2011).
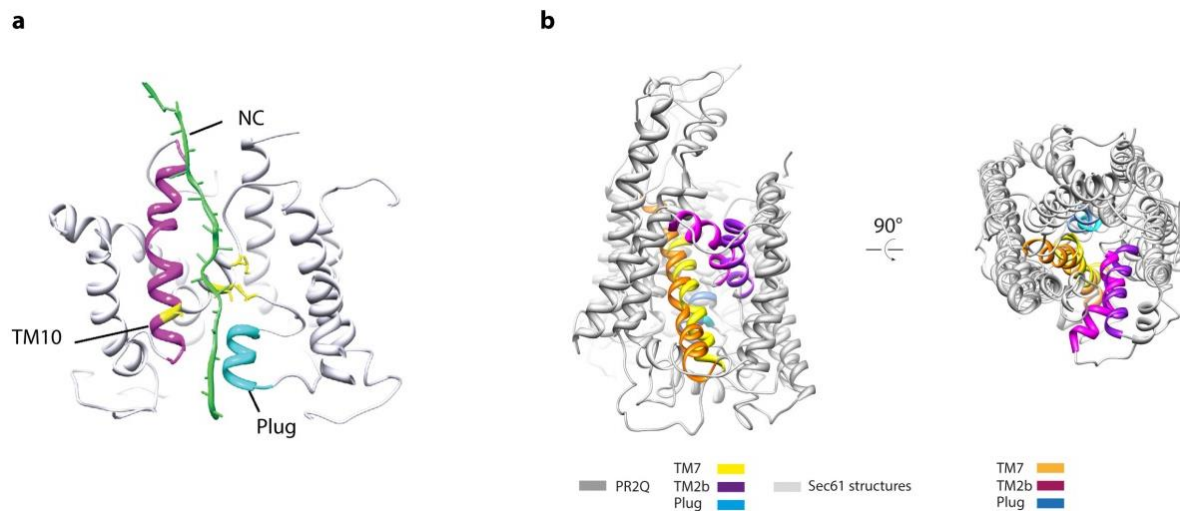


**Figure 46: Comparison of the tightly coupled PR2Q-SecY complex with Sec61 engaging a hydrophilic substrate**
a: Model of a hydrophilic substrate through the Sec61 channel.
b: The overall conformation of the tightly coupled PR2Q-SecY complex is very similar to Sec61 engaging a hydrophilic substrate.
Figure modified from (Bischoff et al., 2014b; Gogala et al., 2014).

Interestingly, we find one of the two substrate helices directly in front of the lateral gate of SecY. Although we miss corroborative evidence, it is assumed that this helix is transmembrane segment 2 of PR. The location of these two helices in front of the lateral gate may correspond to the earlier proposed secondary binding site(s) for nascent polytopic membrane proteins (Sadlish et al., 2005). These peripheral binding sites are thought to stabilize the nascent membrane protein and to assist correct folding.

It should be emphasized that we would exclude that the position in which we find the two substrate helices could be a result of solubilization of the whole complex followed by a spontaneous rebinding of the substrate to SecY. The complex was generated *in vivo* in intact cells and detergent was only applied long after the membrane insertion reaction into the physiological lipid phase took place. Thus, the nascent membrane protein had ample opportunities to partition into the lipid bilayer and it cannot be expected that it would spontaneously re-bind to a defined region of SecY upon detergent solubilization. To the contrary, detergents usually disrupt such interactions. Considering the design of the construct, the TMS theoretically could have bound randomly over the SecY periphery, in which case they would have been averaged out and remain unresolved in our structure.

Although no conclusion could be drawn concerning the path that lead to the position of these two TM segments comparison of our model with other models of SecY/Sec61 engaged with substrates suggests a universally conserved path for proteins partitioning into the membrane. The relative position of signal sequences, signal anchors and a polytopic membrane protein strongly suggest that all of them egress the SecY/Sec61 complex through the lateral gate formed by helices 2b and 7.
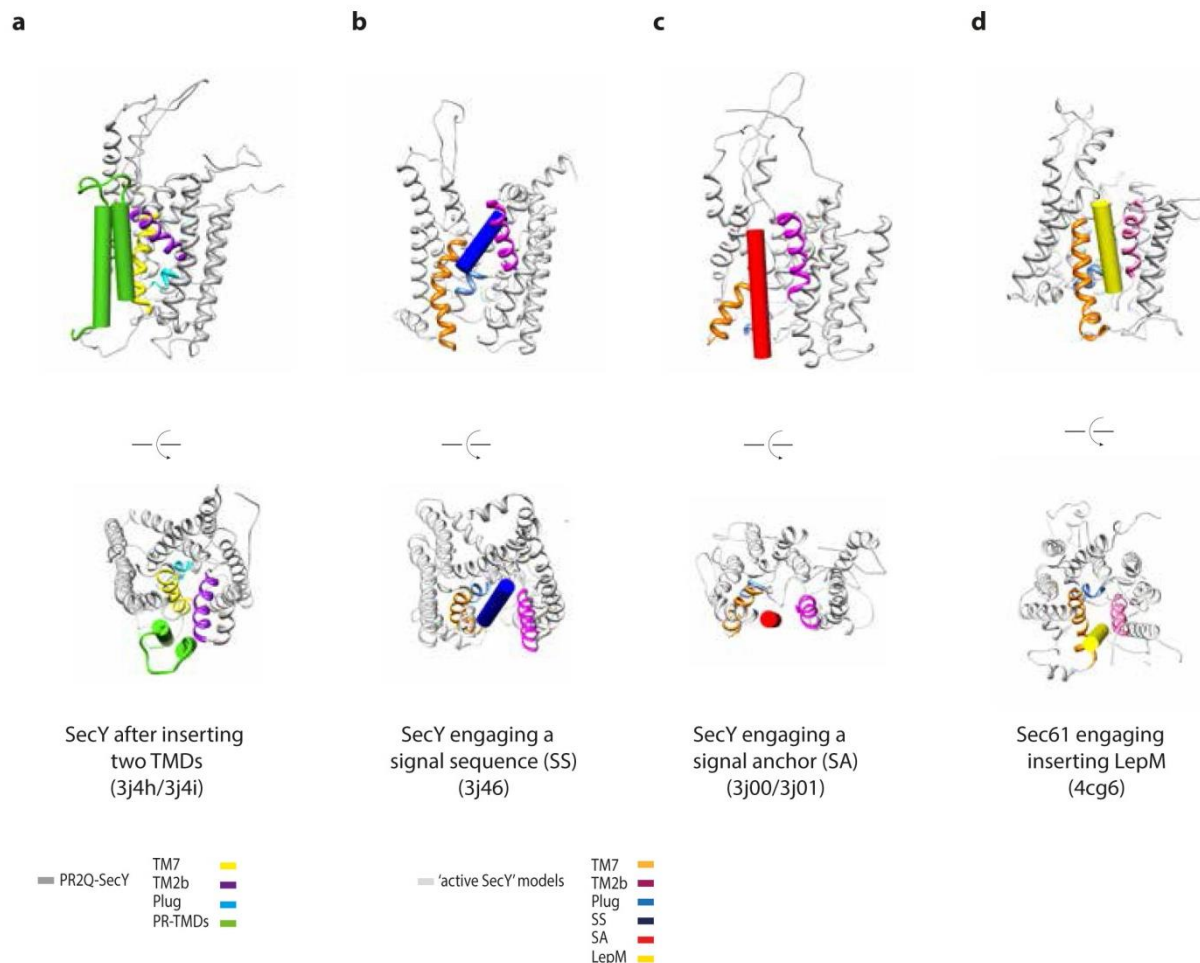


**Figure 47: Position of the nascent transmembrane segments compared to other published models**
The relative position of (a) TM1 and 2 of PR (green) in comparison to a (b) signal sequence (Park et al., 2014) (blue) and a (c) signal anchor (Frauenfeld et al., 2011)(red) engaged with SecY and a (d) hydrophobic TMS of LepM engaged in Sec61 (Gogala et al., 2014) (yellow) suggests a universal path for the membrane insertion of transmembrane segments or signal sequences/anchors.
Figure modified from (Bischoff et al., 2014b).

## 5.2.4 Comparison of our model with a cryo-EM model of SecY cross-linked to a ribosome nascent chain complex

During analysis of the cryo-EM data of the tightly coupled PR2Q-SecY complex presented in this thesis, Park et al. published a cryo-EM model of a RNC-SecY that was isolated applying a

similar strategy as presented here. Park et al. expressed the well-studied secretory protein DsbA that possesses an N-terminal signal sequence followed by a SecM stalling sequence (total length of the construct 100 amino acids) and cross-linked it *in vivo* to the plug domain of SecY using cysteine-cysteine cross-linking. Additionally, Park et al. presented a model of an idle ribosome bound to SecY (Park et al., 2014).
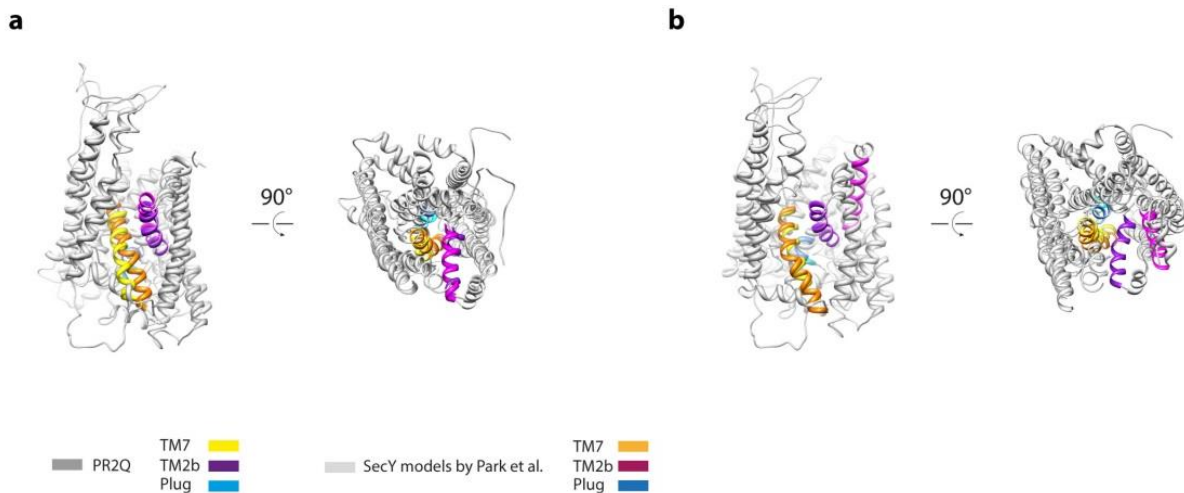


**Figure 48: Comparison of the tightly coupled PR2Q-SecY complex with the idle SecY and the cross-linked DsbA-SecY complexes**
a: Comparison with the idle ribosome SecY complex.
b: Comparison with the active, cross-linked DsbA-SecY complex (Park et al., 2014)(The DsbA signal sequence is removed from the figure for clarity).
Figure modified from (Bischoff et al., 2014b).

Comparison of the idle RNC-SecY complex with our model reveals an overall very similar conformation of SecY, though the lateral gate of the idle RNC-SecY complex is more closed. That suggests that the presence of the hydrophilic loop inside SecY or the two substrate helices in front of the lateral gate trigger that slight opening of the lateral gate observed in our model. In the model of SecY engaging a signal sequence in the lateral gate, the latter is obviously in an open conformation. The position of the signal sequence relative to SecY is depicted in Figure 47. However, the C-terminus remains rather rigid in a very similar conformation compared to our model. However, we refrain from too detailed comparison and interpretation of the structures of Park et al. and the tightly coupled PR2Q SecY complex due to the following reasons: (1) the resolution of the structure of Park et al. is too low for the correct interpretation of secondary structure elements, (2) the authors lack evidence that the DsbA signal sequence is actually present and not already cleaved, (3) the templates ('closed SecY structure for the idle SecY model, 'opened' SecY model for the active SecY model) used to start the MDFF calculations do not ensure an unbiased outcome of the fitting, and (4) the cysteine cross-link between the substrate and SecY formed under

equilibrium conditions does not prove that this path of the nascent chain is of physiological relevance.

