Functional and structural studies of a C-terminally extended YidC

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

Fakultät Naturwissenschaften
Universität Hohenheim

Institut für Mikrobiologie

vorgelegt von

Ines Seitl
aus Esslingen a.N.

2015
Parts of this thesis have been published in


Parts of this thesis have been presented at national and international conferences

Symposium Goethe University Frankfurt, 2010: Evolution of protein translocation systems
Poster presentation “The positively charged C-terminal extensions of marine YidC-homologs are crucial for interaction with E. coli ribosomes”

Gordon Conference, Galveston USA 2012: Protein transport across cell membranes
Poster presentation incl. young investigator talk “Functional analysis of the C-terminal extensions of marine YidC homologs in the SRP-dependent membrane insertion”

EMBO Conference, Dubrovnik 2013: From structure to function of translocation machines
Poster presentation “Characterization of the interaction between the C-terminal domain of a planctomycete YidC homolog and the bacterial ribosome”
# TABLE OF CONTENT

1 INTRODUCTION

1.1 The prokaryotic ribosome

1.1.1 The duty cycle of the bacterial ribosome

1.1.2 The ribosomal exit tunnel: a flexible binding platform

1.1.3 Ribosome nascent chain complexes

1.2 Protein targeting in bacteria

1.2.1 The Signal Recognition Particle and its RNA

1.2.2 Signal sequence recognition by SRP

1.2.3 Co-translational targeting: SRP-FtsY assembly

1.3 The YidC/Oxa1/Alb3 insertase family

1.3.1 Functional specialization in bacteria, mitochondria and chloroplasts

1.3.2 The bacterial membrane insertase YidC

1.3.3 YidC and the ribosome

1.4 Aims of the study

2 MATERIALS AND METHODS

2.1 Culture media and additives

2.2 Bacterial strains

2.3 Plasmids

2.4 PCR primer

2.5 Peptide synthesis

2.6 Methods in molecular biology

2.6.1 Polymerase chain reaction (PCR)

2.6.2 Agarose gel electrophoresis

2.6.3 Restriction digest of DNA

2.6.4 Ligation of DNA fragments

2.6.5 Transformation of chemically competent E. coli cells

2.6.6 Preparation of chemically (RbCl) competent E. coli cells

2.6.7 Isolation of plasmid DNA from bacterial cultures

2.6.8 DNA sequencing (Sanger et al., 1977)

2.6.9 Site-directed mutagenesis (Quikchange)

2.6.10 EtOH precipitation of plasmid DNA

2.6.11 Homologous recombination to generate the E. coli strain MC-Ffh

2.6.12 In vitro transcription
2.7 in vivo complementation assay ................................................................. 58
2.8 Fluorescence microscopy – GFP localization .................................................. 58
2.9 AMS derivatization – pulse chase & immuno-precipitation .............................. 59
2.10 Protein purification .......................................................................................... 61
  2.10.1 Purification of E. coli YidC_NHis10 ..................................................... 61
  2.10.2 Purification of YidC-Rb_NHis10 ...................................................... 62
  2.10.3 Purification of E. coli Ffh & Ffh C406S / M423C ................................. 63
  2.10.4 Purification of ribosomal proteins L24 and L29 .................................... 65
  2.10.5 Purification of the C-terminal insertase domains RbCT, OaCT, O1CT and A3CT ...... 66
2.11 Protein concentration determination according to Scopes ............................... 68
2.12 in vitro ribosome binding assay .................................................................... 68
2.13 Reconstitution of an RNC-YidC complex ..................................................... 69
2.14 Electron microscopy and image processing ................................................... 69
2.15 in vitro pull down assays ................................................................................ 70
2.16 Circular dichroism (CD) - spectroscopy ......................................................... 70
2.17 Microscale thermophoresis (MST) measurements ......................................... 71
3 RESULTS ......................................................................................................... 75
  3.1 Chimera of the membrane insertase YidC with altered C-terminal regions ...... 75
    3.1.1 Complementation of E. coli YidC ...................................................... 76
    3.1.2 Purification of the YidC proteins ...................................................... 77
    3.1.3 Ribosome binding of YidC is improved by the extended C-terminal tails ...... 81
  3.2 Structural studies of the YidC_Rb - ribosome complex ................................. 83
    3.2.1 Cryo-EM structure of the YidC_Rb – RNC complex ............................. 83
    3.2.2 The ribosomal protein L29 binds to the C-terminal tail of YidC-Rb ............ 88
    3.2.3 Structure of the YidC/Oxa1/Alb3 C-terminal domains (CTDs) ............... 91
  3.3 in vivo studies: function of the C-terminal YidC regions .................................. 98
    3.3.1 Insertion of MscL via YidC-Rb in the absence of a functional SRP system .... 98
    3.3.2 Localization of MscL-GFP in the absence of Ffh or FtsY .......................... 102
  3.4 Binding studies of SRP and the SRP signal sequence of KdpD ....................... 104
    3.4.1 Purification and labeling of Ffh from E. coli ........................................ 106
    3.4.2 Reconstitution of a functional SRP protein ......................................... 109
    3.4.3 Binding studies of SRP signal sequences to SRP using microscale thermophoresis 112
4 DISCUSSION .................................................................................................. 119
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>C-terminally extended YidC variants complement <em>E. coli</em> YidC <em>in vivo</em></td>
<td>120</td>
</tr>
<tr>
<td>4.2</td>
<td>The C-terminal tails of marine YidC homologs facilitate ribosome-binding independently of a nascent polypeptide chain</td>
<td>121</td>
</tr>
<tr>
<td>4.3</td>
<td>Visualization of the ribosome-bound insertase complex</td>
<td>123</td>
</tr>
<tr>
<td>4.4</td>
<td>The C-terminal ribosome-binding domains of marine YidC homologs are intrinsically disordered</td>
<td>127</td>
</tr>
<tr>
<td>4.5</td>
<td>Stable ribosome:YidC association can partially substitute for an SRP-mediated targeting process in <em>E. coli</em></td>
<td>130</td>
</tr>
<tr>
<td>4.6</td>
<td>Evolutionary aspects of the accessory YidC domains</td>
<td>135</td>
</tr>
<tr>
<td>4.7</td>
<td>Deciphering the SRP signal sequence of KdpD</td>
<td>139</td>
</tr>
<tr>
<td>SUMMARY</td>
<td></td>
<td>145</td>
</tr>
<tr>
<td>ZUSAMMENFASSUNG</td>
<td></td>
<td>147</td>
</tr>
<tr>
<td>5</td>
<td>REFERENCES</td>
<td>149</td>
</tr>
<tr>
<td>6</td>
<td>ABBREVIATIONS</td>
<td>171</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td></td>
<td>175</td>
</tr>
<tr>
<td>ERKLÄRUNG</td>
<td></td>
<td>177</td>
</tr>
<tr>
<td>CURRICULUM VITAE</td>
<td></td>
<td>179</td>
</tr>
</tbody>
</table>
1 INTRODUCTION

1.1 The prokaryotic ribosome

In all living cells, proteins are constantly being rearranged and degraded, and a steady production of cellular proteins is therefore required. Hence, typical mammalian cells contain more than a million ribosomes and even bacterial cells contain ~ 100 000 ribosomes. Ribosomes are nanomachines which translate the genetic DNA-code into proteins. Prokaryotic ribosomes are constituted of two unequal subunits, the small ribosomal subunit 30S and the large 50S subunit, which assemble during the initiation step of protein biosynthesis to form the active 70S ribosome. The 30S subunit contains an RNA chain (16S rRNA) of about 1540 nucleotides decorated with 20-21 different proteins, whereas the large subunit has two RNA chains (23S and 5S rRNA) of about 2904 nt and 120 nt, respectively, packed with 31 different ribosomal proteins. In bacteria, ribosomes synthesize proteins on a continuous basis at an incredible speed of > 15 peptide bonds formed per second. Despite their size difference (prokaryotic: 70S, ~ 2.5 MDa; eukaryotic: 80S, ~ 4 MDa), ribosomes in all kingdoms of life are functionally conserved, with the highest level of sequence conservation appearing in the functional domains. The central core contains the ribosomal active site within a highly conserved symmetrical region in which 98% of the nucleotides are found in > 95% of sequences from 930 different species in all three domains of life (Agmon et al., 2006).

Due to the recent scientific and technical progress in elucidating ribosome biochemistry and performing macromolecular crystallography, complete atomic-resolution structures for both subunits of archaeal and bacterial ribosomes were solved in 2000 (Ban et al., 2000; Schluenzen et al., 2000; Wimberly et al., 2000). Finally, in 2001 the first structure of the complete 70S ribosome from Thermus thermophilus was published (Yusupov et al., 2001). High resolution structures from various ribosomes and ribosomal complexes, both in eukaryotic and prokaryotic, followed (Jenner et al., 2012; Demeshkina et al., 2012; Yusupova and Yusupov, 2014). A high-resolution structure of the Escherichia coli ribosome was solved this year and reveals unprecedented views of solvation as well as post-transcriptional and post-translational modifications of the ribosome in key functional and antibiotic-binding sites (Noeske et al., 2015). Recent reviews demonstrate the current understanding of the ribosomal function, based on the correlation between functional data and their high resolution crystal structures (Bashan and Yonath, 2008). Deep understanding of the ribosomal assembly process and its protein synthesis machinery...
reveals detailed views also into antibiotic action as many of the clinically relevant antibiotics target the bacterial ribosome (Jenner et al., 2013).

1.1.1 The duty cycle of the bacterial ribosome

This chapter highlights the ribosomes' roles in protein translation and summarizes what is known today about the biochemistry of eubacterial protein synthesis. Ribosomes comprise two ribonucleoprotein subunits, 30S and 50S, that associate to form the functional 70S ribosome. The small subunit contains the decoding center and mediates the base-pairing interactions between the codon triplets of the mRNAs (messenger RNAs) and the appropriate tRNAs (transfer RNAs) that determine the amino acid sequences of the originating proteins (Nirenberg, 1964). The large subunit includes the peptidyl transferase center (PTC) that catalyzes peptide bond formation (Monro, 1967) by transferring the C-terminal carboxylate group of a nascent polypeptide still bound to a tRNA to the α-amino group of an amino acid esterified to a second tRNA molecule. The site on the ribosome occupied by peptidyl tRNAs just before peptide bond transfer occurs is called the P site. The site occupied by aminoacyl tRNAs is called the A site (Monro et al., 1969). The third major site is called E site through which deacylated tRNA molecules pass on their way out of the ribosome to get recycled and reloaded with amino acids (Rheinberger and Nierhaus, 1983). Hence, the protein synthesis cycle of ribosomes can be separated into three parts, the initiation, the elongation and the termination step (Fig. 1.1), whereupon several additional protein factors are required.

The initiation of protein synthesis begins with recruitment of ribosomal subunits from the cellular pool by the three initiation factors IF-1, IF-2 and IF-3 (Gualerzi and Pon, 1990), followed by the binding of an mRNA. The initiation phase ends with an aminoacylated initiator tRNA molecule (fMet tRNA) residing in the P-site of the ribosome, while the A-site is empty. The subsequent elongation step starts with an incoming tRNA complexed with elongation factor Tu (EF-Tu) which is delivered to the A-site driven by GTP hydrolysis: Peptide bond formation then occurs between the aminoacyl tRNA in the A-site and the adjacent peptidyl tRNA bound to the P-site. To reinitialize the ribosome during the so called translocation process that is facilitated by the elongation factor G (EF-G), the deacetylated tRNA in the P site first leaves the ribosome via the E site. This makes the peptidyl tRNA in the A site now move to the empty P site. Thus, the ribosome moves along its mRNA in the 3’direction by one codon so that the next aminoacyl tRNA can be delivered to the empty A site.
Fig. 1.1 The ribosomal protein synthesis cycle. Initiation occurs by two ribosomal subunits being assembled at the start codon of an mRNA molecule together with a loaded tRNA molecule and three initiation factors (IF), and the first peptide bond is formed. The sequence of events that results in the addition of one amino acid residue to a growing peptide chain is called the elongation cycle, and is repeated several times depending on the protein size. When the translating ribosome encounters a stop codon, the nascent polypeptide is released from the ribosome and the two subunits return to the cellular pool, called termination and ribosome recycling. In vivo, the ribosome functions need the assistance of several protein factors. The most important are elongation factor G (EF-G) and elongation factor Tu (EF-Tu). The small ribosomal subunit is shown in gold, and the large subunit in blue. tRNAs and protein factors are consistently color coded, and the shapes assigned to them resemble their 3-dimensional structures. Abbreviations: EF, elongation factor; IF, initiation factor; RF, release factor; RRF, ribosomal recycling factor (Moore, 2012).

At the end of the elongation cycle when the stop codon (UAA, UAG or UGA) in the mRNA has been positioned into the A site, one of two release factors (RF1 or RF2) binds to the A site and promotes the deacylation of the last peptidyl tRNA. RF1 recognizes UAG and RF2 is specific for UGA, but both factors recognize UAA (Youngman et al., 2008). RF3 accelerates the dissociation of RF1 or RF2 from the ribosome. In order to prepare the ribosome for a new synthesis cycle, the ribosome recycling factor (RRF) binds together with EF-G to the ribosome and, driven by GTP hydrolysis, causes the release of the mRNA, removal of the deacylated tRNA from the P-site (RamaKrishnan, 2002) and the dissociation of the two ribosomal subunits (Karimi et al., 1999). However, the exact mechanism of factor-assisted elongation remains unclear since the observation was made
that ribosomes can elongate peptide chains in the absence of both EF-G and EF-TU (Pestka, 1968). Thus we still do not understand how the system uses the free energy released by the GTP hydrolysis that accompanies the process. Single-molecule experiments may provide the data needed to understand not only the details of the mechanism of elongation, but also its design philosophy (Moore, 2012).

1.1.2 The ribosomal exit tunnel: a flexible binding platform

Since its identification in the early 1980s by EM images of Unwin and co-workers (Milligan and Unwin, 1982), the polypeptide exit tunnel has been a subject of extensive study but also speculation concerning its function. The tunnel is located adjacent to the PTC in the large ribosomal subunit (Ban et al., 2000; Harms et al., 2001) and is surrounded mainly by rRNA and only a few ribosomal proteins. This tunnel has a length of 80 - 100 Å with a varying diameter of 10 – 20 Å (Fig. 1.2 B) and possesses the flexibility and dynamics required for interacting with the nascent protein progressing through the tunnel (Fedyukina and Cavagnero, 2011). The tunnel resembles a tube that is able to accommodate a peptide stretch of ~ 30 amino acid residues in extended or up to 40 amino acid residues in an α-helical conformation (Voss et al., 2006; Picking et al., 1992; Malkin and Rich, 1967). In bacteria, the tunnel wall is formed predominantly by 23S rRNA and looped-out segments of the ribosomal proteins L4, L22 and L23 (Fig. 1.2 C). At its distal end, the rim of the exit point at the large ribosomal subunit is composed of RNA, a ring of the four ubiquitously conserved ribosomal proteins L22, L23, L24 and L29 and additional kingdom-specific proteins (Kramer et al., 2009).

