Membrane targeting and insertion of the sensor protein KdpD and the C-tail anchored protein SciP of *Escherichia coli*

Dissertation zur Erlangung des Doktorgrades
der Naturwissenschaften (Dr. rer. nat.)

Fakultät Naturwissenschaften
Universität Hohenheim

Institut für Mikrobiologie

vorgelegt von
Eva Proß
aus
Herrenberg

2019

Dekan:	Prof. Dr. Uwe Beifuß
1. berichtende Person:	Prof. Dr. Andreas Kuhn
2. berichtende Person:	Prof. Dr. Heinz Breer
3. Prüfer:	Prof. Dr. Andreas Schaller

Eingereicht am:	15.08.2019
Mündliche Prüfung am:	18.12.2019

My heart *„SciPs SciPs“* a beat.

Olly Murs

List of publications included in this thesis

- Pross E and Kuhn A (2019) The SRP signal sequence of KdpD. *Sci Rep* 9 (1): 8717

- Pross E, Soussoula L, Seitl I, Lupo D, Kuhn A. (2016) Membrane Targeting and Insertion of the C-tail Protein SciP. *J Mol Biol* 428 (20): 4218-4227

Other publications

- Spann D, Pross E, Chen Y, Dalbey RE, Kuhn A. (2018) Each protomer of a dimeric YidC functions as a single membrane insertase. *Sci Rep* 8 (1): 589

Table of contents

1

GENERAL INTRODUCTION

1.1 Protein sorting

Bacterial cells, even if they seem simple compared to eukaryotic cells, are subdivided into compartments. Gram-negative bacteria consist of four cellular compartments, the cytosol, the inner and the outer membrane and the periplasmic space (periplasm) which is formed between the two membranes (Paramasivam and Linke 2011). All these compartments include several proteins fulfilling their biological functions (Fekkes and Driessen 1999). The challenge of compartmentalization is that all proteins are synthesized by the ribosomes in the cytoplasm. Subsequently they have to be sorted and transported to the different compartments to reach their corresponding target location (Danese and Silhavy 1998). In addition, numerous proteins have their specific function outside of the cell, therefore they have to be secreted across the cell envelope to the extracellular medium. Different molecular mechanisms do exist in the cell to target the proteins to their final destination. This chapter focuses on the various pathways in Gram-negative bacteria, especially in *Escherichia coli* (*E. coli*), for proteins that have to be delivered to or across the plasma membrane.

1.2 Targeting and insertion pathways in Gram-negative bacteria

Proteins which final destination is the cell surface of Gram-negative bacteria can be delivered either in a post- or co-translational manner (Fig. 1). Intrinsic signal elements of the protein chain are recognized and therefore are responsible for the choice of their specific delivery pathway. In the post-translational way, fully synthesized proteins are targeted to the cell envelope to be either inserted into or translocated across the inner membrane. The post-translational export of already folded proteins into the periplasm occurs via the Twin-arginine translocase (Tat), located in the inner membrane and the proton motive force (pmf) (Fig. 1a). This mechanism can also involve cytoplasmic chaperones like DnaK/DnaJ or GroEL/GroES. Unfolded proteins which are intended for the export into the periplasm are targeted by the cytoplasmic chaperone SecB to the membrane embedded SecYEG translocon, associated with the ATPase SecA (Fig. 1a). SecYEG is also able to cooperate with the membrane insertase YidC, located in the inner membrane. SecYEG is not only involved in post-translational export of proteins into the periplasm but also in co-translational insertion of inner membrane proteins. In this case, the insertion is coupled to translation by the ribosome. Co-

translational delivery to the membrane depends on the signal recognition particle (SRP) and its membrane-associated receptor FtsY (Fig. 1b). The co-translational insertion of SRP-dependent membrane proteins by the SecYEG translocon can involve the membrane insertase YidC. However, YidC is also able to react as an independent insertase without SecYEG. This “YidC only” pathway can cooperate with the SRP system for membrane targeting (Fig. 1b) (Cross et al. 2009).

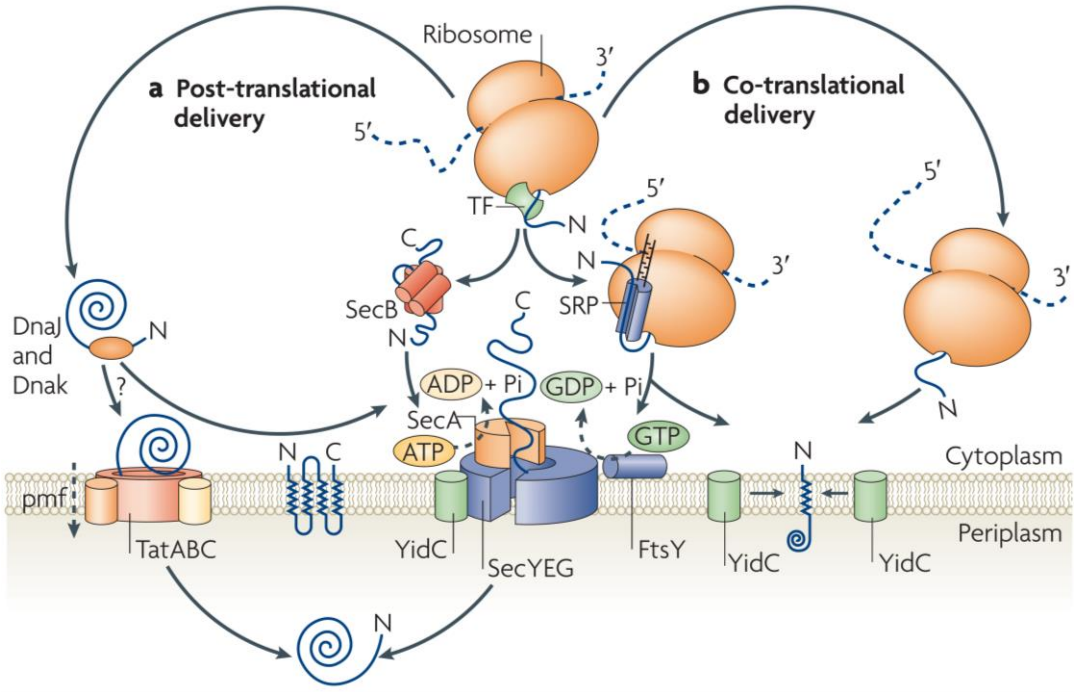


Figure 1: Protein traffic pathways in prokaryotes. Newly synthesized proteins can be targeted either in a post- or co-translational way to the membrane. (a) Post-translational translocation of folded proteins involves chaperones like DnaK/DnaJ or GroEL/GroES and occurs via the Tat translocase and the pmf. Unfolded proteins are targeted post-translationally by the SecB/SecA pathway to the Sec translocon to be inserted under ATP hydrolysis. (b) Co-translational targeting to the Sec translocon is mediated by SRP and its receptor FtsY. The insertion / translocation requires first GTP and further the elongation of the nascent chain. Both, in post- and co-translational insertion / translocation, SecYEG can cooperate with the YidC insertase. However, YidC can work as an independent insertase and thereby a cooperation with the SRP pathway is also possible (Cross et al. 2009).

1.3 Post-translational delivery of folded proteins by the Tat system

The export of fully synthesized proteins into the periplasm can occur via the Tat system localized in the inner membrane. The Tat translocase not only exists in bacteria but also in archaea and chloroplasts of plants (Patel et al. 2014). The peculiarity of this translocase is that already completely folded proteins are translocated across the inner

membrane (Patel et al. 2014). These proteins are often cofactor-containing redox enzymes (Palmer et al. 2005). Because unfolded proteins are incompatible to Tat-dependent translocation there must be a folding quality check point prior to the targeting to the translocase. Not fully folded substrates are prohibited for targeting by chaperone-like proteins encoded within the operon of the respective substrate (Turner et al. 2004; Lee et al. 2006). Additional cytoplasmic chaperones like GroEL/GroES or DnaK/DnaJ/GrpE can assist in the folding or are required to prevent the aggregation of the fully synthesized precursor (Lee et al. 2006). The substrates of the Tat system are recognized and targeted by a N-terminal signal peptide in the precursor sequence including the characteristic twin arginine (RR) motif (Alami et al. 2003; Patel et al. 2014). The Tat system in *E. coli* consists of four inner membrane proteins. TatB and TatC act as receptor complex whereas TatA forms a translocation pore during pmf-dependent oligomerization. TatE probably regulates the TatABC translocase (Dalbey and Kuhn 2012; Cline 2015; Eimer et al. 2015).

1.4 Post-translational delivery to the Sec translocon

The Sec translocon in the inner membrane can do both, insert proteins into or translocate them across the inner membrane. As well as the Tat system, the essential Sec pathway is ubiquitous in the three domains of life and therefore it is found in prokaryotes, archaea and the endoplasmatic reticulum (ER) of eukaryotes (Natale et al. 2008). The Sec translocase of *E. coli* consists of the proteins SecY, SecE and SecG. This stable, heterotrimeric complex is located in the inner membrane (Driessen and Nouwen 2008). The first crystal structure of the protein-conducting channel was solved in 2004 from the archaeon *Methanocaldococcus jannaschii* (former nomenclature *Methanococcus jannaschii)* (Fig. 2) (van den Berg et al. 2004). It is composed of the subunits SecY, SecE and Secβ (SecG in *E. coli*) (van den Berg et al. 2004; Driessen and Nouwen 2008). At that time, only a low-resolution cryo-EM structure of SecYEG from *E. coli* was available for comparison (Breyton et al. 2002) showing that they match in the main features (van den Berg et al. 2004). The channel is built-up of ten transmembrane domains (TMDs) of SecY organized in two halves (TMD 1-5 and TMD 6-10). The pore ring, composed of hydrophobic amino acid (aa) residues is closed by a short α-helical structure (TMD 2a), the movable plug, facing the periplasmic side (van den Berg et al. 2004; Driessen and Nouwen 2008). The channel is opened to the lipid

bilayer forming the so called lateral gate by TMD 2b, TMD 3, TMD 7 and TMD 8 (van den Berg et al. 2004; Kuhn et al. 2017). Differences in the two structures arise from the proteins SecE and SecG. SecE spanning the membrane three times in *E. coli* (single-spanning subunit in *M. jannaschii*) functions as a molecular clamp to stick the two halves of SecY together (van den Berg et al. 2004; Driessen and Nouwen 2008). SecG with two TMDs in *E. coli* (single-spanning β-subunit in *M. jannaschii*) only shows restricted contacts to SecY. The cytosolic loops 4-6 of SecY function as a receptor for proteins involved in targeting the proteins to the translocon (Kudva et al. 2013).

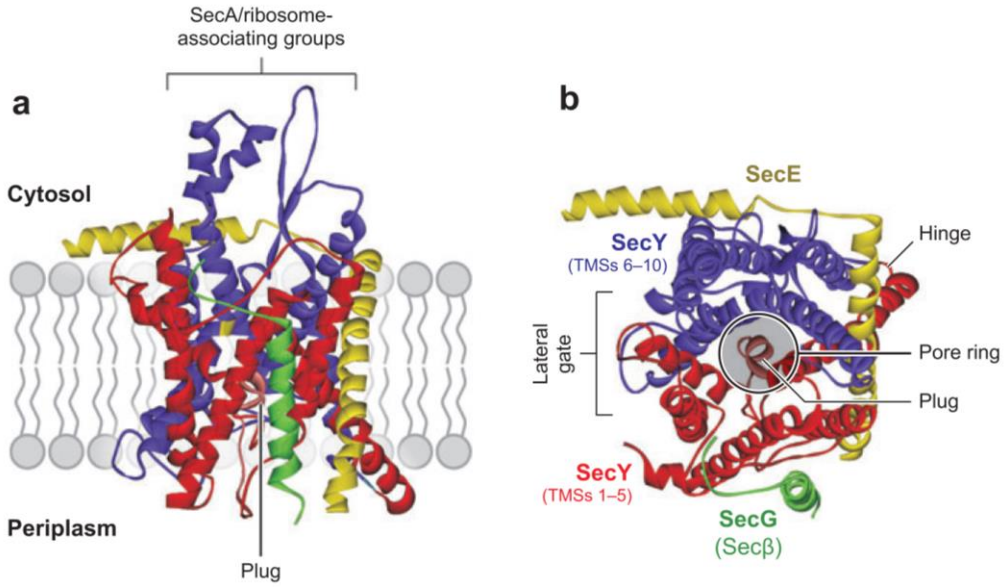


Figure 2: Crystal structure of the SecYEG channel from *M. jannaschii*. (a) Arrangement of the channel in the lipid bilayer and (b) the view from the cytosol. SecY consists of 10 TMDs organized in two-halves (TMD 1-5 in red, TMD 6-10 in blue) forming the translocation channel. The channel is opened towards the bilayer via the lateral gate. The pore ring is lined by hydrophobic amino acids and closed towards the periplasm by the short flexible α-helical TMD 2a (plug). The cytosolic loops 4-6 of SecY are involved in the binding of Sec-associated proteins. The single-spanning membrane protein SecE (in yellow) functions as a molecular clamp whereas the single-spanning membrane protein SecG (Secβ) is peripherally attached. The protein-conducting channel shows an hourglass-like structure in the membrane that can be widened during translocation resulting in the opening towards the periplasm by movement of the plug (Driessen and Nouwen 2008).

In contrast to the export of completely folded proteins via the Tat system, the Sec system translocates fully synthesized but unfolded proteins. This mechanism is used by the majority of the proteins that have to be translocated into the periplasm (Dalbey and Kuhn 2012). Most of the exported proteins are synthesized as preproteins with a cleavable N-terminal signal peptide (Heijne 1990). The signal peptide has a tripartite structure consisting of the N-, H- and C-region. The amino terminal N-region is

positively charged and 1-5 amino acids long, followed by the central hydrophobic region of 7-15 amino acids and subsequently the carboxy terminal polar region of 3-7 amino acids. The C-region also includes the signal peptide cleavage site at position -3 and -1 (Heijne 1990). After targeting to the membrane and during the subsequent translocation the signal peptide is cleaved off by the signal peptidase resulting in the mature protein (Heijne 1990; Danese and Silhavy 1998). In most cases, the post-translational delivery to the membrane-embedded translocase involves the cytoplasmic chaperone SecB and the membrane-associated ATPase SecA, providing an additional energy source to the pmf (Fig. 3B) (Geller et al. 1986; Veenendaal et al. 2004).

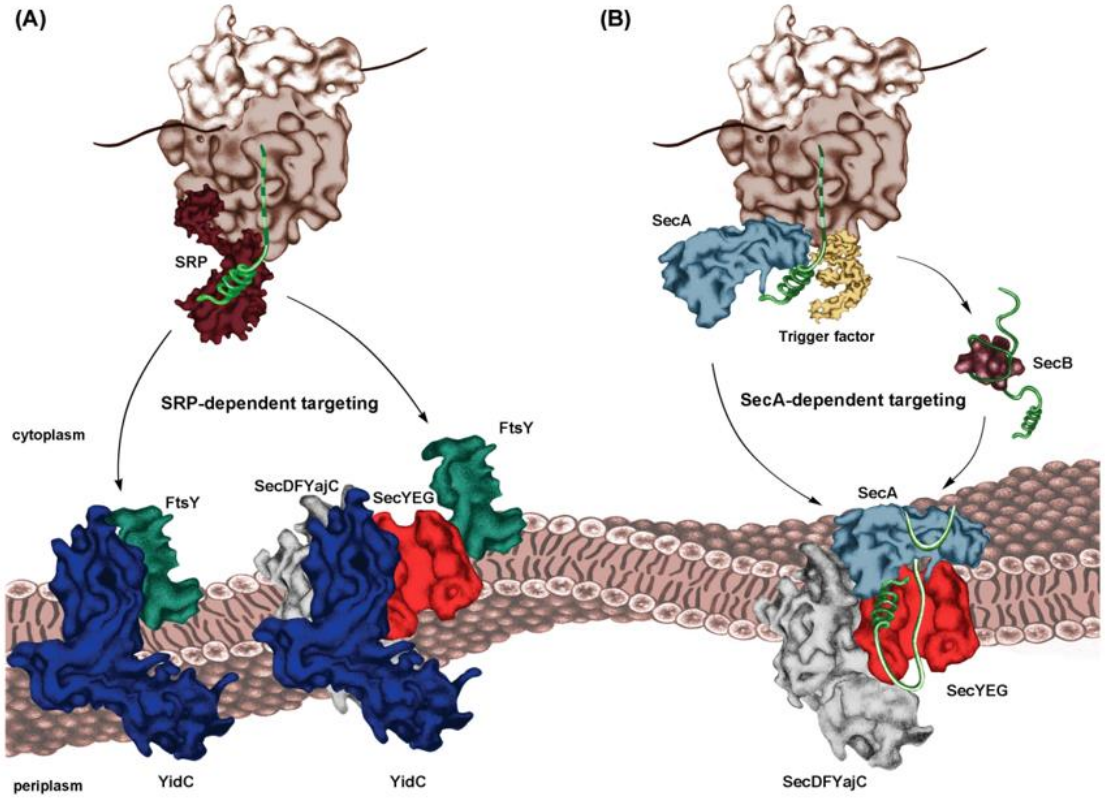


Figure 3: Targeting pathways to the Sec translocon. (A) Co-translational targeting via SRP and FtsY. The signal sequence of an emerging nascent chain is recognized by SRP and targeted to the receptor FtsY. GTP hydrolysis results in the delivery of the nascent chain to the Sec translocon and further to the dissociation of the targeting complex. SRP is also able to interact with the YidC insertase working independently of SecYEG. (B) Delivery to the Sec translocon via SecB and SecA. The nascent chain is first bound by the ribosome-associated chaperone trigger factor, holding the nascent chain in a translocation competent state. Fully synthesized proteins are bound by the cytoplasmic chaperone SecB and are then targeted to the membrane by the interaction with the membrane-associated ATPase SecA. Proteins that do not tend to fold fast or even aggregate can be directly bound by the cytoplasmic form of SecA and targeted to the Sec translocon. Cycles of ATP hydrolysis and the pmf pushes the preprotein through the SecYEG channel (size ratios not taken into account) (Steinberg et al. 2018).

During delivery to SecYEG, it must be guaranteed that the fully synthesized preproteins do not fold up and stay in a translocation competent conformation. During translation this is achieved by binding of the ribosome-associated chaperone trigger factor while the nascent chain has reached a length of about 110 amino acids (Lill et al. 1988; Oh et al. 2011). To keep the fully synthesized precursors stable and unfolded in the cytosol, some are bound by the homotetrameric chaperone SecB (Randall et al. 1998; Driessen and Nouwen 2008; Kudva et al. 2013). For SecB binding, no consensus sequence could be determined, however binding occurs not directly at the signal peptide but at mature regions preventing folding of the preprotein (Randall et al. 1990). Further it is assumed that the signal peptide of the precursor leads to retarded folding, facilitating binding of SecB (Park et al. 1988; Hardy and Randall 1991). In the next step of post-translational delivery, the SecB-precursor complex interacts with the ATPase SecA. SecA is a soluble protein, however about 50% in the cell are membrane-associated (Cabelli et al. 1991). Preproteins, which do not tend to fast folding or to aggregate, are directly bound and targeted by SecA without the initial support of SecB (Fig. 3B) (Crane and Randall 2017). The interaction of membrane-associated SecA with the membrane-embedded SecYEG complex increases the affinity of SecA to the SecB-precursor protein complex (Hartl et al. 1990; Crane and Randall 2017). Thereby, SecA is not only bound to SecB but also to the signal sequence and to the mature domains of the preprotein (van Voorst et al. 2000). These interactions result in the release of SecB and the delivery to the SecYEG complex (Fekkes et al. 1998; Kudva et al. 2013). Conformational changes of the translocon triggered by SecA binding result in an open state (Crane and Randall 2017) and multiple rounds of ATP hydrolysis by SecA pushes another 30-40 amino acids of the preprotein through the SecYEG channel (Schiebel et al. 1991; Kudva et al. 2013). The proposed pushing mode is promoted by ATP-dependent cycles of insertion and deinsertion of a C-terminal 30 kDa domain of SecA (Economou and Wickner 1994). However, not only ATP hydrolysis but the proton motive force is involved in the translocation of the preprotein through the channel (Geller et al. 1986; Veenendaal et al. 2004). Here, it is proposed that the proteins SecDF operate as a proton pump (Tsukazaki et al. 2011). SecDF together with YajC can associate with SecYEG during translocation of secretory proteins. In addition to its role in pmf-dependent translocation of preproteins, SecDFYajC regulates the movement of the preprotein during stabilization of the SecA inserted conformation, thereby acting as translocation supporter (Duong and Wickner 1997b). During

translocation, the signal peptide is bound by the lateral gate of SecY which induces conformational changes, so that the plug facing the periplasmic site moves (Tam et al. 2005; Chatzi et al. 2014). Subsequently the signal peptide is inserted into the membrane whereas the mature domain is translocated into the periplasm (Chatzi et al. 2014). As a last step, the signal peptidase cuts the signal peptide in the C-domain at the -3 and -1 positions (Heijne 1990; Chatzi et al. 2014) resulting in the mature protein.

1.5 Co-translational delivery to the Sec system

In addition to its role in post-translational export of proteins into the periplasm, SecYEG is also responsible for the co-translational insertion of proteins into the inner membrane (Fig. 1) (Cross et al. 2009). Therefore, the Sec translocon cooperates with the universally conserved and essential SRP system, comprised of the signal recognition particle (SRP) and its receptor FtsY (Bernstein et al. 1989; Lutcke 1995; Kuhn et al. 2017). Both SRP and FtsY exhibit a GTPase activity controlling the co-translational targeting process (Bernstein et al. 1989; Yang and Zhang 2011). Probably the biggest advantage of co-translational insertion is that aggregation of the hydrophobic segments of inner membrane proteins, emerging from the ribosome in the aqueous cytosol, is prevented by their direct binding to the co-translational pathway components. However, the cell has to deal with the fact that the insertion is retarded due to the slow translation rate of the ribosome (Steinberg et al. 2018). At the ribosomal exit tunnel, the nascent membrane protein is bound by SRP as soon as a hydrophobic signal sequence emerges (Noriega et al. 2014a). The SRP-ribosome-nascent chain complex (RNC) is then targeted to the inner membrane where SRP interacts with the membrane-associated receptor protein FtsY. After binding to FtsY the nascent chain is forwarded to the membrane embedded SecYEG channel after coordinated GTP hydrolysis by SRP and FtsY and the subsequent dissociation of SRP-FtsY from the nascent chain (Valent et al. 1998; Kuhn et al. 2017). It is assumed that the hydrophobic transmembrane segments are inserted into the bilayer at the lateral gate of SecY whereas periplasmic domains are translocated through the hydrophilic cavity of SecY (van den Berg et al. 2004; Bischoff et al. 2014). SecYEG can cooperate with the SecDFYajC complex and during the insertion of multi-spanning membrane proteins SecYEG is working together with the membrane insertase YidC (Duong and Wickner 1997a; Schulze et al. 2014). Thereby, YidC functions as an assembly site for proteins.

It receives the additional TMDs from SecYEG and is responsible for the folding of the membrane protein into their native tertiary structure (Beck et al. 2001; Nagamori et al. 2004).

1.5.1 The signal recognition particle and its receptor

The idea that secretory proteins contain intrinsic signal elements guiding them to their target destination was first postulated in 1971 by Blobel and Sabatini in their “signal hypothesis” (Blobel and Sabatini 1971). In the following years Blobel and Dobberstein substantiated this theoretical hypothesis by experiments on a reconstituted membrane translocation system (Blobel 1975; Blobel and Dobberstein 1975; Matlin 2002). The components involved in binding and recognition of these postulated intrinsic export signals, SRP and its receptor, were first discovered in 1980 on the endoplasmic reticulum of mammalian cells (Meyer and Dobberstein 1980a, 1980b; Walter and Blobel 1980; Matlin 2002). For a few years, it was not clear if also *E. coli* encodes an SRP system involved in protein translocation, since genetic screens for export mutants did not provide any indication (Bassford et al. 1991). In the late 1980s and early 1990s possible *E. coli* SRP system components were identified by sequence comparison between genes from *E. coli* and the known eukaryotic SRP components (Bernstein et al. 1989; Römisch et al. 1989; Ribes et al. 1990). The most important evidence was provided 1993, by the demonstration that the *E. coli* SRP homolog functionally substitutes the eukaryotic one (Bernstein et al. 1993). As the eukaryotic homolog, the *E. coli* SRP is a ribonucleoprotein consisting of a GTPase protein component, called Ffh (SRP54 in mammalian) and a RNA component, the 4.5S RNA (7S RNA in mammalian) (Fig. 4) (Poritz et al. 1990; Ribes et al. 1990). The *E. coli* SRP interacts with the GTPase domain of its receptor FtsY (SRα in mammalian) (Fig. 4), which is associated with the membrane (Luirink et al. 1994; Egea et al. 2004; Focia et al. 2004). The components of the SRP system are universally conserved among all three kingdoms of life and essential for viability of the cell (Keenan et al. 2001). A big difference between the eukaryotic and prokaryotic SRP pathway is, that in prokaryotes SRP targets mostly inner membrane proteins whereas in eukaryotes the pathway is also used for secretory proteins (Valent et al. 1998; Koch et al. 1999; Kuhn et al. 2017). The targeting of SRP occurs co-translationally with exception of the chloroplast SRP, a unique SRP performing post-translational targeting (Akopian et al. 2013).

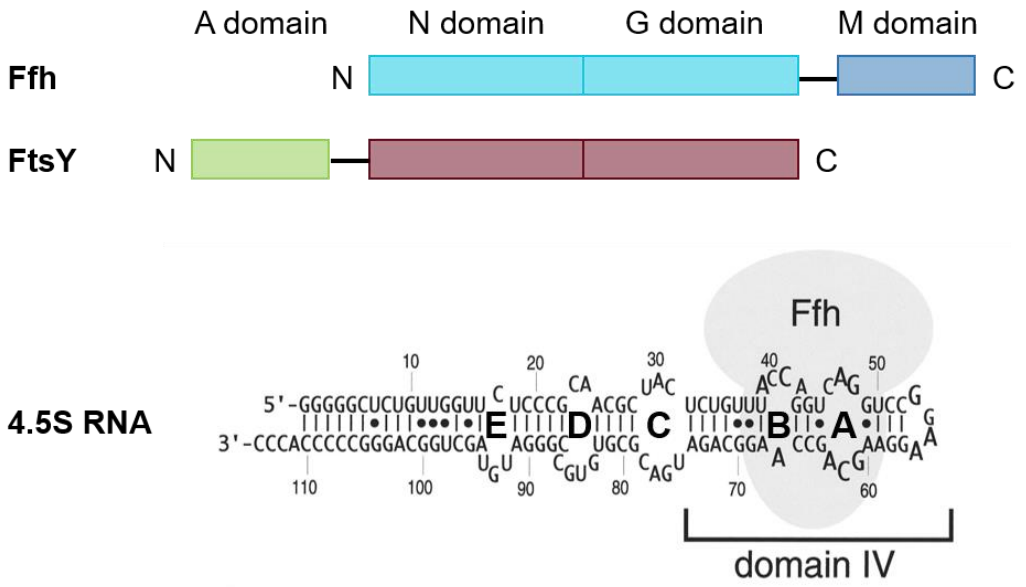


Figure 4: Components and domain organization of the SRP pathway in *E. coli.* The protein component Ffh from SRP and the SRP receptor FtsY consist of the conserved N and G domains. Ffh encodes for an additional C-terminal M domain, responsible for binding of the signal sequence, whereas FtsY encodes for an additional N-terminal A domain, responsible for the attachment to the membrane. The G domains include the GTP binding site. The 4.5S RNA of SRP consists of 115 nt. Ffh is bound via the conserved domain IV of the 4.5S RNA (adapted from Schmitz et al. 1999b; Shan and Walter 2005).

The protein component Ffh of the *E. coli* SRP is divided into three domains, the N, G and M domain (Fig. 4). The amino-terminal N domain forms a four-helix bundle and is involved in ribosome binding and the interaction with the receptor FtsY (Fig. 5B) (Freymann et al. 1997; Gu et al. 2003). The G domain consists of a Ras-like GTPase region for binding and hydrolysis of GTP and is closely associated with the N domain, forming the tightly packed and functional unit, the NG domain (Fig. 5B) (Freymann et al. 1997; Keenan et al. 1998; Buskiewicz et al. 2005). The C-terminal methionine-rich M domain is important for the interaction with the 4.5S RNA and is responsible for binding of the hydrophobic signal sequences (Römisch et al. 1990; Zopf et al. 1990). It is connected to the NG domain via a 30 amino acid long linker, involved in structural rearrangements during RNA binding (Fig. 5) (Buskiewicz et al. 2005).

The C-terminal part of the M domain is built-up of five amphipathic α-helices, αM1-αM5, forming a hydrophobic groove for binding of the signal sequences of SRP substrates (Fig. 5A) (Hainzl et al. 2011; Jomaa et al. 2016; Steinberg et al. 2018). The helices αM1 and αM2 are connected by a finger-loop, which is, together with the flexible αM5, involved in closing of the hydrophobic groove so that the bound signal sequence

is protected from the environment (Keenan et al. 1998; Kuhn et al. 2017). The binding motif for the 115 nucleotide long 4.5S RNA is located in the M domain consisting of a helix-turn-helix motif, a conserved sequence of serine, arginine and glycine residues (Doudna and Batey 2004). This motif is bound to the highly conserved domain IV (about 50 nucleotides) of the 4.5S RNA (Fig. 4) (Schmitz et al. 1999a; Batey 2000).

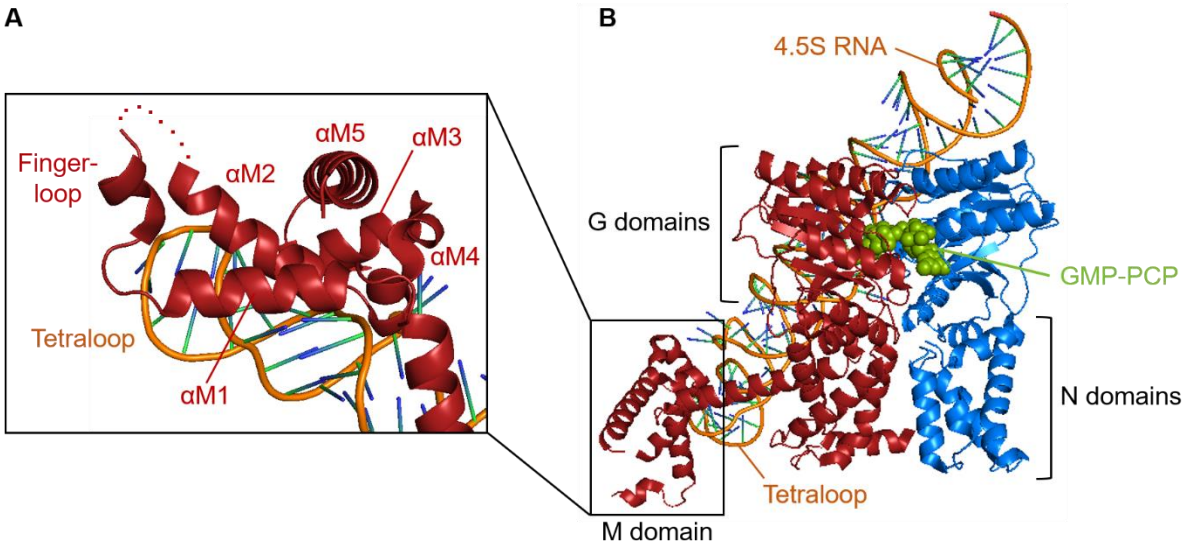


Figure 5: Crystal structure of the bound SRP-FtsY complex in presence of GMP-PCP. Ffh is displayed in red, FtsY is displayed in blue. (A) Insight into the hydrophobic groove of Ffh formed by the five α-helices of the M domain (αM1-αM5). The bound signal sequence is locked in the hydrophobic groove via the αM5 and the fingerloop between αM1 and αM2. The 4.5S RNA is bound to the M domain via a helix-turn-helix motif of αM2 and αM3 and the tetraloop of the RNA. The linker between αM1 and αM2 is disordered and not solved in the crystal structure. (B) Interaction of SRP and FtsY via their homologous NG domains (Ffh NG domain displayed in red, FtsY NG domain displayed in blue). The M domain of Ffh is connected to the NG domain via a 30 amino acid long linker region. The A domain of FtsY is disordered and not solved as well as the C-terminal 21 amino acids of Ffh (Image created with PyMOL 2.3.0, PDB 2XXA, adapted from Steinberg et al. 2018).

Except for domain IV, the SRP RNAs of different organisms vary widely in sequence, structure and size (Samuelsson and Zwieb 1999; Batey 2000). Both, eukaryotic and Gram-positive bacteria SRP RNAs contain an additional Alu-domain, mediating translation arrest until the RNC has arrived at the membrane (Siegel and Walter 1986; Steinberg et al. 2018). This domain and function is missing in Gram-negative bacteria SRP RNAs (Doudna and Batey 2004) and maybe there is some other mechanism involved to compensate the lack of the Alu domain. The fact that the Alu domain is only lost in Gram-negative bacteria assumes that a slow translation rate during targeting could be important. With ribosomal profiling it was shown that translational slowdown

of 69% of the membrane proteins of *E. coli* is triggered by Shine-Dalgarno-like elements between codon 16 and 60 (Fluman et al. 2014). Another study in the same year supports the idea of translational slowdown by the presence of rare non-optimal codon clusters 35-40 codons downstream of the SRP binding site (Fluman et al. 2014; Pechmann et al. 2014; Steinberg et al. 2018).

The 4.5S RNA forms a hairpin-like structure with a conserved GGAA tetraloop at the closed end and five additional internal loops (A-E) (Fig. 4). The binding to the helix-turn-helix motif of Ffh is mediated by the loops A and B (domain IV), a symmetric and an asymmetric loop located near the tetraloop (Batey 2000; Steinberg et al. 2018). Such as the protein component Ffh also the 4.5S RNA is essential in *E. coli* and it is essential for SRP function (Brown and Fournier 1984; Phillips and Silhavy 1992). Binding of the RNA to Ffh not only stabilizes the M domain during signal sequence binding but also functions as a catalytic component in the GTP cycle of SRP and its receptor FtsY (Zheng and Gierasch 1997; Peluso et al. 2001; Steinberg et al. 2018). Since there is an excess of 4.5S RNA (1.3 - 2 µM) compared to the amount of Ffh (0.3 - 0.5 µM) in the cell and the RNA has a picomolar affinity to Ffh it is assumed that all Ffh proteins are bound to its RNA (Batey 2000; Kudva et al. 2013).

The SRP receptor FtsY of *E. coli* is a shorter version of the eukaryotic SRP receptor which consists of two subunits, the soluble SRα and the membrane-integral SRβ subunit (Gilmore 1982). FtsY is homologous to the eukaryotic SRα subunit missing the membrane-integral SRβ subunit of eukaryotes (Gill et al. 1986; Bernstein et al. 1989; Luirink et al. 1994). Thus, FtsY occurs partially soluble in the cytoplasm and partially membrane-associated interacting with the translocon (Luirink et al. 1994; Leeuw et al. 1997; Weiche et al. 2008).

As well as SRP, also the receptor FtsY has a single GTPase subunit, the G domain and a N domain, forming a bundle of four α-helices (Fig. 4 and 5B) (Montoya et al. 1997; Steinberg et al. 2018). The third domain of FtsY, named A domain, is located at the N-terminus of the protein (Fig. 4). FtsY interacts with SRP, membrane lipids and the SecYEG translocon (Miller et al. 1994; Leeuw et al. 2000; Angelini et al. 2005). The N and G domain of FtsY are homologous to the structure of the SRP N and G domain (Egea et al. 2004; Focia et al. 2004; Kuhn et al. 2017). Accordingly, these two domains form the functional NG domain (Bernstein et al. 1989; Römisch et al. 1989). In contrast to SRP, the NG domain of FtsY forms the C-terminal part of the protein. When SRP and FtsY are both bound to one GTP molecule the NG domains of the two proteins

form a stable complex resulting in a parallel symmetrical but mirrored arrangement (Fig. 5B). Through this mirrored arrangement substrate twinning between the two GTP molecules at the protein interface occurs (Egea et al. 2004; Focia et al. 2004; Ataide et al. 2011). In contrast to other GTPases, SRP and FtsY are also stable in the nucleotide-free version (Steinberg et al. 2018). The N-terminal highly acidic A domain of FtsY is responsible for the attachment of FtsY to the membrane via two conserved membrane binding sequences flanking the A domain. These amphiphilic regions at the N-terminus and at the interface between the A and N domain interact with anionic phospholipids of the bilayer (Parlitz et al. 2007; Weiche et al. 2008; Stjepanovic et al. 2011). Interestingly, the A domain is not essential for targeting *in vivo* and *in vitro* and is lacking in Gram-positive bacteria (Eitan and Bibi 2004; Weiche et al. 2008; Steinberg et al. 2018). The deletion of the A domain is compensated through the membrane binding site at the interface of the A and N domain and additional translocon binding sites in the NG domain (Parlitz et al. 2007; Stjepanovic et al. 2011; Kuhn et al. 2015; Steinberg et al. 2018). Until today the reason for the presence of the A domain in Gram-negative bacteria is still unknown (Steinberg et al. 2018).

1.5.2 Signal sequence recognition and targeting via the SRP system

The pathway of co-translational delivery via the SRP system starts at the ribosomal exit tunnel, a multiple binding platform for a number of cellular components involved in protein sorting and nascent chain processing. With a length of 80-100 Å the tunnel harbours about 30-40 amino acids in extended conformation during translation (Voss et al. 2006; Fedyukina and Cavagnero 2011). The edge of the exit tunnel is formed by the ribosomal RNA and the four proteins uL22, uL23, uL24 and uL29 (new nomenclature according to Ban et al., 2014) (Kramer et al. 2009). Of great importance is the protein uL23 since it constitutes the docking site for a lot of ribosome-associated proteins, like trigger factor, SRP or even the Sec translocon (Fig. 6A) (Kramer et al. 2002; Gu et al. 2003; Mitra et al. 2005).

The newly synthesized polypeptides emerge through the exit tunnel and are directly contacted by the different ribosome-associated proteins, mostly in a competitive binding mode. The first step in the SRP cycle is also called the “scanning mode” (Holtkamp et al. 2012). At the very beginning of translation SRP is bound via its NG domain of Ffh to the ribosomal protein uL23 and uL29. Two additional contacts were

observed between the Ffh M domain with uL23 and the ribosomal 23S rRNA. Furthermore, the bacteria specific ribosomal protein bL32 interacts with the SRP 4.5S RNA (Fig. 6B) (Halic et al. 2006; Schaffitzel et al. 2006; Holtkamp et al. 2012; Jomaa et al. 2016).

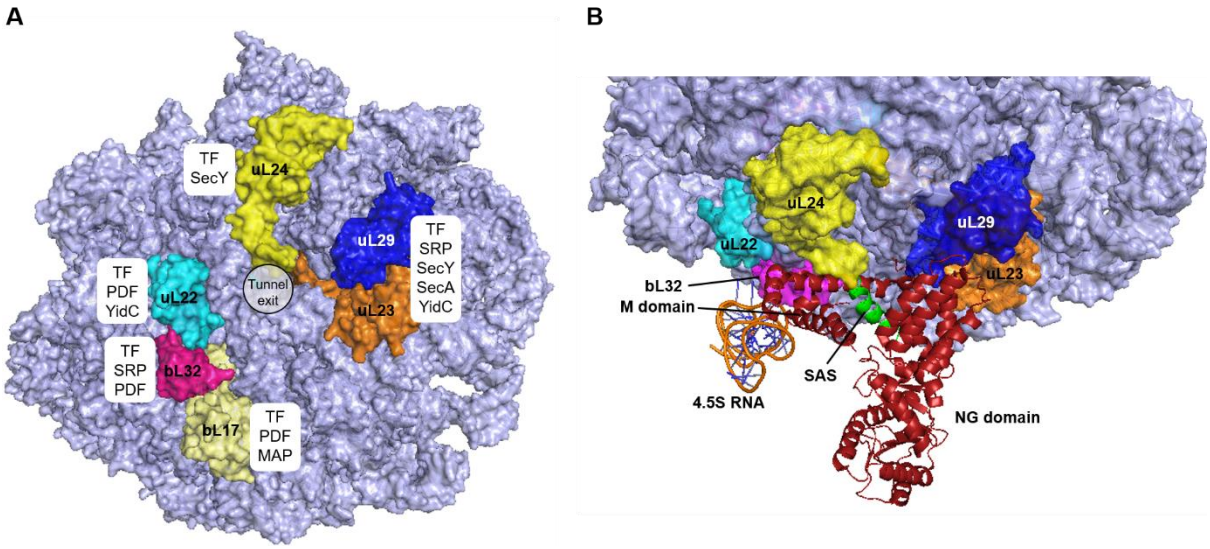


Figure 6: Binding events at the ribosomal exit tunnel. (A) Ribosomal exit tunnel as a multiple binding platform for different cellular components. The edge of the tunnel exit is formed by the proteins uL22, uL23, uL24 and uL29. The proteins bL32 and bL17 are located farther. The different cellular components forming contacts with the ribosomal proteins are indicated. The proteins TF, SRP, SecY, SecA and YidC are involved in protein sorting whereas the enzymes peptidyl formylase (PDF) and methionine amino peptidase (MAP) are responsible for the co-translational removal of the N-terminal methionine residue. (B) Binding of SRP next to the ribosomal exit tunnel early during translation. The ribosome is displayed in light blue, SRP is displayed in red and the signal sequence is displayed in green. The NG domain of Ffh contacts the ribosomal protein uL23 (orange) and uL29 (blue). The protein uL23 also contacts the M domain of Ffh. The 4.5S RNA interacts with the ribosomal protein bL32 (pink) (Image created with PyMOL 2.3.0, PDB 5GAF, adapted from Denks et al. 2014; Steinberg et al. 2018).

It is assumed that SRP scans the ribosomes at the very beginning of translation and binds to RNCs very fast and with high affinity (dissociation constant Kd ≈ 50-80 µM), regardless whether a signal sequence is present or not (Bornemann et al. 2008; Saraogi et al. 2014). However, the dissociation rate of SRP is much slower for RNCs with a strong SRP signal sequence compared to RNCs without a specific SRP signal sequence (Holtkamp et al. 2012; Saraogi et al. 2014; Steinberg et al. 2018). This finding could be a possible explanation for the fundamental question how the small number of SRP molecules are able to distinguish between a correct and incorrect cargo in the huge pool of translating ribosomes in the cell (Steinberg et al. 2018). Another hypothesis is that SRP and the ribosome-associated protein trigger factor (TF), a

chaperone for nascent cytoplasmic, periplasmic and outer membrane proteins (Oh et al. 2011), compete for the nascent chain in an anti-cooperative mode (Oh et al. 2011; Bornemann et al. 2014; Ariosa et al. 2015). Thus, it is assumed that SRP and TF, which is present in excess compared to ribosomes, prevent binding of the other component if an incorrect cargo for them emerges (Bornemann et al. 2014). On the other hand, ribosomal profiling studies indicate that nascent chains are bound at different time points during translation depending on the length of the nascent chain. For TF the strongest binding was observed for nascent chains with a length of more than 100 amino acids (Oh et al. 2011). In contrast, SRP preferably binds to shorter nascent chains with a length of about 50 to 100 amino acids (Schibich et al. 2016).

For a long time, it was assumed that SRP is only able to bind to RNCs as soon as a SRP signal sequence is exposed (Noriega et al. 2014a; Noriega et al. 2014b; Schibich et al. 2016; Steinberg et al. 2018). New studies indicate that SRP reaches into the exit tunnel of so called “vacant” ribosomes sensing the nascent chain at the very beginning of translation (Fig. 7) (Denks et al. 2017; Mercier et al. 2017).

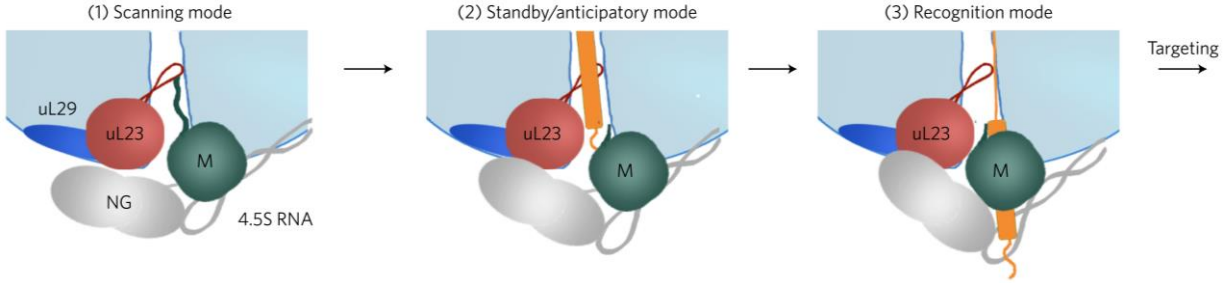


Figure 7: Early steps of the SRP pathway in *E. coli.* (1) Scanning mode. The C-terminus of the M domain (green) of SRP inserts into the ribosomal exit tunnel contacting the intra-tunnel loop of uL23 (red). This allows SRP an early scanning for a nascent chain. The NG domain and the 4.5S RNA are displayed in grey. (2) Standby / anticipatory mode. The contact of the C-terminal part of the M domain is replaced by an emerging nascent chain however the affinity of SRP to the ribosome increases as soon as the nascent chain contacts the loop of uL23. The C-terminal part of the M domain is retracted from the intra tunnel loop. (3) Recognition mode. After the exposure of the nascent chain the latter is bound in the hydrophobic groove of the M domain and subsequently the RNC-SRP complex is targeted to the membrane-bound receptor FtsY (Denks et al. 2017; Steinberg et al. 2018).

This contact is mediated by the C-terminal part of the M domain of SRP and a 19 amino acid long hairpin-like loop of uL23 which protrudes into the exit tunnel. Preliminary data suggest that this interaction is specific for SRP since none of the other tested ribosome-associated proteins reaches the intra-tunnel loop (Denks et al. 2017; Steinberg et al. 2018). This early sensing of the intra tunnel loop could also be a possible explanation

for the fast and efficient cargo recognition (Denks et al. 2017; Steinberg et al. 2018). However, decoding these processes early during translation in detail is still part of the current studies.

In the next step SRP switches from a scanning mode into a recognition mode characterized by a kinetically stable binding of SRP to an exposed signal sequence (Steinberg et al. 2018). This stable complex is characterized by a 20-1000-fold decrease in the dissociation rate and an increase in the binding affinity from 50-80 nM into the sub nanomolar range (< 3 nM) (Bornemann et al. 2008; Saraogi et al. 2014).

Extensive studies on SRP binding to signal sequences have been carried out using stalled ribosomes, exposing varying nascent chains. In *E. coli* ribosome stalling was found during translation of the SecM leader peptide or by the TnaC leader that is used to regulate the expression of the tryptophan operon (Gong and Yanofsky 2001; Nakatogawa and Ito 2002; Woolhead et al. 2006). SecM induced stalling occurs at a C-terminal 17 amino acid long sequence of SecM to adjust the amount of SecA via expression regulation to the translocation activity in the cell (Oliver et al. 1998; Nakatogawa and Ito 2002). Tryptophan-mediated ribosome stalling is induced by the 24 amino acid long TnaC peptide encoded in the *tna* operon. In presence of tryptophan the ribosome stalls at the *tnaC* stop codon since the $tRNA^{Pro}$ in the P site remains bound to the TnaC nascent chain. As a result, the Rho factor transcription termination site is blocked and the transcription of downstream genes for the metabolization of tryptophan occurs (Gong and Yanofsky 2001). A great discovery was that fusing only the *secM* or *tnaC* stalling sequence to the 3' end of a sequence of interest, is sufficient for ribosomal stalling, which facilitates the generation of ribosomes exposing any desired nascent chain (Nakatogawa and Ito 2002; Bischoff et al. 2014). With the help of this method, SRP binding to stalled ribosomes exposing the nascent chain of interest can be analyzed.

In contrast to the signal sequence of secreted proteins in *E. coli*, mostly all SRP signal sequences are non-cleavable signal sequences, remaining in the protein after insertion (Ulbrandt et al. 1997; Koch et al. 1999). In most cases the N-terminal TMD of inner membrane proteins serves as the SRP signal sequence (Lee and Bernstein 2001). For SRP signal sequences no consensus sequence exists since they vary in amino acid composition and sequence length (Lee and Bernstein 2001; Hainzl and Sauer-Eriksson 2015). However, a stretch of at least eight hydrophobic amino acids exceeding a threshold level of hydrophobicity is crucial for SRP recognition (Martoglio

and Dobberstein 1998; Lee and Bernstein 2001; Hegde and Bernstein 2006; Hainzl and Sauer-Eriksson 2015). A study on the presecretory proteins MBP and OmpA that are normally targeted by the SecB pathway can be rerouted into the SRP pathway by increasing the hydrophobicity of their signal peptides or replacing the signal peptide with a TMD (Lee and Bernstein 2001). This study indicated that SRP has a higher affinity for TMDs than for the less hydrophobic signal peptides. The importance of hydrophobicity was also shown by a study with engineered SRP signal sequences, differing in their hydrophobicity. RNCs exposing the nascent chain with the highest grand average of hydropathy (GRAVY) of 1.47 are bound 1000-fold stronger by SRP (Kd = 0.08 nM) compared to vacant 70S ribosomes (Kd = 80-100 nM). This binding affinity decreases to 2.4 nM by lowering the GRAVY from 1.47 to 1.26. In contrast, RNCs exposing the signal sequence of the SecB substrate PhoA, with a GRAVY of only 1.06, resulted in a binding affinity of 40 nM, only 2-fold stronger than to “vacant” ribosomes. Further lowering of the GRAVY to 0.38 abolished measurable SRP binding (Saraogi et al. 2014). Thus, it is suggested that a threshold of hydrophobicity must be exceeded for stable SRP binding and efficient targeting (Lee and Bernstein 2001; Saraogi et al. 2014).

However, it was demonstrated that proteins with moderately hydrophobic signal sequences not reaching the threshold of hydrophobicity can still be switched into the SRP pathway. One possibility could be that this routing into the SRP pathway is achieved by an unusually basic N-terminal part of the signal sequence (Peterson et al. 2003). Studies on MBP and OmpA where the signal peptides were replaced with a signal sequence with four closely spaced basic residues, showed that targeting is now dependent on SRP instead of SecB. It seems that the generation of salt bridges between the SRP RNA and the basic amino acids compensate the less hydrophobic character of the signal sequence (Peterson et al. 2003). In addition to hydrophobicity and basic amino acid residues the absence of helix-breaking amino acids, like proline and glycine, and the presence of rare non-optimal codons resulting in translational slowdown, also facilitate SRP binding (Adams et al. 2002; Fluman et al. 2014; Pechmann et al. 2014). A study on the SRP-dependent membrane protein KdpD showed that SRP is also able to recognize amphiphilic regions (Maier et al. 2008).

The signal sequences are bound by the hydrophobic groove formed by the M domain of Ffh (Keenan et al. 1998). The switch from the scanning mode into the recognition mode is initiated by conformational changes in SRP due to a remodelling on the

ribosome which leads to a more stable complex (Holtkamp et al. 2012; Denks et al. 2017; Steinberg et al. 2018). Recent studies showed that there is an additional step between scanning and recognition, called the anticipatory or standby mode (Fig. 7) (Denks et al. 2017; Mercier et al. 2017; Steinberg et al. 2018). This mode is characterized by the retraction of the contact of SRP with the intra-tunnel loop as soon as the nascent chain reaches a length of about 25 amino acids assuming that SRP awaits the emerging signal sequence (Denks et al. 2017; Steinberg et al. 2018). Due to this retraction, SRP is repositioned at the tunnel exit and the M domain is able to bind to the emerging signal peptide (recognition mode) (Denks et al. 2017; Steinberg et al. 2018). This is in good agreement with earlier findings demonstrating that SRP binds to RNCs with varying nascent chain lengths of up to 27, buried in the ribosomal exit tunnel, with a Kd comparable to vacant ribosomes. Reaching a length of more than 32 amino acids, resulting in the partial exposure of the nascent chain and the retraction of SRP from the intra-tunnel loop, leads to an increase in the binding affinity to ≈ 1 nM (Bornemann et al. 2008; Denks et al. 2017).

The correlation of nascent chain lengths and SRP binding affinities was also shown by another study from Noriega et al. in 2014. A 35 amino acid long nascent chain, not yet fully exposed from the ribosome, showed the weakest affinity in this assay with about 1.5 nM. The affinity increases with the length of the nascent chain reaching the maximum of 0.057 ± 0.01 nM at a nascent chain length of 75 amino acids. With increasing length, the affinity is weakened until finally with a length of 135 amino acids the weakest binding affinity, 5-fold lower, is measured (Noriega et al. 2014a). Reaching a length of about 140 amino acids even prevents targeting of the RNCs by SRP (Siegel and Walter 1988).

Signal sequence binding by the M domain of Ffh leads to conformational changes, which was first shown by resolution of two different crystal structures of SRP from *M. jannaschii* with and without a bound signal sequence (Fig. 8) (Hainzl et al. 2011). Signal sequence binding results in repositioning of the NG domain via the flexible GM linker in a way that the G domain comes in close contact to the tetraloop of the 4.5S RNA priming SRP for further binding to the receptor (Hainzl et al. 2011).

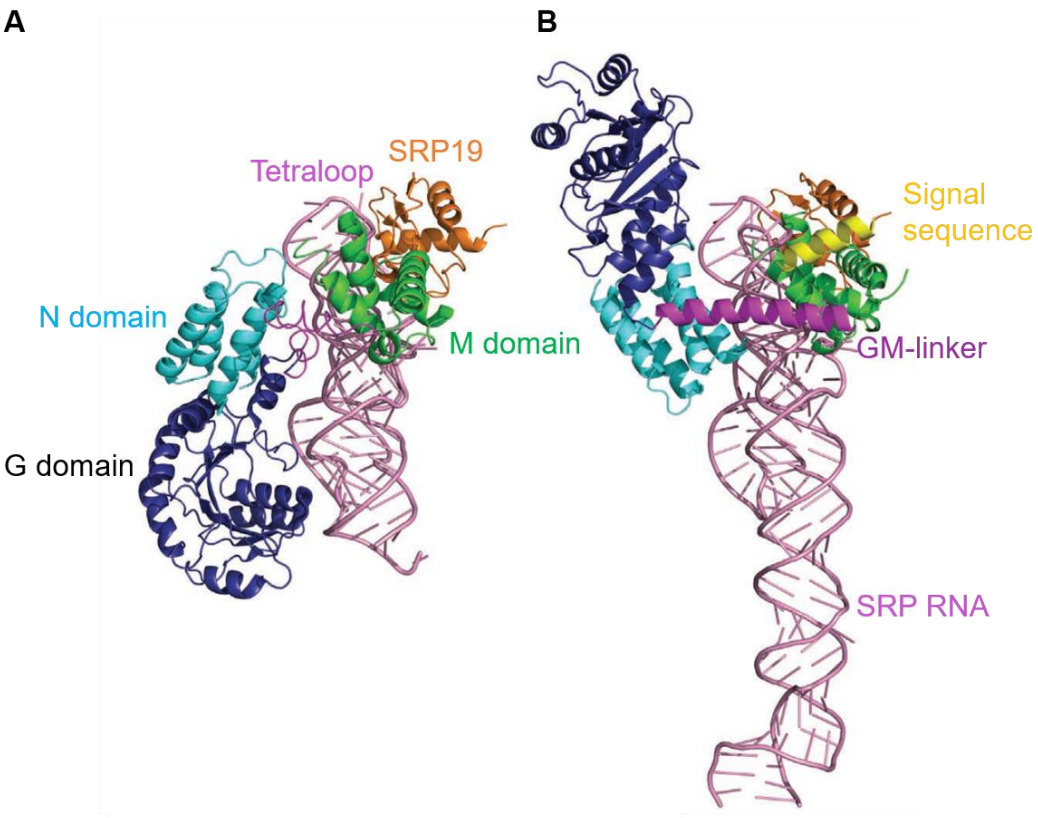


Figure 8: Conformational changes in the *M. jannaschii* SRP during signal sequence binding. SRP54 N domain is displayed in cyan, the SRP54 G domain in dark blue, the SRP54 M domain in green, the SRP54 GM linker in purple, SRP19 in orange, the signal sequence in yellow and the SRP RNA in purple. (A) SRP without a signal sequence. (B) SRP bound to a signal sequence. The binding results in a flip of the NG domain via the flexible GM linker. This rotation brings the G domain in close contact to the tetraloop of the SRP RNA (adapted from Hainzl et al. 2011).

After signal sequence binding by the M domain of Ffh, the RNC-SRP complex is targeted to the membrane via the interaction with the membrane bound receptor FtsY, which is activated for SRP interaction by anionic phospholipid binding (Leeuw et al. 2000). These anionic phospholipids are enriched near the SecYEG translocon, thus increasing the concentration of FtsY next to the translocon (Gold et al. 2010; Prabudiansyah et al. 2015; Steinberg et al. 2018). Furthermore, the interaction of FtsY with anionic phospholipids results in a movement of the A domain apart from the NG domain, exposing the SRP binding site (Draycheva et al. 2016).

Still controversially discussed is the fact that FtsY binding to the RNC-SRP complex was also shown in solution in multiple studies (Jagath et al. 2000; Ataide et al. 2011; Estrozi et al. 2011; Loeffelholz et al. 2013; Loeffelholz et al. 2015; Steinberg et al. 2018). However, *in vivo* the biggest amount of FtsY is located at the membrane (Mircheva et al. 2009) and it was also shown that the interaction of FtsY with lipids and

the SecYEG translocon is required for stable and high-affinity SRP binding (Lam et al. 2010; Draycheva et al. 2016). Therefore, it seems that *in vitro* binding of FtsY to an RNC-SRP complex in solution is possible, however, the formation of a ternary complex is facilitated during activation of FtsY by binding of lipids and SecYEG (Draycheva et al. 2016).

In a recent study with *E. coli* SRP in complex with RNCs, the before mentioned rearrangement of the SRP NG domain was directly coupled with binding of the receptor FtsY via its homologous NG domain resulting in a GTP independent early targeting complex (Fig. 9) (Jomaa et al. 2016). This initial recruitment of FtsY also leads to an increased binding affinity of SRP to the signal sequence into the picomolar range (Saraogi et al. 2014). In this early targeting complex the interaction of SRP and FtsY is mediated by electrostatic interactions between charged clusters on the surface of their N domains and an additional contact between the tetraloop of the 4.5S RNA and the FtsY G domain (Estrozi et al. 2011; Zhang et al. 2011). For the early targeting complex no GTP hydrolysis could be measured probably due to the fact, that the GTPase domains are localized farther from each other (Zhang et al. 2009; Estrozi et al. 2011; Loeffelholz et al. 2013; Loeffelholz et al. 2015). In the early targeting complex, it is possible that incorrect cargos are still bound to SRP. It is assumed that this part of the SRP pathway provides another checkpoint for cargo discrimination via a high dissociation rate of the FtsY-SRP-RNC complex, if bound to an incorrect cargo (Saraogi et al. 2014). The switch from the early to the closed targeting complex, important for the further delivery of the nascent chain to the translocon, is mediated by GTP dependent conformational changes (Fig. 9) (Shan and Walter 2003; Shan et al. 2004). SRP and FtsY belong to the new class of GTPases, not regulated by GTPase-activating proteins but by homodimerization. Due to their low nucleotide affinity and the associated high intrinsic exchange rate, they avoid their dependency on GTPase activating proteins (Gasper et al. 2009). If both proteins are bound to GTP additional extensive contacts between their G domains are found, whereby the transition to the active complex occurs, forming the quaternary SecYEG-FtsY-SRP-RNC complex and further release to the translocon is possible (Fig. 9) (Egea et al. 2004; Shan et al. 2004; Zhang et al. 2009; Saraogi et al. 2014). The GTP dependent switch from the early to the closed complex, driven by FtsY interaction with phospholipids, also weakens the binding of SRP to the signal sequence from the picomolar to the subnanomolar range, providing it for further delivery (Saraogi et al. 2014).

