A novel membrane protein insertion pathway in Escherichia coli
van der Laan, Martin

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2004

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Introduction - On the mechanisms of bacterial membrane protein biogenesis

Martin van der Laan and Arnold J.M. Driessen

Historical introduction

During the last few years the number of successful reports on 3D crystallization of α-helical membrane proteins has increased considerably. High resolution structures of several ion channels, water or glycerol channels, calcium P-type ATPase SERCA, ABC transporters MsbA and BtuCD, secondary transporters AcrB, LacY, GlpT and mitochondrial ADP/ATP carrier, the phototaxis receptor sensory rhodopsin II, the preprotein-conducting SecY/Sec61 translocase and several large energy-transducing complexes have provided exciting new insights in membrane protein structure and function. A collection of membrane protein crystal structures can be found at the regularly updated website “Membrane proteins of known 3D structure” (http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html) hosted by the group of Stephen White. It, however, remains a challenging question how these structurally diverse proteins are asymmetrically incorporated into the lipid bilayer and how they assemble into their functional conformations.

For the mechanism of membrane protein insertion, initially, two models were controversially discussed: the “helical hairpin” and the “signal sequence” hypothesis. The helical hairpin hypothesis postulated that membrane protein insertion occurs spontaneously driven by hydrophobic partitioning of the transmembrane α-helical domains into the lipid bilayer. In energized membranes, this process would be directed by the proton motive force (PMF) (74, 249). Experimental support for this model came from in vitro studies showing the spontaneous insertion of small and simple model membrane proteins, like the major coat proteins of the bacteriophages M13 (83) and Pf3 (199) into protein-free liposomes. Also, a stimulatory and topology-
determining role of the PMF was demonstrated (45, 127, 128). However, it remained difficult to imagine that more complex membrane proteins with multiple transmembrane segments (TMSs) and large hydrophilic domains would insert in the same way. Many different types of TMSs need to be incorporated into the lipid bilayer, while certain domains have to cross the membrane and others are retained in the cytoplasm. Finally, the proteins have to fold into their final tertiary structures, and some even oligomerize and assemble into large membrane protein complexes. In the early Seventies, Blobel and coworkers proposed their signal hypothesis (22). They suggested that proteins that are translocated into the lumen of the endoplasmatic reticulum (ER) in eukaryotic cells are specifically recognized by a targeting factor as soon as a putative signal sequence emerges from the ribosome. This factor would then mediate the binding of the translating ribosome to the ER where the growing polypeptide chain would be pushed across the membrane concomitantly with chain elongation. The signal sequence that directs this event is cleaved off by a specific proteolytic enzyme located in the ER. A few years later, Blobel & Dobberstein (23, 24) proposed the “expanded signal hypothesis” which involved the presence of a translocation channel in the ER membrane formed by specific integral membrane proteins. The authors reasoned that this channel would also be used for the insertion of membrane proteins with the difference that such proteins would not be completely translocated, but instead contain internal signals (“stop transfer” sequences) that halt translocation and induce their lateral release from the protein-conducting channel into the lipid bilayer. Supporting experimental evidence for the signal hypothesis came from in vitro studies on the membrane insertion of the vesicular stomatitis virus (VSV) glycoprotein, which was asymmetrically inserted into canine microsomal vesicles, glycosylated in the lumen of the ER and proteolytically processed upon insertion (123). Moreover, the pathways that mediate membrane protein insertion and translocation of secretory proteins were shown to compete for the same import site in the microsomal membranes (151).

Factors responsible for the co-translational targeting of translating ribosomes to the ER membrane were identified in the early Eighties. Warren and Dobberstein discovered that salt-washed microsomes do not support in vitro protein translocation while re-addition of the salt wash fraction restores activity (256). In 1980, Walter and Blobel purified this activity and showed that it consists of at least six different protein subunits (251). Two years later, it became evident that this complex also contains an
essential 7 S RNA (255). This ribozyme was termed “signal recognition particle” (SRP) and found to bind to ribosomes synthesizing secretory proteins, to arrest their translation (252) and to target the translation-arrested ribosomes to the ER membrane (253). Binding of the SRP-ribosome-nascent chain complex to microsomes releases translation arrest (253), and this activity was later attributed to the SRP receptor (SR) (85, 86, 163). Eukaryotic SRP receptor consists of a small, membrane-integral subunit (SRβ) and a large peripherally associated subunit (SRα) (223), both of which are GTPases (38, 165).

The mechanism of the actual translocation / insertion process and the existence of an aqueous protein-conducting channel has remained a controversial issue for several years. Electrophysiological measurements using planar lipid membranes demonstrated that ER membranes as well as Escherichia coli cytoplasmic membranes contain high-conductivity channels (215). Opening (or formation) of the ER channel requires ribosome binding (216), while the bacterial channel appeared to be gated by signal peptides of secretory proteins (217). Therefore, such channels were believed to represent the protein-conducting channels that were proposed by the expanded signal hypothesis in 1975 (23,24).

In the Gram-negative bacterium Escherichia coli, genetic screens were developed in order to select for mutations that caused pleiotropic protein secretion defects (117, 188) or restored the translocation of secretory proteins with defective signal sequences (13, 73, 200). These genetic studies identified three genes termed secA (prlD), secY (prlA) and secE (prlG). Wickner and coworkers found that mutations in secY and secA not only interfered with the secretion of proteins, but also blocked the membrane insertion of the inner membrane protein leader peptidase (262). The SecA protein turned out to be a peripherally membrane-bound, ATP-driven motor protein (30, 149) that together with the proton motive force (PMF) drives post-translational protein export in bacteria, while SecY and SecE are two interacting membrane proteins. Soon bacterial protein translocation was successfully reconstituted into proteoliposomes with the purified SecYE and SecA proteins (4, 28). An additional small membrane protein, SecG, was discovered that strongly stimulates this process (96). SecG was found to associate with SecY and SecE yielding the heterotrimeric SecYEG complex that is believed to form a protein-conducting channel in the membrane.

A genetic screen for yeast mutants impaired in the secretion of proteins yielded a gene named sec61 (57), which encodes an ER-localized membrane protein. Sec61 mutants
were also found to be defective in the integration of membrane proteins into the ER membrane (219). In addition, *in vitro* crosslinking experiments with photoactivatable amino acid analogues incorporated into nascent secretory proteins (89, 90, 135, 178, 224, 260) and nascent integral membrane proteins (113, 114) demonstrated an interaction with the Sec61 protein. Rapoport and coworkers finally succeeded to reconstitute ER protein translocation and membrane protein insertion with purified components prepared from mammalian microsomes (91). Those studies together with the genetic screens showed that a heterotrimeric membrane protein complex, termed Sec61αβγ is responsible for the translocation activity in ER membranes. Strikingly, the α and γ subunits of this complex are similar in structure and function to the bacterial SecY and SecE proteins, respectively (98).

During the Nineties, the SRP targeting pathway as well as the Sec translocase, both in large parts evolutionary conserved, were extensively characterized biochemically, biophysically and structurally. Several accessory components were described to be involved in particular substeps of the translocation and membrane insertion reactions. In recent years, it has become clear that membrane protein insertion in bacteria is not only mediated by the Sec translocase, but also assisted by another membrane protein termed YidC. YidC is a member of the evolutionary conserved Oxa membrane protein family (140). The small phage coat proteins that were originally believed to insert spontaneously into the membrane, *in vivo* actually use a distinct pathway that involves YidC (205). YidC has also been implicated in Sec-dependent membrane protein insertion. In the following, we will discuss our current understanding of membrane protein targeting and insertion in bacteria.

**Membrane protein targeting**

In *Escherichia coli*, two protein targeting routes converge at the Sec translocase, the co-translational SRP pathway (Fig. 1, pathway A) and the SecB pathway (132, 234). Secretory proteins are synthesized in the cytosol as precursors with an N-terminal cleavable signal sequence. SecB is a molecular chaperone that binds to long nascent secretory proteins and keeps them in an unfolded, translocation-competent state (144). SecB targets these proteins to the translocase by virtue of its affinity for the ATP-
driven motor protein SecA. This interaction triggers an intricate hand-over-hand mechanism that involves binding of the exposed signal sequence of the secretory protein by SecA, which subsequently activates SecB for preprotein transfer to SecA. In an ATP- and PMF-dependent fashion, the preprotein is then translocated through the protein-conducting channel formed by the SecYEG complex (66, 149). The SecB pathway is often referred to as a post-translational pathway, but SecB has also been implicated in the co-translational translocation of secretory proteins (141).