Fig. 1.2 The ribosomal environment for nascent polypeptides. (A) Crystal structure of the E. coli ribosome at 3.5 Å resolution (PDB IDs: 2AVY and 2AW4). The ribosomal RNA of the 50S subunit (23S and 5S rRNA) are shown in turquoise and the 16S rRNA of the small subunit in beige. Ribosomal proteins are shown as ribbons (50S subunit proteins in purple, 30S subunit proteins in green). (B) Structure of the ribosomal exit tunnel (Voss et al., 2006). There is a constriction in the tunnel about 30 Å away from the PTC where loops of L4 and L22 come in close proximity. At the distal end, the ribosomal exit tunnel widens up. (C) Schematic representation of a vertical section of the prokaryotic ribosome, highlighting the ribosomal proteins facing or located near the exit tunnel. The ribosome-associated trigger factor (TF) is shown in pink (adapted from Fedyukina and Cavagnero, 2011).
INTRODUCTION

During protein biosynthesis, the nascent peptide chains on translating ribosomes are scanned by a number of additional proteins, called ribosome-associated protein biogenesis factors (RPBs), such as the chaperone trigger factor (TF), the signal recognition particle (SRP) and the processing enzymes peptide deformylase (PDF) and methionine aminopeptidase (MAP). These RPBs use six ribosomal proteins surrounding the tunnel exit as a confined binding platform of the bacterial 50S ribosomal subunit (Fig. 1.3). The bacterial enzyme peptide deformylase (PDF) is an essential protein that cotranslationally removes the formyl group of nascent chains, a step required for subsequent action of the methionine aminopeptidase (MAP) to excise the N-terminal methionine residue (Fry and Lamborg, 1967; Pine, 1969; Adams, 1968). E. coli PDF binds to a groove between ribosomal proteins L22 and L32, located next to the exit tunnel, thus positioning the active site for interaction with the emerging nascent polypeptides (Bingel-Erlenmeyer et al., 2008), adjacent to the trigger factor binding site (Kramer et al. 2002). The ribosomal protein L17 was identified as the major interaction partner of MAP (Sandikci et al., 2013).

The ribosome-associated signal recognition particle (SRP) is involved in co-translational targeting of nascent inner membrane proteins to the protein-conducting channel (translocon) in the bacterial plasma membrane (Luirink et al., 2005; Bibi, 2011; Grudnik et al., 2009), whereas nascent cytoplasmic, periplasmic and outer membrane proteins interact with the chaperon trigger factor (TF), which transiently associates with the large ribosomal subunit (Oh et al., 2011). Trigger factor forms an arch above the tunnel (Fig. 1.2 C) and thus protects the nascent chain from misfolding and aggregation. It also assists early folding events (Agashe et al., 2004; Hoffmann et al., 2012; Deuerling et al., 1999; Teter et al., 1999; Kramer et al., 2002). Later folding steps often are supported by the
DnaK and GroEL chaperon systems that act independently of the ribosome. SRP and TF both bind to the ribosomal protein L23 (Ferbitz et al., 2004; Kramer et al., 2002; Merz et al., 2008; Gu et al., 2003; Halic et al., 2004) at the peptide exit site (Fig. 1.3). This concurrent binding is accompanied by rearrangements of the complex (Buskiewicz et al., 2004; Ullers et al., 2003). Qualitative and quantitative analyses indicate concurrent binding also of TF and PDF or SRP and PDF to ribosomes (Bornemann et al., 2014; Sandikci et al., 2013), indicating non-competitive binding. Based on these observations a working model was generated for the co-translational processes at the tunnel exit of ribosomes (Fig. 1.4). Fast ribosome binding kinetics facilitates early ribosomal recruitment of SRP at nascent chain lengths of about 30 amino acid residues (Holtkamp et al., 2012; Bornemann et al., 2008). The complex is further stabilized by the emergence of a signal anchor sequence, which finally results in SRP receptor recruitment and targeting the translating ribosomal complex to the translocon. Since SRP does not exclude PDF binding, it is assumed that deformylation of the initiator f-met methionine residue and co-translational targeting can occur simultaneously which is consistent with the almost complete absence of native formylated proteins in *E. coli* (Solbiati et al., 1999; Milligan and Koshland, 1990). However, around 5% of the complete proteome stay formylated. Interestingly, nearly all the formylated proteins are proteins of the inner membrane or proteins involved in protein synthesis and early co-translational events (Bienvenut et al., 2015). It was suggested that the charge of the N-terminus could have influence on the transmembrane topology of some specific proteins (von Heijne, 1989). In this context, the retention of the N-formyl group changes the protein N-terminus charge and might favor transmembrane inclusion. In mitochondria for example, it was shown that COX1 requires an N-terminal formyl methionine for the assembly of cytochrome c oxidase (Hinttala et al., 2015). However, the vast majority of SRP substrates are deformylated, but retain the N-terminal methionine and therefore do not require MAP-mediated processing (Huber et al., 2005). For all other nascent chains, the fast ribosome interaction kinetics of PDF and MAP ensure that N-terminal methionine excision (NME) is completed before trigger factor is recruited to translating ribosomes with much slower kinetics (Rutkowska et al., 2008) and only after 100 residues on average have been synthesized on the growing peptide (Fig. 1.4; Oh et al., 2011).
Recent studies propose that during the interplay of TF and SRP, concurrent binding of both factors to RNCs presenting the respective specific nascent chain for a SRP-dependent or TF-dependent protein leads to weakening of the binding of the respective other ligand. This partial competitive binding behavior of TF and SRP increases therefore the specificity of the small amount of SRP present in the cell in promoting co-translational membrane targeting of those RNCs that synthesize membrane proteins (Bornemann et al., 2014).

In addition to the early processes of RPBs binding, the proteins building the tunnel rim constitute major interaction sites for various factors involved in nascent chain folding, targeting and insertion. Interestingly one of them, the ribosomal protein L23, is involved in almost all interactions of the ribosome with ribosome-associated factors investigated so far, including the translocon SecYEG (Mitra et al., 2005; Frauenfeld et al., 2011) as well as the soluble ATPase SecA (Singh et al., 2014). SecA mediates translocation of periplasmic and outer membrane proteins and channels newly synthesized polypeptides into the post-translational Sec-translocation pathway (Huber et al., 2011; Driessen and Nouwen, 2008). A specific interaction with L23 was also described for the insertases Oxa1 in mitochondria (Jia et al., 2003) and the bacterial YidC (Kohler et al., 2009; Seitl et al., 2014) as well as for the targeting factor SRP (Schaffitzel et al., 2006) and the chaperone trigger factor (Ferbitz et al., 2004; Kramer et al., 2002). It is known that ribosomes can transmit information about the presence of a nascent chain from their interior to the surface, to control the interaction for example with SRP. In E. coli, SRP shows a 100-fold increased affinity for translating ribosomes (Bornemann et al., 2008), whereupon the signal transfer from the inside of the tunnel to the ribosomal surface occurs via a loop structure in L23 that reaches into the exit tunnel (Kramer et al., 2009). Thus, L23 seems to
constitute a general docking platform for various factors that transiently associate with ribosomes to act on nascent chains and could be used for the coordination of these factors in time and space. This even holds for the homologous ribosomal proteins and the ribosome-associated chaperone counterparts in yeast and mammals (Preissler and Deuerling, 2012).

1.1.3 Ribosome nascent chain complexes

Protein synthesis proceeds at variable rates and with different velocities depending on environments and organisms (Zhang and Ignatova, 2010). Since translation rates are faster in prokaryotes (15 – 20 aa / sec) than in eukaryotes (3 – 4 aa / sec) the recombinant expression of eukaryotic proteins in *E. coli* often does not work proper, indicated by misfolding of the recombinantly expressed heterologous protein. An average timescale for the production of a small or medium sized protein in prokaryotes of ~ 10 s is similar to chaperone binding/release times and longer than folding/unfolding timescales of small proteins (Fedyukina and Cavagnero, 2011) suggesting that nascent chains encoding only small proteins may have sufficient time to adopt the proper conformation during synthesis. In contrast, large proteins take much longer and need multiple folding-unfolding rounds to fold properly. In addition, the co-translational folding in *E. coli* seems to be highly protein- and codon-dependent. Rare codon clusters (Clarke and Clark, 2008) seem to be an important feature for an orchestrated translational pausing in order to facilitate co-translational domain folding before synthesis of the following domain is initiated (Zhang *et al.*, 2009). For instance, mutant ribosomes displaying slower translation than wild-type *E. coli* ribosomes enhance the production of active multi-domain proteins of eukaryotic origin in bacteria (Siller *et al.*, 2010).

Characterizing the nature and properties of RNCs exhibits a significant challenge structural and cellular biology experiments. One strategy for defining the structural and dynamic properties of a nascent chain during co-translational folding and targeting up to its interaction with translocons or insertases is the generation of translation arrested RNCs, resulting in the nascent peptide chain being retained on the ribosome. Several nascent peptides stall ribosomes using specific leader peptides during their own translation in both prokaryotes and eukaryotes and can regulate downstream gene expression. Two well characterized leader peptides from *E. coli* cause ribosome stalling and increase the expression of a gene further downstream on the same mRNA. SecM, for example acts as a secretion monitor peptide that regulates *secA* in response to changes in protein translocation activity (Oliver *et al.*, 1998). The 17 residues long SecM arrest peptide corresponds to the intrinsic class of stalling sequences which stalls without any
inducing molecule until the arrest is released as the nascent chain interacts with specific cellular machineries, in this particular case with the Sec translocon. SecM-stalled ribosomes can be rescued by mechanical force generated by the translocon (Goldman et al., 2015). On the other hand, the 24 residues long TnaC leader peptide corresponds to the inducible class of arresting peptides that only stalls in the presence of a specific effector, such as an antibiotic or a metabolite (Ito et al., 2010). TnaC is the regulatory peptide of the tryptophanase tnaCAB operon in E. coli (Gong and Yanofsky, 2002) using a feedback loop requiring the small molecule L-tryptophan (L-Trp). The respective DNA sequence is localized on the E. coli chromosome upstream of the L-Trp catabolizing tryptophanase (TnaA) gene and the tryptophan-specific permease (TnaB) gene. In the presence of inducing levels of free L-Trp, peptide release by RF2 is inhibited (Gong and Yanofsky, 2001) and the ribosome stalls on the TnaC mRNA carrying the Tna leader peptide. As a consequence, the transcription termination factor Rho is blocked from binding to the mRNA region between tnaC and tnaA and, thus, allows the transcription and subsequent translation of TnaB and TnaA (Gong and Yanofsky, 2002) (Fig. 1.5 A). A previous cryo-EM structure of a TnaC-stalled ribosome complex demonstrated that TnaC stalls the ribosome with a peptidyl-tRNA remaining in the ribosomal P-site. The nascent peptide adopts a defined conformation within the ribosomal exit tunnel, and close contacts between TnaC and components of the tunnel wall were identified (Seidelt et al., 2009). A highly resolved cryo-EM structure (Fig. 1.5 B) revealed the formation of two composite binding pockets by the nascent chain and the tunnel wall (Fig. 1.5 C), turning the translating ribosome into an efficient sensor for L-Trp (Bischoff et al., 2014-2).

The observation that only the sequence of the stalling peptide (e.g. for SecM: 17 aa of the 170 residues full length protein) is necessary for proper ribosome stalling, unrelated of downstream sequences (Nakatogawa and Ito, 2002), offers new possibilities to analyze the interaction of substrate-specific RNCs with factors involved in folding, targeting and translocation or insertion of the nascent protein. The SecM and TnaC approaches are preferentially used when RNCs are generated in vivo and purified by affinity chromatography (Bischoff et al., 2014-1).
Fig. 1.5 Cryo-EM structure of a TnaC-stalled ribosome nascent chain complex. (A) Model for TnaCAB operon regulation in E. coli. In the absence of L-Trp the TnaC translation reaches the stop codon UGA, RF2 promotes termination and the ribosome dissociates. Therefore, Rho can bind to its transcription termination site, prior to synthesis of the tnaAB mRNA. If free L-Trp is present, it interacts with two binding pockets in the nascent chain and in the tunnel wall and thus prevents peptide release by RF2. Rho binding is blocked and transcription of tnaAB occurs. (B) Cross-section through the cryo-EM density of the TnaC-RNC. Displaying the 30S (yellow) and 50S (grey) subunit of the E. coli ribosome, carrying a tRNA in the P-site (dark green) and the nascent chain (light green) with free tryptophan molecules (orange) in the ribosomal exit tunnel. (C) Directly within the ribosomal tunnel, 15-20 Å from the PTC, a density for two free L-Trp molecules (W1 and W2) was observed. (Adapted from Bischoff et al., 2014-2)

Taken together, the understanding of the stalling mechanism of RNCs in E. coli and new methods for their synthesis and purification provided the visualization of RNCs in complex with SRP/FtsY (von Loeffelholz et al., 2015), SecA (Singh et al., 2014), SecYEG (Frauenfeld et al., 2011; Bischoff et al., 2014-1) and YidC (Kohler et al., 2009; Seilt et al., 2014; Wickles et al., 2014) by cryo-EM reconstruction and brought new insights into the mechanism of co-translational protein biogenesis.
1.2 Protein targeting in bacteria

The proper biogenesis and homeostasis of proteins are essential to all living cells and require the correct folding, localization, maturation, and quality control of all newly synthesized proteins (Hartl et al., 2011; Powers et al., 2009). As described in the previous chapter, the mechanisms that ensure proper protein homeostasis already begin at the ribosome, where a variety of proteins meet and compete for access to nascent polypeptides. A new era in cell biology was spawned in the 1970s by Günter Blobel’s “signal hypothesis” proposing that newly synthesized proteins carry intrinsic signals, the so-called signal sequences, that encode information about their cellular localization (Blobel and Sabatini, 1971; Blobel, 1980). In the subsequent years, the signal sequences and respective targeting factors were identified and characterized for various organelles particularly those mediating the delivery of proteins to the membranes of the endoplasmic reticulum (ER), nucleus, mitochondria and chloroplasts (Dalbey and von Heijne, 2002). Bacterial cells also contain distinct compartments and specified sites to which newly synthesized proteins must be correctly transported and located, including the inner and outer membranes as well as the periplasmic and the extracellular space. The major protein trafficking route involves the transport of newly synthesized membrane and secretory proteins from the cytosol to the plasma membrane and multiple pathways have evolved for the delivery and sorting of these proteins (Fekkes and Driessen, 1999; Cross et al., 2009).

Protein targeting in bacteria can be divided into two major routes (Fig. 1.6): (a) post-translational delivery after the complete synthesis and release of the protein from the ribosome prior to targeting; (b) the co-translational pathway, in which the targeting and translocation is coupled to the ongoing synthesis by RNCs. In contrast to higher eukaryotes, where the co-translational delivery is the major pathway for secretory proteins, most secretory proteins in bacteria are post-translationally targeted to the plasma membrane. This is, presumably due to most efficiently utilizing the limited number of SecYEG translocation channels, a major translocon in the bacterial inner membrane, since the rate of protein synthesis might be higher than the delivery step (Hegde and Bernstein, 2006). During post-translational delivery, the primary challenge is to keep polypeptides in a translocation-competent state that can be transported into or across the target membrane. Hence, a major function of post-translational delivery factors is to bind the hydrophobic regions of the substrate polypeptide to inhibit misfolding, aggregation and inappropriate interaction with other cytosolic proteins. This task is fulfilled by several molecular chaperones that mostly depend on ATP hydrolysis. For both, the post- and the
co-translational pathway, a cycling mode of action is characteristic for most cellular factors. They can be distinguished by either using regulatory ATP binding components, such as SecA (Fig. 1.6 a) or alternatively by using the highly conserved GTPases of the SRP cycle (Fig. 1.6 b).