### 5.2.5 Interaction of the nascent chain with rRNA helix 59

When building the pseudo-atomic model of the substrate helices we found a continuous density from the rRNA helix 59 to the loop connecting transmembrane helix 1 and 2 of PR as depicted in Figure 27. The presence of four positively charged amino acids in the loop 1/2 and the negatively charged rRNA backbone of the rRNA suggest an ionic interaction between the nascent substrate and the ribosome. The importance of this interaction for the stability of the complex could be shown by mutational analysis of these positively charged amino acids.

ASDYTGVSFWLVTAALLASTVFFF    veRDRvsaKwKTSLT    VSGLVTGIAFWHYMYMRGVWTA

**Figure 49: Amino acid sequence of the first two transmembrane segments and their intermediate cytoplasmic loop**
Boxes sequences correspond to the two transmembrane segments, the unboxed sequence to the cytoplasmic loop. The positively charged residues that could interact with rRNA helix 59 are highlighted in blue.

To date, it is known that the topology of an α helical membrane protein follows the 'positive inside' rule (for review (von Heijne, 2006), thus the positively charged amino acids arginine and lysine are mainly found in cytoplasmic loops compared to periplasmic loops.

Although numerous studies could already provide pieces of data to solve this fundamental problem, the underlying molecular mechanism of the 'positive inside' rule is still obscure. It could be shown that the membrane potential ΔΨ is most likely not responsible for the right topology of membrane proteins (van de Vossenberg et al., 1998). However, it could also be shown that altering the lipid composition of the membrane makes it possible to change the topology of a membrane protein (Bogdanov and Dowhan, 2012). Additionally, evidence was provided that the translocon itself is already involved in arranging nascent transmembrane segments in the correct topology (Prinz et al., 1998; Goder et al., 2004; Junne et al., 2007).

Considering our findings, it is tempting to speculate that a possible interaction between the negative phosphate backbone of rRNA helix 59 in the ribosome and the positive charges in the cytosolic loops of a nascent membrane protein play a crucial role in the topologically correct folding of a membrane protein.

## 5.2.6 Thermodynamic model for the membrane integration of PR2Q

An complementary, thermodynamic view on the integration of membrane proteins was introduced in chapter 1.1.10. The results from biochemical assays on the integration of PR2Q and the results from the built structure can very well be explained by the thermodynamical model for membrane protein integration.

While the 'conventional' model of Sec-dependent membrane protein integration postulates that the transmembrane helices insert into the Sec translocon and are subsequently laterally released into the membrane. In contrast, the thermodynamic model suggests that transmembrane helices insert into the membrane basically by sliding into the membrane along the SecY-lipid interface, similar as suggested for the YidC-dependent membrane protein integration. Only polar periplasmic loops of polytopic membrane proteins would be transported through the hydrophilic interior of the SecY channel. This model would fully satisfy the biochemical observations made for the integration of the proteorhodopsin derivatives presented in this thesis.

Given that hydrophobic domains may never enter SecY it is reasonable that no stable complex between the RNC and the channel would be formed when only integrating hydrophobic transmembrane segments (e.g. PR0-3, FtsQ-PR3TM). However, since a hydrophilic part following the hydrophobic transmembrane segments (PR2Q) has to be translocated through the hydrophilic interior of the channel, a stable RNC-SecY complex is formed.

Nevertheless, more experiments will be needed to fully dissect the mode and the path by which polytopic membrane proteins are integrated into the membrane.
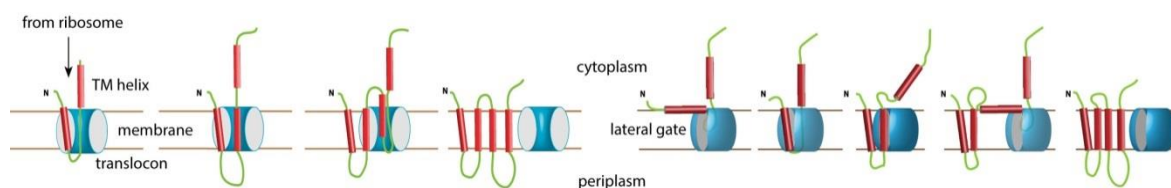


**Figure 50: Models for the integration of polytopic membrane proteins**
Left: Insertion of a polytopic membrane protein via exit of the transmembrane helices through the lateral gate of SecY.
Right: insertion of transmembrane helices on the SecY-lipid interface; translocation of hydrophilic loops through the hydrophilic channel.
Figure modified from (Cymer et al., 2014).

### 5.2.7 Conclusion

On the basis of our experimental data, several conclusions concerning co-translational protein translocation and membrane protein biogenesis in *E.coli* can be drawn.

The binding affinity of SecY to the ribosome exhibits a dual mode, with a tight coupling only in co-translational translocation when inserting a transmembrane segment in type II topology, followed by a hydrophilic loop back to the ribosome. Loose coupling can be observed in the moment of only hydrophobic transmembrane segment inserting into SecY.

Additionally, SecY remains in a 'pre-activated' state after a poly-topic membrane protein egressed, most likely through the lateral gate. The plug helix is only slightly shifting compared to the idle SecY, nevertheless providing enough space for the translocation of an unstructured stretch of amino acids (Lycklama a Nijeholt et al., 2011). Furthermore, an attractive interaction between the rRNA helix 59 and the positively charged amino acids in cytosolic loops of inner membrane proteins is proposed that might assist inner membrane proteins to attain the right topology.

To conclude we complete the model proposed in Figure 51 for the biogenesis of PR2Q:

(1) the signal sequence emerges from the ribosomal exit tunnel and the ribosome is targeted to the membrane most likely by the canonical SRP/FtsY pathway. (2) the correct topology of the signal sequence could be ensured by an electrostatic interaction between positive charges in the N-domain of the signal sequence and rRNA helix 59 in the ribosome (3) at a so far unknown time-point during the insertion, the signal sequence is cleaved. (4) After synthesis of the first TMS of PR2Q the nascent membrane protein is 'tethered' to the ribosome via the newly identified interaction with rRNA helix H59. This interaction might help the next TMS to attain the correct topology (5) after synthesis of the second TMS, when the hydrophilic loop is extended, the binding affinity of the ribosome to SecY decreases due to a still unknown mechanism, leading eventually (6) to a release of the ribosome from SecY allowing SecA to contact and translocate the periplasmic domain of FtsQ leading to the mature membrane protein (7).
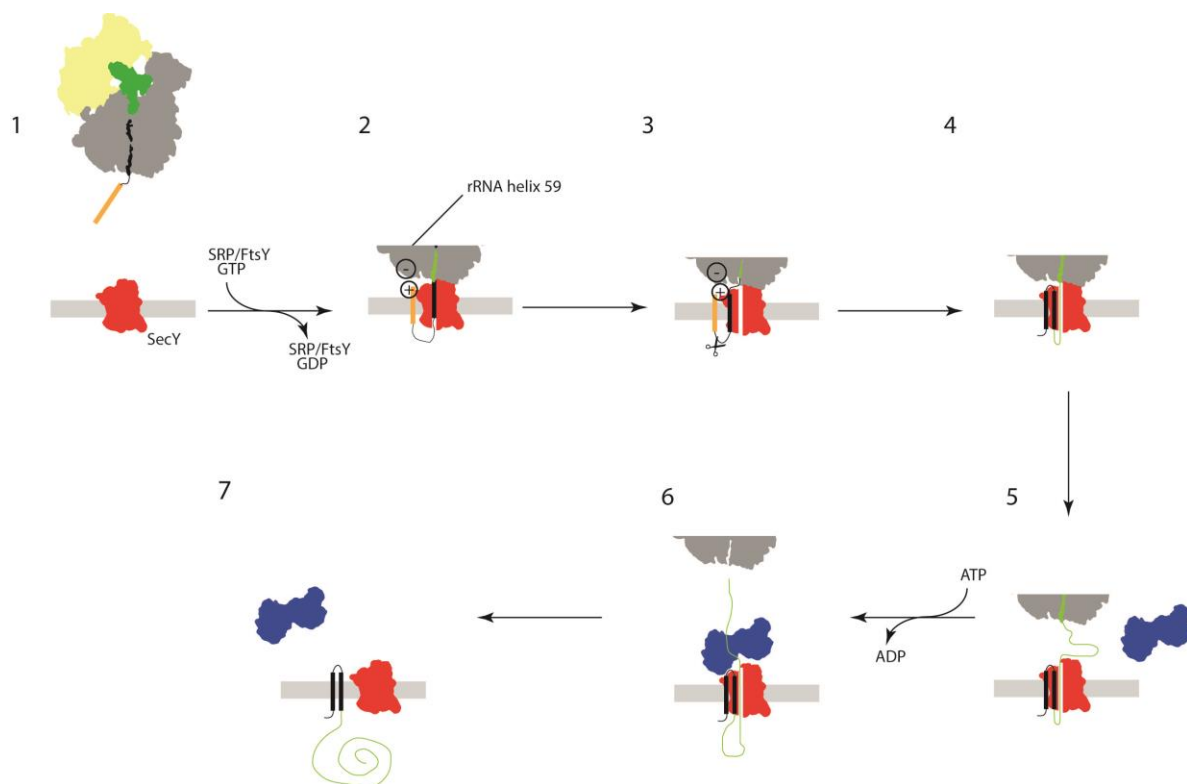
**Figure 51: Schematic model of the biogenesis of PR2Q**

In this model, however, we abstract away from the possibility that another SecY-associated protein assists the insertion of PR2Q. Recent cross-linking data showed that the bacterial membrane protein insertase YidC occupies the lateral gate of SecY and gets replaced by the egressing polytopic membrane protein. In all performed purifications we could not detect any significant amounts of YidC (data not shown); however, we cannot fully exclude that YidC is assisting the folding of PR at some point during its biogenesis. The interplay between SecY and YidC as well as the interplay between SecY and SecA remains subject to further studies.

## 5.3 Site-directed cross-linking

In order to biochemically confirm the position of the two PR helices in the RNC-SecY structure, site directed UV-crosslinking was performed.

The UV-activatable amino acid *p*-bezoyl-L-phenlylalanine was introduced at several positions in the two transmembrane helices of PR, and PR2Q-RNC complexes were purified as described above. Upon UV-irradiation of the isolated complexes, the samples were analyzed by SDS-PAGE and western blotting against the nascent chain and SecY. Although a putative

cross-link of PR2Q(V64pBPA) to SecY could be identified by western blotting (see Figure 28) the results remained inconclusive. While several positions in PR2Q have been tested, only one putative cross-link could be observed.

The anti-SecY western blot shows clear signals for SecY and a UV-dependent extra band that could correspond to a cross-link of SecY to the nascent chain. An extra band at the same size is also present in the anti-HA western blot. However, although the same amount of PR2Q-SecY complexes were loaded in all lanes of the gel, the western blot signals of the UV-treated samples are much stronger. That could be observed in almost all samples that were analyzed (data not shown). The reason for that behavior of the sample upon UV treatment still remains obscure.

The general problem of UV-crosslinking with *p*-BPA lies in its unspecificity. Upon UV-exposure, *p*-BPA enters a diradicaloid triplet state and is very reactive towards C-H bonds (Dorman and Prestwich, 1994). Therefore, it is very likely that intramolecular or other unspecific cross-links to e.g. solvent molecules or detergent molecules are formed, especially because the exact orientation of introduced *p*-BPA is not predictable. Additionally, the *p*-BPA radical can perform an elimination reaction by abstracting two hydrogen atoms of nearby residues. These reactions saturate and therefore inactivate the crosslinking reagent and thus decrease the possibility of a cross-link formation between the substrate and SecY.

Collectively, experiments to support of the position of the two PR helices, as proposed in the cryo-EM based model using the site directed UV-crosslinking did not lead to any conclusive results.