The bacterial SRP pathway seems to be specialized for the targeting of membrane proteins (228). It is functionally similar to the co-translational mechanism of protein targeting to the ER in eukaryotic cells (194). *E. coli* SRP consists of a 4.5 S RNA and a 48 kDa protein which is a homologue of the eukaryotic SRP54 protein and was therefore named “Fifty-four homologue” (Ffh) (192, 198, 202). Ffh contains three distinct structural domains: an N-terminal four-helix bundle (N-domain), a Ras-like GTPase domain (G-domain) and a methionine-rich C-terminal domain (M-domain) (14, 77, 125). The N-domain has been implicated in the regulation of nucleotide binding to the G-domain and in the communication and coordination between GTPase activity, signal sequence recognition and receptor binding (68, 76, 77, 152). It is widely believed that the M-domain together with a conserved domain of the 4.5 S RNA constitutes the signal sequence binding groove (125). However, also the NG domain was shown to interact with signal sequences (37). Major characteristics of the signal sequences recognized by SRP relate to their hydrophobicity and strong α-helical character which are typical features of TMSs of inner membrane proteins. Replacement of the signal sequence of a secretory protein for a TMS of an inner membrane protein re-routes a secretory protein from the SecB-dependent pathway into the SRP pathway (145, 182). The same can be achieved by artificially increasing the hydrophobicity of a secretory protein signal sequence (145, 232) or by replacing a helix-breaking residue for a helix-promoting residue without a change in hydrophobicity (3).

Translation-arrested, ribosome-bound nascent chains of membrane proteins can be targeted to the membrane as soon as the first hydrophobic TMS emerges from the ribosome (116, 230). An early interaction at the ribosome between Ffh and the signal sequence of the model membrane protein FtsQ has been observed (229) (Fig. 3, step 1). Ribosome-bound SRP seems to be positioned in close proximity to the ribosome exit tunnel and interacts with the L23 protein (92, 229). Upon signal sequence
binding, SRP targets the ribosome-nascent chain complex to the membrane where it associates with the membrane-bound receptor FtsY (Fig. 3, step 2). FtsY is a homologue of the eukaryotic SRα subunit (164). In bacteria, no membrane-integral anchor subunit similar to the eukaryotic SRβ has been identified. FtsY contains no obvious hydrophobic transmembrane domain but exhibits a dual localization in the cell, i.e., both at the membrane and in the cytosol (153). The NG-domain organization and structure of FtsY are very similar to Ffh (171). In addition, FtsY contains a strongly acidic domain N-terminal A-domain. Both, NG- and A-domain have been implicated in membrane interaction of FtsY (53, 54, 167, 168, 269). The A-domain can be functionally replaced by an unrelated TMS derived from an inner membrane protein (269). FtsY binds tightly to the membrane surface and adopts a state that is largely non-extractable with urea or alkaline treatment (153). A direct interaction of FtsY with phospholipids has been shown, although conflicting reports have been published concerning the head group specificity of this interaction (54, 168). One study also described a trypsin-sensitive interaction of FtsY with membrane vesicles. Therefore, it has been suggested that a membrane protein contributes to FtsY membrane binding (168). However, the identity of this membrane protein has remained obscure.

Interaction of SRP with FtsY induces the release of the nascent chains from SRP (234) and stimulates the GTPase activity of both proteins (68, 76, 193). GTP hydrolysis is required for the dissociation of the targeting complex as this recycling step is prevented by non-hydrolysable GTP analogues, such as GMP-PNP or GMP-PCP (39). Recently, the crystal structure of a GMP-PCP-stabilized complex formed by the NG-domains of *Thermus aquaticus* Ffh and FtsY has been reported (68, 76). In this highly symmetric complex, the two nucleotides are bound by composite active sites at the extensive interaction surface of the subunits. It has been suggested that GTP binding induces the formation of this interface. This would then trigger the transfer of the signal sequence from the SRP binding groove to the translocation channel (Fig. 3, steps 2-3). GTP hydrolysis is believed to disrupt the interaction surface. The molecular mechanism of substrate transfer to the translocase, however, remains to be elucidated. Direct transfer of the nascent chain from the targeting complex to the translocase implies an, at least transient, interaction between both complexes. However, in bacteria such an interaction has never been demonstrated, although the SecYEG complex can bind ribosomes with high affinity (195). In the
conserved SRP targeting pathway in chloroplasts (44), a stable complex between cpSRP, cpFtsY and cpSecYEG formed in the presence of GMP-PNP has recently been isolated (173). Further analysis of this complex might provide a better understanding of signal sequence transfer. However, the chloroplast SRP pathway shows two unique features: it also mediates post-translational targeting of nuclear-encoded proteins and cpSRP consist of a SRP54 homologue and an additional protein, but does not contain RNA (69). Therefore, major mechanistic aspects might be different as well.

Fig. 1: Overview of membrane protein insertion pathways in Escherichia coli. Most membrane proteins are co-translationally targeted via the SRP pathway to the Sec translocase, where the initial membrane insertion takes place (A). SecA drives the translocation of large hydrophilic domains across the membrane via the SecYEG channel. The YidC protein has been implicated in the lateral release of TMSs from the SecYEG complex into the lipid bilayer. In addition to a putative role in Sec-dependent membrane protein insertion, YidC also functions as a separate membrane protein insertase (B, C). The SRP pathway may deliver certain membrane proteins directly to YidC (B), but so far YidC-dependent membrane insertion of authentic E. coli membrane proteins has been demonstrated independently of the SRP pathway only (C).
Chapter 1

The mammalian Sec61β subunit of the protein-conducting channel plays a key role in the transfer reaction as it acts as guanidine nucleotide exchange factor for the β subunit of the SRP receptor (105). In bacteria, so far no functional equivalent to Sec61β has been identified. SecG, the third subunit of the bacterial protein channel is not homologous to Sec61β, and there is no evidence that it would be involved in SRP-mediated protein targeting.

Recently, an alternative model of SRP/FtsY-mediated membrane protein targeting has been proposed by Bibi and coworkers. They observed in vivo that the targeting of ribosomes to the membrane requires FtsY, but not Ffh (108). Furthermore, evidence was presented that FtsY might be targeted to the membrane co-translationally (110) and that depletion of Ffh or the translocase component SecE leads to the accumulation of membrane-bound FtsY-ribosome complexes (111). It was therefore suggested that Ffh might actually act downstream of FtsY-mediated ribosome binding to the membrane. This model is rather controversial as it proposes a role for SRP, which is distinct from the well-established function of SRP in the mammalian system.

Some small phage proteins such as M13 procoat and Pf3 coat insert into the membrane in a Sec-independent manner. These proteins do not utilize the SRP pathway and this is generally attributed to their small size (Fig. 1, pathway C). Essentially, the synthesis of these proteins is nearly completed by the time SRP can associate with their hydrophobic domains at the ribosome tunnel exit. It is not clear, if there are any authentic E. coli membrane proteins that utilize the SRP pathway while inserting into the membrane in a Sec-independent manner. Certain artificial model proteins have been constructed that do require SRP, but insert into the membrane independently of the Sec translocase (Fig. 1, pathway B). This has led to the suggestion that the SRP pathway might not necessarily deliver membrane proteins to the Sec translocase (40, 49, 79). However, because of the aberrant nature of such proteins it is difficult to estimate if these observations bear any significance for the E. coli cell. Crosslinking of translation-arrested nascent chains to Ffh is generally accepted as strong indication that a particular protein is targeted to the membrane by the SRP pathway. At least for small Sec-independent membrane proteins, there seems to be no such correlation. M13 procoat as well as a Pf3 coat-derived construct have been crosslinked to Ffh (35, 232), but membrane insertion of these proteins does not require SRP (35, 47). In case of Pf3 coat, a possible explanation might be found in the fact that the authors used a tag-extended version of Pf3 coat. Therefore, the interaction
with SRP might not be genuine. This was, however, not the case with the nascent M13 procoat protein that was incubated with a lysate of *E. coli* cells (232). One may argue that the crosslink to SRP arises from the fact that these nascent chains are arrested in a way that hydrophobic domains remain exposed at the ribosomal exit tunnel, thereby strongly increasing the time window for an “nonspecific” interaction with the spatially close SRP. In this respect, nascent chains of the secretory protein proOmpA, a protein known to be translocated independently of the SRP pathway, can be crosslinked to Ffh (70) as well, although those studies strongly rely on cellular lysates enriched with SRP. Overall, it seems that arresting a nascent chain on the ribosome may lead to “forced” interactions with protein factors that do not necessarily occur under conditions of ongoing translation. However, for large proteins, there seems to be a good correlation between the crosslinking to Ffh and the SRP requirement for membrane insertion (48). Nevertheless, it is important to validate, whether putative binding partners are indeed required for the catalytic event of membrane protein insertion or protein secretion.