The homotetrameric SecB chaperone binds with high affinity to nascent secretory precursor proteins with a typical N-terminal cleavable signal sequence (Baars et al., 2006), whereas the export sequence is not bound directly but its presence might extend the time frame for interaction with SecB (Hardy and Randall, 1991). SecB captures and keeps the secretory proteins in a translocation-competent state and mediates the post-translational targeting to the translocation ATPase SecA. SecA tightly associates with the membrane-embedded SecYEG translocon and pushes unfolded substrate protein across it using ATP-driven highly dynamic conformational changes (Fig. 1.6 a) (Cabelli et al., 1988; Gold et al., 2013; Denks et al., 2014). Recent structural analysis of SecA bound to the SecYEG complex reveals the two helix finger domain (Zimmer et al., 2008) driving the substrate through the SecYEG translocation channel by acting as a translocation piston to push the substrate into the periplasmic space (Erlandson et al., 2008). Several reports suggest that a distinct cytosolic pool of SecA can also associate with ribosomes and with the nascent signal sequence of secretory proteins, raising the intriguing possibility that post-translational targeting machineries could also exert some of their actions co-translationally (Eisner et al., 2003; Karamyshev and Johnson, 2005; Huber et al., 2011). Other general chaperones, such as the trigger factor (TF), GroEL or DnaK, may also be involved in maintaining the nascent polypeptides in a translocation-competent unfolded state (Hoffmann et al., 2012).

In an alternative targeting route to the post-translation delivery, a subset of bacterial secretory proteins are translocated in a tightly folded state, in particular cofactor-containing membrane-located redox enzymes (Sargent F., 2007). Substrates for this pathway exhibits a characteristic twin arginine motif in their signal sequences and are translocated via the Tat (twin Arg translocation) complex, a translocon composed of the TatA, TatB and TatC subunits (Palmer and Berks, 2012). The DnaK and DnaJ proteins are cytosolic molecular chaperones of the Hsp70 family that carry important roles in general protein folding and assembly (Genevaux et al., 2007). Despite its role in the export of some SecB-independent substrates (Wild et al., 1992; Qi et al., 2002), DnaK might also play a role in the Tat-pathway. DnaK probably interacts with Tat substrates and promotes their export by sheltering the signal sequence from premature engagement with the TatABC translocon (Graubner et al., 2007; Perez-Rodriguez et al., 2007).
Fig. 1.6 **Protein targeting and export pathways in prokaryotes.** Depending on the specific protein and its final destination, the delivery of proteins from the cytosol to the membrane occurs through distinct routes. **(a)** Post-translational delivery of secretory proteins is mainly mediated by SecB and driven by ATP hydrolysis cycles of SecA into the translocon, which comprises the conserved SecYEG channel together with YidC. YidC is an essential insertase required for the membrane insertion and assembly of a subset of integral inner membrane proteins. The TatABC translocon in cooperation with the chaperones DnaJ and DnaK mediates the export of folded proteins containing a twin arginine motif in their signal sequence. **(b)** The co-translational targeting to SecYEG occurs via SRP and its receptor FtsY in a GTP-dependent manner. For YidC, acting here as an independent insertase without SecYEG, the targeting might occur either following SRP-mediated delivery or through direct binding of the ribosome to YidC. (Cross et al., 2009)

An alternative strategy to maintain the translocation competence of proteins destined for export, instead of using chaperones, is to intimately couple protein synthesis to membrane translocation by delivering the nascent polypeptide on the ribosome to the membrane during translation. This co-translational delivery strategy (Fig. 1.6 b) also prevents the potential and fatal exposure of multiple hydrophobic transmembrane segments of integral membrane proteins to the aqueous cytosolic environment and ensures their coordinated insertion and correct assembly at the membrane. The co-translational targeting starts with the partial competitive exclusion of trigger factor from the ribosomal exit site by the highly conserved ribonucleoprotein complex, the Signal Recognition Particle (SRP) which specifically guides the RNCs translating integral membrane proteins (Ulbrandt et al., 1997; Valent et al., 1997) and some secretory proteins (Emanuelsson and von Heijne, 2001) to the membrane, mainly to the Sec translocon (Keenan et al., 2001). The recruitment of the SRP-RNC to the inner membrane is mediated by the interaction of SRP with its membrane-associated receptor FtsY and driven by GTP hydrolysis (Yang and Zhang, 2011).
The heterotrimeric SecYEG translocon is the primary protein export channel in *E. coli* and it is used for both, post- and co-translationally delivered proteins. The crystal structure from the homologous *Methanocaldococcus jannaschii* SecYEβ provided the first high-resolution insight into the organization and structure of the translocation channel (Fig. 1.7; van den Berg *et al*., 2004).

The SecY protein forms the channel pore by its 10 α-helical transmembrane domains (TMDs) through which proteins might traverse the membrane. TMDs 1-5 and TMDs 6-10 of SecY are pseudo-symmetrically aligned resembling a bivalent shell that forms a lateral gate for the release of hydrophobic domains into the membrane, hinged by SecE (Fig. 1.7). The channel is intrinsically sealed by a plug domain formed by a subregion within the second TMD of SecY (Lycklama *et al*., 2012). While SecYEG is the main site for protein insertion and translocation, the essential insertase YidC was found to act in conjunction with the translocon in integral membrane protein insertion in bacteria (Zhu *et al*., 2013). YidC appears to exist in two pools (Fig. 1.6 b): one that is tightly associated with SecYEG and assists in the integration of polytopic membrane proteins (Scotti *et al*., 2000; Nouwen and Driessen, 2002; Plessis *et al*., 2006; Kol *et al*., 2009), and another species that acts independently of SecYEG to mediate the integration of, mainly small, membrane proteins of *E. coli* (Samuelson *et al*., 2000; van der Laan *et al*., 2003 & 2004; Serek *et al*., 2004). The targeting to YidC is thought to mainly occur via the SRP pathway (Facey *et al*., 2007), although SRP-independent mechanisms have also been described (Gerken *et al*., 2008; Welte *et al*., 2012). The direct interaction of YidC with translating ribosomes was elucidated in several studies during the recent years (Kohler *et al*., 2009; Funes *et al*., 2011; Kedrov *et al*., 2013; Seidl *et al*., 2014; Wickles *et al*., 2014; Geng *et al*., 2015). The functionality and role of this interaction in the targeting and insertion of YidC-only substrates is one main focus of this work. Despite the diversity of trafficking pathways, protein targeting can be divided into three key steps that are common to all pathways: 1.
substrate recognition in the cytosol, 2. targeting and delivery to the respective membrane and 3. passage through or insertion into the membrane lipid bilayer. The SRP pathway comprises these general principles and serves as a paradigm for understanding the molecular basis of protein localization and sorting in all cells.

1.2.1 The Signal Recognition Particle and its RNA

SRP is an ancient and essential ribonucleoprotein particle conserved across all kingdoms of life (Walter and Johnson, 1994; Keenan et al., 2001; Akopian et al., 2013-1). The evolutionary conservation of this pathway was demonstrated experimentally in vitro by the substitution of the eukaryotic homolog with the highly simplified bacterial SRP system. The latter could efficiently mediate the targeting of mammalian substrates to ER microsomes (Bernstein et al., 1993; Powers and Walter, 1997). This allows the bacterial SRP to serve as a universal model system for understanding the fundamental molecular mechanisms of this targeting machine in prokaryotic and eukaryotic cells.

The E. coli SRP is comprised of the protein component Ffh (SRP54 in eukaryotes) and the 4.5S RNA (7S RNA in eukaryotes) (Poritz et al., 1988). The SRP RNA binds with picomolar affinity to the methionine-rich M-domain of SRP (Batey et al., 2000; Jagath et al., 2001; Siu et al., 2007). The M-domain, together with the SRP RNA facilitates the signal sequence recognition and binding (Keenan et al., 1998; Batey et al., 2000; Janda et al., 2010; Hainzl et al., 2011). Besides the C-terminal M-domain, the second structurally and functionally distinct domain of SRP is the NG-domain (Fig. 1.8) composed of an N-terminal helical N-subunit and a centrally located GTPase subunit, the so called G-domain (Freymann et al., 1997; Montoya et al., 1997). The N-domain mediates interactions with the ribosome via binding to the ribosomal protein L23 at the ribosomal tunnel exit site (Pool et al., 2002; Gu et al., 2003; Halic et al., 2006; Schaffitzel et al., 2006), and the G-domain mediates GTP binding and hydrolysis. Collectively, the NG-domain of Ffh interacts directly with the highly conserved NG-domain of the SRP receptor protein FtsY, forming a heterodimeric GTPase core of the SRP targeting complex (Egea et al., 2004; Focia et al., 2004) (Fig. 1.8). In addition to the NG-domain, FtsY contains an N-terminal acidic A-domain, which mediates the peripheral association of the receptor with the phospholipid membrane and with the SecYEG translocon (Parlitz et al., 2007; Weiche et al., 2008).
The discovery of the bacterial SRP RNA (Poritz et al., 1988 & 1990), years later after the discovery of the mammalian SRP RNA by Peter Walter, challenged the view of the RNA component to be a passive scaffold necessary for the correct assembly of the six mammalian SRP protein subunits (Walter and Blobel, 1982). Recent biochemical and structural studies have demonstrated that indeed, the SRP RNA actively mediates domain rearrangement of SRP upon cargo recognition (Buskiewicz et al., 2005), thus allowing also proper interaction with its receptor FtsY (Peluso et al., 2000; Jagath et al., 2001) as well as activating and stimulating the GTPase function of the SRP-receptor complex (Siu et al., 2007). Bacterial SRP shows high homologies to the domain IV of the eukaryotic SRP RNA, a region of ~50 nucleotides, whose highly conserved secondary structure consists of two internal loops that include noncanonical base pairings and unpaired nucleotides (Fig. 1.9 A) and a second hairpin structure that is capped by a highly conserved GGAA tetraloop. Domain IV in bacterial and human SRP stabilizes the ribonucleoparticle and its interaction with the signal peptide (Gowda and Zwieb, 1997; Zheng and Gierasch, 1997). The 4.5S RNA of E. coli binds with picomolar affinity to the M-domain of Ffh (Fig. 1.9 C) via its two internal loops (Fig. 1.9 B) adjacent to the GGAA tetraloop (Batey et al., 2000). The orientation of the M-domain/RNA complex relative to the Ffh NG domain exhibits a high degree of flexibility. At least 4 different structures were found by crystallographic analysis and structural mapping studies, each exhibiting a distinct interdomain arrangement (Keenan et al., 1998; Rosendal et al., 2003; Buskiewicz et al., 2005 1&2; Mainprize et al., 2006; Hainzl et al., 2007). These observations suggest that free SRP is a highly dynamic particle that can undergo substantial structural rearrangements, likely due to the 30 amino acid residues long flexible linker within the Ffh protein, connecting its M- and the NG-domains.
Exempt from domain IV, the sizes, sequences, and secondary structures of SRP RNAs vary widely, even among bacterial species (Samuelsson and Zwieb, 1999), and it is unclear whether these regions of the RNA are essential for SRP function. A recently published crystal structure of the SRP-FtsY/receptor complex provided a possible function for the distal end of the SRP RNA, which trapped the SRP-FtsY GTPase complex ~ 100 Å away from the tetraloop domain (Ataide et al., 2011). Together with biochemical studies and single molecule fluorescence measurements, these results suggest a model in which the SRP-FtsY NG domains, after initial assembly at the tetraloop end of the RNA, relocalize to its distal end where GTP hydrolysis is activated (Shen et al., 2012). The movement of the NG domains to the distal RNA end is negatively regulated by the translating ribosome and restored by the SecYEG complex, providing the first experimental support for a concerted mechanism of cargo handover from SRP to the SecYEG complex by vacating the ribosomal protein L23. This hypothesis provides an attractive mechanism of coupling unloading of the SRP substrate protein cargo to GTP hydrolysis (Shen et al., 2012; Akopian et al., 2013-b). In addition, a recent study showed that the 6S RNA of the SRP from the Gram-positive bacterium Bacillus subtilis facilitates elongation slowdown similarly to the eukaryotic translational arrest by the Alu domain of
SRP (Beckert et al., 2015). Thus the SRP RNA is an active molecular scaffold, mediating protein rearrangements and exchange of distinct factors via multiple protein interaction sites, therefore allowing effective coordination of a complex cellular pathway.

1.2.2 Signal sequence recognition by SRP

In general, N-terminal signal sequences that are recognized by SRP are characterized by a positively charged n-region, an h-region containing the hydrophobic core (8-20 aa) and a polar c-region, preferentially adopting an α-helical structure (von Heijne, 1985; Gierasch, 1989). The unifying characteristic of all signal sequences, and apparently the only essential feature for binding SRP, is an uninterrupted stretch of at least seven hydrophobic residues surpassing a threshold level of hydrophobicity (Lee and Bernstein, 2001). Thus the first transmembrane domain often serves as a signal sequence for SRP. One fundamental question is, how SRP recognizes these signals that are highly divergent in sequence, length and amino acid composition without any known consensus motif (Zheng and Gierasch, 1996; Hegde and Bernstein, 2006). To test the specificity during the cargo recognition step by SRP, several research groups have compared the binding affinities of SRP for RNCs bearing SRP-dependent and SRP-independent substrates (Flanagan et al., 2003; Bornemann et al., 2008; Zhang et al., 2010; Saraogi et al., 2011; Holtkamp et al., 2012). Depending mainly on the hydrophobicity of the signal peptide, SRP tightly binds RNCs displaying an SRP signal sequence with equilibrium dissociation constants ($K_d$) in the nanomolar range (1 – 10 nM). However, yet the weakest cargo or even empty ribosomes bind SRP with significant $K_d$ values in the 50 – 100 nM range (Flanagan et al., 2003; Bornemann et al., 2008; Zhang et al., 2010; Holtkamp et al., 2012). This suggests that SRP can bind rapidly to RNCs bearing both SRP-dependent and -independent substrates, but its dissociation occurs much more slowly from the RNCs displaying a proper cargo. Since the binding of isolated signal peptides to SRP is much weaker, with a dissociation constant in the micromolar range (~ 1.5 µM; Bradshaw et al., 2009), the interaction with the ribosome obviously contributes significantly to the RNC-SRP binding energy and provides an important driving force for SRP-recruitment to nascent polypeptide chains. Regarding the cellular concentrations of SRP (~ 400 nM in bacteria) and ribosomes (40 - 50 µM), it is unlikely that the different affinities in the initial cargo-binding step are sufficient to discriminate against incorrect cargos and to ensure the accurate substrate selection by SRP. Thus, additional fidelity and quality checkpoints are provided by the SRP- and FtsY-GTPases, described in the following chapter. Cross-linking studies (Pool et al., 2002; Gu et al., 2003) and cryo-EM reconstruction data revealed the ribosomal protein L23 as a major contact site for the Ffh N-domain on the ribosome and to a lesser extend L29 in the vicinity of the ribosomal exit tunnel, in both
eukaryotic and bacterial systems (Halic et al., 2004 & 2006; Schaffitzel et al., 2006). The crystal structures of SRP-signal peptide fusions (Janda et al., 2010; Hainzl et al., 2011) clearly showed that the signal sequence binds into a deep, hydrophobic groove in the M-domain. However, different docking modes of the signal peptide arose, highlighting the flexibility of the signal sequence-M domain interaction. Compared to the structures of the respective free SRP (Fig. 1.10 A), without bound signal sequence, it became obvious that upon binding of a signal sequence, huge structural changes in the SRP core are induced (Fig. 1.10 B).