## 5.4 *In vitro* translocation assays to elucidate the translocation mode of PR2Q

We sought of an experiment to elucidate whether PR is targeted to the membrane by the universal SRP pathway, as it would be expected for an inner membrane protein, or if it is targeted by SecA that is known to translocate secretory proteins in a post-translational mode.

In a first *in vitro* assay using cleaned inverted membrane vesicles (IMVs) with over-expressed SecY, translocation in the presence of SRP and FtsY was investigated and successful membrane integration was monitored by cleavage of the signal sequence by endogenous LepB that is still present in the IMVs.

Partial processing of the signal sequence is visible in the samples that contain only SRP or FtsY, which can be explained by residual SRP or FtsY that could not be fully removed during IMV preparation. In samples where SRP and FtsY were added, we observed full processing of the signal sequencing suggesting that PR could be targeted to the membrane by the SRP/FtsY pathway.

To exclude the influence of residual translocation factors as possibly present in the IMVs we performed additional translocation assays in proteoliposomes reconstituted with SecY. Efficient translocation was monitored by protection of the proper inserted substrate against proteinase K.

As expected, no integration could be observed without the addition of any factors or only FtsY added. However, the addition of SRP alone already resulted in a protease protected fragment, suggesting proper membrane insertion. Adding SRP and FtsY to the reaction mixture also resulted in a protease protected fragment, as already observed in the reaction with the IMVs.

Interestingly, addition of SecA/SecB led to the same protease protected fragment as seen in the SRP/FtsY sample. Since the assay was performed under equilibrium conditions and without distinguishing between co- and post-translational translocation, we cannot exclude, that PR is targeted to the membrane post-translational by the SecA/B pathway and then inserted with the help of SecY or spontaneously.

Additionally, given that SRP alone and SecA/B act as chaperone and keep PR2Q in solution during the reaction time, the chance of spontaneous membrane integration could be favored and thus might explain the positive membrane integration in the SecA/B samples.

Furthermore, it is worth mentioning that we cannot exclude completely that the signals that were interpreted as protease protected fragments were artefacts of the incubation with proteinase K. It was shown that proteinase K exhibits a significantly higher activity in the

presence of Triton X100 (Hilz et al., 1975). Hence assuming that the PR2Q is not integrated at all in the proteoliposomes, the presence and absence of signals in the western blot can be only due to the higher efficiency of protease K in detergent containing buffers.

To sum it up, our experiments and available literature on the biogenesis of inner membrane proteins would support that PR is targeted to the membrane by the SRP/FtsY pathway.

However, more experiments a required to finally dissect the correct pathway, how PR is targeted and integrated into the membrane, that will be carried out in cooperation with AG Koch at the University of Freiburg.

## 5.5 Molecular basis for the ribosome functioning as an L-tryptophan sensor

Solving the structure of a TnaC-stalled ribosome by cryo-EM and single particle reconstruction to an average resolution of 3.8 Å enabled us to build a molecular model of the nascent chain, the PTC and the ribosomal tunnel wall. In this model we could identify not only one, but two free L-Trp molecules that are bound in composite binding sites formed by the nascent chain and components of the ribosomal tunnel wall.

Moreover, we find two crucial rRNA bases, U2585 and A2602 in the PTC stabilized in conformation that precludes the accommodation of the GGQ motif of release factor 2. Additionally, that stabilization of R23 of TnaC also hampers the peptide release by RF2. All identified contacts between the nascent chain, the free L-Trp molecules and the ribosomal exit tunnel are additionally summarized in Figure 52.

| TnaC Residue | Contact |
|---|---|
| P24 | rRNA U2585 |
| R23 | rRNA U2505, rRNA C2452, U2506 |
| H22 | rRNA U2505 |
| D21 | rRNA 2062, $W_1$ |
| V20 | $W_1$, $W_2$ |
| I19 | $W_1$ |
| K18 | rRNA U2609-A752 |
| N17 | rRNA U2609-A752 |
| D16 | K90,R92 of L22 |
| I15 | $W_2$ |
| N14 | R61 of uL4 |
| F13 | rRNA A751 |
| W12 | R92 of uL22 |
| $W_1$ | rRNA U2586, TnaC V20 and I19 |
| $W_2$ | rRNA A2058, rRNA 2059, TnaC I19 and I15 |
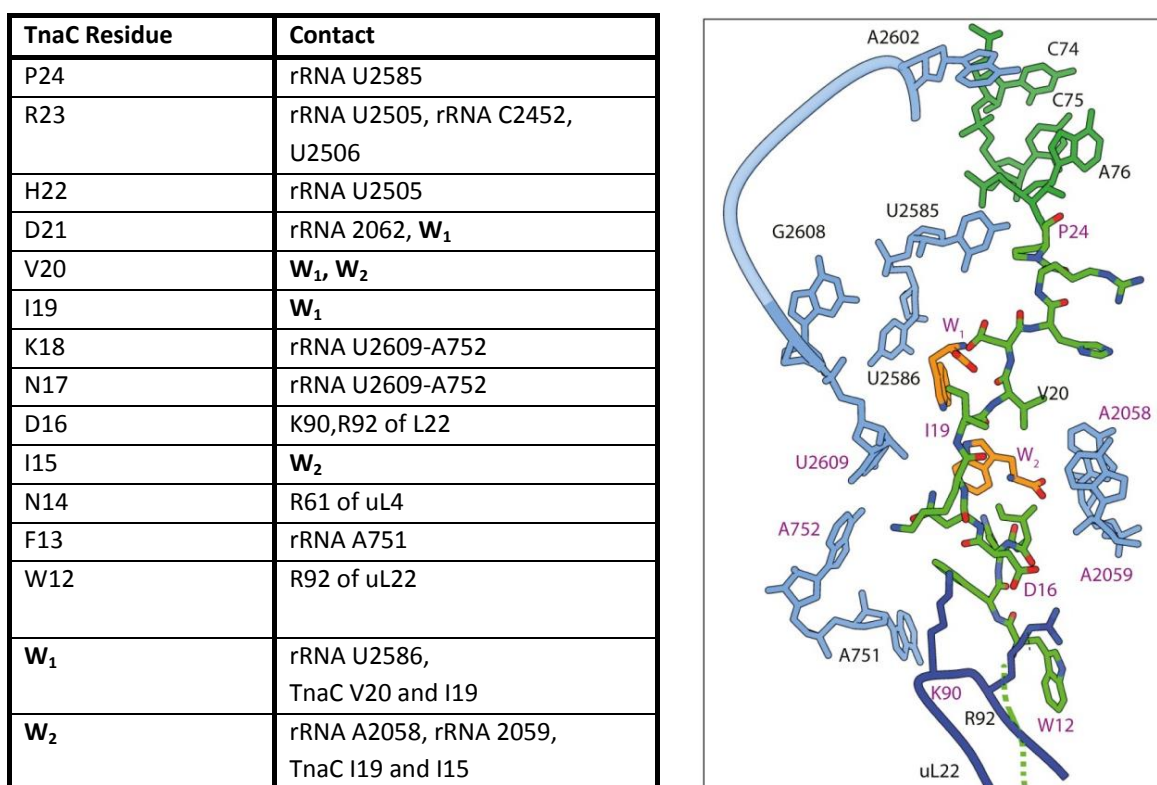


**Figure 52: Table and scheme of the molecular model summarizing all the interactions between TnaC residues and the free L-Trp molecules with components of the ribosomal tunnel wall and the PTC** In the schematic, the CCA-end of the P-site tRNA TnaC peptide is colored in light green, the free L-Trp molecules in orange, the ribosomal RNA is light blue and L22 is blue. Residues identified in mutational studies to be of particular importance for effective TnaC stalling are highlighted in magenta (references given in the text). Figure and Table modified from (Bischoff et al., 2014a).

As depicted in Figure 52, our model supports and explains most of the mutational data available in the literature (for review (Ito and Chiba, 2013)). Nevertheless, the mutational data has to be analyzed with caution, as mutations in the ribosome might induce off-pathways and indirect effects that are not directly related to the molecular mechanisms of TnaC-mediated stalling. For example, a direct involvement of A750, A751 and A754 cannot be supported, since these bases of the ribosome are not engaged in any particular interaction. The effect of mutating these bases thus might result from an overall rearrangement of rRNA caused by these mutations.

A recent study identified TnaC residue I19 to be highly important for effective TnaC stalling (Martinez et al., 2014) and it was proposed to interact with A2058/A2059. Interestingly, our structure does not support this interaction since we cannot assign any electron density to the side chain of I19, indicating high flexibility in this residue that would speak against an interaction with a component of the ribosomal exit tunnel. Instead I19 might be responsible for creating a hydrophobic environment for effective binding of $W_1$ and $W_2$. On the same line we also cannot assign a defined position for I15, thus suggesting a similar role for I15. Well in line, mutation of I15 and I19 to less hydrophobic residues eliminates TnaC stalling (Gish and Yanofsky, 1995; Cruz-Vera and Yanofsky, 2008; Martinez et al., 2014).

Additionally, we find the residues D16 and W12 of TnaC in a 'zipper'-like interaction with K90 and R91 of uL22 of the ribosome. Mutational studies highlighted the importance of all these residues for effective stalling. The possible role of these residues together with P24 of TnaC is discussed further below. An interaction with the ribosomal protein uL4 has also been proposed based on mutational studies and is supported by our molecular model as we find an interaction between N14 of TnaC and R62 of uL4.

However, since the two free L-Trp molecules are bound relatively far away from the PTC the question remains how this specific state of the PTC is induced.
Analysis of several crystal structures of the ribosome revealed that the PTC bases U2585 and A2602 exhibit various conformations depending on the functional state of the ribosome. Additionally, mutational and kinetic studies underlined the importance of these residues for effective function of the PTC. Taken together, this suggests that flexibility of these two bases is of particular importance for the function of the ribosome in different states (Hansen et al., 2002; Schmeing et al., 2005a; Schmeing et al., 2005b; Dunkle and Cate, 2010; Dunkle et al., 2010; Jin et al., 2010; Dunkle et al., 2011).

Hence, we suggest that the major reason for the inability of RF2-mediated peptide release lies in the dramatic loss of flexibility in the PTC that is a direct consequence of L-Trp binding. $W_1$ engaged in a stacking interaction with U2586 might stabilize U2585 in its particular conformation together with the interaction of G2608 and the phosphate backbone of U2585. Intriguingly, mutation of U2586 to a purine base even increased stalling efficiency, indicating the importance of this stacking interaction that is obviously stronger with a purine instead of a pyrimidine base.

Furthermore, $W_2$ interacts with A2058, A2059 in the ribosomal exit tunnel and together with K18 and N17 of TnaC that interact with a base pair A2609/U752 in the ribosomal exit tunnel that forms a 'bridge' spanning over the whole exit tunnel. The importance of these four rRNA bases as well as the importance of a positively charged residue 18 in TnaC was addressed in several mutational studies before (see Table 1).