**The role of the Sec translocase in membrane protein insertion**

Most *E. coli* membrane proteins studied so far are inserted into the membrane by the Sec translocase (recent review: 52) (Fig. 1, pathway A). Upon SRP-mediated targeting, the translating ribosome binds to the translocase and the signal sequence of the nascent membrane protein is inserted into the protein-conducting channel (Fig. 3, step 3). Non-translating ribosomes also exhibit a high affinity for the translocase (195). For the eukaryotic system, however, it has been shown that SRP binding to nascent chains confers a selective advantage to the ribosome so that a competitive inhibition of translocase binding by empty ribosomes or ribosomes that synthesize a cytosolic protein is prevented (196). A similar mechanism might be operational in bacteria.

Leader peptidase (Lep) has been the classical model substrate for Sec-dependent membrane protein insertion. This protein spans the membrane twice with a short N-terminal domain as well as the large, catalytic C-terminal domain located in the periplasm. *In vivo* experiments showed that the translocation of the C-terminal
domain requires SecY and SecA (263), whereas the N-terminal domain was found to be translocated independently of the Sec machinery (146). A recent study, however, demonstrated that ribosome-bound nascent chains of Lep consisting of only the first TMS and the cytoplasmic loop are targeted to the Sec translocase (116). What determines the requirement for particular components of the Sec machinery in the insertion of a membrane protein? Andersson and von Heijne have shown that there is a linear correlation between the length of a translocated region and its degree of dependence on SecA and SecY for lengths between 25 and 55 residues (8). The polytopic membrane protein mannitol permease (MlIA) that lacks large periplasmic domains requires SecY, but not SecA, for membrane insertion (132, 259). Several other membrane proteins that are equipped with a large periplasmic domain, like MalF (225, 227), FtsQ (230, 243) or YidC (134, 231), seem to use SecA for the translocation of these domains (Fig. 3, step 5). However, the initial steps of membrane insertion do not require SecA, as ribosome-bound nascent chains of FtsQ can be targeted to proteoliposomes containing SecYEG in the absence of SecA and the TMS of this insertion intermediate can be crosslinked to SecY (212). Taken together, these studies suggest that the role SecA in membrane protein insertion may be restricted to the translocation of large hydrophilic domains. Although the number of membrane proteins analyzed so far is limited, it seems that membrane proteins that are SecA-dependent for membrane insertion also require SecG (134, 231, 243), while membrane proteins that insert independently of SecA also do not seem to require SecG (133). This supports the idea that the function of SecG is to facilitate the SecA reaction cycle (62, 184).

On the structure and the dynamics of the catalytically active Sec translocase

In recent years considerable insight has been obtained in the structure and function of the protein-conducting channel. However, these studies have also fuelled a controversy as to the functional oligomeric state and the conformational dynamics of this channel. Initial electron microscopy (EM) analysis of the purified canine Sec61 complex showed oligomeric ring structures with a central ~20 Å indentation that were
formed only upon the addition of ribosomes, but remained stable after ribosome removal (97). EM pictures of the 80 S ribosome bound to the ER translocon suggested that the ribosomal exit tunnel is aligned with the central pore of an oligomeric ring formed by three to four Sec61αβγ protomers (18). Images with improved resolution showed at a later stage that the exit tunnel is located rather at the periphery of the ring and not in its centre (19, 174). Collisional quenching experiments using intact microsomal membranes indicated that the diameter of the channel formed by the Sec61 translocon varies from 9-15 Å in the resting, ribosome-free state (95) to 40-60 Å in the active, translocating state (94). Manting et al. obtained negative staining images of a tetrameric *E. coli* SecYEG translocase with a central indentation of ~50 Å in diameter, possibly a pore (156). Such structures are formed when SecYEG is incubated with SecA and a non-hydrolysable ATP analog or when a secretory preprotein is trapped as an intermediate in the channel. Two-dimensional crystals of SecYEG, however, showed a dimeric form after three-dimensional reconstruction (27). Based on this reconstruction, the protein-conducting channel was suggested to localize at the dimer interface. At this interface two highly tilted helices seem to close this putative channel at the periplasmic face of the membrane. Therefore, it was suggested that this structure resembles the closed state of the channel. The dimeric forms were also found in the negative stain images of SecYEG, and these were attributed to unassembled SecYEG complexes. Recently, a 3D crystal structure of the Sec translocon from the archaeon *Methanococcus jannaschii* has been reported (Fig. 2) (237). The structure suggests the presence of a central pore within one single Sec61αβγ heterotrimer that constitutes a narrow opening formed by a ring of hydrophobic amino acyl side chains. This pore-like opening is closed on the periplasmic face of the membrane by a re-entrance loop that corresponds to the first periplasmic domain of Sec61α. It was suggested that the structure resembles the closed conformation of the channel that is sealed by a plug formed by this periplasmic loop (237).

Biochemical studies confirm the presence of monomeric, dimeric, and tetrameric *E. coli* SecYEG complexes, but disagree on the active oligomeric state of the membrane domain of the translocon (21, 65, 124, 156, 245, 247, 248, 268). Based on the structural studies with the *M. jannaschii* Sec61αβγ complex, it has been suggested
Fig. 2: The position of the putative lateral gate in the monomeric archaeal Sec61αβγ complex involved in the release of TMS from the protein-conducting channel. The ribbon structure was adapted from van der Berg et al. (237). The structure of the monomeric Methanococcus jannaschii Sec61αβγ complex is believed to represent the closed state of the channel. A top view from the cytosolic side of the membrane is shown. The position of the small subunits Sec61β and Sec61γ, and the large subunit Sec61α are indicated. The dashed arrow shows the putative lateral exit from the Sec61αβγ channel. It is located between TMS 2 and TMS 3 in the N-terminal half of Sec61α and TMS 7 and TMS 8 in the C-terminal half. The black circle indicates the position in which the secY40 mutation is located (A363S) in SecY homologs. This mutation specifically blocks membrane protein integration, but still allows the SecA-dependent translocation of secretory proteins.

that each subunit of such oligomeric constitutes a separate protein-conducting channel, but that one channel is used at a time (237). Obviously, one wonders about the functional role of the oligomeric forms, and it was suggested that they serve to stabilize the individual channels or allow for regulation of channel activity. Moreover,
it is difficult to imagine, how a single Sec61αβγ protomer could form a channel with a diameter around 40-60 Å as measured for the active ER translocon in its native membrane environment (94). A possible explanation for this apparent discrepancy is that the archaeal Sec61αβγ complex visualized in the crystal structure represents an inactive, “stand-by” conformation and it can only be speculated, how the translocon rearranges upon substrate binding. This rearrangement might involve additional proteins, which are present in the native microsomal membranes used for the collisional quenching experiments, but lack in the crystallized complex (6).

It should be emphasized that the route of a secretory protein or a secreted hydrophilic domain through the translocon might differ from the route a TMS takes through the channel. In the latter case, the TMS has to leave the channel laterally into the lipid bilayer (see below). The membrane insertion and translocation modes of the translocase might also correspond to different oligomeric states. The trigger to adopt one or the other complex form could be the different targeting routes delivering the substrate to the translocase.