**Fig. 1.10** Domain rearrangement of the SRP core upon signal sequence binding. (A) Free form of the M. jannaschii SRP54-SRP19-S domain RNA complex (Hainzl et al., 2011). (B) Signal-sequence (SS) bound form of (A) reveals an SRP54 domain rearrangement in the SRP core generated by a 90° rotation and a 180° flip of the NG domain. Additionally, the GM-linker forms a continuous helix compared to an extended conformation in the free SRP core. Ribbon representations of the SRP / SRP-signal sequence complex with the color code as follows: M domain, green; N domain, cyan; G domain, dark blue; SRP19, orange; signal sequence, yellow and the GM linker in magenta. (adapted from Hainzl et al., 2011)

The two distinct structures of free SRP and SRP bound to a signal sequence suggest that the binding of a signal sequence triggers the local restructuring and an α-helical formation of the, until then disordered, GM-linker (compare Fig. 1.10 A and B), causing the repositioning of the NG domain in an RNA dependent manner (Fig. 1.10 B). This brings
the GTPase domain into proximity of the universally conserved tetraloop, preparing the
NG-domain for receptor (FtsY) interaction (Hainzl et al., 2011).

1.2.3 Co-translational targeting: SRP-FtsY assembly

The SRP receptor, called FtsY in bacteria mediates the membrane targeting of cargo-
bound SRP via interaction through their GTPase modules. The acidic A domain of FtsY is
supposed to anchor the targeting complex to the membrane (Parlitz et al., 2007; Weiche
et al., 2008) in a dynamic mode (Luirink et al., 1994; Rubio et al., 2005; Mircheva et al.,
2009) also involving membrane lipids, preferentially anionic phospholipids (phosphatidyl
glycerol (PG) and cardiolipin (CL)) (de Leeuw et al., 2000; Parlitz et al., 2007; Lam et al.,
2010). In addition, to entail the targeting complex into the vicinity of the translocon, a
direct interaction of FtsY with the SecYEG translocon was found (Angelini et al., 2005 &
2006; Kuhn et al., 2011). The GTPase domains of SRP and FtsY contain the four
conserved sequence motifs of the GTPase superfamily and share the classic P-loop
GTPase-fold. Two additional structural features were identified, comprising a flexible
insertion box domain (IBD) loop with multiple catalytic residues and the N-domain, which
is a four-helix bundle that packs tightly against the G-domain to form the structural and
functional NG-domain unit (Freymann et al., 1997; Montoya et al., 1997). Compared to
classic signaling GTPases which are regulated by specific G-protein activating factors
(GAPs, GEFs), the SRP and FtsY proteins belong to a new class of GTPases which are
activated by nucleotide-dependent dimerization (GADs) (Gasper et al., 2009).

Free SRP and FtsY have their GTPase domains, even with bound GTP, in an inactive
open conformation exhibiting low nucleotide binding affinity as their nucleotide-binding
pocket is wide open, allowing free exchange of nucleotides (Montoya et al., 1997;
Freymann et al., 1997 & 1999). In this state their catalytic IBD loops are not correctly
positioned provoking only a low basal GTPase activity (Peluso et al., 2001). Driven by
dimerization, the GTPase cycles of SRP and its receptor are initiated by a series of
discrete conformational changes that causes activation. SRP and FtsY can initially
associate to form a transient 'early' intermediate (Fig. 1.11 [2]) independently of GTP (Kd ~
4-10 µM), primarily driven by electrostatic interaction of the N-domains without stable
contacts between their G-domains (Estrozi et al., 2011; Zhang et al., 2011). A stable
closed complex (Fig. 1.11 [3]) with extensive interactions between the G-domains (~ 16-
30 nM) is formed by GTP dependent rearrangements, primarily involving readjustment of
the NG-domain interface (Shan and Walter 2003; Egea et al., 2004; Focia et al., 2004;
Shan et al., 2004) and repositioning of the inhibitory N-terminal helix N1 of FtsY
(Shepotinovskaya and Freymann 2002; Gawronski-Salerno and Freymann 2007; Neher et
al., 2008). The final rearrangement in the GTPase cycle involves repositioning of the catalytic residues in the IBD loops at the active site and results in an activated complex (Fig. 1.11 [4]) that triggers efficient GTP hydrolysis (Egea et al., 2004; Focia et al., 2004; Shan et al., 2004). This, in turn completes the cycle by causing the disassembly and release of the SRP-FtsY complex (Fig. 1.11 [5]) (Connolly et al., 1991; Peluso et al., 2001). The conformational changes during the GTPase cycle are versatilely regulated both by the cargo protein and the target membrane (Zhang et al., 2009; Lam et al., 2010; Akopian et al., 2013-b) to enhance the targeting efficiency and to provide multiple fidelity checkpoints to reject incorrect cargos (Fig. 1.11 red arrows).

![Conformational changes in SRP and FtsY GTPase domains are coupled to incorrect cargo rejection.](image)

**Fig. 1.11** Conformational changes in SRP and FtsY GTPase domains are coupled to incorrect cargo rejection. [1] The binding of an RNC exposing a proper signal sequence (purple) to SRP (blue) positions the NG domain onto the top of the SRP RNA tetraloop. [2] The cargo bound SRP forms a stabilized early intermediate with FtsY (SR; green) due to interaction between their N-domains. [3] FtsY association to anionic phospholipids in the membrane strongly accelerates the rearrangement of the early targeting complex to the closed state. [4] Interaction with the SecYEG translocon (dark grey) induces the formation of the activated state in which the NG-domain complex relocates to the distal end of SRP RNA. [5] Hydrolysis of GTP triggers disassembly of the SRP-GTPase complex while the cargo is transferred from SRP to the Sec translocon. At each step it is possible to retain (black arrows) or reject (red arrows) the cargo from the SRP pathway (Zhang and Shan, 2014). The lower panels show molecular models of the SRP (blue) – FtsY (green) complex in the early (left) and activated (right) state (adapted from Saraogi and Shan, 2013).
The first quality checkpoint in the SRP targeting cycle is the binding of SRP to RNCs displaying a correct cargo. This results in a stabilization of the early intermediate by more than 100-fold (Zhang et al., 2009), allowing the SRP-RNC-FtsY complex to be sufficiently stable to accumulate under physiological conditions (Fig. 1.11 [1]). Thereby, the assembly of the closed SRP-FtsY state is accelerated by 1000-fold (Zhang et al., 2009), accomplishing the rapid delivery to the target membrane once a correct cargo is loaded on SRP. In contrast, the less stable targeting complexes formed with incorrect cargos are much more likely to disassemble before they arrive at the membrane (Fig. 1.11 [2+3]). Another important step is the timing and regulation of GTP hydrolysis to avoid abortive targeting reactions before the complex successfully transfers the cargo to the Sec-translocation machinery. The so called kinetic proofreading by GTP hydrolysis solves this problem since a correct cargo delays the conformational rearrangement that leads to GTPase activation (Fig. 1.11 [4]). This retards the GTP hydrolysis rate by 6-8 fold (Zhang et al., 2009) and consequently extends the lifetime of the targeting complex from < 1 s to ~ 5 s. However, the delayed GTPase activation can be reversed by anionic phospholipids (Lam et al., 2010; Braig et al., 2011) and the interaction with the SecYEG complex (Akopian et al., 2013-b), indicating that GTP hydrolysis is tightly coupled to the unload of cargo at the target membrane (Fig. 1.11 [5]). In the last step of co-translational protein targeting the loaded ribosome has to be transferred to the SecYEG translocon without abortive loss of cargo. This challenging task is solved by the use of overlapping binding sites of SRP and SecYEG on the ribosome. Biochemical and genetic studies (Cheng et al., 2005; Ménétret et al., 2007) together with cryo-EM reconstitutions of the SecYEG-RNC-translocon complex (Beckmann et al., 2001; Mitra et al., 2005; Becker et al., 2009; Frauenfeld et al., 2011), showed that highly conserved basic residues in the cytosolic loops C4 and C5 of SecY interact with the ribosomal protein L23. Thus, the binding of SRP and SecYEG to RNCs is expected to be mutually exclusive and SRP has to detach from the RNC prior to its stable engagement with the translocon, involving a coordinated mechanism of cargo transfer. This process occurs via formation of a quaternary complex of RNC-SRP-FtsY-SecYEG (Akopian et al., 2013-b) and the movement of the NG dimer to the RNA distal end (Shen et al., 2012; Akopian et al., 2013-a,b) that vacates the ribosomal protein L23, thereby making it accessible for SecYEG.

There are first evidences that this could also hold for other translocation and insertion machineries since similar specific binding sites on the ribosome were found for the YidC insertase (Kohler et al., 2009; Funes et al., 2011; Kedrov et al., 2013; Seidl et al., 2014; Wickles et al., 2014; Geng et al., 2015), suggesting a general mechanism for the cargo handover of SRP dependent substrates to their final insertase or translocase.
1.3 The YidC/Oxa1/Alb3 insertase family

In all living cells, biological membranes form hydrophobic, though semipermeable, barriers that separate the intracellular compartment from the extracellular environment. These membranes contain numerous intrinsic proteins that play vital roles for the cell and make up approximately 30% of the proteome in an organism (Poetsch and Wolters, 2008). In eukaryotic cells, proteins are sorted from their source of origin in the cytoplasm to membrane-enclosed organelles such as the nucleus, endoplasmic reticulum, peroxisomes, mitochondria and chloroplasts. In prokaryotes, membrane proteins are also transported from the cytoplasm into or across the inner phospholipid bilayer. Due to the prokaryotic origin of mitochondria and chloroplasts, specialized membrane protein translocation systems have evolved that are exclusively found in these organelles and in bacteria. One of those unique protein families is the YidC/Oxa1/Alb3 insertase family which promotes membrane protein integration into bacterial inner membranes as well as into mitochondrial and plastid membranes, utilizing an evolutionarily conserved pathway for protein insertion.

1.3.1 Functional specialization in bacteria, mitochondria and chloroplasts

The YidC/Oxa1/Alb3 protein family is widely spread throughout all three domains of life (Yen et al., 2001; Zhang et al., 2009) and its members direct the insertion of integral membrane proteins into the lipid bilayer of the bacterial plasma membrane, the inner mitochondrial membrane, and the chloroplast thylakoid membrane (Saller et al., 2012; Dalbey et al., 2014). Insertase-dependent membrane proteins mainly consist of components of the respiratory chain or other energy transducing protein complexes (Luirink et al., 2001). Most bacteria contain only one YidC insertase gene copy, but some Gram-positive bacteria such as Bacilli, Lactobacilli, Actinobacteria and some Clostridia, often harbor two YidC homologs (Zhang et al., 2009; Funes et al., 2009 & 2011). The number of YidC homologs in eukaryotic organelles varies between different organisms with a maximum number of six in the plant Arabidopsis thaliana (Alb chloroplast proteins). This is probably caused by separate gene duplication and / or secondary losses of Oxa/Alb3 proteins leading to a rich diversity and to the distinct number of homologs found in each organism. The existence of YidC homologs in archaea has been proposed based on phylogenetic analysis within the Euryarchaeota but not in Nanoarchaeota and Crenarchaeota (Pohlschröder et al., 2005; Zhang et al., 2009). Recently, the first crystal structure of a YidC-like protein (DUF) from the archaeal plasma membrane of Methanocaldococcus jannaschii was solved and provided first experimental evidence for a structural and functional homology to YidC/Oxa1/Alb3 family members (Borowska et al.,
Unlike eukaryotic insertases, the archaeal YidC-like protein could not complement for YidC insertases in eubacteria, suggesting an only distantly related homology of the DUF proteins.

The YidC/Oxa1/Alb3 proteins are thought to share a conserved core domain of approximately 250-300 residues in length comprising five transmembrane (TM) domains (Saaf et al., 1998; Luirink et al., 2001; Kuhn et al., 2003; Kumazaki et al., 2014-a & -b) exhibiting the insertase activity of the proteins (Fig. 1.12). However, the conserved transmembrane core found in the shorter archaeal proteins of the DUF106 family is predicted to contain only three TM helices (Borowska et al., 2015). Extensive studies showed that those hydrophobic core domains can be functionally exchanged between different non-archaeal members of this family (Jiang et al., 2002; Funes et al., 2004-2; Preuss et al., 2005; van Bloois et al., 2007; Dong et al., 2008). Noticeably, the homology within this insertase family is mainly a kind of functional conservation rather than mere sequence homology. The YidC/Oxa1/Alb3 proteins only share a low level of primary sequence similarity and their lengths vary over a four-fold range (~200 – 800 amino acid residues; Yen et al., 2001; Jiang et al., 2003). The highest homology is found in the core TM domains 1, 2 and 5 (corresponding to TM 2,3 and 6 in E. coli and other Gram-negative bacteria) (Yen et al., 2001), whereas the N- and C-terminal flanking regions of the core domain exhibit a huge variety in length, composition and structure (Fig. 1.12). Unique to the YidC proteins of Gram-negative bacteria is an additional N-terminal TM domain connected via a large periplasmic loop to TM2, indicating an organism-specific function (Oliver and Paetzel, 2008; Ravaud et al., 2008). On the other hand, Oxa1, Alb3, Alb4 from organelles and YidC2 from the Gram-positive bacterium Streptococcus mutans contain elongated, highly positively charged C-terminal tails mediating specialized functions during protein insertion in the respective organism (Fig. 1.12).

Mitochondria of animals, fungi and plants consistently contain two insertase homologs, Oxa1 and Cox18 (Funes et al., 2004-1). Oxa1 facilitates the insertion of the mitochondrial-encoded inner membrane proteins of the respiratory chain complex as well as of the ATP synthase subunits ATP6, 4 and 9, related to the bacterial subunits F\textsubscript{0}a, F\textsubscript{0}b and F\textsubscript{0}c, respectively (Bonnefoy et al., 1994; Bauer et al., 1994; Hell et al., 2001; Jia et al., 2007). Also some nuclear encoded proteins like the ABC-transporter Mdl1 depend on Oxa1 for their correct membrane assembly (Bohnert et al., 2010; Hildenbeutel et al., 2012). The membrane insertion of mitochondrial translation products occurs in a co-translational process (Hell et al., 1997 & 1998), mediated by Oxa1 with its C-terminal α-helical region binding mitochondrial ribosomes at the exit tunnel via the ribosomal proteins Mrp20 and Mrp40. These are homologs of the bacterial L23 and L24 ribosomal proteins, respectively.
Electrophysiological measurements with purified Oxa1 showed an ability to form substrate-gated pores in lipid bilayers (Krueger et al., 2012). The second insertase in mitochondria, Cox18, is an essential assembly factor of cytochrome c oxidase (Souza et al., 2000). Compared to Oxa1 (Fig. 1.12), Cox18 lacks a C-terminal ribosome binding domain and presumably functions exclusively in a post-translational manner (Funes et al., 2004-1) downstream of Oxa1 (Saracco and Fox, 2002).