Interestingly, this 'bridge' resembles the ketolide antibiotic telithromycin (Dunkle et al., 2010) that, similar to erythromycin (Dunkle et al., 2010), bridges the ribosomal exit tunnel in the very same position and also induces translation arrest. This could lead to the general conclusion that the area of A2058, A2059 and the base pair U2609/A752 plays a crucial role in the allosteric control of the ribosome by serving as a binding platform for various small molecules.
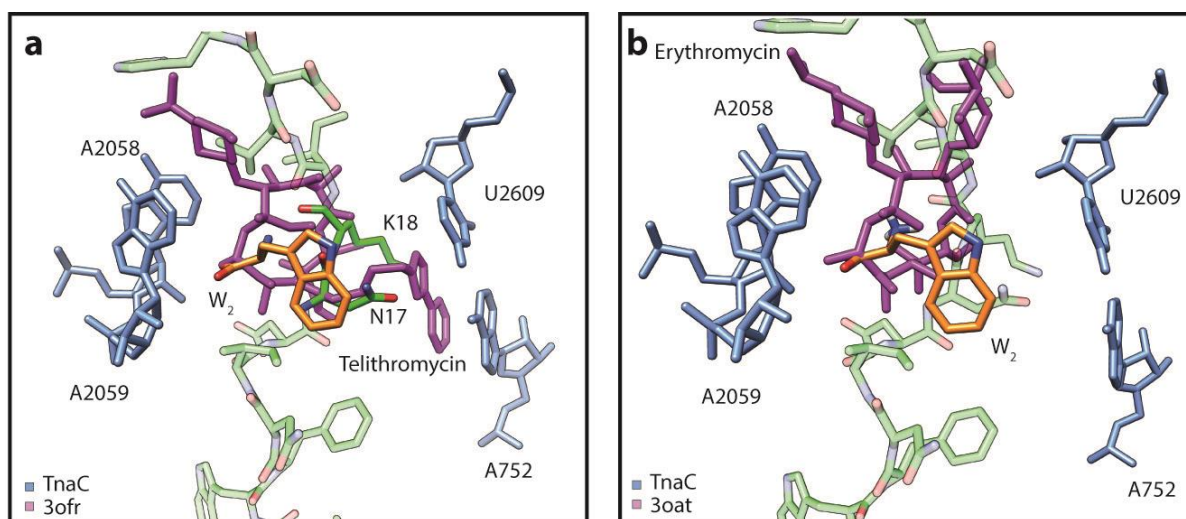


**Figure 53: Comparing the position of ketolide and macrolide antibiotics with the position of one free L-tryptophan molecule in the ribosomal exit tunnel**
a: $W_2$ together with K18 and N17 of TnaC bridges the ribosomal exit tunnel with high similarity to the ketolide telithromycin.
b: Overlay of $W_2$ with the macrolide erythromycin also shows high similarity in the binding site.
Figure modified from (Bischoff et al., 2014a).

Of additional interest is the role of P24 of TnaC which is engaged in an interaction with U2585 and is absolutely crucial for efficient TnaC stalling. We suggest that the general

relatively slow incorporation of proline by the ribosome might play a crucial role, as it opens a time window for L-Trp binding and the other interactions to form. Additionally, the distance between P24, D16 and W12 of TnaC is of particular importance.

In contrast to the PTC inhibiting RF2 accommodation, the accommodation of an A-site tRNA is not inhibited by the stabilized bases of the TnaC-PTC. Although the mentioned bases are stabilized by L-Trp and TnaC, under the equilibrium conditions in the cell incorporation of an additional amino acid in place of the stop codon would alleviate stalling, which has been shown in biochemical studies (Cruz-Vera et al., 2006). However, placing a rare isoleucine codon (AUA) in the place of the stop codon induces stalling on the TnaC peptide (Cruz-Vera et al., 2006). That suggests a general mechanism of kinetic competition between the binding of L-Trp in the tunnel, establishment of all crucial interactions and accommodation of A-site ligands, like release factors, and aa-tRNAs.

Together with our model and available mutational data, we would put forward the following model for L-Trp induced stalling on the TnaC peptide:

The TnaC peptide is translated in the presence of inducing levels of L-Trp. The relatively slow incorporation of P24 leads to a slowdown in translation (Pavlov et al., 2009). This enables the conserved TnaC residues D16 and W12 to engage in the reported interactions that additionally stabilize the nascent chain in the tunnel. This slow-down and stabilization allow a collective engagement of other TnaC residues with components of the ribosomal exit tunnel. Consequently, two composite binding pockets of the TnaC peptide and the ribosomal exit tunnel are formed that allow binding of two free L-Trp molecules, thus turning the translating ribosome in an effective tryptophan sensor. Whether the two L-Trp molecules bind in a cooperative fashion still remains to be elucidated. Binding of the L-Trp molecules finally induces allosteric silencing of the PTC as described above.

It will be interesting to see, whether this is a unique mechanism or if other small molecule sensing regulatory peptides act in a similar way to affect the translating ribosome.

# 6. Summary and outlook

To study the mechanisms of membrane protein insertion we established a protocol that allows isolation of *in vivo* assembled ribosome nascent chain complexes (RNCs) from *E. coli* in high yield and quality. To investigate the interaction of SecY with a translating ribosome, model membrane proteins of different length and topology were over-expressed and the respective RNCs were isolated under mild conditions to allow co-purification of the SecY complex. Analysis of the interaction of RNCs with SecY *in vivo* suggested that, as expected, a tight engagement of the ribosome and SecY is only established for nascent chains that are translocated co-translationally. We observed that SecY and the RNC do not form a stable complex at the moment of hydrophobic transmembrane segments inserting in the translocon. However, a stable engagement of the RNC with SecY was observed, when inserting a transmembrane segment with a type II topology into SecY followed by a hydrophilic loop of a certain length which allows the isolation of this complex. That suggested a dual binding mode of tight and loose coupling of SecY to the translating ribosome dependent on the nature of the nascent substrate.

We present the first three dimensional structure of an *in vivo* assembled, tightly coupled polytopic RNC-SecYE complex at 7.2 Å solved by cryo-EM and single particle reconstruction.

A molecular model based on the cryo-EM structure reveals that SecYE could be trapped in a post-insertion state, with the two substrate helices interacting with the periphery of SecY, while still translocating the hydrophilic loop. The lateral gate of SecY remains in a 'pre-opened' conformation during the translocation of the hydrophilic loop. The interaction sites of SecY with the ribosome were found as described. Remarkably, we could also reveal an interaction of helix 59 in the ribosome with nascent membrane protein via positively charged residues in the first cytoplasmic loop of the substrate. It is tempting to speculate that this interaction contributes to the positive inside rule.

Though, we provided an unprecedented snapshot of an inserting polytopic membrane protein, the exact path of the nascent chain and the molecular mechanism of the actual insertion could not be solved so far.

Further structural studies on different substrates (e.g. PR constructs having 4 or 6 TMS) trapped in different time-points of their insertion will provide a more comprehensive insight into the mechanisms of membrane protein biogenesis.

Expression of the *E. coli* tryptophanase (TnaA) operon is triggered by ribosome stalling during translation of the upstream TnaC leader peptide. Notably, this stalling is strictly dependent on the presence of tryptophan that acts in a hitherto unknown way. Here, we

present a cryo-EM reconstruction of the stalled nascent TnaC leader peptide in the ribosomal exit tunnel.

The structure of the TnaC-stalled ribosome was solved to an average resolution of 3.8 Å by cryo-EM and single particle analysis. It reveals the conformation of the silenced peptidyl-transferase center as well as the exact path of the stalled nascent peptide and its contacts in detail. Furthermore, we clearly resolve not a single but two free tryptophan molecules in the ribosomal exit tunnel. The nascent TnaC peptide chain together with distinct rRNA bases in the ribosomal exit tunnel creates two hydrophobic binding pockets for the tryptophan coordination. One tryptophan molecule is coordinated by V20 and I19 of TnaC and interacts with U2586 of the rRNA, the second tryptophan is bound between I19 and I15 in the area of A2058 and A2059 of the rRNA. Interestingly, the latter is also the binding platform for macrolide antibiotics. Engagement of L-Trp in these composite binding pockets leads to subtle conformational changes in residues of the ribosomal tunnel wall that are translated to the PTC eventually resulting in silencing by stabilizing the conformations of the conserved nucleotides A2602 and U2585. These conformations of the two nucleotides in the PTC are incompatible with the correct accommodation of the GGQ motive of release factor 2, thus inhibiting the peptide release.

In a future study we plan to solve the three dimensional structure of release factor 2 bound to the TnaC-stalled ribosome, to further characterize how the silenced PTC inhibits the release of the peptide.

101

# 7. References

Akimaru, J., S. Matsuyama, H. Tokuda, and S. Mizushima. 1991. Reconstitution of a protein translocation system containing purified SecY, SecE, and SecA from Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*. 88:6545-6549.

Almeida, P.F., A.S. Ladokhin, and S.H. White. 2012. Hydrogen-bond energetics drive helix formation in membrane interfaces. *Biochimica et biophysica acta*. 1818:178-182.

Amunts, A., A. Brown, X.C. Bai, J.L. Llacer, T. Hussain, P. Emsley, F. Long, G. Murshudov, S.H. Scheres, and V. Ramakrishnan. 2014. Structure of the yeast mitochondrial large ribosomal subunit. *Science*. 343:1485-1489.

Andersson, H., and G. von Heijne. 1994. Membrane protein topology: effects of delta mu H+ on the translocation of charged residues explain the 'positive inside' rule. *The EMBO journal*. 13:2267-2272.

Angelini, S., D. Boy, E. Schiltz, and H.G. Koch. 2006. Membrane binding of the bacterial signal recognition particle receptor involves two distinct binding sites. *The Journal of cell biology*. 174:715-724.

Angelini, S., S. Deitermann, and H.G. Koch. 2005. FtsY, the bacterial signal-recognition particle receptor, interacts functionally and physically with the SecYEG translocon. *EMBO reports*. 6:476-481.

Anger, A.M., J.P. Armache, O. Berninghausen, M. Habeck, M. Subklewe, D.N. Wilson, and R. Beckmann. 2013. Structures of the human and Drosophila 80S ribosome. *Nature*. 497:80-85.

Arenz, S., S. Meydan, A.L. Starosta, O. Berninghausen, R. Beckmann, N. Vazquez-Laslop, and D.N. Wilson. 2014a. Drug Sensing by the Ribosome Induces Translational Arrest via Active Site Perturbation. *Molecular cell*.

Arenz, S., H. Ramu, P. Gupta, O. Berninghausen, R. Beckmann, N. Vazquez-Laslop, A.S. Mankin, and D.N. Wilson. 2014b. Molecular basis for erythromycin-dependent ribosome stalling during translation of the ErmBL leader peptide. *Nature communications*. 5:3501.

Ataide, S.F., N. Schmitz, K. Shen, A. Ke, S.O. Shan, J.A. Doudna, and N. Ban. 2011. The crystal structure of the signal recognition particle in complex with its receptor. *Science*. 331:881-886.

Bai, X.C., I.S. Fernandez, G. McMullan, and S.H. Scheres. 2013. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *eLife*. 2:e00461.

Bauer, B.W., and T.A. Rapoport. 2009. Mapping polypeptide interactions of the SecA ATPase during translocation. *Proceedings of the National Academy of Sciences of the United States of America*. 106:20800-20805.

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

Becker, T., S. Bhushan, A. Jarasch, J.P. Armache, S. Funes, F. Jossinet, J. Gumbart, T. Mielke, O. Berninghausen, K. Schulten, E. Westhof, R. Gilmore, E.C. Mandon, and R. Beckmann. 2009. Structure of monomeric yeast and mammalian Sec61 complexes interacting with the translating ribosome. *Science*. 326:1369-1373.

Beckmann, R., D. Bubeck, R. Grassucci, P. Penczek, A. Verschoor, G. Blobel, and J. Frank. 1997. Alignment of conduits for the nascent polypeptide chain in the ribosome-Sec61 complex. *Science*. 278:2123-2126.

Beckmann, R., C.M. Spahn, J. Frank, and G. Blobel. 2001. The active 80S ribosome-Sec61 complex. *Cold Spring Harbor symposia on quantitative biology*. 66:543-554.

Beja, O., L. Aravind, E.V. Koonin, M.T. Suzuki, A. Hadd, L.P. Nguyen, S.B. Jovanovich, C.M. Gates, R.A. Feldman, J.L. Spudich, E.N. Spudich, and E.F. DeLong. 2000. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. *Science*. 289:1902-1906.

Beja, O., E.N. Spudich, J.L. Spudich, M. Leclerc, and E.F. DeLong. 2001. Proteorhodopsin phototrophy in the ocean. *Nature*. 411:786-789.

Belin, D., G. Plaia, Y. Boulfekhar, and F. Silva. 2015. Escherichia coli SecG Is Required for Residual Export Mediated by Mutant Signal Sequences and for SecY-SecE Complex Stability. *Journal of bacteriology*. 197:542-552.