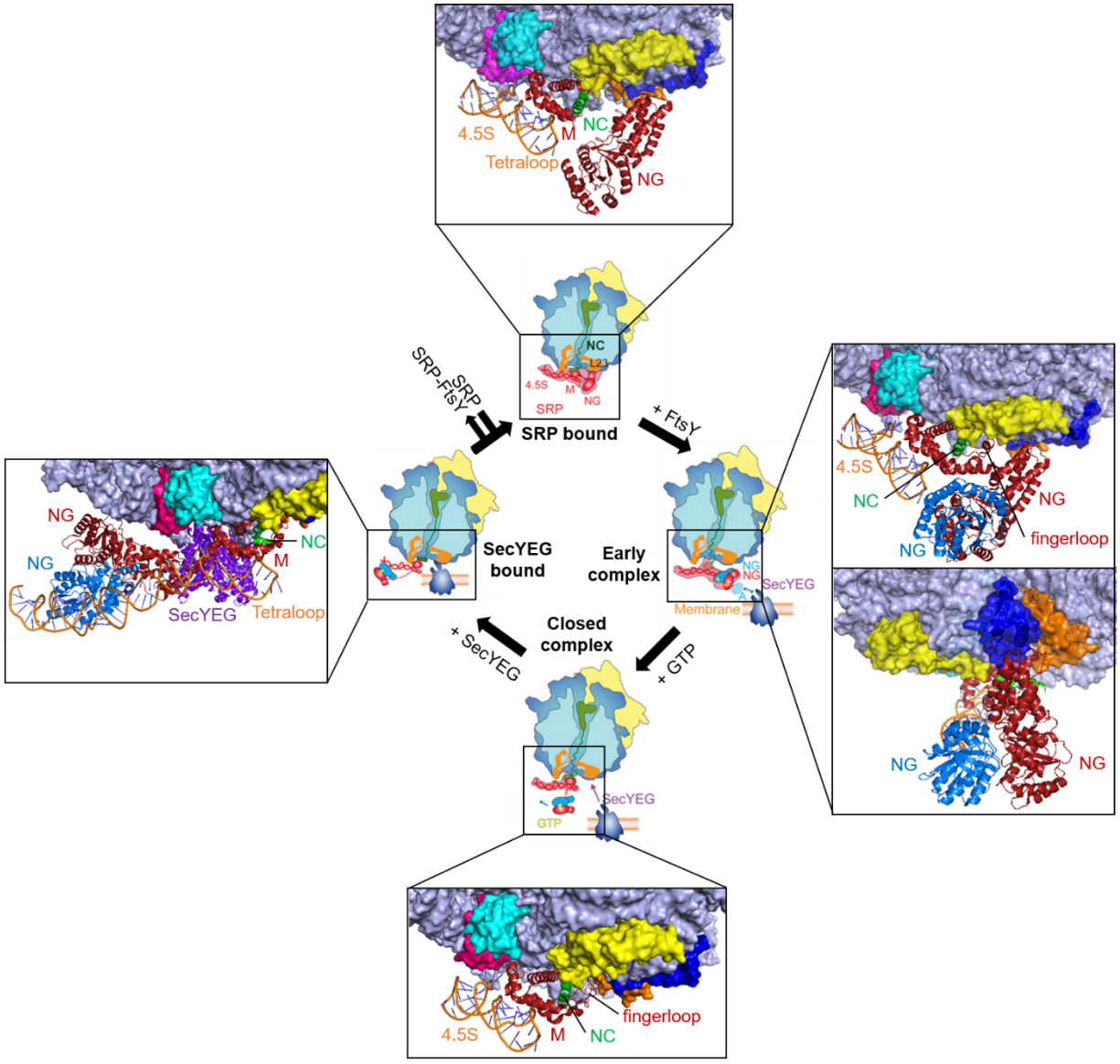


Figure 9: Conformational changes during the SRP pathway. SRP bound stage: SRP (red) is bound to the ribosome (lightblue) contacting uL23 (orange), uL29 (dark blue) and bL32 (pink). Nascent chain (NC, green) binding of SRP via the M domain repositions the NG domain via the flexible GM linker, so that the G domain contacts the tetraloop of the 4.5S RNA. The fingerloop of the M domain is still disordered (PDB 5GAF). Early complex: SRP and FtsY (blue) are bound via their NG domains (lower panel). The affinity of the M domain to the NC increases (indicated by the solved structure of the fingerloop; upper panel) (PDB 5GAD). GTP binding initiated formation of the closed complex. The NG domains move from the tetraloop to the distal end of the 4.5S RNA (not solved in the crystal structure due to their flexibility) exposing the SecYEG binding site on the ribosome (PDB 5GAG). SecYEG (purpleblue) bound stage: Stable RNC-SecYEG complex due to the remaining of the NG domains at the distal end of the 4.5S RNA. The affinity of SRP to the NC decreases and by GTP hydrolysis, the RNC-SRP-FtsY complex dissociates and delivers the NC to the Sec translocon, positioned at the ribosomal exit tunnel (Image created with PyMOL 2.3.0, adapted from Loeffelholz et al. 2015).

1.5.3 Nascent chain handover and insertion by the Sec translocon

The delivery of the nascent chain to the Sec translocon requires additional conformational changes of the closed targeting complex resulting in the activated complex. FtsY interacts with the Sec translocon, both with the A and the NG domain independently of each other (Kuhn et al. 2015). On the translocon the contact sites for FtsY are the cytoplasmic loops C4 and C5 of SecY (Kuhn et al. 2011). This interaction is characterized by a high binding affinity of about 0.18 µM (Kuhn et al. 2015). The GTP dependent rearrangements of SRP and FtsY include a reorientation of the NG heterodimer resulting in the exposure of the ribosomal protein uL23, the SecYEG binding site on the ribosome (Kuhn et al. 2015; Loeffelholz et al. 2015) (Fig. 9). Thus, a stable RNC-SecYEG complex is formed (Kuhn et al. 2017). GTP hydrolysis by the SRP-FtsY complex is initiated by additional conformational changes of the quaternary complex which leads to the dissociation of SRP from the nascent chain and the latter is further delivered to the translocon for insertion (Fig. 9). The delivery of the nascent chain to SecYEG seems to occur in a stepwise mechanism. In a first step, the NG domain of SRP is removed from the ribosome due to SecYEG binding to the ribosomal protein uL23 placing the ribosomal exit tunnel close to the SecYEG channel (Saraogi et al. 2014; Kuhn et al. 2015). This removal is also facilitated by the progressing translation resulting in ever-growing nascent chains (Saraogi et al. 2014). In the next step, the signal sequence is released from the SRP M domain and forwarded to SecYEG. During this transfer process, the coupled GTP hydrolysis is initiated by the movement of the NG domains of SRP and FtsY from the tetraloop of the 4.5S RNA over 100 Å to the distal end of the 4.5S RNA (Ataide et al. 2011; Shen et al. 2012). This provides evidence for the important function of the 4.5S RNA in the SRP pathway acting as a scaffold for productive exchange of the targeting and translocation factors (Shen et al. 2012; Voigts-Hoffmann et al. 2013; Steinberg et al. 2018).

The driving force for translocation of the nascent chain seems to involve two mechanisms. The initial step, the hand over and insertion of the first hydrophobic segment, seems to depend on GTP hydrolysis (Knyazev et al. 2018). Thereby, the nascent chain opens the SecY channel and is inserted forming a loop inside the channel, thus the hydrophobic segment interacts with the lateral gate (Knyazev et al. 2018). After dissociation of the SRP-FtsY complex, further translocation is dependent

on the translational chain elongation (Knyazev et al. 2018). Since on the outside of the channel, the nascent chain forms an additional loop on the cytoplasmic surface before entering the channel, a direct pushing force by the ribosome, as was initially assumed, is excluded (Park et al. 2014; Knyazev et al. 2018). Instead, it is suggested that a pulling force may be involved in the efficient translocation (Park et al. 2014). One possible candidate providing a pulling force could be SecDF which is driven by the pmf and was shown to be involved in ATP-independent translocation (Tsukazaki et al. 2011; Park et al. 2014). SecDF acts as a membrane-integrated chaperone and therefore binding, folding and movement of periplasmic domains of the translocated nascent chain could be involved in pulling the chain through the channel (Tsukazaki et al. 2011; Park et al. 2014). However, the detailed mechanism is not finally clarified but it is assumed that the pulling force reaches a limit if nascent chains with extended periplasmic domains have to be translocated (Knyazev et al. 2018; Steinberg et al. 2018). In this case, additional energy for translocation of SRP substrates could be provided by the aid of SecA (Knyazev et al. 2018; Steinberg et al. 2018). This was shown for the single-spanning membrane protein FtsQ, the double-spanning membrane protein leader peptidase and a hybrid protein (signal sequence of the SRP substrate MtlA fused to mature part of SecA-dependent OmpA). These proteins are targeted co-translationally by SRP and involve SecA (Neumann-Haefelin et al. 2000; Urbanus et al. 2001; van der Laan et al. 2004; Kuhn et al. 2017). Hence, it seems that the SRP pathway is able to cooperate with SecA in a sequential mode (Neumann-Haefelin et al. 2000).

SRP is not only able to interact with the Sec translocon it can also target RNCs to the membrane insertase YidC, acting independently of SecYEG.

1.6 The membrane insertase YidC

YidC belongs to the universally conserved YidC/Oxa1/Alb3 family (Yen et al. 2001). The *E. coli* YidC, essential for cell viability, consists of 548 amino acids and spans the membrane six times (Sääf et al. 1998; Samuelson et al. 2000). It is involved in the insertion of proteins with small periplasmic domains. Typically, these proteins are short and only consist of one or two transmembrane helices (Kuhn et al. 2017). Nevertheless, it was shown that also some polytopic membrane proteins can be inserted via the YidC only pathway *in vitro*, if they bear only small periplasmic domains

as it is the case for the six-spanning membrane proteins TatC and MtlA (Welte et al. 2012). Previously, it was thought that in bacteria, proteins which do not depend on the Sec system, are directly inserted into the membrane without the help of other protein components (Geller and Wickner 1985; Gier et al. 1998). In 1994, sequence analysis of the Oxa1 protein from *Saccharomyces cerevisiae*, a membrane insertase found in mitochondria, showed significant similarity with a putative *E. coli* protein, later named YidC (Bonnefoy et al. 1994). In 2000, studies on the M13 phage procoat protein, a Sec-independent membrane protein, demonstrated for the first time that YidC is involved in membrane insertion processes (Samuelson et al. 2000). It was also shown that the insertion of the Sec-dependent multi-spanning membrane protein ProW is affected in the absence of YidC, indicating a cooperation between the Sec system and YidC (Samuelson et al. 2000). To date the YidC only pathway includes the single-spanning Pf3 coat protein and the C-tail anchored protein SciP (TssL) as well as the double-spanning M13 procoat protein, MscL and subunit c of F1Fo ATP synthase. In combination with the Sec translocon, YidC is involved in the insertion of the single-spanning membrane protein FtsQ, the double-spanning membrane protein Lep and the polytopic membrane proteins CyoA (3 TMDs), subunit a of F1Fo ATP synthase (5 TMDs) and MalF (8 TMDs) (Kuhn et al. 2017).

All members of the insertase family share common structural features. These involve at least 5 TMDs (Fig. 10), forming a characteristic helix bundle in the membrane and a coiled-coil domain in the cytoplasm (Kiefer and Kuhn 2018). In addition to these common features the members can have unique characteristic elements resulting in functional specialization in different organisms (Fig. 10) (Funes et al. 2011). YidC homologs in Gram-positive bacteria, mitochondria and chloroplast consist of an elongated C-terminal cytoplasmic domain, involved in ribosome binding (Fig. 10) (Funes et al. 2011). YidC of *E. coli* shows reduced affinity to ribosomes, probably due to the shorter C-terminal cytoplasmic domain (Kuhn et al. 2017). Recently it was shown that C-terminal extensions are not limited to the homologs found in Gram-positive bacteria and mitochondria since also homologs in marine Gram-negative bacteria share this feature. Fused to *E. coli* YidC, the C-terminal extensions of these marine YidC homologs increased the binding affinity to ribosomes (Seitl et al. 2014). A unique feature of the *E. coli* YidC is an additional TMD at the N-terminus followed by a large periplasmic domain of 319 residues (Fig. 10) (Sääf et al. 1998; Funes et al. 2011). This domain forms a β-supersandwich with an elongated cleft, linked to the TMD 2 via a C-

terminal α-helical linker region (Fig. 11) (Ravaud et al. 2008). The TMD 1 as well as 90% of the periplasmic domain are not essential for the insertase function, however, they are involved in binding to the SecYEG associated protein SecF (Jiang et al. 2003; Xie et al. 2006). In contrast, deletion of the C-terminal part of the periplasmic domain, including an amphipathic helix (EH1), impairs the function of membrane insertion which suggests that the conserved TMDs 2-6 are critical for the insertase function (Jiang et al. 2003; Xie et al. 2006).

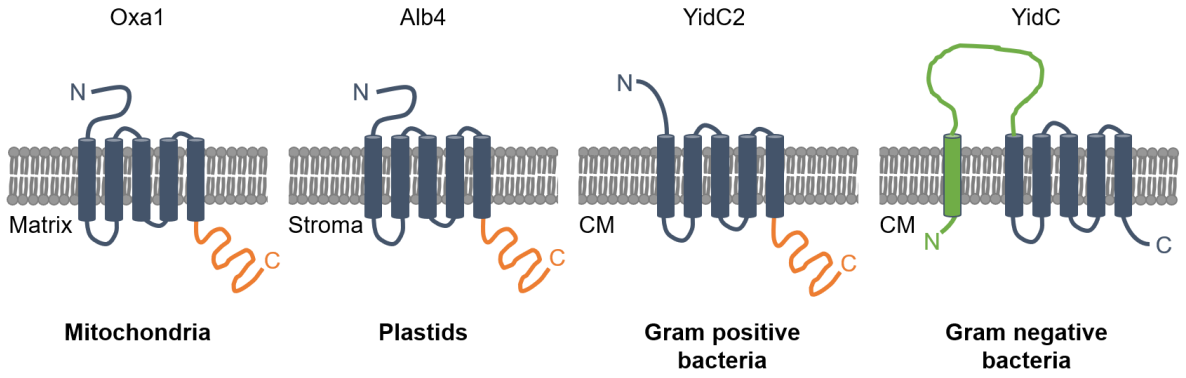


Figure 10: Members of the YidC/Oxa/Alb family. The common feature of this protein family is that they have at least five TMDs. In addition, the individual members may have specific structural characteristics. The YidC homologs Oxa1 in mitochondria, Alb4 in plastids and YidC2 in Gram-positive bacteria consist of an elongated C-terminal domain, involved in ribosome binding. YidC from Gram-negative bacteria contains an additional N-terminal transmembrane segment connected to the second transmembrane segment via a large periplasmic domain (adapted from Funes et al. 2011).

For a long time, it has been a controversial discussion if YidC is located in the membrane as a monomer or dimer. Previous studies have indicated that YidC forms a channel by the face-to-face interaction of two monomers (Lotz et al. 2008; Boy and Koch 2009; Kohler et al. 2009). This is due to the fact that YidC was purified from inner membrane vesicles as a dimer and in cryo-electron microscopy structures of the mitochondrial homolog Oxa1 and *E. coli* YidC bound to RNCs are found as dimers at the ribosomal exit tunnel (Boy and Koch 2009; Kohler et al. 2009). In contrast, another study with detergent-solubilized and membrane-embedded YidC showed that YidC in its monomeric form is sufficient for binding to a translating ribosome (Kedrov et al. 2013).

Novel insights into the insertion mechanism by YidC and its conformation in the membrane were gained in 2014 by solving the crystal structure of the *E. coli* YidC

(Kumazaki et al. 2014b). The structure supports the fact that the minimal functional unit of YidC, responsible for membrane insertion, is monomeric (Kumazaki et al. 2014b). It reveals a hydrophilic groove formed by the 5-helix bundle of TMD 2-6 (Fig. 11) with a positively charged surface in the groove, generated by the highly conserved Arg366 (Kumazaki et al. 2014b). Although this Arg366 is conserved and important for the function of the *Bacillus halodurans* YidC, structurally solved also in 2014, it is not essential in *E. coli* YidC (Chen et al. 2014; Kumazaki et al. 2014a). It is assumed that the hydrophilic regions of translocating substrates are shielded by the groove from the hydrophobic surrounding. Probably the size of the groove is crucial for the ability that only small periplasmic domains can be translocated by YidC. The groove is opened towards the cytoplasm and the membrane bilayer through a gap between TMD 3 and TMD 5. However, it is closed towards the periplasm by the tightly packed five transmembrane helices in the periplasmic leaflet and the amphipathic helix EH1 from the P1 loop, located parallel to the surface of the membrane (Fig. 11) (Kumazaki et al. 2014a; Kumazaki et al. 2014b).

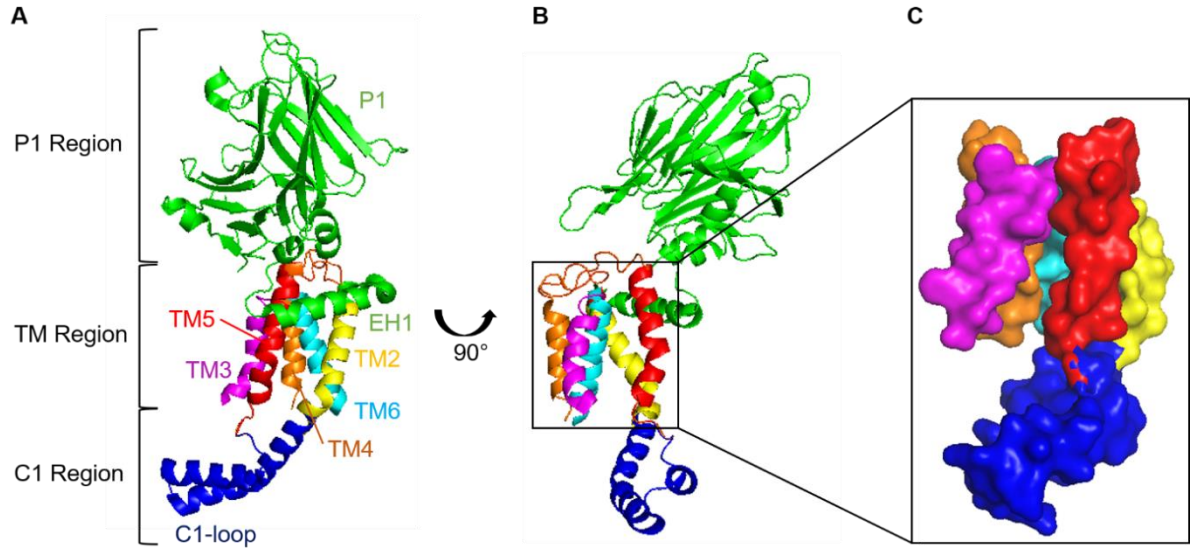


Figure 11: Crystal structure of the *E. coli* YidC. (A) The TMDs form a five-helix bundle (TMD 2 in yellow, TMD 3 in magenta, TMD 4 in orange, TMD 5 in red and TMD 6 in cyan) with a hydrophilic groove, open towards the cytoplasm and the bilayer between TMD 3 and TMD 5. The hydrophilic groove is closed towards the periplasm by the helix EH1 (green) of the large P1 loop (green), forming a β-supersandwich. The C1 loop consists of 2 α-helices forming a hairpin-like structure. (B) View from (A) rotated 90° to the right. The opening of the hydrophilic groove towards the cytoplasm by the gap between TMD 3 and TMD 5 is shown. (C) Detailed view with surface representation of the five-helix bundle with the hydrophilic groove and the opening towards the bilayer via TMD 3 and TMD 5. TMD 1 and the N-terminal part of the P1 loop is not solved in the crystal (Image created with PyMOL 2.3.0, PDB 3WVF, adapted from Kumazaki et al. 2014b).

This leads to the still unanswered question how the periplasmic domains, harboured in the hydrophilic groove, enter the periplasm. One possibility could be that YidC undergoes conformational changes during translocation, which has already been observed by the binding of RNCs to YidC (Kedrov et al. 2016).

Substrate binding by YidC results in a more opened structure towards the periplasm via movement of the N-terminal helices TMD 2 and TMD 3 and a displacement of the helix EH1 into the hydrophobic membrane core. The structural changes in the hydrophilic groove due to the rearrangements of TMD 2/TMD 3 and EH1 may facilitate the release of the nascent chain (Kedrov et al. 2016).

An additional possibility could be that the hydrophilic regions can enter the periplasm easily by a YidC induced thinning of the membrane due to the short lipid-exposed transmembrane helices 3, 4 and 5 (Chen et al. 2017; Kiefer and Kuhn 2018).

Another important feature for YidC function is the positively charged first cytoplasmic loop (C1) (Fig. 11), since deletion of the entire C1 loop leads to a non-functional YidC (Chen et al. 2014; Dalbey and Kuhn 2014). The C1 loop appears to be very flexible and consists of two antiparallel α-helices forming a hairpin-like structure located at the entrance of the hydrophilic groove and thus may be involved in substrate binding (Fig. 12 and 13) (Chen et al. 2014; Dalbey and Kuhn 2014; Kumazaki et al. 2014b; Kiefer and Kuhn 2018).

In addition to the information gained from structural studies, cross-linking experiments with several Sec-dependent and independent YidC substrates provide insights into the insertion mechanism. A universal mechanism of substrate recognition and insertion is assumed since both Sec-dependent and independent proteins are contacted by the same region of YidC (Yu et al. 2008). Cross-linking studies revealed that the transmembrane domains of YidC substrates contact the helices of TMD 3, TMD 4 and TMD 5 and also the non-essential helix of TMD 1 during the insertion process (Klenner et al. 2008; Yu et al. 2008; Klenner and Kuhn 2012). TMD 2 showed no contacts to substrates in the cross-linking studies. This could be due to the fact that the crystal structure of *E. coli* YidC decodes that TMD 2 wraps around the helix bundle and is therefore located at the outside of the membrane bundle (Kumazaki et al. 2014b; Kuhn et al. 2017). The contacts observed with TMD 3 and TMD 5 represent all hydrophobic residues located at only one face of the helices (Klenner et al. 2008; Klenner and Kuhn 2012). These studies reveal that the membrane embedded five-helix bundle makes strong contacts with the substrates during the insertion process and that these contacts

are due to hydrophobic interactions (Kuhn et al. 2017).

Based on the information gained from structural and crosslinking studies, two different insertion models for YidC substrates were developed (Fig. 12 and 13). Kumazaki et al. forwarded the linear, electrostatic insertion model for single-spanning membrane proteins with an extracellular acidic N-terminal tail like the YidC substrate MifM (Fig. 12). This model is based on the finding that the conserved arginine residue in the hydrophilic groove is essential for the insertase function of the *Bacillus subtilis* orthologue SpoIIIJ (Kumazaki et al. 2014a). In addition, several YidC substrates possess an acidic extracellular N-tail that could possibly interact with the positively charged hydrophilic groove (Dalbey et al. 2014; Kumazaki et al. 2014a).

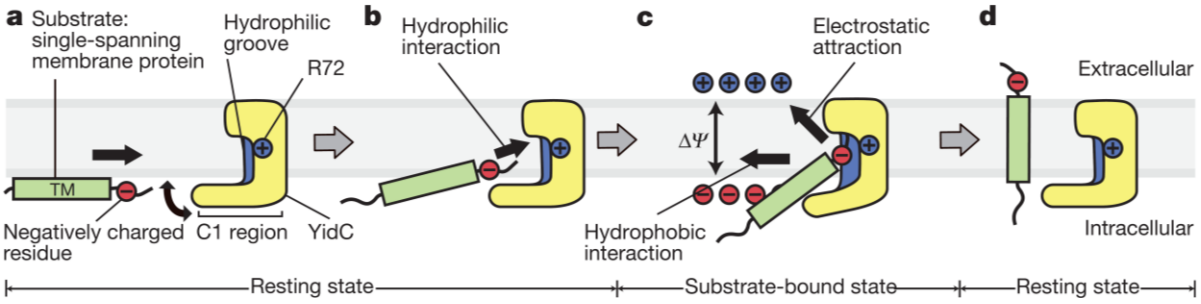


Figure 12: Linear, electrostatic insertion model for single-spanning membrane proteins with acidic extracellular N-tails. (a) The single-spanning membrane protein is attached to the membrane surface and subsequently recognized by the hairpin-like structure of the C1 loop. (b) The N-terminal hydrophilic domain is captured in the hydrophilic groove probably involving the interaction of the acidic N-tail with the conserved arginine residue in the groove. (c and d) The hydrophobic TMD is released into the bilayer whereas the hydrophilic N-tail is translocated into the extracellular region in a process which involves electrostatic attractions due to the membrane potential (Kumazaki et al. 2014a).

Initially the substrate is recognized by the positively charged hairpin-like structure of the C1 loop (Kumazaki et al. 2014a). The N-terminal hydrophilic domain is assumed to be captured in the hydrophilic groove of YidC since MifM could be cross-linked to residues of the groove. Substrate recognition and binding probably involves the positively charged residue in the hydrophilic groove and the negatively charged residues in the substrates N-tail. The involvement of electrostatic interactions between the groove and the substrates arise from the fact that, except for a substitution into a lysine residue, substitutions into a neutral or a negatively charged residue abolished the insertion of the substrate MifM. Simultaneously, the hydrophobic transmembrane domain is then released into the bilayer whereas the hydrophilic N-tail is translocated into the extracellular region. This last step, the release of the substrate can be promoted by both, hydrophobic interactions and electrostatic attractions due to the

membrane potential (Fig. 12) (Kumazaki et al. 2014a).

This model could also apply to the YidC protein of the Gram-positive bacteria *Streptococcus mutans*. The conserved arginine residue was shown to be necessary for the insertion of the single-spanning pf3 Lep fusion protein (pf3 fused at the C-terminus to amino acids 23-323 of leader peptidase) (Chen et al. 2014). However, the studies on YidC from *S. mutans* highlighted that the insertion model no longer applies to double-spanning membrane proteins and needs to be expanded (Chen et al. 2014). The substitution of the conserved arginine residue into a neutral charged residue did not abolished the insertion of the double-spanning model substrate M13 procoat fused to leader peptidase (Chen et al. 2014). This leads to the assumption that the electrostatic model only matches to some YidC proteins and specific YidC substrates.

In contrast, the positively charged arginine is not essential for the function of the *E. coli* YidC, suggesting a different insertion mechanism (Chen et al. 2014). Cross-linking studies on the YidC substrate Pf3 coat led to the development of the hydrophobic slide mechanism model (Fig. 13) (Klenner et al. 2008; Klenner and Kuhn 2012; Kiefer and Kuhn 2018)

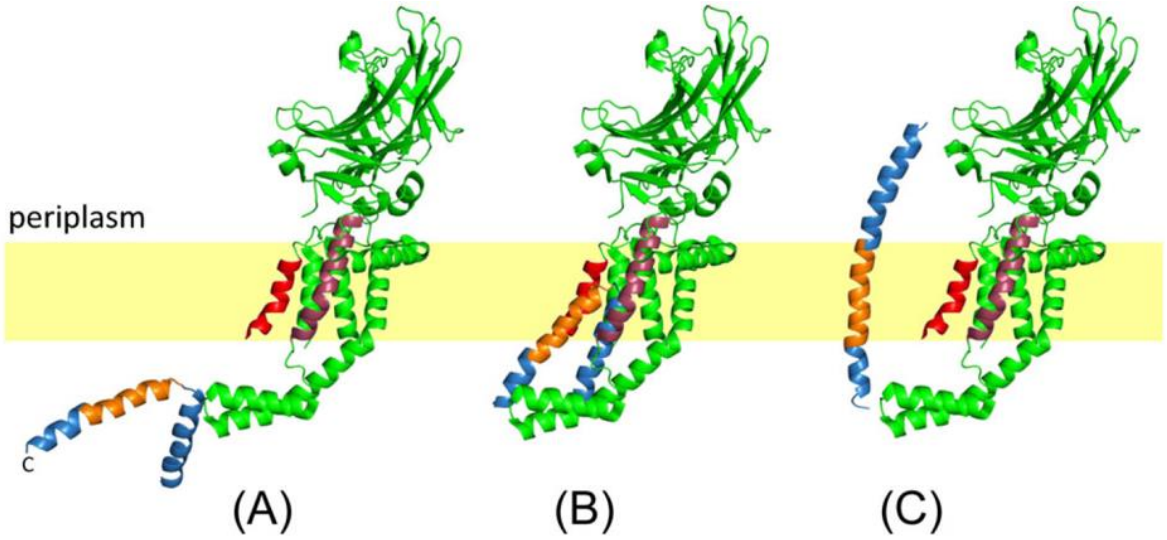


Figure 13: Hydrophobic slide mechanism for the insertion of the single-spanning membrane protein Pf3 coat. (A) Pf3 coat is attached to the membrane surface and subsequently recognized by the hairpin-like structure of the C1 loop. (B) The TMD of Pf3 coat (gold) interacts with the cytoplasmic part of the YidC greasy slide (TMD 3 = red and TMD 5 = purple). During the sliding mechanism of the pf3 coat TMD along the TMD 3 and 5 of YidC, the N-terminal hydrophilic domain is captured in the hydrophilic groove (C) The transmembrane segment is released into the bilayer between TMD 3 and TMD 5 of YidC whereas the hydrophilic N-tail is translocated to the extracellular side (Kiefer and Kuhn 2018).

After the interaction with the C1 domain, the hydrophobic TMD of the single-spanning membrane protein slides along the TMDs 3 and 5 of YidC to be integrated into the bilayer. This sliding mechanism pulls the hydrophilic N-tail of the substrate into the hydrophilic groove of YidC. After the sliding mechanism the hydrophilic part of the substrate, captured in the groove, is released into the periplasm, a step which could also involve the membrane potential (Fig. 13) (Dalbey and Kuhn 2014; Kuhn et al. 2017; Kiefer and Kuhn 2018). Most likely, the hydrophilic groove generates an environment sufficient for incorporation and release of the extracellular N-tail without the assistance of electrostatic interactions (Chen et al. 2014). The hydrophobic slide mechanism is also supposed for the insertion of the double-spanning M13 procoat, with the particularity that the two hydrophobic segments are probably inserted simultaneously forming a hairpin-like structure (Kuhn et al. 1986; Eisenhawer et al. 2001; Samuelson et al. 2001; Kuhn et al. 2017). Interestingly, M13 procoat can be inserted by the *S. mutans* YidC without the assistance of the important positively charged arginine in the hydrophilic groove (Chen et al. 2014). This is in contrast to Pf3 coat, where a substitution of the arginine residue to a neutral or a negatively charged residue prevented the insertion by the *S. mutans* YidC (Chen et al. 2014). It is reasonable, that the hydrophobic slide mechanism also matches for the *S. mutans* YidC during insertion of the double-spanning substrate M13 procoat. Probably, the higher hydrophobic driving force is sufficient for the insertion into the membrane (Chen et al. 2014). Cross-linking studies with the double-spanning membrane protein MscL also reveal that it forms a hairpin-like structure during the insertion process, since both hydrophobic segments are contacted by TMD 3 and TMD 5 of YidC (Neugebauer et al. 2012). Therefore, a similar sliding mechanism as for M13 procoat is assumed (Kuhn et al. 2017).

In contrast to the phage proteins M13 procoat and Pf3 coat, MscL is targeted to the YidC insertase via the SRP system (van Bloois et al. 2004). Therefore, it was shown for the first time that YidC is also able to interact with the SRP system although the Sec translocon is not involved. The interaction of YidC with the co-translational SRP pathway is mediated by contacts between RNCs and the cytoplasmic loop C2 and the C-terminal part of YidC (Kohler et al. 2009; Kedrov et al. 2013; Kedrov et al. 2016). In addition, direct contacts between SRP / FtsY and the C1 loop of YidC as well as the C-terminal part of YidC were found (Welte et al. 2012; Petriman et al. 2018). The weak binding affinity of non-translating ribosomes to YidC, is increased by the presence of a

nascent chain (Kohler et al. 2009; Kedrov et al. 2013). Therefore, it is assumed that the strongest interaction probably takes place at the nascent chain, since it is known that the low binding affinity for non-translating ribosomes of the *E. coli* YidC is due to the shorter C-terminal domain, compared to the YidC homologs from Gram-positive bacteria and mitochondria, as mentioned above (Kedrov et al. 2013). This suggests that the insertion of the hydrophobic signal sequence is responsible for a strong ribosome-YidC complex (Kuhn et al. 2017). In addition to the contact with RNCs, YidC is also able to bind directly to the targeting components of the SRP system (Welte et al. 2012). This contact is mediated by the C1 loop and the C-terminal part of YidC, however the interaction with SRP is stronger than the interaction with the receptor FtsY (Welte et al. 2012; Petriman et al. 2018). The detailed mechanism of the communication between the SRP system and YidC is still unknown and has to be analyzed in future studies.

Further novel insights into the YidC dependent insertion of inner membrane proteins were gained in 2012, when it was shown that YidC is responsible for the insertion of the C-tail anchored protein SciP from enteroaggregative *E. coli* (EAEC) (Aschtgen et al. 2012). The unique feature of C-tail anchored proteins is a large N-terminal cytoplasmic domain and a transmembrane domain located at the extreme C-terminal part of the protein. To date little is known about the insertion of C-tail anchored proteins in *E. coli*, however, detailed information about the insertion mechanism of this protein class in eukaryotes are available. The insertion of tail anchored (TA) proteins in the endoplasmatic reticulum (ER) occurs post-translational via the GET pathway (Fig. 14). After their synthesis the TMD is bound by a pre-targeting complex consisting of the proteins Get4, Get 5 and Sgt2. This pre-targeting complex transfers the TA protein to the ATP-bound Get3, delivering the protein to the membrane-bound Get1/2 complex through the interaction with the cytoplasmic domain of the membrane protein Get2. The protein is released from Get3 by ATP hydrolysis and further inserted into the ER membrane by Get1 (Fig. 14) (Denic et al. 2013).

Recent studies suggested that the Yeast Get1 protein is a conformational analog of the Oxa1 protein, therefore also the yeast Get pathway involves a protein from the conserved YidC/Oxa/Alb family (Anghel et al. 2017; Chen and Dalbey 2018). Hence, it is plausible that YidC is also involved in membrane insertion of C-tail anchored proteins in *E. coli.* However, in contrast to the other YidC substrates, little is known about the exact insertion mechanism of this special protein class.

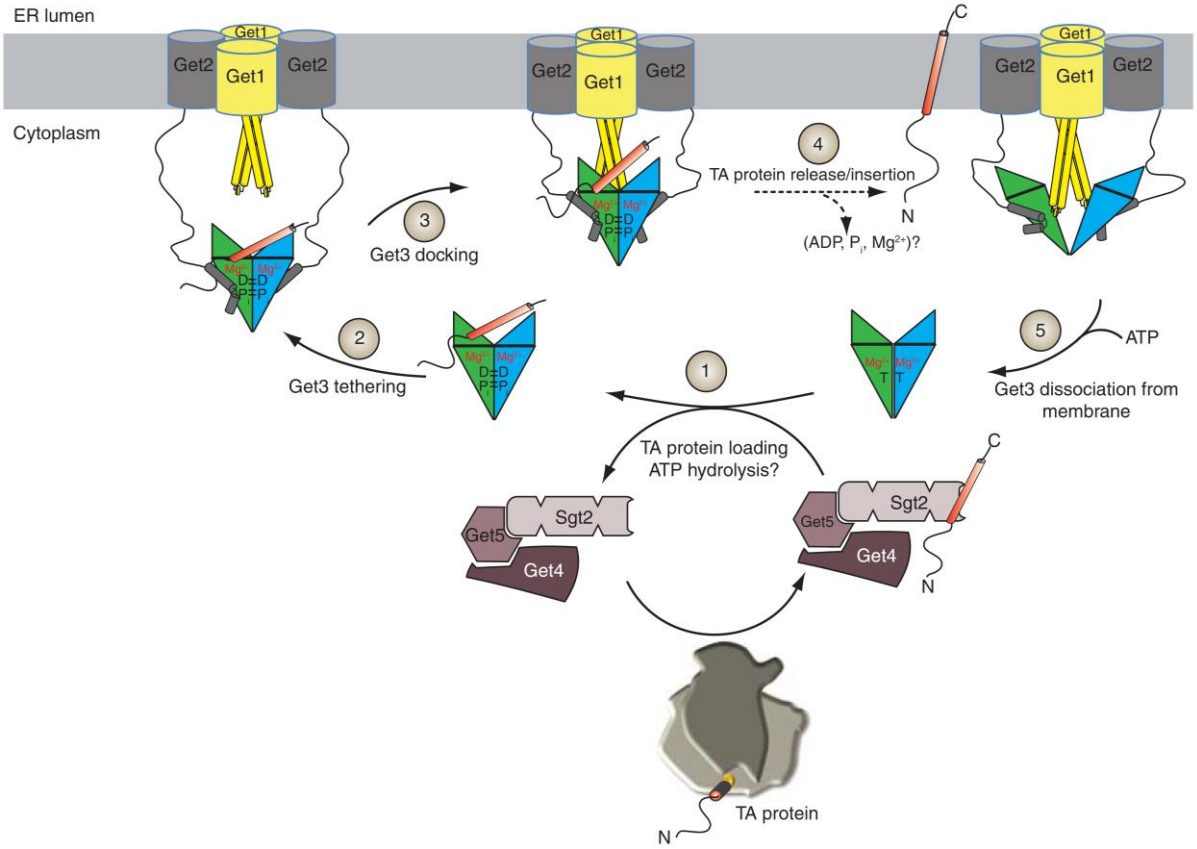


Figure 14: Model of the yeast GET pathway. After the release from the ribosome the newly synthesized tail-anchored (TA) protein is bound in the cytoplasm by the pre-targeting complex Sgt2/Get4/Get5 (1). The pre-targeting complex transfers the TA protein to the ATP-bound Get3 protein which is then targeted to the membrane via docking to the cytoplasmic domain of Get2 (2). Subsequently ATP hydrolysis results in the release of the TA protein from Get3 and it is forwarded to Get1 for membrane insertion (4). The Get3 protein is activated for the next targeting cycle by the release from Get1/2 through ATP binding (5) (Denic et al. 2013).

1.7 The C-tail anchored protein SciP from EAEC

The protein SciP from EAEC is part of the type 6 secretion system (T6SS) found in many pathogenic bacteria, like *Vibrio cholerae, Salmonella enterica* or the enteroaggregative *E. coli* strain (Boyer et al. 2009; Aschtgen et al. 2010a). T6SS is a multiprotein complex involved in the delivery of effectors, like small molecules, proteins and DNA, into target cells or the environment (Costa et al. 2015). This takes place as a direct response of the communication between bacteria and their environment as well as of the competition with other bacteria for space and resources (Durand et al. 2015). T6SS is one of the main molecular weapon of bacteria, that can target both eukaryotic and prokaryotic cells (Jani and Cotter 2010; Schwarz et al. 2010; Silverman et al. 2012; Durand et al. 2015). In Gram-negative bacteria there are six different

secretion systems present, named type I-VI secretion systems (Fig. 15) (Abdallah et al. 2007).

Some of these secretion systems share components of the previous discussed translocation or insertion systems used in bacteria, like the Sec or the Tat translocon, found in the inner membrane (Fig. 15) (Green and Mecsas 2016). The two translocons are found in those secretion systems that secrete the effector molecules in a two step mechanism across the cell envelope (Green and Mecsas 2016). In this two step secretion the effectors are first translocated into the periplasm and further across the outer membrane (Green and Mecsas 2016). The T2SS and T5SS are members of the two step secretion pathway (Fig. 15). In contrast T1SS, T3SS, T4SS and T6SS are members of the one step secretion pathway, delivering proteins from the cytoplasm directly to the outside of the cell (Fig. 15) (Abdallah et al. 2007; Green and Mecsas 2016). A special feature of T3SS, T4SS and T6SS is that the effectors cannot only be secreted across the own cell envelope but also across the membrane of the target cell directly into the cytoplasm (Fig. 15) (Green and Mecsas 2016).

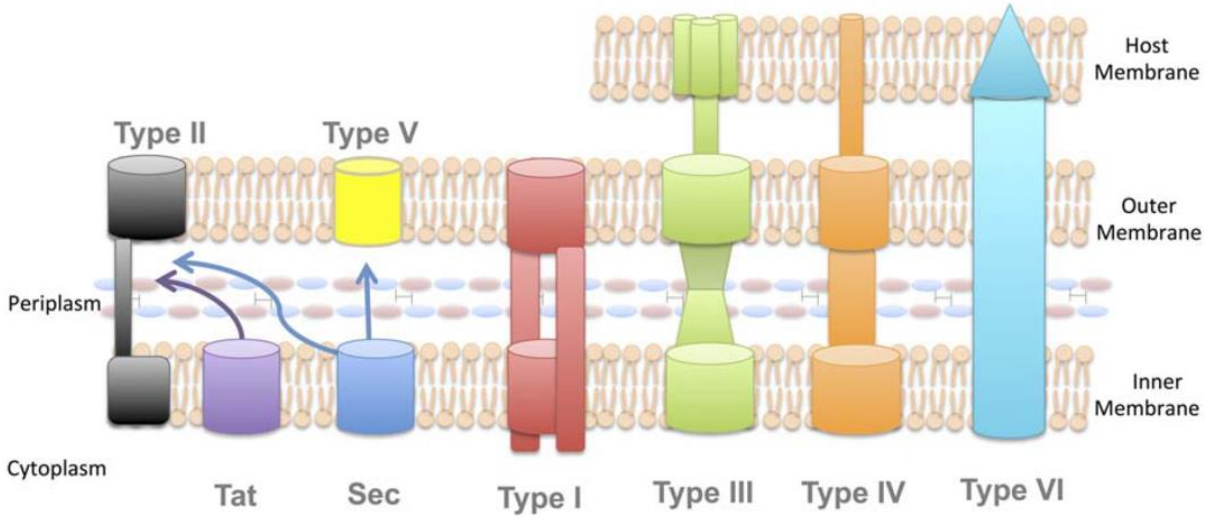


Figure 15: Secretion systems in Gram-negative bacteria. In Gram-negative bacteria there are 6 different secretion systems named type I-VI. They are divided into two classes, the one step and the two step secretion systems. Type II and type V belong to the two step secretion systems which involve the Sec or Tat translocon for the translocation of effectors into the periplasm. Type I, III, IV and VI belong to the one step secretion systems, which deliver effectors directly from the cytoplasm to the outside of the cell. Type III, IV and VI are able to transport the effectors across the membrane of a target cell (Green and Mecsas 2016).

T6SS is the latest discovered secretion system in Gram-negative bacteria (Mougous et al. 2006; Pukatzki et al. 2006). It is comprised of at least 13 components, the so-called core components, usually encoded by clustered genes (Shalom et al. 2007;

Cascales 2008). Among the different T6SS, the number of the components widely varies and most gene clusters encode for up to 25 components (Filloux et al. 2008; Aschtgen et al. 2012). T6SS forms a tubular puncturing device with a length of about 600 nm which shares structural similarities with bacteriophage T4 tail and baseplate components (Fig. 16) (Records 2011; Basler et al. 2012; Zoued et al. 2016). Thus, an evolutionary relationship is suspected (Records 2011).

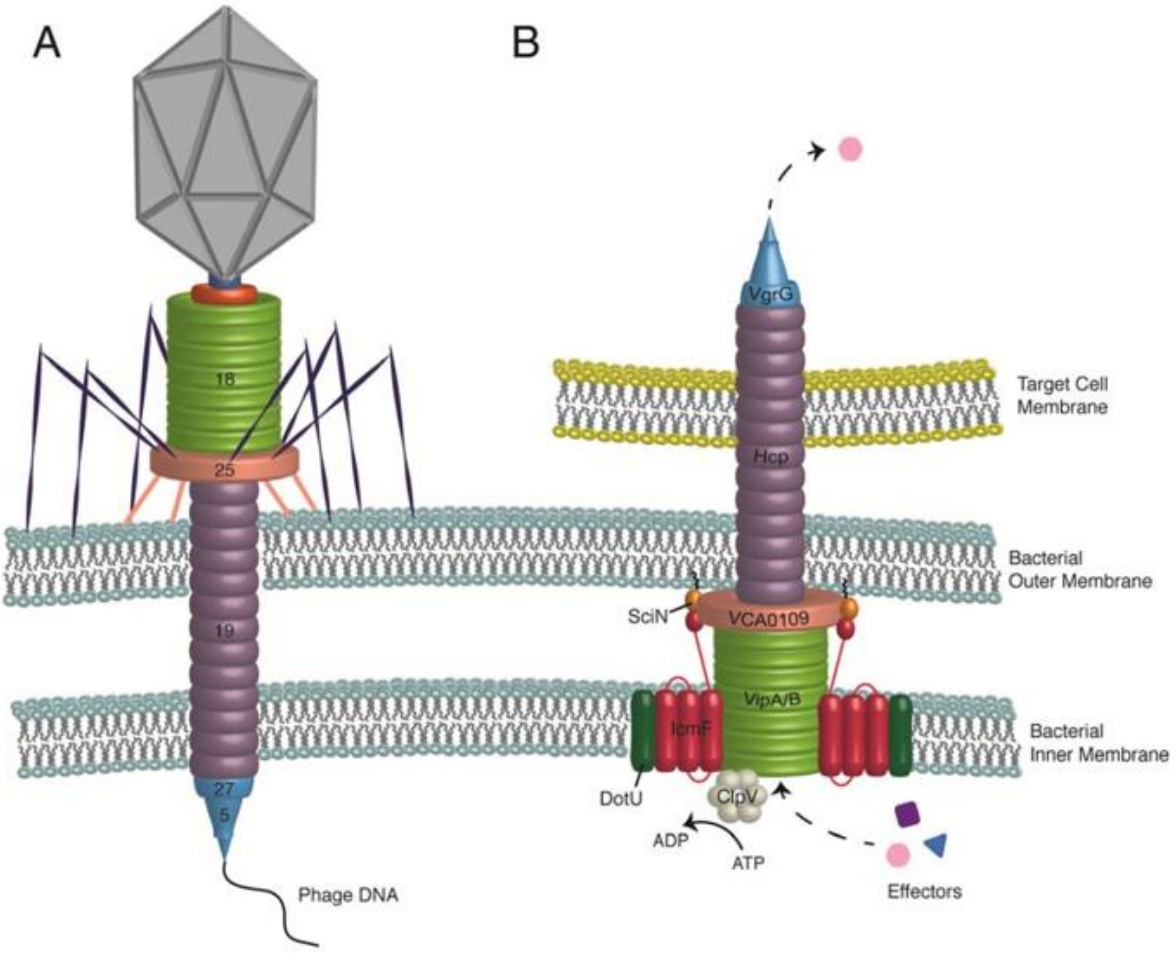


Figure 16: Structure and similarity of T4 bacteriophage and T6SS. T6SS looks like the inverted injection apparatus of T4 phage. The inner tube of T6SS consists of the protein Hcp and is wrapped by the sheath formed by the proteins VipA/B (TssB/TssC) which share structural similarities to gp19 and gp18 from T4 phage, respectively. The protein VgrG has a needle-like structure and is similar to gp27/gp5 of T4. The assembly of the sheath occurs on the baseplate build by the protein VCA0109 (TssE) or gp25 of T4 phage. The tail of T6SS is anchored to the membrane via the inner membrane proteins DotU (TssL, SciP), IcmF (TssM) and TagL (not shown in this picture) and the outer membrane protein SciN (TssJ) (Records 2011).

The inner tube is formed by stacked hexameric rings of the protein Hcp, similar to the protein gp19 of the T4 bacteriophage (Records 2011; Brunet et al. 2014). The inner

tube is wrapped by the sheath built by the proteins TssB and TssC (VipA and VipB), similar to the protein gp18 of T4 (Records 2011; Basler et al. 2012). At the top, the protein VgrG forms a needle-like structure, similar to gp5/gp27 of T4 (Records 2011; Shneider et al. 2013). This tubular structure has a diameter of about 300 Å, with a channel of about 100 Å in diameter (Bönemann et al. 2009). The sheath is assembled on the baseplate, similar to the baseplate of T4 build by gp25, consisting of the protein TssE (VCA0109) (Records 2011; Zoued et al. 2014). The tail is anchored to the membrane via the membrane complex build by the inner membrane proteins TssL (DotU, SciP), TssM (IcmF) and TagL and the outer membrane protein TssJ (SciN) (Aschtgen et al. 2010a; Aschtgen et al. 2010b). This trans-envelope complex serves as a channel for the inner tube after sheath contraction (Fig. 16) (Zoued et al. 2014; Durand et al. 2015).

The protein TssL or SciP represents an essential component of T6SS, since the secretion system is not assembled in absence of SciP (Aschtgen et al. 2012). SciP from EAEC is encoded by the gene *tssL* localized in the *sci1* gene cluster (Aschtgen et al. 2012; Durand et al. 2012). It consists of 217 amino acids with a molecular weight of about 24 kDa (Aschtgen et al. 2012). Topological studies showed that SciP has a N_{in} / C_{out} topology with a TMD at the very C-terminal part of the protein and a short periplasmic domain (Fig. 17) (Aschtgen et al. 2012). This topology does not depend on the assembly of T6SS, since SciP has the same topology in *E. coli* K12, a strain not coding for T6SS (Aschtgen et al. 2012). Software-based analysis predict the TMD between amino acid 184 and 206, resulting in a very short periplasmic domain of only 11 amino acids (Software tool TMHMM Server v. 2.0). Thus, SciP belongs to the special class of C-tail anchored proteins. In 2012, the crystal structure of the large cytoplasmic domain of SciP was solved (Fig. 17), indicating that SciP serves as a structural component, since no catalytic sites were found (Durand et al. 2012). In addition, no structurally related proteins that could indicate a molecular function, were found (Durand et al. 2012). The cytoplasmic domain of SciP is composed of 8 α-helices, forming two bundles of three helices (h1-h3 and h6-8), connected by two short helices (h4 and h5). Both bundles are forming a separate globular structure. It is notable that helix 7 and 8 are connected by a relatively large loop of 10 amino acids, with a central part consisting of charged amino acids (RSQDDE). The helix bundle is connected to the TMD by an elongated unstructured linker (amino acids 158-178) (Fig. 17).

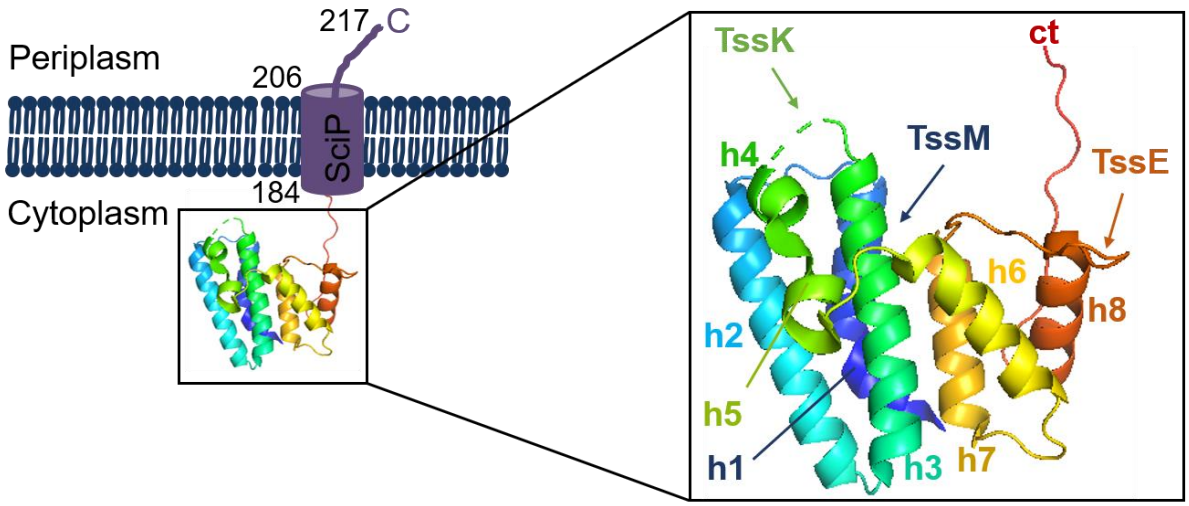


Figure 17: The C-tail anchored protein SciP from EAEC. SciP has one TMD between amino acid 184 and 206, a short periplasmic domain of 11 amino acids and a large N-terminal cytoplasmic domain of 183 amino acids. The crystal structure of the cytoplasmic domain shows 8 helices organized in two 3-helix bundles which are connected by the two shorter helices h4 and h5 (rainbow coloring from blue = N-terminus to red = C-terminus). The core region of the cytoplasmic domain is connected with the TMD via an elongated linker (red). The binding sites for the proteins TssK, TssM and TssE are indicated with arrows (Image created with PyMOL 2.3.0, PDB 3U66, adapted from Durand et al. 2012; Zoued et al. 2016).

Surface analysis indicates a scattered and balanced charge distribution over the whole surface and a huge cavity of about 1000 Å^3 between the base of the linker and the two helix bundles of the core domain. It is assumed that this groove at the interface of the two helix bundles is structurally important, since a large area of conserved residues within SciP homologs faces the groove (Durand et al. 2012). Recent studies showed that the cytoplasmic domain is also involved in binding to the T6SS components, TssE, TssM and TssK, necessary for the function of T6SS (Zoued et al. 2016). In addition, the anchoring of SciP in the inner membrane via the C-terminal TMD, which is also responsible for dimerization of the protein, is required for T6SS function and the stability of the protein (Aschtgen et al. 2012; Zoued et al. 2018).

However, little is known to date about the mechanism involved in the insertion of SciP into the membrane. A study from Aschtgen et al. in 2012 suggested that the insertion of SciP is dependent on YidC and modulated by DnaK but independent of SecYEG and SRP (Aschtgen et al. 2012). However, these first results have to be confirmed by further experiments since little is known about the insertion mechanism of C-tail anchored proteins in *E. coli* and only 11 known proteins in the *E. coli* strain K12 show

this topology (Borgese and Righi 2010). It is assumed that special insertion mechanisms are required for the insertion of SciP due to the large cytoplasmic domain and the TMD located 183 amino acids away from the N-terminus. This special feature is also found in another inner membrane protein, which is not part of the C-tail anchored proteins, but contains its first TMD 400 amino acids away from the cytoplasmic N-terminus. This protein is the potassium sensor protein KdpD from *E. coli* (Zimmann et al. 1995).

1.8 The sensor kinase protein KdpD

The sensor protein KdpD from *E. coli* is part of the high affinity potassium uptake system Kdp (Ballal et al. 2007). The *kdpDE* operon encodes for two proteins, which belong to the group of sensor kinase/response regulator systems (Parkinson and Kofoid 1992; Walderhaug et al. 1992; Zimmann et al. 1995). These proteins are responsible for the regulation of the *kdpFABC* operon encoding for the Kdp-ATPase which consists of the three membrane-bound proteins KdpA, KdpB and KdpC (Laimins et al. 1978; Polarek et al. 1992). Potassium is the major intracellular cation responsible for essential physiological processes like turgor homeostasis, pH regulation, gene expression and activation of cellular enzymes (Booth 1985; Epstein 1986; Suelter 1970; Prince and Villarejo 1990; Ballal et al. 2007). To maintain the turgor in bacterial cells there are several types of potassium regulation systems. These systems respond to the changes in the osmolarity of the medium by either the influx or efflux of potassium ions to adjust the intracellular potassium concentration (Ballal et al. 2007). *E. coli* encodes for three constitutive active potassium uptake systems, named TrkH, TrkG and Kup (TrkD) which are capable of maintaining the required intracellular potassium concentration under normal growth conditions (Dosch et al. 1991; Ballal et al. 2007). However, they have a low affinity for potassium ions (Ballal et al. 2007). The high affinity Kdp system is expressed as an emergency system during potassium limitation or osmotic upshock. In these cases the three constitutive active systems are no longer sufficient to maintain the required cellular potassium concentration (Zimmann et al. 1995; Ballal et al. 2007). The induction of the Kdp system starts with the recognition of a stimulus, like potassium limitation, by the membrane-bound sensor kinase KdpD (Fig. 18).

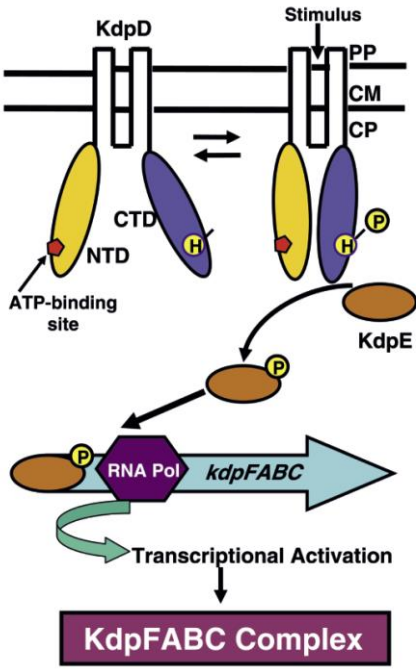


Figure 18: The Kdp system of *E. coli.* Recognition of a stimulus like potassium limitation by the membrane protein KdpD results in the autophosphorylation of KdpD. The phosphate group is then transferred to KdpE, which activates the transcription of the *KdpFABC* operon (Ballal et al. 2007).

It is assumed that the N-terminal domain of KdpD, including two motifs similar to the Walker A and Walker B motif of classical ATP binding sites, are involved in stimulus recognition. After autophosphorylation of KdpD the phosphate group is transferred to KdpE, a cytoplasmic response regulator. The phosphorylated KdpE protein recognizes and binds to a specific DNA sequence located upstream of the *kdpFABC* operon. Binding of KdpE induces the transcription of the genes in the *kdpFABC* operon encoding for the Kdp-ATPase (Fig. 18) (Ballal et al. 2007).

The sensor protein KdpD consists of 894 amino acids and is a four-spanning inner membrane protein with a large N-terminal (amino acid 1-400) and C-terminal (amino acid 499-894) cytoplasmic domain. The four TMDs are closely spaced in the center of the protein (Fig. 19). TMD 1 and TMD 2 as well as TMD 3 and TMD 4 are connected by short periplasmic loops of 4 and 10 amino acids (Zimmann et al. 1995; Facey and Kuhn 2003). The formation of a homodimer in the membrane is proposed to be important for the kinase activity (Heermann et al. 1998). Protease accessibility assays showed that the insertion of KdpD into the inner membrane is independent of SecA, SecE and YidC but dependent on the membrane potential (Facey and Kuhn 2003). KdpD can probably insert without the help of other protein components since only short

periplasmic loops have to be translocated which require less energy (Facey and Kuhn 2003). Interestingly the extension of the second periplasmic loop from 10 to 27 amino acids resulted in the requirement of YidC and SecYEG for insertion. The extension of the first periplasmic loop from 4 to 19 amino acids had no effect (Facey and Kuhn 2003). This confirms the assumption that only large periplasmic domains require the help of other components for efficient translocation. In the latest study it was shown that not only the pmf but also the SRP system is required for KdpD insertion into the inner membrane (Maier et al. 2008). This was surprising since SRP usually recognizes N-terminal TMDs located in the first 100 amino acids of an emerging nascent chain. However, KdpD has the first TMD 400 amino acids away from the cytoplasmic N-terminus (Maier et al. 2008). This study showed for the first time that SRP is also able to bind to short hydrophobic sequences apart from the TMD in the N-terminal cytoplasmic part of a protein. This then allows an early co-translational targeting of the nascent KdpD-RNC. The SRP signal sequence of KdpD encompasses the amphiphilic region between amino acid 22-48 (Fig. 19).