Plastids of Arabidopsis thaliana and other plants contain also two paralogs, Alb3 and Alb4, which differ in their function (Sundberg et al., 1997; Gerdes et al., 2006). Alb3 is critical for the post-translational insertion of LHCP (light-harvesting chlorophyll-binding
proteins) into the thylakoids (Moore et al., 2000; Bellafiore et al., 2002), the most abundant membrane protein system on earth. The nuclear-encoded LHCP-precursors are synthesized in the plant cell cytosol and subsequently targeted and imported into the chloroplast stroma where they bind to the plastid SRP (cpSRP) (Schuenemann et al., 1998). Instead of an SRP RNA, the plant cpSRP contains a specialized subunit, cpSRP43, mediating the interaction with the extended C-terminal domain of Alb3 (Fig. 1.12). Consequently, the C-terminal domain of Alb3 is crucial for the targeting of LHCPs to the thylakoid membrane (Moore et al., 2003; Falk et al., 2010), representing a plant-specific factor for the biogenesis of LHCPs. Alb4 does not interact with cpSRP43 and therefore cannot compensate for the loss of Alb3. Interestingly, however upon expression in mitochondria, Alb4 can functionally replace Oxa1 (Funes et al., 2004-1) indicating that its C-terminal extended domain can facilitate ribosome binding (Fig. 1.12). In plastids, Alb4 interacts with subunits of the CF$_0$CF$_1$-ATPase (Benz et al., 2009), similar to YidC and Oxa1 mediating the membrane assembly of the F$_0$F$_1$-ATPase.

The genomes of most Gram-positive bacteria encode two slightly different YidC proteins (Fig. 1.12) (Yen et al., 2001; Funes et al., 2009) with obviously overlapping activities, since only the simultaneous deletion of both leads to a lethal phenotype (Murakami et al., 2002; Tjalsma et al., 2003; Saller et al., 2009). In Bacillus subtilis the YidC proteins are termed SpoIIIJ, a protein that was found to be essential for sporulation, and YqjG that is upregulated under conditions at which the levels of SpoIIIJ are not sufficient to fully promote protein insertion (Chiba et al., 2009). Streptococcus mutans exhibits two functional orthologs of E. coli YidC, YidC1 (SpoIIIJ) and YidC2 (YqjG) (Hasona et al., 2005), each complementing multiple defects of YidC depletion in E. coli. The S. mutans YidC1 and 2 proteins can mediate insertion of both, Sec-dependent and -independent YidC substrates (Dong et al., 2008). However, deletion of YidC1 alone has only little obvious effect, whereas the phenotype of a YidC2 deletion is strikingly similar to that of SRP pathway mutants, suggesting that YidC1 and YidC2 are functionally distinct (Hasona et al., 2005). One reason for these differences is probably the C-terminal ribosome binding domain of YidC2 which is similar to that of Oxa1. This domain is lacking in YidC1 (Dong et al., 2008). Interestingly, deletion of SRP is not lethal in S. mutans until simultaneous deletion of YidC2, suggesting that YidC2 supports a co-translational insertion pathway in Gram-positive bacteria that does not solely rely on the SRP pathway. This hypothesis was supported by the observation that YidC2 expression in yeast mitochondria promotes co-translational integration of proteins into the mitochondrial inner membrane in the absence of Oxa1 (Funes et al., 2009). Thus, a gene duplication of YidC
proteins in Gram-positive bacteria obviously allowed the specialization of one paralog for a co-translational protein insertion route (Funes et al., 2009).

Gram-negative bacteria usually compromise only one YidC protein with a short C-terminal tail and no essential function could yet be ascribed to this domain (Jiang et al., 2003) (Fig. 1.12). However, genomic analyses and sequence comparisons show that YidC homologs in many marine and extremophilic eubacteria have C-terminal hydrophilic and often highly charged extensions comparable to the eukaryotic organellar YidC-like proteins (Kiefer and Kuhn, 2007). In this work, evidence is presented for a possible (ancestral) function of these extended tails by analyzing the function of the YidC C-tail regions of two marine prokaryotes in membrane targeting. One of those bacteria, the marine planctomycete Rhodopirellula baltica was first isolated from the Baltic Sea (Schlesner et al., 2004) as a unique bacterium which can anaerobically oxidize ammonium in specialized cellular compartments, the anammoxomes. Rhodopirellulas genome of 7.15 Mbp is one of the largest found in bacteria (Glöckner et al., 2003). The YidC homolog of R. baltica has a size of 90 kDa and a hydrophilic C-terminal tail region of 80 amino acid residues with a predicted isoelectric point (pI) of 10.45. Oceanicaulis alexandrii, the second marine bacterium used in this study, is a stalked, aerobic α-proteobacterium, isolated originally as a symbiont from the dinoflagellate Alexandrium tamarense (Strömpl et al., 2003). Its YidC homolog is 83 kDa in size and has a large, positively charged C-terminal tail of 99 residues with an estimated pI of 9.85. In comparison, E. coli YidC is 61 kDa in size and its C-terminal region is only 16 amino acid residues long. Marine prokaryotic organisms are, in many cases, unique in their genomic organization (Serres et al., 2009). Particularly in the planctomycete group, several gene duplications and protein motifs are found that only have counterparts in archaeal or even eukaryotic phyla (Studholme et al., 2004). Thus, studying marine bacterial groups may be a key to understand how complex cellular processes like protein translocation and membrane targeting may have evolved and spread over the whole organismic world.

Taken together, the gene duplications broaden the physiological range of substrates using YidC members during cell evolution and revealed a phylogenetic tree of YidC/Oxa1/Alb3 membrane insertases with 6 subbranches that evolved from a common ancestor (Funes et al., 2011). In this phylogenetic analysis two main bacterial branches exist (YidC, YidC1, SpoIIIJ and YidC2, YqjG) where the YidC gene first arose in proteobacteria and only later in Gram-positive bacteria where it was duplicated in the firmicutes subphylum to form the YidC1 and YidC2 subbranches (Zhang et al., 2009). Since the proteobacteria and cyanobacteria, being most closely related to mitochondria and chloroplasts respectively,
have only a single copy of the YidC gene, gene duplications must have occurred after the endosymbiotic event that leads to mitochondria and chloroplasts. This resulted in two mitochondrial branches (Oxa1 and Cox18) and two chloroplast branches (Alb3 and Alb4). The YidC/Oxa1/Alb3 homologs of eubacteria and eukaryotes are more closely related to each other than to the highly diverged archaeal DUF106 proteins (Luirink et al., 2001; Zhang et al., 2009). Their lack of two of the five TMs present in the other YidC homologs may reflect differences in substrate specificity or in the mechanism of action in the archaeal DUF insertases (Borowska et al., 2015). This is consistent with the finding that many of the YidC/Oxa1/Alb3 paralogues, particularly the various insertase proteins in organelles and in Gram-positive bacteria, differ in their substrate preference and therefore differ in their distinct function (Saller et al., 2009; Benz et al., 2009 & 2013) representing organism specific, specialized membrane insertases.

### 1.3.2 The bacterial membrane insertase YidC

In 1994 it was discovered that the mitochondrial Oxa1 has homologous genes in Gram-negative and Gram-positive bacteria (Bonnefoy et al., 1994). Six years later the function of the bacterial Oxa1 homolog, termed YidC, as an essential membrane insertase in *E. coli* was shown by Samuelson et al., (Samuelson et al., 2000). Its Sec-dependent membrane protein inserting function was discovered by Scotti et al. (Scotti et al., 2000) in the same year. Since then, the membrane insertase YidC of *E. coli* was the focus of many structural and functional studies to understand the new membrane insertion mechanism and is one of the best-studied YidC/Oxa1/Alb3 family members. The 548 amino acid residues comprising YidC of *E. coli* is a six-spanning membrane protein with 61 kDa in size and is integrated into the inner membrane by the SecYEG/SecA translocase as well as in an SRP dependent manner. This suggests that YidC, in contrast to its mitochondrial orthologue Oxa1, cannot engage a SecYEG-independent protein conducting channel (Koch et al., 2002).

A large number of substrate proteins was identified that are inserted by the bacterial YidC alone or in cooperation with the Sec translocon. This results in a much broader substrate spectrum of YidC compared to the rather limited number of substrates inserted by Oxa1 or Alb3 in organelles. In cooperation with the Sec translocon, YidC promotes the insertion of the F$_1$F$_0$ ATP synthase subunits F$_0$A and F$_0$B, the NADH-quinone oxidoreductase subunit K NuoK, and CyoA, the subunit 2 of the cytochrome bo oxidase (Yi et al., 2004; du Plessis et al., 2006; van Bloois et al., 2006; Celebi et al., 2006; Kol et al., 2009; Price and Driessen 2010). In addition to its insertase function, YidC is also required for the folding of LacY (Nagamori et al., 2004; Zhu et al., 2013) and MalF (Wagner et al., 2008) in the
membrane bilayer as well as for the membrane assembly of the oligomeric maltose transporter MalFGK (Wagner et al., 2008). As part of the holo-translocon (HTL) SecYEG-SecDF-YajC-YidC (Scotti et al., 2000; Schulze et al., 2014), YidC interacts directly with SecY involving the residues Gly 355 in TM2 and Met 471 in TM4 of E. coli YidC (Li et al., 2014). Supposedly, by this interaction YidC promotes the dislodge of the transmembrane segments of inserting membrane proteins from the Sec channel through its lateral gate (Zhu et al., 2012), facilitating their integration into the lipid bilayer (Urbanus et al., 2001), thus serving as an assembly site for multi-spanning membrane proteins (Beck et al., 2001). With approximately 2500 to 3000 copies per cell, YidC is much more abundant than the SecYEG complex (Urbanus et al., 2002), and therefore probably only a portion of the YidC protein pool forms a constituent of the HTL. Recently, the isolation of a stable holo-translocon revealed two distinct versions of the HTL. One containing a single copy of SecYEG, YidC, SecDF and YajC each, that functions preferably in membrane protein insertion, while the second HTL species is comprised of a SecYEG dimer only which functions in protein export (Schulze et al., 2014).

As an independent insertase, YidC facilitates the insertion of F_{0}C, the subunit c of F_{1}F_{0} ATP synthase, the mechanosensitive channel protein MscL, the phage proteins M13 procoat and Pf3 coat, and the tail anchored membrane protein TssL (Samuelson et al., 2000 & 2001; Chen et al., 2002; Yi et al., 2003; van Bloois et al., 2004; Facey et al., 2007; Aschtgen et al., 2012). Interestingly, all these substrate proteins comprise less than 200 amino acid residues and have only short periplasmic loops (< 30 residues). The region of the five conserved TM domains 2 - 6 of E. coli YidC is crucial for its insertase function, whereas the unique features of Gram-negative YidC homologs, the additional TM1 and the large periplasmic loop 1 (P1), are dispensable (Jiang et al., 2003). However, the additional N-terminal TM helix serves as a signal sequence for inserting YidC correctly into the inner membrane. The P1 domain interacts with multiple components of the Sec machinery and is considered to facilitate stable complex formation particularly through binding to the SecF-subunit (Xie et al., 2006; Sachelaru et al., 2013). The large periplasmic P1 loop was the first domain of the E. coli YidC whose structure was solved. It revealed a β-super sandwich fold with short α-helical domains in the interior as well as at the edge of the sandwich. A flexible linker of two α-helices at the very C-terminus connects the P1 loop to the downstream core domain (Oliver and Paetzel, 2008).

In the year 2014, Kumazaki et al. reported the first crystal structure of a bacterial YidC protein. The molecular structure from Bacillus halodurans YidC1 was solved to 2.4 Å resolution (Kumazaki et al., 2014-a), contributing a landmark to the membrane biology
field. In late 2014, the X-ray structure of *E. coli* YidC was also solved to 3.2 Å by Kumazaki *et al*. (Kumazaki *et al*., 2014-b), providing new insight into the mechanism by which YidC inserts proteins into the bacterial inner membrane. Recently, Borowska *et al*., (2015) reported the crystal structure and provide experimental evidence that the DUF106 (Mj0480) protein from *M. jannaschii* represents a YidC-like archaeal membrane insertase. This provides the evidence that those membrane insertases are present in all three domains of life. All of these structures show closed topologies towards the periplasmic / extracellular space and form a hydrophilic groove that is open to the cytoplasm and also to the lipids of the membrane (Fig. 1.13), providing a new mechanism of membrane protein insertion. In contrast to the Sec translocase, YidC has no channel that is controlled by a hydrophobic pore ring and a periplasmic lid, but rather provides a platform to bind small proteins in an amphiphilic groove and reduces the required translocation energy to cross the membrane (Dalbey and Kuhn, 2014). Both bacterial YidC structures showed that the five conserved TM domains are tightly packed in the periplasmically exposed half of the membrane and spread out in the cytoplasmic half. Strikingly, the three TMs of the much smaller (23 kDa) Mj0480 DUF-like protein form a structure similarly folded as parts of the YidC hydrophilic groove of *B. halodurans* YidC1 and the *E. coli* YidC (Fig. 1.13).

Another feature that is conserved in all family members is the accumulation of methionine residues at the cytoplasmic surface region of the YidC proteins, similar to the methionine rich M-domain of SRP which facilitates the binding of the signal peptide (Bernstein *et al*., 1989). This suggests that these methionine residues might also play a role in binding the inserting TM segments of the substrate proteins (Borowska *et al*., 2015). This methionine rich region includes the first cytoplasmic loop (C1) which contains a dynamic coiled coil motif (CH1+2 Fig. 1.13) being located near the cytosolic entrance of the hydrophilic groove and was shown to be essential for YidC function *in vivo* (Chen *et al*., 2014; Kumazaki *et al*., 2014-a). The corresponding region in Mj0480 is predicted to form also a coiled coil but is disordered and therefore not shown in the crystal structure (Fig. 1.13).

Despite the similarities of the membrane-embedded core regions, the insertases from the archaeon *M. jannaschii*, the Gram-positive *B. halodurans*, and the Gram-negative *E. coli* show an increasing complexity having gained additional features like the large periplasmic domain (P1 Fig. 1.13) of the *E. coli* YidC. The archaeal version of YidC possibly describes the minimal core motif forming the insertase domain. A subsequent step in characterizing the role of the DUF106 protein in the archaeal membrane-protein biogenesis is to find an endogenous substrate and test the ability to cooperate with the Sec translocase.
Fig. 1.13 Structures of the archaeal, Gram-positive and Gram-negative membrane insertases from *M. jannaschii*, *B. halodurans* and *E. coli*. Cartoon representation of the homologous YidC proteins depicting a transverse profile through the membrane. The respective homologous TM domains are displayed in the same color code and cytosolic and periplasmic/extracellular domains in green, respectively. CH1+2: cytoplasmic helix 1 and 2; EH1+2: extracellular helix 1 and 2; P1: periplasmic loop 1. The coiled-coil C1 region of the *M. jannaschii* YidC, as well as TM1 and the N-terminal part of the P1 domain of the *E. coli* YidC are not visible in the crystal structure due to their high flexibility (Kumazaki et al., 2014-b; Borowska et al., 2015). The three structures of YidC insertases demonstrate the conservation of the basic structural elements of the membrane-embedded core domain, as well as the additional features that were added on or were lost during evolution (adapted from Dalbey and Kuhn, 2015).