Ben-Shaul, A., N. Ben-Tal, and B. Honig. 1996. Statistical thermodynamic analysis of peptide and protein insertion into lipid membranes. *Biophysical journal*. 71:130-137.

Ben-Tal, N., A. Ben-Shaul, A. Nicholls, and B. Honig. 1996. Free-energy determinants of alpha-helix insertion into lipid bilayers. *Biophysical journal*. 70:1803-1812.

Bessonneau, P., V. Besson, I. Collinson, and F. Duong. 2002. The SecYEG preprotein translocation channel is a conformationally dynamic and dimeric structure. *The EMBO journal*. 21:995-1003.

Bhushan, S., M. Gartmann, M. Halic, J.P. Armache, A. Jarasch, T. Mielke, O. Berninghausen, D.N. Wilson, and R. Beckmann. 2010a. alpha-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. *Nature structural & molecular biology*. 17:313-317.

Bhushan, S., T. Hoffmann, B. Seidelt, J. Frauenfeld, T. Mielke, O. Berninghausen, D.N. Wilson, and R. Beckmann. 2011. SecM-stalled ribosomes adopt an altered geometry at the peptidyl transferase center. *PLoS biology*. 9:e1000581.

Bhushan, S., H. Meyer, A.L. Starosta, T. Becker, T. Mielke, O. Berninghausen, M. Sattler, D.N. Wilson, and R. Beckmann. 2010b. Structural basis for translational stalling by human cytomegalovirus and fungal arginine attenuator peptide. *Molecular cell*. 40:138-146.

Bischoff, L., O. Berninghausen, and R. Beckmann. 2014a. Molecular basis for the ribosome functioning as an L-tryptophan sensor. *Cell reports*. 9:469-475.

Bischoff, L., S. Wickles, O. Berninghausen, E.O. van der Sluis, and R. Beckmann. 2014b. Visualization of a polytopic membrane protein during SecY-mediated membrane insertion. *Nature communications*. 5:4103.

Blobel, G., and B. Dobberstein. 1975a. Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. *The Journal of cell biology*. 67:835-851.

Blobel, G., and B. Dobberstein. 1975b. Transfer of proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. *The Journal of cell biology*. 67:852-862.

Bogdanov, M., and W. Dowhan. 2012. Lipid-dependent generation of dual topology for a membrane protein. *The Journal of biological chemistry*. 287:37939-37948.

Boyd, D., and J. Beckwith. 1990. The role of charged amino acids in the localization of secreted and membrane proteins. *Cell*. 62:1031-1033.

Boyd, D., C. Manoil, and J. Beckwith. 1987. Determinants of membrane protein topology. *Proceedings of the National Academy of Sciences of the United States of America*. 84:8525-8529.

Bradatsch, B., C. Leidig, S. Granneman, M. Gnadig, D. Tollervey, B. Bottcher, R. Beckmann, and E. Hurt. 2012. Structure of the pre-60S ribosomal subunit with nuclear export factor Arx1 bound at the exit tunnel. *Nature structural & molecular biology*. 19:1234-1241.

Bradshaw, N., and P. Walter. 2007. The signal recognition particle (SRP) RNA links conformational changes in the SRP to protein targeting. *Molecular biology of the cell*. 18:2728-2734.

Braig, D., C. Bar, J.O. Thumfart, and H.G. Koch. 2009. Two cooperating helices constitute the lipid-binding domain of the bacterial SRP receptor. *Journal of molecular biology*. 390:401-413.

Breyton, C., W. Haase, T.A. Rapoport, W. Kuhlbrandt, and I. Collinson. 2002. Three-dimensional structure of the bacterial protein-translocation complex SecYEG. *Nature*. 418:662-665.

Brown, A., A. Amunts, X.C. Bai, Y. Sugimoto, P.C. Edwards, G. Murshudov, S.H. Scheres, and V. Ramakrishnan. 2014. Structure of the large ribosomal subunit from human mitochondria. *Science*. 346:718-722.

Brundage, L., J.P. Hendrick, E. Schiebel, A.J. Driessen, and W. Wickner. 1990. The purified E. coli integral membrane protein SecY/E is sufficient for reconstitution of SecA-dependent precursor protein translocation. *Cell*. 62:649-657.

Calhoun, K.A., and J.R. Swartz. 2006. Total amino acid stabilization during cell-free protein synthesis reactions. *Journal of biotechnology*. 123:193-203.

Cannon, K.S., E. Or, W.M. Clemons, Jr., Y. Shibata, and T.A. Rapoport. 2005. Disulfide bridge formation between SecY and a translocating polypeptide localizes the translocation pore to the center of SecY. *The Journal of cell biology*. 169:219-225.

Castanie-Cornet, M.P., N. Bruel, and P. Genevaux. 2014. Chaperone networking facilitates protein targeting to the bacterial cytoplasmic membrane. *Biochimica et biophysica acta*. 1843:1442-1456.

Chatzi, K.E., M.F. Sardis, A. Economou, and S. Karamanou. 2014. SecA-mediated targeting and translocation of secretory proteins. *Biochimica et biophysica acta*. 1843:1466-1474.

Chen, J.Z., and N. Grigorieff. 2007. SIGNATURE: a single-particle selection system for molecular electron microscopy. *Journal of structural biology*. 157:168-173.

Chin, J.W., A.B. Martin, D.S. King, L. Wang, and P.G. Schultz. 2002. Addition of a photocrosslinking amino acid to the genetic code of Escherichiacoli. *Proceedings of the National Academy of Sciences of the United States of America*. 99:11020-11024.

Cruz-Vera, L.R., M. Gong, and C. Yanofsky. 2006. Changes produced by bound tryptophan in the ribosome peptidyl transferase center in response to TnaC, a nascent leader peptide. *Proceedings of the National Academy of Sciences of the United States of America*. 103:3598-3603.

Cruz-Vera, L.R., A. New, C. Squires, and C. Yanofsky. 2007. Ribosomal features essential for tna operon induction: tryptophan binding at the peptidyl transferase center. *Journal of bacteriology*. 189:3140-3146.

Cruz-Vera, L.R., S. Rajagopal, C. Squires, and C. Yanofsky. 2005. Features of ribosome-peptidyl-tRNA interactions essential for tryptophan induction of tna operon expression. *Molecular cell*. 19:333-343.

Cruz-Vera, L.R., R. Yang, and C. Yanofsky. 2009. Tryptophan inhibits Proteus vulgaris TnaC leader peptide elongation, activating tna operon expression. *Journal of bacteriology*. 191:7001-7006.

Cruz-Vera, L.R., and C. Yanofsky. 2008. Conserved residues Asp16 and Pro24 of TnaC-tRNAPro participate in tryptophan induction of Tna operon expression. *Journal of bacteriology*. 190:4791-4797.

Cymer, F., and G. von Heijne. 2013. Cotranslational folding of membrane proteins probed by arrest-peptide-mediated force measurements. *Proceedings of the National Academy of Sciences of the United States of America*. 110:14640-14645.

Cymer, F., G. von Heijne, and S.H. White. 2014. Mechanisms of Integral Membrane Protein Insertion and Folding. *Journal of molecular biology*.

Dalbey, R.E., A. Kuhn, L. Zhu, and D. Kiefer. 2014. The membrane insertase YidC. *Biochimica et biophysica acta*. 1843:1489-1496.

Demirci, E., T. Junne, S. Baday, S. Berneche, and M. Spiess. 2013. Functional asymmetry within the Sec61p translocon. *Proceedings of the National Academy of Sciences of the United States of America*. 110:18856-18861.

Derman, A.I., J.W. Puziss, P.J. Bassford, Jr., and J. Beckwith. 1993. A signal sequence is not required for protein export in prlA mutants of Escherichia coli. *The EMBO journal*. 12:879-888.

Ding, H., J.F. Hunt, I. Mukerji, and D. Oliver. 2003. Bacillus subtilis SecA ATPase exists as an antiparallel dimer in solution. *Biochemistry*. 42:8729-8738.

Dorman, G., and G.D. Prestwich. 1994. Benzophenone photophores in biochemistry. *Biochemistry*. 33:5661-5673.

Doura, A.K., and K.G. Fleming. 2004. Complex interactions at the helix-helix interface stabilize the glycophorin A transmembrane dimer. *Journal of molecular biology*. 343:1487-1497.

Dowhan, W., and M. Bogdanov. 2009. Lipid-dependent membrane protein topogenesis. *Annual review of biochemistry*. 78:515-540.

Driessen, A.J., and N. Nouwen. 2008. Protein translocation across the bacterial cytoplasmic membrane. *Annual review of biochemistry*. 77:643-667.

du Plessis, D.J., N. Nouwen, and A.J. Driessen. 2006. Subunit a of cytochrome o oxidase requires both YidC and SecYEG for membrane insertion. *The Journal of biological chemistry*. 281:12248-12252.

Dudek, J., S. Pfeffer, P.H. Lee, M. Jung, A. Cavalie, V. Helms, F. Forster, and R. Zimmermann. 2014. Protein Transport into the Human Endoplasmic Reticulum. *Journal of molecular biology*.

Dunkle, J.A., and J.H. Cate. 2010. Ribosome structure and dynamics during translocation and termination. *Annual review of biophysics*. 39:227-244.

Dunkle, J.A., L. Wang, M.B. Feldman, A. Pulk, V.B. Chen, G.J. Kapral, J. Noeske, J.S. Richardson, S.C. Blanchard, and J.H. Cate. 2011. Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. *Science*. 332:981-984.

Dunkle, J.A., L. Xiong, A.S. Mankin, and J.H. Cate. 2010. Structures of the Escherichia coli ribosome with antibiotics bound near the peptidyl transferase center explain spectra of drug action. *Proceedings of the National Academy of Sciences of the United States of America*. 107:17152-17157.

Duong, F. 2003. Binding, activation and dissociation of the dimeric SecA ATPase at the dimeric SecYEG translocase. *The EMBO journal*. 22:4375-4384.

Egea, P.F., and R.M. Stroud. 2010. Lateral opening of a translocon upon entry of protein suggests the mechanism of insertion into membranes. *Proceedings of the National Academy of Sciences of the United States of America*. 107:17182-17187.

Elofsson, A., and G. von Heijne. 2007. Membrane protein structure: prediction versus reality. *Annual review of biochemistry*. 76:125-140.

Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography*. 60:2126-2132.

Erlandson, K.J., S.B. Miller, Y. Nam, A.R. Osborne, J. Zimmer, and T.A. Rapoport. 2008. A role for the two-helix finger of the SecA ATPase in protein translocation. *Nature*. 455:984-987.

Eswar, N., D. Eramian, B. Webb, M.Y. Shen, and A. Sali. 2008. Protein structure modeling with MODELLER. *Methods in molecular biology*. 426:145-159.

Eswar, N., B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M.Y. Shen, U. Pieper, and A. Sali. 2006. Comparative protein structure modeling using Modeller. *Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.]*. Chapter 5:Unit 5 6.

Eswar, N., B. Webb, M.A. Marti-Renom, M.S. Madhusudhan, D. Eramian, M.Y. Shen, U. Pieper, and A. Sali. 2007. Comparative protein structure modeling using MODELLER. *Current protocols in protein science / editorial board, John E. Coligan ... [et al.]*. Chapter 2:Unit 2 9.

Facey, S.J., S.A. Neugebauer, S. Krauss, and A. Kuhn. 2007. The mechanosensitive channel protein MscL is targeted by the SRP to the novel YidC membrane insertion pathway of Escherichia coli. *Journal of molecular biology*. 365:995-1004.

Farrell, I.S., R. Toroney, J.L. Hazen, R.A. Mehl, and J.W. Chin. 2005. Photo-cross-linking interacting proteins with a genetically encoded benzophenone. *Nature methods*. 2:377-384.

Fernandez, I.S., X.C. Bai, G. Murshudov, S.H. Scheres, and V. Ramakrishnan. 2014. Initiation of translation by cricket paralysis virus IRES requires its translocation in the ribosome. *Cell*. 157:823-831.