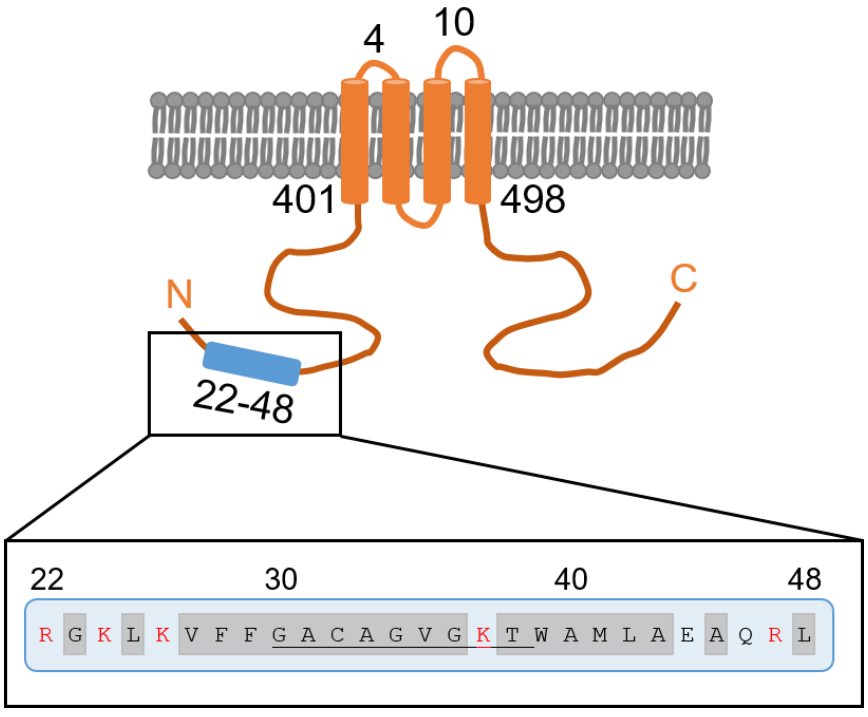


Figure 19: Topology of the sensor kinase KdpD from *E. coli*. The protein consists of 4 closely spaced TMDs and a large N- and C-terminal cytoplasmic domain of 400 and 396 amino acids, respectively. The N-terminal cytoplasmic domain contains an amphiphilic region between amino acid 22-48 which constitutes the SRP signal sequence. This sequence contains 3 positively charged residues (red), followed by a stretch of 10 hydrophobic amino acids (shaded grey) and another 6 hydrophobic amino acids. The Walker A motif (amino acid 30-38) is underlined.

Interestingly, this region also includes the Walker A motif (aa 30-38) required for the ATP-binding function of KdpD (Maier et al. 2008). The region contains 3 positively charged amino acids (aa 22, 24 and 26), followed by a stretch of 10 hydrophobic amino acids (aa 27-36) and another 6 hydrophobic residues (aa 38-43) (Fig. 19). Thus, the identified SRP signal sequence correlates well with other known SRP signal sequences.

For KdpD insertion it is still unclear how switching from targeting to insertion occurs, since weakening of the SRP-FtsY complex usually results in its dissociation and supports the binding of RNCs to the translocon (Draycheva et al. 2018). However, as mentioned before, KdpD is inserted independently of SecYEG and YidC and therefore, no translocon is necessarily involved in the insertion process. Further studies have to be performed to find out which mechanism triggers the SRP-FtsY dissociation in this special insertion pathway.

1.9 Aim of the thesis

The insertion of proteins into the inner *E. coli* membrane represents an essential process for cell viability. Over 30% of the proteome of all three domains of life has to be inserted into or translocated / secreted across the membrane (Wallin and Heijne 1998; Papanikou et al. 2007). In *E. coli,* several insertion or translocation pathways have been identified. Decoding the underlying insertion mechanisms in detail is still part of the current studies. In general, the insertion of inner membrane proteins can be divided into two main mechanisms, the post- or co-translational insertion (Cross et al. 2009). The advantage of co-translational insertion is that the emerging hydrophobic TMDs are protected from aggregation in the cytosol since the translating ribosome is already located at the membrane. This is mediated by the early binding of the emerging hydrophobic TMD by the signal recognition particle and the subsequent targeting of the ribosome nascent chain complex (RNC) to the inner membrane (Gier et al. 1996; Ulbrandt et al. 1997; Valent et al. 1998; Lee and Bernstein 2001).

However, the four-spanning *E. coli* inner membrane protein KdpD is targeted co-translationally by SRP although its first TMD is located 400 amino acids away from the cytoplasmic N-terminus (Maier et al. 2008). A short amphiphilic region in the N-terminal cytoplasmic part was identified as a SRP signal sequence (Maier et al. 2008). In the present study, the novel SRP signal sequence of KdpD was used to decode the crucial features of a signal sequence to be recognizable for SRP. Several signal sequence mutants were generated and their targeting abilities were monitored with *in vivo* subcellular localisation studies using fluorescence microscopy. In addition, stalled ribosomes exposing the SRP signal sequence of KdpD or the mutant signal sequences were purified to determine the binding affinity to SRP or a preincubated SRP-FtsY complex using microscale thermophoresis. The aim was to decipher whether the general binding mechanism of SRP is also valid for cytoplasmic SRP signal sequences.

The exceptional topology of KdpD of having a large N-terminal cytoplasmic domain, is shared by a special protein class, the C-tail anchored proteins. These proteins are characterized by a TMD at the extreme C-terminal part resulting in only short periplasmic domains. If the hydrophobic TMD emerges from the ribosome the protein is already almost completely synthesized so it is assumed that they are targeted and inserted in a post-translational way. How the *E. coli* cells cope with this special feature

and prevent the aggregation of the C-tail anchored proteins after transmembrane exposure is not yet understood. First results indicate that the membrane insertase YidC is involved in the insertion of C-tail anchored proteins (Aschtgen et al. 2012), but it is still unclear if they are targeted in a co- or post-translational way.

In this thesis, the aim was to decode the targeting and insertion mechanisms of the C-tail anchored proteins in *E. coli* by the use of the protein SciP since the insertion of this special protein class in prokaryotes is still unknown. Pioneering for the studies on SciP were the results obtained from the protein KdpD which showed that a large cytoplasmic domain does not rule out an early co-translational SRP-dependent targeting. The components involved in the targeting and insertion of SciP were deciphered with *in vivo* translocation studies using protease accessibility assays and periplasmic cysteine modifications in various depletion strains. In addition, the detailed membrane targeting mechanism of SciP was analyzed with different SciP-sfGFP fusion proteins and *in vivo* subcellular localisation studies. Further, artificially stalled ribosomes exposing various SciP sequences are used to analyze whether C-tail anchored proteins are interacting with SRP during their synthsis.

Taken together these studies provide further insights into the targeting and insertion mechanisms of inner membrane proteins in *E. coli*.

2

The SRP signal sequence of KdpD

Eva Pross & Andreas Kuhn

Institute of Microbiology and Molecular Biology,
University of Hohenheim, 70599 Stuttgart, Germany

Scientific Reports 2019 9 (1): 8717

Abstract

KdpD is a four-spanning membrane protein that has two large cytoplasmic domains at the amino- and at the carboxyterminus, respectively. During its biogenesis KdpD binds to the signal recognition particle (SRP) of *Escherichia coli* that consists of a 48-kDa protein Ffh and a 4.5S RNA. The protein is targeted to the inner membrane surface and is released after contacting the SRP receptor protein FtsY. The information within the KdpD protein that confers SRP interaction was found in the amino-terminal cytoplasmic domain of KdpD, particularly at residues 22-48. Within this sequence a Walker A motif is present at residues 30-38. To determine the actual sequence specificity to SRP, a collection of mutants was constructed. When the KdpD peptides (residues 22-48) were fused to sfGFP the targeting to the membrane was observed by fluorescence microscopy. Further, nascent chains of KdpD bound to ribosomes were purified and their binding to SRP was analysed by microscale thermophoresis. We found that the amino acid residues R22, K24 and K26 are important for SRP interaction, whereas the residues G30, G34 and G36, essential for a functional Walker A motif, can be replaced with alanines without affecting the affinity to SRP-FtsY and membrane targeting.

Introduction

In *Escherichia coli,* the cytoplasmic signal recognition particle (SRP) mediates co-translational targeting of membrane proteins by binding to a “signal sequence” and generating a ribosome nascent chain complex (RNC)[1,2]. The RNC then binds to a membrane-associated SRP receptor, FtsY for membrane targeting[3]. SRP is universally conserved in its core region that consists of a ribonucleoprotein particle, the SRP RNA (4.5S RNA in *E. coli*) and the protein component SRP54 (Ffh in *E. coli* for “fifty-four-homolog”) that binds to the conserved RNA domain IV[4]. However, the composition of SRP varies among the different organisms with the most evolved version found in eukaryotes[5]. In contrast to the signal sequences of exported proteins, the bacterial SRP signal sequences are more hydrophobic and are mostly “uncleaved signal sequences” present in membrane proteins that remain in the final protein-chain as transmembrane anchor sequences. In general, the signal for insertion into the inner bacterial membrane is located in the first hydrophobic transmembrane domain and insertion is catalysed by the Sec translocase and / or YidC insertase.

SRP is bound to the ribosome and is ready to interact with a nascent protein chain. The amino-terminal NG domain of SRP is bound to the ribosomal proteins uL23 and uL29, next to the tunnel exit and the carboxy-terminal M domain to the ribosomal 23S RNA[6]. SRP at the ribosomal exit tunnel scans a nascent chain for bearing a hydrophobic SRP signal sequence[7]. The presence of such an emerging SRP signal sequence causes a tight binding to the hydrophobic groove of the SRP M domain[8]. This RNC-SRP complex then interacts with the membrane-associated SRP receptor, FtsY[3]. There, SRP and FtsY engage into a tight complex[3,9] which results in the coordinated activation of the SRP/FtsY GTPase activities in Ffh and in FtsY at the membrane, essential for protein translocation[10,11]. After GTP hydrolysis, the RNC-SRP-FtsY complex is dissociated and the nascent chain is then released[12] to further interact with the membrane, the Sec translocase or the YidC insertase[13,14]. In addition, SRP and FtsY are programmed for the next targeting cycle.

In this “SRP pathway”, an open question is how a signal sequence is recognized since no consensus motif exists and is relevant for SRP binding. With cross-linking studies, it could be shown that the hydrophobicity of the signal sequence is crucial for SRP binding, since lowering the hydrophobicity resulted in less efficient cross-linking[15,16]. The increase of the hydrophobicity in the signal sequences of *E. coli* presecretory proteins makes it possible to re-route the SecA dependent preproteins into the SRP targeting cycle[17]. Furthermore, it is suggested that the binding of SRP to the signal sequence is promoted by the presence of basic amino acid residues through electrostatic interactions[18]. The methionine-rich M domain of the Ffh protein binds hydrophobic residues of the substrate protein and accommodates the SRP-signal sequence in a hydrophobic groove[19]. A high resolution structure of Ffh from *E. coli* with a bound signal sequence showed that it was bound to the M-domain[6]. The structure shows that the signal sequence is sandwiched between the αM1 and αM4 helices of the Ffh protein and a hairpin loop of the ribosomal protein uL24. This binding involves both, hydrophobic- and electrostatic-mediated contacts.

Some inner membrane proteins have a large N-terminal cytoplasmic domain, like the sensor protein KdpD with the first transmembrane segment (TMS) starting at amino acid position 400. Previous data have suggested that KdpD has its signal sequence in the cytoplasmic domain, in a short amphiphilic sequence (aa 22-48) that targets the SRP complex to the inner membrane[20].

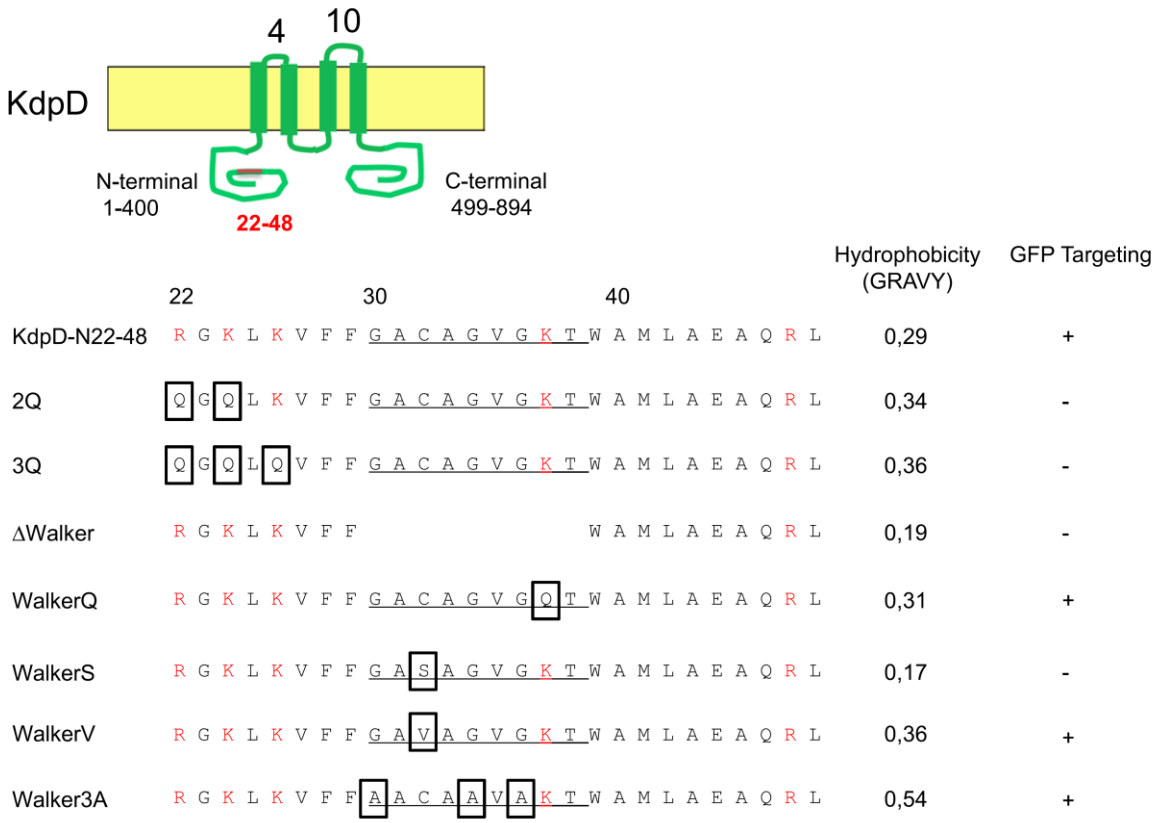


Figure 1. Topology of KdpD and the amino acid sequence of the N-terminal region (aa 22-48) of KdpD that is involved in SRP interaction. The KdpD protein is a four-spanning membrane protein with two large cytoplasmic domains at the N-terminus and the C-terminus. The SRP signal peptide of KdpD is located in the N-terminal domain at residues 22-48 and contains five positively charged residues (red letters) and a Walker A motif (underlined). The mutations studied here are marked with a box. Site-directed mutagenesis was used to alter the amino acids in the peptide. The hydrophobicity of each mutant is indicated.

After targeting, the amphiphilic sequence is released from SRP and folds into an ATP binding domain. During the ongoing translation the transmembrane segments of KdpD are exiting the ribosome and can readily insert into the membrane even if YidC or SecYEG have been depleted in cells[21]. The KdpD signal sequence contains five positively charged residues, in which three are closely spaced (aa 22-26) and a stretch of 10 hydrophobic residues (aa 27-36) which is too short to span the membrane. The peptide contains also a Walker A motif, which is similar to classical ATP binding sites[22,23]. In the present study, we investigated in detail the involvement of the Walker A motif and the positively charged residues in the signal sequence binding to SRP and the membrane targeting using sfGFP fusion proteins. We found that when the glycine residues at positions 30, 34 and 36 are replaced by alanines the sfGFP fusion protein was still targeted to the membrane surface although these mutations destroy the Walker Box. In contrast, when the positively charged

residues outside the Walker box at positions 22, 24 and 26 are replaced by glutamines membrane targeting was inhibited.

Results

Membrane targeting of N22-48-sfGFP depends on SRP.

The integral sensor protein KdpD consists of two large cytoplasmic domains, located at the N- and C-terminus (1-400 and 499-894), which are separated by four closely spaced transmembrane segments (401-498). Recently, it was shown that the amino acid residues 22-48 at the beginning of the N-terminal cytoplasmic region of KdpD (Fig. 1) serve as the SRP signal sequence[20]. In order to visualize the localization of N22-48 of KdpD we fused the N-terminal peptide (22-48) to sfGFP. To verify the involvement of SRP in the membrane targeting of the N22-48-sfGFP, we now analysed the localization of the N22-48-sfGFP in the Ffh-depletion strain MCΔFfh and found that SRP is required for membrane targeting of N22-48-sfGFP. The MC1061 cells expressing N22-48-sfGFP were examined under the fluorescence microscope and found at the inner membrane (Fig. 2a). Likewise, when the MCΔFfh cells were grown in presence of arabinose to express Ffh we also found the fusion protein at the surface of the inner membrane (panel b). However, when the cells were grown without arabinose and did not express Ffh (panel c, Fig. S3) most of the fluorescent signal appeared as punctuates in the cells that are indicative of aggregate formation. For controls, the sfGFP without the KdpD peptide remained in the cytoplasm under all the conditions (Fig. S1).

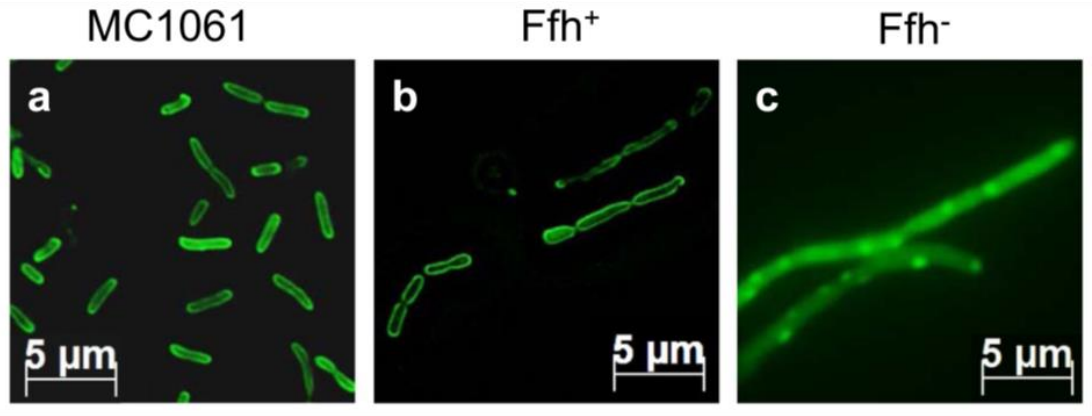


Figure 2. Membrane targeting of the KdpD signal peptide depends on SRP. Expression of N22-48-sfGFP in *E. coli* MC1061 (a) and in the depletion strain MCΔFfh (b, c) was induced for 30 min (a) and 20 min (b, c) at 37 °C and followed by fluorescence microscopy. Under arabinose growth conditions Ffh is present in the cells and allow the N22-48-GFP fusion protein to target to the membrane surface (b). Under glucose conditions, Ffh is depleted and the fluorescence signal of N22-48-sfGFP is patchy or distributed throughout the cytoplasm (c). Non-fused sfGFP was always distributed throughout the cytoplasm, regardless whether Ffh was present or not (Fig. S1).

SRP signal mutants of KdpD-N22-48-sfGFP.

Mutants were generated to explore the function of the N22-48 sequence listed in Fig. 1. The positively charged residues at 22, 24 and 26 were substituted with glutamyl residues giving the 2Q and 3Q mutants. In addition to hydrophobicity, it is assumed that the presence of basic amino acid residues promotes signal sequence binding to SRP[18]. The KdpD signal element contains a Walker A motif (residues 30 to 38), which is very similar to the classical ATP binding site. We constructed a mutant where the entire Walker box element was deleted (ΔWalker). The influence of the residues within the Walker box on SRP-dependent membrane targeting of KdpD was investigated by a number of mutants generated in the Walker motif. The conserved lysyl residue in the Walker box was exchanged with a glutamyl in the WalkerQ mutant, the cysteine at residue 32 was substituted with a serine or valine and the conserved glycines at 30, 34 and 36 were exchanged with alanines in the Walker3A mutant. All mutant proteins were expressed in *E. coli* MC1061 for 30 min and analysed by SDS-PAGE (Fig. S2).

To analyse the effects of substituting the positively charged residues of 2Q and 3Q, the sfGFP fusion proteins were expressed in MC1061 (Fig. 3). The cells were grown to a density of 0.5 at OD_{600}, induced with 1 mM IPTG and grown for 30 min at 37 °C. The cells were then applied for the fluorescence microscopy. In contrast to KdpD22-48 the fluorescence of the 2Q and 3Q proteins was detected uniformly in the cytoplasm (a, b). This shows for both mutants that the membrane targeting of the sfGFP fusion proteins was clearly inhibited. We conclude that the positively charged residues in the SRP signal are essential to target sfGFP to the membrane surface.

The residues 30 to 38 resemble the Walker motif. When these residues were deleted, the cells accumulated the sfGFP fusion protein in large aggregates at the poles of the cells (Fig. 3c). With the WalkerQ mutant where the conserved lysyl residue was substituted by a glutamyl we observed a clear fluorescence signal at the membrane (Fig. 3d).

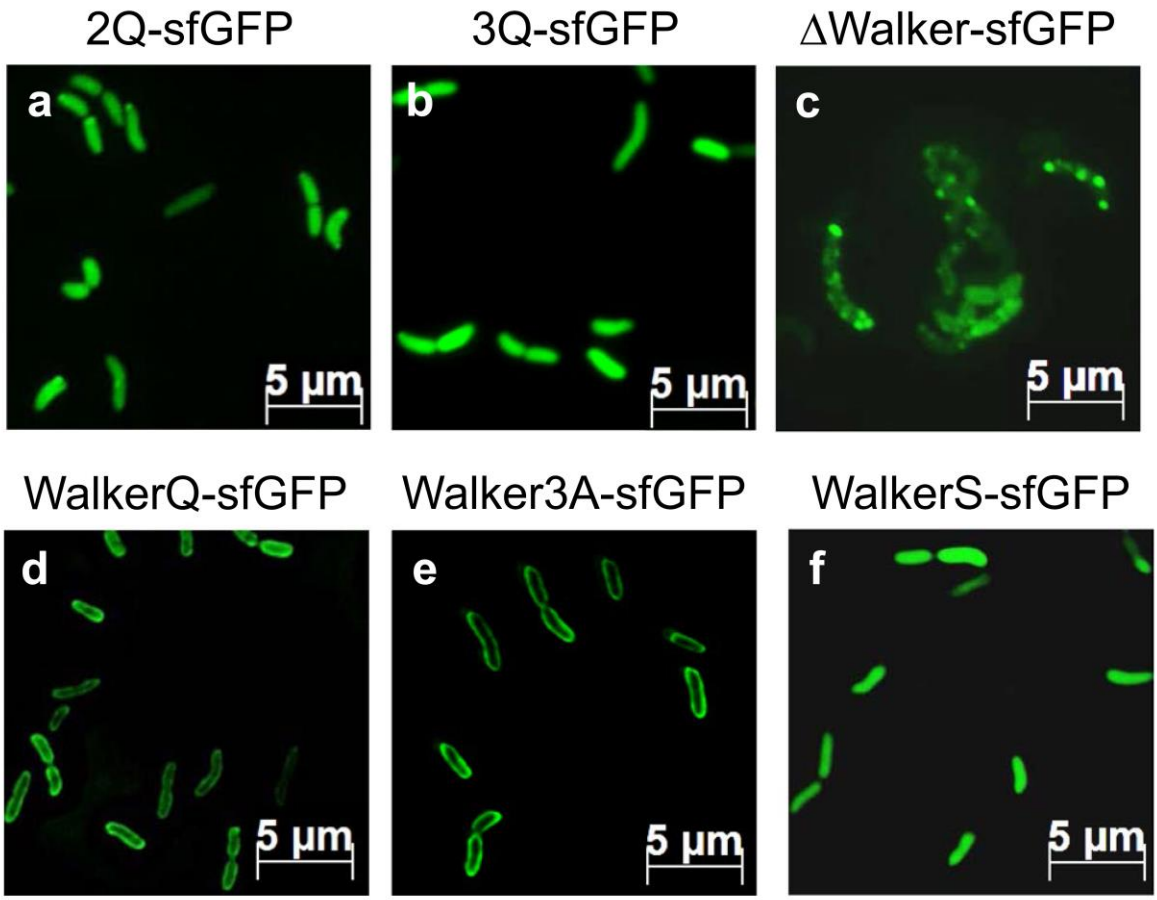


Figure 3. Targeting of the KdpD-GFP mutants. The KdpD-sfGFP fusion mutants 2Q (a) and 3Q (b) have 2 or 3 positively charged residues mutated to glutamines, respectively. The KdpD-sfGFP fusion mutant ΔWalker (c) and WalkerQ (d) have all residues of the Walker box deleted, WalkerQ has the conserved lysine residue 37 mutated into glutamine (d), Walker3A has the 3 glycine residues at residues 30, 34 and 36 mutated to alanines (e) and WalkerS has the cysteine residue 32 mutated to a serine (f), respectively. They were expressed in *E. coli* MC1061 at 37 °C and induced with 1 mM IPTG for 30 min. Except for the WalkerQ and the Walker3A mutant, the fluorescent signal was found distributed throughout the cytoplasm or in patches.

Next, the function of the glycyl residues in the Walker box was investigated. The Walker3A mutant with the conserved glycines exchanged to alanines was expressed in MC1061 (Figs 3e, 4a) and in MCΔFfh cells (Fig. 4b, c). When Ffh is present as is in MC1061 (Fig. 4a) and when the MCΔFfh cells were grown with arabinose (panel b), the fluorescence was found at the membrane indicating that the mutant protein was clearly targeted to the membrane. The targeting of this mutant was also sensitive to the depletion of Ffh (panel c). We conclude that the WalkerQ and the Walker3A mutants are fully functional to interact with SRP and are targeted to the membrane surface. When SRP is depleted the targeting of these mutants was inhibited similar to the wild-type sequence (Fig. 2).

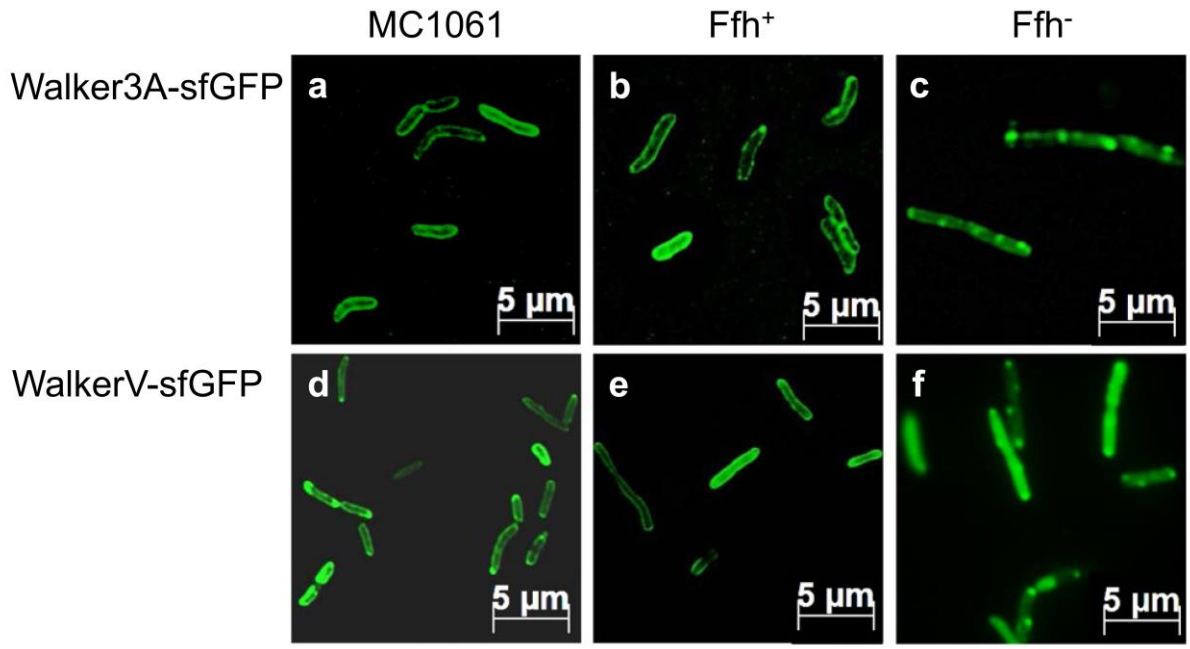


Figure 4. SRP-dependence of the functional Walker mutants. Localization of the Walker3A-sfGFP (a–c) and the WalkerV-sfGFP (d–f) mutant proteins in *E. coli* MC1061 (a, d) and MCΔFfh (b, c, e, f) after induction with 1 mM IPTG for 30 min (a, d) and 20 min (b, c, e, f) at 37 °C. When the cells were grown to express Ffh the fusion protein was localized at the membrane, whereas after depletion of Ffh (c, f) the fluorescence signal was found patchy at the cell poles or distributed in the cytoplasm. The depletion of Ffh was verified by Western blot (Fig. S3). The mutation of cysteine in valine at position 32 restores the targeting defect of the WalkerS mutant in *E. coli* MC1061 (Fig. 3f), where the cysteine residue was mutated into serine.

When the cysteine residue at 32 in the Walker box was replaced by a serine the targeting to the membrane was inhibited (Fig. 3f). This might be because the hydrophobicity of the signal sequence is reduced (Fig. 1). Replacing the serine residue by the more hydrophobic residue valine restored the membrane targeting of the mutant (Fig. 4d). Also, the targeting of the WalkerV mutant was still sensitive to the depletion of Ffh (panel f). We conclude that not the cysteine residue at position 32 is important for membrane targeting but the hydrophobicity since mutation to valine with a comparable hydrophobicity restored the membrane targeting defect of the WalkerS mutant.

Binding of KdpD22-48 to the hydrophobic groove of SRP.

To explore whether the SRP signal sequence of KdpD is recognized by the hydrophobic groove formed by the M domain of SRP, cross-linking experiments with purified SRP and synthesized KdpD peptides 22-48 were performed. Since the wildtype Ffh has a cysteine residue at position 406 in the M domain disulfide cross-

linking was tested with the *in vitro* synthesized KdpD22-48 peptide that has a cysteine residue at position 32. In addition, two mutant Ffh proteins with a cysteine residue at position 181 in the G domain or at 423 in the M domain pocket were analysed. Both mutants were combined with a serine at 406. After mixing the proteins cross-linking with copper phenanthroline was performed for 1 h on ice. The samples were analysed on an SDS-PAGE with and without DTT, respectively (Fig. 5). When the KdpD peptide was incubated with the wildtype SRP having a cysteine residue in the M domain or at position 423 in the M domain an additional band appeared indicating that the peptide was cross-linked to SRP (lane 5, 7). The cross-link was sensitive to the reducing agent DTT (Fig. 5, lane 12, 14). The shifted band did not appear when KdpD or SRP alone was incubated with the cross-linking agent. In contrast to SRP and SRP 423, the addition of the copper phenanthroline to the KdpD peptide with SRP L181C (in the G domain) showed no cross-linked band (lane 3).

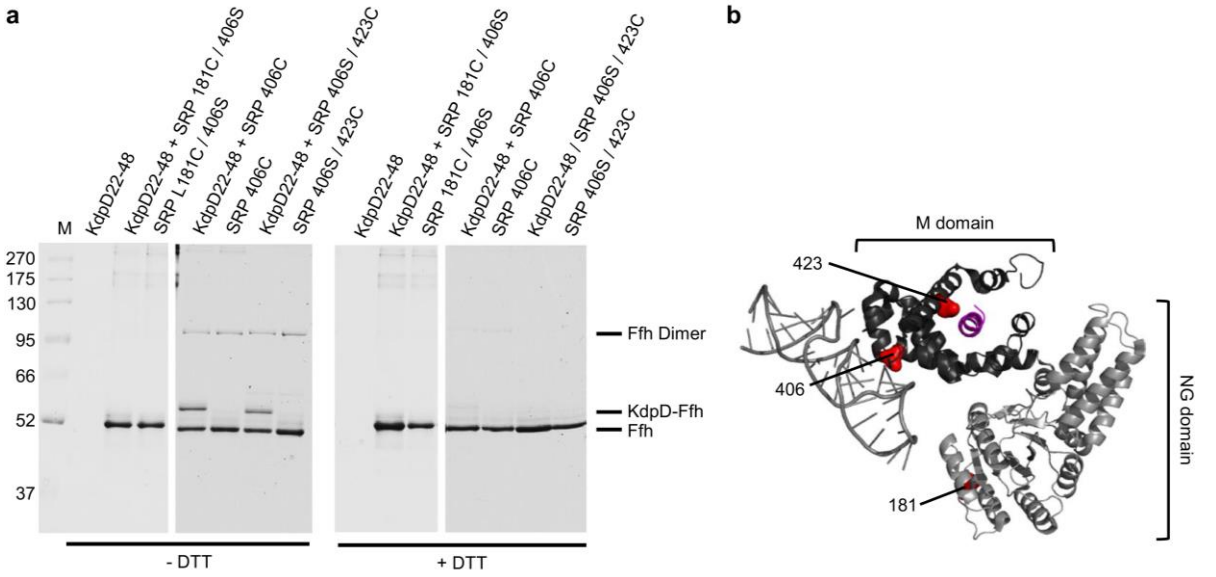


Figure 5. *In vitro* cross-linking of KdpD22-48 with SRP. (a) Copper phenanthroline crosslinking of the *in vitro* synthesized KdpD22-48 peptide with wildtype Ffh (406C), Ffh 181C/C406S or Ffh 406S/423C. 2 µM of reconstituted SRP was incubated with 20 µM of the peptide and 1 mM copper phenanthroline for 1 h on ice. Samples were TCA precipitated, resuspended in buffer with or without DTT and analysed by SDS-PAGE. The crosslinked peptide to Ffh 406C and 423C in the M domain generated a shifted band (KdpD-Ffh) but not for the 181C mutant in the G domain of SRP. For controls, the peptide and the proteins alone were analyzed. (b) The crystal structure of *E. coli* SRP (Pdb: 5GAD, the ribosome and FtsY are not shown for clarity) is displayed in grey, the signal sequence in purple and the positions of the cysteine residues are highlighted as red dots.

In conclusion, our results are consistent with the KdpD signal sequence binding in the same position of SRP as other and canonical signal sequences do[6,24].

Binding of the KdpD mutants to SRP.

To test whether the observed membrane targeting abilities of the mutants correlate with binding capabilities to SRP we analysed the interaction of the 22-48 peptide with the wild-type sequence, the 3Q mutant and the Walker3A mutant with purified SRP employing microscale thermophoresis (MST).

Ribosome-nascent chains (RNC) were designed by introducing the sequence for KdpD22-74 or the mutant sequences in a TnaC stalling plasmid resulting in tryptophan dependent ribosome stalling. Since there are about 30 amino acids in the ribosomal exit tunnel the KdpD peptide 22-48 was expected to be exposed outside the tunnel, so that the accessibility is guaranteed. The different his-tagged RNCs were purified and labeled with the fluorescent dye NT-647. The interaction studies of RNCs and purified SRP were analyzed using microscale thermophoresis with a fixed concentration of 5 nM for the labeled RNCs and varying concentrations of 1 µM to 0.5 nM for unlabeled SRP. First, RNCs encoding 22-74 as nascent chain were incubated with SRP resulting in a binding event comparable with the positive control, where RNCs with amino acids 4-85 of the SRP substrate FtsQ as nascent chain were used (Fig. 6a). As a negative control, ribosomes with amino acids 2-50 of the cytoplasmic protein firefly luciferase as a nascent chain were incubated with SRP, which showed no binding in this concentration range (a). In addition, RNCs with only a short nascent chain of about 13 amino acids, mimicking ribosomes at the very beginning of translation, were used as an additional control. Like for Luc2-50-RNCs, also these short-chain ribosomes showed no binding to SRP under these conditions (a). The exchange of three conserved glycines in alanines in the Walker A motif in the KdpD nascent chain did not affect the SRP binding and was similar to that of KdpD22-74 (Fig. 6b). In contrast, the exchange of three positively charged amino acids at position 22, 24 and 26 resulted in a slightly different binding indicating a reduced affinity to SRP. The exchange of the cysteine at position 32 in the Walker A motif to a serine residue led to a very weak SRP binding resulting in the fact that no saturation could be reached in the measured concentration range. We conclude that both mutants that were inhibited for membrane targeting were affected in binding. Therefore, the positively charged N-terminal part of the peptide and the hydrophobicity of the core region are critical for SRP binding.

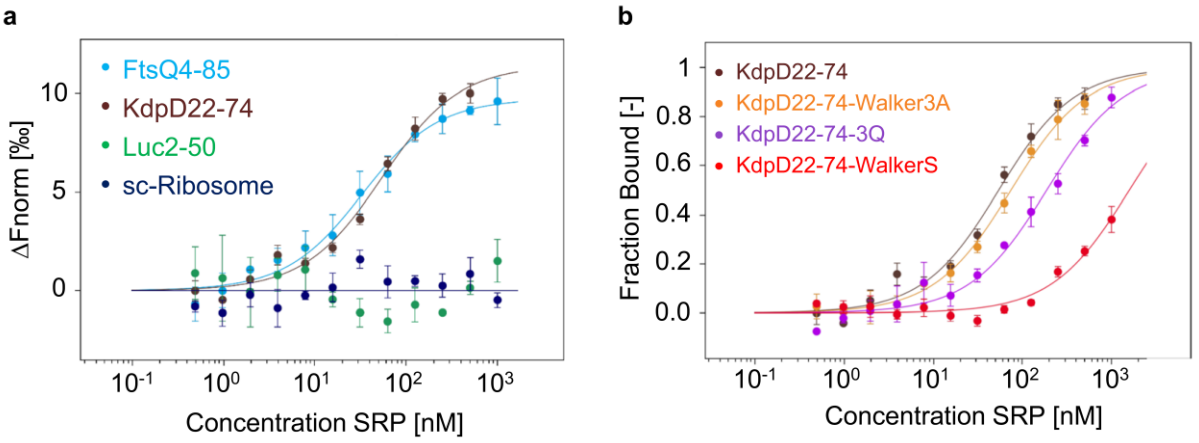


Figure 6. Binding of ribosome nascent chains (RNC) to SRP. Microscale thermophoresis (MST) measurements of unlabeled SRP (1 µM to 0.49 nM) with labeled RNCs (5 nM) of (a) FtsQ4-85 (blue dots), KdpD22-74 (brown dots), Luc2-50 (green dots), short-chain-ribosomes (dark blue dots) and (b) with KdpD22-74-Walker3A (orange dots), KdpD22-74-3Q (purple dots), KdpD22-74-WalkerS (red dots). After a 5 min incubation on ice the dilutions were filled into Premium capillaries (NanoTemper Technologies) for the MST measurements. In (a), the change in normalized fluorescence (ΔFnorm in ‰) and in (b), the fraction bound, is plotted against the concentration of unlabeled SRP (3 independently pipetted measurements, error bars represent the standard deviation). The raw data of the measurements are shown in Figs S4 and S5.

Binding of the KdpD mutants to SRP-FtsY.

After the SRP-dependent membrane targeting the SRP-nascent-chain complex is expected to interact with the SRP receptor FtsY. The binding of FtsY to SRP results in conformational changes leading to a stronger binding of SRP to the nascent chain[10]. This only occurs if SRP is bound to a correct cargo. Therefore, we analyzed whether the different KdpD-nascent chains are able to bind to a closed SRP-FtsY complex containing GTP. For the MST binding experiments, the reconstituted SRP was incubated with purified FtsY in a molar ratio of 1:4 in presence of 200 µM of the non-hydrolysable GTP analogue GppNHp. MST experiments were done with a fixed concentration of 5 nM for the labeled RNCs and varying concentrations of 250 nM/ 500 nM to 0.12/0.24 nM for unlabeled SRP-FtsY complex. First, RNCs with amino acids 22-74 of KdpD as a nascent chain were incubated with the closed SRP-FtsY complex. The MST measurements showed that the SRP-FtsY complex is able to bind to the RNCs exposing residues 22-48 of KdpD (Fig. 7a). In contrast, the negative control with residues 2-50 of firefly luciferase showed no binding with the SRP-FtsY complex under these conditions (Fig. 7a). From this we conclude that RNCs with residues 22-74 of KdpD represent a correct cargo for SRP. Also, the interaction with

KdpD-RNCs where the three conserved glycines in the Walker A motif were mutated in alanines showed an efficient binding to the SRP-FtsY complex (Fig. 7b). Thus, the exchange of the glycines to alanines known to be essential for the function of the Walker motif did not affect SRP-FtsY binding. In contrast, the mutant where the three positively charged amino acids at position 22, 24 and 26 were substituted with glutamines showed a weaker binding to the SRP-FtsY complex as was with only SRP (Fig. 7b). Since even in this case binding to a closed SRP-FtsY complex was observed we conclude that SRP-FtsY binding is not inhibited but is weakend after the exchange of the three positively charged amino acids. The WalkerS mutant with the lowest hydrophobicity (GRAVY = 0.17; GRAVY 22-48 = 0.29) showed the weakest binding to SRP. Therefore, the substitution of the cysteine residue with a serine residue prevents binding of the SRP-FtsY complex (Fig. 7b). Taken together, our experiments show that both, the positively charged N-terminal part of the sequence and the hydrophobicity of the core region are critical for SRP binding and, in addition, for FtsY recruitment.

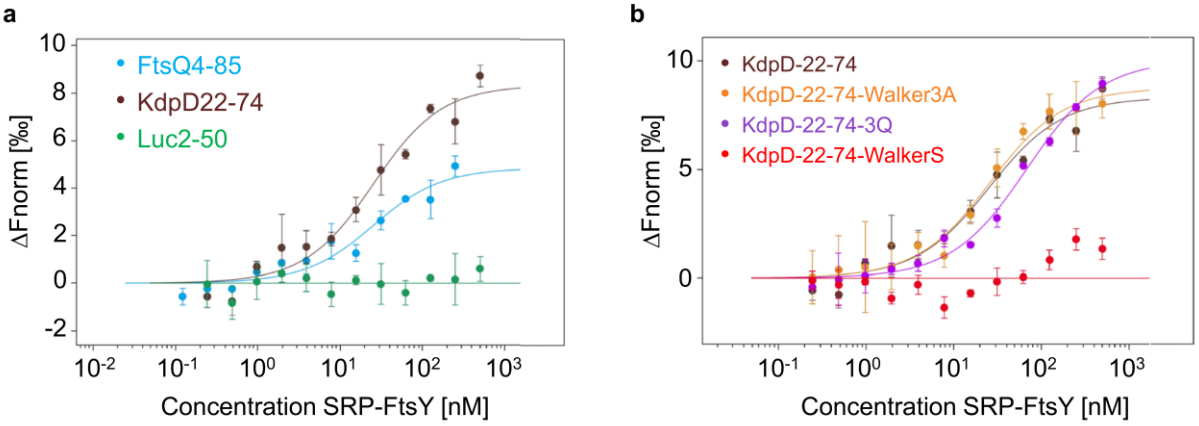


Figure 7. Binding of ribosome nascent chains (RNC) to SRP-FtsY. Microscale thermophoresis (MST) measurements of unlabeled SRP-FtsY complex (500/250 nM to 0.25/0.12 nM) with labeled RNCs (5 nM) of (a) FtsQ4-85 (blue dots), KdpD22-74 (brown dots), Luc2-50 (green dots) and (b) with KdpD22-74-Walker3A (orange dots), KdpD22-74-3Q (purple dots), KdpD22-74-WalkerS (red dots). After a 5 min incubation on ice the dilutions were filled into Premium capillaries (NanoTemper Technologies) for the MST measurements. The change in normalized fluorescence (ΔFnorm in ‰) is plotted against the concentration of the unlabeled SRP-FtsY complex (3 independently pipetted measurements, error bars represent the standard deviation). The raw data of the measurements are shown in Figs S6 and S7.

Discussion

The amino-terminal region of the KdpD sensor protein has two functions. In the fully folded protein it binds ATP that might modulate the communication of KdpD with its response regulator protein KdpE[23]. ATP is bound by a classical Walker box where the Walker A motif is located at the residues 30 to 38 and a potential Walker B motif at residues 105 to 110. The second function of the amino-terminal region is early during its synthesis in the ribosome. For the targeting to the membrane the amino-terminal region of KdpD at the residues 22-48 is first contacting the SRP[20]. The RNC-SRP complex is then transported to the receptor protein FtsY to ensure that the nascent protein chain is close to the membrane surface. We show here that an artificially stalled KdpD22-74 RNC binds to SRP comparable with RNCs exposing the positive control protein FtsQ4-85, a well-studied substrate of SRP (Fig. 6)[25]. In addition, also a preassembled SRP-FtsY complex which was activated by GTP could be bound to KdpD22-74 RNC as for FtsQ4-85 RNC (Fig. 7). These results show that the residues 22-48 efficiently bind to SRP and allow the nascent chain of KdpD an early targeting to the membrane surface before translation has synthesized the full 894 long protein. The binding to SRP was analysed with microscale thermophoresis. Fluorescently labeled RNCs at 5 nM were mixed 1:1 with purified and assembled SRP (Ffh and 4.5 RNA) in different dilutions (0.5 nM to 1 µM) and applied to capillaries. For the binding to SRP-FtsY the assembled SRP was mixed with purified FtsY and 200 µM GppNHp prior to the addition of RNCs. Previous affinity determinations were performed with fluorescence equilibrium titrations using fluorescence-labeled SRP with leader peptidase RNCs[26] or by fluorescence anisotropy of fluorescein-labeled SRP with EspP RNCs[27] resulted in comparable binding events. Since the SRP-bound cargo has to pass a number of checkpoints, the binding to a preassembled SRP-FtsY complex with high affinity is an important step in the pathway[27].

The early targeting of KdpD at the membrane surface ensures that the 4 transmembrane segments can readily insert when they appear at the ribosome tunnel exit[28]. Previous studies have shown that the insertion of KdpD into the inner membrane can occur even in the absence of SecYEG or YidC[21]. Most likely, the cotranslational membrane insertion directly into the bilayer is possible because the periplasmic loops are very short with 4 and 10 residues, respectively (Fig. 1).

C-tailed membrane proteins share with KdpD the fact that a large hydrophilic domain is released from the ribosome before a membrane anchoring segments is exposed at

the tunnel exit. We have recently shown that the membrane targeting and insertion of the C-tail protein SciP involves SRP and the SRP receptor[29]. Similar as in KdpD the TMS of SciP is far from the N-terminus (residues 184–206). We found that two short hydrophobic regions at residues 12-20 and 62-71 of SciP have the potential to interact with SRP. When these peptides were fused with sfGFP the fusion protein was targeted the membrane surface. We conclude that the amino-terminal sequences found in SciP and KdpD are functional SRP signal sequences and are located close to the N-terminus in membrane proteins that contain large cytoplasmic domains. In both cases, the SRP signal sequence is separate from the first membrane spanning segment and therefore allows an early binding of SRP before the membrane segments are translated by the ribosome.

The function of the SRP-signal sequence can be tested in a fusion protein with the green fluorescent protein GFP[20]. Here, we show that the assay can be improved when the super folding GFP (sfGFP) is used for the fusion. The fluorescence in cells can be observed already 20 min after induction. With this assay we tested a collection of mutants that affected the function of the 22-48 sequence as a SRP signal and/or Walker A element. The key lysine of the Walker A element cannot be replaced by another amino acid without affecting the function of the Walker element[30]. When we substituted the key lysine with a glutamine in the WalkerQ mutant, or when the conserved glycine residues at position 30, 34 and 36 in the Walker3A mutant were replaced by alanine residues, membrane localization was not affected (Fig. 3d, e). Likewise, the binding events of the Walker3A mutant to SRP and to SRP-FtsY corresponded to the binding we had obtained with the wild-type signal sequence of 22-48 (Figs 6, 7). Therefore, these modified sequences fully function as a SRP targeting signal.

The deletion of the Walker A element (residues 30-38) resulted in a patchy appearance of the fusion protein in the cells (Fig. 3c). A similar phenotype was observed for the wild-type sequence when Ffh was depleted (Fig. 2c). We assume that the patches form because of aggregation in the cytoplasm. Interestingly, the sfGFP itself did not form such aggregates, even when Ffh was depleted (Fig. S1). It is surprising that small 19 to 28 residues long N-terminal extensions of sfGFP have such a big effect on the solubility. The binding affinity of the WalkerS mutant to SRP was the lowest we had measured and there was no binding to SRP-FtsY. It is possible that the lower hydrophobicity of the cysteine is critical in allowing that the sequence is still able to

bind to SRP. The malfunctioning of this mutant in membrane targeting was supported by the fluorescence microscopy (Fig. 3f). Interestingly, mutation of the cysteine residue into the more hydrophobic valine restored membrane targeting indicating that not the cysteine residue but the hydrophobicity is critical for SRP binding (Fig. 4d).

Finally, the positively charged residues at 22, 24 and 26 were investigated for membrane targeting and SRP binding. The sfGFP fusion protein showed that the fluorescence was evenly distributed in the cytoplasm suggesting that the binding to SRP is affected. Indeed, the binding to SRP and to SRP-FtsY was lower compared to the wild-type signal sequence of 22-48 respectively (Figs 6, 7). This shows that the positively charged residues play an important role for the interaction with SRP.

Taken together, this study shows that a SRP signal sequence is not restricted to a transmembrane segment but can be localized in a cytoplasmic region. The results obtained with the mutants of the signal sequence underline the importance of the positively charged N-terminal part and the hydrophobic C-terminal part of the signal sequence to allow the interaction with the M-pocket of SRP and membrane targeting.

Methods

Bacterial strains and culture conditions.

E. coli MC1061 (hsdR mcrB araD139 (*ΔaraABC-leu*)*7697 lacX74 galU galK rspL thi*) was decribed[31]. For Ffh depletion, *E. coli* MCΔFfh[32] was grown overnight in LB medium containing 0.2% (w/v) arabinose and 0.4% (w/v) glucose, washed in medium lacking arabinose and back-diluted 1:100 in LB medium containing 0.4% (w/v) glucose. Media preparation and bacterial manipulations were performed according to standard methods[33]. Where appropriate, ampicillin (100 µg/mL, final concentration) was added to the medium. *E. coli* BL21 (DE3)[34] and *E. coli* KC6[35] were grown overnight in LB with ampicillin (100 µg/mL) and back diluted 1:100 in fresh LB with ampicillin (100 µg/mL).

Construction of SRP-signal mutants of KdpD-N22-48-sfGFP.

All oligonucleotides used in this study are listed in the Supplementary Table S1 and the used plasmids in the Supplementary Table S2. The amino acid residues 22 to 48 of KdpD were amplified flanking the PCR product with a *Hind*III and *Bam*HI recognition site. The PCR product was then cloned in a pMS119EH derivate containing the sfGFP gene (own collection) using the restriction enzymes *Hind*III and *Bam*HI. The different SRP-signal mutants of KdpD-N22-48-sfGFP were constructed by a PCR-based

mutagenesis of N22-48-sfGFP. Substitution of the basic residues at position 22 and 24 of N22-48-sfGFP into glutamine resulted in the mutant pMS-KdpD22-48-2Q-sfGFP (named 2Q). Substitution of three of the closely spaced basic amino acid residues in N22-48-sfGFP were substituted into glutamines resulted in the mutant pMS-KdpD22-48-3Q-sfGFP (named 3Q). Substitution of lysine 37 within the Walker A motif with a glutamine resulted in mutant pMS-KdpD22-48-WalkerQ-sfGFP (named WalkerQ). To construct mutant pMS-KdpD22-48-ΔWalker-sfGFP (named ΔWalker), amino acids 30-38 were deleted with site-directed mutagenesis. To change the conserved motif of the Walker box into a no Walker-motif a site-directed mutagenesis was done to generate pMS-KdpD22-48-Walker3A-sfGFP (named Walker3A). Substitution of cysteine 32 within the Walker A motif with a serine or valine resulted in mutant pMS-KdpD22-48-WalkerS-sfGFP (named WalkerS) and mutant pMS-KdpD22-48-WalkerV-sfGFP (named WalkerV). The coding regions of all constructs were verified by DNA sequence analysis.

Construction of plasmids for ribosome-nascent-chain (RNC) interaction studies.

For cloning of the TnaC-stalling sequence, pBAT4-MscL[115] (kindly provided by R. Beckmann, Munich) was digested with *Xba*I and *Hind*III and cloned into pMS119EH[36]. To remove the MscL sequence a *Nco*I and *Mfe*I recognition site was introduced with site-directed mutagenesis resulting in pMS-MscL[115]-TnaC. The sequence encoding for amino acids 22-74 of KdpD was amplified flanking the PCR product with a *Mfe*I and *Nco*I recognition site. The digested PCR product was cloned into pMS-MscL[115]-TnaC using *Nco*I and *Mfe*I (pMS-KdpD22-74-TnaC). For cloning of pMS-KdpD22-74-W3A-TnaC and pMS-KdpD22-74-3Q-TnaC the sequence was amplified using the respective primers and template DNA and cloned into pMS-MscL[115]-TnaC using *Nco*I and *Mfe*I. Plasmid pMS-KdpD22-74-C32S-TnaC was generated by site-directed mutagenesis on plasmid pMS-KdpD22-74-TnaC. The plasmid pMS-His-HA-TnaC (named short-chain (sc)-ribosome) was cloned by digesting a pMS-TnaC plasmid (institute collection) with *Mfe*I and *Pst*I, blunted with T4 DNA polymerase and religated. This results in RNCs with exposing a short nascent chain of about 13 amino acids.

The plasmid encoding the first 50 amino acids of the cytoplasmic protein firefly luciferase was generated by amplification from plasmid pUC19-T7-Luc[50] (kindly provided by Shu-ou Shan; Caltech). The PCR product was flanked with an *Eco*RI and a *Nco*I recognition site producing compatible sticky ends with *Mfe*I and *Nco*I digested plasmid resulting in plasmid pMS-Luc2-50-TnaC. The plasmid encoding the first 85

amino acids of FtsQ (pEM36-3C) was kindly provided by R. Beckmann, Munich.

Fluorescence microscopy.

Strains were grown overnight at 37 °C, diluted in fresh LB medium and grown to an OD_{600} of 0.5. IPTG was then added to a final concentration of 1 mM. The cells were incubated for 20 min (MCΔFfh) or 30 min (MC1061) at 37 °C under continuous shaking. The cells were treated as described[29] and collected by centrifugation, washed twice with LB medium and resuspended in 2 mM EDTA, 50 mM Tris-HCl, pH 8.0. The cell suspension (3 µL) was applied to a polylysine-coated cover glass (Sigma-Aldrich) and examined immediately by fluorescence microscopy with the Zeiss AxioImager M1 fluorescence microscope. Emission was detected with filters specific for GFP. Analysis was done by using the AxioVision Software (Zeiss). The expression of the different KdpD22-48-sfGFP mutants was analysed on a 12% SDS-PAGE (after TCA precipitation with 10% TCA) and immunoblotting with an α-GFP and α-rabbit antibody.

Purification of Ffh and FtsY.

E. coli Ffh was purified essentially as described by Seitl[37]. Wild-type Ffh and the mutant L181C and M423C were expressed from pMS-Ffh-C-Strep, pMS-Ffh-L181C-C-Strep or pMS-Ffh-M423C-C-Strep in BL21 (DE3) cells. 2 L LB with 100 µg/mL ampicillin were inoculated 1:100 from an overnight culture and grown at 37°C until an OD_{600} of 0.5. The cells were induced with 1 mM IPTG for 3 h at 37°C, harvested, resuspended in buffer A_{Ffh} (20 mM Hepes pH 8, 350 mM NaCl, 10 mM $MgCl_2$, 10 mM KCl, 10% glycerol) and lysed using the OneShot at 1.23 kbar. Before cell disruption 0.2 mM PMSF was added. The lysate was centrifuged 2x for 30 min at 20 000 g and the supernatant was loaded onto 3 mL Strep-Tactin matrix by gravity flow. After washing the matrix with 50 mL buffer W_{Ffh} (20 mM Hepes pH 8, 500 mM NaCl, 10 mM $MgCl_2$, 100 mM KCl) the protein was eluted with 20 mL buffer E_{Ffh} (20 mM Hepes pH 8, 350 mM NaCl, 10 mM $MgCl_2$, 10 mM KCl, 10% glycerol, 2.5 mM desthiobiotin) in 2 mL fractions. The elution fractions were further purified using the Äkta-purifier System on a Superdex 75 16/60 column in buffer GF_{Ffh} (20 mM Hepes pH 8, 200 mM NaCl, 10 mM $MgCl_2$, 10 mM KCl, 10% glycerol).

E. coli FtsY was expressed from plasmid pTrc99-FtsY-His (kindly provided by HG Koch, Freiburg) in BL21 (DE3) cells. 2 L LB with 100 µg/mL ampicillin were inoculated with an overnight culture (1:100) and grown at 37°C until an OD_{600} of 0.5. The culture was induced with 1 mM IPTG for 4 h at 37°C, the cells were harvested and

resuspended in buffer A_{FtsY} (50 mM Hepes pH 7.6, 1 M NH_4Ac, 10 mM $Mg(OAc)_2$, 10% glycerol, 1 mM DTT). After the addition of 0.2 mM PMSF the cells were lysed using the OneShot at 1.23 kbar, the lysate was centrifuged for 30 min at 4300 g and the supernatant was centrifuged in a Beckman Ti60 rotor for 1 h at 38 000 rpm. The supernatant was incubated with 2 mL Ni-NTA in Buffer A_{FtsY} + 30 mM imidazole for 1 h at 4 °C on a rotary wheel. The matrix was washed with 20 mL buffer W_{FtsY} (50 mM Hepes pH 7.6, 1 M NH_4Ac, 10 mM $Mg(OAc)_2$, 1 mM DTT, 30 mM imidazole) and the protein was eluted with 20 mL buffer E_{FtsY} (50 mM Hepes pH 7.6, 1 M NH_4Ac, 10 mM $Mg(OAc)_2$, 400 mM imidazole, 10% glycerol) in 2 mL fractions. The elution fractions were further purified using the Äkta-purifier System on a Superdex 200 16/60 column in buffer GF_{FtsY} (100 mM Hepes pH 7.6, 200 mM KOAc, 20 mM $Mg(OAc)_2$, 2 mM DTT, 10% glycerol).