Several substrate contacts were previously reported for the *E. coli* YidC to Sec-dependent and -independent substrate proteins in *in vivo* and *in vitro* cross-linking assays (Yu et al., 2008; Klenner and Kuhn, 2012). Combined with the new insights of the *E. coli* YidC crystal structure, they are located in the groove between TM3 and TM5, on the exterior region of the TM as well as in the hydrophilic groove (Kumazaki et al., 2014-b). This leads to the assumption that residues facing the hydrophilic cavity contact the translocated, hydrophilic regions of the substrate proteins while those facing the membrane bilayer assist the insertion and lateral integration of the substrate's hydrophobic segments into the membrane (Hennon et al., 2015). Recently it was shown that the evolutionarily conserved positively charged arginine residue in the hydrophilic groove of YidC is crucial only in
some members of the YidC/Oxa1/Alb3 family, suggesting that there are different insertion requirements based on the characteristics of the substrates (Chen et al., 2014). Finally, the crystal structures of YidC provide a first insight into the unique mechanism by which YidC is inserting proteins into the membrane lipid bilayer as an independent, monomeric insertase (Kumazaki et al., 2014-a). In the initial step, a single spanning membrane protein with an acidic N-terminal extracellular region, such as the Pf3 coat protein, binds electrostatically to the membrane inner surface and contacts the flexible C1 region of YidC (Fig. 1.14 A). Subsequently, the negatively charged residues in the N-tail interact, in some cases, with the conserved positively charged arginine residue in the hydrophilic groove of YidC and is then transiently captured in its aqueous cavity (Fig. 1.14 B). This binding step can induce structural changes in the hydrophilic groove to accommodate the various substrate proteins. Catalyzed by electrophoretic attraction of the positively charged membrane potential side in the periplasmic space on the negatively charged residues in the translocated region of the substrate (Chen et al., 2002; Zhu et al., 2013) and hydrophobic interaction between the substrate TM segment and YidC (Fig. 1.14 C), the N-tail is then released from the groove and crosses the outer leaflet of the membrane lipid bilayer (Fig. 1.14 D).

**Fig. 1.14 Model for the insertion of a single-spanning membrane protein via YidC.** (A) Prior to the interaction with YidC in the resting state, the substrate protein is peripherally associated with the membrane. (B) The hydrophilic N-terminal region of the substrate would then be transiently captured in the hydrophilic substrate groove of YidC, possibly mediated by interaction with a cytoplasmic loop (C1) and hydrophilic interaction, resulting in the substrate-bound state of YidC. (C) Substrate insertion is believed to be facilitated by the proton motive force, hydrophobic interaction between the TM regions and the aliphatic chains, as well as possibly thinning of the membrane by YidC. (D) After successful translocation and insertion into the bilayer via a greasy slide, the substrate protein is laterally released from YidC and YidC returns to the resting state. (adapted from Kumazaki et al., 2014-a and Hennon et al., 2015)
INTRODUCTION

It is proposed that the inserting substrate TM segment moves along a greasy slide formed by TM3 and TM5 of the *E. coli* YidC in order to form a proper transmembrane configuration (Dalbey and Kuhn, 2014). To which extent this mechanism holds also true for the insertion of other YidC substrate classes like double-spanning membrane proteins, remains to be elucidated. In addition, further biological and structural studies are required to understand how the individual YidC homologs differ in their insertion mechanism and in their determinants of substrate properties that govern the YidC-dependent integration into the membrane.

In *E. coli*, the conserved arginine residue in TM2 is not essential for insertion or cell survival (Jiang *et al.*, 2003), whereas the substitution of the conserved arginine residue R73 in *B. halodurans* YidC1 abolished the insertion activity of the endogenous substrate protein MifM (Kumazaki *et al.*, 2014-a). As already mentioned, this underlines the fact that the conserved positively charged residue in the hydrophilic groove is critical for the insertase function only in certain YidC homologs (Chen *et al.*, 2014) and some YidC substrates may use a different combination of specific structural features for insertion. Two recent studies analyzed the influence of TM segment composition and charge distribution of the translocated regions in guidance of the substrate membrane proteins into the YidC, SecYEG or YidC/SecYEG insertion pathway, respectively. Zhu *et al.* (2013) and Soman *et al.* (2014) proposed that increasing hydrophobicity in TM segments and decreasing polarity and number of charges in the translocated loop or tail region lowers the requirement of YidC or SecYEG for insertion. In addition, they showed that a positive charge in a translocated loop requires YidC and SecYEG, while a negative charge in the loop region only required YidC. These findings are in good agreement with the proposed insertion mechanism by Kumazaki *et al.* (2014-a) (Fig. 1.14). Earlier, it had been shown that Oxa1 in mitochondria was important for the insertion of proteins with highly charged domains and negative charges in the translocated region (Herrmann and Bonnefoy, 2004), proposing a conserved and specialized insertion mechanism of the YidC/Oxa1/Alb3 insertase family as respiratory proteins, their major substrates often have negatively charged residues that are essential for function (Price and Driessen, 2010). Although it was shown that charge-unbalanced TM segments more likely depend on YidC for insertion (Gray *et al.*, 2011), many of the identified YidC-dependent proteins did not have unbalanced TM segments and other features obviously also influence the recruitment to and insertion via YidC.
1.3.3 YidC and the ribosome

Despite recent structural insights, our understanding of how ribosomes associate with the components of the insertion machinery, especially with YidC, is still relative scarce. Hence, further investigation of the coupling of co-translational targeting to YidC-mediated insertion might be an important step forward to understand the exact molecular mechanism of YidC-mediated membrane protein biogenesis. Due to the hydrophobic nature of the YidC substrate proteins, those nascent membrane proteins have a high risk to form insoluble, nonfunctional and non-translocation-competent aggregates. Thus, a tight coupling of the synthesis to the insertion of membrane proteins appears to be crucial. The C-terminal regions of the mitochondrial Oxa1 (89 residues; total charge +24) and S. mutants YidC2 (64 residues; +14) proteins have been shown to be pivotal for direct ribosome interaction (Szyrach et al., 2003; Jia et al., 2003; Wu et al., 2013), and deletions within these domains compromised the protein insertion efficiency. In mitochondria it is known that the C-terminal tail of Oxa1 tethers the mitochondrial ribosomes permanently to the membrane, presumably as an evolutionary consequence of the elimination of genes in the mitochondrial genome that code for hydrophilic proteins (Szyrach et al., 2003; Jia et al., 2003). Moreover this spurred the loss of an SRP-like pathway (Funes et al., 2009). In E. coli, the hydrophilic, C-terminal segment of YidC is moderately positively charged and substantially shorter (16 residues; +8) than the mitochondrial counterpart. Its role in facilitating ribosome binding was analyzed in several studies (Kohler et al., 2009; Funes et al., 2009; Kedrov et al., 2013; Seidl et al., 2014; Geng et al., 2015) and is also an important scope of this work.

Cross-linking studies indicated that during membrane protein synthesis, YidC and bacterial ribosomes are in close contact (de Gier and Luirink 2003; Welte et al., 2012) and binding of a detergent-solubilized E. coli YidC dimer to translating ribosomes was visualized by cryo-EM for the first time in 2009 (Kohler et al., 2009), proposing an Oxa1-like insertase-ribosome complex. However, subsequent studies revealed that the addition of just six histidine residues to the C-terminus of E. coli YidC, as was used for the cryo-EM studies by Kohler et al., is sufficient to bind YidC to ribosomes (Kedrov et al., 2013). Moreover, Kedrov et al. showed that YidC binds only to ribosomes if they expose a YidC substrate from their exit tunnel and that the C-terminal domain of YidC is involved but not essential for ribosome recruitment in E. coli. In contrast to the data published by Kohler et al. (2009) describing a YidC dimer binding to ribosomes, it was shown that a single YidC copy is sufficient to bind a substrate-translating ribosome, thus being the minimal functional unit (Kedrov et al., 2013; Seidl et al., 2014). This suggests that E. coli ribosomes do not intrinsically associate with YidC but only when synthesizing YidC substrates or
when additional interacting sequences are added to its C-terminus (Kohler et al., 2009; Seitl et al., 2014). Other supplementary factors, particularly the bacterial SRP and its receptor, might further contribute to this interaction in vivo. Since the C-terminal tail of the E. coli YidC was thought not to be the only determinant for its interaction with the ribosome, the positively charged cytosolic loops (C1 and C2) were suggested as potential binding partners (Kedrov et al., 2009; Wickles et al., 2014). Geng et al. (2015) applied both in vivo and in vitro analyses to investigate the roles played by these two cytosolic domains of E. coli YidC. They demonstrate that the cytosolic loop C2 and the C-terminal tail of YidC may determine binding to RNCs, while the C1 loop is essential for cell viability but not for ribosome interaction. The approximately 10 amino acid residues long C2 loop of YidC is proposed to be flexible, since the conformation of this region could not be resolved in the crystal structures of YidC (Kumazaki et al., 2014-a & 2014-b). The C2 loop bears one conserved charged residue (Asp488) which is essential in this position for cell viability (Wickles et al., 2014), and is proposed to ensures the stable docking of a translating ribosome to YidC. Further studies will be necessary to elucidate the precise interplay between YidC, SRP and the ribosome nascent chain complexes in E. coli.
1.4 Aims of the study

For the insertion of inner membrane proteins in *E. coli*, the membrane translocase SecYEG and the membrane insertase YidC, as well as the signal recognition particle (SRP) targeting system are of central importance. Insertion of integral membrane proteins into the lipid bilayer takes place co-translationally at membrane-associated ribosomes (Luirink and Sinning, 2004; Egea *et al.*, 2005; Halic and Beckmann, 2005). Not fully clear, however, is how SRP recognizes the specific features of nascent chains of hydrophobic proteins and how the ribosome nascent chain (RNC) complexes are transferred to the respective translocase or insertase, especially in the YidC-only insertion pathway.

The primary subject of this work was to characterize and analyze in detail the function of a C-terminal ribosome-binding domain on YidC and its role in SRP mediated membrane insertion in Gram-negative bacteria. Therefore, the C-terminal ribosome-binding domains of YidC homologs from the marine Gram-negative bacteria *Rhodopirellula baltica* and *Oceanicaulis alexandrii* were fused to the short C-terminal region of the *E. coli* YidC. In order to understand the mechanism by which the YidC insertase is coupled to translating ribosomes, the determination of a cryo-EM structure of RNCs in complex with the C-terminally extended YidC was one aim of this study. Despite the structural analysis, different biochemical approaches were performed to identify the interaction mode of the C-terminal ribosome binding domains with particular ribosomal proteins of the polypeptide exit tunnel. Furthermore, *in vivo* studies in *E. coli* were performed to investigate a possible role of the C-terminally extended YidC in SRP-mediated co-translational membrane insertion of the YidC-dependent MscL protein.

The second part focuses on the interaction of the signal recognition particle with SRP signal sequences to determine the specificity of SRP recognition in proteins. The interaction studies were established in an *in vitro* system and binding affinities were determined by microscale thermophoresis (MST), a new approach that enables immobilization-free, in-solution kinetic measurements of biomolecular interactions.

Taken together, this study contributes to the understanding of the molecular mechanisms of co-translational membrane protein biogenesis in bacteria and also highlights evolutionary aspects of this complex vital cellular process.
2 MATERIALS AND METHODS

2.1 Culture media and additives

2.1.1 LB medium (Luria-Bertani)

10 g/l tryptone, 5 g/l yeast extract, 5 g/l NaCl

The pH was adjusted to 7.5 with NaOH. For the preparation of LB-agar plates, 15 g/l Agar No1 were added prior to autoclaving. Antibiotics and other additives are added after autoclaving when the medium has cooled down to approximately 50 °C.

2.1.2 Antibiotics

<table>
<thead>
<tr>
<th>antibiotic</th>
<th>stock solution</th>
<th>working concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ampicillin</td>
<td>200 mg/ml (ddH2O)</td>
<td>200 µg/ml</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>25 mg/ml (EtOH)</td>
<td>25 µg/ml</td>
</tr>
<tr>
<td>Kanamycin</td>
<td>25 mg/ml (ddH2O)</td>
<td>25 µg/ml</td>
</tr>
<tr>
<td>Streptomycin</td>
<td>10 mg/ml (ddH2O)</td>
<td>10 µg/ml</td>
</tr>
</tbody>
</table>

2.1.3 Sugars

<table>
<thead>
<tr>
<th>Sugar</th>
<th>Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabinose</td>
<td>20 % 0,2 %</td>
</tr>
<tr>
<td>Glucose</td>
<td>40 % 0,4 %</td>
</tr>
</tbody>
</table>

2.2 Bacterial strains

2.2.1 Escherichia coli XL1-Blue

*lac thi1 gyrA96 endA1 hsdR17 relA1 supE44 recA1 [F’ proAB lacI9 lacZ ΔM15 Tn10]*

*E. coli* XL1-Blue is a cloning strain with recombination deficiency, used for the amplification and subsequent isolation of plasmid DNA. The mutated lacZ gene in this strain allows blue-white screening with appropriate vectors.

2.2.2 Escherichia coli BL21 (DE3)

*E. coli* B F’ dcm ompT hsdS (rB mB) gal λ (DE3)

*E. coli* BL21 contains an integrated bacteriophage λ DE3 on its chromosome, carrying the T7 RNA-polymerase gene under the control of the *lac* promoter. This strain is used for
intense overexpression of genes which are cloned into expression vectors under the control of the bacteriophage T7 promoter. The gene expression is induced by the addition of IPTG.

2.2.3 Escherichia coli C41 (DE3) (Miroux and Walker, 1996)

E. coli B F− dcm ompT hsdS (rB− mB−) gal λ (DE3); additional uncharacterized mutations

E. coli C41 (DE3) is a derivative of E. coli BL21 (DE3), discovered by Miroux and Walker during screening studies of BL21 to find mutants that allow overexpression of proteins which are toxic to BL21 cells.

2.2.4 Escherichia coli C43 (DE3) (Miroux and Walker, 1996)

E. coli B F− dcm ompT hsdS (rB− mB−) gal λ (DE3); additional uncharacterized mutations

E. coli C43 (DE3) is developed from C41 during further screening studies. This strain allows over expression of particular heterologous membrane proteins which are toxic to E. coli C41 (DE3) cells.

2.2.5 Escherichia coli MK6S (Klenner et al., 2008)

E. coli F− araD139 Δ(ara-leu) 7696 galE15 galK16 Δ(lac) X74 rpsL (Strr) hsdR2 (rB− mB−) mcrA mcrB1 para YidC

MK6S is a derivative of the E. coli strain MC1061 (Casadaban and Cohen, 1980). In this strain the promoter region of yidc on the chromosome was replaced by an araC-araBAD promoter cassette. This allows regulation of the YidC expression by the addition of arabinose (YidC expression) or glucose (YidC depletion).

2.2.6 Escherichia coli MC-Ffh (Seitl et al., 2014)

MC1061-Kan-AraCP-ffh

The Ffh depletion strain MC-Ffh was derived from E. coli MC1061 by homologous recombination of the araC-araBAD operator region, containing a kan resistance cassette for screening, into the chromosome upstream of the ffh gene (2.6.11).