Fernandez, J.J., D. Luque, J.R. Caston, and J.L. Carrascosa. 2008. Sharpening high resolution information in single particle electron cryomicroscopy. *Journal of structural biology*. 164:170-175.

Flower, A.M., L.L. Hines, and P.L. Pfennig. 2000. SecG is an auxiliary component of the protein export apparatus of Escherichia coli. *Molecular & general genetics : MGG*. 263:131-136.

Frank, J., P. Penczek, R. Grassucci, and S. Srivastava. 1991. Three-dimensional reconstruction of the 70S Escherichia coli ribosome in ice: the distribution of ribosomal RNA. *The Journal of cell biology*. 115:597-605.

Frank, J., M. Radermacher, P. Penczek, J. Zhu, Y. Li, M. Ladjadj, and A. Leith. 1996. SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *Journal of structural biology*. 116:190-199.

Frauenfeld, J., J. Gumbart, E.O. Sluis, S. Funes, M. Gartmann, B. Beatrix, T. Mielke, O. Berninghausen, T. Becker, K. Schulten, and R. Beckmann. 2011. Cryo-EM structure of the ribosome-SecYE complex in the membrane environment. *Nature structural & molecular biology*. 18:614-621.

Gavel, Y., J. Steppuhn, R. Herrmann, and G. von Heijne. 1991. The 'positive-inside rule' applies to thylakoid membrane proteins. *FEBS letters*. 282:41-46.

Gavel, Y., and G. von Heijne. 1992. The distribution of charged amino acids in mitochondrial inner-membrane proteins suggests different modes of membrane integration for nuclearly and mitochondrially encoded proteins. *European journal of biochemistry / FEBS*. 205:1207-1215.

Gilmore, R., P. Walter, and G. Blobel. 1982. Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor. *The Journal of cell biology*. 95:470-477.

Gish, K., and C. Yanofsky. 1995. Evidence suggesting cis action by the TnaC leader peptide in regulating transcription attenuation in the tryptophanase operon of Escherichia coli. *Journal of bacteriology*. 177:7245-7254.

Goder, V., T. Junne, and M. Spiess. 2004. Sec61p contributes to signal sequence orientation according to the positive-inside rule. *Molecular biology of the cell*. 15:1470-1478.

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

Gollnick, P., and C. Yanofsky. 1990. tRNA(Trp) translation of leader peptide codon 12 and other factors that regulate expression of the tryptophanase operon. *Journal of bacteriology*. 172:3100-3107.

Gong, F., K. Ito, Y. Nakamura, and C. Yanofsky. 2001. The mechanism of tryptophan induction of tryptophanase operon expression: tryptophan inhibits release factor-mediated cleavage of TnaC-peptidyl-tRNA(Pro). *Proceedings of the National Academy of Sciences of the United States of America*. 98:8997-9001.

Gong, F., and C. Yanofsky. 2001. Reproducing tna operon regulation in vitro in an S-30 system. Tryptophan induction inhibits cleavage of TnaC peptidyl-tRNA. *The Journal of biological chemistry*. 276:1974-1983.

Gong, F., and C. Yanofsky. 2002a. Analysis of tryptophanase operon expression in vitro: accumulation of TnaC-peptidyl-tRNA in a release factor 2-depleted S-30 extract prevents Rho factor action, simulating induction. *The Journal of biological chemistry*. 277:17095-17100.

Gong, F., and C. Yanofsky. 2002b. Instruction of translating ribosome by nascent peptide. *Science*. 297:1864-1867.

Gorlich, D., and T.A. Rapoport. 1993. Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. *Cell*. 75:615-630.

Grudnik, P., G. Bange, and I. Sinning. 2009. Protein targeting by the signal recognition particle. *Biological chemistry*. 390:775-782.

Halic, M., T. Becker, M.R. Pool, C.M. Spahn, R.A. Grassucci, J. Frank, and R. Beckmann. 2004. Structure of the signal recognition particle interacting with the elongation-arrested ribosome. *Nature*. 427:808-814.

Hansen, J.L., P.B. Moore, and T.A. Steitz. 2003. Structures of five antibiotics bound at the peptidyl transferase center of the large ribosomal subunit. *Journal of molecular biology*. 330:1061-1075.

Hansen, J.L., T.M. Schmeing, P.B. Moore, and T.A. Steitz. 2002. Structural insights into peptide bond formation. *Proceedings of the National Academy of Sciences of the United States of America*. 99:11670-11675.

Henderson, R. 1992. Image contrast in high-resolution electron microscopy of biological macromolecules: TMV in ice. *Ultramicroscopy*. 46:1-18.

Henderson, R. 1995. The potential and limitations of neutrons, electrons and X-rays for atomic resolution microscopy of unstained biological molecules. *Quarterly reviews of biophysics*. 28:171-193.

Hessa, T., H. Kim, K. Bihlmaier, C. Lundin, J. Boekel, H. Andersson, I. Nilsson, S.H. White, and G. von Heijne. 2005. Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature*. 433:377-381.

Hessa, T., N.M. Meindl-Beinker, A. Bernsel, H. Kim, Y. Sato, M. Lerch-Bader, I. Nilsson, S.H. White, and G. von Heijne. 2007. Molecular code for transmembrane-helix recognition by the Sec61 translocon. *Nature*. 450:1026-1030.

Hessa, T., J.H. Reithinger, G. von Heijne, and H. Kim. 2009. Analysis of transmembrane helix integration in the endoplasmic reticulum in S. cerevisiae. *Journal of molecular biology*. 386:1222-1228.

Hildebrand, A., M. Remmert, A. Biegert, and J. Soding. 2009. Fast and accurate automatic structure prediction with HHpred. *Proteins*. 77 Suppl 9:128-132.

Hilz, H., U. Wiegers, and P. Adamietz. 1975. Stimulation of proteinase K action by denaturing agents: application to the isolation of nucleic acids and the degradation of 'masked' proteins. *European journal of biochemistry / FEBS*. 56:103-108.

Hizlan, D., A. Robson, S. Whitehouse, V.A. Gold, J. Vonck, D. Mills, W. Kuhlbrandt, and I. Collinson. 2012. Structure of the SecY complex unlocked by a preprotein mimic. *Cell reports*. 1:21-28.

Hong, H., and J.U. Bowie. 2011. Dramatic destabilization of transmembrane helix interactions by features of natural membrane environments. *Journal of the American Chemical Society*. 133:11389-11398.

Huber, D., N. Rajagopalan, S. Preissler, M.A. Rocco, F. Merz, G. Kramer, and B. Bukau. 2011. SecA interacts with ribosomes in order to facilitate posttranslational translocation in bacteria. *Molecular cell*. 41:343-353.

Hunt, J.F., S. Weinkauf, L. Henry, J.J. Fak, P. McNicholas, D.B. Oliver, and J. Deisenhofer. 2002. Nucleotide control of interdomain interactions in the conformational reaction cycle of SecA. *Science*. 297:2018-2026.

Ismail, N., R. Hedman, N. Schiller, and G. von Heijne. 2012. A biphasic pulling force acts on transmembrane helices during translocon-mediated membrane integration. *Nature structural & molecular biology*. 19:1018-1022.

Ito, K., and S. Chiba. 2013. Arrest peptides: cis-acting modulators of translation. *Annual review of biochemistry*. 82:171-202.

Jha, S.S., and A.A. Komar. 2012. Using SecM arrest sequence as a tool to isolate ribosome bound polypeptides. *Journal of visualized experiments : JoVE*.

Jiang, F., M. Chen, L. Yi, J.W. de Gier, A. Kuhn, and R.E. Dalbey. 2003. Defining the regions of Escherichia coli YidC that contribute to activity. *The Journal of biological chemistry*. 278:48965-48972.

Jin, H., A.C. Kelley, D. Loakes, and V. Ramakrishnan. 2010. Structure of the 70S ribosome bound to release factor 2 and a substrate analog provides insights into catalysis of peptide release. *Proceedings of the National Academy of Sciences of the United States of America*. 107:8593-8598.

Joly, J.C., M.R. Leonard, and W.T. Wickner. 1994. Subunit dynamics in Escherichia coli preprotein translocase. *Proceedings of the National Academy of Sciences of the United States of America*. 91:4703-4707.

Junne, T., T. Schwede, V. Goder, and M. Spiess. 2007. Mutations in the Sec61p channel affecting signal sequence recognition and membrane protein topology. *The Journal of biological chemistry*. 282:33201-33209.

Kamath, A.V., and C. Yanofsky. 1997. Roles of the tnaC-tnaA spacer region and Rho factor in regulating expression of the tryptophanase operon of Proteus vulgaris. *Journal of bacteriology*. 179:1780-1786.

Kedrov, A., I. Kusters, V.V. Krasnikov, and A.J. Driessen. 2011. A single copy of SecYEG is sufficient for preprotein translocation. *The EMBO journal*. 30:4387-4397.

Kedrov, A., M. Sustarsic, J. de Keyzer, J.J. Caumanns, Z.C. Wu, and A.J. Driessen. 2013. Elucidating the native architecture of the YidC: ribosome complex. *Journal of molecular biology*. 425:4112-4124.

Keenan, R.J., D.M. Freymann, P. Walter, and R.M. Stroud. 1998. Crystal structure of the signal sequence binding subunit of the signal recognition particle. *Cell*. 94:181-191.

Kiefer, D., X. Hu, R. Dalbey, and A. Kuhn. 1997. Negatively charged amino acid residues play an active role in orienting the Sec-independent Pf3 coat protein in the Escherichia coli inner membrane. *The EMBO journal*. 16:2197-2204.

Kohler, R., D. Boehringer, B. Greber, R. Bingel-Erlenmeyer, I. Collinson, C. Schaffitzel, and N. Ban. 2009. YidC and Oxa1 form dimeric insertion pores on the translating ribosome. *Molecular cell*. 34:344-353.

Konan, K.V., and C. Yanofsky. 1997. Regulation of the Escherichia coli tna operon: nascent leader peptide control at the tnaC stop codon. *Journal of bacteriology*. 179:1774-1779.

Konan, K.V., and C. Yanofsky. 1999. Role of ribosome release in regulation of tna operon expression in Escherichia coli. *Journal of bacteriology*. 181:1530-1536.

Kucukelbir, A., F.J. Sigworth, and H.D. Tagare. 2014. Quantifying the local resolution of cryo-EM density maps. *Nature methods*. 11:63-65.

Kuhlbrandt, W. 2014. Biochemistry. The resolution revolution. *Science*. 343:1443-1444.

Kuhn, P., B. Weiche, L. Sturm, E. Sommer, F. Drepper, B. Warscheid, V. Sourjik, and H.G. Koch. 2011. The bacterial SRP receptor, SecA and the ribosome use overlapping binding sites on the SecY translocon. *Traffic*. 12:563-578.

Kumazaki, K., S. Chiba, M. Takemoto, A. Furukawa, K. Nishiyama, Y. Sugano, T. Mori, N. Dohmae, K. Hirata, Y. Nakada-Nakura, A.D. Maturana, Y. Tanaka, H. Mori, Y. Sugita, F. Arisaka, K. Ito, R. Ishitani, T. Tsukazaki, and O. Nureki. 2014a. Structural basis of Sec-independent membrane protein insertion by YidC. *Nature*. 509:516-520.

Kumazaki, K., T. Kishimoto, A. Furukawa, H. Mori, Y. Tanaka, N. Dohmae, R. Ishitani, T. Tsukazaki, and O. Nureki. 2014b. Crystal structure of Escherichia coli YidC, a membrane protein chaperone and insertase. *Scientific reports*. 4:7299.

Kumazaki, K., T. Tsukazaki, T. Nishizawa, Y. Tanaka, H.E. Kato, Y. Nakada-Nakura, K. Hirata, Y. Mori, H. Suga, N. Dohmae, R. Ishitani, and O. Nureki. 2014c. Crystallization and preliminary X-ray diffraction analysis of YidC, a membrane-protein chaperone and insertase from Bacillus halodurans. *Acta crystallographica. Section F, Structural biology communications*. 70:1056-1060.