During co-translational protein integration into the ER membrane, the ribosome remains bound to the Sec61 translocon in such a way that also cytosolic domains of the inserting membrane protein make contact with the Sec61 complex (176). However, the release of hydrophilic domains into the cytosol requires a breakage of the ion-tight ribosome-translocon seal that is formed upon targeting of a ribosome-nascent chain complex (42, 43). Early after targeting as well as in the ribosome-free state, the luminal site of the translocon is closed the Hsp70 chaperone BiP (95). After the initiation of translocation, the luminal site of the pore is open allowing the nascent protein to pass through (148). As soon as a TMS is detected (see below), BiP closes the channel again and briefly after this event, the pore becomes accessible from the cytosolic site indicating that the ribosome-translocon junction has been opened to allow the passage of a protein domain into the cytosol (148). These observations reveal that the Sec61 channel is never in an aqueous continuum with both the cytosol and the ER lumen. It is not known, how exactly the permeability barrier is maintained during bacterial membrane protein insertion as bacteria as well as archaea do not have a BiP-like protein to seal the translocon at the trans site of the membrane. On the cytosolic site, the evolutionary conserved ribosome-translocon interactions are mainly mediated by the large ribosomal subunit RNA (19, 195). However, additional proteins appear to be required to establish the ion-tight ribosome-translocon junction, which
are not present in the purified ribosome-translocon complexes examined by EM, as such a tight seal was not observed in these images (18, 19).

In bacteria, true co-translational translocation of large hydrophilic domains driven by chain elongation does not seem to occur. Instead, SecA is required to mediate the translocation step (Fig. 3, step 5). The cytoplasmic domains 5 and 6 of SecY have been implicated in SecA binding to the translocase (175). Recent data suggests that SecA and the translating ribosome can bind simultaneously and noncompetitively to the translocase (270) (Fig. 3). Under translocation conditions, SecA seems to possess a ribosome-translocase dissociation activity that is coupled to the ATP-dependent membrane cycling of SecA (Fig. 3, step 5). The observation that ribosome-bound nascent chains can be crosslinked to SecA is also consistent with a simultaneous binding of SecA and the ribosome to the SecYEG complex. However, with respect to the dimensions of the SecYEG channel, the ribosome and SecA, it is difficult to imagine how this is achieved.

The role of the proton motive force in membrane protein biogenesis

An important factor in the biogenesis of membrane proteins is the proton motive force (PMF) across the bacterial cytoplasmic membrane. For Lep, it has been shown that the PMF is required for the translocation of the C-terminal periplasmic domain (262). The PMF promotes the translocation of negatively charged residues across the membrane (9, 31). Other studies suggest a similar role for the PMF in controlling the transmembrane topology of membrane proteins, not only for those that use the Sec machinery for integration but also for Sec-independent proteins (56, 127, 128). In addition, the presence of a PMF is inhibitory for the translocation of domains with positive charges across the membrane (211), and it has been suggested that the PMF contributes to the so-called “positive inside rule” (250). Positively charged residues are predominantly found in cytosolic domains of membrane proteins, and excluded from extracellular loops. However, the electrostatic interaction of positively charged residues with anionic phospholipids might be more important in this context (246) as the membrane proteins of some thermoacidophilic archaea also adhere to “positive inside rule” whereas the transmembrane potential ($\Delta\psi$) across these membrane bears an opposite polarity (outside positive instead of negative) (236).
Recently, the Sec-dependent insertion of FtsQ (Fig. 3) was reconstituted into proteoliposomes with purified SecYEG (243). FtsQ consists of a small cytoplasmic N-terminal domain, one TMS and a large periplasmic C-terminal domain. Translocation of the C-terminal domain into proteoliposomes was shown to require the Sec machinery as well as a $\Delta \psi$. The $\Delta \psi$ might function as an additional driving force for the transversal movement of this domain (Fig. 3, step 5). It could also mediate the correct positioning of the TMS ($N_{\text{in}}-C_{\text{out}}$) of FtsQ prior to translocation of the C-terminal domain (Fig. 3, step 3). The charge distribution around this TMS and the helix dipole moment suggest that such a mechanism may be valid. Goder and Spiess recently showed that in the ER translocon a TMS initially always orients in an $N_{\text{trans}}-C_{\text{cis}}$ conformation and that an inversion into the opposite orientation occurs within a few seconds (88). The efficiency of this re-orientation was found to depend on the charge distribution around the TMS. In bacteria, a similar mechanism might be supported by the $\Delta \psi$. The PrlA4 mutant of SecY has been shown to suppress the PMF dependence of protein translocation (185). Interestingly, with this mutant the membrane insertion of FtsQ remains strictly $\Delta \psi$-dependent, which suggests a role of the PMF in membrane protein insertion that is different from that in protein secretion.

**Lateral release of TMSs from the translocase into a lipid environment**

It is generally believed that the Sec translocase forms an aqueous protein-conducting channel in the membrane (see above). When we assume that membrane proteins use the same pathway, it will be necessary to laterally release the inserting TMSs into the lipid bilayer. Such a lateral opening towards the lipid bilayer was first demonstrated for the Sec61 translocase by Martoglio and Dobberstein (158). Generally, three classes of TMSs can be distinguished. Type I TMSs induce the translocation of an N-terminal domain. Such TMSs appeared to be membrane-inserted in a Sec-independent manner as for example the first TMS of ProW (260) and Lep (146). The observation that the type I TMS of Lep is targeted to the Sec translocase *in vitro* has challenged this view (116). Type II TMSs or signal anchor sequences promote the translocation of a C-terminal domain comparable to the cleavable signal sequence of secretory
proteins. In some cases, like the Sec-independent M13 procoat protein or cytochrome o oxidase subunit A (CyoA, see below), the N-terminus of the mature membrane protein is translocated to the periplasm with the help of a cleavable signal anchor sequence. Once C-terminal translocation has been initiated by a type II signal, the next TMS is a so-called stop-transfer sequence. Therefore, most Sec-dependent polytopic membrane proteins can be seen as a series of alternating, independent signal anchor and stop transfer elements that sequentially insert into the membrane. However, for some proteins, the insertion mechanism seems to be more complex. In case of the tetracycline resistance protein (TetA), the efficient insertion of an odd-numbered TMS requires the presence of the following even-numbered segment indicating that they might insert in pairs, probably as helical hairpins (93). Similar examples of cooperative TMS insertion are known for the ER membrane (recent review: 6).

Currently, it is unclear, if TMSs leave the translocase one by one, in pairs or all together after the termination of translation. For the eukaryotic Sec61 complex, it has been shown that before the translation of a membrane protein is terminated, a TMS can laterally exit the translocation channel and enter the lipid environment (176). Heinrich et al. suggested that the Sec61 complex together with a translocon-associated protein named TRAM forms a hydrophobic partitioning surface along which the TMSs diffuse into the lipid bilayer (100). The equilibration of a TMS would depend on its hydrophobicity and its distance from the ribosomal peptidyl transferase center. Heinrich and Rapoport studied the integration of the bacterial protein Lep into the ER membrane (101). While the first, long and hydrophobic TMS (H1) inserts into the lipid bilayer on its own, the second, shorter TMS (H2) collaborates with the previously inserted H1 during its integration. H1 diffuses away from the Sec61 complex in the absence of H2, but is still close to Sec61 when H2 arrives inside the channel. This data also shows that the requirement for a downstream TMS does not exclude a sequential insertion mode. Other crosslinking studies, however, indicated that the nascent TMSs are rather bound to translocon proteins at distinct, fixed positions during membrane insertion, where they remain for a prolonged time, and that their lateral release might be actively regulated by the Sec61 complex and TRAM (58, 161, 162). A possible accumulation of TMSs in the translocon before they are released into the lipid bilayer has been demonstrated as well. For instance, the first halves of newly synthesized P-glycoprotein (26) and LacY
(179) remain sensitive to extraction with urea until the proteins are completely synthesized. This suggests that these proteins remain in a hydrophilic environment and are released into the lipid bilayer only when the ribosome has released the polypeptide chain. However, it is difficult to imagine how a large helix-bundle would fit into the aqueous lumen of the protein-conducting pore, especially if one anticipates that the Sec translocase functions as a single heterotrimer with a central channel. How does the secretion machinery recognize a TMS? It has been shown that the presence of hydrophobic stretches in secretory proteins delays their translocation (208). Hydrophobic sequences with increasing length introduced into proOmpA halt translocation and cause partitioning into the lipid bilayer (63). Thus, it seems that hydrophobic TMSs first induce a pausing of transversal movement and this creates a time window for lateral escape of the TMS into the lipid bilayer. It has been suggested that this process is kinetically determined, i.e., slow translocation would allow more time for this lateral release and, thus, increase the likelihood of membrane partitioning. The opening or closing of each end of the protein-conducting channel must be tightly controlled in a sequence that does not compromise the membrane's permeability barrier (see above). Liao et al. (148) demonstrated that in the ER, these structural changes in the Sec61 complex are induced by a TMS in a nascent membrane protein already from inside the ribosome. Thus, the ribosome, not the translocase, first recognizes the TMS and triggers conformational changes within the translocase that may be involved in shifting its function from translocation to integration. Fluorescence resonance energy transfer experiments revealed that a TMS within a nascent membrane protein folds into a compact conformation near the peptidyl transferase center of the ribosome, which is compatible with a \( \alpha \)-helix (265). This folding is induced and stabilized by the ribosome, as it disappears upon emerging from the exit tunnel. Moreover, it coincides in time with the sequential opening and closing of the two ends of the translocon pore (see above).