2.2.7 Escherichia coli lY26

BW25113-Kan-AraCP-ftsY

The FtsY depletion strain lY26 was obtained from E. Bibi. FtsY is under the control of the araBAD promoter and operator.
2.3 Plasmids

2.3.1 Complementation assay (A)

<table>
<thead>
<tr>
<th>Nr.</th>
<th>Vector</th>
<th>Gene</th>
<th>Modification</th>
<th>Cloning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pGZ119 EH</td>
<td><em>Escherichia coli</em> YidC</td>
<td>-</td>
<td>Diploma thesis; Sandra Grenz, 2009</td>
</tr>
<tr>
<td>2</td>
<td>pGZ119 EH</td>
<td><em>Rhodopirellula baltica</em> YidC</td>
<td>-</td>
<td>EcoR1 fragment from pGEM-Teasy</td>
</tr>
<tr>
<td>3</td>
<td>pGZ119 EH</td>
<td>Oceanicaulis YidC</td>
<td>-</td>
<td>Nde fragment from pGEM-T (diploma thesis; Ines Seitl, 2010)</td>
</tr>
<tr>
<td>4</td>
<td>pGZ119 EH</td>
<td><em>Saccharomyces cerevisiae</em> mOxa1</td>
<td>mature Oxa1, without mitochondrial targeting sequence</td>
<td>Nde / BamH1 from pUC18</td>
</tr>
<tr>
<td>5</td>
<td>pGZ119 EH</td>
<td>YidC-Rb</td>
<td><em>E. coli</em> YidC 1-540 and <em>R. baltica</em> YidC 748-827</td>
<td>Nde (diploma thesis Ines Seitl, 2010)</td>
</tr>
<tr>
<td>6</td>
<td>pGZ119 EH</td>
<td>YidC-Oa</td>
<td><em>E. coli</em> YidC 1-540 and <em>O. alexandrii</em> YidC 575-673</td>
<td>Nde</td>
</tr>
<tr>
<td>7</td>
<td>pGZ119 EH</td>
<td>YidC-Oxa</td>
<td><em>E. coli</em> YidC 1-540 and <em>S. cerevisiae</em> Oxa1 272-360</td>
<td>Nde</td>
</tr>
</tbody>
</table>

Genomic DNA from *R. baltica* and *O. alexandrii* was used to amplify the homologous *yidC* genes with an Nde restriction site, serving as start codon. The second Nde site was
obtained by cloning the PCR product into pGEM-T / pGEM-T easy vectors. Based on this constructs, the yidC genes can be cloned by Nde restriction in several other vectors.

Three chimeric proteins of the *E. coli* YidC were constructed (YidC-Rb, YidC-Oa, YidC-Oxa) with the C-terminal region of YidC replaced. First, an NsiI restriction site at the codon representing the amino acid position 539 of the *E. coli* YidC was introduced using site-directed mutagenesis. The C-terminal extensions of both marine YidC homologs and Oxa1 were amplified with flanking Nsi sites. The resulting PCR products were digested with Nsi and cloned into *E. coli yidC* resulting in YidC-Rb, YidC-Oa and YidC-Oxa fusion proteins.

### 2.3.2 Protein overexpression and purification (B)

<table>
<thead>
<tr>
<th>Nr. B</th>
<th>Vector</th>
<th>Gene</th>
<th>Modification</th>
<th>Cloning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pET16b</td>
<td><em>Escherichia coli</em> YidC (Ec-YidC)</td>
<td>N-terminal His\textsubscript{10} Tag</td>
<td>Nde</td>
</tr>
<tr>
<td>2</td>
<td>pET16b</td>
<td><em>Oceanicaulis alexandrii</em> YidC (Oa-YidC)</td>
<td>N-terminal His\textsubscript{10} Tag</td>
<td>Nde from pGEM-T</td>
</tr>
<tr>
<td>3</td>
<td>pET16b</td>
<td>YidC-Rb; chimeric gene EcYidC-RbCT</td>
<td>N-terminal His\textsubscript{10} Tag</td>
<td>Nde from pGZ 119 EH</td>
</tr>
<tr>
<td>4</td>
<td>pET16b</td>
<td>YidC-Oa; chimeric gene EcYidC-OaCT</td>
<td>N-terminal His\textsubscript{10} Tag</td>
<td>Nde from pGEM-T</td>
</tr>
<tr>
<td>5</td>
<td>pET16b</td>
<td>RbCT</td>
<td><em>R. baltica</em> YidC 748-827 N-terminal His\textsubscript{10} Tag</td>
<td>Nde-Xho (PCR)</td>
</tr>
<tr>
<td>6</td>
<td>pET16b</td>
<td>OaCT</td>
<td><em>O. alexandrii</em> YidC 575-673 N-terminal His\textsubscript{10} Tag</td>
<td>Nde-Xho (PCR)</td>
</tr>
</tbody>
</table>
The PCR primers for the generation of the C-terminal YidC peptides from *R. baltica* YidC (RbCT), *O. alexandrii* YidC (OaCT), *A. thaliana* Alb3 (A3CT) and *S. cerevisiae* Oxa1 (O1CT) contained flanking Nde and Xho restriction sites. This allowed direct cloning into the respective expression vectors (pET16b and pET22b).

The ribosomal proteins L23, L24 and L29 from *E. coli* were cloned by Emmanuelle Mboubi Kouaga during her Master thesis work in this lab (Strukturelle Dynamik der ribosomalen Proteine L23, L24 und L29 bei der Interaktion mit Membraninsertasen, 2013).
### 2.3.3 Pulse chase assay (C)

<table>
<thead>
<tr>
<th>Nr. C</th>
<th>Vector</th>
<th>Gene</th>
<th>Modification</th>
<th>Cloning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pGZ119 HE</td>
<td><em>Oceanicaulis alexandrii</em> YidC</td>
<td>N-terminal His&lt;sub&gt;10&lt;/sub&gt; Tag</td>
<td>Xba-EcoR1 from pET16b</td>
</tr>
<tr>
<td>2</td>
<td>pGZ119 HE</td>
<td><em>Escherichia coli</em> YidC</td>
<td>N-terminal His&lt;sub&gt;10&lt;/sub&gt; Tag</td>
<td>Nde from pET16b into pGZ119 HE <em>Oceanicaulis alexandrii</em> YidC</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>YidC-Rb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>YidC-Oa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>YidC-Oxa</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>pGZ119 HE</td>
<td><em>Escherichia coli</em> Ffh</td>
<td>N-terminal His&lt;sub&gt;10&lt;/sub&gt; Tag</td>
<td>Nde from pGEM-T into pGZ119 HE <em>Oceanicaulis alexandrii</em> YidC</td>
</tr>
<tr>
<td>7</td>
<td>pMS119 EH</td>
<td><em>Escherichia coli</em> MscL I68C; single cys mutant</td>
<td>N-terminal His&lt;sub&gt;10&lt;/sub&gt; Tag</td>
<td>Facey <em>et al.</em>, 2007 (pSF147)</td>
</tr>
</tbody>
</table>

For the pulse chase analyses (2.9) the YidC constructs were cloned into the pGZ119 HE vector by Nde restriction. In this vector the *yidC* genes are under the control of the *tac* promoter and the CoID replicon allowed coexpression of the proteins with pMS119EH-MscLI68C.

### 2.3.4 Fluorescence microscopy (D)

<table>
<thead>
<tr>
<th>Nr. D</th>
<th>Vector</th>
<th>Gene</th>
<th>Modification</th>
<th>Cloning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pMS 119 EH</td>
<td><em>E. coli</em> MscL</td>
<td>C-terminal GFP fusion ; N-terminal His&lt;sub&gt;10&lt;/sub&gt; Tag</td>
<td>Xba-HindIII from pSF195 into pSF147 (Facey <em>et al.</em>, 2007)</td>
</tr>
</tbody>
</table>
2.3.5 *In vitro* transcription (E)

<table>
<thead>
<tr>
<th>Nr. E</th>
<th>Vector</th>
<th>Gene</th>
<th>Modification</th>
<th>Cloning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pUC18</td>
<td>4.5S RNA of <em>E. coli</em> SRP</td>
<td>T7 promoter sequence</td>
<td>I. Sinning group; BZH</td>
</tr>
</tbody>
</table>

2.4 PCR primer

Primers for cloning of *Rhodopirellula baltica* YidC (Rb-YidC)

**RhocaNde:** [5′ - 3′]: A GGA GGT TGA CAT ATG GCC GAC AGT TCA GTG
GAA CGC CGA C

**RhocaChistag(Hind)** [5′ - 3′]: C GAA AAG CTT CTA ATG GTG ATG GTG ATG GTG
ATG GTG ATG GTG ACC GCG GCG TTT CTT GTT GCC GGG

Template DNA: Rb-YidC in pGEM-T

Primers for cloning of YidC-Rb

**RiboNsi-fw:** [5′ - 3′]: GGC GGC CGC ATG CAT CCAT CGA AGC CTC ACT TC

**RiboNsi-rv:** [5′ - 3′]: C GCG GCC GCA TGC ATG CAT GCG GCG TTT CTT GTT G

Template DNA: Rb-YidC in pGEM-T

Primers for cloning of *Oceanicaulis alexandrii* YidC (Oa-YidC)

1) Ocean-fw: [5′ - 3′]: CGC CCT CAA AAC CGT AAG GGA CCG

**Ocean-rev:** [5′-3′]: CCT GGG CTT CGG GAG AGG GCT C

Template DNA: chromosomal DNA from *O. alexandrii*

2) Ocean- *Nde*-fw: [5′-3′]: CAT ATG GGT GAG AAC CGC AAT TTC CTG

**Ocean-TAG-rev:** [5′-3′]: CTA TTT CTT CTT GCC GCC GCG

Template DNA: Oa-YidC from 1) in pGEM-T
Primers for cloning of **YidC-Oa**

Ocean-Rib-fw (Nsi): \[5’- 3’\]: GGC GGC CGC ATG CAT CGT CAG GGC GTG GTG
ACC

Ocean-Rib-rev (Nsi): \[5’- 3’\]: CGC GGC CGC ATG CAT CTA TTT CTT CTT GCC
GCC

Template DNA: Oa-YidC from 1) in pGEM-T

Primers for cloning of **YidC-Oxa**

O1CT-Nsi fw: \[5’- 3’\]: GA TAC ATG CAT ATG ATT TTG AGA AAC AAA TGG

O1CT-Nsi rev: \[5’- 3’\]: GC ACG ATG CAT TCA TTT TTT GTT ATT AAT GAA G

Template DNA: *Saccharomyces cerevisiae* mOxa1 in pUC18

Primers for cloning of **RbCT N-His\textsubscript{10}**

Rhoba-Nde: \[5’- 3’\]: GG AAA AAA CGT CAT ATG CAT CCA TCG AAG

Rhoba-Xho-Stop: \[5’- 3’\]: G ATA TTC CTC GAG CTA GCG GCG TTT C

Template DNA: YidC-Rb in pGZ119EH

Primers for cloning of **RbCT C-His\textsubscript{6}**

Rhoba-Nde: \[5’- 3’\]: GG AAA AAA CGT CAT ATG CAT CCA TCG AAG

Rhoba-Xho-2: \[5’- 3’\]: G ATA TTC CTC GAG GCG GCG TTT C

Template DNA: YidC-Rb in pGZ119EH

Primers for cloning of **OaCT N-His\textsubscript{10}**

Ocean-Nde: \[5’- 3’\]: G GAA AAA CGT CAT ATG CAT CGT CAG GG

Ocean-Xho-Stop: \[5’- 3’\]: GAT ATC CTC GAG CTA TTT CTT CTT GCC
G

Template DNA: YidC-Oa in pGZ119EH
Primers for cloning of **OaCT C-His**

Ocean-**Nde**: [5’- 3’]: G GAA AAA CGT CAT ATG CAT CGT CAG GG

Ocean-**Xho-2**: [5’- 3’]: GAT ATC **CTC GAG** TTT CTT CTT GCC G

Template DNA: YidC-Oa in pGZ119EH

Primers for cloning of **A3CT N-His**

Alb-**Nde**: [5’- 3’]: GAT ATC **CAT ATG** AAT GTA CTT AGT ACC GCC

Alb-**Xho-Stop**: [5’- 3’]: GAT ATC **CTC GAG CTA** TAC AGT GCG TTT CCG

Template DNA: Alb3CT in pET21d provided by I. Sinning, University of Heidelberg

Primers for cloning of **A3CT C-His**

Alb-**Nde**: [5’- 3’]: GAT ATC **CAT ATG** AAT GTA CTT AGT ACC GCC

Alb-**Xho-2**: [5’- 3’]: GAT ATC **CTC GAG** TAC AGT GCG TTT CCG

Template DNA: Alb3CT in pET21d provided by I. Sinning, University of Heidelberg

Primers for cloning of **O1CT N-His**

Oxa1-**Nde**: [5’- 3’]: GAT ATC **CAT ATG** ATT TTG AGA AAC AAA TGG

Oxa1-**Xho-Stop**: [5’- 3’]: GCA TGC **CTC GAG TCA** TTT TTT GTT ATT AAT G

Template DNA: *Saccharomyces cerevisiae* mOxa1 in pUC18

Primers for cloning of **Ffh N-Strep**

Ffh **Nco** Strep **fw**: [5’- 3’]: **CC ATG** GCA AGC TGG AGC CAC CCG CAG TTC GAA AAG GGT TTT GAT AAT TTA ACC GAT CG

Ffh **Eco-Bam** rev: [5’- 3’]: **GGA TCC** GAA **TTC** TTA GCG ACC AGG GAA GCC TGG

Template DNA: Ffh NHis in pGZ119HE
Primers for cloning of **L23, L24 and L29**

Master thesis MSc Emmanuelle Mboubi, 2013; University of Hohenheim

Primers for cloning of **Ffh C-Strep / Ffh C-His\textsubscript{10} Prescision**

Ffh \textit{Nco} fw: \[5`-3`]: CCATG GCA TTT GAT AAT TTA ACC

Ffh \textit{Mun-Strep-Eco} rev: \[5`-3`]: GAA TTC TTA CTT TTC GAA CTG CGG GTG GCT CCA ATT GCC GCG ACC AGG GAA GCC TGG

Template DNA: Ffh N\textsubscript{His} in pGZ119HE

2.5 Peptide synthesis

The N-terminal KdpD-peptides for SRP binding studies via Microscale Thermophoresis (2.17) were synthesized by the Custom peptide synthesis services from GENOSPHERE Biotechnologies (France).

The 27 amino acid residues long peptides were synthesized with N-terminal acetylation and C-terminal amidation and a purity of > 95 %.

The following three synthesized peptides were shipped and stored in lyophilized form at -80 °C:

- \textbf{N22-48}\textsubscript{pep}: N-acetyl RGKLKVFFGACAGVGKTWAMLAEAQL C-amid
- \textbf{W3A}\textsubscript{pep}: N-acetyl RGKLKVFFAAACAAVAKTWAMLAEAQL C-amid
- \textbf{3Q}\textsubscript{pep}: N-acetyl QGQLQVFFGACAGVGKTWAMLAEAQL C-amid

Solution conditions: 3 mg/ml in ddH\textsub{2}O

Molecular Weight: ~ 2 951 Da

\(\varepsilon = 5 500\) approx.
2.6 Methods in molecular biology

2.6.1 Polymerase chain reaction (PCR)

Polymerase chain reaction was used to amplify DNA fragments for cloning in vitro using the GoTaq polymerase (Promega) from *Thermus aquaticus* according to the manufacturer’s protocol.