Ladokhin, A.S., and S.H. White. 1999. Folding of amphipathic alpha-helices on membranes: energetics of helix formation by melittin. *Journal of molecular biology*. 285:1363-1369.

Leapman, R.D., and S. Sun. 1995. Cryo-electron energy loss spectroscopy: observations on vitrified hydrated specimens and radiation damage. *Ultramicroscopy*. 59:71-79.

Liao, M., E. Cao, D. Julius, and Y. Cheng. 2013. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature*. 504:107-112.

Lu, J., and C. Deutsch. 2005. Secondary structure formation of a transmembrane segment in Kv channels. *Biochemistry*. 44:8230-8243.

Lu, P., X.C. Bai, D. Ma, T. Xie, C. Yan, L. Sun, G. Yang, Y. Zhao, R. Zhou, S.H. Scheres, and Y. Shi. 2014. Three-dimensional structure of human gamma-secretase. *Nature*. 512:166-170.

Lycklama a Nijeholt, J.A., Z.C. Wu, and A.J. Driessen. 2011. Conformational dynamics of the plug domain of the SecYEG protein-conducting channel. *The Journal of biological chemistry*. 286:43881-43890.

Mackenzie, K.R. 2006. Folding and stability of alpha-helical integral membrane proteins. *Chemical reviews*. 106:1931-1977.

Mackinnon, A.L., V.O. Paavilainen, A. Sharma, R.S. Hegde, and J. Taunton. 2014. An allosteric Sec61 inhibitor traps nascent transmembrane helices at the lateral gate. *eLife*. 3:e01483.

Martinez, A.K., E. Gordon, A. Sengupta, N. Shirole, D. Klepacki, B. Martinez-Garriga, L.M. Brown, M.J. Benedik, C. Yanofsky, A.S. Mankin, N. Vazquez-Laslop, M.S. Sachs, and L.R. Cruz-Vera. 2014. Interactions of the TnaC nascent peptide with rRNA in the exit tunnel enable the ribosome to respond to free tryptophan. *Nucleic acids research*. 42:1245-1256.

McMullan, G., S. Chen, R. Henderson, and A.R. Faruqi. 2009a. Detective quantum efficiency of electron area detectors in electron microscopy. *Ultramicroscopy*. 109:1126-1143.

McMullan, G., A.T. Clark, R. Turchetta, and A.R. Faruqi. 2009b. Enhanced imaging in low dose electron microscopy using electron counting. *Ultramicroscopy*. 109:1411-1416.

McMullan, G., A.R. Faruqi, R. Henderson, N. Guerrini, R. Turchetta, A. Jacobs, and G. van Hoften. 2009c. Experimental observation of the improvement in MTF from backthinning a CMOS direct electron detector. *Ultramicroscopy*. 109:1144-1147.

Menetret, J.F., J. Schaletzky, W.M. Clemons, Jr., A.R. Osborne, S.S. Skanland, C. Denison, S.P. Gygi, D.S. Kirkpatrick, E. Park, S.J. Ludtke, T.A. Rapoport, and C.W. Akey. 2007. Ribosome binding of a single copy of the SecY complex: implications for protein translocation. *Molecular cell*. 28:1083-1092.

Meyer, T.H., J.F. Menetret, R. Breitling, K.R. Miller, C.W. Akey, and T.A. Rapoport. 1999. The bacterial SecY/E translocation complex forms channel-like structures similar to those of the eukaryotic Sec61p complex. *Journal of molecular biology*. 285:1789-1800.

Mitra, K., C. Schaffitzel, T. Shaikh, F. Tama, S. Jenni, C.L. Brooks, 3rd, N. Ban, and J. Frank. 2005. Structure of the E. coli protein-conducting channel bound to a translating ribosome. *Nature*. 438:318-324.

Montoya, G., C. Svensson, J. Luirink, and I. Sinning. 1997. Crystal structure of the NG domain from the signal-recognition particle receptor FtsY. *Nature*. 385:365-368.

Murphy, C.K., and J. Beckwith. 1994. Residues essential for the function of SecE, a membrane component of the Escherichia coli secretion apparatus, are located in a conserved cytoplasmic region. *Proceedings of the National Academy of Sciences of the United States of America*. 91:2557-2561.

Neumann-Haefelin, C., U. Schafer, M. Muller, and H.G. Koch. 2000. SRP-dependent co-translational targeting and SecA-dependent translocation analyzed as individual steps in the export of a bacterial protein. *The EMBO journal*. 19:6419-6426.

Nissen, P., J. Hansen, N. Ban, P.B. Moore, and T.A. Steitz. 2000. The structural basis of ribosome activity in peptide bond synthesis. *Science*. 289:920-930.

Norousi, R., S. Wickles, C. Leidig, T. Becker, V.J. Schmid, R. Beckmann, and A. Tresch. 2013. Automatic post-picking using MAPPOS improves particle image detection from cryo-EM micrographs. *Journal of structural biology*. 182:59-66.

Nouwen, N., and A.J. Driessen. 2002. SecDFyajC forms a heterotetrameric complex with YidC. *Molecular microbiology*. 44:1397-1405.

Ojemalm, K., S.C. Botelho, C. Studle, and G. von Heijne. 2013. Quantitative analysis of SecYEG-mediated insertion of transmembrane alpha-helices into the bacterial inner membrane. *Journal of molecular biology*. 425:2813-2822.

Or, E., D. Boyd, S. Gon, J. Beckwith, and T. Rapoport. 2005. The bacterial ATPase SecA functions as a monomer in protein translocation. *The Journal of biological chemistry*. 280:9097-9105.

Osborne, A.R., W.M. Clemons, Jr., and T.A. Rapoport. 2004. A large conformational change of the translocation ATPase SecA. *Proceedings of the National Academy of Sciences of the United States of America*. 101:10937-10942.

Paetzel, M. 2014. Structure and mechanism of Escherichia coli type I signal peptidase. *Biochimica et biophysica acta*. 1843:1497-1508.

Papanikolau, Y., M. Papadovasilaki, R.B. Ravelli, A.A. McCarthy, S. Cusack, A. Economou, and K. Petratos. 2007. Structure of dimeric SecA, the Escherichia coli preprotein translocase motor. *Journal of molecular biology*. 366:1545-1557.

Park, E., J.F. Menetret, J.C. Gumbart, S.J. Ludtke, W. Li, A. Whynot, T.A. Rapoport, and C.W. Akey. 2014. Structure of the SecY channel during initiation of protein translocation. *Nature*. 506:102-106.

Park, E., and T.A. Rapoport. 2011. Preserving the membrane barrier for small molecules during bacterial protein translocation. *Nature*. 473:239-242.

Park, E., and T.A. Rapoport. 2012a. Bacterial protein translocation requires only one copy of the SecY complex in vivo. *The Journal of cell biology*. 198:881-893.

Park, E., and T.A. Rapoport. 2012b. Mechanisms of Sec61/SecY-mediated protein translocation across membranes. *Annual review of biophysics*. 41:21-40.

Pavlov, M.Y., R.E. Watts, Z. Tan, V.W. Cornish, M. Ehrenberg, and A.C. Forster. 2009. Slow peptide bond formation by proline and other N-alkylamino acids in translation. *Proceedings of the National Academy of Sciences of the United States of America*. 106:50-54.

Pettersen, E.F., T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, and T.E. Ferrin. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of computational chemistry*. 25:1605-1612.

Pfeffer, S., F. Brandt, T. Hrabe, S. Lang, M. Eibauer, R. Zimmermann, and F. Forster. 2012. Structure and 3D arrangement of endoplasmic reticulum membrane-associated ribosomes. *Structure*. 20:1508-1518.

Pfeffer, S., J. Dudek, M. Gogala, S. Schorr, J. Linxweiler, S. Lang, T. Becker, R. Beckmann, R. Zimmermann, and F. Forster. 2014. Structure of the mammalian oligosaccharyl-transferase complex in the native ER protein translocon. *Nature communications*. 5:3072.

Plath, K., W. Mothes, B.M. Wilkinson, C.J. Stirling, and T.A. Rapoport. 1998. Signal sequence recognition in posttranslational protein transport across the yeast ER membrane. *Cell*. 94:795-807.

Popot, J.L., and D.M. Engelman. 2000. Helical membrane protein folding, stability, and evolution. *Annual review of biochemistry*. 69:881-922.

Powers, T., and P. Walter. 1997. Co-translational protein targeting catalyzed by the Escherichia coli signal recognition particle and its receptor. *The EMBO journal*. 16:4880-4886.

Prinz, W.A., D.H. Boyd, M. Ehrmann, and J. Beckwith. 1998. The protein translocation apparatus contributes to determining the topology of an integral membrane protein in Escherichia coli. *The Journal of biological chemistry*. 273:8419-8424.

Ramirez, U.D., G. Minasov, P.J. Focia, R.M. Stroud, P. Walter, P. Kuhn, and D.M. Freymann. 2002. Structural basis for mobility in the 1.1 A crystal structure of the NG domain of Thermus aquaticus Ffh. *Journal of molecular biology*. 320:783-799.

Reckel, S., D. Gottstein, J. Stehle, F. Lohr, M.K. Verhoefen, M. Takeda, R. Silvers, M. Kainosho, C. Glaubitz, J. Wachtveitl, F. Bernhard, H. Schwalbe, P. Guntert, and V. Dotsch. 2011. Solution NMR structure of proteorhodopsin. *Angewandte Chemie*. 50:11942-11946.

Reyes, C.L., E. Rutenber, P. Walter, and R.M. Stroud. 2007. X-ray structures of the signal recognition particle receptor reveal targeting cycle intermediates. *PloS one*. 2:e607.

Rutz, C., W. Rosenthal, and R. Schulein. 1999. A single negatively charged residue affects the orientation of a membrane protein in the inner membrane of Escherichia coli only when it is located adjacent to a transmembrane domain. *The Journal of biological chemistry*. 274:33757-33763.

Saaf, A., M. Monne, J.W. de Gier, and G. von Heijne. 1998. Membrane topology of the 60-kDa Oxa1p homologue from Escherichia coli. *The Journal of biological chemistry*. 273:30415-30418.

Sachelaru, I., N.A. Petriman, R. Kudva, P. Kuhn, T. Welte, B. Knapp, F. Drepper, B. Warscheid, and H.G. Koch. 2013. YidC occupies the lateral gate of the SecYEG translocon and is sequentially displaced by a nascent membrane protein. *The Journal of biological chemistry*. 288:16295-16307.

Sadlish, H., D. Pitonzo, A.E. Johnson, and W.R. Skach. 2005. Sequential triage of transmembrane segments by Sec61alpha during biogenesis of a native multispanning membrane protein. *Nature structural & molecular biology*. 12:870-878.

Saller, M.J., Z.C. Wu, J. de Keyzer, and A.J. Driessen. 2012. The YidC/Oxa1/Alb3 protein family: common principles and distinct features. *Biological chemistry*. 393:1279-1290.

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

Saraogi, I., and S.O. Shan. 2014. Co-translational protein targeting to the bacterial membrane. *Biochimica et biophysica acta*. 1843:1433-1441.

Scheres, S.H., and S. Chen. 2012. Prevention of overfitting in cryo-EM structure determination. *Nature methods*. 9:853-854.

Schmeing, T.M., K.S. Huang, D.E. Kitchen, S.A. Strobel, and T.A. Steitz. 2005a. Structural insights into the roles of water and the 2' hydroxyl of the P site tRNA in the peptidyl transferase reaction. *Molecular cell*. 20:437-448.

Schmeing, T.M., K.S. Huang, S.A. Strobel, and T.A. Steitz. 2005b. An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA. *Nature*. 438:520-524.

Scotti, P.A., M.L. Urbanus, J. Brunner, J.W. de Gier, G. von Heijne, C. van der Does, A.J. Driessen, B. Oudega, and J. Luirink. 2000. YidC, the Escherichia coli homologue of mitochondrial Oxa1p, is a component of the Sec translocase. *The EMBO journal*. 19:542-549.