Depending on the oligomeric state of the translocase, different mechanisms of lateral opening are possible. An oligomer forming a central translocation pore could either transiently dissociate or open like a hinge to release a TMS into the lipid bilayer. Based on the structure of the \( \textit{M. jannaschii} \) Sec61\( \alpha\beta\gamma \) complex that constitutes a single copy of the heterotrimer, the putative central pore will need to open laterally to release the TMS. It has been suggested that this lateral gate of the translocase is located between TMS 2 and TMS 3 in the N-terminal half of SecY/Sec61\( \alpha \) and TMS
7 and TMS 8 in the C-terminal half (Fig. 2) (237). The inserting TMS would be positioned at the interface between the pore and lipids, at the site that is also suggested to bind the signal sequences of secretory proteins. Interestingly, a point mutation in secY (secY40) has been described that is selectively impaired in membrane protein insertion (183). This mutation causes an alanine to serine substitution in the fifth cytosolic loop of SecY which is adjacent to TMS 8 (Fig. 2). It might affect the conformational flexibility of this region suggesting that the membrane insertion phenotype of secY40 is related to a defect in lateral release of TMSs. However, the precise path of TMSs through the Sec translocase and the molecular mechanism of TMS release are far from understood.

The role of YidC in Sec-dependent membrane protein insertion

YidC is an essential membrane protein in E. coli, and a member of the Cytochrome Oxidase Biogenesis (Oxa) membrane protein family (140). Oxa1p is located in the inner membrane of mitochondria (126), and mediates the insertion of membrane proteins synthesized in the mitochondrial matrix (104). Albino3 (Alb3) is a homolog of the chloroplast thylakoids (221) and is needed for the membrane insertion of subunits of the light-harvesting complex (172). Although the homology between YidC and mitochondrial Oxa1 was known for some time, it took until 2000 that Scotti et al. (213) demonstrated by a photocrosslinking approach that the TMS of a ribosome-bound, membrane-inserted nascent FtsQ is in close contact with YidC (Fig. 3, step 4). Furthermore, they showed that a part of the cellular YidC pool is associated with the Sec translocase. This work was followed by the demonstration that YidC is essential for the membrane insertion of the Sec-independent M13 procoat protein (205). Further crosslinking experiments revealed that YidC contacts many Sec-dependent membrane proteins such as Lep (115, 205), MtIA (17) and YidC itself (134, 231).

An unresolved question is how the Sec-translocase and YidC cooperate in the insertion of membrane proteins. Some of the cellular YidC protein is associated with the Sec translocase. This interaction is probably mediated by the SecDFyajC complex (187), which in turn is loosely associated with the SecYEG complex. The exact role
Introduction

of the SecDFyajC complex is not known. It stimulates preprotein secretion, probably by modulating the SecA reaction cycle (61, 62), but various other roles have been suggested as SecD and SecF are also present in archaea that lack a SecA homolog. A close association of YidC and the Sec translocase also follows from crosslinking experiments that probe the order of interactions of a growing nascent chain. Using FtsQ insertion intermediates of various sizes, it was shown that the insertion of the single TMS occurs via a number of defined steps whereby the protein first interacts with SecY before it gains access to YidC and lipids (Fig. 3, steps 3-4) (230, 241). Lep, however, appears to insert into a combined SecY/YidC environment immediately after the targeting event (116). The first TMS of Lep remains bound to YidC, until the second TMS arrives in the translocase. Consistently, the first three TMSs of MtlA were shown to accumulate in close proximity to YidC during their insertion (17). These data led to the hypothesis that YidC might facilitate the lateral release of TMSs from the Sec translocase into the lipid bilayer, comparable to the TRAM protein in the ER membrane. The crosslinking data are in agreement with distinct interaction patterns of TMSs with translocon components and a regulated release into the lipid bilayer (see above).

Although TMSs of various membrane proteins have been crosslinked to YidC, functional studies indicate that YidC is not essential for their membrane insertion. The in vivo depletion of YidC has no or little effect on the membrane insertion of FtsQ (230), while its effect on Lep insertion appears to be more pronounced (205). However, even with Lep, membrane integration is not completely blocked upon YidC depletion. YidC also is not involved in the translocation of secretory protein in vivo (205) and in vitro (241). Interestingly, overexpression of Lep under YidC depletion conditions, leads to an accumulation of proOmpA in the cytosol (206). A possible explanation for this phenomenon could be that in the absence of YidC, jammed TMSs block the translocase. These in vivo experiments, however, should be carefully interpreted. Depletion of an essential membrane protein in combination with the high level overexpression of another membrane protein imposes an extreme stress on the cell, which can lead to strong, often unrelated, pleiotropic effects. For instance, YidC depletion causes a dramatic reduction of the ability of the cells to generate a PMF (242). This is due to defects in the functional assembly of cytochrome o oxidase and the F1F0 ATP synthase complex, which is reminiscent of the effects of OXAl deletion in yeast on energy transducing complexes in the mitochondrial inner membrane (see
Fig. 3: Model of the Sec-dependent membrane insertion of FtsQ. SRP binds to the hydrophobic signal sequence of FtsQ as it emerges from the ribosome (step 1). Upon interaction of SRP with its membrane-bound receptor FtsY (step 2), the ribosome-nascent chain complex is transferred to the SecYEG translocase, where the initial membrane insertion takes place (step 3). At this stage, the TMS of nascent FtsQ is in close proximity to SecY. Concomitant with chain elongation, the TMS moves from a SecY environment to a position in which it can be crosslinked to SecY, YidC and phospholipids (step 4). YidC appears to facilitate the lateral release of the TMS from the SecYEG complex into the lipid bilayer. SecA mediates the translocation of the large hydrophilic C-terminal domain of FtsQ across the membrane (step 5). The Δψ that is required for membrane insertion of FtsQ may act on the correct positioning of the TMS (N-in-C-out) in the SecYEG channel (step 2-3) and / or on the SecA-mediated translocation step. Finally, after termination of translation, the ribosome is released from the SecYEG complex and FtsQ acquires its functional conformation in the membrane (step 6).

below). As the PMF is required for the membrane insertion of FtsQ and Lep, the effects of YidC depletion on the biogenesis of these proteins might rather be indirect. The recent reconstitution of FtsQ membrane insertion into SecYEG proteoliposomes clearly established that YidC is not required for this process (243). The TMS of FtsQ was shown to stably integrate into the lipid bilayer even in the absence of YidC while the C-terminal domain is translocated in a Δψ-dependent manner. The latter explains the weak effect of YidC depletion on FtsQ insertion in vivo. The reconstituted system
allowed for the first time to clearly separate the contributions of YidC and the PMF and will be a valuable tool for further studies on Sec-dependent membrane insertion. It is still possible that YidC is involved in the hydrophobic partitioning of Sec-dependent TMSs, but at least for FtsQ, it seems that such effects more relate to the kinetics of the reaction as there is no strict YidC dependence. Membrane proteins or even specific TMSs of certain membrane proteins might differ in their YidC requirement, but the features of such TMSs remain to be elucidated.