***In vitro* cross-linking with copper phenanthroline.**

The KdpD peptide (RGKLKVFFGACAGVGKTWAMLAEAQRL) was synthesized by the Custom peptide synthesis services from GENOSPHERE Biotechnologies (France) with N-terminal acetylation and C-terminal amidation and a purity of >95%. For the *in vitro* cross-linking experiment the plasmid pMS-Ffh-L181C-C-Strep was generated by site-directed mutagenesis on plasmid pMS-Ffh-C-Strep.

For *in vitro* cross-linking 2 µM of purified Ffh (406C), Ffh L181C/406S or Ffh 406S/423C were mixed with 3 µM of 4.5S RNA to get a functional SRP in $buffer_{SRP}$ (20 mM Hepes pH 7.2, 50 mM KOAc, 5 mM $Mg(OAc)_2$). Each of the reconstituted SRP was mixed with 20 µM of synthesized KdpD peptide in $buffer_{Crosslink}$ (50 mM Tris pH 7.4, 150 mM NaCl, 10 mM $MgCl_2$) and 1 mM copper phenanthroline was added. The mixture was incubated for 1 h on ice, TCA precipitated, resuspended in buffer with or without DTT (1 mM) and loaded on a 10% SDS-PAGE.

Construction and purification of ribosome-nascent-chains (RNCs).

The plasmids pEM36-3C (encoding for a N-terminal His-Tag, a C3-protease cleavage site, the first 85 amino acids of FtsQ, a C-terminal HA-Tag and the TnaC-stalling sequence), pMS-Luc2-50-TnaC, pMS-KdpD22-74-TnaC, pMS-KdpD22-74-W3A-TnaC and pMS-KdpD22-74-3Q-TnaC were transformed in *E. coli* KC6[35] cells. 2 L LB containing 100 µg/mL ampicillin were inoculated 1:100 from an overnight culture and grown until an OD_{600} = 0.5. For induction, 1 mM IPTG was added and the cells were grown for another hour at 37°C. After harvesting, cells were resuspended in buffer A_{RNC}

(20 mM Hepes pH 7.2, 250 mM KOAc, 25 mM $Mg(OAc)_2$, 2 mM L-tryptophane, 0.1% DDM), 0.2 mM PMSF was added and the cells were lysed using the OneShot at 1.23 kbar. The lysate was centrifuged at 31000 g for 20 min, the supernatant was loaded onto 750 mM sucrose in buffer A_{RNC} and the ribosomes were pelleted at 25 000 rpm for 20 h in a Beckman Ti45 rotor. The ribosomal pellet was resuspended in buffer A_{RNC} and loaded onto 3 mL Ni-NTA (blocked with 10 µg/mL *E. coli* tRNA for 1 h at 4 °C) and batched for 1 h at 4°C on a rotary wheel. The matrix was washed with 30 mL Buffer B_{RNC} (50 mM Hepes pH 7.2, 500 mM KOAc, 25 mM $MgCl_2$, 2 mM L-tryptophane, 0.1% DDM) and the His-tagged ribosomes were eluted with 3 mL buffer B_{RNC} + 150 mM imidazole and 3 mL buffer B_{RNC} with 300 mM imidazole in 1 mL fractions. The elution fractions were loaded onto a linear sucrose gradient (10–40% sucrose in buffer B_{RNC}) and centrifuged at 30 000 rpm for 3.5 h in a Beckman SW40 rotor. The gradient was collected in 1 mL fractions, the ribosome containing fractions were identified measuring the absorbance at 260 nm, pooled and centrifuged at 40 000 rpm for 4 h in a Beckman Ti60 rotor. The ribosomal pellet was resuspended in buffer C_{RNC} (20 mM Hepes pH 7.2, 50 mM KOAc, 5 mM $Mg(OAc)_2$, 2 mM L-tryptophane) and stored at −80 °C.

***In vitro* 4.5S RNA synthesis and SRP reconstitution.**

To get a functional SRP the purified Ffh protein was reconstituted with the *in vitro* synthesized 4.5S RNA as described by Seitl[37]. Therefore, plasmid pUC18-4.5S RNA (kindly provided by Irmgard Sinning, Heidelberg) was used where the 4.5S RNA sequence was placed downstream of the T7 promoter. First, the plasmid was linearized with *Bam*HI and gel purified. For *in vitro* transcription with the HiScribe™ T7 High Yield RNA Synthesis Kit (NEB) 1 µg linearized plasmid DNA was used. After incubation for 16 h at 37°C the 4.5S RNA was purified using the RNA Clean & Concentrator™ -25 Kit (Zymo Research), analyzed on a 2% agarose gel in 1x Tris-borate-EDTA buffer and stored at −80°C. Prior to the reconstitution, the RNA was heated to 75°C for 2 min and chilled on ice for 1 min. Ffh and 1.5-fold molar excess of 4.5S RNA were mixed in MST-buffer (20 mM Hepes pH 7.2, 50 mM KOAc, 5 mM $Mg(OAc)_2$, 2 mM L-tryptophane, 0.05% Tween-20) and incubated for 10 min at 20 °C.

Labeling of RNCs.

The different RNCs were labeled with the cysteine-reactive fluorescent dye NT-647 and the RED-maleimide labeling kit (NanoTemper Technologies). The RNCs were adjusted to a concentration of 2 µM in 100 µL buffer C_{RNC} and mixed with 3-fold molar

excess of the dye in 100 µL labeling buffer provided in the kit. After incubation for 30 min at RT in the dark the labeling reaction was purified to remove the free dye using the column provided in the kit. The concentration of the labeled RNCs after purification was calculated measuring the absorbance at 260 nm and they were stored at −80°C.

Microscale thermophoresis.

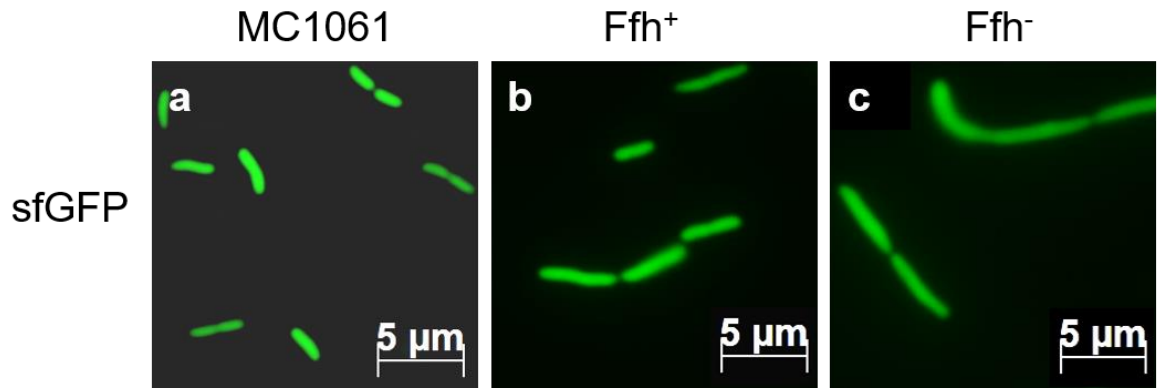
For the interaction studies between the RNCs and SRP, the labeled RNCs were adjusted to 10 nM in MST buffer. With the reconstituted SRP a series of 1:1 dilutions was prepared in MST buffer with a concentration ranging from 1 µM to 0.49 nM. The ligand dilutions were mixed with one volume of labeled RNCs resulting in a RNC concentration of 5 nM. After incubation for 5 min on ice, the dilutions were filled in Monolith NT Premium Treated capillaries (NanoTemper Technologies) and measured using the Monolith NT.115 instrument. During measurement the temperature was kept constant at 22°C. Thermophoresis was measured with 5 s laser off, 20-30 s laser on and 5 s laser off, a LED power of 40-60% and the MST Power “Low”. The data of three independently pipetted measurements were merged and analyzed using the software MO.Affinity Analysis v2.3 (NanoTemper Technologies) using the manual evaluation (Cold region start/end: −1 s/0 s; Hot region start/end: 5.01 s/10.08 s).

For the interaction studies between the RNCs and SRP-FtsY a preincubated closed SRP-FtsY complex was used. Therefore, the reconstituted SRP was mixed with 4-fold molar excess of FtsY in MST buffer containing 200 µM of the non-hydrolysable GTP-analogue GppNHp. After incubation for 10 min at 25°C the stable SRP-FtsY complex was incubated on ice. For measurements a series of 1:1 dilutions was prepared in MST buffer with 200 µM GppNHp with a complex concentration ranging from 500 nM/250 nM to 0.25/0.12 nM. The ligand dilutions were mixed with one volume of labeled RNCs resulting in a RNC concentration of 5 nM. After incubation for 5 min on ice, the dilutions were filled in Monolith NT capillaries and measured using the Monolith NT.115 instrument. During measurement the temperature was kept constant at 22°C. Thermophoresis was measured with 5 s laser off, 20-30 s laser on and 5 s laser off, a LED power of 60% and the MST Power „low”. The data of three independently pipetted measurements were merged and analyzed using the software MO.Affinity Analysis v2.3 (NanoTemper Technologies) using the manual evaluation (Cold region start/end: −1 s/0 s; Hot region start/end: 5.01 s/10.02 s).

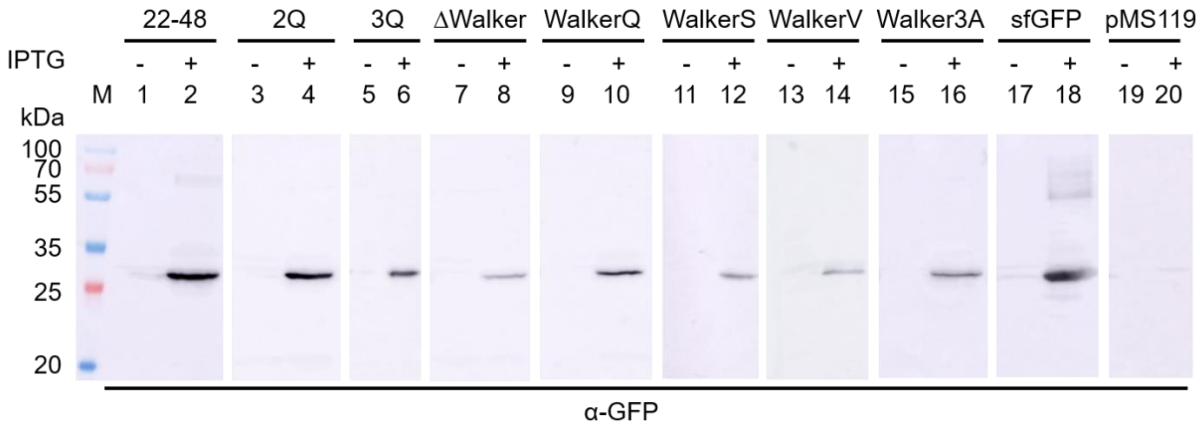
Data Availability

All data generated or analysed during this study are included in this published article (and its Supplementary Information Files) or are available from the corresponding author on reasonable request.

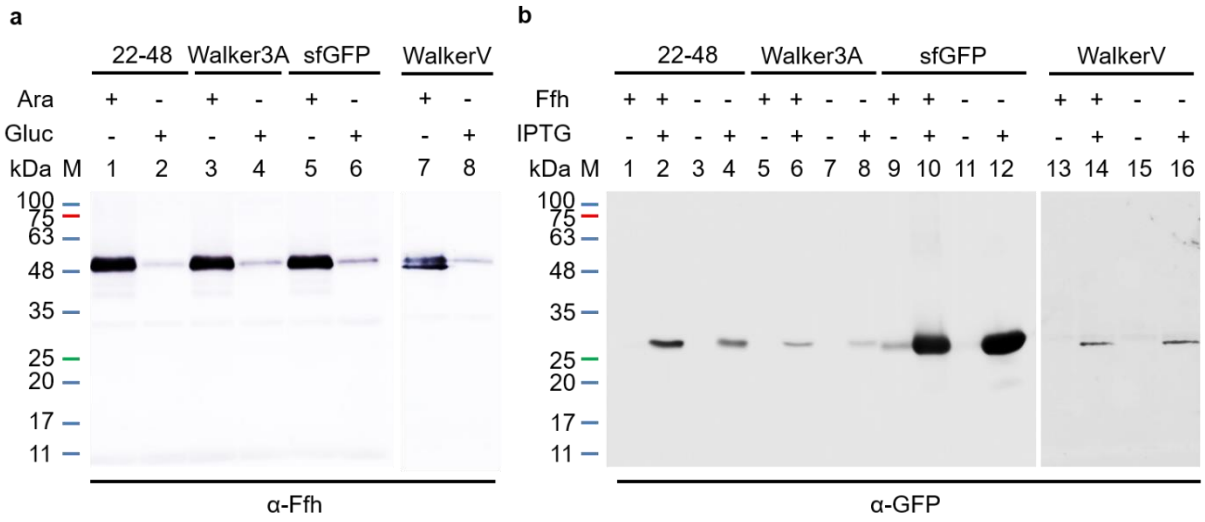
Supplementary Information



Supplementary Figure S1: Cellular localization of sfGFP. Expression of sfGFP in *E. coli* MC1061 (a) and in the depletion strain MCΔFfh (b, c) was induced for 30 min (a) and 20 min (b, c) at 37°C and followed by fluorescence microscopy. Non-fused sfGFP was always distributed throughout the cytoplasm, regardless whether Ffh was present or not.

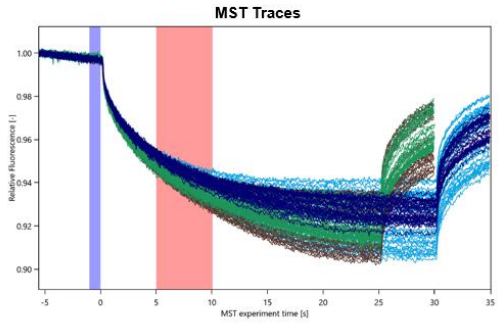


Supplementary Figure S2: Expression of the different KdpD-sfGFP fusion proteins in *E. coli* MC1061. Expression was induced at an OD_{600} of 0.5 for 30 min, the cells were TCA precipitated and loaded on a 12% SDS-PAGE. After Western transfer, immune detection was carried out with α-GFP and α-rabbit antibodies. Due to induction, a band between 25 and 35 kDa corresponding to the different KdpD-sfGFP fusion proteins and for the non-fused sfGFP was detected. The empty plasmid was analysed as a control.

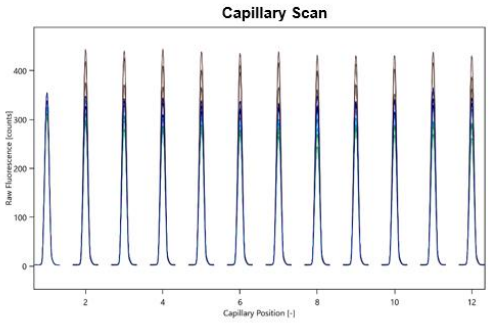


Supplementary Figure S3: Expression of the different KdpD-sfGFP fusion proteins in *E. coli* MCΔFfh. (a) The amount of Ffh before the induction of N22-48-sfGFP, Walker3A-sfGFP, non-fused sfGFP and WalkerV-sfGFP under arabinose (0.2%) and glucose (0.4%) growth conditions. The amount of Ffh was verified by Western blot with α-Ffh and α-rabbit antibodies. Growth in glucose leads to depleted Ffh levels. (b) Expression of N22-48-sfGFP, Walker3A-sfGFP, non-fused sfGFP and WalkerV-sfGFP in *E. coli* MCΔFfh. Expression was induced at an OD_{600} of 0.5 for 30 min, the cells were TCA precipitated and loaded on a 12% SDS-PAGE. After Western transfer, immuno detection was carried out with α-GFP and α-rabbit antibodies. Due to induction, a band between 25 and 35 kDa corresponding to the different KdpD-sfGFP fusion proteins and for the non-fused sfGFP was detected.

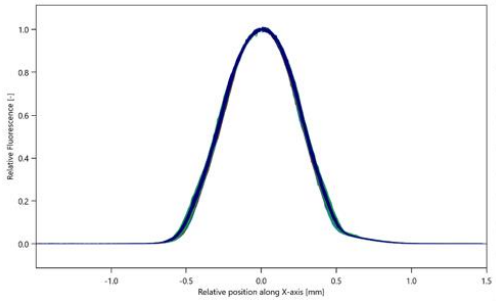
a



b



c

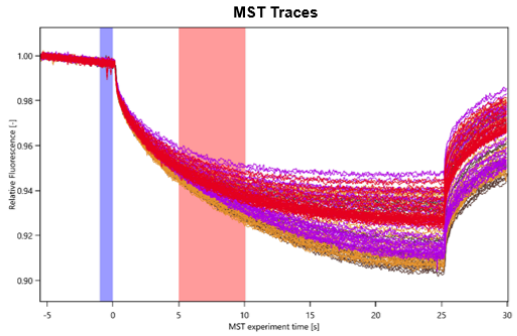


d

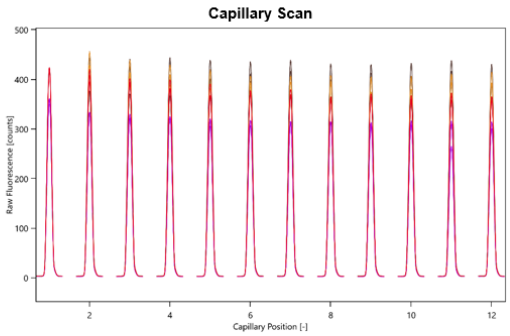
Name	FtsQ4-85 vs SRP	KdpD22-74 vs SRP	Luc2-50 vs SRP	sc-Ribosome
Graph Color:	●	●	●	●
Target Concentration:	5 nM	5 nM	5 nM	5 nM
Ligand Concentration:	1E+03 to 0.488 nM	500 to 0.488 nM	1E+03 to 0.488 nM	1E+03 to 0.488 nM
n:	3	3	3	3
Excitation Power:	50%	60%	40%	50%
MST Power:	20%	20%	20%	20%
Temperature:	22.0°C	22.0°C	22.0°C	22.0°C
Kd:	30.348 nM	51.458 nM	22 [Fixed]	22 [Fixed]
Kd Confidence:	± 5.8684 nM	± 10.947 nM		
Response Amplitude:	9.710854	11.423748	0	0
Target Conc.:	5 nM [Fixed]	5 nM [Fixed]	5 nM [Fixed]	5 nM [Fixed]
Unbound:	941.33	938.2	942.79 [Fixed]	948.38 [Fixed]
Bound:	951.04	949.63	942.79 [Fixed]	948.38 [Fixed]
Std. Error of Regression:	0.49584508	0.54083863	0.94909443	0.75369782
Reduced χ^2:	3.970388	8.3759013	38.80484	2.7063745
Signal to Noise:	21.651426	23.615433	0	0

Supplementary Figure S4: Raw data of the MST measurements with FtsQ-4-85, KdpD-22-74 and Luc2-50 as a nascent chain and short-chain (sc)-ribosomes with SRP. Microscale thermophoresis measurements of unlabelled SRP (1 µM to 0.49 nM) with labelled RNCs (5 nM) of FtsQ-4-85 (blue), KdpD-22-74 (brown), Luc-2-50 (green) and sc-ribosomes (dark blue). After a 5 min incubation on ice the dilutions were filled into Premium capillaries (NanoTemper Technologies) for the MST measurements. In a, the MST traces (relative fluorescence [-] plotted against the MST experimental time [s]), in b, the capillary scan (raw fluorescence [counts] plotted against the capillary position [-]) and in c, the capillary shape (relative fluorescence [-] plotted against the relative position along x-axis [mm] is shown. In d, the dataset overview is listed.

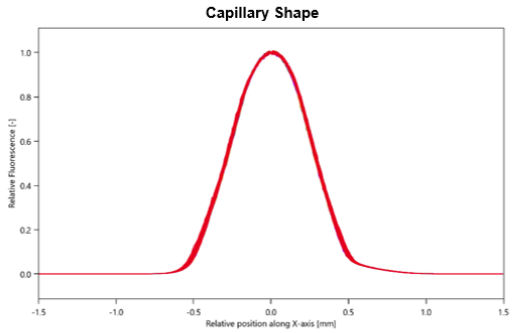
a



b



c

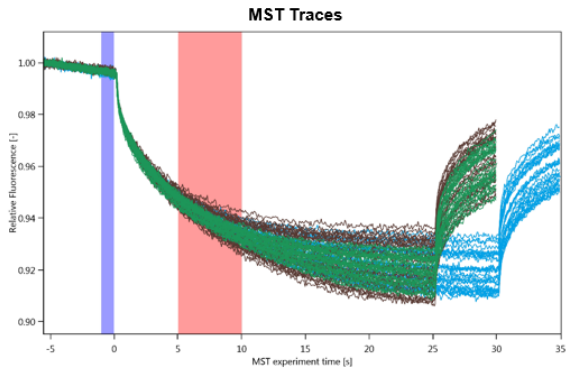


d

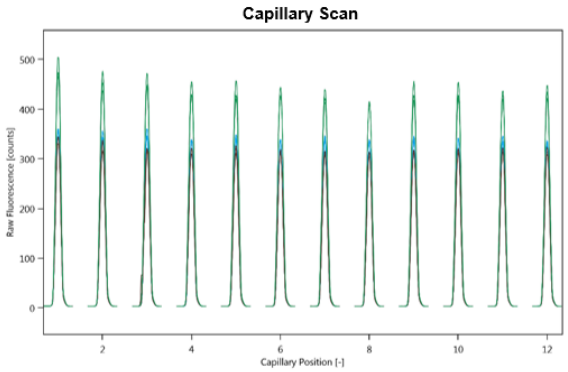
Name	KdpD22-74 vs SRP	KdpD22-74-Walker3A vs SRP	KdpD22-74-3Q vs SRP	KdpD22-74-WalkerS vs SRP
Graph Color:	●	●	●	●
Target Concentration n:	5 nM	5 nM	5 nM	5 nM
Ligand Concentration:	500 to 0.488 nM	500 to 0.488 nM	1E+03 to 0.488 nM	1E+03 to 0.488 nM
n:	3	3	3	3
Excitation Power:	60%	60%	50%	60%
MST Power:	20%	20%	20%	20%
Temperature:	22.0°C	22.0°C	22.0°C	22.0°C
Kd:	51.458 nM	71.57 nM	184.01 nM	1574.2 nM
Kd Confidence:	± 10.947 nM	± 7.4236 nM	± 41.6 nM	± 1171.9 nM
Response Amplitude:	11.423748	15.851063	15.022232	22.252655
Target Conc.:	5 nM [Fixed]	5 nM [Fixed]	5 nM [Fixed]	5 nM [Fixed]
Unbound:	938.2	937.78	941.98	945.42
Bound:	949.63	953.63	957.01	967.67
Std. Error of Regression:	0.54083863	0.34194973	0.68846451	0.69027065
Reduced χ^2:	8.3759013	1.9073639	73.239789	3.1654293
Signal to Noise:	23.615433	51.826412	24.122815	35.639986

Supplementary Figure S5: Raw data of the MST measurements with KdpD-22-74, KdpD-22-74-Walker3A, KdpD-22-74-3Q and KdpD-22-74-WalkerS as a nascent chain with SRP. Microscale thermophoresis measurements of unlabelled SRP (1 µM to 0.49 nM) with labelled RNCs (5 nM) of KdpD-22-74 (brown), KdpD-22-74-Walker3A (orange), KdpD-22-74-3Q (purple) and KdpD-22-74-WalkerS (red). After a 5 min incubation on ice the dilutions were filled into Premium capillaries (NanoTemper Technologies) for the MST measurements. In a, the MST traces (relative fluorescence [-] plotted against the MST experiment time [s]), in b, the capillary scan (raw fluorescence [counts] plotted against the capillary position [-]) and in c, the capillary shape (relative fluorescence [-] plotted against the relative position along x-axis [mm] is shown. In d, the dataset overview is listed.

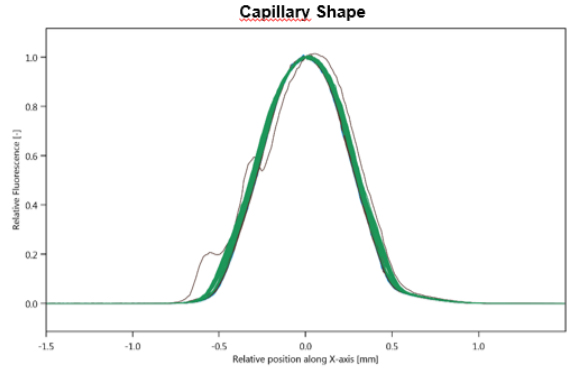
a



b



c

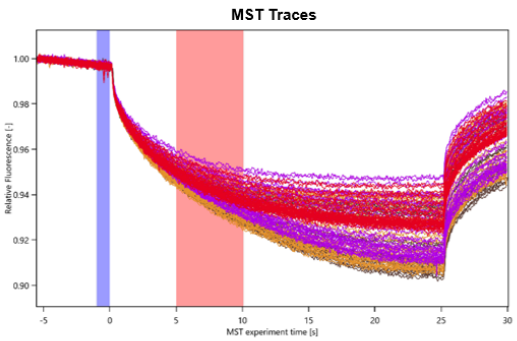


d

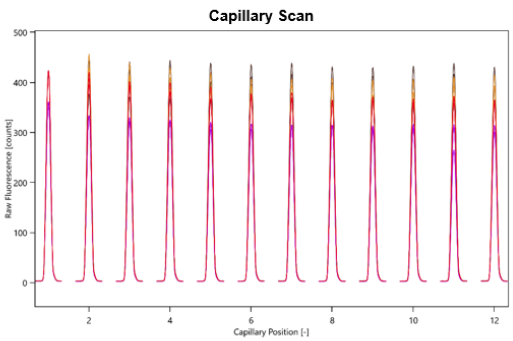
Name	FtsQ4-85 vs SRP-FtsY	KdpD22-74 vs SRP-FtsY	Luc2-50 vs SRP-FtsY
Graph Color:	●	●	●
Target Concentration:	5 nM	5 nM	5 nM
Ligand Concentration:	250 to 0.122 nM	500 to 0.244 nM	500 to 0.244 nM
n:	3	3	3
Excitation Power:	60%	60%	60%
MST Power:	20%	20%	20%
Temperature:	22.0°C	22.0°C	22.0°C
Kd:	22.932 nM	22.826 nM	11 [Fixed]
Kd Confidence:	± 10.75 nM	± 7.4535 nM	
Response Amplitude:	4.8508007	8.3330603	0
Target Conc.:	5 nM [Fixed]	5 nM [Fixed]	5 nM [Fixed]
Unbound:	941.56	939.22	942.85 [Fixed]
Bound:	946.41	947.56	942.85 [Fixed]
Std. Error of Regression:	0.50948147	0.68351503	0.38100858
Reduced χ^2:	6.8864189	2.6936454	0.80878355
Signal to Noise:	10.525922	13.478189	0

Supplementary Figure S6: Raw data of the MST measurements with FtsQ-4-85, KdpD-22-74 and Luc2-50 as a nascent chain with a SRP-FtsY complex. Microscale thermophoresis measurements of unlabelled SRP-FtsY (500 / 250 nM to 0.244 / 0.122 nM) with labelled RNCs (5 nM) of FtsQ-4-85 (blue), KdpD-22-74 (brown) and Luc-2-50 (green). After a 5 min incubation on ice the dilutions were filled into Premium capillaries (NanoTemper Technologies) for the MST measurements. In a, the MST traces (relative fluorescence [-] plotted against the MST experiment time [s]), in b, the capillary scan (raw fluorescence [counts] plotted against the capillary position [-]) and in c, the capillary shape (relative fluorescence [-] plotted against the relative position along x-axis [mm] is shown. In d, the dataset overview is listed.

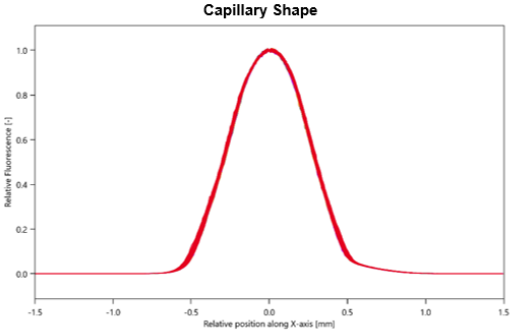
a



b



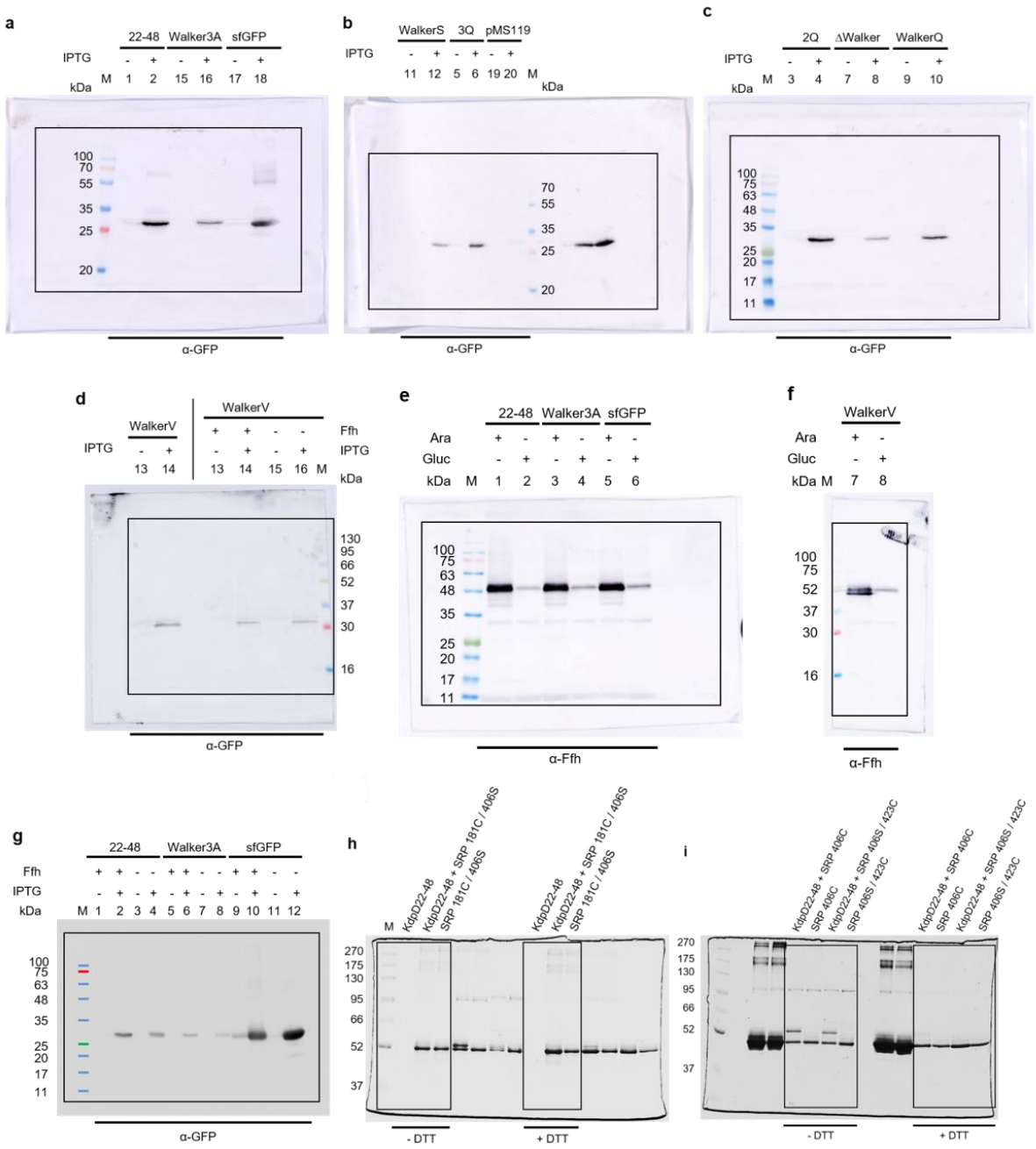
c



d

Name	KdpD22-74 vs SRP	KdpD22-74-Walker3A vs SRP	KdpD22-74-3Q vs SRP	KdpD22-74-WalkerS vs SRP
Graph Color:	●	●	●	●
Target Concentration n:	5 nM	5 nM	5 nM	5 nM
Ligand Concentration:	500 to 0.488 nM	500 to 0.488 nM	1E+03 to 0.488 nM	1E+03 to 0.488 nM
n:	3	3	3	3
Excitation Power:	60%	60%	50%	60%
MST Power:	20%	20%	20%	20%
Temperature:	22.0°C	22.0°C	22.0°C	22.0°C
Kd:	51.458 nM	71.57 nM	184.01 nM	1574.2 nM
Kd Confidence:	± 10.947 nM	± 7.4236 nM	± 41.6 nM	± 1171.9 nM
Response Amplitude:	11.423748	15.851063	15.022232	22.252655
Target Conc.:	5 nM [Fixed]	5 nM [Fixed]	5 nM [Fixed]	5 nM [Fixed]
Unbound:	938.2	937.78	941.98	945.42
Bound:	949.63	953.63	957.01	967.67
Std. Error of Regression:	0.54083863	0.34194973	0.68846451	0.69027065
Reduced χ^2:	8.3759013	1.9073639	73.239789	3.1654293
Signal to Noise:	23.615433	51.826412	24.122815	35.639986

Supplementary Figure S7: Raw data of the MST measurements with KdpD-22-74, KdpD-22-74-Walker3A, KdpD-22-74-3Q and KdpD-22-74-WalkerS as a nascent chain with a SRP-FtsY complex. Microscale thermophoresis measurements of unlabelled SRP (500 nM to 0.244 nM) with labelled RNCs (5 nM) of KdpD-22-74 (brown), KdpD-22-74-Walker3A (orange), KdpD-22-74-3Q (purple) and KdpD-22-74-WalkerS (red). After a 5 min incubation on ice the dilutions were filled into Premium capillaries (NanoTemper Technologies) for the MST measurements. In a, the MST traces (relative fluorescence [-] plotted against the MST experiment time [s]), in b, the capillary scan (raw fluorescence [counts] plotted against the capillary position [-]) and in c, the capillary shape (relative fluorescence [-] plotted against the relative position along x-axis [mm] is shown. In d, the dataset overview is listed.



Supplementary Figure S8: Uncropped Western Blot and SDS-PAGE of supplementary figure S2 and S3 and of figure 5. The labelling of the lanes in a, b, c and d corresponds to the labelling in figure S2, the labelling of the lanes in d, e, f and g corresponds to the labelling in figure S3 and the labelling of the lanes in h and i corresponds to figure 5.

Supplementary Table S1: Oligonucleotides used in this study.

Primer name	Sequence (5' to 3')	Usage	Plasmid
KdpD22-48_HindIII	GGCAAGCTTATGCGCGGGAAGCTGAAAG	Amplification	pMS-KdpD22-48-sfGFP
KdpD22-48_BamHI	CGGGGATCCCAGTCGCTGGGCTTCTGC		
KdpD-R22QK24Q-1	GATATACATATGCAGGGGCAGCTGAAAGTTTTC	Site-directed mutagenesis	pMS-KdpD22-48-2Q-sfGFP
KdpD-R22QK24Q-2	GAAAACTTTCAGCTGCCCCTGCATATGTATATC		
KdpD-R22QK24QK26Q-1	GCAGGGGCAGCTGCAAGTTTTCTTCGG	Site-directed mutagenesis	pMS-KdpD22-48-3Q-sfGFP
KdpD-R22QK24QK26Q-2	CCGAAGAAAACTTGCAGCTGCCCCTGC		
KdpD-K37Q-1	GCAGGCGTCGGGCAGACCTGGGCGATG	Site-directed mutagenesis	pMS-KdpD22-48-WalkerQ-sfGFP
KdpD-K37Q-2	CATCGCCCAGGTCTGCCCGACGCCTGC		
KdpD∆30-38-1	GCTGAAAGTTTTCTTCTGGGCGATGCTGGCAGAAG	Site-directed mutagenesis	pMS-KdpD22-48-∆Walker-sfGFP
KdpD∆30-38-2	CTTCTGCCAGCATCGCCCAGAAGAAAACTTTCAGC		
G30A-G34A-G36A-1	GTTTTCTTCGCTGCCTGTGCAGCCGTCGCGAAGACCTGGGCG	Site-directed mutagenesis	pMS-KdpD22-48-Walker3A-sfGFP
G30A-G34A-G36A-2	CGCCCAGGTCTTCGCGACGGCTGCACAGGCAGCGAAGAAAAC		
KdpD-C32S-1	CTTCGGTGCCAGCGCAGGCGTCGGG	Site-directed mutagenesis	pMS-KdpD22-48-WalkerS-sfGFP pMS-KdpD-22-74-C32S-TnaC
KdpD-C32S-2	CCCGACGCCTGCGCTGGCACCGAAG		
KdpD-C32V-1	CTTCGGTGCCGTTGCAGGCGTCG	Site-directed mutagenesis	pMS-KdpD22-48-WalkerV-sfGFP
KdpD-C32V-2	CGACGCCTGCAACGGCACCGAAG		
Ffh-L181C-1	CTGAAAGAAGCCAAATGCAAATTCTACGACGTG	Site-directed mutagenesis	pMS-Ffh-L181C-C-Strep
Ffh-L181C-2	CACGTCGTCGAATTTGCATTTGGCTTCTTTCAG		

MscL115-TnaC-MfeI-1	CTGTATTTTCAGGGACAATTGAGCATTATTAAAGAATTTCG	Site-directed mutagenesis	pMS-MscL^{115}
MscL-TnaC-MfeI-2	CGAAATTCTTTAATAATGCTCAATTGTCCCTGAAAATACAG		
MscL-TnaC-NdeI-1	CCGCACCTGCACCATGGTGTGTGACCTCAAAATGG	Site directed mutagenesis	pMS-MscL^{115}
MscL-TnaC-NdeI-2	CCATTTTGAGGTCACACACCATGGTGCAGGTGCGG		
KdpD22-74-MfeI	GGCCAATTGCGGGGGAAGCTGAAAGTTTTC	Amplification	pMS-KdpD-22-74-TnaC pMS-KdpD-22-74-W3A-TnaC
KdpD22-74-NdeI	CGGCCATGGCCCCTCGAGCATGGC		
KdpD22-74-3Q-MfeI	GGCCAATTGCAGGGGCAGCTGCAAGTTTTC	Amplification	pMS-KdpD-22-74-3Q-TnaC
Luc2-50-EcoRI	GCGAATTCGAAGACGCCAAAAACATAAAG	Amplification	pMS-Luc2-50-TnaC
Luc2-50-NdeI	CGGCCATGGGTTCACCTCGATATGTGCATC		

Supplementary Table S2: Plasmids used in this study.

Plasmids	Name	Characteristics	Reference or source
pMS119EH		Ap^R, P_{tac}, expression vector	Balzer *et al.*, 1992
pMS119EH-sfGFP	sfGFP	Ap^R, pMS119EH, *sfgfp*	Pross *et al.,* 2016
pMS-KdpD22-48-sfGFP	N22-48-sfGFP	Ap^R, pMS119EH, *kdpD*$_{\Delta 1\text{-}21:\Delta 49\text{-}894}$-*sfgfp*	This study
pMS-KdpD22-48-2Q-sfGFP	2Q	Ap^R, pMS119EH, *kdpD*$_{\Delta 1\text{-}21:R22Q,K24Q;\Delta 49\text{-}894}$-*sfgfp*	This study
pMS-KdpD22-48-3Q-sfGFP	3Q	Ap^R, pMS119EH, *kdpD*$_{\Delta 1\text{-}21:R22Q,K24Q,K26Q;\Delta 49\text{-}894}$-*sfgfp*	This study
pMS-KdpD22-48-ΔWalker-sfGFP	ΔWalker	Ap^R, pMS119EH, *kdpD*$_{\Delta 1\text{-}21;\Delta 30\text{-}38;\Delta 49\text{-}894}$-*sfgfp*	This study
pMS-KdpD22-48-WalkerQ-sfGFP	WalkerQ	Ap^R, pMS119EH, *kdpD*$_{\Delta 1\text{-}21:K37Q;\Delta 49\text{-}894}$-*sfgfp*	This study
pMS-KdpD22-48-WalkerS-sfGFP	WalkerS	Ap^R, pMS119EH, *kdpD*$_{\Delta 1\text{-}21:C32S;\Delta 49\text{-}894}$-*sfgfp*	This study
pMS-KdpD22-48-Walker3A-sfGFP	Walker3A	Ap^R, pMS119EH, *kdpD* $_{\Delta 1\text{-}21;}$ $_{G30A\text{-}G34A\text{-}G36A;\ \Delta 49\text{-}894}$ -*sfgfp*	This study
pMS-KdpD22-48-WalkerV-sfGFP	WalkerV	Ap^R, pMS119EH, *kdpD*$_{\Delta 1\text{-}21:C32V;\Delta 49\text{-}894}$-*sfgfp*	This study
pMS-Ffh-C-Strep	Ffh	Ap^R, pMS119EH, C-terminal Strep tag, *Ffh*	Seitl, 2016
pMS-Ffh-L181C-C-Strep	Ffh 181C / 406S	Ap^R, pMS119EH, C-terminal Strep tag, *Ffh*$_{L181C,\ C406S}$	This study
pMS-Ffh-M423C-C-Strep	Ffh 406S / 423C	Ap^R, pMS119EH, C-terminal Strep tag, *Ffh*$_{C406S,\ M423C}$	Seitl, 2016
pTrc99a-FtsY	FtsY	Ap^R, pTrc99a, *FtsY*, C-terminal His_6 tag	Kuhn *et al.,* 2015
pBAT4-MscL^{115}		ApR, P_{T7}, pBAT4, N-terminal His_6 and HA tag, TEV site, *mscL*$_{\Delta 116\text{-}136}$, $TnaC_{7\text{-}24}$	Seidelt *et al.,* 2009
pMS-MscL^{115}		Ap^R, pMS119EH, N-terminal His_6 and HA tag, TEV site, *mscL*$_{\Delta 116\text{-}136}$, $TnaC_{7\text{-}24}$	This study
pMS-KdpD-22-74-TnaC	KdpD-22-74	Ap^R, pMS119EH, N-terminal His_6 and HA tag, TEV site, *kdpD*$_{\Delta 1\text{-}21:\Delta 75\text{-}894}$, $TnaC_{7\text{-}24}$	This study

pMS-KdpD-22-74-W3A-TnaC	KdpD-22-74-W3A	Ap^R, pMS119EH, N-terminal His_6 and HA tag, TEV site, *kdpD* Δ1-21; G30A-G34A-G36A; Δ75-894, $TnaC_{7-24}$	This study
pMS-KdpD-22-74-3Q-TnaC	KdpD-22-74-3Q	Ap^R, pMS119EH, N-terminal His_6 and HA tag, TEV site, *kdpD*Δ1-21:R22Q,K24Q,K26Q;Δ75-894, $TnaC_{7-24}$	This study
pMS-KdpD-22-74-WalkerS-TnaC	KdpD-22-74-WalkerS	Ap^R, pMS119EH, N-terminal His_6 and HA tag, TEV site, *kdpD*Δ1-21:C32S;Δ75-894, $TnaC_{7-24}$	This study
pEM36-3C	FtsQ-4-85	Ap^R, N-terminal His_6 tag, C3-protease cleavage site, *FtsQ*$_{Δ1-3; Δ86-276}$, HA-tag, $TnaC_{7-24}$	Bischoff *et al.,* 2014
pUC19-T7-Luc[50]		Ap^R, pUC19, N-terminal strep tag, *luc*$_{Δ51-550}$, $SecM_{41-78}$, His_6 tag	Zhang *et al.,* 2010
pMS-Luc2-50-TnaC	Luc2-50	Ap^R, pMS119EH, N-terminal His_6 and HA tag, TEV site, *luc*$_{Δ51-550}$, $TnaC_{7-24}$	
pMS-His-HA-TnaC	sc-ribosome	Ap^R, pMS119EH, N-terminal His_6 and HA tag, TEV site, $TnaC_{7-24}$	This study

References Supplementary Information

Balzer, D., Ziegelin, G., Pansegrau, W., Kruft, V. & Lanka, E. KorB protein of promiscuous plasmid RP4 recognizes inverted sequence repetitions in regions essential for conjugative plasmid transfer. *Nucleic Acids Research* **20**, 1851-1858 (1992).

Bischoff, L., Wickles, S., Berninghausen, O., van der Sluis, EO. & Beckmann, R. Visualization of a polytopic membrane protein during SecY-mediated membrane insertion. *Nat Commun*. **5**, 4103 (2014).

Kuhn, P. et al. Ribosome binding induces repositioning of the signal recognition particle receptor on the translocon. *J Cell Biol*. **211**, 91-104 (2015).

Pross, E., Soussoula, L. Seitl, I., Lupo, D. & Kuhn, A. Membrane targeting and insertion of the C-tail protein SciP. *J. Mol. Biol*. **428**, 4218-4227 (2016).

Seidelt, B. et al. Structural insight into nascent polypeptide chain-mediated translational stalling. *Science* **326**, 1412-1415 (2009).

Seitl, I. Functional and structural studies of C-terminally extended YidC. *Dissertation University of Hohenheim* opus1191 (2016).

Zhang X, Rashid R, Wang K, Shan SO. Sequential checkpoints govern substrate selection during cotranslational protein targeting. *Science* **328**, 757-60 (2010).

References

1. Akopian, D., Dalal, K., Shen, K., Duong, F. & Shan, S. O. SecYEG activates GTPases to drive the completion of cotranslational protein targeting. *J. Cell Biol.* **200**, 397–405 (2013).

2. Kuhn, A., Koch, H. G. & Dalbey, R. E. Targeting and insertion of membrane proteins. *EcoSal plus* **7**, 1–27 (2017).

3. Seluanov, A. & Bibi, E. FtsY, the prokaryotic signal recognition particle receptor homologue, is essential for biogenesis of membrane proteins. *J. Biol. Chem.* **272**, 2053–2055 (1997).

4. Doudna, J. A. & Batey, R. T. Structural insights into the signal recognition particle. *Ann. Rev. Biochem.* **73**, 539–557 (2004).

5. Luirink, J. & Sinning, I. SRP-mediated protein targeting: structure and function revisited. *Biochim. Biophys. Acta* **1694**, 17–35 (2004).

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

7. Denks, K. et al. The signal recognition particle contacts uL23 and scans substrate translation inside the ribosomal tunnel. *Nature Microbio*. **2**, 16265 (2017).

8. Zopf, D., Bernstein, H. D., Johnson, A. E. & Walter, P. The methionine-rich domain of the 54 kd protein subunit of the signal recognition particle contains an RNA binding site and can be crosslinked to a signal sequence. *EMBO J*. **9**, 4511–4517 (1990).

9. Peluso, P. et al. Role of 4.5S RNA in assembly of the bacterial signal recognition particle with its receptor. *Science* **288**, 1640–1643 (2000).

10. Egea, P. F. et al. Substrate twinning activates the signal recognition particle and its receptor. *Nature* **427**, 215–221 (2004).

11. Shan, S. O., Chandrasekar, S. & Walter, P. Conformational changes in the GTPase modules of the signal reception particle and its receptor drive initiation of protein translocation. *J. Cell Biol.* **178**, 611–620 (2007).

12. Connolly, T., Rapiejko, P. J. & Gilmore, R. Requirement of GTP hydrolysis for dissociation of the signal recognition particle from its receptor. *Science* **252**, 1171–1173 (1991).

13. Keenan, R. J., Freymann, D. M., Stroud, R. M. & Walter, P. The signal recognition particle. *Ann. Rev*. Biochem. **70**, 755–775 (2001).

14. Welte, T. et al. Promiscuous targeting of polytopic membrane proteins to SecYEG or YidC by the *Escherichia coli* signal recognition particle. *Mol Biol Cell* **23**, 464–479 (2012).

15. Valent, Q. A. et al. Nascent membrane and presecretory proteins synthesized in *Escherichia coli* associate with signal recognition particle and trigger factor. *Mol Microbiol* **25**, 53–64 (1997).

16. Valent, Q. A. et al. The *Escherichia coli* SRP and SecB targeting pathways converge at the translocon*. EMBO J.* **17**, 2504–2512 (1998).

17. Lee, H. C. & Bernstein, H. D. The targeting pathway of *Escherichia coli* presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. *Proc. Natl. Acad. Sci. USA* **98**, 3471–3476 (2001).

18. Peterson, J. H., Woolhead, C. A. & Bernstein, H. D. Basic amino acids in a distinct subset of signal peptides promote interaction with the signal recognition particle. *J. Biol. Chem.* **278**, 46155–46162 (2003).

19. Batey, R. T., Rambo, R. P., Lucast, L., Rha, B. & Doudna, J. A. Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science* **287**, 1232–1239 (2000).

20. Maier, K. S. et al. An amphiphilic region in the cytoplasmic domain of KdpD is recognized by the signal recognition particle and targeted to the *Escherichia coli* membrane. *Mol. Microbiol.* **68**, 1471–1484 (2008).

21. Facey, S. J. & Kuhn, A. The sensor protein KdpD inserts into the *E. coli* membrane independent of the Sec translocase and YidC. *Eur. J. Biochem*. **270**, 1724–1734 (2003).

22. Jung, K. & Altendorf, K. Truncation of amino acids 12-128 causes deregulation of the phosphatase activity of the sensor kinase KdpD of *Escherichia coli*. *J. Biol. Chem*. **273**, 17406–17410 (1998).

23. Heermann, R., Altendorf, K. & Jung, K. The N-terminal input domain of the sensor kinase KdpD of *Escherichia coli* stabilizes the interaction between the cognate response regulator KdpE and the corresponding DNA-binding site. *J. Biol. Chem.* **278**, 51277–51284 (2003).

24. Hainzl, T. & Sauer-Eriksson, E. Signal-sequence induced conformational changes in the signal recognition particle. *Nat. Commun.* **6**, 7163.

25. Ullers, R. S. et al. Interplay of signal recognition particle and trigger factor at L23 near the nascent chain exit site on the *Escherichia coli* ribosome. *J. Cell. Biol.* **161**, 679–684 (2003).

26. Bornemann, T., Jöckel, J., Rodnina, M. V. & Wintermeyer, W. Signal sequence-independent membrane targeting of ribosomes containing short nascent peptides within the exit tunnel. *Nat. Struct. Mol. Biol.* **15**, 494–499 (2008).

27. Zhang, X., Rashid, R., Wang, K. & Shan, S. O. Sequential checkpoints govern substrate selection during cotranslational protein targeting. *Science* **328**, 757–60 (2010).

28. Facey, S. & Kuhn, A. Biogenesis of bacterial inner-membrane proteins. *Cell. Mol. Life Sci.* **67**, 2343–2362 (2010).

29. Pross, E., Soussoula, L., Seitl, I., Lupo, D. & Kuhn, A. Membrane targeting and insertion of the C-tail protein SciP. *J. Mol. Biol.* **428**, 4218–4227 (2016).

30. Loh, B., Haase, M., Müller, L., Kuhn, A. & Leptihn, S. The transmembrane morphogenesis protein gp1 of filamentous phages contains Walker A and Walker B motifs essential for phage assembly. *Viruses* **9**, 73 (2017).

31. Wertman, K. F., Wyman, A. R. & Botstein, D. Host/vector interactions which affect the viability of recombinant phage lambda clones. *Gene* **49**, 253–62 (1986).

32. Seitl, I., Wickles, S., Beckmann, R., Kuhn, A. & Kiefer, D. The C-terminal regions of YidC from *Rhodopirellula baltica* and *Oceanicaulis alexandrii* bind to ribosomes and partially substitute for SRP receptor function in *Escherichia coli*. *Mol Microbiol* **91**, 408–421 (2014).

33. Maniatis T., Fritsch, E. F. & Sambrook, J. Molecular Cloning: A Laboratory Manual. *Cold Spring Harbour Press* (1982).

34. Studier, F. W. & Moffatt, B. A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **189**, 113–30 (1986).

35. Calhoun, K. A. & Swartz, J. R. Total amino acid stabilization during cell-free protein synthesis reactions. *J. Biotechnol.* **123**, 193–203 (2006).

36. Balzer, D., Ziegelin, G., Pansegrau, W., Kruft, V. & Lanka, E. *Nucl. Acid Res.* **20**, 1851–1858 (1992).

37. Seitl, I. Functional and structural studies of C-terminally extended YidC. *Dissertation University of Hohenheim* opus 1191 (2016).

Acknowledgements

We are grateful to Katja Maier and Sandra Facey who had initiated this project. We would like to thank Ines Seitl for providing pMS-Ffh-C-Strep and pMS-Ffh-M423C-C-Strep, the KdpD peptide, support and discussions, Cathleen Panek and Bettina

Schmidt for assistance in the lab, Samuel Kraus for plasmid pMS-Ffh-L181C-C-Strep and R Beckmann, SO Shan and HG Koch for strains and plasmids. This work was supported by the Deutsche Forschungsgemeinschaft Ku749/5 and by a stipend of the Landesgraduiertenförderung Baden-Württemberg.

Author Contributions

The experiments were designed by A.K. and performed by E.P. The results were discussed and the manuscript was written by A.K. and E.P.

3

Membrane Targeting and Insertion of the C-Tail Protein SciP

Eva Pross, Lavinia Soussoula, Ines Seitl,
Domenico Lupo and Andreas Kuhn

Institute of Microbiology and Molecular Biology,
University of Hohenheim, 70599 Stuttgart, Germany

Journal of Molecular Biology 428 (20): 4218-4227

Abstract

C-tailed membrane proteins insert into the bilayer post-translationally because the hydrophobic anchor segment leaves the ribosome at the end of translation. Nevertheless, we find here evidence that the targeting of SciP to the membrane of *Escherichia coli* occurs co-translationally since signal elements in the N-terminal part of the SciP protein sequence are present. Two short hydrophobic sequences were identified that targeted a green fluorescent protein-SciP fusion protein to the membrane involving the signal recognition particle. After targeting, the membrane insertion of SciP is catalyzed by YidC independent of the SecYEG translocase. However, when the C-terminal tail of SciP was extended to 21 aa residues, we found that SecYEG becomes involved and makes its membrane insertion more efficient.

Introduction

Most membrane proteins use the Sec translocase and the YidC insertase for their insertion reaction. Thereby, the transmembrane (TM) region of the substrate first interacts with SecY, most likely at the lateral gate and possibly together with YidC [1]. The polar surface of the translocase center easily allows the transfer of large hydrophilic protein regions into the periplasm, while the TM region can interact with the hydrophobic lateral gate. In general, bacterial membrane proteins are co-translationally inserted and delivered to the translocase early during their biosynthesis by the ribosome. This ensures that the hydrophobic TM regions can readily interact with the translocase, preventing their aggregation. However, the C-tail proteins are synthesized with a large cytoplasmic domain first before the TM segment has left the exit tunnel of the ribosome. In the endoplasmic reticulum (ER), the C-tail proteins are first recognized by the Get3/Get4 complex in the cytoplasm. This complex guides the protein to Get1/2 for membrane insertion, an ER membrane complex [2]. In *Escherichia coli*, no homologs of the Get proteins exist and it has been found that the membrane insertase YidC inserts the C-tail-anchored protein TssL [3].

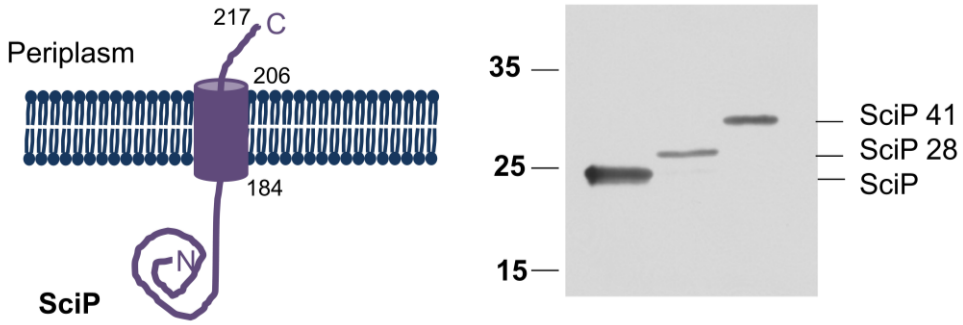
SciP (TssL) is a central component of the type six secretion system (T6SS) [4] that is found in the pathogenic strains of enteroaggregative *E. coli* (EAEC), *Salmonella enterica*, and *Vibrio cholerae*. The T6SS core complex consists of about 13 components and is assembled in the inner membrane. Many of the proteins show a striking structural similarity to phage T4 tail and baseplate proteins [5]. This secretion system is involved in

the delivery of toxic effectors into target cells and in biofilm formation [6]. The biosynthetic assembly pathway of the T6SS is well understood and defines a major membrane-based subcomplex [7], of which SciP (TssL) is a part of. Here, we have studied the biosynthesis of SciP and its targeting to the membrane and its insertion into the membrane. We show with protease mapping and with periplasmic accessibility for cysteine modification that the C-terminal tail is exposed in the periplasm and that the presence of YidC is required to achieve this topology. We also show that it is possible to extend the C-terminal tail by 10 additional amino acid residues, which then involves SecYEG in the membrane insertion process. However, an extension of more than 10 residues at the C terminus prevented membrane insertion of SciP.

Results

Membrane insertion of SciP requires the membrane insertase YidC

SciP has a large cytoplasmic domain at the N terminus of 183 aa residues, followed by a 25-residue-long membrane anchor and a short C-terminal tail of 11 residues in the periplasm. To investigate the insertion of the C-tail-anchored protein SciP in detail, we constructed a collection of mutants with C-terminal extensions (Fig. 1).



SciP	- ADQVARLTGQL
SciP-C	- ADQVARLTGQLC
SciP20	- ADQVARLTGLMASMTGGCMG
SciP21	- ADQVARLTGQLMASMTGGCMG
SciP22	- ADQVARLTGQLMASMTGGQQMG
SciP26	- ADQVARLTGQLASGSYPYDVPDYAQL
SciP28	- ADQVARLTGQLMASMTGGQQMGSGSAQL
SciP41	- ADQVARLTGQLASGSYPYDVPDYAQASGSYPYDVPDYAQL

Fig. 1. Variants of the SciP protein. The topology of the C-tail protein SciP with its TM segment between residues 184 and 206 is depicted. The amino acid sequences of the C-tail (207–217) and its derivatives are shown in the single-letter code. Expression of SciP and the variants SciP28 and SciP41 that contain an amino-terminal His_{10}-tag was analyzed by SDS-PAGE on a Western blot. The numbers on the left margin refer to the molecular weight in kDa.

The plasmid-borne expression of the N-terminally His_{10}-tagged protein and the extended variants is documented by their respective apparent molecular weight on a polyacrylamide gel.

First, the membrane insertion of SciP with a non-modified C-tail was investigated by protease mapping in the *E. coli* strain MK6 that either expressed YidC or was YidC-depleted (Fig. 2). The cells were pulse-labeled with ^{35}S-methionine for 3 min, converted to spheroplasts, and acid precipitated. The SciP protein was immunoprecipitated and analyzed by SDS-PAGE. When YidC was present, the addition of proteinase K to the outside of the cells resulted in a partial digestion of the protein (lane 2), whereas in the absence of YidC, the SciP protein remained intact, indicating a cytoplasmic location of the protein (lane 5). In both cases, lysis of the cells with detergent led to the complete digestion of the SciP protein (lanes 3 and 6). For a control, the cytoplasmic GroE protein was monitored for the protease inaccessibility (lane 8) and the OmpA protein for its accessibility (lane 10) in the YidC-depleted cells. Our results therefore confirm the observation of Aschtgen *et al.* [3] showing that the membrane insertion of SciP requires YidC.