Seidelt, B., C.A. Innis, D.N. Wilson, M. Gartmann, J.P. Armache, E. Villa, L.G. Trabuco, T. Becker, T. Mielke, K. Schulten, T.A. Steitz, and R. Beckmann. 2009. Structural insight into nascent polypeptide chain-mediated translational stalling. *Science*. 326:1412-1415.

Seppala, S., J.S. Slusky, P. Lloris-Garcera, M. Rapp, and G. von Heijne. 2010. Control of membrane protein topology by a single C-terminal residue. *Science*. 328:1698-1700.

Serek, J., G. Bauer-Manz, G. Struhalla, L. van den Berg, D. Kiefer, R. Dalbey, and A. Kuhn. 2004. Escherichia coli YidC is a membrane insertase for Sec-independent proteins. *The EMBO journal*. 23:294-301.

Shan, S.O., and P. Walter. 2005. Molecular crosstalk between the nucleotide specificity determinant of the SRP GTPase and the SRP receptor. *Biochemistry*. 44:6214-6222.

Shao, S., and R.S. Hegde. 2011. Membrane protein insertion at the endoplasmic reticulum. *Annual review of cell and developmental biology*. 27:25-56.

Shen, K., S. Arslan, D. Akopian, T. Ha, and S.O. Shan. 2012. Activated GTPase movement on an RNA scaffold drives co-translational protein targeting. *Nature*. 492:271-275.

Skach, W.R. 2009. Cellular mechanisms of membrane protein folding. *Nature structural & molecular biology*. 16:606-612.

Smith, M.A., W.M. Clemons, Jr., C.J. DeMars, and A.M. Flower. 2005. Modeling the effects of prl mutations on the Escherichia coli SecY complex. *Journal of bacteriology*. 187:6454-6465.

Soding, J., A. Biegert, and A.N. Lupas. 2005. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic acids research*. 33:W244-248.

Taufik, I., A. Kedrov, M. Exterkate, and A.J. Driessen. 2013. Monitoring the activity of single translocons. *Journal of molecular biology*. 425:4145-4153.

Trabuco, L.G., C.B. Harrison, E. Schreiner, and K. Schulten. 2010. Recognition of the regulatory nascent chain TnaC by the ribosome. *Structure*. 18:627-637.

Tsukazaki, T., H. Mori, S. Fukai, R. Ishitani, T. Mori, N. Dohmae, A. Perederina, Y. Sugita, D.G. Vassylyev, K. Ito, and O. Nureki. 2008. Conformational transition of Sec machinery inferred from bacterial SecYE structures. *Nature*. 455:988-991.

Ulmschneider, M.B., J.P. Ulmschneider, N. Schiller, B.A. Wallace, G. von Heijne, and S.H. White. 2014. Spontaneous transmembrane helix insertion thermodynamically mimics translocon-guided insertion. *Nature communications*. 5:4863.

Urbanus, M.L., P.A. Scotti, L. Froderberg, A. Saaf, J.W. de Gier, J. Brunner, J.C. Samuelson, R.E. Dalbey, B. Oudega, and J. Luirink. 2001. Sec-dependent membrane protein insertion: sequential interaction of nascent FtsQ with SecY and YidC. *EMBO reports*. 2:524-529.

van de Vossenberg, J.L., S.V. Albers, C. van der Does, A.J. Driessen, and W. van Klompenburg. 1998. The positive inside rule is not determined by the polarity of the delta psi (transmembrane electrical potential). *Molecular microbiology*. 29:1125-1127.

Van den Berg, B., W.M. Clemons, Jr., I. Collinson, Y. Modis, E. Hartmann, S.C. Harrison, and T.A. Rapoport. 2004. X-ray structure of a protein-conducting channel. *Nature*. 427:36-44.

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. *The Journal of cell biology*. 165:213-222.

Vassylyev, D.G., H. Mori, M.N. Vassylyeva, T. Tsukazaki, Y. Kimura, T.H. Tahirov, and K. Ito. 2006. Crystal structure of the translocation ATPase SecA from Thermus thermophilus reveals a parallel, head-to-head dimer. *Journal of molecular biology*. 364:248-258.

Vitrac, H., M. Bogdanov, and W. Dowhan. 2013. In vitro reconstitution of lipid-dependent dual topology and postassembly topological switching of a membrane protein. *Proceedings of the National Academy of Sciences of the United States of America*. 110:9338-9343.

von Heijne, G. 1986. Towards a comparative anatomy of N-terminal topogenic protein sequences. *Journal of molecular biology*. 189:239-242.

von Heijne, G. 1990. The signal peptide. *The Journal of membrane biology*. 115:195-201.

von Heijne, G. 1992. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *Journal of molecular biology*. 225:487-494.

von Heijne, G. 2006. Membrane-protein topology. *Nature reviews. Molecular cell biology*. 7:909-918.

Voorhees, R.M., I.S. Fernandez, S.H. Scheres, and R.S. Hegde. 2014. Structure of the mammalian ribosome-Sec61 complex to 3.4 A resolution. *Cell*. 157:1632-1643.

Voss, N.R., M. Gerstein, T.A. Steitz, and P.B. Moore. 2006. The geometry of the ribosomal polypeptide exit tunnel. *Journal of molecular biology*. 360:893-906.

Walter, P., and G. Blobel. 1980. Purification of a membrane-associated protein complex required for protein translocation across the endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America*. 77:7112-7116.

White, S.H., and W.C. Wimley. 1998. Hydrophobic interactions of peptides with membrane interfaces. *Biochimica et biophysica acta*. 1376:339-352.

White, S.H., and W.C. Wimley. 1999. Membrane protein folding and stability: physical principles. *Annual review of biophysics and biomolecular structure*. 28:319-365.

Wickles, S., A. Singharoy, J. Andreani, S. Seemayer, L. Bischoff, O. Berninghausen, J. Soeding, K. Schulten, E.O. van der Sluis, and R. Beckmann. 2014. A structural model of the active ribosome-bound membrane protein insertase YidC. *eLife*. 3:e03035.

Wilson, D.N., and R. Beckmann. 2011. The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. *Current opinion in structural biology*. 21:274-282.

Wimley, W.C., K. Hristova, A.S. Ladokhin, L. Silvestro, P.H. Axelsen, and S.H. White. 1998. Folding of beta-sheet membrane proteins: a hydrophobic hexapeptide model. *Journal of molecular biology*. 277:1091-1110.

Wimley, W.C., and S.H. White. 1996. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nature structural biology*. 3:842-848.

Wong, W., X.C. Bai, A. Brown, I.S. Fernandez, E. Hanssen, M. Condron, Y.H. Tan, J. Baum, and S.H. Scheres. 2014. Cryo-EM structure of the Plasmodium falciparum 80S ribosome bound to the anti-protozoan drug emetine. *eLife*. 3.

Woodbury, R.L., S.J. Hardy, and L.L. Randall. 2002. Complex behavior in solution of homodimeric SecA. *Protein science : a publication of the Protein Society*. 11:875-882.

Wu, Z.C., J. de Keyzer, A. Kedrov, and A.J. Driessen. 2012. Competitive binding of the SecA ATPase and ribosomes to the SecYEG translocon. *The Journal of biological chemistry*. 287:7885-7895.

Xie, K., T. Hessa, S. Seppala, M. Rapp, G. von Heijne, and R.E. Dalbey. 2007. Features of transmembrane segments that promote the lateral release from the translocase into the lipid phase. *Biochemistry*. 46:15153-15161.

Xie, K., D. Kiefer, G. Nagler, R.E. Dalbey, and A. Kuhn. 2006. Different regions of the nonconserved large periplasmic domain of Escherichia coli YidC are involved in the SecF interaction and membrane insertase activity. *Biochemistry*. 45:13401-13408.

Yang, R., L.R. Cruz-Vera, and C. Yanofsky. 2009. 23S rRNA nucleotides in the peptidyl transferase center are essential for tryptophanase operon induction. *Journal of bacteriology*. 191:3445-3450.

Yi, L., N. Celebi, M. Chen, and R.E. Dalbey. 2004. Sec/SRP requirements and energetics of membrane insertion of subunits a, b, and c of the Escherichia coli F1F0 ATP synthase. *The Journal of biological chemistry*. 279:39260-39267.

Zhang, X., and S.O. Shan. 2014. Fidelity of cotranslational protein targeting by the signal recognition particle. *Annual review of biophysics*. 43:381-408.

Zhou, F.X., H.J. Merianos, A.T. Brunger, and D.M. Engelman. 2001. Polar residues drive association of polyleucine transmembrane helices. *Proceedings of the National Academy of Sciences of the United States of America*. 98:2250-2255.

Zimmer, J., Y. Nam, and T.A. Rapoport. 2008. Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature*. 455:936-943.

# 8. Curriculum Vitae

| | |
|---|---|
| Name | Lukas Sebastian Bischoff |
| Date of birth | 20.08.1985 |
| Place of birth | Graefelfing, Germany |
| Nationality | German |

## Education

| | |
|---|---|
| Since April 2011 | **PhD Student;** <br> **Member of the International Max Planck Research School for Molecular and Cellular Life Sciences (IMPRS-LS)** <br> Laboratory of Prof. R. Beckmann, <br> Ludwig-Maximilians University Munich, Germany |
| Apr 2009 - Dec 2010 | **Master of Science in Biochemistry** <br> Ludwig-Maximilians University Munich, Germany |
| Oct 2005 - Mar 2009 | **Bachelor of Science in Chemistry and Biochemistry** <br> Ludwig-Maximilians University Munich, Germany |
| Sept 2004 - May 2005 | **Alternative Civilian Service** <br> Bavarian Red Cross, Starnberg |
| Sep 1995 - Jun 2004 | **Gymnasium** <br> Gymnasium Starnberg |

## Publications

Visualization of a polytopic membrane protein during SecY-mediated membrane insertion.
**Bischoff L**, Wickles S, Berninghausen O, van der Sluis EO, Beckmann R.
Nature Communications. 2014 Jun 10;5:4103. doi: 10.1038/ncomms5103.

Molecular Basis for the Ribosome Functioning as an L-Tryptophan Sensor.
**Bischoff L**, Berninghausen O, Beckmann R.
Cell Reports. 2014 Oct 7. pii: S2211-1247(14)00779-7. doi: 10.1016/j.celrep.2014.09.011.

A structural model of the active ribosome-bound membrane protein insertase YidC.
Wickles S, Singharoy A, Andreani J, Seemayer S, **Bischoff L**, Berninghausen O, Soeding J, Schulten K, van der Sluis EO, Beckmann R.
Elife. 2014 Jul 10;3:e03035. doi: 10.7554/eLife.03035.

Dynamic behavior of Trigger Factor on the ribosome
Deeng J, Chan KY, van der Sluis EO, **Bischoff L**, Berninghausen O, Han W, Gumbart J, Schulten K, Beatrix B. and Beckmann R.
*SUBMITTED*

## International Conferences

Kuo Symposium on 3D cryo-EM Molecular Imaging,
26.07.14 – 30.07.14, Shanghai, China, poster presentation

Gordon Research Conference: 'Three-Dimensional Electron Microscopy',
22.06.14 – 27.06.14, Girona, Spain, poster presentation

Gordon Research Conference: 'Protein Transport Across Cell Membranes',
09.03.14 – 14.03.14 Galveston, USA, oral and poster presentation

EMBO Conference Series: 'From Structure to Function of Translocation Machines', 13.04.13 – 17.04.13, Dubrovnik, Croatia, poster presentation

Gordon Research Conference: 'Protein Transport Across Cell Membranes',
11.03.12 – 16.03.12 Galveston, USA, oral and poster presentation

## Honors and Awards

2014   Award for outstanding scientific performance at the 'Kuo Symposium on 3D cryo-EM Molecular Imaging', Shanghai, China

2011   Award of the 'Dr. Klaus Römer-Stiftung' for excellent performance during the Master Thesis

# 9. Acknowledgement