Besides affecting the kinetics of the release of certain TMSs from the translocase, YidC might more-or-less nonspecifically interact with any TMSs while they emerge from the translocase and protect them from non-productive interactions or stabilize intermediates of membrane protein folding pathways. Therefore, YidC might not be essential for membrane integration per se, but still play an important role as a kind of chaperone in membrane protein assembly in vivo. Experimental support for this hypothesis comes from a recent study by Nagamori et al., who have shown that YidC is required for the correct folding of the polytopic membrane protein LacY, while it is dispensable for its membrane integration per se (180).

In addition, Fröderberg et al. demonstrated that YidC depletion inhibits the Sec-dependent translocation of two lipoproteins, i.e. murein lipoprotein (Lpp) and bacteriocin release protein (Brp) (80). Nascent chains of both proteins could be crosslinked to Ffh, SecA, SecY and YidC, which is very similar to the typical crosslinking pattern obtained with nascent membrane proteins.

In the Gram-positive bacterium Bacillus subtilis, two YidC homologs were found, i.e. SpoIIIJ and YpjG (177, 226). The presence of either SpoIIIJ or YqjG is required for cell viability. Whereas SpoIIIJ is needed for sporulation, YqjG is dispensable for this developmental process. Interestingly, SpoIIIJ- and YqjG-limiting conditions result in a strong post-translocational defect in the stability of secretory proteins, whereas the effects on membrane protein biogenesis are only mild (226). However, it should be emphasized that also in E. coli, YidC depletion does not cause a general decrease in steady-state levels of membrane proteins, but rather specific defects in the biogenesis of certain membrane protein complexes (see below). In addition, it is not clear, if the observed protein secretion defect in B. subtilis is directly caused by SpoIIIJ- and YqjG-limiting conditions or rather due to an unrecognized assembly defect of a specific membrane protein or by an impairment of the PMF.
YidC mediates insertion of membrane proteins independently of the Sec translocase

Surprisingly, Samuelson et al. observed that the Sec-independent insertion of the phage protein M13 procoat into the inner membrane of E. coli strictly requires YidC in vivo (Fig. 1, pathway C) (205). M13 procoat membrane insertion was long believed to be a spontaneous event and a wealth of literature has been published that underscored this hypothesis (reviewed in 139). M13 procoat consists of two antiparallel transmembrane helices; the first TMS represents a cleavable signal peptide while the second TMS is a structural unit of the final coat. The hydrophobic domain acts as a scaffold for the assembly of the coat, and therefore does not represent an authentic membrane protein but rather a membrane-inserted intermediate of a complex assembly reaction. In the membrane-inserted, but unprocessed protein, the N- and C-termini remain in the cytosol, whereas the short loop connecting the TMSs is translocated to the periplasm (Fig. 5 B). The latter process is PMF-dependent (45), which raised the question, whether the effect of YidC depletion on M13 procoat insertion could be indirect. However, also PMF-independent point mutants of M13 procoat appear to require YidC for efficient membrane insertion in vivo, although it is remarkable that the insertion of these proteins is less dramatically affected under YidC depletion conditions (206). Temperature-sensitive variants of YidC exhibit the M13 procoat insertion defect already briefly after shift to the non-permissive temperature, earlier than the PMF defects are observed upon YidC depletion (36). This indicates that YidC indeed mediates the membrane insertion of M13 procoat in vivo. Chen et al. further demonstrated that a second Sec-independent phage protein, Pf3 coat, requires YidC for membrane insertion both, in vivo and in vitro (35). A direct interaction between YidC and Pf3 coat was demonstrated by crosslinking. Pf3 coat is a polypeptide of only 44 amino acids that spans the membrane once with the N-terminus translocated to the periplasm. Therefore, it resembles the second TMS of M13 procoat. Insertion of Pf3 coat is also strongly dependent on the PMF (128). Recently, the reconstitution of Pf3 coat membrane insertion into YidC proteoliposomes was demonstrated (214).

What might be the mechanistic role of YidC in Sec-independent membrane protein insertion? For M13 procoat, early models subdivided the insertion process in four
Introduction

steps: 1) electrostatic binding to the membrane via positively charged residues in the N- as well as the C-terminal region; 2) hydrophobic interaction of the TMSs with the acyl chains of the membrane phospholipids; 3) PMF-driven, electrophoretic translocation of the negatively charged periplasmic loop across the membrane; and 4) the final cleavage in this loop by Lep (Fig. 5 B) (136). It was shown, that both hydrophobic domains are essential for the translocation of the central loop indicating that an interaction between the two TMSs facilitates membrane insertion (31, 137). This helical hairpin formation is required due to the high abundance of charges in the loop that needs be translocated (201) (Fig. 5 B). In turn, however, the negatively charged residues are not essential for the translocation of the loop. Interestingly, although the negative charges acquire a topogenic character in the presence of a PMF (31), the synergistic insertion of the two TMSs by hydrophobic interactions alone can drive the translocation of positive charges (138). Therefore, it appears that hydrophobic forces and the PMF contribute independently to membrane insertion and that both parameters can at least partly compensate a defect or reduction in the other (Fig. 5 B). It seems that electrostatic, electrophoretic and hydrophobic forces together are needed to realize a net negative free energy of the membrane insertion reaction. *In vivo*, depletion of YidC does not affect membrane binding of M13 procoat (206), the initial step of M13 procoat insertion. In addition, PMF-independent mutants of M13 procoat remain YidC-dependent (206, see above) for their membrane insertion indicating that YidC and the PMF mediate different steps of the insertion reaction. Thus, it appears that YidC acts on the insertion step that involves hydrophobic forces and helical hairpin formation.

For Pf3 coat, a somewhat different scenario seems to exist as the protein contains only one TMS. Mutations in YidC have been identified that specifically affect either M13 procoat or Pf3 coat insertion (36). The negatively charged residues in the N-terminus of Pf3 coat appear to move electrophoretically across the membrane as this process is driven by the PMF (127). The PMF therefore controls the final transmembrane topology. A mutant Pf3 coat protein (Pf3-4N) that lacks charged residues fails to insert into the membrane (128). However, when the hydrophobicity of the TMS is increased (Pf3-4N3L), the hydrophobic forces can overcome the energetic barrier imposed by the polar phospholipids headgroups and drive the spontaneous membrane insertion without topological control (128). Pf3-3L does not require YidC for insertion (214). Similar results have been obtained for the membrane insertion of a
Sec-independent fusion protein that consists of the N-terminal periplasmic domain of Pf3 coat and the first TMS of Lep (56). These results confirm the synergistic action of the PMF and hydrophobic forces as described for M13 procoat. In the absence of YidC in vivo, ~60% of wild-type Pf3 coat does not partition into the membrane (35). The remaining 40% which are not extractable from the membrane by alkaline treatment fails to acquire the correct transmembrane topology. This again might be caused be the collapse of the PMF under YidC depletion conditions. However, the interaction of YidC with the Pf3 coat nascent chains is independent of the PMF which supports the notion that YidC and the PMF act on different steps of the insertion reaction (35).

The data on M13 procoat as well as Pf3 coat strongly suggest that the function of YidC is to facilitate the hydrophobic interaction of TMSs with the membrane lipids. It seems likely that YidC provides a special amphiphilic surface that lowers the activation energy of the hydrophobic partitioning event for instance by overcoming the repulsion of the hydrophobic protein segments by polar lipid head groups (Fig. 5). In addition, YidC might shield polar residues on their way through the hydrocarbon core of the membrane and promote the folding of hydrophobic regions within a membrane protein into its transmembrane conformation (214). As in case of M13 procoat, this process might involve helical hairpin formation.

Recently, a mutagenesis study has been carried out to identify regions in YidC that are critical for function (121). As a substrate to test YidC activity, a YidC-dependent fusion protein consisting of the C-terminal catalytic domain of Lep fused to the M13 procoat protein was used. E. coli YidC consists of an N-terminal TMS followed by a large periplasmic domain and five closely-spaced C-terminal TMSs. It appears that the large periplasmic domain and residues in the exposed hydrophilic loops or C-terminal tail are not critical for YidC activity. Rather, the five C-terminal TMSs that contain the three consensus sequences in the YidC/Oxa1/Alb3 family are important for its function, as deletions in this part of the protein are not tolerated (121). However, the mutagenesis study also showed that the precise sequence of these TMSs is not essential for YidC activity. Therefore, it was suggested that the TMSs of YidC function as a supportive platform for the membrane-inserting substrate protein. It is, however, very difficult to reconcile this hypothesis with the specific role of YidC in the membrane insertion of a number of small integral membrane proteins. On the other hand, the thylakoidal Albino3 protein can functionally complement YidC in E.
Introduction

coli, which further suggests conserved functions in membrane protein integration (120).