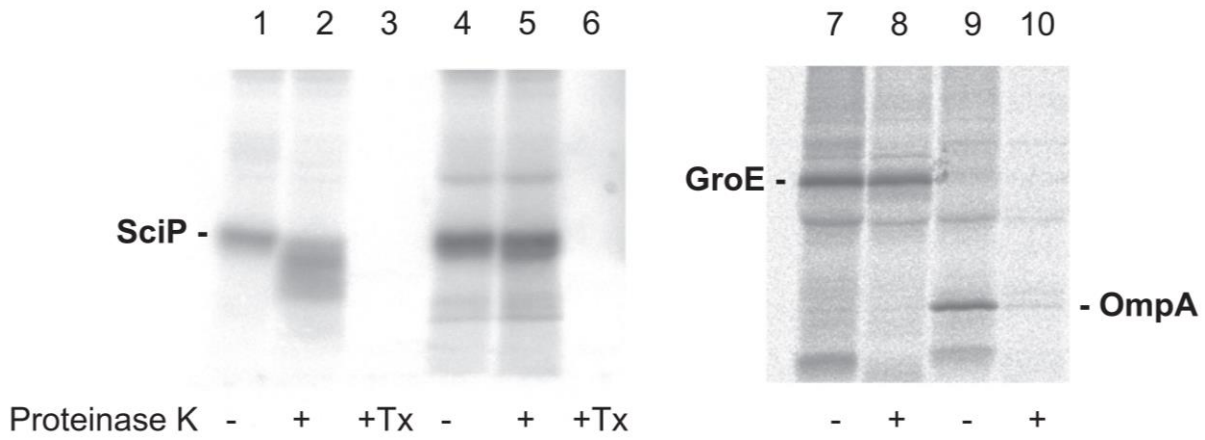


Fig. 2. Membrane insertion of SciP. SciP was pulse-labeled in *E. coli* MK6 cells with ^{35}S-methionine for 3 min and analysed by protease mapping. The immunoprecipitated protein was partially accessible by proteinase K (lane 2). Lysis of the cells with the detergent Triton X-100 resulted in the complete digestion of SciP (lanes 3 and 6). When the cells were depleted for YidC (lanes 4-6), the membrane insertion of SciP was inhibited (lane 5). The cytoplasmic protein GroE and the outer membrane protein OmpA served as control proteins (right panel).

Targeting and translocation factors

Since the protease mapping of SciP resulted in an array of smaller protein bands (Fig. 2, lane 2), we decided to investigate the membrane insertion of SciP with another method. To do so, we generated a site-directed mutant named SciP-C that has an additional cysteine residue at the C terminus at position 218. The endogenous cysteine residue at position 200 was mutated to an alanine. First, the contribution of the Sec translocase was tested in CM124, which depletes SecE and hence also SecY [8] when no arabinose is present in the medium. The cells were pulse-labeled for 2 min and chased with non-radioactive methionine for 10 min. To investigate whether the C-terminal cysteine residue was exposed in the periplasm, we added 200 mM 4-acetoamido-4′-maleimidylstilbene- 2,2′disodium sulfonate (AMS) 1 min before pulse labeling. Since the amount of the modifiable protein differed among the depletion strains used, a direct control with the same strain under non-depleted conditions was necessary. Depletion of SecE did not show any effect on the translocation of the C-terminal tail of SciP-C (Fig. 3a). For a control, we analyzed the processing of proOmpA (lanes 5 and 6), which is known to be inhibited in the absence of SecE [9]. However, when YidC was depleted in MK6 cells, SciP-C showed no derivatization of its C-terminal cysteine (Fig. 3b, lane 4), verifying the importance of YidC. In addition, a possible involvement of the cytoplasmic signal recognition particle (SRP; composed of Ffh and 4.5S RNA) was tested (Fig. 3c). SciP-C was expressed in MCΔFfh, an arabinose-dependent depletion strain for the SRP protein. When the cells were grown in glucose medium, Ffh was depleted after 3 h, whereas growth in arabinose allowed the expression of Ffh. When Ffh was present, about 41% of the labeled protein was derivatized by AMS and shifted on the polyacrylamide gel by about 0.5 kDa, indicating a periplasmic location of the cysteine residue (lane 2). In contrast, in the absence of Ffh, very little (8%) shifted protein was detectable (lane 4). Similarly, SciP-C insertion was affected after the depletion of FtsY in *E. coli* IY26 (Fig. 3d). Taken together, the data show that the C-terminal tail of SciP is translocated by YidC but is independent of SecYEG, whereas the targeting of SciP is supported by SRP.

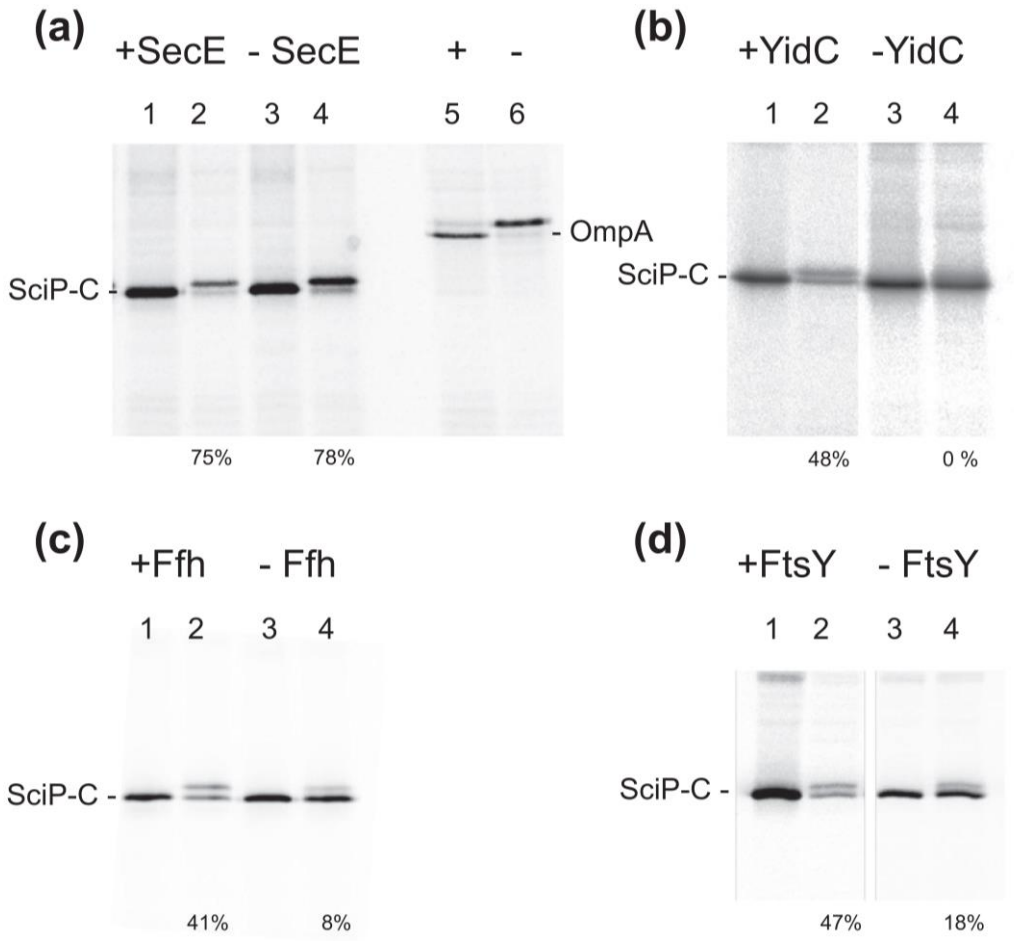


Fig. 3. Translocation of the C-tail of SciP-C. The 11-residue-long C-tail region of SciP was extended by a cysteine residue and expressed in the depletion strains for SecE: (a) CM124, (b) YidC: MK6, (c) Ffh: MCΔFfh, and (d) FtsY: IY26. In the even-numbered lanes, AMS was added for 10 min. The cells were then pulse labeled with ^{35}S-methionine for 2 min and chased for 10 min with non-labeled methionine. The samples were immunoprecipitated and analyzed by SDS-PAGE and fluorography. The SciP-C protein was labeled under non-depleting conditions (lanes 1 and 2) or under depleted conditions (lanes 3 and 4). The modification of the C-terminal cysteine by AMS results in a mobility shift of the protein band indicating the periplasmic exposure of the cysteine residue. The relative amount of the modified protein was quantified and is indicated at the bottom of the gel. For a control, the SecE depletion was tested by analyzing the inhibited processing of proOmpA to OmpA (lanes 5 and 6).

The translocation of an extended C-tail is YidC and SecYEG dependent

To explore whether the C-terminal tail of SciP can be extended without affecting its translocation efficiency into the periplasm, we generated several mutants (Fig. 1). The plasmid-encoded variants were transformed into MK6 and induced with 1 mM IPTG for 10 min. The extended tails allowed the use of the protease mapping technique.

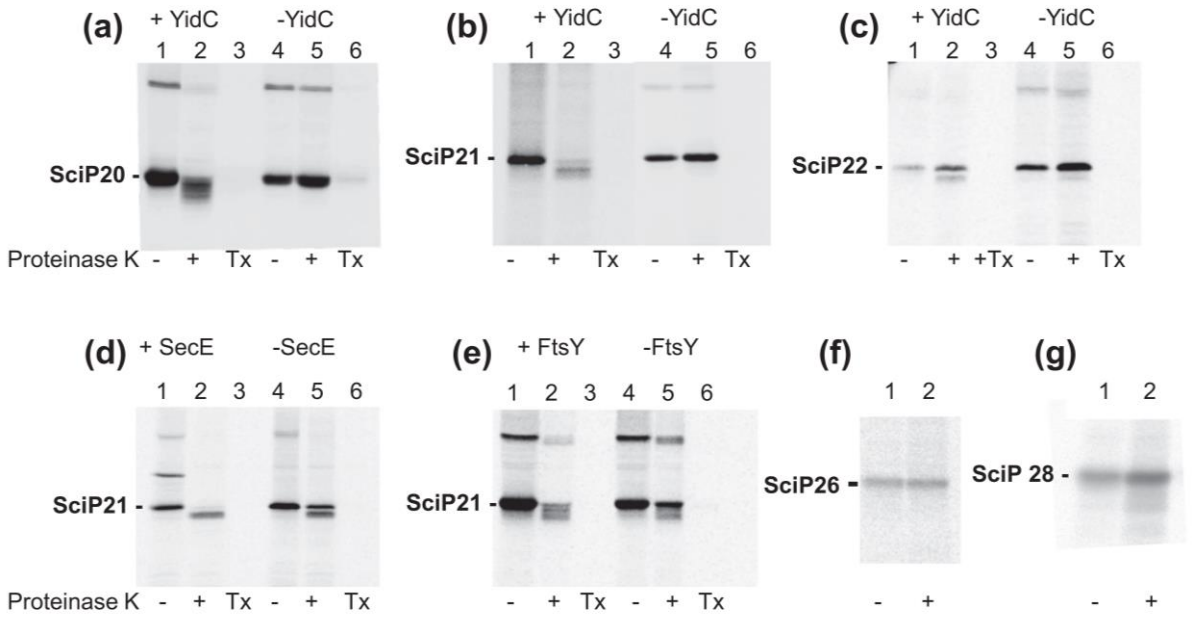


Fig. 4. C-tail extensions of SciP. SciP with extended tails of 20, 21, 22, 26, and 28 residues was expressed in MK6 under (a–c, f, g) non-depleted or (a–c; lanes 4–6) YidC-depleted conditions. The proteins were analyzed for membrane insertion as described for Fig. 2. Whereas (a) SciP20 and (b) SciP21 inserted into the membrane quite efficiently and were dependent on YidC, (c) SciP22 was partially inhibited for membrane insertion. The translocation of the C-tail of SciP21 was analyzed after (d) SecE depletion in CM124 and (e) FtsY depletion in IY26 (f and g). SciP with extended tails of 26 and 28 aa residues (see Fig. 1 for their sequence) was expressed in MK6 cells in 0.2% arabinose. The topology of the proteins was analyzed with proteinase K as described in Fig. 2. The accessibility of OmpA and the protection of GroE were used as a control (Fig. S4).

^{35}S-methionine was added to the cells for 3 min, followed by a chase with non-labeled methionine for 10 min, and the cells were processed for protease mapping. An extended C-tail of 20 aa residues (with 9 added residues) was readily translocated (Fig. 4a) in the presence of YidC but not in its absence (lanes 2 and 5). This was also the case for SciP21 (Fig. 4b). In addition, the translocation of the C-tail of SciP21 was partially inhibited when SecYE (Fig. 4d) or FtsY (Fig. 4e) was depleted, respectively. However, when the C-tail was extended to 22 residues that include two hydrophilic glutamines, membrane insertion was observed but it was inefficient (Fig. 4c, lane 2). Only about 50% of SciP22 was converted to the proteolyzed form, considering the number of the remaining methionine residues. When the C-tail was extended to 26, 28 (Fig. 4f and g), or 41 residues (not shown), SciP was not translocated at all. We conclude from our results that YidC translocates C-terminal small periplasmic protein domains and is supported by SecYEG to translocate C-tails up to 22 residues.

Translocation factors investigated by green fluorescent protein (GFP)-mediated localization

A plasmid encoding a fusion protein of the GFP and SciP was constructed and expressed in various depletion strains to explore the effects of Ffh, FtsY, YidC, and SecYE on the localization of the fusion protein. When the strains were grown with arabinose, in which the targeting and translocation components were expressed, the strains showed normal growth. In these cases, the GFP-SciP fusion protein appeared at the membrane surface in the fluorescence microscope (Fig. 5, upper panels). When the cells were depleted for Ffh (Fig. 5b), FtsY (Fig. 5d), and YidC (Fig. 5f), the fluorescence was no longer evenly distributed at the membrane but appeared as distinct spots probably caused by the aggregated fusion protein. In contrast, in the Sec-depleted cells, the fluorescent signal was detected at the membrane, suggesting a normal membrane targeting of GFP-SciP (Fig. 5h). The depletion of the respective proteins was verified in each culture by analyzing a sample on a Western blot taken at the time point of induction (Fig. S1).

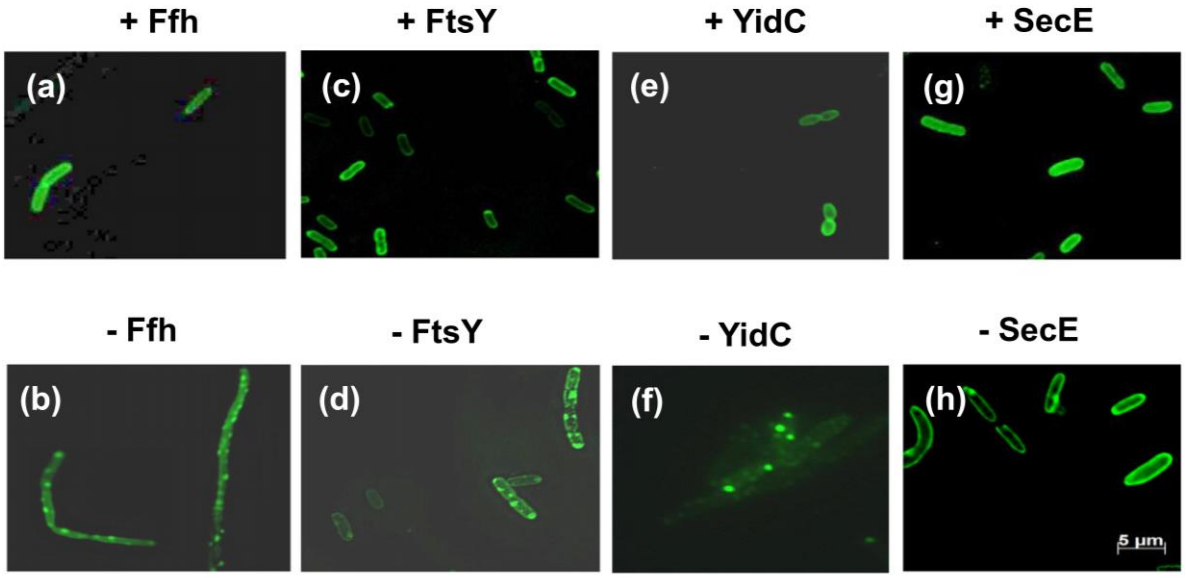


Fig. 5. GFP-SciP requires targeting and insertion factors. The cellular distribution of the fluorescent signal of GFP-SciP was investigated in non-depleted conditions (upper panels) or depleted conditions (lower panels). The depletion strains (a and b) MCΔFfh, (c and d) IY26, (e and f) MK6, and (g and h) CM124 expressing GFP-SciP were grown to an OD_{600} of 0.5 at 37°C, induced with 1 mM IPTG for 1 h at 30°C, and applied onto poly-L-lysine-coated cover slides for fluorescence microscopy.

DnaK and DnaJ are not involved in SciP-C membrane targeting and insertion

Since a previous study [3] had suggested a strong involvement of DnaK and DnaJ, we carefully tested the localization and membrane insertion of SciP-C in the DnaK deletion strain GP365 and in the DnaJ deletion strain BB4312. The absence of DnaK and DnaJ was verified on a Western blot (Fig. S1), and the membrane insertion of SciP-C was not impaired in these strains (Fig. S2A and B). Membrane targeting was also tested in the two deletion strains with the GFP-SciP fusion protein. The fluorescent signal of the GFP-SciP was clearly found at the membrane under *dnaK*⁻ and *dnaJ*⁻ conditions (Fig. S2D and F).

Targeting to the membrane depends on N-terminal region

To test whether the membrane targeting of SciP occurs early after synthesis involving a targeting signal within its first 100 residues, we applied a hydrophobicity analysis according to Kyte and Doolittle [10]. Inspection of the N-terminal 100 residues of SciP showed two regions (residues 12-20 and residues 62-71) with a higher hydrophobicity score (Fig. S3), suggesting that they could function as targeting signals. To test whether the hydrophobic domains in the N-terminal region of SciP are involved in membrane targeting, we made a deletion mutant in GFP-SciP that had the two hydrophobic regions missing (Fig. 6). The fluorescence of this mutant was evenly distributed in the cytoplasm (Fig. 6b) when compared to the non-deleted GFP-SciP (Fig. 6a), suggesting that the hydrophobic regions are involved in targeting the protein to the membrane surface.

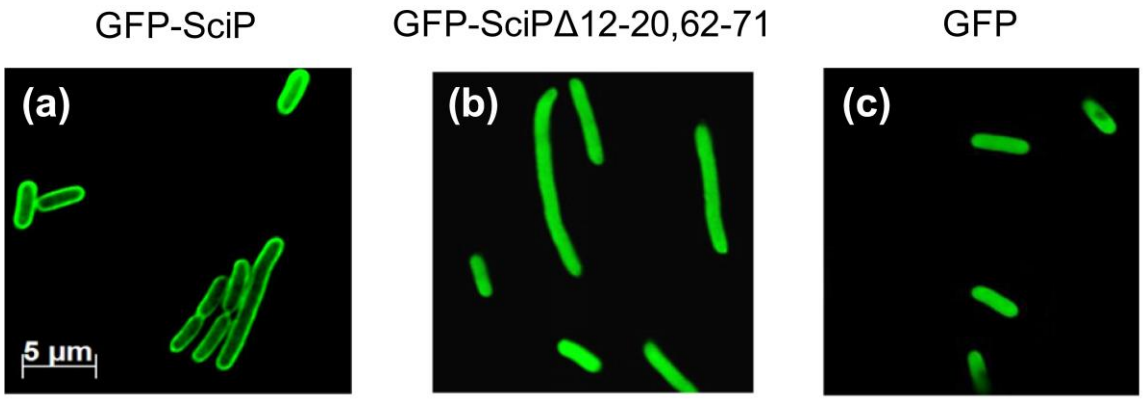


Fig. 6. Hydrophobic regions 12-20 and 62-71 in the N-terminal domain of SciP are required for membrane targeting. The cellular localization of (a) GFP-SciP, (b) GFP-SciPΔ12-20/62-71, and (c) GFP in MC4100 cells was analyzed with fluorescence microscopy. Whereas GFP-SciP was localized at the membrane, GFP-SciPΔ12-20/62-71 and GFP alone were found distributed in the cytoplasm.

To test whether the short hydrophobic regions (see Fig. S3 for their amino acid sequences) might resemble a membrane targeting signal, we investigated each of the two regions with GFP fusion proteins for its capability to direct GFP to the membrane surface.

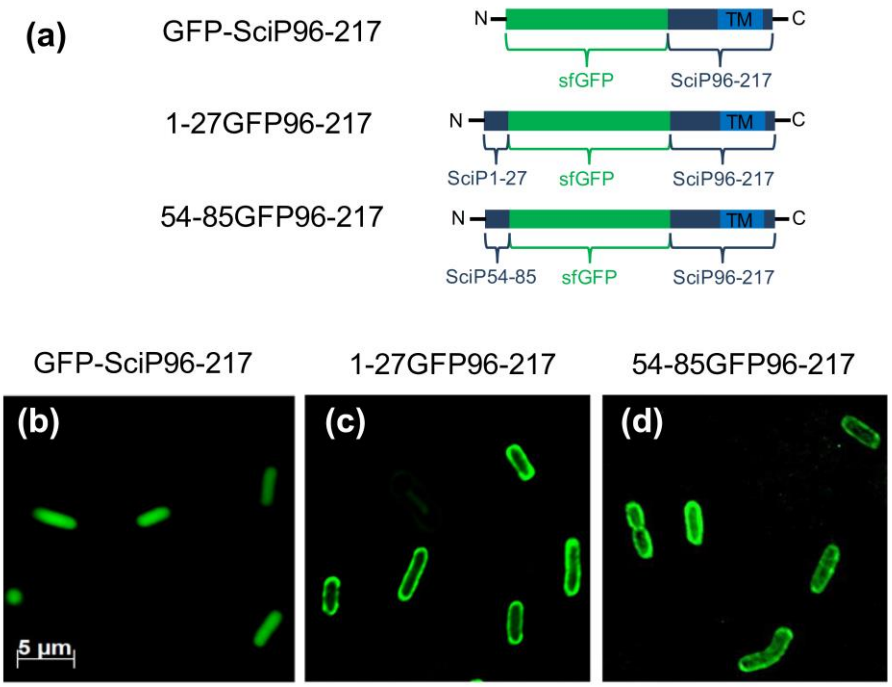


Fig. 7. Residues 1-27 and 54-85 in the N-terminal domain of SciP exhibit membrane targeting signals. (a) The suspected targeting signals 1-27 and 54-85 were fused to the N terminus of GFP-SciP96-217. The cellular localization of (b) GFP-SciP96-217, (c) 1-27GFP-SciP96-217, and (d) 54-85GFP-SciP96-217 in MC4100 cells was analyzed by fluorescence microscopy. Whereas GFP-SciP96-217 was distributed in the cytoplasm, 1-27GFP-SciP96-217 and 54-85GFP-SciP96-217 were localized at the membrane.

When GFP had residues 96-217 of SciP at its C terminus, the fluorescent signal was distributed in the cytoplasm (Fig. 7b). Residues 1-27 or 54-85 of SciP were then fused to the N terminus of GFP96-217 by directed mutagenesis (Fig. 7a). Expression of 1-27GFP96-217 and 54-85GFP96-217, respectively, was then induced in MC4100 cells, and the cellular localization of both fusion proteins was followed by fluorescence microscopy (Fig. 7c and d). Clearly, the fluorescent signal appeared at the membrane, suggesting a membrane targeting function of both hydrophobic regions, 1-27 and 54-85. Taken together, we conclude that SciP is targeted to the membrane early during its synthesis, allowing a swift membrane insertion of its C-terminal membrane anchor region by YidC.

Discussion

The membrane insertion of C-tail proteins poses a conundrum that the major part of the protein is synthesized and has left the ribosome before the hydrophobic TM region is exposed at the ribosomal exit tunnel. Therefore, membrane targeting of these proteins might involve signals that are separate from the TM region and reside within the N-terminal part of the protein. Indeed, we found that two short hydrophobic sequences (Fig. S3) within the N-terminal cytoplasmic domain of SciP could play this role. When we deleted the amino acid residues 12-20 and 62-71 of SciP, the targeting of a GFP-SciP fusion protein to the membrane was inhibited (Fig. 6), suggesting that these protein regions might interact with SRP during their synthesis and co-translationally target the protein to the membrane. Such a targeting mechanism has been observed earlier with the sensor protein KdpD, a four-spanning membrane protein that has an N-terminal large domain of 400 residues in the cytoplasm [11]. In this domain, a small hydrophobic sequence was found that showed affinity to SRP, explaining the involvement of SRP in the membrane insertion of KdpD.

Intriguingly, the C-tail protein SciP was independent of the Sec translocase for its membrane insertion. Also, the C-tailed proteins of the ER do not use the Sec translocase, and their membrane insertion is catalyzed by the Get pathway [2]. In *E. coli*, the C-tail-anchored protein SciP is inserted by the membrane insertase YidC [3]. YidC is a six-spanning membrane protein that has a hydrophilic groove in the cytoplasmic leaflet of the bilayer [12]. This groove binds the substrate protein domain that is to be translocated into the periplasm. In addition, YidC contacts the hydrophobic region of the substrates, Pf3 coat protein, and MscL between TM3 and TM5 [13,14].

When the translocated regions of MscL and M13 procoat protein were extended, their membrane insertion became dependent on Sec translocase [15,16]. However, not only the length of the translocated region is important but also its hydrophilicity [17,18]. Sec and YidC share the lateral gate [19], which then allows the translocation of a large hydrophilic domain by the Sec translocase. When the C-tail of SciP was extended by 10 residues (in SciP21), we observed that the Sec translocase supported its membrane insertion. However, when the periplasmic region of SciP was extended by more than 15 residues, the protein was no longer inserted into the membrane. Although other sequence extensions should be tested, our data suggest that the size and polarity of the extended region are limited. Whereas the extended C-tail of SciP21 by 10 residues was readily translocated, an 11-residue extension in SciP22, where the cysteine residue was

replaced by two glutamines, was already inefficient (Fig. 4c).

Previously, the membrane insertion of SciP/TssL has been studied with membrane fractionation in different deletion and depletion strains [3]. We have extended this study and found that the membrane insertion of the protein requires not only YidC but also Ffh, a component of the bacterial SRP. In addition, the membrane insertion of SciP was also severely affected when FtsY was depleted (Fig. 3D). This latter result is in contrast to Aschtgen *et al.* [3] who followed TssL after cell fractionation on a Western blot. Since our study is based on pulse-chase experiments, the fast kinetics of membrane targeting is better resolved. Also, a GFP-SciP fusion protein was inhibited to form a membrane staining in the absence of Ffh or FtsY, but it resulted in patches in the cells (Fig. 5b and d). Such patches were also observed when YidC was depleted (Fig. 5f). Most likely, the patches resemble the non-inserted protein aggregates that might fractionate with the membrane. In addition, we investigated the membrane insertion in CM124, a depletion strain for SecE and consequently for SecY [8], showing that Sec translocase is not involved in the translocation of the C-tail of SciP.

The role of DnaJ and DnaK, the members of the Hsp70 chaperone family remains less clear. Aschtgen *et al.* [3] had observed a reduction of the membrane localization by about 20% when DnaJ or DnaJ and DnaK were deleted. Although we could not detect any contribution of DnaK and DnaJ to membrane targeting and insertion, we cannot exclude that these chaperones might support the folding of the protein in the cytoplasm (Fig. S2). We suspect that in the previous study of Aschtgen *et al.* [3], TssL was overproduced such that the correct folding of the protein and consequently its membrane insertion depended on DnaK and DnaJ. To prevent overexpression, we induced the cells for only 10 min and followed radiochemical amounts (Fig. S2A and B). Our data suggest that membrane targeting of SciP is mediated by SRP co-translationally when the N-terminal part leaves the ribosome. Particularly, C-tail membrane proteins can use the time when the N-terminal cytoplasmic part of the protein leaves the ribosome for interaction with SRP and membrane targeting. This ensures that the protein is already localized at the membrane surface when the hydrophobic membrane anchor sequence at the C terminus leaves the ribosome.

Materials and Methods

Plasmid construction

The *sciP* gene was amplified from plasmid pIBA37-P (pIBA-TssL; kindly provided by Eric Cascales, Marseille, France) using the primers 5′ ACA CAT ATG AAT AAA CCT GTT ATC TCC CGG GC 3′ (forward) and 5′ ACA CTC GAG TTA CAA TTG CCC GGT AAG CCG TGC 3′ (reverse) flanking the PCR product with an *Nde*I and *Mun*I/*Xho*I restriction sites. The PCR product was cloned into the vector pET16b (Novagen) using *Nde*I and *Xho*I (pET16b-SciP), which placed a His_{10} tag at the N-terminal part of SciP and in a pGZ119EH [20] derivative using *Nde*I and the compatible restriction enzymes *Xho*I/*Sal*I. From this pGZ-SciP plasmid, pMS-SciP was constructed using *Eco*RI and *Sph*I.

The plasmids pGZ-SciP26 and pGZ-SciP28 carrying a C-terminal HA tag or T7 tag were constructed using the *Mun*I-linearized pGZ-SciP plasmid and the oligonucleotides 5′ AAT TAG CGT CAG GTT CAT ATC CGT ATG ACG TTC CGG ATT ATG CAC 3′ (forward) and 5′ GTG CAT AAT CCG GAA CGT CAT ACG GAT ATG AAC CTG ACG CTA ATT 3′ (reverse) for HA tag and 5′ AAT TAA TGG CGT CCA TGA CCG GTG GTC AAC AGA TGG GTT CAG GTT CAG CGC (forward) and 5′ GCG CTG AAC CTG AAC CCA TCT GTT GAC CAC CGG TCA TGG ACG CCA TTA ATT 3′ (reverse) for T7 tag.

The plasmid pGZ-SciP22 (periplasmic domain of 22 aa) was constructed by replacing the *sciP-T7* gene on pGZ-SciP28 with the PCR product amplified from this plasmid using the primers 5′ ACA CAT ATG AAT AAA CCT GTT ATC TCC CGG GC 3′ (forward) and 5′ CGG CAA TTG TTA ACC CAT CTG TTG ACC ACC GGT CAT G 3′ (reverse) and the restriction enzymes *Nde*I and *Mun*I. The plasmid pMS-SciP21 was constructed using the QuikChange site-directed mutagenesis method on plasmid pGZ-SciP22 with the oligonucleotides 5′ CC ATG ACC GGT GGT TGC ATG GGT TAA CAA TTG 3′ and 5′ CAA TTG TTA ACC CAT GCA ACC ACC GGT CAT GG 3′ and was subcloned in pMS119EH [21] with *Eco*RI and *Sph*I. Plasmid pMS-SciP20 was constructed on plasmid pMS-SciP21 with the primers 5′ GTG GCA CGG CTT ACC GGG TTA ATG GCG TCC ATG ACC GG 3′ and 5′ CC GGT CAT GGA CGC CAT TAA CCC GGT AAG CCG TGC CAC 3′.

For AMS derivatization, the cysteine residue at position 200 in the SciP protein was replaced by an alanine residue by site-directed mutagenesis on pMS-SciP using the

oligonucleotides 5′ GCT GGT CTC TGG GCT GTG CTT TCT TCT 3′ (forward) and 5′ AGA AGA AAG CAC AGC CCA GAG ACC AGC 3′ (reverse). Then, a cysteine codon at position 218 of the SciP protein was inserted using the oligonucleotides 5′ CGG GCA ATT GTG CTA ACT CGA CC 3′ (forward) and 5′ GG TCG AGT TAG CAC AAT TGC CCG 3′ (reverse; pMS-SciP-C). The gene for SciP-C was subcloned into pLZ1 [22] using *Nco*I and *Hin*dIII.

For the localization studies, SciP was fused at the N-terminal part with the superfolder GFP (sfGFP) [23]. The *sfGFP* gene was amplified from a pET24d derivative (kindly provided by Gunter Stier, BZH Heidelberg, Germany) using the primers 5′ G GCG AAT TCA GGA GGT ATA CAT ATG GGA TCC 3′ (forward) and 5′ CGG TCT AGA GCT GCC TTT ATA CAG TTC ATC C 3′ (reverse) flanking the PCR product with an *Eco*RI and *Xba*I restriction site. The *sciP* gene was amplified from the plasmid pMS-SciP using the oligonucleotides 5′ GGC TCT AGA AAT AAA CCT GTT ATC TCC CGG GCT G 3′ and 5′ CGG AAG CTT TTA CAA TTG CCC GGT AAG CCG 3′ flanking the PCR product with an *Xba*I and *Hin*dIII restriction site. The two PCR products were cloned in the pMS119EH [21] vector using the respective restriction enzymes. Plasmid pLZ1-GFP-SciP was constructed using the restriction enzymes *Nsb*I and *Xho*I and was blunted with T4 DNA polymerase. Tac-GFP-SciP-rrnBT from pMS-GFP-SciP was cloned into the blunted pLZ1 plasmid using the enzyme *Ssp*I.

For the localization studies of the SciP deletion mutants, the QuikChange site-directed mutagenesis method was performed on plasmid pMS-GFP-SciP using the primers 5′ CT GTT ATC TCC CGG GCT GAA CAG AGC CAG CTG CGC AGC 3′/5′ GCT GCG CAG CTG GCT CTG TTC AGC CCG GGA GAT AAC AG 3′ (Δ 12–20) and 5′ GCA GGA TTC AGT CAG AAA AGC AGT GAC GAC GAG AGT GTA CTG AAC CGC G 3′/5′ C GCG GTT CAG TAC ACT CTC GTC GTC ACT GCT TTT CTG ACT GAA TCC TGC 3′ (Δ62–71).

Plasmid pMS-GFP-SciP96-217 was constructed using the restriction enzymes *Xma*I and *Pst*I blunted with T4 DNA polymerase. Two different SciP fragments (SciP1-27 and SciP54-85) were subcloned into a pET24d derivative using the oligonucleotides 5′ GGC AAG CTT ATG AAT AAA CCT GTT ATC TCC CGG G 3′ (forward) and 5′ CGG GGA TCC TTG CCC GCT GCG CAG C 3′ (reverse), respectively, and the oligonucleotides 5′ GGC AAG CTT ATG GGA TTC AGT CAG AAA AGC AGT GAC 3′ (forward) and 5′ CGG GGA TCC CCA GCC ATC GTC TGT TTT TTC G 3′ (reverse) and were ligated into the *Hin*dIII and *Bam*HI sites. Then, the two SciP fragments were cloned into pMS-

GFP-SciP96-217 using the restriction enzymes *Eco*RI and *Bam*HI.

Bacterial strains and growth conditions

YidC depletion was carried out using the *E. coli* strain MK6 [13], where the *yidC* promoter had been exchanged by the *araC-araBAD* promoter, thus allowing the depletion of the chromosomally encoded *yidC* by the addition of 0.4% glucose.

The FtsY depletion strain IY26 (BW25113-Kan-AraCP- ftsY) was obtained from Eitan Bibi (Weizman Institute of Science, Rehovot, Israel). In the IY26 strain, the *ftsY* gene is under the control of the *araBAD* promoter and operator, which allow the depletion of the chromosomally encoded *ftsY* by the addition of 0.4% glucose.

Ffh depletion was carried out using the *E. coli* strain MCΔFfh (MC1061-Kan-AraCP-ffh) [24]. In the MCΔFfh strain, the *ffh* gene is under the control of the *araC-araBAD* operator region. For Ffh, depletion cells were grown in the presence of 0.4% glucose.

For SecE depletion, the *E. coli* CM124 [25] was used. If SecE is depleted in cells, SecY is rapidly degraded [8]. Therefore, SecE depletion also causes a depletion of SecY (Fig. S1).

For DnaK deletion, the *E. coli* strain GP365 was used. This strain is a derivative of *E. coli* GP355 [26] obtained by an insertion of the chloramphenicol resistance gene into *dnaK52.* The deletion of *dnaJ* was a tetracycline resistance insertion into MC4100, resulting in *E. coli* BB4312 (both strains were kindly provided by Bernd Bukau, Heidelberg, Germany).

The *E. coli* depletion strains MK6, MCΔFfh, and IY26, respectively, were cultivated in LB supplemented with 0.2% arabinose and 0.4% glucose. The overnight culture was washed twice with fresh LB to remove the arabinose and glucose and was back diluted (MK6 1:200; MCΔFfh 1:50; IY26 1:200) either in fresh LB with 0.2% arabinose to allow YidC, Ffh, or FtsY expression or in LB with 0.4% glucose to deplete cells for YidC, Ffh, or FtsY, respectively.

For pulse-chase labeling, cells were grown for 3 h to an OD_{600} of 0.5, washed and resuspended in M9 minimal media lacking methionine, supplemented with either 0.2% arabinose or 0.4% glucose, and were grown for 45 min prior to pulse-chase labeling.

The *E. coli* strain CM124 (SecE depletion) was cultivated in M9 supplemented with 0.2% arabinose and 0.4% glucose. The overnight culture was washed twice with M9 to remove the arabinose and was back diluted 1:20 in fresh M9 either with 0.2% arabinose and 0.4% glucose to allow SecE expression or with 0.4% glucose to deplete cells for SecE. For pulse-chase labeling, the cells were grown to an OD_{600} of 0.5, washed and

resuspended in M9 minimal media lacking methionine, supplemented with 0.2% arabinose and 0.4% glucose or 0.4% glucose, and were grown for 45 min prior to the pulse-chase labeling.

The *E. coli* strains MC4100, BB4312, GP355, and GP365 were grown overnight in LB and were back diluted 1:100 in fresh LB. For pulse-chase labeling, the cells were grown to an OD_{600} of 0.5, washed and resuspended in M9 minimal media lacking methionine, supplemented with 0.2% fructose, and were grown for 45 min prior to the pulse-chase labeling.

Protease mapping with proteinase K

E. coli MK6 was transformed with pGZ-SciP, pGZ-SciP22, pGZ-SciP26, pGZ-SciP28, pMS-SciP20, and pMS-SciP21, respectively. As a control for the YidC depletion, a portion of cells were acid precipitated at the time of induction and immunoblotted using YidC antiserum. *E. coli* CM124 was transformed with pLZ1-SciP21 and *E. coli* IY26 with pMS-SciP21. The expression of the SciP variants was induced with 1 mM IPTG for 10 min, the cells were radiolabeled with ^{35}S methionine (15-30 µCi / ml culture) for 3 min, and (where indicated) non-radioactive methionine was added for 10 min. Spheroplasts were generated by centrifugation of the radiolabeled cells at 4 °C and by resuspension in ice-cold spheroplast buffer [33 mM Tris–HCl (pH 8) and 40% (wt/vol) sucrose]. Then, 1 mM EDTA and lysozyme (5 µg/ml) were added and the samples were kept on ice for 15 min. Spheroplasts were incubated for 1 h on ice either with proteinase K (0.5 mg/ml) or without. As a control for proteinase K activity, a portion of spheroplasts were lysed with 3% Triton-X-100 for 1 h on ice in the presence of proteinase K (0.5 mg/ml). The samples were acid precipitated, resuspended in 10 mM Tris/2% SDS, and immunoprecipitated with anti-His, anti-GroEL, or anti-OmpA antiserum. The samples were analyzed on a 14% SDS-PAGE and examined using a phosphorimager. Quantification was done with the program ImageJ [27].

AMS derivatization

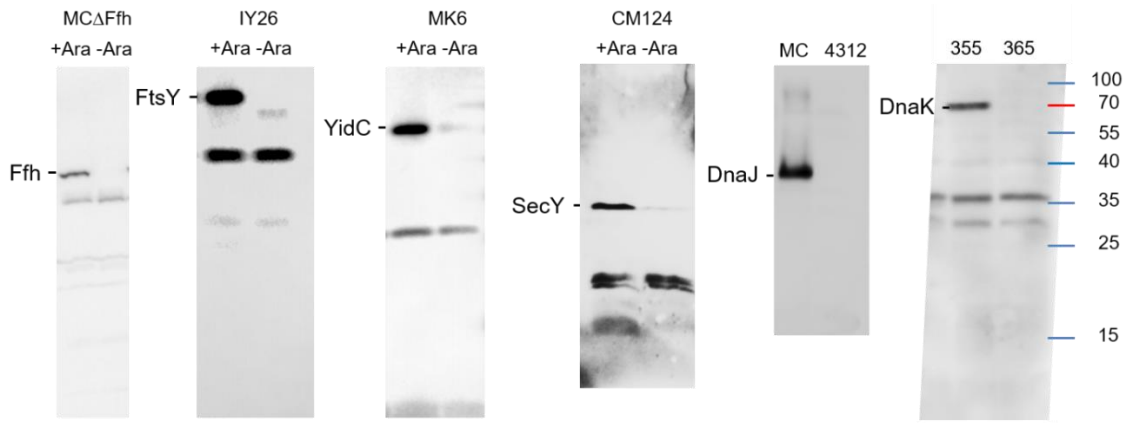
The *E. coli* strains MK6, MCΔFfh, IY26, CM124, MC4100, BB4312, GP355, and GP365 were transformed with the plasmid pMS-SciP-C or pLZ-SciP-C, respectively. As controls for the Ffh, FtsY, SecE, and YidC depletion and for the DnaJ and DnaK deletion, a portion of cells was acid precipitated at the time of induction and immunoblotted using the respective antibodies (Fig. S1). The expression of the SciP cysteine mutant was induced with 1 mM IPTG for 10 min; the cells were incubated with 2.5 mM AMS

(Molecular Probes) for 1 min and were radiolabelled with ^{35}S methionine (15-30 µCi / ml culture) for 2 min. Non-radioactive L-methionine was added for 10 min, and the AMS reaction was quenched by the addition of 10 mM DTT for a further 10 min. Samples were acid precipitated and resuspended in 10 mM Tris / 2% SDS prior to immunoprecipitation with an anti-His antibody. The samples were analyzed by SDS-PAGE and examined by phosporimaging. Quantification was done with the program ImageJ [27].

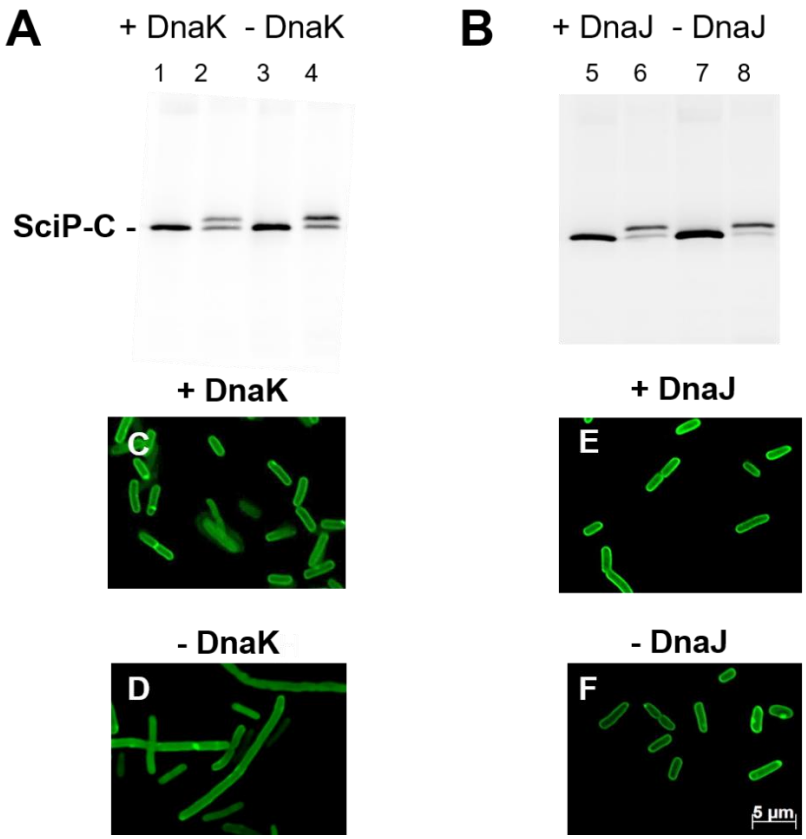
Fluorescence microscopy

For localization analysis, the SciP protein was fused at the N-terminal part with sfGFP and expressed from the plasmid pMS119EH or pLZ1 in the *E. coli* strains MK6, MCΔFfh, IY26, CM124, MC4100, BB4312, GP355, and GP365. As controls for the YidC, Ffh, FtsY, and SecE depletion and for the DnaJ and DnaK deletion, a portion of cells were acid precipitated at the time point of induction. The presence of all these components was analyzed after PAGE and immunoblotting using the respective antibodies (Fig. S1). The expression of the GFP-SciP fusion construct was induced with 1 mM IPTG at 30°C for the time indicated. Cells were harvested by centrifugation, washed twice with LB, resuspended in 2 mM EDTA and 50 mM Tris-HCl (pH 8), and were applied to poly-L-lysine-coated cover slides. The examination was performed with a filter set for GFP by a Zeiss AXIO Imager M1 fluorescence microscope. For the localization studies of the SciP deletion mutant and for the analysis of the membrane targeting of the two hydrophobic regions, *E. coli* MC4100 was transformed with pMS-GFP-SciPΔ12-20,62-71, pMS-SciP1-27-GFP-SciP96-217, or pMS-SciP54-85-GFP-SciP96-217 or with the plasmid pMS-GFP-SciP96-217 as a control. The expression of the SciP constructs was induced with 1 mM IPTG for 1 h at 30°C, and the further steps were carried out as described above.

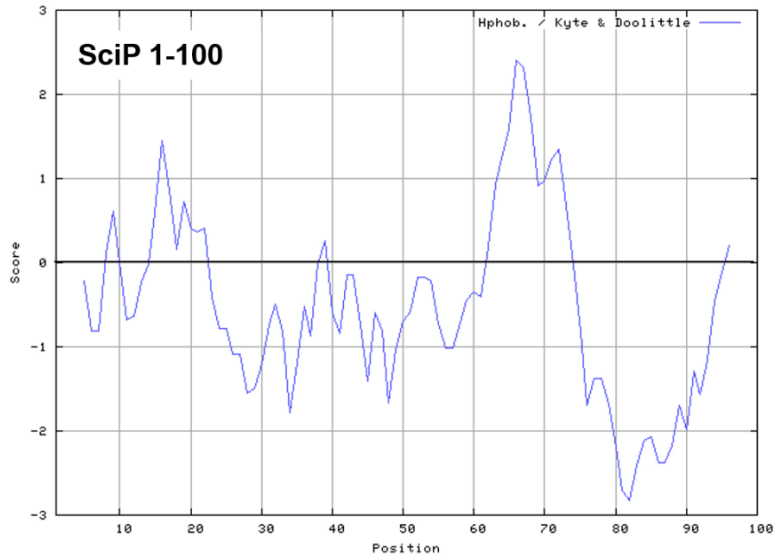
Supplementary Data



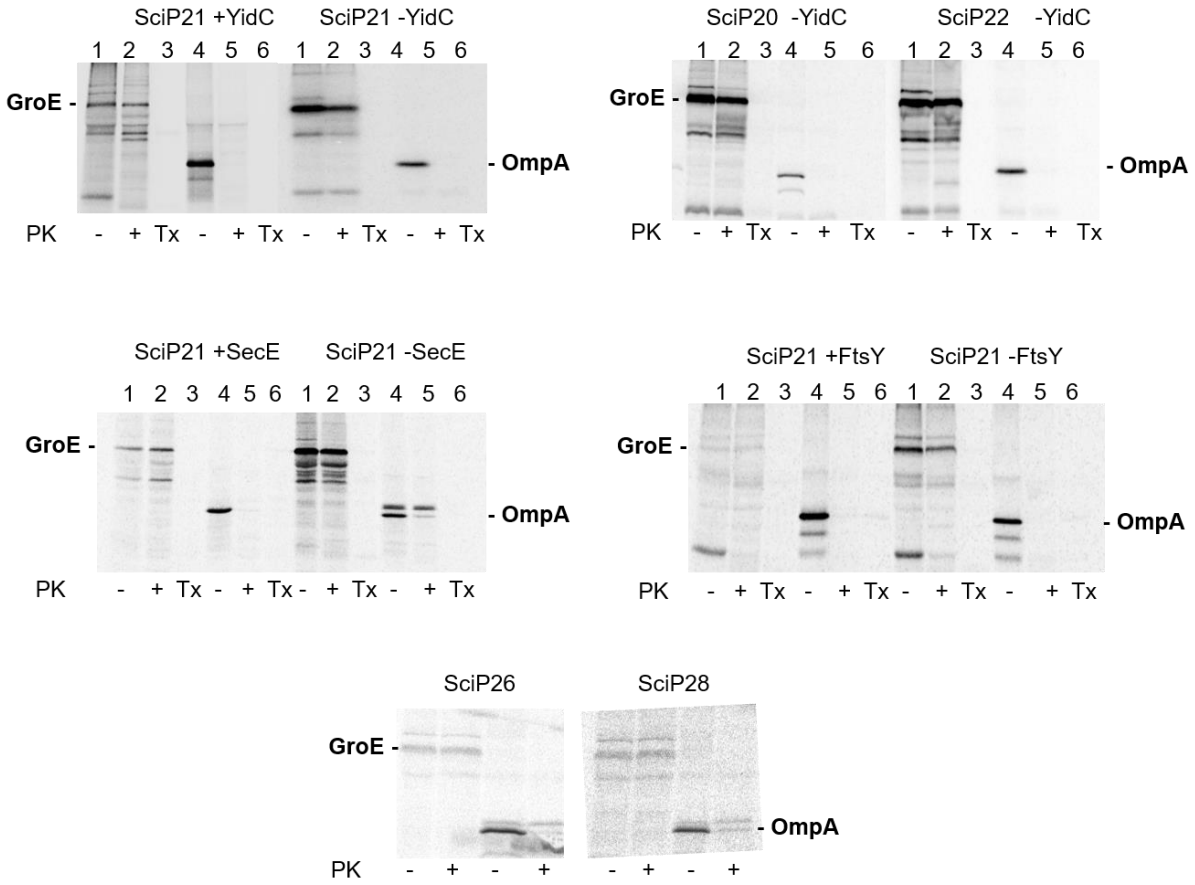
S1 Fig. Controls of the expression levels of Ffh, FtsY, YidC, SecY, DnaJ, and DnaK under non-depleted (left lanes) and depleted (right lanes) conditions. Samples of the cultures shown in Fig. 3, Fig. 4, Fig. 5, and S2 were processed for Western blotting with the respective antiserum.



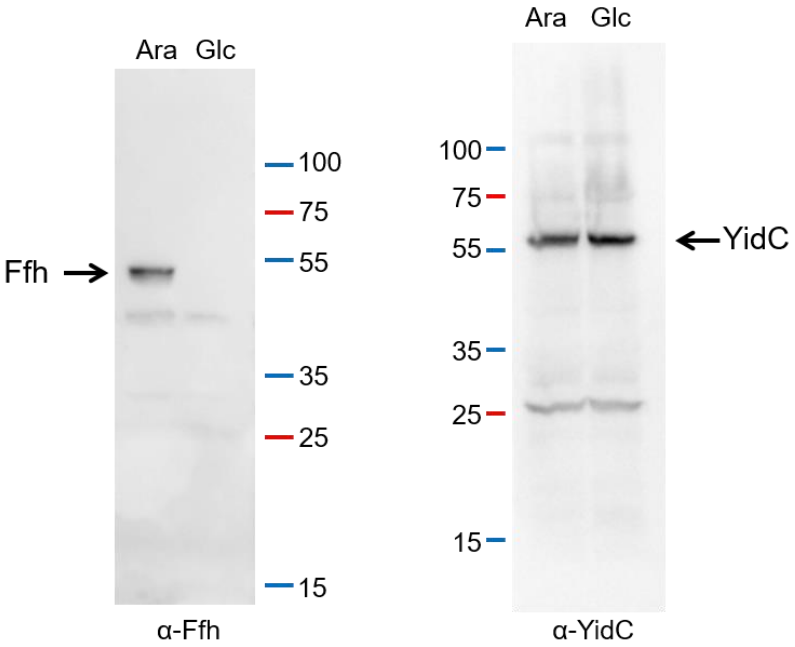
S2 Fig. Targeting and membrane insertion in the absence of DnaK and DnaJ. (A) Membrane insertion of SciP-C was analyzed in *E. coli* GP355 (lanes 1 and 2), GP365 (lanes 3 and 4), MC4100 (lanes 5 and 6), and BB4312 (lanes 7 and 8) by AMS modification as described in Fig. 3. (C–F) Membrane targeting of GFP-SciP was analyzed by fluorescence microscopy in *E. coli* (C) GP355, (D) GP365, (E) MC4100, and (F) BB4312.



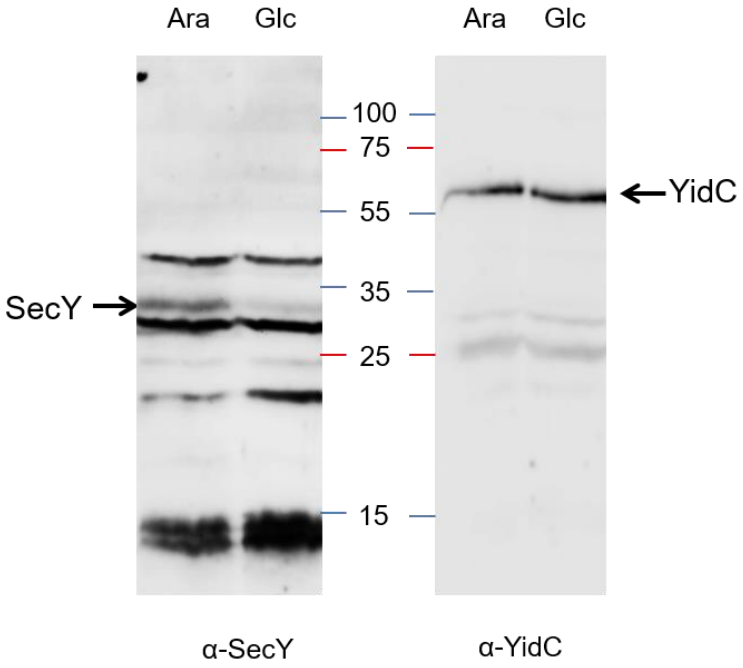
S3 Fig. Hydrophobicity plot according to Kyte and Doolittle [10] of the amino-terminal residues 1-100 of SciP. Two hydrophobic regions between residues 12 and 20 and between residues 62 and 71 are shown. They have the following sequence: 12-IFYPGWLMV-20 and 62-IMLYAFCALL-71.



S4 Fig. Controls of the proteinase mapping experiments shown in Fig. 4. The samples of the cultures shown in Fig. 4 were immunoprecipitated with antiserum to GroEL and OmpA, respectively.



S5 Fig. Analysis of cellular YidC levels in Ffh-depleted cells. The Ffh-depleted culture shown in Fig. 3C was analyzed by Western blotting with an antiserum to Ffh and YidC, respectively.



S6 Fig. Analysis of cellular SecY and YidC levels in SecE-depleted cells. The SecY-depleted culture shown in Fig. 3A was analyzed by Western blotting with an antiserum to SecY and YidC, respectively.

Acknowledgments

We thank Eric Cascales (Marseille) for providing the plasmid pIBA37-P containing *sciP,* Gunter Stier for providing the plasmid pET24d-sfGFP, and Bernd Bukau providing the deletion strains GP365 and BB4312.

This work was supported by the Deutsche Forschungsgemeinschaft (DFG grants KU749/5 and KU749/6). E.P. is holding a fellowship by the Landesgraduiertenförderung Baden-Württemberg.

Author contributions

I.S., D.L., and A.K. conceived and designed the experiments; E.P. and L.S. performed the experiments; E.P., D.L., and A.K. analyzed the data. A.K. wrote the paper.

The authors declare no conflict of interest.

Keywords

membrane insertase YidC; signal recognition particle (SRP); protein localization; protein targeting; fluorescence microscopy

Data availability statement

All relevant data are within the paper and its supporting information files.

Abbreviations used

TM, transmembrane; ER, endoplasmic reticulum; T6SS, type six secretion system; AMS, 4-acetoamido-4′-maleimidylstilbene-2,2′disodium sulfonate; SRP, signal recognition particle; GFP, green fluorescent protein; sfGFP, superfolder GFP.

References

[1] L. Bischoff, S. Wickles, O. Berninghausen, E.O. Vandersluis, R. Beckmann, Visualization of a polytopic membrane protein during SecY-mediated membrane insertion, Nat. Commun. 5 (2014) 4103.

[2] F. Wang, C. Chan, N.R. Weir, V. Denic, The Get1/2 transmembrane complex is an endoplasmic-reticulum membrane protein insertase, Nature 512 (2014) 441–444.

[3] M.S. Aschtgen, A. Zoued, R. Lloubes, L. Journet, E. Cascales, The C-tail anchored TssL subunit, an essential component of the enteroaggregative *Escherichia coli* Sci-1 type VI secretion system, is inserted by YidC, Microbiol. Open 1 (2012) 71–82.

[4] F. Boyer, G. Fichant, J. Berthod, Y. Vandenbrouck, I. Attree, Dissecting the bacterial type VI secretion system by a genome wide *in silico* analysis: what can be learned from available microbial genomic resources? BMC Genomics 10 (2009) 104.
[5] M.M. Shneider, S.A. Buth, B.T. Ho, M. Basler, J.J. Mekalanos, P.G. Leiman, PAAR-repeat proteins sharpen and diversify the type VI secretion system spike, Nature 500 (2013) 350–353.
[6] E. Durand, A. Zoued, S. Spinelli, P.J. Watson, M.S. Aschtgen, L. Journet, C. Cambillau, E. Cascales, Structural characterization and oligomerization of the TssL protein, a component shared by bacterial type VI and type IVb secretion systems, J. Biol. Chem. 287 (2012) 14,157–14,168.
[7] A. Zoued, E. Durand, Y.R. Brunet, S. Spinelli, B. Douzi, M. Guzzo, N. Flaugnatti, P. Legrand, L. Journet, R. Fronzes, T. Mignot, C. Cambillau, E. Cascales, Priming and polymerization of a bacterial contractile tail structure, Nature 531 (2016) 59–63.
[8] T. Taura, T. Baba, Y. Akiyama, K. Ito, Determinants of the quantity of stable SecY complex in the *Escherichia coli* cell, J. Bacteriol. 175 (1993) 7771–7775.
[9] N. Celebi, L. Yi, S.J. Facey, A. Kuhn, R.E. Dalbey, Membrane biogenesis of subunit II of cytochrome bo oxidase: contrasting requirements for insertion of N-terminal and C-terminal domains, J. Mol. Biol. 357 (2006) 1428–1436.
[10] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, J. Mol. Biol. 157 (1982) 105–132.
[11] K.S. Maier, S. Hubich, H. Liebhart, S. Krauss, A. Kuhn, S.J. Facey, An amphiphilic region in the cytoplasmic domain of KdpD is recognized by SRP and targeted to the *Escherichia coli* membrane, Mol. Microbiol. 68 (2008) 1471–1484.
[12] K. Kumazaki, T. Kishimoto, A. Furukawa, H. Mori, Y. Tanaka, N. Dohmae, R. Ishitani, T. Tsukazaki, O. Nureki, Crystal structure of *Escherichia coli* YidC, a membrane protein chaperone and insertase, Sci. Rep. 4 (2014) 7299.
[13] C. Klenner, J. Yuan, R.E. Dalbey, A. Kuhn, The Pf3 coat protein contacts TM1 and TM3 of YidC during membrane biogenesis, FEBS Lett. 582 (2008) 3967–3972.
[14] S.A. Neugebauer, A. Baulig, A. Kuhn, S.J. Facey, Membrane protein insertion of variant MscL proteins occurs at YidC and SecYEG of *Escherichia coli*, J. Mol. Biol. 416 (2012) 375–386.
[15] A. Kuhn, Alterations in the extracellular domain of M13 procoat protein make its membrane insertion dependent on *secA* and *secY*, Eur. J. Biochem. 177 (1988) 267–271.

[16] T. Roos, D. Kiefer, S. Hugenschmidt, A. Economou, A. Kuhn, Indecisive M13 procoat protein mutants bind to SecA but do not activate the translocation ATPase, J. Biol. Chem. 276 (2001) 37,909–37,915.

[17] M. Chen, K. Xie, J. Yuan, L. Yi, S.J. Facey, N. Pradel, L.F. Wu, A. Kuhn, R.E. Dalbey, Involvement of SecDF and YidC in the membrane insertion of M13 procoat mutants, Biochemistry 44 (2005) 10,741–10,749.

[18] R. Soman, J. Yuan, A. Kuhn, E.R. Dalbey, Polarity and charge of the periplasmic loop determine the YidC and Sec translocase requirement for the M13 procoat Lep protein, J. Biol. Chem. 289 (2014) 1023–1032.

[19] I. Sachelaru, N.A. Petriman, R. Kudva, P. Kuhn, T. Welte, B. Knapp, F. Drepper, B. Warscheid, H.G. Koch, YidC occupies the lateral gate of the SecYEG translocon and is sequentially displaced by a nascent membrane protein, J. Biol. Chem. 288 (2013) 16,295–16,307.

[20] M. Lessl, D. Balzer, R. Lurz, V.L. Waters, D.G. Guiney, E. Lanka, Dissection of IncP conjugative plasmid transfer: definition of the transfer region Tra2 by mobilization of the Tra1 region *in trans*, J. Bacteriol. 174 (1992) 2493–2500.

[21] D. Balzer, G. Ziegelin, W. Pansegrau, V. Kruft, E. Lanka, KorB protein of promiscuous plasmid RP4 recognizes inverted sequence repetitions in regions essential for conjugative plasmid transfer, Nucleic Acids Res. 20 (1992) 1851–1858.

[22] L. Zhu, C. Klenner, A. Kuhn, R.E. Dalbey, Both YidC and SecYEG are required for translocation of the periplasmic loops 1 and 2 of the multispanning membrane protein TatC, J. Mol. Biol. 424 (2012) 354–367.

[23] J.D. Pédelacq, C. Cabantous, T. Tran, T.C. Terwilliger, G.S. Waldo, Engineering and characterization of a superfolder green fluorescent protein, Nat. Biotechnol. 24 (2006) 79–88.

[24] I. Seitl, S. Wickles, R. Beckmann, A. Kuhn, D. Kiefer, The C-terminal regions of YidC from *Rhodopirellula baltica* and *Oceanicaulis alexandrii* bind to ribosomes and partially substitute for SRP receptor function in *Escherichia coli*, Mol. Microbiol. 91 (2014) 408–421.

[25] T. Traxler, C. Murphy, Insertion of the polytopic membrane protein MalF is dependent on the bacterial secretion machinery, J. Biol. Chem. 271 (1996) 12,394–12,400.

[26] K.F. Jensen, The *Escherichia coli* K-12 “wild types” W3110 and MG1655 have a rph frameshift mutation that leads to pyrimidine starvation due to low *pyrE* expression levels, J. Bacteriol. 175 (1993) 3401–3407.
[27] C.A. Schneider, W.S. Rasband, K.W. Eliceiri, NIH image to ImageJ: 25 years of image analysis, Nat. Methods 9 (2012) 671–675.

4

The two SRP signal sequences within the amino-terminal domain of the C-tailed protein SciP

Introduction

In chapter 3 it was shown that the C-tail anchored protein SciP from enteroaggregative *E. coli* is targeted by the universally conserved SRP system, involving the cytoplasmic ribonucleoprotein SRP and the membrane-associated SRP receptor FtsY. After targeting to the membrane, the C-terminal TMD is inserted into the bilayer by the YidC insertase (chapter 3). Due to the special topology of SciP, with a TMD at the extreme C-terminal part of the protein, the insertion of the hydrophobic TMD has to occur in a post-translational way. Nevertheless, there are indications that the SRP-dependent targeting occurs co-translationally via two short hydrophobic regions in the N-terminal cytoplasmic part of the protein (Fig. 21). The fact that SRP signal sequences are not limited to TMDs was also shown for the potassium sensor protein KdpD (Maier et al. 2008). A short amphiphilic region located between amino acids 22-48 of KdpD was shown to function as the SRP signal sequence for an early co-translational targeting. It is reasonable that a similar mechanism enables co-translational targeting of the C-tail anchored protein SciP, due to the special topology. Clear indications for a co-translational targeting arise from the results described in chapter 3. Deletion of the hydrophobic regions in a sfGFP-SciP fusion construct abolished the targeting of SciP to the membrane. In addition, it was shown that the amino acids 1-27 and 54-85 of SciP including the short hydrophobic regions (amino acids 12-20 and 62-71) are both able to target a sfGFP fusion construct to the membrane, indicating a targeting function of both regions. However, no SRP substrate with more than one SRP signal sequence is described in the literature to date.

The question, whether both regions actually serve as the SRP signal sequence and interact with SRP was addressed in this chapter to gain more precise information about the targeting mechanism of SciP. Cysteine-accessibility assays show that the translocation of the C-tail is only impaired if both hydrophobic regions are deleted. The deletion of only one of these regions has no influence on the insertion of the protein. *In vitro* interaction studies of purified SRP and TnaC stalled ribosomes using microscale thermophoresis demonstrate that both hydrophobic regions of SciP are bound with similar affinity by SRP or by a preincubated SRP-FtsY complex. In addition, cross-linking studies with *in vitro* synthesized SciP peptides reveal that they are bound by the SRP M domain, similar to other known SRP substates.

Results

Deletion of the hydrophobic regions impairs the translocation of the C-terminal tail.

Fluorescence microscopy studies with sfGFP fusion constructs suggested that the hydrophobic regions between amino acids 12-20 and 62-71 of SciP are involved in the co-translational SRP-dependent targeting of the protein. However, fluorescence microscopy studies only allow a statement about the localisation of the respective protein in the cell. If the deletion of these regions also impairs the insertion and the translocation of the C-tail was analyzed with cysteine-accessibility assays. Therefore, the SciP-C mutant described in chapter 3, with an additional cysteine residue at the C-terminus in the periplasmic domain (218C) was used. Either one or both hydrophobic regions of SciP-C were deleted by site-directed mutagenesis. The C-tail translocation of the different mutants was monitored in an *E. coli* wildtype strain (MC4100) via cysteine modification with 4-acetoamido-4′-maleimidylstilbene-2,2′disodium sulfonate (AMS). After induction with 1 mM IPTG for 10 minutes, 2.5 mM AMS was added 1 minute prior to pulse labeling with radioactive methionine. Translocation of the periplasmic part leads to the accessibility of the cysteine residue for AMS and a mobility shift of about 0.5 kDa on a SDS-PAGE. As a direct control, the insertion of SciP-C without the deletions was analyzed in MC4100 showing that about 43% of the protein was modified and shifted on the SDS-PAGE by AMS (Fig. 20, lane 1). Deletion of the first or the second hydrophobic region (aa 12-20 or 62-71) had no effect on the translocation of the C-tail since the same amount was modified with AMS (lane 2 and 3). In contrast, the combined deletion of both regions resulted in a decrease to 28% (lane 4).

This result strengthens the hypothesis that the protein contains two targeting signals, however, *in vivo* one region is sufficient for effective SRP-dependent targeting since the targeting was only impaired if both regions were deleted.

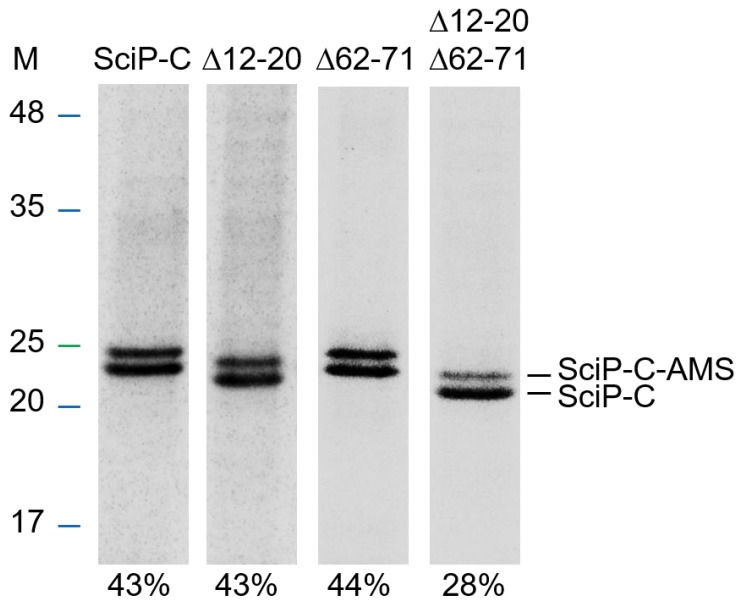
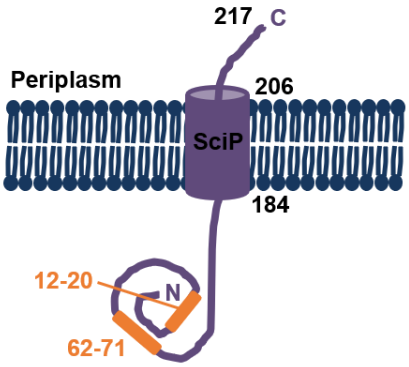


Figure 20: Cysteine accessibility assay of SciP-C, SciP-CΔ12-20 (Δ12-20), SciP-CΔ62-71 (Δ62-71) and SciP-CΔ12-20, Δ62-71 (Δ12-20 Δ62-71) with AMS. The C-tail of the SciP variants was extended by a cysteine residue at position 218 and expressed in MC4100 cells. The expression was induced with 1 mM IPTG for 10 min, 2.5 mM AMS was added 1 min prior to pulse labeling with ^{35}S-methionine / cysteine (15-30 µCi / ml culture) for 2 min. The cells were chased for 10 min with non-labeled methionine / cysteine, immunoprecipitated against the N-terminal His_{10}-Tag and analyzed on a 14% SDS-PAGE. The modification of the cysteine residue at position 218 leads to a mobility shift due to its exposure in the periplasm. The relative amount of the AMS-modified protein is indicated.

Alterations in the hydrophobicity of the second hydrophobic region of SciP affect SRP-dependent targeting

Crucial for an SRP binding to a signal sequence is its hydrophobicity which has to exceed a threshold level. In chapter 2 it was confirmed that SRP binding to the signal sequence of KdpD was also sensitive to alterations in the hydrophobicity. A decrease in the hydrophobicity by the substitution of the cysteine residue at position 32 of KdpD with a serine residue resulted in targeting defects and a weaker SRP binding to the KdpD signal sequence. Interestingly, the second hydrophobic region of SciP also contains a cysteine residue at position 68 located in the hydrophobic core region (Fig. 21). To test whether a substitution of the cysteine residue with less hydrophobic residues impairs the targeting ability, subcellular localisation studies with fluorescence microscopy were performed. Therefore, the SciP54-85sfGFP96-217 fusion construct (named SciP54-85) described in chapter 3 was used. With site-directed mutagenesis the cysteine at position 68 was substituted with a serine, resulting in a decrease of the hydrophobicity from 2.32 to 1.96 GRAVY (Fig. 21). For the determination of the grand

average of hydropathy (GRAVY) index of the amino acid sequences the calculation according to Kyte and Doolittle was used (Kyte and Doolittle 1982).



		Hydrophobicity (GRAVY)
	1 10 20 27	
SciP-1-27	M N K P V I S R A E Q **I F Y P G W L M V** S Q L R S G Q	1.44
	54 60 70 80 85	
SciP-54-85	G F S Q K S S D **I M L Y A F** [C] **A L L** D E S V L N R E K T D D G W	2.32
	54 60 70 80 85	
SciP-54-85 C68S	G F S Q K S S D **I M L Y A F** [S] **A L L** D E S V L N R E K T D D G W	1.96
	54 60 70 80 85	
SciP-54-85 C68M	G F S Q K S S D **I M L Y A F** [M] **A L L** D E S V L N R E K T D D G W	2.26

Figure 21: Amino acid sequence of the cytoplasmic hydrophobic regions of SciP and two mutant SciP proteins. SciP has a transmembrane domain between amino acid 184 to 206 resulting in a large cytoplasmic N-terminal domain. This region includes two hydrophobic parts between amino acid 12-20 and 62-71. The amino acid sequence of the two hydrophobic regions and their respective GRAVY are listed as well as the two derivatives used for fluorescence microscopy.

For the localisation studies with fluorescence microscopy *E. coli* MC4100 cells were used. The protein expression was induced with 1 mM IPTG and the localisation of the mutant was analyzed after 1h at 30°C with a specific filter set for GFP. In contrast to SciP54-85 (Fig. 22A) the fluorescence signal of the C68S mutant was no longer located at the membrane, but appeared as distinct spots indicating aggregation of the mutant protein (Fig. 22B). A substitution of the cysteine with a methionine residue which reduces the hydrophobicity only to 2.26 GRAVY (Fig. 21) resulted in a mixed phenotype with a signal at the membrane and additional spots mainly at the cell poles (Fig. 22C). It seems that the substitution with a methionine residue as well impairs the targeting, however, the effects are less pronounced than observed in the C68S mutant. From that it is concluded that the hydrophobicity of the second region is a crucial feature for SRP-dependent targeting of the SciP54-85-sfGFP fusion construct.

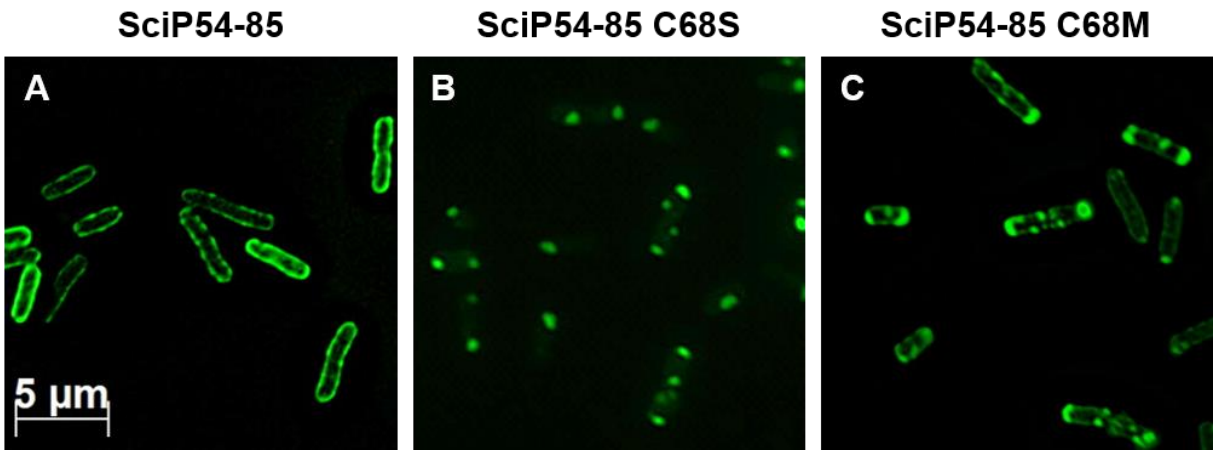


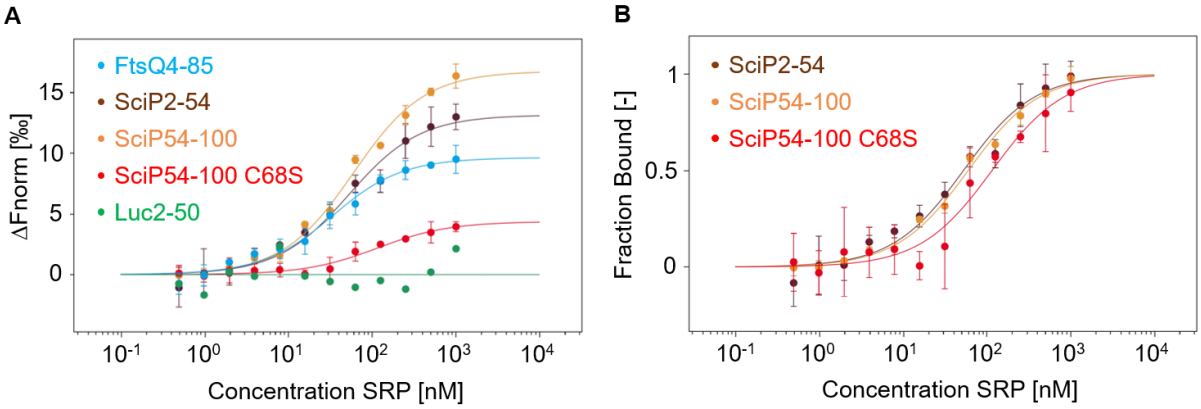
Figure 22: Fluorescence microscopy of variant SciP-sfGFP fusion proteins. MC4100 cells bearing the respective plasmids were induced with 1 mM IPTG and applied for the fluorescence microscopy after 1h at 30°C. (A) For SciP54-85 the fluorescent signal was clearly found at the membrane, indicating a targeting function of this region. (B) Substitution of the cysteine residue at position 68 with a serine residue resulted in distinct spots probably due to the aggregation of the protein. (C) Substitution of the cysteine with a methionine residue resulted in a mixed phenotype with signals found at the membrane and as distinct spots mainly at the cell poles**.**

Binding of the hydrophobic regions to SRP

To investigate if both hydrophobic regions of SciP can bind SRP, microscale thermophoresis with TnaC stalled ribosomes and purified SRP was carried out to determine the dissociation constant (Kd). The studies with RNCs exposing the SRP signal sequence of KdpD, described in chapter 2 served as a basis for the experiments with SciP.

Artificially stalled ribosomes were generated by introducing the sequence encoding for amino acids 2-54 or 54-100 of SciP to the TnaC stalling sequence. The fused sequences were extended due to the fact that about 30 amino acids are buried in the ribosomal exit tunnel resulting in a guaranteed exposure of the hydrophobic regions. For purification of the different RNCs the fused sequences encode for an N-terminal His-Tag and were further purified on a sucrose gradient. For the microscale thermophoresis measurements one interaction partner has to be labeled with a fluorescent dye whereas the other interaction partner is unlabeled and added in serial dilutions. Therefore, the different RNCs were labeled with the cysteine reactive dye NT-647 (NanoTemper Technologies). 10 nM of labeled RNCs were mixed 1:1 with the purified and reconstituted SRP in serial dilutions, resulting in a final SRP concentration from 1 µM to 0.5 nM and a fixed RNC concentration of 5 nM. To discriminate between a correct and incorrect cargo for SRP a positive and a negative control was used. As

described in chapter 2, RNCs with amino acids 4-85 of the well-studied SRP substrate FtsQ were used as a positive control resulting in a Kd of 30.6 ± 6 nM (Fig. 23). In contrast, RNCs with amino acids 2-50 of the cytoplasmic protein luciferase served as a negative control which showed no binding under these conditions (Fig. 23). First, RNCs with amino acids 2-54 of SciP were incubated with SRP resulting in a clear binding event (Fig. 23A, brown dots). The determined Kd value of 49.7 ± 12 nM is close to the Kd of the positive control with RNCs exposing amino acids 4-85 of FtsQ (blue dots) (Fig. 23C). Since in this assay, ribosomes exposing amino acids 2-50 of the cytoplasmic protein luciferase are not bound by SRP (green dots), it can be concluded that the amino acids 2-54 of SciP represent a correct target for SRP.



C

RNCs	Kd_{SRP} [nM]
FtsQ4-85	30.6 ± 6
SciP2-54	49.7 ± 12
SciP54-100	57 ± 7.6
SciP54-100 C68S	112.1 ± 33.2
Luc2-50	no binding

Figure 23: MST measurements of RNCs with SRP. 5 nM of labeled RNCs were incubated with unlabeled SRP (1 µM to 0.49 nM) for 5 min on ice and filled into Premium capillaries (NanoTemper Technologies). (A) The change in normalized fluorescence (ΔFnorm in ‰) of FtsQ4-85 (blue dots), SciP2-54 (brown dots), SciP54-100 (orange dots), SciP54-100 C68S (red dots) and Luc2-50 (green dots) is plotted against the concentration of unlabeled SRP. (B) The fraction bound of SciP2-54 (brown dots), SciP54-100 (orange dots) and SciP54-100 C68S (red dots) is plotted against the concentration of unlabeled SRP. Three independently pipetted measurements are merged and the error bars represent the standard deviation. (C) Summary of the Kd values.

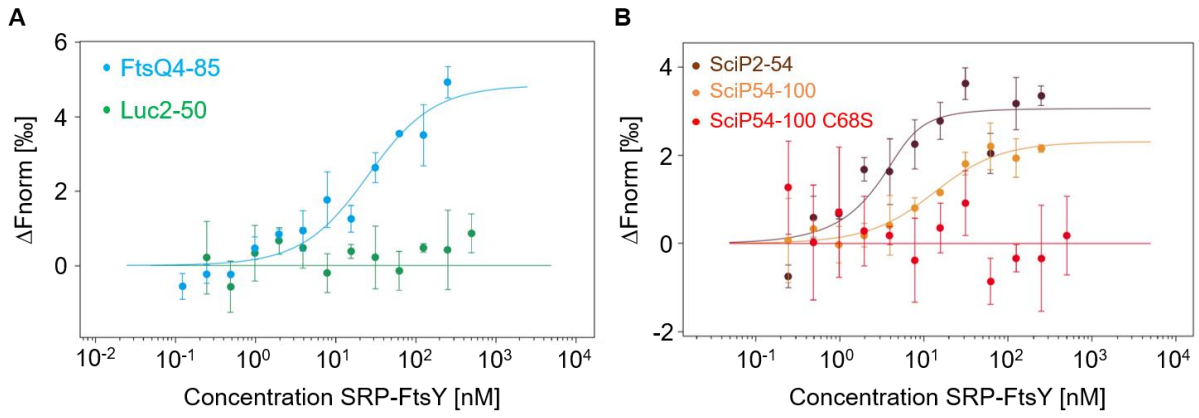
Next, RNCs encompassing amino acids 54-100 of SciP, including the second hydrophobic region were incubated with SRP. Also in this case, a clear binding event with a Kd of 57 ± 7.6 nM was determined (Fig. 23A, orange dots) which also correlates with the positive control FtsQ. By blotting the fraction bound against the concentration of the unlabeled SRP the curves of RNCs with the first and the second hydrophobic region of SciP, respectively, match almost perfectly (Fig. 23B). From that it is concluded that both regions separately are able to be recognized and bound by SRP and therefore both could serve as the SRP signal sequence.

As shown in figure 22, the substitution of the cysteine at position 68 of SciP with a serine residue, impaired the targeting ability of the second hydrophobic region. Therefore, it would be interesting to see whether the altered hydrophobic region is also impaired in its ability to be bound by SRP. With site-directed mutagenesis the C68S mutation was incorporated into the sequence encoding for stalled ribosomes with amino acids 54-100 of SciP. In this case, binding of SRP was also detected, although with a reduced affinity of 112.1 ± 33.2 nM (Fig. 23, red dots) compared with the non-mutated RNCs with a Kd of 57 ± 7.6 nM (Fig. 23, orange dots). This clearly shows that the mutation of the cysteine into a serine residue reduces the binding affinity of SRP to the second hydrophobic region of SciP.

Binding of the hydrophobic regions to a preincubated SRP-FtsY complex

A crucial step in the SRP-dependent targeting is the interaction of SRP with its receptor FtsY. Several conformational changes in both proteins occur during the interaction of the two GTPases and are required for the completion of the SRP cycle (Shan and Walter 2003; Shan et al. 2004). This step represents another checkpoint for the discrimination between correct and incorrect SRP substrates. A stable RNC-SRP-FtsY complex can only arise if SRP is bound to a correct cargo (Saraogi et al. 2014). With microscale thermophoresis it was attempted to clarify if the different stalled ribosomes with nascent chains of SciP are binding to a preincubated SRP-FtsY complex. This could further support the findings obtained by the binding studies with SRP only. To generate a stable SRP-FtsY complex, 200 µM of the non-hydrolysable GTP analogue GppNHp was used. 10 nM of the labeled RNCs were incubated 1:1 with serial dilutions of the GppNHp-SRP-FtsY complex resulting in a concentration ranging from 500/250 nM to 0.24/0.12 nM and a fixed concentration of 5 nM for the labeled RNCs. Both, RNCs with amino acids 2-54 or 54-100 of SciP show a clear binding curve with a Kd of 0.8 ± 1.2 nM and 10.8 ± 3.6 nM, respectively (Fig. 24, brown and orange dots). The

determined Kd values are close to the positive control FtsQ with a Kd of 22.9 ± 10.8 nM (Fig. 24, blue dots). However, SRP seems to have a slightly higher affinity for the first hydrophobic region (Fig. 24B, brown dots). The negative control, RNCs with amino acids 2-50 of luciferase, was not bound by the SRP-FtsY complex (Fig. 24A, green dots), which shows again that both hydrophobic regions of SciP represent a correct target for SRP. Interestingly, the mutation of the cysteine residue at position 68 into a serine residue in the second hydrophobic region of SciP prevented SRP-FtsY binding to the nascent chain (Fig. 24B, red dots). Taken together, it is concluded that both hydrophobic regions of SciP can serve as the SRP signal sequence during co-translational targeting. In addition, the substitution of the cysteine residue at position 68 with a serine residue destroys the second SRP signal sequence.



C

RNCs	$Kd_{SRP-FtsY}$ [nM]
FtsQ4-85	22.9 ± 10.8
SciP2-54	0.8 ± 1.2
SciP54-100	10.8 ± 3.6
SciP54-100 C68S	no binding
Luc2-50	no binding

Figure 24: MST measurements of RNCs with a preincubated SRP-FtsY complex. 5 nM of labeled RNCs were incubated with unlabeled SRP-FtsY (500/250 nM to 0.24/0.12 nM) for 5 min on ice and filled into Premium capillaries (NanoTemper Technologies). (A) The change in normalized fluorescence (ΔFnorm in ‰) of FtsQ4-85 (blue dots) and Luc2-50 (green dots) is plotted against the concentration of unlabeled SRP-FtsY. (B) The change in normalized fluorescence (ΔFnorm in ‰) of SciP2-54 (brown dots), SciP54-100 (orange dots) and SciP54-100 C68S (red dots) is plotted against the concentration of unlabeled SRP-FtsY. Three independently pipetted measurements are merged and the error bars represent the standard deviation. (C) Summary of the Kd values.

Binding of SciP1-27 and SciP54-85 to the SRP M domain

SRP signal sequences are bound after their exposure at the ribosomal exit tunnel by the hydrophobic groove of the SRP M domain (Keenan et al. 1998). With microscale thermophoresis and artificially stalled ribosomes it was shown that both regions are binding to SRP when exposed outside of the ribosomal tunnel. To investigate if the hydrophobic regions of SciP are binding within the hydrophobic groove of the SRP M domain, *in vitro* disulphide cross-linking studies with copper phenanthroline were done. The idea was to use synthesized peptides of SciP1-27 and SciP54-85 containing a cysteine residue for the cross-linking studies with single cysteine mutants of SRP. The SciP peptide 54-85 already contains a cysteine residue at position 68 which can be used for disulphide-crosslinking experiments (Fig. 25A). For SciP1-27 a cysteine residue was incorporated into the sequence at position 16, since it does not contain an endogenous cysteine residue (Fig. 25A). Position 16 is located in the hydrophobic core region and displays the same orientation as the cysteine residue at position 68 of the second peptide (Fig. 25B). The substitution of the glycine with a cysteine residue results in an increase of the hydrophobicity of the first hydrophobic region (amino acid 12-20) from 1.44 to 1.76 GRAVY. To exclude that the substitution has an effect on the targeting ability of this sequence, the mutation was also incorporated into the plasmid encoding for SciP1-27GFP96-217 (named SciP1-27) described in chapter 3. The subcellular localisation of this fusion construct, named SciP1-27 G16C, was analysed with fluorescence microscopy. The expression of SciP1-27 G16C was induced with 1 mM IPTG for 1h at 30°C in *E. coli* MC4100 cells. The fluorescence signal was located at the membrane, indicating that the mutation of G16C has no influence on membrane targeting and therefore the peptide can be used for disulphide cross-linking studies with SRP (Fig. 25C). The SciP peptide 1-27 with the cysteine at position 16 and the SciP peptide 54-85 with the cysteine at position 68 (Fig. 25A) were synthesized *in vitro* by the custom peptide synthesis services from GENOSPHERE Biotechnologies.

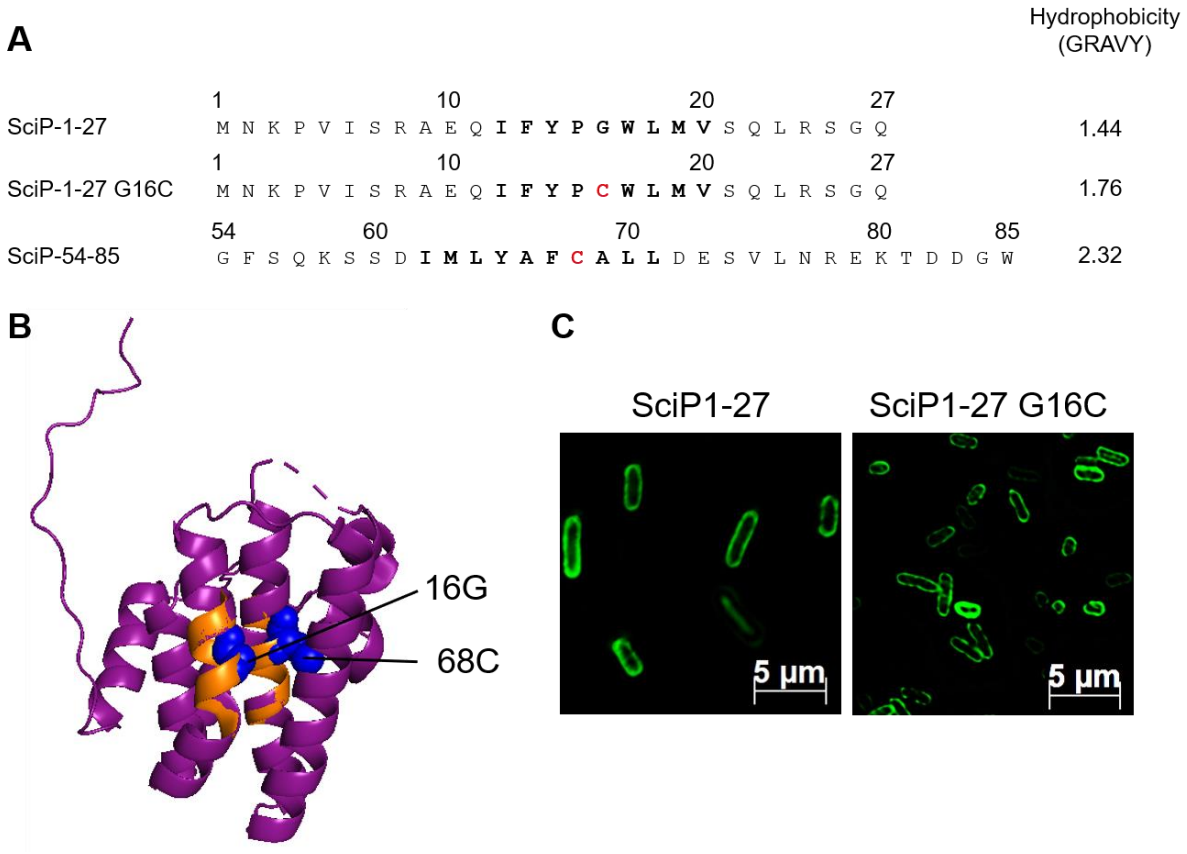


Figure 25: SciP peptides for the disulphide cross-linking studies. (A) Amino acid sequence of SciP1-27 and SciP54-85 including the hydrophobic regions from amino acid 12-20 and 62-71 (bold letters). For disulphide cross-linking studies the glycine at position 16 of SciP1-27 was mutated into cysteine (coloured red). The GRAVY of the peptides is listed. (B) Crystal structure of the cytoplasmic domain of SciP represented as cartoon. The hydrophobic regions (12-20 and 62-71) are coloured orange. The glycine residue at postion 16 and the cysteine residue at position 68 are highlighted as blue spheres (Image created with PyMOL 2.3.0, PDB 3U66). (C) Subcellular localisation studies of SciP1-27GFP96-217 described in chapter 3 (named SciP1-27) and the mutant SciP1-27GFP96-217 G16C (named SciP1-27 G16C) with fluorescence microscopy.

To analyze whether the SciP peptides are bound by the SRP M domain the single cysteine mutant SRP 406S/423C described in chapter 2 was used. This mutant has a cysteine residue in the hydrophobic groove formed by the SRP M domain (Fig. 26B) and is involved in signal sequence binding. As a negative control the SRP single cysteine mutant 181C/406S, also described in chapter 2 was used, which has a cysteine residue in the NG domain (Fig. 26B) that is not involved in signal sequence binding. In both mutants the endogenous cysteine residue at position 406 was substituted with a serine residue.

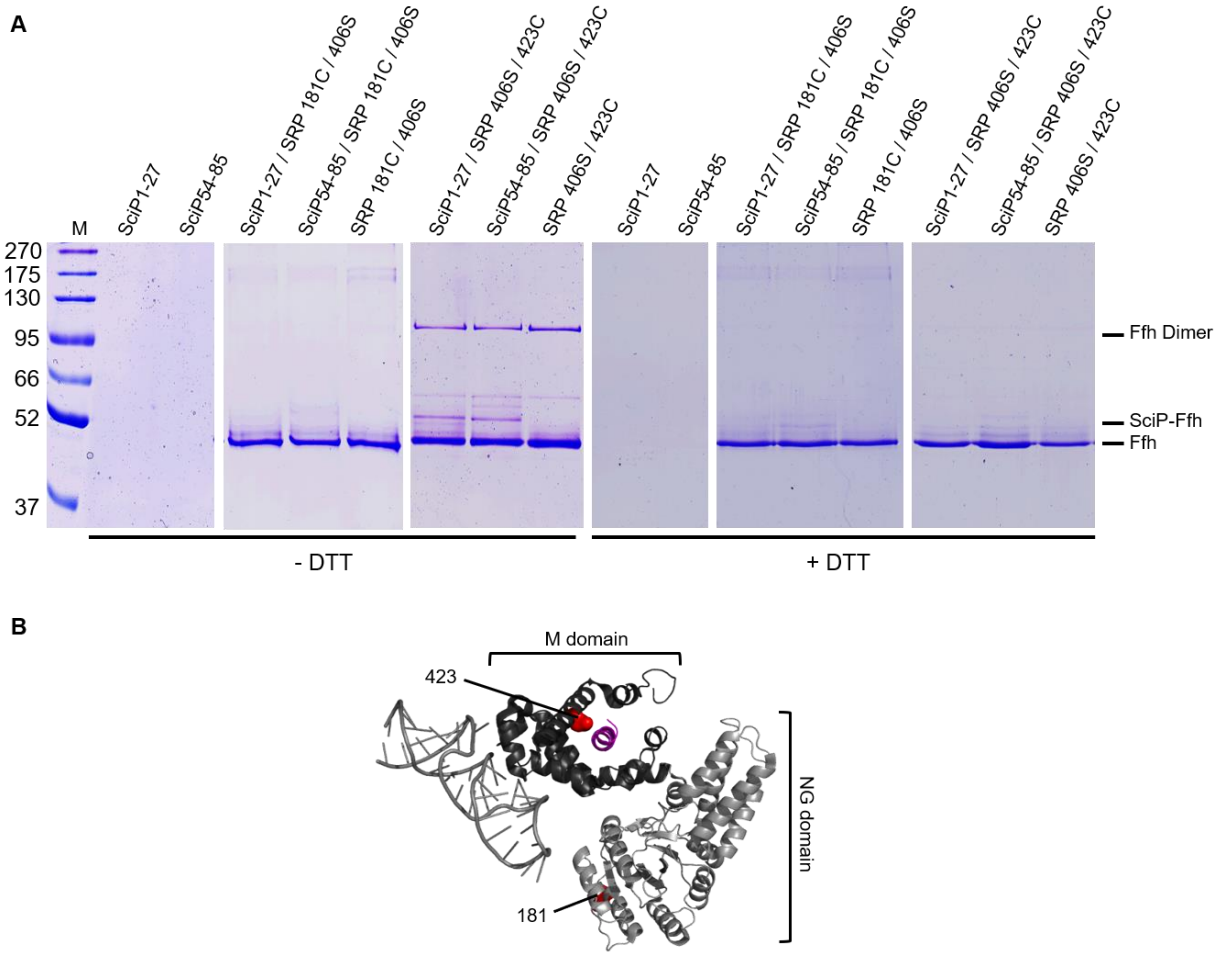


Figure 26: *In vitro* disulphide cross-linking of SciP1-27 / SciP54-85 and SRP. (A) 2 µM of reconstituted SRP L181C/406S or SRP 406S/423C was incubated with 20 µM of the SciP1-27 or SciP54-85 peptide and 1 mM copper phenanthroline for 1h on ice. After TCA precipitation the samples were resuspended in buffer with or without DTT and loaded on a 10% SDS-PAGE. A cross-linking band (SciP-Ffh) appeared when the peptides were incubated with SRP 406S/423C. The shifted bands were sensitive to the reducing agent DTT. In contrast, no cross-linking band can be detected with SRP 181C/406S. The SRP mutants and the SciP peptides alone were analyzed for control. (B) Crystal structure of the *E. coli* SRP bound to a signal sequence, displayed as a cartoon. The NG domain and the 4.5S RNA are colored in light grey, the M domain in dark grey and the signal sequence in purple. The positions of the incorporated cysteine residues are displayed as red dots (Image created with PyMOL 2.3.0, PDB 5GAD).

For cross-linking studies 2 µM of reconstituted SRP was incubated with 20 µM of the SciP peptides and 1 mM copper phenanthroline was added. After 1h on ice, the samples were TCA precipitated, resuspended in buffer with or without DTT and loaded on a 10% SDS-PAGE. When the two peptides were incubated with SRP 406S/423C an additional band appeared at about 52 kDa (Fig. 26, lane 7 and 8). The two peptides or the SRP mutants alone did not show this additional band (Fig. 26, lanes 2, 3, 6 and 9). The addition of DTT to these samples resulted in the disappearance of the additional band (Fig. 26, lane 15 and 16). This shows that both SciP peptides can be

cross-linked to the cysteine residue at position 423, indicating that they are both bound by the hydrophobic groove of the SRP M domain. Due to the small size of the SciP peptides with only 3 kDa, they cannot be detected on the 10% SDS-PAGE. When the peptides were incubated with SRP having a cysteine residue in the NG domain at position 181, no shift appeared (Fig. 26, lane 4 and 5), indicating that both peptides do not interact with the NG domain. From that it is concluded, that also the cytoplasmic SRP signal sequences of SciP are bound by the hydrophobic groove of SRP.

Discussion

In general, inner membrane proteins are targeted via their N-terminal TMD by the SRP system to the membrane-integrated SecYEG translocase or YidC insertase. However, the 217 amino acid long protein SciP is a member of the class of C-tail anchored proteins with a TMD at the extreme C-terminal part of the protein. Due to its topology an early SRP-dependent co-translational targeting by its TMD can be ruled out. However, as has previously been shown for the potassium sensor protein KdpD an early SRP-dependent targeting is not limited to a N-terminal TMD. KdpD is recognized by SRP as soon as a short amphiphilic helix located between the amino acids 22-48 is exposed outside of the ribosomal exit tunnel (Maier et al. 2008). In this chapter the question is addressed, whether both regions serve as SRP signal sequences and are required for membrane insertion. Therefore, cysteine accessibility assays, *in vitro* cross-linking experiments and protein-protein interaction studies with microscale thermophoresis were done.

In chapter 3, the effect on membrane targeting was monitored with fluorescence microscopy studies when the hydrophobic regions were deleted in a sfGFP-SciP fusion construct. To analyze if the translocation of the C-tail of SciP is impaired by the deletion was addressed with cysteine-accessibility assays (Fig. 20). Interestingly, the deletion of either one of the potential SRP signal sequence of SciP did not affect the translocation of the C-tail. In contrast, deletion of both regions resulted in an impaired translocation of the C-tail. From that it is concluded that both regions can mediate membrane targeting independently of each other and only if both regions are missing, targeting of SciP is impaired. The fact that the translocation was not completely abolished when the SRP-mediated targeting of the protein is impaired is not surprising. Also in the Ffh depletion strain a minor amount of the SciP protein was still inserted by YidC (chapter 3). It is possible that in this case, other mechanisms like protein diffusion

in the cytoplasm allows the inefficient insertion of a small amount of the SciP protein into the inner membrane.

The results of the binding studies with different stalled RNCs using microscale thermophoresis clearly indicate that the amino-terminal hydrophobic regions of SciP are recognized by SRP and therefore function as SRP signal sequences. The results also demonstrate that the protein can be bound early during translation by SRP since only short nascent chains of SciP were capable for SRP interaction. Interestingly, both hydrophobic regions of SciP show nearly the same binding affinity to SRP as determined with microscale thermophoresis (Kd 49.7 ± 12 and 57 ± 7.6 nM). These binding affinities correlate with the estimated Kd of the well studied SRP substrate FtsQ (Kd 30.6 ± 6 nM). Previous affinity determinations with SRP and RNCs exposing the strong SRP signal sequence of leader peptidase or an engineered phoA SRP signal sequence revealed a Kd of < 3 nM (Bornemann et al. 2008; Saraogi et al. 2014; Ariosa et al. 2015). These binding affinities were calculated with fluorescence equilibrium titrations using Alexa Fluor 555 labeled SRP or Förster resonance energy transfer (FRET) with Cm-labeled RNCs and BODIPY FL labeled SRP (Bornemann et al. 2008; Saraogi et al. 2014). With FRET, it was also shown that SRP binds to RNCs exposing the SRP signal sequence of FtsQ with a Kd of 1.1 nM (Ariosa et al. 2015). The variations in the estimated Kd values for SRP binding to its substrates probably arise from the different binding assays used to determine the binding affinity. In contrast to the previous studies, unlabeled SRP was used for microscale thermophoresis which could also contribute to the differences in the Kd values. In addition, the buffer of the MST assay differs from that of the study of Ariosa et al. in 2015. In the MST assay a buffer with 20 mM HEPES pH 7.2, 50 mM KOAc, 5 mM $Mg(OAc)_2$, 2 mM L-tryptophane and 0.05% Tween-20 whereas in the study from Ariosa et al. a buffer with 50 mM HEPES pH 7.5, 150 mM KOAc, 10 mM $Mg(OAC)_2$ and 2 mM DTT was used. The fact that MST-based Kd values strongly depend on the buffer conditions was already shown in a previous study (Baaske et al. 2010). The affinity of a DNA-aptamer to the protein thrombin in a 20 mM Tris buffer with pH 7.4 was calculated with 25 ± 2 nM, however in a buffer where Tris was exchanged with 15 mM sodium citrate the affinity decreased to 190 ± 20 nM (Baaske et al. 2010). It is concluded that all these mentioned differences result in the decrease of the binding affinity of SRP to FtsQ-exposing RNCs from about 1 nM to 30 nM. However, due to the fact that the conditions were kept constant during the MST assays, it is possible to compare the Kd values of the different stalled

ribosomes with SRP calculated in this study.

The results obtained from the MST measurements indicate that SciP is the first described protein with two SRP signal sequences. If both regions are active *in vivo* during SRP-dependent targeting or if one region is preferred as a SRP binding site and the other resembles a kind of “back up” signal is unknown. Possibly, the two regions have additional functions for SciP, for example in the folding of the protein as is assumed from the crystal structure of the cytoplasmic domain of SciP (Fig. 25). The architecture of the cytoplasmic domain provides a huge cavity, which is assumed to be structurally important (Durand et al. 2012). In addition, the cytoplasmic domain is responsible for the interaction with the T6SS components TssE, TssK and TssM which enables the formation of the type VI secretion apparatus (Zoued et al. 2016).

As soon as a SRP signal sequence is exposed at the outside of the ribosomal exit tunnel, it is bound by the hydrophobic groove of the SRP M domain, shielding it from the aqueous environment. As observed in cross-linking studies, also the SRP signal sequence of KdpD was harboured in the hydrophobic groove (chapter 2). To identify the exact binding site of the signal sequences of SciP, *in vitro* cross-linking studies with synthesized peptides and SRP single cysteine mutants were performed. The results obtained from the cross-linking studies demonstrated that both SciP peptides are bound by SRP and the binding site was identified in the SRP M domain. This clearly shows that SRP uses the same binding mechanism for cytoplasmic signal sequences as for TMDs.

For SRP signal sequences no consensus motif has been established. A unique feature of all sequences is a stretch of hydrophobic residues exceeding a threshold level of hydrophobicity (Martoglio and Dobberstein 1998; Lee and Bernstein 2001; Hegde and Bernstein 2006; Hainzl and Sauer-Eriksson 2015). It seems that the hydrophobicity is also crucial for the second SRP signal sequence of SciP, since lowering the hydrophobicity in the C68S mutant reduced the binding affinity to SRP and even abolished the ability to bind a preincubated SRP-FtsY complex (Fig. 23 and 24). In contrast, a substitution of the cysteine residue at position 68 of SciP with a methionine residue did not abolish the membrane targeting ability of the second hydrophobic region as shown with fluorescence localisation studies (Fig. 22). The fluorescence signal was partially located at the membrane and at the cell poles, indicating that membrane targeting by SRP is still possible. Probably the mixed phenotype of the SciP

C68M mutant is a result of the possibility to be targeted to the membrane but then the interaction with the membrane-associated FtsY is impaired. This may lead to additional fluorescence signals at the cell poles due to the aggregation of the fusion protein as a result of the impaired SRP-FtsY interaction.

Taken together, the results described in this chapter clearly show that SciP is targeted in a co-translational manner to the membrane-integrated YidC insertase, despite its special topology. Two SRP signal sequences in the N-terminal cytoplasmic part are present to route the protein to the SRP pathway. This mechanism could be favourable since the protein is already located at the membrane close to YidC. As soon as the fully synthesized protein, including the C-terminal TMD, is released from the ribosome YidC is able to directly receive the TMD, preventing the aggregation in the cytosol. It would be very interesting to know if this mechanism is valid for other C-tail anchored proteins in *E. coli*, since to date little is known about their general insertion mechanisms.

Material und Methods

Bacterial strains and culture conditions

E. coli MC4100, BL21 (DE3) and KC6 were grown overnight in LB with ampicillin (100 µg/ml) and back diluted in fresh LB with ampicillin (100 µg/ml).

Construction of SciP mutants for the cysteine accessibility assay

The sequences encoding for amino acids 12-20 and 62-71 of SciP were deleted by site-directed mutagenesis on the plasmid pMS-SciP-C (chapter 3) using the oligonucleotides 5′ CT GTT ATC TCC CGG GCT GAA CAG AGC CAG CTG CGC AGC 3′ / 5′ GCT GCG CAG CTG GCT CTG TTC AGC CCG GGA GAT AAC AG 3′ (∆12–20) and 5′ GCA GGA TTC AGT CAG AAA AGC AGT GAC GAC GAG AGT GTA CTG AAC CGC G 3′ / 5′ C GCG GTT CAG TAC ACT CTC GTC GTC ACT GCT TTT CTG ACT GAA TCC TGC 3′ (∆ 62–71) resulting in plasmids pMS-SciP-CΔ12-20, pMS-SciP-CΔ62-71 and pMS-SciP-CΔ12-20,Δ62-71 respectively. The coding regions of the constructs were verified by sequence analysis.

Construction of SciP-sfGFP-mutants

Mutations in the second hydrophobic region of SciP were generated with site-directed mutagenesis on plasmid pMS-SciP54-85sfGFP96-217 (chapter 3). The

oligonucleotides 5‘ C ATG TTG TAT GCC TTC AGC GCC CTG CTG G 3‘ and 5’ C CAG CAG GGC GCT GAA GGC ATA CAA CAT G 3’ were used to generate pMS-SciP54-85sfGFP96-217 C68S (named SciP54-85 C68S) whereas the oligonucleotides 5‘ C ATG TTG TAT GCC TTC ATG GCC CTG CTG GAC G 3‘ and 5’ C GTC CAG CAG GGC CAT GAA GGC ATA CAA CAT G 3’ resulted in the construction of pMS-SciP54-85sfGFP96-217 C68M (named SciP54-85 C68M). The substitution of the glycine residue at position 16 with a cysteine residue (SciP1-27 G16C) was done with site-directed mutagenesis on plasmid pMS-SciP1-27sfGFP96-217 (chapter 3) using the oligonucleotides 5‘ CAG ATT TTT TAT CCC TGC TGG CTG ATG GTC 3‘ and 5’ GAC CAT CAG CCA GCA GGG ATA AAA AAT CTG 3’. The coding regions of the constructs were verified by sequence analysis.

Construction of plasmids for microscale thermophoresis measurements

For cloning of the stalled ribosome-nascent chain complexes the plasmid pMS-MscL^{115}-TnaC (chapter 2) was used. The sequence which encodes for the amino acids 2-54 of SciP was amplified from pMS-SciP-C with the oligonucleotides 5‘ GGC CAA TTG AAT AAA CCT GTT ATC TCC CGG GC 3‘ and 5‘ CGG CCA TGG TCC TGC TTC GGC CAG CTC TTC ACG 3‘ flanking the PCR product with a *Mfe*I and *Nco*I restriction site. To generate stalled ribosomes with a chain of amino acids 54-100 of SciP the oligonucleotides 5‘ GGC CAA TTG GGA TTC AGT CAG AAA AGC AGT GAC 3‘ and 5‘ CGG CCA TGG CGT ACC AAA AAA ATG AGC CTG CAG C 3‘ were used on plasmid pMS-SciP-C. Both PCR products and the plasmid pMS-MscL^{115}-TnaC were digested with *Mfe*I and *Nco*I and ligated which results in the exchange of the MscL sequence with the two SciP sequences, respectively. Plasmid pMS-SciP54-100-TnaC was used for site-directed mutagenesis with the oligonucleotides 5‘ C ATG TTG TAT GCC TTC AGC GCC CTG CTG G 3‘ and 5’ C CAG CAG GGC GCT GAA GGC ATA CAA CAT G 3’ to generate pMS-SciP54-100-C68S-TnaC. The coding regions of the constructs were verified by sequence analysis.

The construction of plasmid pMS-Luc2-50-TnaC is described in chapter 2. The plasmid encoding amino acids 4-85 of FtsQ (pEM36-3C) (chapter 2) was provided by R. Beckmann, Munich.

AMS-Derivatisation

The *E. coli* strain MC4100 was transformed with the plasmids pMS-SciP-C, pMS-SciP-CΔ12-20 (named Δ12-20), pMS-SciP-CΔ62-71 (named Δ62-71) and

pMS-SciP-CΔ12-20, Δ62-71 (named Δ12-20 Δ62-71), respectively. After the growth to an OD_{600} of 0.5, the cells were washed twice and resuspended in M9 minimal media lacking methionine and cysteine. After 1h at 37°C, the expression of the SciP variants was induced with 1 mM IPTG for 10 min followed by the addition of 2.5 mM AMS and after 1 min the cells were pulse-labeled with ^{35}S-methionine / cysteine for 2 min. Non-radioactive methionine / cysteine was added for 10 min followed by the quenching of the AMS reaction with 10 mM DTT for another 10 min. The samples were precipitated with 20% TCA, resuspended in 10 mM Tris / 2% SDS, immunoprecipitated with anti-His antibody and loaded on a 14% SDS-PAGE followed by the examination by phosphorimaging. The quantification was done with the program ImageJ.

Purification and labeling of ribosome-nascent chain complexes

The plasmids pMS-SciP2-54-TnaC, pMS-SciP54-100-TnaC, pMS-SciP54-100-C68S-TnaC, pMS-Luc2-50-TnaC and pEM36-3C were transformed in *E. coli* KC6. The purification of the different stalled ribosomes and the labeling with the fluorescent dye NT-647 was exactly done as described in chapter 2. All labeled RNCs were stored at -80°C until the measurements.

Purification of Ffh and FtsY

The plasmids pMS-Ffh-C-Strep, pMS-Ffh-L181C-C-Strep, pMS-Ffh-M423C-C-Strep and pTrc99-FtsY-His (kindly provided by HG Koch, Freiburg) were transformed in *E. coli* BL21 (DE3) cells and purified as described in chapter 2.

***In vitro* 4.5S RNA synthesis and SRP reconstitution**

The purified Ffh protein has to be reconstituted with the 4.5S RNA prior to the experiments to get a functional SRP. The *in vitro* transcription and the reconstitution were done as described in chapter 2.

Microscale thermophoresis

The MST measurements of the stalled RNCs with SRP and SRP-FtsY were done as described in the method section of chapter 2.

Peptide synthesis

The two peptides SciP1-27 G16C (MNKPVISRAEQIFYPCWLMVSQLRSGQ) and SciP54-85 (GFSQKSSDIMLYAFCALLDESVLNREKTDDGW) were synthesized *in vitro* by the Custom peptide synthesis services from GENOSPHERE Biotechnologies

(France) with N-terminal acetylation and C-terminal amidation and a purity of >95%. The SciP1-27 G16C peptide was dissolved in H_2O:CH_3CN (3:1) whereas the SciP54-85 peptide was dissolved in DMSO.

Cross-linking studies with copper phenanthroline

For cross-linking studies, the SciP1-27 G16C peptide was synthesized with a cysteine residue at position 16 (G16C). The reconstitution of SRP 181C/406C and SRP 406S/423C as well as the cross-linking studies were done exactly as described in the material and method section of chapter 2.

CROSS-CHAPTER DISCUSSION

5

5.1 General discussion

In this study, the insertion mechanism of a special inner membrane protein, belonging to the class of C-tail anchored proteins, was explored in *E. coli* in detail. To date, only little is known about the targeting and insertion pathways of this special protein class in prokaryotes, however, the insertion mechanism of tail-anchored proteins in eukaryotes is well understood.

As a model protein to study C-tail anchored protein insertion, the 217 amino acid long protein SciP from enteroaggregative *E. coli* was used. SciP has a single TMD located at the extreme C-terminal part of the protein (Aschtgen et al. 2012). This results in a large cytoplasmic domain of 183 amino acids and a short periplasmic domain of only 11 amino acids. Interestingly, there is another *E. coli* protein, not part of the class of C-tail-anchored proteins, which shares a commonality with SciP, the potassium sensor protein KdpD. The 894 amino acid long protein KdpD is a four-spanning membrane protein with a large N- and C-terminal cytoplasmic domain of about 400 amino acids, respectively (Zimmann et al. 1995; Facey and Kuhn 2003). This means that, just like the protein SciP, the first TMD of KdpD is located far away from the cytoplasmic N-terminus. This special feature would make one assume to rule out an early co-translational targeting and insertion mechanism. Nevertheless, it was shown for the protein KdpD, that the component of the co-translational targeting pathway, the signal recognition particle (SRP), is still involved. A short amphiphilic region between amino acid 22-48 in the cytoplasmic domain of KdpD is recognized by SRP and therefore serves as a signal for co-translational targeting (Maier et al. 2008). Beside the SRP system no other protein component like SecYEG or YidC is involved in the insertion of KdpD into the inner membrane (Facey and Kuhn 2003).

In this study it was explored whether the structural commonalities between SciP and KdpD of having a large cytoplasmic domain at the N-terminus also result in similarities in their targeting and insertion pathways. Therefore, the previous studies on the SRP-dependent targeting of KdpD was expanded. Sequential mutations in the SRP signal sequence were analyzed in relation to their localisation with sfGFP fusion constructs and fluorescence microscopy as well as the affinity of SRP or SRP-FtsY to bind to artificially stalled ribosomes using microscale thermophoresis (chapter 2). This was done to decode the detailed mechanism of the SRP-dependent targeting of KdpD. In

this regard, KdpD served as a pioneer for the studies on the targeting and insertion of the C-tail anchored protein SciP. The results obtained from cysteine-accessibility assays and sfGFP fusion proteins clearly show that SciP also depends on the SRP system, however YidC but not SecYEG is necessary for the insertion (chapter 3). Two short hydrophobic regions (aa 12-20 and 62-71) were identified that qualify as SRP signal sequences, since both are able to target the cytoplasmic protein sfGFP to the *E. coli* inner membrane (chapter 3). In chapter 4, microscale thermophoresis measurements with TnaC stalled ribosomes and purified SRP point out that SRP is able to bind both hydrophobic regions with similar affinity. *In vitro* disulphide cross-linking studies demonstrate that the hydrophobic regions of SciP can bind to the SRP M domain, similar to other known SRP substrates. Therefore, SciP is the first described protein with two SRP signal sequences in the N-terminal cytoplasmic domain. This poses the question why SciP, in contrast to other SRP substrates, contains two SRP targeting signals.

The data obtained in this study indicate a novel class of SRP substrates exists, that are not recognized by hydrophobic TMDs but by short hydrophobic regions in a N-terminal cytoplasmic domain. This allows an early targeting during translation even if the TMD is far away from the cytoplasmic N-terminus. Also, it was confirmed that in contrast to KdpD, SciP needs the help of the YidC insertase for its insertion into the inner membrane. The involvement of SRP and YidC in the insertion of SciP are first indications how the insertion of a C-tail anchored protein occurs.

In addition to the studies on SRP-mediated targeting, the insertion of SciP via YidC was analyzed in more detail. Protease-accessibility assays demonstrated that the YidC-mediated insertion of SciP depends on the characteristic of the short periplasmic part. Sequential changes in length and hydrophilicity of this region also gained information about the requirements of proteins to use the YidC only pathway (chapter 3).

This study addresses general open questions in the insertion pathways of inner membrane proteins in *E. coli* and also provides further insights into the not well studied delivery pathway of the C-tail anchored protein SciP.

5.2 A novel SRP signal sequence in the protein KdpD

In contrast to most inner membrane proteins that are targeted co-translationally by exposing their N-terminal hydrophobic TMD, KdpD and SciP are recognized by SRP via short hydrophobic regions in the N-terminal cytoplasmic part. For KdpD, this SRP signal sequence was identified at the beginning of the protein between amino acid 22-48 (Maier et al. 2008). To gain more information about the conditions of the amino acid composition necessary for SRP binding, several mutations in the signal sequence were analyzed with sfGFP fusion constructs and fluorescence microscopy (chapter 2). These studies revealed that basic amino acids at the N-terminal part of the signal sequence as well as the hydrophobicity of the core region are crucial for SRP-dependent membrane targeting (chapter 2). To study the interaction of SRP with the signal sequence of KdpD as well as to the signal sequence mutants the amino acids 22-74 of KdpD were fused to the TnaC stalling sequence resulting in artificially stalled ribosomes with a nascent chain length of 100 amino acids (chapter 2). Since about 30 amino acids are buried in the ribosomal exit tunnel, the amino acids 22-63 of KdpD are probably exposed and therefore accessible for binding of SRP via its hydrophobic groove (Fig. 27A).