On the physiological role of YidC

As discussed, YidC probably plays a kinetic role in the process of Sec-dependent membrane protein biogenesis and might act as a chaperone in the organization of helix-helix-interactions without being absolutely required for the insertion of Sec-dependent substrate proteins (at least the ones tested so far). In addition, M13 procoat and Pf3 coat strictly depend on YidC for membrane insertion, but these phage proteins do not represent authentic E. coli substrates nor are they essential for E. coli viability. This raises the question why is YidC an essential protein in E. coli (205). To address this question, the physiological role of YidC was studied using a YidC depletion strain (242). It was observed that YidC depletion causes a strong induction of the membrane stress protein PspA. High-level expression of PspA signals an impairment of the PMF (130). Further studies revealed that cells grown under YidC depletion conditions as well as inner membrane vesicles prepared from YidC-depleted cells indeed exhibit a strongly reduced ability to generate a PMF. This was observed with both ATP and oxidizable respiratory chain substrates like NADH. The former correlates with decreased levels of functional F1F0 ATP synthase. In particular, YidC depletion leads to strongly reduced amounts of the small, ring-forming rotor subunit c (F0c) of the membrane-integral F0 subcomplex (recent reviews: 33, 257) in the membrane (Fig. 4 A). The reduced respiratory capacity was found to be due to decreased levels of cytochrome o oxidase. In particular, the levels of the ubiquinol-binding subunit a (CyoA) of cytochrome o oxidase (Fig. 4 B) in the membrane are severely reduced under YidC-depleting conditions. Since these two enzymes are the key energy-transducing protein complexes of aerobically grown E. coli, it seems that the essential role of YidC for E. coli viability is related to the biogenesis of these complexes. Yi et al. reported that F1F0 ATP synthase subunits F0c as well as F0 subunit a (F0a) and the Sec translocase subunit SecE require YidC for membrane insertion in vivo (268). However, F0a does not stably integrate into the membrane in the absence of F0c (106), and therefore it is not clear, if the insertion defect of F0a is directly related to the loss of YidC. Rather it is more likely that the unassembled F0a
becomes proteolyzed by FtsH, a membrane-bound protease of the AAA family that is involved in quality control of newly inserted membrane protein complexes (5). Also, the observed insertion defect of SecE has to be assessed critically. SecY cannot assemble stably into the membrane in the absence of SecE, and SecE is essential for the functioning of the Sec-translocase. A membrane insertion defect of SecE would imply a strong pleiotropic effect on protein secretion and membrane protein insertion in general. However, this is not observed in the YidC depletion strain.

A *Saccharomyces cerevisiae* strain, in which the mitochondrial YidC homologue Oxa1p (see above) is defective, shows a similar reduction in the amount of functional F$_1$F$_0$ ATP synthase and cytochrome $c$ oxidase (7, 25). In the mitochondrial inner membrane, Oxa1p is required for the PMF-dependent insertion of membrane proteins that are encoded by the mitochondrial DNA (104). These proteins are subunits of large energy-transducing membrane protein complexes. Cytochrome $c$ oxidase subunit II (Cox2p), which is a homolog of CyoA (Fig. 4 B), was the first inner membrane protein shown to require Oxa1 (15, 99, 102). Direct interaction of Cox2p with Oxa1p was demonstrated by crosslinking (103). Also in *Neurospora crassa*, an Oxa1p homolog is involved in the biogenesis of respiratory chain membrane protein complexes (181). Recently, a second homolog named Oxa2p was identified in *N. crassa*, which is also required for the biogenesis of cytochrome $c$ oxidase (81). Like Oxa1p, Oxa2p interacts with the cytochrome $c$ oxidase subunits Cox2p and Cox3p (Fig. 4 B). In addition, *N. crassa* Oxa2p complements yeast mutants that are deficient in the cytochrome $c$ oxidase biogenesis protein Cox18p (207, 218) suggesting at least partially overlapping functions of both proteins. In case of yeast F$_1$F$_0$ ATP synthase, the amounts of the F$_0$ subunits Atp6p (subunit a), Atp17p (subunit f) and Atp4p (subunit b) (Fig. 4 A) were strongly reduced in a yeast OXA1 deletion strain, while a mild reduction of Atp9p (subunit c) levels was observed (147). Interestingly, the F$_1$F$_0$ ATP synthase defect could be complemented by a co-deletion of Yme1p, a protease of the AAA family that has been implicated in the degradation of mis-folded inner membrane proteins. This observation suggests that in the biogenesis of mitochondrial F$_1$F$_0$ ATP synthase, Oxa1p exerts a chaperone function by protecting unfolded or unassembled proteins from rapid proteolysis.
Introduction

Fig. 4: Transmembrane topologies of the membrane-integral F₀ ATP synthase subcomplex and cytochrome o oxidase subunits. In brackets, the names of the homologous proteins in yeast are given.

The F₀ domain of the *E. coli* F₁F₀ ATP synthase consists of the subunits F₀a, F₀b and F₀c (stoichiometry a₁b₂c₁₀) (A). F₀a is involved in vectorial proton flow. Together with the F₁ domain-interacting F₀b, it constitutes a part of the stator complex. F₀c assembles into ring-like oligomeric structures that together with the F₁ subunits γ and ε form the rotor domain of the F₁F₀ ATP synthase (Capaldi & Aggeler, 2000). *E. coli* cytochrome o oxidase consists of three evolutionary conserved core subunits named CyoA, CyoB and CyoC (B). CyoA is synthesized as a precursor protein. The first TMS represents a cleavable signal peptide, which is proteolytically removed upon membrane insertion. Subsequently, the N-terminal cysteine of the mature bacterial protein is lipid-modified. In the large C-terminal periplasmic domain, CyoA contains the ubiquinol binding site of the complex. CyoB contains two heme cofactors and a copper atom. It binds and reduces oxygen and couples the electron flow from ubiquinol to oxygen to the vectorial transfer of protons across the membrane. CyoC is required for the assembly of functional cytochrome o oxidase, but its function is not understood (82).
Recently, it was shown that mitochondrial ribosomes bind directly to the C-terminal domain of Oxa1p suggesting a co-translational insertion mechanism (119, 222). This C-terminal domain is not present in *E. coli* YidC but also not in Oxa2p. This might indicate that in terms of substrate recognition and mechanism, YidC is more similar to Oxa2p. In addition, Oxa1p has also been implicated in the membrane insertion of nuclear-encoded proteins, like Oxa1p itself, which appear to be initially translocated into the mitochondrial matrix and subsequently inserted into the inner membrane (107).

Mitochondria do not contain a Sec-translocase, thus it appears that only the Sec-independent function of YidC/Oxa1p is evolutionary conserved. On the other hand, the thylakoid membranes of plant chloroplasts contain both, a Sec translocase and the YidC/Oxa1p homolog Alb3 (221). A direct interaction between the thylakoid Sec translocase and Alb3 has been demonstrated (131). However, the only thylakoid membrane proteins so far shown to require Alb3 for insertion are nuclear-encoded light harvesting chlorophyll binding proteins (LHCPs) (20, 172, 264). These subunits do not require the Sec translocase for membrane insertion (172). Interestingly, LHCPs seem to be post-translationally targeted to Alb3 via the SRP pathway (173). Taken together, it appears that the physiological role of the YidC/Oxa1p/Alb3 membrane protein family primarily concerns the insertion and assembly of energy-transducing membrane protein complexes.

The biogenesis of bacterial $F_0c$ was recently studied in detail by van der Laan *et al.* (244). $F_0c$ has a molecular mass of 8.3 kDa and consists of two TMSs connected by a small polar loop that is exposed to the cytoplasm (Fig. 4 A). Insertion of *in vitro* synthesized $F_0c$ into membranes derived from cells grown under YidC-depleting conditions was strongly reduced compared to wild-type membranes. Since $F_0c$ inserts into the membrane independently of the PMF, this effect seems to relate directly to a role of YidC. Importantly, YidC proteoliposomes mediate the efficient and topologically oriented membrane insertion of $F_0c$, while the Sec translocase and the SRP targeting pathway are not required. Upon insertion into YidC proteoliposomes, $F_0c$ assembles into large oligomeric complexes in an analogous manner as *in vivo*. These data demonstrate that YidC functions as a membrane protein insertase for the $F_0c$ subunit (Fig. 5 A), which represents the first discovered physiological substrate of this novel insertion pathway. The membrane insertion requirements of $F_0c$ are very similar to that described for M13 procoat and Pf3 coat (see above, Fig 5). Therefore, it
**Fig. 5: Model of YidC-mediated membrane protein insertion.** The physiological substrate $F_{0c}$ initially binds to the membrane via electrostatic interactions (*A, step 1*). Subsequently, YidC provides an amphiphilic surface that facilitates the hydrophobic partitioning of $F_{0c}$ into the membrane and helical hairpin formation (*A, step 2*). These hydrophobic interactions drive the translocation of the N- and C-terminal domains across the membrane. Positively charged amino acid residues anchor the hydrophilic loop that connects the two TMSs at the cytosolic surface of the membrane. For the insertion of M13 procoat, the common principle of electrostatic membrane binding and hydrophobic partitioning facilitated by YidC is identical to $F_{0c}$ insertion (*B, step 1 & 2*). However, the charge distribution in M13 procoat is inverted. Therefore, the N- and C-terminal domains are retained in the cytosol and the translocation of the negatively charged hydrophilic loops is strongly stimulated by the PMF.
seems that these phage coat proteins employ an essential host cell membrane protein biogenesis pathway in order to become inserted into the membrane, a prerequisite for phage particle assembly. The only difference is that F₀c insertion does not depend on the PMF. The short N-terminal, periplasmic tail of F₀c contains two negative charges (Fig. 5 A). However, a possible electrophoretic contribution to their translocation is obviously not required.

Interestingly, the stepwise re-addition of charged amino acids into an uncharged Pf3 coat variant (Pf3-4N) demonstrated that negatively charged residues in the periplasmic N-terminal domain show a clear electrophoretic response only when the hydrophobicity of the TMS is limiting (128). F₀c, however, is a very hydrophobic protein. This could explain why YidC-mediated hydrophobic interactions in combination with helical hairpin formation seems to be the sole driving force for F₀c membrane insertion. A possible electrostatic role of the positively charged residues in the cytoplasmic loop of F₀c remains to be elucidated. These positive charges might help to prevent translocation of this domain and to anchor it to the cytoplasmic face of the membrane by the interaction with negatively charged phospholipids (Fig. 5 A).

Future perspectives

Understanding the mechanistic role of YidC and its homologs in the insertion and assembly of large energy-transducing membrane protein complexes will be a major challenge for the future. How are other membrane-integral subunits of cytochrome o oxidase and F₁F₀ ATP synthase inserted? For instance, CyoA contains a large, C-terminal periplasmic domain, and this likely requires SecYEG and SecA for its correct localization (Fig. 4 B). It is therefore well possible that YidC and SecYEG cooperate in the membrane insertion of CyoA, and this would define CyoA as another class of YidC-dependent proteins next to the F₀c, which requires YidC alone. Many open questions concern the mechanism of Sec-independent membrane protein insertion via YidC. For example, how does YidC interact with its substrates and how does it facilitate their hydrophobic partitioning into the membrane? A three-dimensional structure of YidC would be very helpful to address these issues.
Moreover, it remains to be elucidated, how YidC interacts and cooperates with the Sec translocase. Where is YidC positioned relative to the SecYEG channel during membrane protein insertion and how does it affect the kinetics of TMS release from the Sec translocase? Another possible function for YidC is to act as a chaperone in the folding of membrane proteins. The observation that YidC is required for folding and stability of the polytopic membrane protein LacY points into this direction. It is an intriguing question, whether YidC is also involved in the assembly of large oligomeric membrane protein complexes.

A general concern with studies on membrane protein insertion is the unrestricted use of tagged proteins or artificial fusion constructs instead of the naturally occurring *E. coli* inner membrane proteins. Membrane proteins contain many elements that contribute to the targeting information and/or topogenesis. Such signals in the primary sequence and secondary structure may synergistically control the integration into the membrane. Evidently, the introduction of tags may interfere with these signals, and results may be obtained that do not tell much about the *in vivo* membrane insertion process. A co-evolution of membrane proteins and their respective membrane insertion machineries may be assumed. Therefore the physiological significance of data obtained with artificial model proteins may be limited. In addition, *in vivo* studies often make use of depletion strains, in which plasmid-encoded model proteins are expressed in order to determine their requirements for membrane insertion. The depletion of essential proteins, like SecY, SecE, Ffh or YidC, however, imposes an extreme stress on the cells, which may lead to all kinds of pleiotropic effects that are difficult to distinguish from the direct consequences of the loss of a certain protein. Consequently, also these results have to be interpreted with great care, and their validation *in vitro* is absolutely required.

It will be of great interest to study a much larger number of *E. coli* proteins from several structural classes to understand the physiological importance of the different membrane protein insertion pathways. For example, what determines if a (Sec-dependent or Sec-independent) membrane protein requires YidC and how does YidC recognize its substrates? The reconstitution of both, the Sec-dependent and the YidC-dependent membrane protein insertion pathway will be a valuable tool in these studies.
Chapter 1

Scope of this thesis

The goal of the work described in this thesis was to analyze the catalytic role of the Sec translocase and the YidC protein in membrane protein insertion. To precisely define the minimal requirements for the insertion of different classes of membrane proteins, an in vitro reconstitution approach was chosen. For this purpose, proteoliposomes reconstituted with purified SecYEG and/or YidC were analyzed for their ability to catalyze the insertion of membrane proteins. In addition, a YidC depletion strain was used to determine the essential physiological function of YidC.

The first part of this thesis deals with the reconstitution of Sec-dependent membrane protein insertion using the model membrane protein FtsQ as a substrate. Chapter 2 describes the purification of YidC and its functional co-reconstitution with the SecYEG complex. A sequential interaction of the TMS of nascent FtsQ with first SecY and then YidC is demonstrated by means of a crosslinking approach.

In Chapter 3, it is shown that proteoliposomes containing only SecYEG catalyze the Δψ-dependent membrane insertion of full-length FtsQ. Although YidC interacts with nascent FtsQ, it is not essential for the membrane insertion reaction as such. Furthermore, it is demonstrated that membrane insertion of FtsQ requires the SRP targeting pathway and the SecA motor protein.

In Chapter 4, the physiological consequences of a depletion of YidC are analyzed. This study aimed to identify native membrane protein substrates of YidC to explain the essential role of YidC in E. coli. It is shown that YidC depletion leads to a strongly reduced ability to generate a PMF, which is caused by assembly defects of the major energy-transducing membrane protein complexes, F1F0 ATP synthase and cytochrome o oxidase. In particular, the steady state levels of the ring-forming, membrane-integral subunit c of F1F0 ATP synthase (F0c) and the CyoA subunit of cytochrome o oxidase were found to be severely reduced upon YidC depletion.

Chapter 5 describes the reconstitution of F0c membrane insertion. It is shown that proteoliposomes reconstituted with only YidC mediate the efficient membrane insertion of F0c. The SRP pathway, SecYEG and the PMF are not required for this process. F0c inserted into YidC proteoliposomes has acquired the correct transmembrane topology and, just like in vivo, assembles into large oligomeric
complexes. Therefore, this study defines a novel membrane protein insertion pathway in *E. coli*, in which YidC plays an exclusive role.

In **Chapter 6**, a possible role of the SecDFyajC complex in the maintenance of the PMF is re-examined. It has been reported that SecDFyajC-depleted membranes are deficient in PMF generation with succinate (11). It is shown that this effect is due to reduced levels of the respiratory chain enzyme succinate dehydrogenase. This repression is caused by the presence of glucose in the growth medium of the depleted cells. With NADH as electron source, SecDFyajC-depleted membranes generate a normal PMF.