Microscale thermophoresis measurements demonstrated that the binding of SRP to the nascent chain of KdpD correlates with the binding of SRP to ribosomes exposing the well-studied SRP substrate FtsQ. The SRP signal sequence of FtsQ is located within the first TMD, between amino acid 28-48 (Valent et al. 1995; Schaffitzel et al. 2006). Due to the construction of the stalled RNCs also this sequence is fully exposed and accessible for SRP (Fig. 27B). Selective ribosomal profiling studies identified 488 inner membrane proteins, including KdpD, that interact with SRP (Schibich et al. 2016). On average, when the N-terminus of the first TMD of these proteins reaches a distance of 42 amino acids from the peptidyl transferase centre (PTC) in the ribosome, SRP binding was observed, with the maximal binding at a distance of 55 amino acids. On average, SRP does not bind to emerging substrates when the N-terminus of the first TMD reaches a distance of about 100 amino acids from the PTC (Schibich et al. 2016). In the case of the used KdpD22-74-RNCs and the FtsQ4-85-RNCs, the N-termini of the predicted SRP signal sequences are located 73 or 93 amino acids away from the PTC, respectively (Fig. 27A, B). Even if this is at the very edge, both signal sequences lie within the described SRP binding range (distance of 42-100 aa from the PTC).

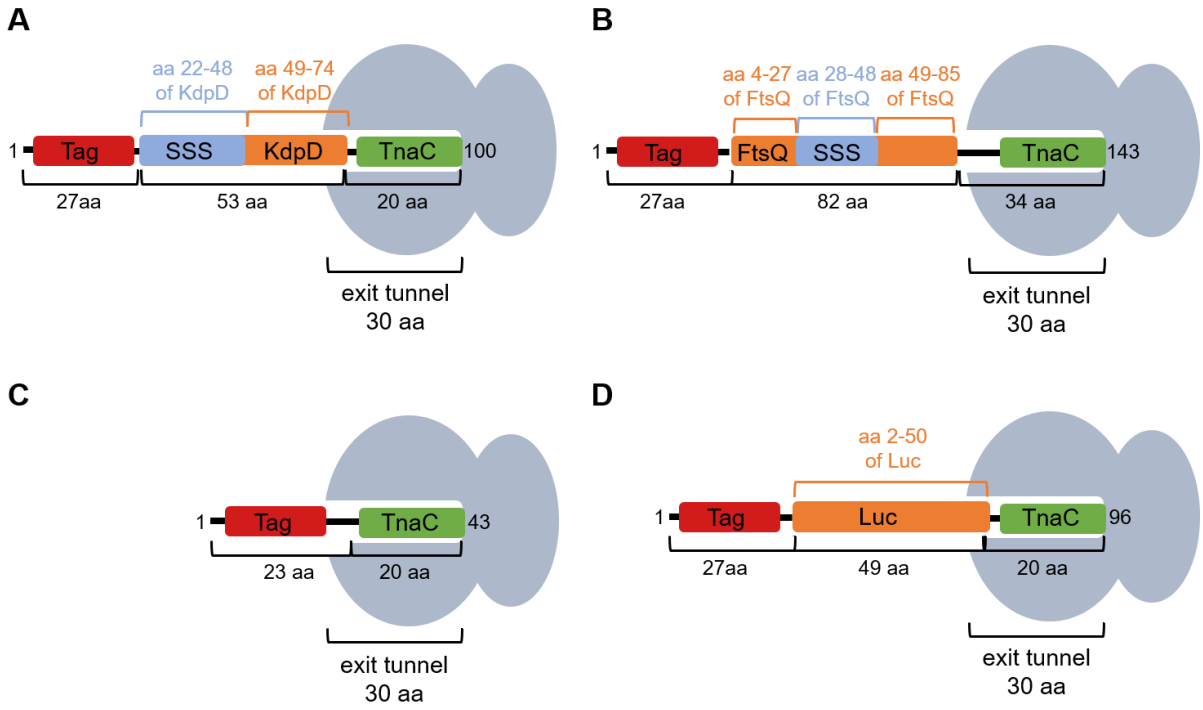


Figure 27: Schematic representation of the TnaC-stalled ribosomes used in this study. For purification all RNCs (50S and 30S subunit in gray) have a N-terminal His-Tag (red). The amino acid (aa) length of the different components of the nascent chain is indicated. The nascent chains of KdpD, FtsQ and Luciferase are colored orange, the (putative) SRP signal sequences (SSS) are indicated in blue and the TnaC stalling sequences, buried in the ribosomal exit tunnel, are colored in green. Stalled ribosomes with nascent chains of (A) KdpD22-74, (B) FtsQ4-85, (C) 23 amino acid long unspecific sequence including the His-Tag (short-chain ribosomes) and (D) Luciferase2-50 are shown.

MST measurements of stalled ribosomes with a total nascent chain length of 96 amino acids including amino acids 2-50 of the cytoplasmic protein Luciferase (Fig. 27D) did not show a binding to SRP. This is in good agreement with a previous study on the SRP substrate leader peptidase (Houben et al. 2005). The exchange of 4 hydrophobic leucine residues with a helix-breaking proline residue and charged residues (L7E, L9K, L14P, L19R) abolished cross-linking of SRP to leader peptidase (Houben et al. 2005). An unspecific, nascent-chain independent binding of SRP to ribosomes was only observed when the nascent chain was still buried in the ribosomal exit tunnel (Bornemann et al. 2008). Also, the ribosomes with a short nascent chain of 43 amino acids (named short-chain ribosomes), containing no SRP signal sequence showed no binding to SRP (Fig. 27C and chapter 2). Considering the fact that about 30 amino acids are buried in the ribosomal exit tunnel, these RNCs already exposed about 13 amino acids outside the tunnel, consisting of the N-terminal His_{10} and HA tag. At that point, *in vivo,* SRP is already retracted from the inside of the ribosomal tunnel and is

in the recognition mode if a correct cargo is exposed or is already dissociated from the ribosome in the case of a non-SRP-substrate (Denks et al. 2017). This leads to the conclusion that, as soon as the nascent chain reaches a length of about 40 amino acids, SRP can discriminate between a correct and incorrect cargo and this discrimination is traced by the used microscale thermophoresis method. In conclusion, amino acids 22-74 of KdpD represent a correct cargo for SRP.

The finding that SRP can also bind to amphiphilic cytoplasmic regions outside a TMD was confirmed by studies on the heat shock transcription factor σ^{32}. During negative feedback regulation in the heat shock response, σ^{32} is targeted to the membrane via the SRP system for degradation by FtsH (Lim et al. 2013). With disulphide cross-linking, an amphiphilic α-helix in the homeostatic control region of σ^{32}, located between amino acids 35-67, was identified as SRP binding site (Miyazaki et al. 2016). That confirms that, as already assumed for KdpD, amphiphilic helices provide a hydrophobic surface for SRP binding. Further, the control region of σ^{32} could be cross-linked to the signal peptide binding region of SRP (Miyazaki et al. 2016), which lies within the C-terminal M domain. Cross-linking studies with the synthesized KdpD22-48 peptide complemented these findings, since the peptide was also cross-linked to the hydrophobic groove of SRP. This clearly shows that the binding site of cytoplasmic SRP signal sequences are shared with those of transmembrane embedded SRP signal sequences.

5.3 The role of SRP in the targeting of the C-tail anchored protein SciP

The fact that SRP binding is not limited to TMDs, is also strengthened by the targeting and insertion studies on the C-tail anchored protein SciP. As KdpD, also SciP is recognized by SRP and the results obtained in this study indicate that the targeting also occurs early during translation, due to the presence of a SRP signal sequence in the N-terminal cytoplasmic domain. The fact that not the TMD but a cytoplasmic signal in the protein is responsible for SRP-dependent targeting was obtained with subcellular localisation studies using fluorescence microscopy. The fusion of the amino acids 96-217 of SciP, including the TMD, to the C-terminus of sfGFP was not functional for membrane targeting (chapter 3). This clearly shows that the TMD of SciP cannot be recognized by SRP, probably due to the long distance from the N-terminus of the protein. From that it was concluded, that a targeting signal has to be present in the N-

terminal part of the protein to route it into the SRP pathway, despite its special topology. In contrast to KdpD, there is strong evidence that SciP contains two individually acting targeting signals. This conclusion arises from different experiments described in chapter 3 and 4. First, the deletion of two short hydrophobic stretches from amino acid 12-20 and 62-71 in a sfGFP-SciP fusion construct impaired the targeting of the protein, as shown with fluorescence microscopy. In addition, the translocation of the C-tail is impaired if both regions are deleted, monitored via cysteine accessibility assays (chapter 4). In contrast, the amino acids 1-27 and 54-85 each contain a hydrophobic stretch and are able to target the sfGFP-SciP96-217 fusion protein to the membrane (chapter 3). This indicates that both regions resemble a functional targeting signal. When synthezised as individual peptides, the amino acids 1-27 and 54-85 of SciP bind to the SRP M domain, as was shown with *in vitro* disulphide cross-linking studies (chapter 4). As had been shown for the KdpD22-48 peptide (chapter 2) the SciP peptides were cross-linked to the cysteine at position 423 of the SRP M domain. *In vitro* interaction studies of stalled ribosomes with amino acids 2-54 or 54-100 of SciP as a nascent chain and SRP demonstrated that both nascent chains are bound with similar affinity by SRP or by a preincubated SRP-FtsY complex (chapter 4). From that it is concluded that both regions can be recognized and bound by SRP during co-translational targeting. With cysteine accessibility assays it was shown that the deletion of only one of the hydrophobic regions can be compensated by the presence of the remaining one (chapter 4). This strengthens the hypothesis that both regions act independently of each other in membrane targeting. Possibly, the presence of two targeting signals functions as a back-up system if one region is skipped during the SRP scanning and recognition mode. If there is a preference of SRP for one of the two regions has to be addressed with additional experiments. Global profiling studies revealed that RNCs are bound multiple times by SRP and these binding events correlate with the exposure of TMDs of the substrate (Schibich et al. 2016). It is assumed that some RNCs exposing an only moderately hydrophobic TMD sometimes lose the contact to the translocon. Hence, it is suggested that they are re-targeted by SRP due to the binding of an exposed downstream TMD (Schibich et al. 2016). The presence of two SRP binding sites in the SciP protein may enable a re-targeting mechanism in the case that the SciP-RNCs lose the contact to the YidC insertase. This would make the targeting step more efficient.

The microscale thermophoresis measurements with artificially stalled RNCs, exposing either one of the hydrophobic regions and purified SRP, as described for KdpD, clearly show that the interaction with SRP occurs early during translation (chapter 4). For the measurements, the sequence encoding for the amino acids 2-54 or 54-100 of SciP were fused to the TnaC stalling sequence resulting in a nascent chain length of 100 and 93 amino acids, respectively (Fig. 28). As described, about 30 amino acids are buried in the ribosomal exit tunnel which leads to the exposure of amino acids 2-44 or 54-90 of SciP including the exposure of the hydrophobic stretch, respectively. Since SRP is able to bind to these nascent chains it is concluded that both regions represent a SRP signal sequence. The N-termini of the two SRP signal sequences are located 73 or 66 amino acids away from the PTC, respectively. These distances also match with the findings of Schibich et al. in 2016 who defined the SRP binding range with a distance of 42-100 amino acids away from the PTC. Both nascent chains are bound by SRP with similar affinity and this affinity is also similar to the one observed for FtsQ. From that it is concluded that SRP can bind to both hydrophobic regions in the protein and this binding can occur independently of one another (chapter 4). As a result, SciP is targeted to the inner membrane in a co-translational manner due to the binding of SRP early during translation to the short hydrophobic regions in the N-terminal cytoplasmic part.

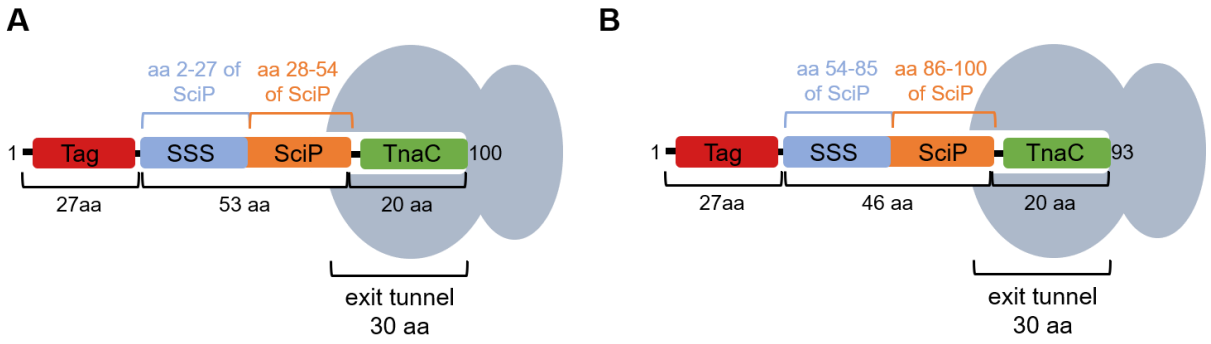


Figure 28: Schematic representation of the SciP-TnaC-stalled ribosomes used in this study. For purification the RNCs (50S and 30S subunit in gray) consist of an N-terminal His-Tag (red). The amino acid (aa) length of the different components of the nascent chain is indicated. The nascent chains of SciP are colored orange, the (putative) SRP signal sequences (SSS) are indicated in blue and the TnaC stalling sequences, buried in the ribosomal exit tunnel, are colored in green. Stalled ribosomes with nascent chains of (A) SciP2-54 and (B) SciP54-100 are shown.

This finding is in contrast to another study on two additional C-tail anchored proteins from *E. coli,* DjlC and Flk (Peschke et al. 2018). The significantly larger proteins DjlC

(483 amino acids) and Flk (331 amino acids) have short periplasmic tails of only 1 amino acid and the authors concluded that the targeting also occurs via the SRP system but in a post-translational way. This hypothesis arises from fusion constructs where the C-terminal TMD of DjlC or Flk was added to the C-terminus of the fluorescent protein mNeonGreen. In both cases, the TMD was sufficient for a localization of the fusion construct at the membrane and this localisation was impaired in absence of SRP. However, Peschke at al. only analyzed the dependency on SRP using artificial mNeonGreen fusion constructs and subcellular localisation with fluorescence microscopy. The artificial mNeonGreen fusion proteins miss the whole context of the large cytoplasmic part of the protein. Furthermore, no dependency on SRP was shown for the full-length protein. It is also not clear if the absence of SRP also impairs the insertion of the fusion proteins into the inner membrane, since C-tail translocation studies are only shown under wild-type conditions and not in the absence of SRP. In addition to the subcellular localisation studies, Peschke et al. showed that Ffh, the protein component of SRP, can be cross-linked *in vivo* to the TMDs of DjlC and Flk. By virtue of these experiments, using the C-terminal TMDs, the authors concluded that targeting of DjlC and Flk by SRP must occur post-translationally (Peschke et al. 2018). This is different to the C-tail anchored protein SciP, where the fusion of the TMD to sfGFP was not sufficient for a localization at the membrane (chapter 3), indicating that in the case of SciP, the TMD does not contain information for the targeting of the protein to the inner membrane. In contrast, it seems that the TMDs of DjlC and Flk do contain such a targeting information.

Even if the three proteins DjlC, Flk and SciP are classified as C-tail anchored proteins, it seems that different insertion pathways may exist in *E. coli.* Also for eukaryotic tail anchored (TA) proteins different targeting and insertion pathways were described. In mammalian cells and in yeast there is evidence that some of them can insert spontaneously without the assistance of other protein components. *In vitro* studies with the TA proteins $Cytb_5$ and tyrosine phosphatase 1B showed that they can insert into liposomes without the assistance of other proteins or nucleotides (Brambillasca et al. 2005; Brambillasca et al. 2006; Colombo et al. 2009) . Besides that, it is assumed that the majority uses a chaperone-mediated pathway, divided in the SRP, the Hsc70/Hsp40, the SND or the GET pathway (Rabu et al. 2009; Borgese and Righi 2010; Johnson et al. 2013). The choice of delivery pathway is assumed to be dependent on the hydrophobicity of the tail anchor. However, a unique feature of

eukaryotic TA proteins is that they are all inserted post-translationally (Rabu et al. 2009; Borgese and Righi 2010; Johnson et al. 2013).

The results obtained in this study regarding the targeting and insertion of the C-tail anchored protein SciP, are the first indications for a SRP-dependent co-translational targeting of TA proteins. The fact that SciP presumably uses a different pathway than DjlC and Flk could also depend on the hydrophobicity of the TMD.

In contrast to SciP, Flk has a very hydrophobic tail anchor of 2.83 GRAVY and DjlC a moderately hydrophobic one of 1.95 GRAVY (Peschke et al. 2018). In contrast, the tail anchor of SciP even has a lower hydrophobicity than DjlC with only 1.47 GRAVY (Table 1). The hydrophobicity plot of the full-length SciP protein demonstrates that the GRAVY of the first hydrophobic region (aa 12-20) lies within the same range (1.44 GRAVY) as the TMD. However, the second hydrophobic region (62-71) is even much more hydrophobic than the TMD with 2.54 GRAVY.

Table 1: Comparison of length and hydrophobicity of the C-tail anchored proteins SciP, DjlC and Flk. Abbreviations: aa = amino acid, GRAVY = grand average of hydropathicity, TMD = transmembrane domain.

Protein	Length (aa)	GRAVY TMD
SciP	217	1.47
DjlC	483	1.95
Flk	331	2.83

Possibly due to the low hydrophobicity of the TMD, SciP is not compatible for an SRP-dependent post-translational targeting and needs an additional hydrophobic region to be successfully routed to the co-translational SRP pathway. Hydrophobicity plots of DjlC and Flk revealed no hydrophobic patches in the cytoplasmic first 100 amino acids. Maybe due to their higher hydrophobicity of the TMDs such hydrophobic regions are not necessary for targeting. It cannot be ruled out that, in addition to both hydrophobic regions of SciP, also the TMD is binding to SRP at later stages during the insertion mechanism, as observed for DjlC and Flk (Peschke et al. 2018). This hypothesis was raised due to the ribosomal profiling study of Schibich et al. in 2016. Their results indicate that about 77% of all SRP-dependent inner membrane proteins are binding

several times to SRP indicating multiple cycles of SRP-dependent targeting. These multiple binding events correlate with the exposure of the TMDs of multi-spanning membrane proteins (Schibich et al. 2016).

In conclusion and in contrast to DjlC and Flk, SciP is targeted by SRP in a co-translational way probably due to the recognition of the two hydrophobic regions in the cytoplasmic domain of SciP. They actually function as a SRP signal sequence and can both be bound during co-translational targeting of the protein.

The hypothesis that DjlC and Flk use a different delivery pathway than SciP is also strengthened by the fact that the chaperone DnaK is involved (Peschke et al. 2018), whereas DnaK was not found to be important for the targeting and insertion of SciP (chapter 3). Assuming a post-translational targeting and insertion of DjlC and Flk, it is reasonable that chaperones like DnaK are needed to prevent the aggregation of the C-terminal hydrophobic TMD in the cytoplasm. A recent study identified that the yeast homolog Ssa1 captures artificial model TA proteins (non-cleavable SUMO domain fused to the TMDs of Sbh1 and Bos1, respectively) to keep them in a soluble and translocation-competent state (Cho and Shan 2018). After the capturing by Ssa1 the proteins are then delivered to the Get pathway components for targeting and insertion (Cho and Shan 2018). For SciP maybe such a mechanism is redundant, since it is targeted co-translationally and the nascent chain is already located close to the insertase as soon as the TMD is exposed outside of the ribosome. Due to the co-translational targeting the C-terminal TMD can be received directly by YidC preventing its aggregation.

The differences in the targeting and insertion of DjlC, Flk and SciP indicate once again that the targeting and insertion of proteins into the inner membrane of *E. coli* is a very complex mechanism where quite a few new fascinating discoveries can be expected.

5.4 Decoding the novel cytoplasmic SRP signal sequences

The studies with the protein KdpD revealed that co-translational targeting by SRP is not limited to the recognition and binding of an emerging transmembrane domain. With protease accessibility assays it was shown that the deletion of the N-terminal first 50 amino acids of KdpD prevented the insertion of the protein (Maier et al. 2008). In addition, pull-down assays with SRP and KdpD1-48 fused to GFP demonstrated that residues 1-48 of KdpD are bound by SRP (Maier et al. 2008). Here, the novel SRP

signal sequence of KdpD was further investigated to decode the conditions needed for the binding of a cytoplasmic signal sequence to SRP. The aim was to test if cytoplasmic SRP signal sequences correspond in their main features with those located in the TMDs of SRP substrates. Crucial for an SRP binding is a stretch of hydrophobic residues exceeding a threshold level of hydrophobicity (Martoglio and Dobberstein 1998; Lee and Bernstein 2001; Hegde and Bernstein 2006; Hainzl and Sauer-Eriksson 2015). In addition, basic amino acids are assumed to promote SRP binding (Peterson et al. 2003). The novel SRP signal sequence of KdpD contains a hydrophobic core region with a stretch of 10 hydrophobic residues followed by a lysine residue and additional six hydrophobic residues with a total GRAVY of 1.33 (Fig. 29). The amino acids 22-48 of KdpD fused to the N-terminus of the fluorescent protein sfGFP resulted in a SRP-dependent localisation at the membrane (chapter 2). Several substitutions in the sfGFP fusion construct, regarding hydrophobicity and charge were analyzed with fluorescence microscopy to identify the crucial features (chapter 2). In addition to the *in vivo* experiments, the SRP binding was also investigated *in vitro* with TnaC stalled ribosomes exposing the SRP signal sequence of KdpD (chapter 2). Both *in vivo* and *in vitro*, the decrease in the hydrophobicity of the SRP signal sequence of KdpD (1.33 to 1.14 GRAVY in the Walker S mutant) resulted in a weaker SRP binding and targeting defects (chapter 2).

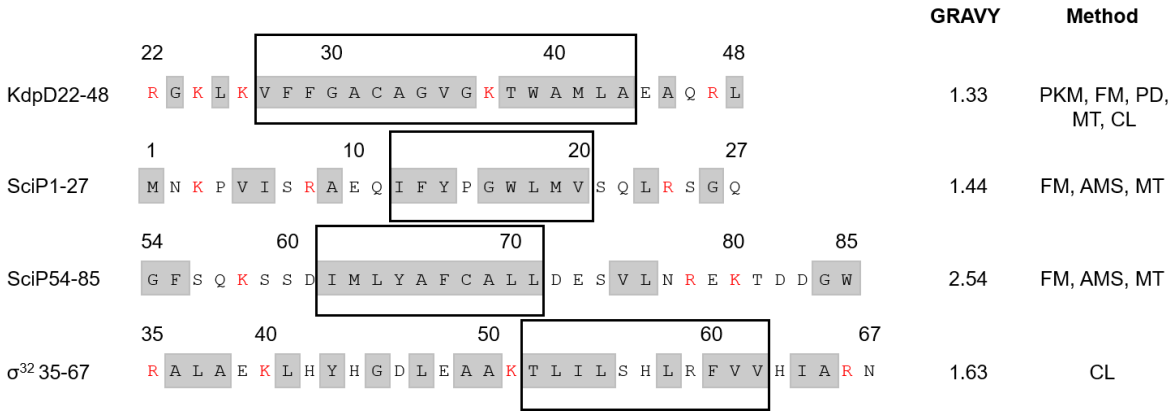


Figure 29: Features of the SRP signal sequence of KdpD and σ^{32} and the two signal sequences of SciP. The number represents the position of the amino acid in the full-length protein. Hydrophobic amino acids are shaded gray and positively charged amino acids are colored red. The GRAVY of the amino acids in the black boxes and the methods used to analyze the sequences is listed (PKM = Proteinase-K Mapping, FM = fluorescence microscopy, PD = pull-down assay, MT = microscale thermophoresis, CL = cross-linking studies, AMS = cysteine accessibility assay with AMS).

This was also shown for the SRP binding site of σ^{32}, since the exchange of a hydrophobic into a hydrophilic amino acid residue (I54N) (1.63 to 0.9 GRAVY) reduced the amount of cross-linked Ffh (Miyazaki et al. 2016). Both studies show that also N-terminal cytoplasmic SRP signal sequences must exceed a threshold level of hydrophobicity to be successfully recognized by SRP.

Compared to the hydrophobicity of the well-studied SRP substrates FtsQ (GRAVY 2.17) or leader peptidase (GRAVY 2.6), the SRP binding sites of KdpD and σ^{32} are only moderately hydrophobic. The lower hydrophobicity may be compensated by the presence of the positively charged amino acids at the N-terminal part of the hydrophobic core regions, respectively (Fig. 29). This could be due to the generation of salt bridges between the SRP RNA and the basic amino acids of the signal sequence (Peterson et al. 2003). Hence, the three positively charged amino acids at the N-terminal part of the KdpD SRP signal sequence were exchanged into neutral amino acids. These alterations resulted in a weaker SRP binding as monitored with microscale thermophoresis and an impaired targeting to the membrane analysed with fluorescence microscopy (chapter 2). The fact that moderately hydrophobic regions can route proteins to the SRP pathway by added basic amino acids was shown in the study from Peterson et al. in 2003. The replacement of the signal peptides of the SecB-dependent proteins MBP and OmpA with a moderately hydrophobic signal sequence with four closely spaced basic residues led to a SRP-dependent targeting (Peterson et al. 2003). In conclusion, a lower hydrophobicity of the core region may be compensated by the presence of positively charged residues to allow SRP binding.

A closer look at the two SRP signal sequences of SciP reveal that these regions correlate well with the characteristics of the KdpD and σ^{32} signal sequences (Fig. 29). The hydrophobicity of the first cytoplasmic region of SciP (GRAVY 1.44) deciphers that it is only moderately hydrophobic but it also contains 2 positively charged residues (Fig. 29). In contrast, the second region of SciP is very hydrophobic with a GRAVY of 2.54 and contains only one positively charged residue (Fig. 29). As already mentioned, the addition of basic amino acids to moderately hydrophobic signal sequences was shown to route proteins (MBP and OmpA) to the SRP pathway (Peterson et al. 2003). Perhaps, the only moderately hydrophobic region (aa 12-20) is also sufficiently recognized by SRP due to the presence of the two basic amino acid residues, facilitating a stable binding by the generation of salt bridges with the SRP RNA.

With fluorescence microscopy and stalled ribosomes, it was found out that a decrease

in the hydrophobicity of the second hydrophobic region of SciP has a similar effect as already observed for the KdpD and σ^{32} signal sequences. The substitution of the cysteine residue at position 68 with a serine residue (2.54 to 1.96 GRAVY) led to targeting defects of the sfGFP fusion construct (chapter 4). In contrast, the substitution with the more hydrophobic methionine residue (2.54 to 2.26 GRAVY) had only minor effects on the targeting of the fusion construct (chapter 4). This indicates that not a specific amino acid residue but the hydrophobicity is crucial for SRP binding.

The results obtained with fluorescence microscopy were confirmed with SRP binding assays to the nascent chain containing the second hydrophobic region of SciP with the serine residue. The C68S mutant showed a significant weaker binding to SRP compared to the non-mutated one (chapter 4). The substitution with a serine residue blocked the binding to a preincubated SRP-FtsY complex completely (chapter 4). This indicates that the level of hydrophobicity was not sufficient in the C68S mutant, which makes it no longer recognizable for SRP.

It would be interesting to see if alterations in the first SRP signal sequence (aa 12-20), regarding the hydrophobicity, have similar effects on SRP binding and targeting as observed for the second hydrophobic region. Furthermore, the influence of the basic amino acid residues in the hydrophobic regions of SciP on SRP binding and targeting could also be investigated with substitutions to neutral amino acid residues. This could further contribute to the understanding of the function of N-terminal cytoplasmic SRP signal sequences.

Another similarity of the SRP binding sites of KdpD and σ^{32} is that they represent short amphipathic helices. The crystal structure of the cytoplasmic domain of SciP (aa 6-178) reveals that the two SRP signal sequences fold into an α-helical structure, respectively (Fig. 30A). Helical wheel projections of the two helices identified that also the less hydrophobic region (aa 12-20) folds an amphipathic helix (Fig. 30B) in contrast to the second, the strong hydrophobic one, where no amphipathic character is recognizable (Fig. 30C). Crucial for the SRP binding to a signal sequence is an uninterrupted stretch of at least 8 hydrophobic residues (Martoglio and Dobberstein 1998; Lee and Bernstein 2001; Hegde and Bernstein 2006; Hainzl and Sauer-Eriksson 2015). This uninterrupted stretch is present in the second hydrophobic region of SciP (Fig. 29), providing a potential binding site for SRP. In contrast, in the first hydrophobic region of SciP as well as in the SRP recognition sequence of σ^{32}, the hydrophobic parts are interrupted by hydrophilic amino acid residues (Fig. 29). For σ^{32} it was assumed

that the formation of an amphiphilic helix compensates for the absence of the consecutive hydrophobic amino acids and allows the recognition by SRP (Miyazaki et al. 2016). From that it is concluded that also the first hydrophobic region of SciP is able to be recognized by SRP due to the formation of an amphiphilic helix, providing a hydrophobic surface for SRP binding.

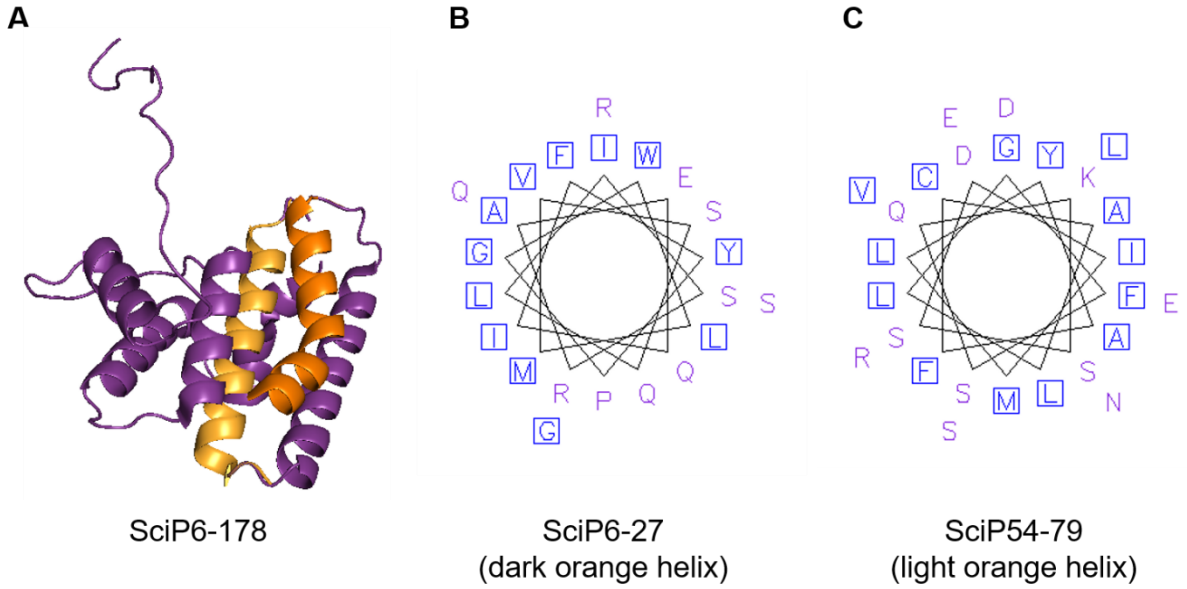


Figure 30: Analysis of the hydrophobic regions of SciP. (A) Crystal structure of the cytoplasmic amino acids 6-178 of SciP (purple). Amino acids 6-27 of SciP are colored in dark orange, the amino acids 54-79 are colored in light orange (Image created with PyMOL 2.3.0, PDB 3U66). (B and C) Helical wheel projection of the amino acids (B) 6-27 and (C) 54-79 of SciP. Hydrophobic residues are marked as blue squares, hydrophilic amino acids are colored purple (made with “emboss pepcoil”).

Taken together, the two SRP signal sequences of SciP correlate well with the features of the signal sequence of KdpD necessary for SRP binding. Both regions of SciP represent an SRP targeting signal, however, it still has to be tested if both regions are active in SRP binding *in vivo* and if one region is preferred as a SRP binding site.

5.5 Binding to the SRP M domain

The analysis of the novel SRP signal sequences of KdpD and SciP pointed out that they match in their structure with other known SRP signal sequences located in the TMD of the protein. Whether they are bound with a similar mechanism by SRP was addressed in this study. During co-translational targeting the SRP signal sequences

are harboured within a hydrophobic groove formed by the SRP M domain (Keenan et al. 1998). This groove protects the first hydrophobic TMD, which normaly functions as the targeting signal (Lee and Bernstein 2001), from the aggregation in the aqueous environment (Fig. 31A). With disulphide cross-linking studies it was tested if the cytoplasmic SRP signal sequences of KdpD and SciP are harboured in the hydrophobic groove during targeting. Therefore, the methionine residue at position 423, located towards the inside of the hydrophobic groove, was substituted with a cysteine residue (Fig. 31A).

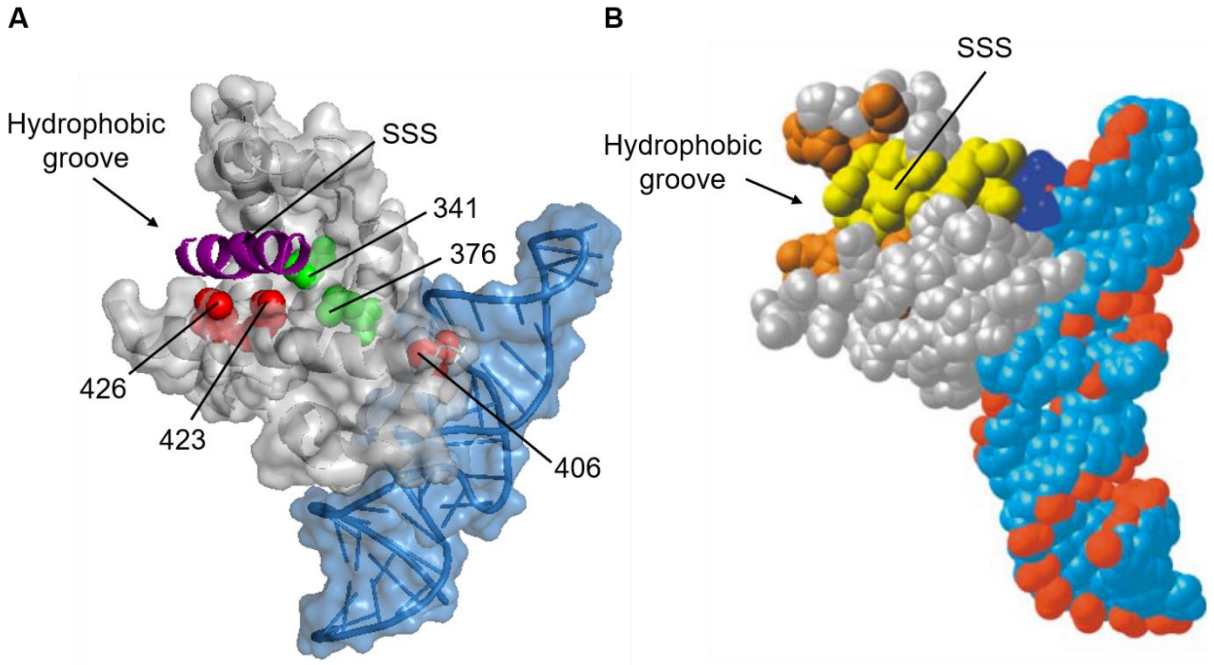


Figure 31: Signal sequence binding by the SRP M domain. (A) The SRP signal sequence (SSS) of a substrate (purple) is harboured within the hydrophobic groove formed by the SRP M domain (gray). The SRP RNA is colored in blue. The endogenous cysteine at position 406 is colored red. For disulphide cross-linking studies the methionine residue located towards the hydrophobic groove at position 423 (red) was substituted with a cysteine residue. The amino acid residues 341, 376 (green) and 426 (red) were cross-linked to σ^{32} (Image created with PyMOL 2.3.0, PDB 5GAD). (B) Docking of a hypothetical signal sequence (SSS) (yellow) into the hydrophobic groove (orange) of the SRP M domain (gray) indicating the proposed interaction of the N-terminal basic amino acid residues of the SSS (dark blue) with the phosphate groups (red) of the SRP RNA (blue) (Rupert and Ferré-D'amaré 2000).

The KdpD22-48 peptide as well as the SciP1-27 G16C and SciP54-85 peptides were cross-linked to the cysteine residue at position 423, located towards the hydrophobic groove (chapter 2 and 4). This indicates that also the novel cytoplasmic SRP signal sequences are harboured in the groove. This is in good agreement with a previous

study on the SRP recognition sequence of σ^{32} (Miyazaki et al. 2016). With *in vivo* photo cross-linking it was shown that also the SRP recognition sequence of σ^{32} was located in the hydrophobic groove of the SRP M domain. The SRP recognition site of σ^{32} was cross-linked to the amino acid residues 341, 376 and 426 of SRP, lining the hydrophobic groove (Miyazaki et al. 2016).

For the KdpD peptide a further cross-link was observed with the endogenous cysteine residue at position 406 of SRP (chapter 2). This was first thought to be unlikely, since the distance is to large for a disulphide bond by the oxidizing reagent copper phenanthroline. However, it was shown that signal sequence binding by SRP results in conformational changes of the *M. jannaschii* SRP, which results in a flip of the NG domain via the flexible GM linker and a reorientation of the M domain (Hainzl et al. 2011; Hainzl and Sauer-Eriksson 2015). Therefore, it is conceivable that the cysteine residue at position 406 moves more towards the signal sequence allowing the formation of a disulphide bond. In addition, the modelling of a hypothetical signal sequence with a N-terminal basic region into the hydrophobic groove suggest that the N-terminal part of the signal sequence comes in close contact to the RNA binding motif, including the amino acid resiude 406 (Fig. 31B). A similar orientation of the signal sequence of KdpD in the hydrophobic groove, like the hypothetical one, may allows a cross-link to the cysteine at position 406. However, this hypothesis has to be analyzed with further cross-linking studies.

As mentioned, KdpD22-48 and SciP1-27 fold amphiphilic helices. This may promote SRP binding as is suggested by the structure of the hydrophobic groove (Fig. 31). Probably, the hydrophobic half of the amphiphilic helix is surrounded by the hydrophobic residues of SRP that line up the groove. In contrast, the hydrophilic half of the amphiphilic helix is located towards the opening facing the cytoplasm.

The binding of SRP signal sequences by the hydrophobic groove is a crucial step in the SRP targeting cycle, since signal sequence binding induces conformational changes in SRP (Hainzl et al. 2011). These conformational changes prime SRP for the interaction with its receptor FtsY (Jomaa et al. 2016). Therefore, it seems reasonable, that also the novel SRP signal sequences are bound by the SRP hydrophobic groove to initiate the next targeting step, the interaction with the membrane-bound FtsY. In conclusion, the results obtained in this study show that SRP uses the same binding mechanism for transmembrane located and cytoplasmic signal sequences.

5.6 Interplay of SRP and YidC during protein insertion

After the co-translational SRP-dependent targeting of inner membrane proteins to the membrane-associated receptor FtsY, the proteins are normally inserted by the Sec translocase and / or the YidC insertase (Cross et al. 2009; Petriman et al. 2018). Membrane-associated FtsY adopts an open conformation during the interaction with SecYEG, preparing it for the interaction with the nascent-chain-bound SRP (Draycheva et al. 2016). The interaction of SRP and FtsY then results in coordinated GTP hydrolysis and the dissociation of SRP from the nascent chain (Connolly et al. 1991; Egea et al. 2004). However, in the targeting and insertion pathway of KdpD no translocase or insertase is involved (Facey and Kuhn 2003), which was thought to be a crucial step for SRP-FtsY interaction and further nascent chain release. How this mechanism is then bypassed and how the KdpD nascent chain is released from SRP is not yet understood.

In contrast to KdpD, SciP is inserted into the inner membrane by the YidC insertase (Aschtgen et al. 2012). Cysteine-accessibility assays with SciP and different YidC mutant proteins demonstrated that the second helix of the first cytoplasmic loop (CH2) of YidC is important for the insertion of SciP (Spann et al. 2018). The C1 domain forms a helical hairpin structure at the entrance of the hydrophilic groove of YidC and is essential for YidC function. The importance of the first cytoplasmic loop (C1) was also shown with two other YidC-dependent model proteins, the M13 procoat and the Pf3 coat protein fused to leader peptidase. The insertion of both proteins was highly impaired in cells expressing the YidC CH2 mutant (Chen et al. 2014). Previous studies have shown that YidC directly interacts with the components of the SRP system via the C1 loop (Petriman et al. 2018). So it is assumed that the C1 loop could play a role in the recruitment of the YidC substrates (Petriman et al. 2018). This is a possible explanation why the insertion of SciP is impaired in the CH2 mutant due to the fact that the SRP-YidC interaction is restricted. However, the YidC substrates M13 procoat and Pf3 coat are not SRP-dependent and the insertion is prohibited in the YidC CH2 mutant as well. Further studies are needed to adress the essential role of the C1 domain in the YidC dependent insertion mechanism of SciP and other YidC substrates.

YidC not only interacts directly with SRP and FtsY but also with RNCs exposing a YidC substrate. In contrast, non-translating ribosomes only show a weak binding affinity to YidC (Welte et al. 2012; Kedrov et al. 2013; Kedrov et al. 2016). The interaction with

RNCs could already be observed at the beginning of translation where only about 30 amino acids of the nascent chain are exposed. However, the maximal binding affinity of YidC to nascent chains was reached as soon as the first TMD of the YidC substrate, the subunit c of F1Fo ATP synthase, was fully exposed (Kedrov et al. 2016). This could be the explanation for the interplay between SRP and YidC in the targeting and insertion mechanism of SciP. The co-translationally targeted nascent chain of SciP is received by YidC at an early stage of translation preparing it for the post-translational binding to the C-terminal TMD.

As shown for SciP also the C-tail anchored proteins DjlC and Flk are assumed to be inserted by YidC, however these conclusions only arise from localisation and cross-linking studies with mutant mNeonGreen fusion proteins (Peschke et al. 2018). To be sure, the insertion of wildtype DjlC and Flk should directly be analyzed in YidC depleted cells. Nevertheless, there is evidence that YidC is a universal component of the insertion pathway of C-tail anchored proteins in *E. coli.* In eukaryotes most TA proteins use the Get pathway to be inserted into the membrane of the endoplasmatic reticulum (Schuldiner et al. 2008). A recent study now suggested, that the Get1 protein is a structural homolog to YidC and therefore belongs to the Oxa1 superfamiliy (Anghel et al. 2017). It seems that bacteria lack the Get system but have YidC instead which probably is the ancestor of Get1 and was adapted during evolution to insert eukaryotic C-tail anchored proteins in the ER (Anghel et al. 2017). In prokaryotes, it is therefore likely that YidC still fulfills the function of C-tail anchored protein insertion.

Previous results on other YidC-only substrates indicate that they all share the feature of only short periplasmic domains with a small number of hydrophilic residues and the absence of charged clusters (Soman et al. 2014; Shanmugam et al. 2019). This restriction could be due to the size and capacity of the hydrophilic cavity of YidC. SciP with a C-tail of 11 amino acids including only 3 hydrophilic amino acids and a neutral net charge was shown to be inserted by YidC (chapter 3, Table 2). Extending the tail to 20 amino acids including 4 hydrophilic amino acids does not change the dependency on YidC (chapter 3, Table 2). However, an extension to 21 amino acids including 5 hydrophilic amino acids led to a dependency on SecYEG (chapter 3, Table 2). Interestingly, the increase to 22 amino acids with 7 hydrophilic amino acids led to an inefficient translocation of the C-tail (chapter 3, Table 2). This leads to the conclusion, that also in the case of SciP the C-tail length and the number of hydrophilic amino acids is a determinant for YidC and / or SecYEG dependent insertion. A further increase of

the C-tail to 26 or 28 amino acids with 6 or 10 hydrophilic amino acids, respectively, completely abolished the translocation of the C-tail (chapter 3, Table 2). This is in agreement with a previous study on the insertion of a M13-procoat leader peptidase fusion protein. Beyond a critical length of the periplasmic part, these fusion proteins could not be inserted anymore (Shanmugam et al. 2019). Longer periplasmatic loops, observed for example in the proteins MalF or MBP need the additional help of SecA for translocation (Wagner et al. 2008; Shanmugam et al. 2019). It is assumed that downstream reverse signals are necessary for the recruitment of SecA to assist in the translocation of longer periplasmatic domains (Wagner et al. 2008; Shanmugam et al. 2019). Since SciP is SecA independent, no downstream signals for SecA recruitment exist and so it is assumed that this is the reason why SciP proteins with longer tails cannot be inserted anymore. However, the exact mechanism of YidC-dependent insertion of the C-tail anchored protein SciP and also the interplay between SRP and YidC have to be studied more detailed in the future.

Table 2: Characteristics of the different C-tails of the protein SciP and their corresponding insertion capability and dependency on YidC and / or SecYEG. Hydrophilic amino acids are coloured green, charged residues are coloured red.

Protein	Sequence C-terminal tail	Number of amino acids			Net charge	Insertion	Dependency
		total	charged	hydrophilic			
SciP	ADQVARLTGQL	11	2	3	0	++	YidC
SciP-C	ADQVARLTGQLC	12	2	3	0	++	YidC
SciP20	ADQVARLTGLMASMTGGCMG	20	2	4	0	++	YidC
SciP21	ADQVARLTGQLMASMTGGCMG	21	2	5	0	++	YidC / SecYEG
SciP22	ADQVARLTGQLMASMTGGQQMG	22	2	7	0	+	YidC / SecYEG
SciP26	ADQVARLTGQLASGSYPYDVPDYAQL	26	4	6	-3	-	-
SciP28	ADQVARLTGQLMASMTGGQQMGSGSAQL	28	2	10	0	-	-

Taken together, the dependency of the C-tail anchored protein SciP on YidC was confirmed in this work. In addition, the studies employing extensions of the periplasmic C-tails gained more information about the conditions needed for a YidC and / or SecYEG dependent insertion. However, more precise analyses are necessary to understand the mechanism of YidC dependent insertion of SciP and the interplay of YidC with the SRP system during its insertion.

SUMMARY & REFERENCES

6

Zusammenfassung

In *E. coli* werden die meisten integralen Proteine co-translational durch das universell konservierte SRP an die Membran geleitet (Bernstein et al. 1989; Valent et al. 1998; Schibich et al. 2016). SRP scannt translatierende Ribosomen und bindet mit hoher Affinität an eine exponierte SRP Signalsequenz der naszierenden Proteinkette (Bornemann et al. 2008; Holtkamp et al. 2012; Saraogi et al. 2014). Nach dem Transport zum Membran-assoziierten SRP Rezeptor FtsY wird das naszierende Membranprotein an die Sec Translokase oder YidC zur Insertion in den Bilayer übergeben (Miller et al. 1994; Cross et al. 2009; Welte et al. 2012; Akopian et al. 2013). Allgemein sind die Transport- und Insertionswege von Innenmembranproteinen in *E. coli* bereits sehr gut untersucht. Zu einer speziellen Proteinklasse, mit nur wenigen Proteinen in *E. coli*, gehören die sogenannten „C-tail anchored“ Proteine, deren Insertionsweg in Prokaryoten bislang jedoch unbekannt ist. Um den Insertionsweg zu erforschen, wurde das „C-tail anchored“ Protein SciP als Modellprotein ausgewählt. SciP aus dem enteroaggregativen *E. coli* stellt eine strukturelle Komponente des Typ 6 Sekretionssystems dar und besitzt eine Transmembrandomäne (TMD) am C-Terminus von Aminosäure 184 bis 206. Somit weist SciP eine große N-terminale cytoplasmatische Domäne von 183 Aminosäuren auf. Ein weiteres *E. coli* Protein, das Kalium-Sensorprotein KdpD, teilt mit SciP die Gemeinsamkeit einer großen N-terminalen cytoplasmatischen Domäne. KdpD ist ein vier-spänniges Membranprotein, bei dem die erste TMD nach 400 Aminosäuren beginnt. Zunächst wurde angenommen, dass die spezielle Topologie dieser beiden Proteine einen co-translationalen SRP-abhängigen Transport ausschließt. Dennoch konnte gezeigt werden, dass KdpD durch ein cytoplasmatisches Transportsignal zwischen den Aminosäuren 22 und 48 durch SRP erkannt und co-translational an die Membran gebracht wird (Maier et al. 2008).

In dieser Studie wurde gezeigt, dass ebenfalls das „C-tail anchored“ Protein SciP bereits zu Beginn der Translation durch SRP an die Membran geleitet wird. Mit Hilfe der Fluoreszenzmikroskopie und sfGFP-SciP Fusionskonstrukten wurden zwei kurze hydrophobe Bereiche in der N-terminalen cytoplasmatischen Domäne (Aminosäuren 12-20 und 62-71) identifiziert, die wichtig für den Transport sind. Mit translations-pausierten Ribosomen, die jeweils eines der beiden Transportsignale von SciP exponieren, konnte durch Microscale Thermophorese entschlüsselt werden, dass

beide Transportsignale während der Translation von SRP und einem SRP-FtsY Komplex gebunden werden. Über Cystein-Modifikationsstudien wurde bestätigt, dass SciP das erste beschriebene Protein mit zwei SRP-Signalen darstellt, da die Deletion eines der beiden Signale durch die Anwesenheit des jeweils anderen *in vivo* ausgeglichen werden kann.

Um die ausschlaggebenden Merkmale der neuartigen cytoplasmatischen SRP-Signalsequenzen von KdpD und SciP zu identifizieren, wurden die Signalsequenzen modifiziert. Die Auswirkungen wurden mit Hilfe von sfGFP-Fusionskonstrukten und der Fluoreszenzmikroskopie sowie mit der Microscale Thermophorese und künstlich pausierten Ribosomen analysiert. Damit wurde gezeigt, dass auch die neuartigen Signalsequenzen ein kritisches Hydrophobizitätslevel überschreiten müssen, um von SRP erkannt und gebunden zu werden und daraufhin zum Transport von sfGFP an die Membran führen. Desweiteren konnte gezeigt werden, dass drei positiv geladene Aminosäuren in der SRP Signalsequenz von KdpD die SRP Bindung verstärken. Um den Bindemechanismus von SRP an die Signalsequenzen zu charakterisieren, wurden *in vitro* Disulfid-Crosslinking-Studien mit synthetisierten KdpD22-48, SciP1-27 und SciP54-85 Peptiden durchgeführt. Alle drei Peptide konnten an die hydrophobe Bindetasche der M-Domäne von SRP gebunden werden, was mit der Bindung anderer SRP Substrate übereinstimmt. In dieser Arbeit wurde gezeigt, dass die SRP Bindung nicht auf eine TMD eines Proteins beschränkt ist, sondern dass auch kurze hydrophobe Bereiche in cytoplasmatischen Domänen als SRP Signale agieren können.

Cystein-Modifikationsstudien mit dem Protein SciP zeigten, dass nicht nur SRP sondern auch YidC am Insertionsprozess beteiligt ist. Mit nur 11 Aminosäuren in der periplasmatischen Domäne stimmt diese Charakteristik mit anderen YidC Substraten überein. Durch die Verlängerung der periplasmatischen Domäne von SciP zeigte sich, dass es eine kritische Länge von 20 Aminosäuren gibt und dass die Überschreitung die Insertion Sec-abhängig werden lässt. Anhand der Studien mit den verlängerten periplasmatischen Domänen von SciP konnten allgemeine Informationen über den YidC Insertionsprozess erlangt werden. Die Ergebnisse mit SciP sind die ersten Anhaltspunkte, wie die Insertion der speziellen „C-tail anchored“ Proteine in *E. coli* stattfinden könnte. Es wird angenommen, dass das SRP System und die YidC Insertase das Fehlen des eukaryotischen Get-Systems, welches eukaryotische „tail-anchored“ Proteine inseriert, ersetzen kann.

Summary

In *E. coli,* most inner membrane proteins are targeted in a co-translational manner by the universally conserved signal recognition particle (Bernstein et al. 1989; Valent et al. 1998; Schibich et al. 2016). SRP scans the translating ribosomes and binds with high affinity to an exposed SRP signal sequence, present in the nascent chain (Bornemann et al. 2008; Holtkamp et al. 2012; Saraogi et al. 2014). After targeting to the membrane-associated SRP receptor FtsY, the nascent membrane protein is forwarded to the Sec translocase or to the YidC insertase to be integrated into the bilayer (Miller et al. 1994; Cross et al. 2009; Welte et al. 2012; Akopian et al. 2013).
In general, the targeting and insertion pathways of inner membrane proteins in *E. coli* are already well studied. However, there is a special class of proteins, the C-tail anchored proteins with only a few members in *E. coli*, whose insertion mechanisms are unknown in prokaryotes to date. To study those insertion mechanisms, the C-tail anchored protein SciP was used as a model protein. SciP from the enteroaggregative *E. coli* is a structural component of the type 6 secretion system and contains a transmembrane domain (TMD) at the extreme C-terminal part from amino acid 184 to 206. This results in a large N-terminal cytoplasmic domain of 183 amino acids. In *E. coli,* there is another protein, the potassium sensor protein KdpD which shares with SciP the commonality of a large N-terminal cytoplasmic domain. KdpD is a four-spanning membrane protein with the first TMD starting at amino acid position 400. For both proteins, with the TMD being located far away from the cytoplasmic N-terminal part, it was thought that they cannot use the co-translational SRP pathway. However, it was shown that KdpD is targeted co-translationally by SRP and a cytoplasmic targeting signal located between amino acids 22-48 was identified (Maier et al. 2008).

In this study it was shown that the C-tail anchored protein SciP is also targeted early during translation by SRP. With fluorescence microscopy studies and sfGFP-SciP fusion constructs, two short hydrophobic regions in the N-terminal cytoplasmic domain (amino acids 12-20 and 62-71) were identified as being important for membrane targeting. With artificially stalled ribosomes exposing each of the targeting signal, microscale thermophoresis meausurements decoded that both signals bind to SRP and to a preincubated SRP-FtsY complex, mimicking the next targeting step. Cysteine-accessibilty assays demonstrated that SciP is the first described protein with two

targeting signals since the deletion of one of the hydrophobic regions was compensated by the other remaining one *in vivo*.

To decipher the crucial features of the novel cytoplasmic SRP signal sequences of KdpD and SciP alterations in the signal sequences were analyzed with fluorescence microscopy using sfGFP fusion constructs and microscale thermophoresis measurements using stalled ribosomes. These studies revealed that the novel signal sequences have to exceed a threshold level of hydrophobicity to be recognized and bound by SRP and target sfGFP to the membrane. In addition, three positively charged amino acids in the KdpD SRP signal sequence were identified to promote SRP binding. To characterize the binding mechanism of SRP to the signal sequences, *in vitro* disulphide cross-linking studies with synthesized KdpD22-48, SciP1-27 and SciP54-85 peptides were performed. All three peptides could be cross-linked to the hydrophobic groove of SRP formed by the M domain, which correlates with the binding of SRP to other substrates. Taken together, the results show that SRP binding is not limited to the TMDs of proteins. SRP is also able to recognize short hydrophobic stretches in the cytoplasmic domain of inner membrane proteins.

Cysteine-accessibility assays with the C-tail anchored protein SciP decoded that not only SRP is involved in the delivery pathway but also the insertase YidC. With only 11 amino acids in the periplasmic domain SciP matches with the characteristics of other known YidC only substrates. By extending the C-tail of SciP it was found out that a critical length of 20 amino acids exists and that the exceed of this limit makes the insertion of SciP dependent on the Sec translocase. The studies with the extended C-tails of SciP helped to gain more general information about the YidC dependent insertion of proteins. The results obtained with the protein SciP are first indications about how the insertion of C-tail anchored proteins occurs in *E. coli*. It is assumed that the SRP system and the insertase YidC compensate the absence of the eukaryotic Get system, responsible for the insertion of eukaryotic tail-anchored proteins.

References

Abdallah AM, van Gey Pittius NC, Champion PAD, Cox J, Luirink J, Vandenbroucke-Grauls CMJE, Appelmelk BJ, Bitter W (2007): Type VII secretion--mycobacteria show the way. In *Nat Rev Microbiol* 5 (11), 883–891.

Adams H, Scotti PA, Cock H de, Luirink J, Tommassen J (2002): The presence of a helix breaker in the hydrophobic core of signal sequences of secretory proteins prevents recognition by the signal-recognition particle in Escherichia coli. In *Eur J Biochem* 269 (22), 5564–5571.

Akopian D, Shen K, Zhang X, Shan S-o (2013): Signal recognition particle: an essential protein-targeting machine. In *Annu Rev Biochem* 82, 693–721.

Alami M, Lüke I, Deitermann S, Eisner G, Koch H-G, Brunner J, Müller M (2003): Differential interactions between a twin-arginine signal peptide and its translocase in Escherichia coli. In *Mol Cell* 12 (4), 937–946.

Angelini S, Deitermann S, Koch H-G (2005): FtsY, the bacterial signal-recognition particle receptor, interacts functionally and physically with the SecYEG translocon. In *EMBO Rep* 6 (5), 476–481.

Anghel SA, McGilvray PT, Hegde RS, Keenan RJ (2017): Identification of Oxa1 Homologs Operating in the Eukaryotic Endoplasmic Reticulum. In *Cell Rep* 21 (13), 3708–3716.

Ariosa A, Lee JH, Wang S, Saraogi I, Shan S-o (2015): Regulation by a chaperone improves substrate selectivity during cotranslational protein targeting. In *Proceedings of the National Academy of Sciences* 112 (25), E3169-78.

Aschtgen M-S, Gavioli M, Dessen A, Lloubès R, Cascales E (2010a): The SciZ protein anchors the enteroaggregative Escherichia coli Type VI secretion system to the cell wall. In *Mol Microbiol* 75 (4), 886–899.

Aschtgen M-S, Thomas MS, Cascales E (2010b): Anchoring the type VI secretion system to the peptidoglycan: TssL, TagL, TagP… what else? In *Virulence* 1 (6), 535–540.

Aschtgen M-S, Zoued A, Lloubès R, Journet L, Cascales E (2012): The C-tail anchored TssL subunit, an essential protein of the enteroaggregative Escherichia coli Sci-1 Type VI secretion system, is inserted by YidC. In *Microbiologyopen* 1 (1), 71–82.

Ataide SF, Schmitz N, Shen K, Ke A, Shan S-o, Doudna JA, Ban N (2011): The crystal structure of the signal recognition particle in complex with its receptor. In *Science* 331 (6019), 881–886.

Baaske P, Wienken CJ, Reineck P, Duhr S, Braun D (2010): Optical thermophoresis for quantifying the buffer dependence of aptamer binding. In *Angew Chem Int Ed Engl* 49 (12), 2238–2241.

Ballal A, Basu B, Apte SK (2007): The Kdp-ATPase system and its regulation. In *J Biosci* 32 (3), 559–568.

Basler M, Pilhofer M, Henderson GP, Jensen GJ, Mekalanos JJ (2012): Type VI secretion requires a dynamic contractile phage tail-like structure. In *Nature* 483 (7388), 182–186.

Bassford P, Beckwith J, Ito K, Kumamoto C, Mizushima S, Oliver D, Randall L, Silhavy T, Tai PC, Wickner B (1991): The primary pathway of protein export in E. coli. In *Cell* 65 (3), 367–368.

Batey RT (2000): Crystal Structure of the Ribonucleoprotein Core of the Signal Recognition Particle. In *Science* 287 (5456), 1232–1239.

Beck K, Eisner G, Trescher D, Dalbey RE, Brunner J, Müller M (2001): YidC, an assembly site for polytopic Escherichia coli membrane proteins located in immediate proximity to the SecYE translocon and lipids. In *EMBO Rep* 2 (8), 709–714.

Bernstein HD, Poritz MA, Strub K, Hoben PJ, Brenner S, Walter P (1989): Model for signal sequence recognition from amino-acid sequence of 54K subunit of signal recognition particle. In *Nature* 340 (6233), 482–486.

Bernstein HD, Zopf D, Freymann DM, Walter P (1993): Functional substitution of the signal recognition particle 54-kDa subunit by its Escherichia coli homolog. In *Proceedings of the National Academy of Sciences* 90 (11), 5229–5233.

Bischoff L, Wickles S, Berninghausen O, van der Sluis EO, Beckmann R (2014): Visualization of a polytopic membrane protein during SecY-mediated membrane insertion. In *Nat Commun* 5, 4103.

Blobel G (1975): Transfer of proteins across membranes. I. Presence of proteolytically processed and unprocessed nascent immunoglobulin light chains on membrane-bound ribosomes of murine myeloma. In *J Cell Biol* 67 (3), 835–851.

Blobel G, Dobberstein B (1975): Transfer of proteins across membranes. II. Reconstitution of functional rough microsomes from heterologous components. In *J Cell Biol* 67 (3), 852–862.

Blobel G, Sabatini DD (1971): Ribosome-Membrane Interaction in Eukaryotic Cells. In Lionel A. Manson (Ed.): Biomembranes, vol. 56. Boston, MA: Springer US, 193–195.

Bönemann G, Pietrosiuk A, Diemand A, Zentgraf H, Mogk A (2009): Remodelling of VipA/VipB tubules by ClpV-mediated threading is crucial for type VI protein secretion. In *EMBO J* 28 (4), 315–325.

Bonnefoy N, Chalvet F, Hamel P, Slonimski PP, Dujardin G (1994): OXA1, a Saccharomyces cerevisiae nuclear gene whose sequence is conserved from prokaryotes to eukaryotes controls cytochrome oxidase biogenesis. In *J Mol Biol* 239 (2), 201–212.

Booth IR (1985): Regulation of cytoplasmic pH in bacteria. In *Microbiol Rev* 49 (4), 359–378.

Borgese N, Righi M (2010): Remote origins of tail-anchored proteins. In *Traffic* 11 (7), 877–885.

Bornemann T, Holtkamp W, Wintermeyer W (2014): Interplay between trigger factor and other protein biogenesis factors on the ribosome. In *Nat Commun* 5, 4180.

Bornemann T, Jöckel J, Rodnina MV, Wintermeyer W (2008): Signal sequence-independent membrane targeting of ribosomes containing short nascent peptides within the exit tunnel. In *Nat Struct Mol Biol* 15 (5), 494–499.

Boy D, Koch H-G (2009): Visualization of distinct entities of the SecYEG translocon during translocation and integration of bacterial proteins. In *Mol Biol Cell* 20 (6), 1804–1815.

Boyer F, Fichant G, Berthod J, Vandenbrouck Y, Attree I (2009): Dissecting the bacterial type VI secretion system by a genome wide in silico analysis: what can be learned from available microbial genomic resources? In *BMC Genomics* 10, 104.

Brambillasca S, Yabal M, Makarow M, Borgese N (2006): Unassisted translocation of large polypeptide domains across phospholipid bilayers. In *J Cell Biol* 175 (5), 767–777.

Brambillasca S, Yabal M, Soffientini P, Stefanovic S, Makarow M, Hegde RS, Borgese N (2005): Transmembrane topogenesis of a tail-anchored protein is modulated by membrane lipid composition. In *EMBO J* 24 (14), 2533–2542.

Breyton C, Haase W, Rapoport TA, Kühlbrandt W, Collinson I (2002): Three-dimensional structure of the bacterial protein-translocation complex SecYEG. In *Nature* 418 (6898), 662–665.

Brown S, Fournier MJ (1984): The 4.5 S RNA gene of Escherichia coli is essential for cell growth. In *J Mol Biol* 178 (3), 533–550.

Brunet YR, Hénin J, Celia H, Cascales E (2014): Type VI secretion and bacteriophage tail tubes share a common assembly pathway. In *EMBO Rep* 15 (3), 315–321.

Buskiewicz I, Peske F, Wieden H-J, Gryczynski I, Rodnina MV, Wintermeyer W (2005): Conformations of the signal recognition particle protein Ffh from Escherichia coli as determined by FRET. In *J Mol Biol* 351 (2), 417–430.

Cabelli RJ, Dolan KM, Qian LP, Oliver DB (1991): Characterization of membrane-associated and soluble states of SecA protein from wild-type and SecA51(TS) mutant strains of Escherichia coli. In *J Biol Chem* 266 (36), 24420–24427.

Cascales E (2008): The type VI secretion toolkit. In *EMBO Rep* 9 (8), 735–741.

Chatzi KE, Sardis MF, Economou A, Karamanou S (2014): SecA-mediated targeting and translocation of secretory proteins. In *Biochim Biophys Acta* 1843 (8), 1466–1474.

Chen Y, Capponi S, Zhu L, Gellenbeck P, Freites JA, White SH, Dalbey RE (2017): YidC Insertase of Escherichia coli: Water Accessibility and Membrane Shaping. In *Structure* 25 (9), 1403-1414.e3.

Chen Y, Dalbey RE (2018): Oxa1 Superfamily: New Members Found in the ER. In *Trends Biochem Sci* 43 (3), 151–153.

Chen Y, Soman R, Shanmugam SK, Kuhn A, Dalbey RE (2014): The role of the strictly conserved positively charged residue differs among the Gram-positive, Gram-negative, and chloroplast YidC homologs. In *J Biol Chem* 289 (51), 35656–35667.

Cho H, Shan S-o (2018): Substrate relay in an Hsp70-cochaperone cascade safeguards tail-anchored membrane protein targeting. In *EMBO J* 37 (16).

Cline K (2015): Mechanistic Aspects of Folded Protein Transport by the Twin Arginine Translocase (Tat). In *J Biol Chem* 290 (27), 16530–16538.

Colombo SF, Longhi R, Borgese N (2009): The role of cytosolic proteins in the insertion of tail-anchored proteins into phospholipid bilayers. In *J Cell Sci* 122 (Pt 14), 2383–2392.

Connolly T, Rapiejko PJ, Gilmore R (1991): Requirement of GTP hydrolysis for dissociation of the signal recognition particle from its receptor. In *Science* 252 (5009), 1171–1173.

Costa TRD, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, Waksman G (2015): Secretion systems in Gram-negative bacteria: structural and mechanistic insights. In *Nat Rev Microbiol* 13 (6), 343–359.

Crane JM, Randall LL (2017): The Sec System: Protein Export in Escherichia coli. In *EcoSal Plus* 7 (2).

Cross BCS, Sinning I, Luirink J, High S (2009): Delivering proteins for export from the cytosol. In *Nat Rev Mol Cell Biol* 10 (4), 255–264.

Dalbey RE, Kuhn A (2012): Protein traffic in Gram-negative bacteria--how exported and secreted proteins find their way. In *FEMS Microbiol Rev* 36 (6), 1023–1045.

Dalbey RE, Kuhn A (2014): How YidC inserts and folds proteins across a membrane. In *Nat Struct Mol Biol* 21 (5), 435–436.

Dalbey RE, Kuhn A, Zhu L, Kiefer D (2014): The membrane insertase YidC. In *Biochim Biophys Acta* 1843 (8), 1489–1496.

Danese PN, Silhavy TJ (1998): Targeting and assembly of periplasmic and outer-membrane proteins in Escherichia coli. In *Annu Rev Genet* 32, 59–94.

Denic V, Dötsch V, Sinning I (2013): Endoplasmic reticulum targeting and insertion of tail-anchored membrane proteins by the GET pathway. In *Cold Spring Harb Perspect Biol* 5 (8), a013334.

Denks K, Sliwinski N, Erichsen V, Borodkina B, Origi A, Koch H-G (2017): The signal recognition particle contacts uL23 and scans substrate translation inside the ribosomal tunnel. In *Nat Microbiol* 2, 16265.

Denks K, Vogt A, Sachelaru I, Petriman N-A, Kudva R, Koch H-G (2014): The Sec translocon mediated protein transport in prokaryotes and eukaryotes. In *Mol Membr Biol* 31 (2-3), 58–84.

Dosch DC, Helmer GL, Sutton SH, Salvacion FF, Epstein W (1991): Genetic analysis of potassium transport loci in Escherichia coli: evidence for three constitutive systems mediating uptake potassium. In *J Bacteriol* 173 (2), 687–696.

Doudna JA, Batey RT (2004): Structural insights into the signal recognition particle. In *Annu Rev Biochem* 73, 539–557.

Draycheva A, Bornemann T, Ryazanov S, Lakomek N-A, Wintermeyer W (2016): The bacterial SRP receptor, FtsY, is activated on binding to the translocon. In *Mol Microbiol* 102 (1), 152–167.

Draycheva A, Lee S, Wintermeyer W (2018): Cotranslational protein targeting to the membrane: Nascent-chain transfer in a quaternary complex formed at the translocon. In *Sci Rep* 8 (1), 9922.

Driessen AJM, Nouwen N (2008): Protein translocation across the bacterial cytoplasmic membrane. In *Annu Rev Biochem* 77, 643–667.

Duong F, Wickner W (1997a): Distinct catalytic roles of the SecYE, SecG and SecDFyajC subunits of preprotein translocase holoenzyme. In *EMBO J* 16 (10), 2756–2768.

Duong F, Wickner W (1997b): The SecDFyajC domain of preprotein translocase controls preprotein movement by regulating SecA membrane cycling. In *EMBO J* 16 (16), 4871–4879.

Durand E, van Nguyen S, Zoued A, Logger L, Péhau-Arnaudet G, Aschtgen M-S, Spinelli S, Desmyter A, Bardiaux B, Dujeancourt A, Roussel A, Cambillau C, Cascales E, Fronzes R (2015): Biogenesis and structure of a type VI secretion membrane core complex. In *Nature* 523 (7562), 555–560.

Durand E, Zoued A, Spinelli S, Watson PJH, Aschtgen M-S, Journet L, Cambillau C, Cascales E (2012): Structural characterization and oligomerization of the TssL protein, a component shared by bacterial type VI and type IVb secretion systems. In *J Biol Chem* 287 (17), 14157–14168.

Economou A, Wickner W (1994): SecA promotes preprotein translocation by undergoing ATP-driven cycles of membrane insertion and deinsertion. In *Cell* 78 (5), 835–843.

Egea PF, Shan S-o, Napetschnig J, Savage DF, Walter P, Stroud RM (2004): Substrate twinning activates the signal recognition particle and its receptor. In *Nature* 427 (6971), 215–221.

Eimer E, Fröbel J, Blümmel A-S, Müller M (2015): TatE as a Regular Constituent of Bacterial Twin-arginine Protein Translocases. In *J Biol Chem* 290 (49), 29281–29289.

Eisenhawer M, Cattarinussi S, Kuhn A, Vogel H (2001): Fluorescence resonance energy transfer shows a close helix-helix distance in the transmembrane M13 procoat protein. In *Biochemistry* 40 (41), 12321–12328.

Eitan A, Bibi E (2004): The core Escherichia coli signal recognition particle receptor contains only the N and G domains of FtsY. In *J Bacteriol* 186 (8), 2492–2494.

Epstein W (1986): Osmoregulation by potassium transport in Escherichia coli. In *FEMS Microbiol Lett* 39 (1-2), 73–78.

Estrozi LF, Boehringer D, Shan S-o, Ban N, Schaffitzel C (2011): Cryo-EM structure of the E. coli translating ribosome in complex with SRP and its receptor. In *Nat Struct Mol Biol* 18 (1), 88–90.

Facey SJ, Kuhn A (2003): The sensor protein KdpD inserts into the Escherichia coli membrane independent of the Sec translocase and YidC. In *Eur J Biochem* 270 (8), 1724–1734.

Fedyukina DV, Cavagnero S (2011): Protein folding at the exit tunnel. In *Annu Rev Biophys* 40, 337–359.

Fekkes P, Driessen AJ (1999): Protein targeting to the bacterial cytoplasmic membrane. In *Microbiol Mol Biol Rev* 63 (1), 161–173.

Fekkes P, Wit JG de, van der Wolk JP, Kimsey HH, Kumamoto CA, Driessen AJ (1998): Preprotein transfer to the Escherichia coli translocase requires the co-operative binding of SecB and the signal sequence to SecA. In *Mol Microbiol* 29 (5), 1179–1190.

Filloux A, Hachani A, Bleves S (2008): The bacterial type VI secretion machine: yet another player for protein transport across membranes. In *Microbiology (Reading, Engl)* 154 (Pt 6), 1570–1583.

Fluman N, Navon S, Bibi E, Pilpel Y (2014): mRNA-programmed translation pauses in the targeting of E. coli membrane proteins. In *Elife* 3.

Focia PJ, Shepotinovskaya IV, Seidler JA, Freymann DM (2004): Heterodimeric GTPase core of the SRP targeting complex. In *Science* 303 (5656), 373–377.

Freymann DM, Keenan RJ, Stroud RM, Walter P (1997): Structure of the conserved GTPase domain of the signal recognition particle. In *Nature* 385 (6614), 361–364.

Funes S, Kauff F, van der Sluis EO, Ott M, Herrmann JM (2011): Evolution of YidC/Oxa1/Alb3 insertases: three independent gene duplications followed by functional specialization in bacteria, mitochondria and chloroplasts. In *Biol Chem* 392 (1-2), 13–19.

Gasper R, Meyer S, Gotthardt K, Sirajuddin M, Wittinghofer A (2009): It takes two to tango: regulation of G proteins by dimerization. In *Nat Rev Mol Cell Biol* 10 (6), 423–429.

Geller BL, Movva NR, Wickner W (1986): Both ATP and the electrochemical potential are required for optimal assembly of pro-OmpA into Escherichia coli inner membrane vesicles. In *Proceedings of the National Academy of Sciences* 83 (12), 4219–4222.

Geller BL, Wickner W (1985): M13 procoat inserts into liposomes in the absence of other membrane proteins. In *J Biol Chem* 260 (24), 13281–13285.

Gier J-WL de, Mansournia P, Valent QA, Phillips GJ, Luirink J, Heijne G von (1996): Assembly of a cytoplasmic membrane protein in Escherichia coli is dependent on the signal recognition particle. In *FEBS Lett* 399 (3), 307–309.

Gier J-WL de, Scotti PA, Sääf A, Valent QA, Kuhn A, Luirink J, Heijne G von (1998): Differential use of the signal recognition particle translocase targeting pathway for inner membrane protein assembly in Escherichia coli. In *Proceedings of the National Academy of Sciences* 95 (25), 14646–14651.

Gill DR, Hatfull GF, Salmond GP (1986): A new cell division operon in Escherichia coli. In *Mol Gen Genet* 205 (1), 134–145.

Gilmore R (1982): Protein translocation across the endoplasmic reticulum. II. Isolation and characterization of the signal recognition particle receptor. In *J Cell Biol* 95 (2), 470–477.

Gold VAM, Robson A, Bao H, Romantsov T, Duong F, Collinson I (2010): The action of cardiolipin on the bacterial translocon. In *Proceedings of the National Academy of Sciences* 107 (22), 10044–10049.

Gong F, Yanofsky C (2001): Reproducing tna operon regulation in vitro in an S-30 system. Tryptophan induction inhibits cleavage of TnaC peptidyl-tRNA. In *J Biol Chem* 276 (3), 1974–1983.

Green ER, Mecsas J (2016): Bacterial Secretion Systems: An Overview. In *Microbiol Spectr* 4 (1).

Gu S-Q, Peske F, Wieden H-J, Rodnina MV, Wintermeyer W (2003): The signal recognition particle binds to protein L23 at the peptide exit of the Escherichia coli ribosome. In *RNA* 9 (5), 566–573.

Hainzl T, Huang S, Meriläinen G, Brännström K, Sauer-Eriksson AE (2011): Structural basis of signal-sequence recognition by the signal recognition particle. In *Nat Struct Mol Biol* 18 (3), 389–391.

Hainzl T, Sauer-Eriksson AE (2015): Signal-sequence induced conformational changes in the signal recognition particle. In *Nat Commun* 6.

Halic M, Blau M, Becker T, Mielke T, Pool MR, Wild K, Sinning I, Beckmann R (2006): Following the signal sequence from ribosomal tunnel exit to signal recognition particle. In *Nature* 444 (7118), 507–511.

Hardy SJ, Randall LL (1991): A kinetic partitioning model of selective binding of nonnative proteins by the bacterial chaperone SecB. In *Science* 251 (4992), 439–443.

Hartl FU, Lecker S, Schiebel E, Hendrick JP, Wickner W (1990): The binding cascade of SecB to SecA to SecY/E mediates preprotein targeting to the E. coli plasma membrane. In *Cell* 63 (2), 269–279.

Heermann R, Altendorf K, Jung K (1998): The turgor sensor KdpD of Escherichia coli is a homodimer. In *Biochim Biophys Acta* 1415 (1), 114–124.

Hegde RS, Bernstein HD (2006): The surprising complexity of signal sequences. In *Trends Biochem Sci* 31 (10), 563–571.

Heijne G von (1990): The signal peptide. In *J. Membrain Biol.* 115 (3), 195–201.

Holtkamp W, Lee S, Bornemann T, Senyushkina T, Rodnina MV, Wintermeyer W (2012): Dynamic switch of the signal recognition particle from scanning to targeting. In *Nat Struct Mol Biol* 19 (12), 1332–1337.

Houben ENG, Zarivach R, Oudega B, Luirink J (2005): Early encounters of a nascent membrane protein: specificity and timing of contacts inside and outside the ribosome. In *J Cell Biol* 170 (1), 27–35.

Jagath JR, Rodnina MV, Wintermeyer W (2000): Conformational changes in the bacterial SRP receptor FtsY upon binding of guanine nucleotides and SRP. In *J Mol Biol* 295 (4), 745–753.

Jani AJ, Cotter PA (2010): Type VI secretion: not just for pathogenesis anymore. In *Cell Host Microbe* 8 (1), 2–6.

Jiang F, Chen M, Yi L, Gier J-W de, Kuhn A, Dalbey RE (2003): Defining the regions of Escherichia coli YidC that contribute to activity. In *J Biol Chem* 278 (49), 48965–48972.

Johnson N, Powis K, High S (2013): Post-translational translocation into the endoplasmic reticulum. In *Biochim Biophys Acta* 1833 (11), 2403–2409.

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

Kedrov A, Sustarsic M, Keyzer J de, Caumanns JJ, Wu ZC, Driessen AJM (2013): Elucidating the native architecture of the YidC: ribosome complex. In *J Mol Biol* 425 (22), 4112–4124.

Kedrov A, Wickles S, Crevenna AH, van der Sluis EO, Buschauer R, Berninghausen O, Lamb DC, Beckmann R (2016): Structural Dynamics of the YidC:Ribosome Complex during Membrane Protein Biogenesis. In *Cell Rep* 17 (11), 2943–2954.

Keenan RJ, Freymann DM, Stroud RM, Walter P (2001): The signal recognition particle. In *Annu Rev Biochem* 70, 755–775.

Keenan RJ, Freymann DM, Walter P, Stroud RM (1998): Crystal structure of the signal sequence binding subunit of the signal recognition particle. In *Cell* 94 (2), 181–191.

Kiefer D, Kuhn A (2018): YidC-mediated membrane insertion. In *FEMS Microbiol Lett* 365 (12).

Klenner C, Kuhn A (2012): Dynamic disulfide scanning of the membrane-inserting Pf3 coat protein reveals multiple YidC substrate contacts. In *J Biol Chem* 287 (6), 3769–3776.

Klenner C, Yuan J, Dalbey RE, Kuhn A (2008): The Pf3 coat protein contacts TM1 and TM3 of YidC during membrane biogenesis. In *FEBS Lett* 582 (29), 3967–3972.

Knyazev DG, Kuttner R, Zimmermann M, Sobakinskaya E, Pohl P (2018): Driving Forces of Translocation Through Bacterial Translocon SecYEG. In *J. Membrain Biol.* 251 (3), 329–343.

Koch HG, Hengelage T, Neumann-Haefelin C, MacFarlane J, Hoffschulte HK, Schimz KL, Mechler B, Müller M (1999): In vitro studies with purified components reveal signal recognition particle (SRP) and SecA/SecB as constituents of two independent protein-targeting pathways of Escherichia coli. In *Mol Biol Cell* 10 (7), 2163–2173.

Kohler R, Boehringer D, Greber B, Bingel-Erlenmeyer R, Collinson I, Schaffitzel C, Ban N (2009): YidC and Oxa1 form dimeric insertion pores on the translating ribosome. In *Mol Cell* 34 (3), 344–353.

Kramer G, Boehringer D, Ban N, Bukau B (2009): The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. In *Nat Struct Mol Biol* 16 (6), 589–597.

Kramer G, Rauch T, Rist W, Vorderwülbecke S, Patzelt H, Schulze-Specking A, Ban N, Deuerling E, Bukau B (2002): L23 protein functions as a chaperone docking site on the ribosome. In *Nature* 419 (6903), 171–174.

Kudva R, Denks K, Kuhn P, Vogt A, Müller M, Koch H-G (2013): Protein translocation across the inner membrane of Gram-negative bacteria: the Sec and Tat dependent protein transport pathways. In *Res Microbiol* 164 (6), 505–534.

Kuhn A, Koch H-G, Dalbey RE (2017): Targeting and Insertion of Membrane Proteins. In *EcoSal Plus* 7 (2).

Kuhn A, Kreil G, Wickner W (1986): Both hydrophobic domains of M13 procoat are required to initiate membrane insertion. In *EMBO J* 5 (13), 3681–3685.

Kuhn P, Draycheva A, Vogt A, Petriman N-A, Sturm L, Drepper F, Warscheid B, Wintermeyer W, Koch H-G (2015): Ribosome binding induces repositioning of the signal recognition particle receptor on the translocon. In *J Cell Biol* 211 (1), 91–104.

Kuhn P, Weiche B, Sturm L, Sommer E, Drepper F, Warscheid B, Sourjik V, Koch H-G (2011): The bacterial SRP receptor, SecA and the ribosome use overlapping binding sites on the SecY translocon. In *Traffic* 12 (5), 563–578.

Kumazaki K, Chiba S, Takemoto M, Furukawa A, Nishiyama K-i, Sugano Y, Mori T, Dohmae N, Hirata K, Nakada-Nakura Y, Maturana AD, Tanaka Y, Mori H, Sugita Y, Arisaka F, Ito K, Ishitani R, Tsukazaki T, Nureki O (2014a): Structural basis of Sec-independent membrane protein insertion by YidC. In *Nature* 509 (7501), 516–520.

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

Kyte J, Doolittle RF (1982): A simple method for displaying the hydropathic character of a protein. In *J Mol Biol* 157 (1), 105–132.

Laimins LA, Rhoads DB, Altendorf K, Epstein W (1978): Identification of the structural proteins of an ATP-driven potassium transport system in Escherichia coli. In *Proceedings of the National Academy of Sciences* 75 (7), 3216–3219.

Lam VQ, Akopian D, Rome M, Henningsen D, Shan S-o (2010): Lipid activation of the signal recognition particle receptor provides spatial coordination of protein targeting. In *J Cell Biol* 190 (4), 623–635.

Lee HC, Bernstein HD (2001): The targeting pathway of Escherichia coli presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal. In *Proceedings of the National Academy of Sciences* 98 (6), 3471–3476.

Lee PA, Tullman-Ercek D, Georgiou G (2006): The bacterial twin-arginine translocation pathway. In *Annu Rev Microbiol* 60, 373–395.

Leeuw E de, Poland D, Mol O, Sinning I, Hagen-Jongman CM ten, Oudega B, Luirink J (1997): Membrane association of FtsY, the E. coli SRP receptor. In *FEBS Lett* 416 (3), 225–229.

Leeuw E de, te Kaat K, Moser C, Menestrina G, Demel R, Kruijff B de, Oudega B, Luirink J, Sinning I (2000): Anionic phospholipids are involved in membrane association of FtsY and stimulate its GTPase activity. In *EMBO J* 19 (4), 531–541.

Lill R, Crooke E, Guthrie B, Wickner W (1988): The "trigger factor cycle" includes ribosomes, presecretory proteins, and the plasma membrane. In *Cell* 54 (7), 1013–1018.

Lim B, Miyazaki R, Neher S, Siegele DA, Ito K, Walter P, Akiyama Y, Yura T, Gross CA (2013): Heat Shock Transcription Factor σ32 Co-opts the Signal Recognition Particle to Regulate Protein Homeostasis in E. coli. In *PLoS Biol* 11 (12).

Loeffelholz O von, Jiang Q, Ariosa A, Karuppasamy M, Huard K, Berger I, Shan S-o, Schaffitzel C (2015): Ribosome-SRP-FtsY cotranslational targeting complex in the closed state. In *Proceedings of the National Academy of Sciences* 112 (13), 3943–3948.

Loeffelholz O von, Knoops K, Ariosa A, Zhang X, Karuppasamy M, Huard K, Schoehn G, Berger I, Shan S-o, Schaffitzel C (2013): Structural basis of signal sequence surveillance and selection by the SRP-FtsY complex. In *Nat Struct Mol Biol* 20 (5), 604–610.

Lotz M, Haase W, Kühlbrandt W, Collinson I (2008): Projection structure of yidC: a conserved mediator of membrane protein assembly. In *J Mol Biol* 375 (4), 901–907.

Luirink J, Hagen-Jongman CM ten, van der Weijden CC, Oudega B, High S, Dobberstein B, Kusters R (1994): An alternative protein targeting pathway in Escherichia coli: studies on the role of FtsY. In *EMBO J* 13 (10), 2289–2296.

Lutcke H (1995): Signal Recognition Particle (SRP), a Ubiquitous Initiator of Protein Translocation. In *Eur J Biochem* 228 (3), 531–550.

Maier KS, Hubich S, Liebhart H, Krauss S, Kuhn A, Facey SJ (2008): An amphiphilic region in the cytoplasmic domain of KdpD is recognized by the signal recognition particle and targeted to the Escherichia coli membrane. In *Mol Microbiol* 68 (6), 1471–1484.

Martoglio B, Dobberstein B (1998): Signal sequences: more than just greasy peptides. In *Trends in Cell Biology* 8 (10), 410–415.

Matlin KS (2002): The strange case of the signal recognition particle. In *Nat Rev Mol Cell Biol* 3 (7), 538–542.

Mercier E, Holtkamp W, Rodnina MV, Wintermeyer W (2017): Signal recognition particle binds to translating ribosomes before emergence of a signal anchor sequence. In *Nucleic Acids Res* 45 (20), 11858–11866.

Meyer DI, Dobberstein B (1980a): A membrane component essential for vectorial translocation of nascent proteins across the endoplasmic reticulum: requirements for its extraction and reassociation with the membrane. In *J Cell Biol* 87 (2 Pt 1), 498–502.

Meyer DI, Dobberstein B (1980b): Identification and characterization of a membrane component essential for the translocation of nascent proteins across the membrane of the endoplasmic reticulum. In *J Cell Biol* 87 (2 Pt 1), 503–508.

Miller JD, Bernstein HD, Walter P (1994): Interaction of E. coli Ffh/4.5S ribonucleoprotein and FtsY mimics that of mammalian signal recognition particle and its receptor. In *Nature* 367 (6464), 657–659.

Mircheva M, Boy D, Weiche B, Hucke F, Graumann P, Koch H-G (2009): Predominant membrane localization is an essential feature of the bacterial signal recognition particle receptor. In *BMC Biol* 7, 76.

Mitra K, Schaffitzel C, Shaikh T, Tama F, Jenni S, Brooks CL, Ban N, Frank J (2005): Structure of the E. coli protein-conducting channel bound to a translating ribosome. In *Nature* 438 (7066), 318–324.

Miyazaki R, Yura T, Suzuki T, Dohmae N, Mori H, Akiyama Y (2016): A Novel SRP Recognition Sequence in the Homeostatic Control Region of Heat Shock Transcription Factor σ32. In *Sci Rep* 6.

Montoya G, Svensson C, Luirink J, Sinning I (1997): Crystal structure of the NG domain from the signal-recognition particle receptor FtsY. In *Nature* 385 (6614), 365–368.

Mougous JD, Cuff ME, Raunser S, Shen A, Zhou M, Gifford CA, Goodman AL, Joachimiak G, Ordoñez CL, Lory S, Walz T, Joachimiak A, Mekalanos JJ (2006): A virulence locus of Pseudomonas aeruginosa encodes a protein secretion apparatus. In *Science* 312 (5779), 1526–1530.

Nagamori S, Smirnova IN, Kaback HR (2004): Role of YidC in folding of polytopic membrane proteins. In *J Cell Biol* 165 (1), 53–62.

Nakatogawa H, Ito K (2002): The ribosomal exit tunnel functions as a discriminating gate. In *Cell* 108 (5), 629–636.

Natale P, Brüser T, Driessen AJM (2008): Sec- and Tat-mediated protein secretion across the bacterial cytoplasmic membrane--distinct translocases and mechanisms. In *Biochim Biophys Acta* 1778 (9), 1735–1756.

Neugebauer SA, Baulig A, Kuhn A, Facey SJ (2012): Membrane protein insertion of variant MscL proteins occurs at YidC and SecYEG of Escherichia coli. In *J Mol Biol* 417 (4), 375–386.

Neumann-Haefelin C, Schäfer U, Müller M, Koch HG (2000): SRP-dependent co-translational targeting and SecA-dependent translocation analyzed as individual steps in the export of a bacterial protein. In *EMBO J* 19 (23), 6419–6426.

Noriega TR, Chen J, Walter P, Puglisi JD (2014a): Real-time observation of signal recognition particle binding to actively translating ribosomes. In *Elife* 3.

Noriega TR, Tsai A, Elvekrog MM, Petrov A, Neher SB, Chen J, Bradshaw N, Puglisi JD, Walter P (2014b): Signal recognition particle-ribosome binding is sensitive to nascent chain length. In *J Biol Chem* 289 (28), 19294–19305.

Oh E, Becker AH, Sandikci A, Huber D, Chaba R, Gloge F, Nichols RJ, Typas A, Gross CA, Kramer G, Weissman JS, Bukau B (2011): Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. In *Cell* 147 (6), 1295–1308.

Oliver D, Norman J, Sarker S (1998): Regulation of Escherichia coli secA by cellular protein secretion proficiency requires an intact gene X signal sequence and an active translocon. In *J Bacteriol* 180 (19), 5240–5242.

Palmer T, Sargent F, Berks BC (2005): Export of complex cofactor-containing proteins by the bacterial Tat pathway. In *Trends Microbiol* 13 (4), 175–180.

Papanikou E, Karamanou S, Economou A (2007): Bacterial protein secretion through the translocase nanomachine. In *Nat Rev Microbiol* 5 (11), 839–851.

Paramasivam N, Linke D (2011): ClubSub-P: Cluster-Based Subcellular Localization Prediction for Gram-Negative Bacteria and Archaea. In *Front Microbiol* 2, 218.

Park E, Ménétret J-F, Gumbart JC, Ludtke SJ, Li W, Whynot A, Rapoport TA, Akey CW (2014): Structure of the SecY channel during initiation of protein translocation. In *Nature* 506 (7486), 102–106.

Park S, Liu G, Topping TB, Cover WH, Randall LL (1988): Modulation of folding pathways of exported proteins by the leader sequence. In *Science* 239 (4843), 1033–1035.

Parkinson JS, Kofoid EC (1992): Communication modules in bacterial signaling proteins. In *Annu Rev Genet* 26, 71–112.

Parlitz R, Eitan A, Stjepanovic G, Bahari L, Bange G, Bibi E, Sinning I (2007): Escherichia coli signal recognition particle receptor FtsY contains an essential and autonomous membrane-binding amphipathic helix. In *J Biol Chem* 282 (44), 32176–32184.

Patel R, Smith SM, Robinson C (2014): Protein transport by the bacterial Tat pathway. In *Biochim Biophys Acta* 1843 (8), 1620–1628.

Pechmann S, Chartron JW, Frydman J (2014): Local slowdown of translation by nonoptimal codons promotes nascent-chain recognition by SRP in vivo. In *Nat Struct Mol Biol* 21 (12), 1100–1105.

Peluso P, Shan SO, Nock S, Herschlag D, Walter P (2001): Role of SRP RNA in the GTPase cycles of Ffh and FtsY. In *Biochemistry* 40 (50), 15224–15233.

Peschke M, Le Goff M, Koningstein GM, Karyolaimos A, Gier J-W de, van Ulsen P, Luirink J (2018): SRP, FtsY, DnaK and YidC Are Required for the Biogenesis of the E. coli Tail-Anchored Membrane Proteins DjlC and Flk. In *J Mol Biol* 430 (3), 389–403.

Peterson JH, Woolhead CA, Bernstein HD (2003): Basic amino acids in a distinct subset of signal peptides promote interaction with the signal recognition particle. In *J Biol Chem* 278 (46), 46155–46162.

Petriman N-A, Jauß B, Hufnagel A, Franz L, Sachelaru I, Drepper F, Warscheid B, Koch H-G (2018): The interaction network of the YidC insertase with the SecYEG translocon, SRP and the SRP receptor FtsY. In *Sci Rep* 8 (1), 578.

Phillips GJ, Silhavy TJ (1992): The E. coli ffh gene is necessary for viability and efficient protein export. In *Nature* 359 (6397), 744–746.

Polarek JW, Williams G, Epstein W (1992): The products of the kdpDE operon are required for expression of the Kdp ATPase of Escherichia coli. In *J Bacteriol* 174 (7), 2145–2151.

Poritz MA, Bernstein HD, Strub K, Zopf D, Wilhelm H, Walter P (1990): An E. coli ribonucleoprotein containing 4.5S RNA resembles mammalian signal recognition particle. In *Science* 250 (4984), 1111–1117.

Prabudiansyah I, Kusters I, Caforio A, Driessen AJM (2015): Characterization of the annular lipid shell of the Sec translocon. In *Biochim Biophys Acta* 1848 (10 Pt A), 2050–2056.

Prince WS, Villarejo MR (1990): Osmotic control of proU transcription is mediated through direct action of potassium glutamate on the transcription complex. In *J Biol Chem* 265 (29), 17673–17679.

Pukatzki S, Ma AT, Sturtevant D, Krastins B, Sarracino D, Nelson WC, Heidelberg JF, Mekalanos JJ (2006): Identification of a conserved bacterial protein secretion system in Vibrio cholerae using the Dictyostelium host model system. In *Proceedings of the National Academy of Sciences* 103 (5), 1528–1533.

Rabu C, Schmid V, Schwappach B, High S (2009): Biogenesis of tail-anchored proteins: the beginning for the end? In *J Cell Sci* 122 (Pt 20), 3605–3612.

Randall LL, Hardy SJ, Topping TB, Smith VF, Bruce JE, Smith RD (1998): The interaction between the chaperone SecB and its ligands: evidence for multiple subsites for binding. In *Protein Sci* 7 (11), 2384–2390.

Randall LL, Topping TB, Hardy SJ (1990): No specific recognition of leader peptide by SecB, a chaperone involved in protein export. In *Science* 248 (4957), 860–863.

Ravaud S, Stjepanovic G, Wild K, Sinning I (2008): The crystal structure of the periplasmic domain of the Escherichia coli membrane protein insertase YidC contains a substrate binding cleft. In *J Biol Chem* 283 (14), 9350–9358.

Records AR (2011): The type VI secretion system: a multipurpose delivery system with a phage-like machinery. In *Mol Plant Microbe Interact* 24 (7), 751–757.

Ribes V, Römisch K, Giner A, Dobberstein B, Tollervey D (1990): E. coli 4.5S RNA is part of a ribonucleoprotein particle that has properties related to signal recognition particle. In *Cell* 63 (3), 591–600.

Römisch K, Webb J, Herz J, Prehn S, Frank R, Vingron M, Dobberstein B (1989): Homology of 54K protein of signal-recognition particle, docking protein and two E. coli proteins with putative GTP-binding domains. In *Nature* 340 (6233), 478–482.

Römisch K, Webb J, Lingelbach K, Gausepohl H, Dobberstein B (1990): The 54-kD protein of signal recognition particle contains a methionine-rich RNA binding domain. In *J Cell Biol* 111 (5 Pt 1), 1793–1802.

Rupert PB, Ferré-D'amaré AR (2000): SRPrises in RNA-protein recognition. In *Structure* 8 (5), R99-104.

Sääf A, Monné M, Gier J-W de, Heijne G von (1998): Membrane Topology of the 60-kDa Oxa1p Homologue from Escherichia coli. In *J Biol Chem* 273 (46), 30415–30418.

Samuelson JC, Chen M, Jiang F, Möller I, Wiedmann M, Kuhn A, Phillips GJ, Dalbey RE (2000): YidC mediates membrane protein insertion in bacteria. In *Nature* 406 (6796), 637–641.

Samuelson JC, Jiang F, Yi L, Chen M, Gier JW de, Kuhn A, Dalbey RE (2001): Function of YidC for the insertion of M13 procoat protein in Escherichia coli: translocation of mutants that show differences in their membrane potential dependence and Sec requirement. In *J Biol Chem* 276 (37), 34847–34852.

Samuelsson T, Zwieb C (1999): The Signal Recognition Particle Database (SRPDB). In *Nucleic Acids Res* 27 (1), 169–170.

Saraogi I, Akopian D, Shan S-o (2014): Regulation of cargo recognition, commitment, and unloading drives cotranslational protein targeting. In *J Cell Biol* 205 (5), 693–706.

Schaffitzel C, Oswald M, Berger I, Ishikawa T, Abrahams JP, Koerten HK, Koning RI, Ban N (2006): Structure of the E. coli signal recognition particle bound to a translating ribosome. In *Nature* 444 (7118), 503–506.

Schibich D, Gloge F, Pöhner I, Björkholm P, Wade RC, Heijne G von, Bukau B, Kramer G (2016): Global profiling of SRP interaction with nascent polypeptides. In *Nature* 536 (7615), 219–223.

Schiebel E, Driessen AJ, Hartl FU, Wickner W (1991): Delta mu H+ and ATP function at different steps of the catalytic cycle of preprotein translocase. In *Cell* 64 (5), 927–939.

Schmitz U, Behrens S, Freymann DM, Keenan RJ, Lukavsky P, Walter P, James TL (1999a): Structure of the phylogenetically most conserved domain of SRP RNA. In *RNA* 5 (11), 1419–1429.

Schmitz U, James TL, Lukavsky P, Walter P (1999b): Structure of the most conserved internal loop in SRP RNA. In *Nat Struct Biol* 6 (7), 634–638.

Schuldiner M, Metz J, Schmid V, Denic V, Rakwalska M, Schmitt HD, Schwappach B, Weissman JS (2008): The GET Complex Mediates Insertion of Tail-Anchored Proteins into the ER Membrane. In *Cell* 134 (4), 634–645.

Schulze RJ, Komar J, Botte M, Allen WJ, Whitehouse S, Gold VAM, Lycklama a Nijeholt JA, Huard K, Berger I, Schaffitzel C, Collinson I (2014): Membrane protein insertion and proton-motive-force-dependent secretion through the bacterial holo-translocon SecYEG–SecDF–YajC–YidC. In *Proceedings of the National Academy of Sciences* 111 (13), 4844–4849.

Schwarz S, Hood RD, Mougous JD (2010): What is type VI secretion doing in all those bugs? In *Trends Microbiol* 18 (12), 531–537.

Seitl I, Wickles S, Beckmann R, Kuhn A, Kiefer D (2014): The C-terminal regions of YidC from Rhodopirellula baltica and Oceanicaulis alexandrii bind to ribosomes and partially substitute for SRP receptor function in Escherichia coli. In *Mol Microbiol* 91 (2), 408–421.

Shalom G, Shaw JG, Thomas MS (2007): In vivo expression technology identifies a type VI secretion system locus in Burkholderia pseudomallei that is induced upon invasion of macrophages. In *Microbiology (Reading, Engl)* 153 (Pt 8), 2689–2699.

Shan S-o, Stroud RM, Walter P (2004): Mechanism of association and reciprocal activation of two GTPases. In *PLoS Biol* 2 (10), e320.

Shan S-o, Walter P (2003): Induced nucleotide specificity in a GTPase. In *Proceedings of the National Academy of Sciences* 100 (8), 4480–4485.

Shan S-o, Walter P (2005): Co-translational protein targeting by the signal recognition particle. In *FEBS Lett* 579 (4), 921–926.

Shanmugam SK, Backes N, Chen Y, Belardo A, Phillips GJ, Dalbey RE (2019): New Insights into Amino-Terminal Translocation as Revealed by the Use of YidC and Sec Depletion Strains. In *J Mol Biol* 431 (5), 1025–1037.

Shen K, Arslan S, Akopian D, Ha T, Shan S-o (2012): Activated GTPase movement on an RNA scaffold drives co-translational protein targeting. In *Nature* 492 (7428), 271–275.

Shneider MM, Buth SA, Ho BT, Basler M, Mekalanos JJ, Leiman PG (2013): PAAR-repeat proteins sharpen and diversify the type VI secretion system spike. In *Nature* 500 (7462), 350–353.

Siegel V, Walter P (1986): Removal of the Alu structural domain from signal recognition particle leaves its protein translocation activity intact. In *Nature* 320 (6057), 81–84.

Siegel V, Walter P (1988): The affinity of signal recognition particle for presecretory proteins is dependent on nascent chain length. In *EMBO J* 7 (6), 1769–1775.

Silverman JM, Brunet YR, Cascales E, Mougous JD (2012): Structure and regulation of the type VI secretion system. In *Annu Rev Microbiol* 66, 453–472.

Soman R, Yuan J, Kuhn A, Dalbey RE (2014): Polarity and charge of the periplasmic loop determine the YidC and sec translocase requirement for the M13 procoat lep protein. In *J Biol Chem* 289 (2), 1023–1032.

Spann D, Pross E, Chen Y, Dalbey RE, Kuhn A (2018): Each protomer of a dimeric YidC functions as a single membrane insertase. In *Sci Rep* 8 (1), 589.

Steinberg R, Knüpffer L, Origi A, Asti R, Koch H-G (2018): Co-translational protein targeting in bacteria. In *FEMS Microbiol Lett* 365 (11).

Stjepanovic G, Kapp K, Bange G, Graf C, Parlitz R, Wild K, Mayer MP, Sinning I (2011): Lipids trigger a conformational switch that regulates signal recognition particle (SRP)-mediated protein targeting. In *J Biol Chem* 286 (26), 23489–23497.

Suelter CH (1970): Enzymes activated by monovalent cations. In *Science* 168 (3933), 789–795.

Tam PCK, Maillard AP, Chan KKY, Duong F (2005): Investigating the SecY plug movement at the SecYEG translocation channel. In *EMBO J* 24 (19), 3380–3388.

Tsukazaki T, Mori H, Echizen Y, Ishitani R, Fukai S, Tanaka T, Perederina A, Vassylyev DG, Kohno T, Maturana AD, Ito K, Nureki O (2011): Structure and function of a membrane component SecDF that enhances protein export. In *Nature* 474 (7350), 235–238.

Turner RJ, Papish AL, Sargent F (2004): Sequence analysis of bacterial redox enzyme maturation proteins (REMPs). In *Can J Microbiol* 50 (4), 225–238.

Ulbrandt ND, Newitt JA, Bernstein HD (1997): The E. coli Signal Recognition Particle Is Required for the Insertion of a Subset of Inner Membrane Proteins. In *Cell* 88 (2), 187–196.

Urbanus ML, Scotti PA, Froderberg L, Saaf A, Gier JW de, Brunner J, Samuelson JC, Dalbey RE, Oudega B, Luirink J (2001): Sec-dependent membrane protein insertion: sequential interaction of nascent FtsQ with SecY and YidC. In *EMBO Rep* 2 (6), 524–529.

Valent QA, Scotti PA, High S, Gier JW de, Heijne G von, Lentzen G, Wintermeyer W, Oudega B, Luirink J (1998): The Escherichia coli SRP and SecB targeting pathways converge at the translocon. In *EMBO J* 17 (9), 2504–2512.

van Bloois E, Jan Haan G, Gier J-W de, Oudega B, Luirink J (2004): F(1)F(0) ATP synthase subunit c is targeted by the SRP to YidC in the E. coli inner membrane. In *FEBS Lett* 576 (1-2), 97–100.

van den Berg B, Clemons WM, Collinson I, Modis Y, Hartmann E, Harrison SC, Rapoport TA (2004): X-ray structure of a protein-conducting channel. In *Nature* 427 (6969), 36–44.

van der Laan M, Nouwen N, Driessen AJM (2004): SecYEG proteoliposomes catalyze the Deltaphi-dependent membrane insertion of FtsQ. In *J Biol Chem* 279 (3), 1659–1664.

van Voorst F, Vereyken IJ, Kruijff B de (2000): The high affinity ATP binding site modulates the SecA-precursor interaction. In *FEBS Lett* 486 (1), 57–62.

Veenendaal AKJ, van der Does C, Driessen AJM (2004): The protein-conducting channel SecYEG. In *Biochim Biophys Acta* 1694 (1-3), 81–95.

Voigts-Hoffmann F, Schmitz N, Shen K, Shan S-o, Ataide SF, Ban N (2013): The structural basis of FtsY recruitment and GTPase activation by SRP RNA. In *Mol Cell* 52 (5), 643–654.

Voss NR, Gerstein M, Steitz TA, Moore PB (2006): The geometry of the ribosomal polypeptide exit tunnel. In *J Mol Biol* 360 (4), 893–906.

Wagner S, Pop OI, Pop O, Haan G-J, Baars L, Koningstein G, Klepsch MM, Genevaux P, Luirink J, Gier J-W de (2008): Biogenesis of MalF and the MalFGK(2) maltose transport complex in Escherichia coli requires YidC. In *J Biol Chem* 283 (26), 17881–17890.

Walderhaug MO, Polarek JW, Voelkner P, Daniel JM, Hesse JE, Altendorf K, Epstein W (1992): KdpD and KdpE, proteins that control expression of the kdpABC operon, are members of the two-component sensor-effector class of regulators. In *J Bacteriol* 174 (7), 2152–2159.

Wallin E, Heijne G von (1998): Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. In *Protein Sci* 7 (4), 1029–1038.

Walter P, Blobel G (1980): Purification of a membrane-associated protein complex required for protein translocation across the endoplasmic reticulum. In *Proceedings of the National Academy of Sciences* 77 (12), 7112–7116.

Weiche B, Bürk J, Angelini S, Schiltz E, Thumfart JO, Koch H-G (2008): A cleavable N-terminal membrane anchor is involved in membrane binding of the Escherichia coli SRP receptor. In *J Mol Biol* 377 (3), 761–773.

Welte T, Kudva R, Kuhn P, Sturm L, Braig D, Müller M, Warscheid B, Drepper F, Koch H-G (2012): Promiscuous targeting of polytopic membrane proteins to SecYEG or YidC by the Escherichia coli signal recognition particle. In *Mol Biol Cell* 23 (3), 464–479.

Woolhead CA, Johnson AE, Bernstein HD (2006): Translation arrest requires two-way communication between a nascent polypeptide and the ribosome. In *Mol Cell* 22 (5), 587–598.

Xie K, Kiefer D, Nagler G, Dalbey RE, Kuhn A (2006): Different regions of the nonconserved large periplasmic domain of Escherichia coli YidC are involved in the SecF interaction and membrane insertase activity. In *Biochemistry* 45 (44), 13401–13408.

Yang M-J, Zhang X (2011): Molecular dynamics simulations reveal structural coordination of Ffh-FtsY heterodimer toward GTPase activation. In *Proteins* 79 (6), 1774–1785.

Yen MR, Harley KT, Tseng YH, Saier MH (2001): Phylogenetic and structural analyses of the oxa1 family of protein translocases. In *FEMS Microbiol Lett* 204 (2), 223–231.

Yu Z, Koningstein G, Pop A, Luirink J (2008): The conserved third transmembrane segment of YidC contacts nascent Escherichia coli inner membrane proteins. In *J Biol Chem* 283 (50), 34635–34642.

Zhang X, Lam VQ, Mou Y, Kimura T, Chung J, Chandrasekar S, Winkler JR, Mayo SL, Shan S-o (2011): Direct visualization reveals dynamics of a transient intermediate during protein assembly. In *Proceedings of the National Academy of Sciences* 108 (16), 6450–6455.

Zhang X, Schaffitzel C, Ban N, Shan S-o (2009): Multiple conformational switches in a GTPase complex control co-translational protein targeting. In *Proceedings of the National Academy of Sciences* 106 (6), 1754–1759.

Zheng N, Gierasch LM (1997): Domain Interactions in E. coli SRP: Stabilization of M Domain by RNA Is Required for Effective Signal Sequence Modulation of NG Domain. In *Mol Cell* 1 (1), 79–87.

Zimmann P, Puppe W, Altendorf K (1995): Membrane topology analysis of the sensor kinase KdpD of Escherichia coli. In *J Biol Chem* 270 (47), 28282–28288.

Zopf D, Bernstein HD, Johnson AE, Walter P (1990): The methionine-rich domain of the 54 kd protein subunit of the signal recognition particle contains an RNA binding site and can be crosslinked to a signal sequence. In *EMBO J* 9 (13), 4511–4517.

Zoued A, Brunet YR, Durand E, Aschtgen M-S, Logger L, Douzi B, Journet L, Cambillau C, Cascales E (2014): Architecture and assembly of the Type VI secretion system. In *Biochim Biophys Acta* 1843 (8), 1664–1673.

Zoued A, Cassaro CJ, Durand E, Douzi B, España AP, Cambillau C, Journet L, Cascales E (2016): Structure-Function Analysis of the TssL Cytoplasmic Domain Reveals a New Interaction between the Type VI Secretion Baseplate and Membrane Complexes. In *J Mol Biol* 428 (22), 4413–4423.

Zoued A, Duneau J-P, Durand E, España AP, Journet L, Guerlesquin F, Cascales E (2018): Tryptophan-mediated Dimerization of the TssL Transmembrane Anchor Is Required for Type VI Secretion System Activity. In *J Mol Biol* 430 (7), 987–1003.

APPENDIX 7

Danksagung

Ich möchte mich ganz herzlich bei Herrn Prof. Dr. Kuhn bedanken, für die Möglichkeit die Promotion in seinem Institut durchzuführen. Vielen Dank für Ihre Unterstützung und dass ich meine Arbeit bei internationalen Konferenzen vorstellen durfte. Ich habe es sehr geschätzt, dass Ihre Türe immer offenstand, für jedes Anliegen das ich hatte.

Des Weiteren möchte ich mich ganz besonders bei Domenico bedanken. Danke für deine immer gute Laune, deine Motivation, wenn es mal nicht so gut gelaufen ist und für die tolle Zeit in der „AG Pupo“. Ein großer Dank geht an die gute Seele des Labors 334, Gerda, für die vielen tollen Gespräche und die Hilfe über all die Jahre. Besonders bedanken möchte ich mich bei der besten Laborpartnerin Farina für die wahnsinnig schöne Zeit, die vielen witzigen Momente, deine Unterstützung, die zahlreichen aufschlussreichen Gespräche und Diskussionen und natürlich für deine unendliche Hilfsbereitschaft. Du wirst mir wirklich fehlen.

Bei dir liebe Ines, möchte ich mich ganz besonders bedanken, wobei Worte dafür eigentlich gar nicht ausreichen. Du warst immer ein Vorbild für mich, durch dein Wissen und deine tolle Art. Danke, dass du die Leidenschaft für SRP mit mir geteilt hast, für die vielen schönen Feierabendtreffen und deine grenzenlose Hilfsbereitschaft und Unterstützung.

Ein ganz herzlicher Dank geht an Doro, für die Unterstützung beim Stipendium und der Arbeit an der landwirtschaftlichen Schule. Danke, dass ich immer mit allen Problemen zu dir kommen konnte.
Ein weiterer großer Dank für eine wirklich schöne Zeit und ein tolles familiäres Arbeitsklima geht an meine lieben (Ex-) Kollegen und Kolleginnen, Chris, Dirk, Betty, Max, Sandra, Sabrina, Anja, Renate, Lutz, Gisela, Markus, Simone, Su und Beate. Betty, vielen Dank für deine Hilfe und Unterstützung im Labor. Du warst mir eine sehr große Hilfe. Max, selbst du wirst mir irgendwie fehlen :)

Vielen Dank an Fabian und meine Familie für die Unterstützung und die Motivation während der gesamten Zeit. Danke, dass ich mich immer auf euch verlassen kann.

Abbreviations

aa	amino acid
Alb	Albino
AMS	4-acetamido-4´-maleimidylstilbene-2,2´-disulfonic acid
ara	arabinose
ATP	adenosine triphosphate
Å	Angstrom
B. halodurans	*Bacillus halodurans*
B. subtilis	*Bacillus subtilis*
C-terminal	carboxy-terminal
C	cytoplasmic domain
cryo-EM	cryo-electron microscopy
CyoA	Cytochrome bo ubiquinol oxidase subunit 2
DNA	deoxyribonucleic acid
DDM	n-dodecyl-β-d-maltoside
DTT	dithiothreitol
ER	endoplasmatic reticulum
EAEC	enteroaggregative *Escherichia coli*
E. coli	*Eschericia coli*
EDTA	ethylenediaminotetraacetic acid
et al.	*et alii,* and others
Ffh	fifty-four-homolog
FtsQ	filamentous temperature sensitive Q
FtsY	filamentous temperature sensitive Y
g	acceleration of gravity (9.81 m/s^2)
GFP	green fluorescent protein
gluc	glucose
gp	gene product
GRAVY	grand average of hydropathy
GTP	guanosine triphosphate
H	hour
HCl	hydrochloric acid
Hcp	Hemolysin-coregulated protein

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
his-tag	histidine-tag
IM	inner membrane
IMP	inner membrane protein
IPTG	isopropyl-β-D-thiogalactoside
Kd	dissociation constant
kDa	Kilodalton
LB	Luria-Bertani
Lep	leaderpeptidase
MAP	methionine amino peptidase
MBP	maltose-binding protein
MifM	membrane protein insertion and folding monitor
min	minute
mM	millimolar
MscL	mechanosensitive channel of large conductance
MST	microscale thermophoresis
MtlA	mannitol permease
MW	molecular weight
M. jannaschii	*Methanocaldococcus jannaschii*
N-terminal	amino-terminal
NaCl	sodium chloride
nm	nanometer
nM	nanomolar
OD	optical density
OM	outer membrane
OmpA	outer membrane protein A
Oxa1	Oxidase assembly 1
P	periplasmic domain
PhoA	Alkaline Phosphatase
PAGE	Polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	protein data bank
PDF	peptidyl formylase
pmf	proton motive force

PMSF	phenylmethylsulfonyl fluoride
PTC	peptidyl transferase center
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
RNC	ribosome nascent chain
rpm	revolutions per minute
RR	twin arginine
S. cerevisiae	*Saccharomyces cerevisiae*
Sec	Secretory
SDS	sodium dodecyl sulfate
sfGFP	superfolder GFP
SRP	Signal recognition particle
Strep-tag	streptavidin-tag
S. mutans	*Streptococcus mutans*
T6SS	type VI secretion system
TA	tail anchored
Tat	Twin arginine translocation
TCA	trichloracetic acid
TF	Trigger factor
TM	transmembrane
TMD	transmembrane domain
TMS	transmembrane segment
Tris	trishydroxymethylaminomethane
tRNA	Transfer ribonucleic acid
Tss	Type six secretion
VgrG	valine-glycine repeat protein
α	anti, alpha
µm	micrometer
µM	micromolar

Curriculum vitae

Personal data

Name:	Eva Proß
Date and Place of birth:	09.02.1990 in Herrenberg
Nationality:	German

Employment

From 2019 / 09	PostDoc, University of Hohenheim, Institute of Food Science and Biotechnology, Prof. Dr. Lutz Fischer

Education

2015 / 02 – to date	PhD student, University of Hohenheim, Institute for Microbiology, Prof. Dr. Andreas Kuhn
2012 / 10 – 2015 / 02	Studies in Biology, Master of Science, University of Hohenheim
2010 / 10 – 2012 / 07	Studies in Biology, Bachelor of Science, Ludwig-Maximilians-Universität München
2009 / 10 – 2010 / 08	Studies in Biology, Bachelor, Eberhard Karls Universität Tübingen
2000 / 09 – 2009 / 07	Abitur at Otto-Hahn-Gymnasium, Nagold

Secondary Employment

2015 / 09 – 2017 / 07	Teacher for Microbiology, vocational college for Laboratory Assistants in Agriculture and environmental analysis, Landwirtschaftliche Schule Hohenheim

Scholarship

2015 / 05 – 2017 / 04	Promotion Scholarship, Landesgraduiertenförderung

Place, Date

Signature

Publications

Pross E, Soussoula L, Seitl I, Lupo D, Kuhn A. (2016)
Membrane Targeting and Insertion of the C-tail Protein SciP. *J Mol Biol* 428 (20): 4218-4227

Cheng et al. (2016)
Combined B, T and NK Cell Deficiency accelerates Atherosclerosis in BALB/c Mice. *PLoS One* 11 (8): e0157311

Spann D, Pross E, Chen Y, Dalbey RE, Kuhn A. (2018) Each protomer of a dimeric YidC functions as a single membrane insertase. *Sci Rep* 8 (1): 589

Pross E and Kuhn A (2019)
The SRP signal sequence of KdpD. *Sci Rep* 9 (1): 8717

Presentations at international conferences

Gordon Conference, Galveston USA 2016: Protein transport across cell membranes.
Poster presentation “Co-translational targeting of the C-tailed protein SciP”.

BPE, Leuven, Belgium 2018: Bacterial Protein Export.
Poster presentation “Novel SRP targeting signals in the membrane proteins KdpD and SciP in *E. coli*”.

Eidesstattliche Versicherung gemäß § 7 Absatz 7 der Promotionsordnung der Universität Hohenheim zum Dr. rer. nat.

1. Bei der eingereichten Dissertation zum Thema

   Membrane targeting and insertion of the sensor protein KdpD and the C-tail anchored protein SciP of *Escherichia coli*

   Handelt es sich um meine eigenständig erbrachte Leistung.

2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.
3. Ich habe nicht die Hilfe einer kommerziellen Promotionsvermittlung oder -beratung in Anspruch genommen.
4. Die Bedeutung der eidesstattlichen Versicherung und der strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Die Richtigkeit der vorstehenden Erklärung bestätige ich: Ich versichere an Eides statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

_____ _____

Ort und Datum Unterschrift

Erklärung zum Eigenanteil

Pross E, Soussoula L, Seitl I, Lupo D, Kuhn A. (2016)
Membrane Targeting and Insertion of the C-tail Protein SciP. J Mol Biol 428 (20): 4218-4227
Durchführung der Versuche (ausgenommen der Expressionsexperimente mit SciP, SciP28 und SciP41 und der PK Mapping Experimente mit SciP, SciP26 und SciP28 in MK6), Interpretation der Ergebnisse, Generierung der Abbildungen und Verfassung des Manuskriptteils „Material und Methoden“.

Pross E and Andreas Kuhn (2019)
The SRP signal sequence of KdpD. Sci Rep 9 (1): 8717
Durchführung aller experimenteller Arbeiten, Interpretation der Ergebnisse, Generierung der Abbildungen und Verfassung des Manuskripts.

_____ _____

Ort und Datum Unterschrift Herr Prof. Dr. Kuhn