The cycle of denaturation (95 °C), annealing (X °C) and elongation (72 °C) is repeated for 30 steps. The annealing temperature varies according to the melting temperatures ($T_m$) of the used primer pair (approximately 5°C below the calculated melting point of the primer with the lowest $T_m$). The duration of the elongation step depends on the DNA polymerase and on the length of the DNA fragment to be amplified.

Reaction mixture:

<table>
<thead>
<tr>
<th>Component</th>
<th>Final Volume</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>5X Green GoTaq reaction buffer</td>
<td>10 µl</td>
<td>1X (1.5 mM MgCl$_2$)$^2$</td>
</tr>
<tr>
<td>dNTP Mix (10 mM each)</td>
<td>1 µl</td>
<td>0.2 mM each dNTP</td>
</tr>
<tr>
<td>upstream primer (10 µM)</td>
<td>3 µl</td>
<td>0.6 µM</td>
</tr>
<tr>
<td>downstream primer (10 µM)</td>
<td>3 µl</td>
<td>0.6 µM</td>
</tr>
<tr>
<td>GoTaq Polymerase (5 u/µl)</td>
<td>0.25 µl</td>
<td>1.25 u</td>
</tr>
<tr>
<td>template DNA</td>
<td>X µl</td>
<td>&lt;0.5 µg / 50µl</td>
</tr>
<tr>
<td>nuclease-free water to</td>
<td>50 µl</td>
<td></td>
</tr>
</tbody>
</table>

Thermal cycling conditions:

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
<th>Number of Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>95 °C</td>
<td>2 min</td>
<td>1</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95 °C</td>
<td>30 sec</td>
<td></td>
</tr>
<tr>
<td>Annealing</td>
<td>X °C</td>
<td>30 sec</td>
<td>30</td>
</tr>
<tr>
<td>Elongation</td>
<td>72 °C</td>
<td>1 min/kb</td>
<td></td>
</tr>
<tr>
<td>Final Elongation</td>
<td>72 °C</td>
<td>5 min</td>
<td>1</td>
</tr>
<tr>
<td>Soak</td>
<td>7 °C</td>
<td>∞</td>
<td>1</td>
</tr>
</tbody>
</table>
The PCR products were then separated by agarose gel electrophoresis in the presence of EtBr under UV light. Reactions performed with the 5X Green GoTaq reaction buffer were loaded onto the gel directly after amplification.

2.6.2 Agarose gel electrophoresis

The concentration used for gels in this work was between 0.7 % and 1 % (w/v) agarose (Biozym) in 1X TAE buffer (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA) and 2 µl of an ethidium bromide solution (Sigma, 10 mg/ml). DNA samples were mixed 5:1 with 6X loading dye (0.25 % bromophenol blue, 0.25 % Xylene cyanol, 15 % Pharmacia Ficoll400, 60 mM EDTA) prior to loading into the wells of the gel and run for 60-90 min at a constant voltage of 70 V. DNA bands were visualized at a wavelength of 302 nm and documented using the UV light documentation system Diana v1.6 (raytest, Straubenhardt).

2.6.3 Restriction digest of DNA

For cloning of defined DNA fragments into vectors it is necessary to digest the DNA with restriction endonucleases (Fermentas). These enzymes cleave DNA strands near or at specific recognition sequences and create blunt or sticky ends.

In general, an 20 µl restriction mixture was composed of 18 µl DNA, 1 µl conventional (10 U/µl) or FastDigest restriction enzyme (1 FDU/µl) and 1 µl of the appropriate 10x reaction buffer. The restriction digest was incubated at 37 °C for 30 min using FastDigest enzymes or 2 h using conventional endonucleases.

The restriction reactions were subjected to agarose gel electrophoresis and the DNA fragments of interest were cut out from the gel. For the isolation of DNA from agarose gels, the GeneJET™ Gel Extraction Kit (Thermo Fisher Scientific) was used according to the manufacturer’s protocol.

2.6.4 Ligation of DNA fragments

Ligation of digested DNA fragments after isolation from an agarose gel was performed with T4 DNA ligase (Fermentas). For a standard 10 µl ligation reaction, a threefold molar excess of the insert DNA-fragment was added to the linearized vector and mixed with 1 µl 10x ligation buffer and 1 µl T4 DNA ligase. The ligation reaction was incubated for 2 hours at 22 °C or overnight at 4°C and transformed into chemically competent E. coli XL1 cells (5 µl ligation mixture / 50 µl XL1).
2.6.5 Transformation of chemically competent *E. coli* cells

For transformation of chemically competent *E. coli* cells (2.2) 1 µl of plasmid DNA or 5 µl of ligation mixture (2.6.4) was incubated with 50 µl thawed cells for 20 min on ice. A heat shock at 42 °C for 90 sec allows the DNA to enter the bacterial cells. After incubation for 5 min on ice, 1 ml LB medium was added and incubated for 1 h at 37 °C on a shaker. Prior to centrifugation (2 min; 3000 rpm) 100 µl of the sample was directly plated onto an antibiotic containing LB agar plate. The cell pellet was resuspended in 100 µl LB medium and also plated for overnight growth at 37 °C.

2.6.6 Preparation of chemically (RbCl) competent *E. coli* cells

The rubidium chloride (RbCl) method was used to prepare competent cloning and expression strains of *E. coli*, since this method allows long term storage of the competent cells at -80 °C without significant loss of competence.

100 ml LB medium was inoculated 1:100 from an overnight culture of the respective strain and grown to an OD$_{600}$ of 0.6 at 37 °C. The cells were harvested in a sterile JLA14.500 tube for 15 min at 3000 g and 4 °C. The cell pellet was carefully resuspended in 10 ml ice cold TFB1 buffer (100 mM RbCl, 50 mM MnCl$_2$, 30 mM KAc, 10 mM CaCl$_2$, 15 % glycerol; pH 5.8) on ice and incubated for 90 min. After a second centrifugation (15 min, 3000 g, 4 °C) in a sterile JA25.50 tube, the cell pellet was resuspended in 2,5 ml ice cold TFB2 buffer (10 mM MOPS, 10 mM RbCl, 75 mM CaCl$_2$, 15 % glycerol; pH 8). The competent cells were aliquoted and shock frosted in liquid nitrogen for storage at -80 °C.

2.6.7 Isolation of plasmid DNA from bacterial cultures

The Zyppy™ Plasmid Miniprep Kit (Zymo Research) was used to isolate the plasmid DNA according to the manufacturer’s protocol with a few modifications.

2 ml LB medium containing the appropriate antibiotics was inoculated with a single colony picked from the transformation plate and incubated overnight at 37 °C. The cells were pelleted (1 min, 13 000 rpm), resuspended by vortexing in 600 µl 1x TE buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA) and mixed carefully with lysis buffer and neutralization buffer. After two washing steps, a centrifugation step of the empty column was performed to remove the ethanol contained in the second wash buffer completely from the column. The DNA was eluted by 50 µl preheated ddH$_2$O (56 °C).
2.6.8 DNA sequencing (Sanger et al., 1977)

The Thermo Sequenase™ Cycle Sequencing kit from USB was used to sequence DNA by a modified dideoxy chain termination method according to Sanger with fluorescently labeled primers.

The sample was divided into four separate amplification reactions, containing all of the deoxynucleotides (dNTPs). Additionally each reaction contained one of the four dideoxynucleotides (ddATP, ddCTP, ddGTP, ddTTP) which causes the chain termination when incorporated into the extending DNA strand (termination mix). This resulted in a mixture of different elongated DNA fragments which were separated by polyacrylamide-urea gel electrophoresis and detected by an infrared laser at 700 nm or 800 nm, respectively (primer dependent).

Four PCR tubes were prepared by adding 4 µl of the respective termination mix and 4 µl of the master mix (13 µl template DNA, 2 µl reaction buffer, 1 µl primer [2 pMol], 1,5 µl thermosequenase [4 U/µl]).

**Thermal cycling conditions:**

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
<th>Number of Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>95 °C</td>
<td>2 min</td>
<td>1</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95 °C</td>
<td>30 sec</td>
<td></td>
</tr>
<tr>
<td>Annealing</td>
<td>X °C</td>
<td>30 sec</td>
<td>40</td>
</tr>
<tr>
<td>Elongation</td>
<td>72 °C</td>
<td>1 min</td>
<td></td>
</tr>
<tr>
<td>Final Elongation</td>
<td>72 °C</td>
<td>3 min</td>
<td>1</td>
</tr>
<tr>
<td>Soak</td>
<td>7 °C</td>
<td>∞</td>
<td>1</td>
</tr>
</tbody>
</table>

The sequencing reaction was terminated by adding 3 µl of stopping buffer (95 % formamide, 10 mM EDTA, 0,1 % Basic Fuchsins, 0,01 % Bromophenol Blue, pH 9). The samples were heated to 72 °C for 2 min and 1,7 µl of each reaction was loaded onto the sequencing gel (1X TBE [89 mM Tris-HCl pH 8, 89 mM boric acid, 2 mM EDTA], 6 % acrylamide [Long Ranger], 7 M Urea, 0,05 % APS, 0,05 % TEMED).

2.6.9 Site-directed mutagenesis (Quikchange)

The Quikchange mutagenesis is a method that allows the introduction of site-specific mutations into plasmid DNA. Two complementary primers, containing the desired
mutation in the middle, are used to linearly amplify the entire plasmid with a non-strand-displacing DNA polymerase (Pfu-Ultra II, Stratagene).

This method was used to create the single cysteine mutant of Ffh (2.3 plasmid B14) with the following primers (the respective complementary primer is not shown):

- **Ffh C406S**: removal of the wildtype cysteine residue at position 406
  
  5` CGT ATT GCT GCC GGT TCC GGT ATG CAG GTG CAG G 3`

- **Ffh M423C**: introduction of a cysteine residue at position 423
  
  5` CAG TTC GAC GAC TGC CAG CGC ATG ATG 3`

To mutate both amino acid residues at once, a double mutagenesis with both primer pairs Ffh C406S and Ffh M423C in one PCR reaction, was performed.

**Reaction mixture:**

<table>
<thead>
<tr>
<th>Component</th>
<th>Final Volume</th>
<th>Final Concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pfu Ultra II buffer (10x)</td>
<td>5 µl</td>
<td>1X</td>
</tr>
<tr>
<td>dNTP Mix (10 mM each)</td>
<td>1 µl</td>
<td>0,2 mM each dNTP</td>
</tr>
<tr>
<td>each primer (10 µM)</td>
<td>0,5 µl</td>
<td>0,1 µM</td>
</tr>
<tr>
<td>Pfu Ultra II Polymerase</td>
<td>1 µl</td>
<td></td>
</tr>
<tr>
<td>template DNA X µl</td>
<td></td>
<td>200 ng / 50 µl</td>
</tr>
<tr>
<td>Nuclease-Free Water to 50 µl</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Thermal cycling conditions:**

<table>
<thead>
<tr>
<th>Step</th>
<th>Temperature</th>
<th>Time</th>
<th>Number of Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial Denaturation</td>
<td>95 °C</td>
<td>3 min</td>
<td>1</td>
</tr>
<tr>
<td>Denaturation</td>
<td>95 °C</td>
<td>1 min</td>
<td></td>
</tr>
<tr>
<td>Annealing</td>
<td>56 °C</td>
<td>1 min</td>
<td>20</td>
</tr>
<tr>
<td>Elongation</td>
<td>68 °C</td>
<td>15 min</td>
<td></td>
</tr>
<tr>
<td>Final Elongation</td>
<td>68 °C</td>
<td>10 min</td>
<td>1</td>
</tr>
<tr>
<td>Soak</td>
<td>12 °C</td>
<td>∞</td>
<td>1</td>
</tr>
</tbody>
</table>
The methylated template plasmid DNA was digested by adding 1 µl DpnI and incubating for 2 h or overnight at 37 °C. Immediately, 10 µl of the reaction were transformed into 100 µl competent XL-1 cells, while the rest the sample was precipitated with EtOH (2.6.10).

2.6.10 EtOH precipitation of plasmid DNA

The DNA sample was transferred into a 1,5 ml Eppendorf tube and 1/10 Vol. 3M Na acetate pH 5.2 and 2,5 Vol. EtOH (100 %) were added. The precipitate was incubated at -20 °C for at least 1 h or overnight, centrifuged for 20 min at 17 000 g, 4 °C and dried in the speedvac for 30 min. The precipitated DNA was solved in 10 µl 1x TE (10 mM Tris-HCl pH 7.5, 1 mM EDTA) and transformed into 100 µl competent XL-1 cells.

2.6.11 Homologous recombination to generate the E. coli strain MC-Ffh

The Ffh depletion strain MC-Ffh (2.2.6) was derived from E. coli strain MC1061 (Casadaban and Cohen, 1980) by homologous recombination of the araC-araBAD operator region together with a kanamycin cassette as a selective marker into the chromosome upstream of the ffh gene (Datsenko and Wanner, 2000).

1. Preparation of P1wt-phage lysate

For the preparation of the P1wt-lysate an overnight culture of MC 4100 was used to inoculate 2,5 ml LB medium 1:100. The culture was grown to an OD600 of 0.4 then 10 mM CaCl2 and 10 µl P1wt (P1vir) lysate were added. Continuing growth for about 2,5 h at 37 °C led to the complete lysis of the cells. To isolate the phages the lysed culture was mixed with a few drops of chloroform and stood still for 10 min until the chloroform separated. 2 ml from the supernatant was transferred to an Eppendorf tube and centrifuged for 5 min at 17 000 g. The supernatant contained the P1wt –lysate and was transferred to a new Eppendorf tube. The lysate was stored at 4 °C.

2. Amplification of the kan-ara cassette for promoter exchange of ffh

To integrate the araC-araBAD operator region into the chromosomal ffh promoter region of the E. coli strain DY330 by homologous recombination, a PCR reaction was performed to amplify the kan-ara cassette with primers containing flanking regions of the upstream ffh promoter sequence and the start region of the ffh gene.

The ffh gene is monocistronic so that no other genes are affected by the promoter exchange (Fig. 2.1).
The following primers were used for the amplification of the **kan-ara** cassette:

**ffh-prom (fw):** $T_m$ match $66 \, ^\circ C$

```
5' gct cca acc gtt gca taa caa caa aga ggc ctc CCA TAT GAA TAT CCT CCT TAG TTC C 3'
```

**ffh-start (rev):** $T_m$ match $64 \, ^\circ C$

```
5' ggc tgc ggc aac aac gat cgg tta aat tat caa aca tGG TGA ATT CCT CCT GCT AGC 3'
```

The PCR reaction was performed according to the standard protocol (2.6.1) with additional 1 mM MgCl$_2$. The annealing temperature was set to 62 °C and as template DNA an isolated **kan-ara-cassette** fragment from IY26 was used. The expected PCR fragment had a size of 2500 bp, including the **kan-ara-cassette** (~2400 bp) with flanking regions of the upstream promoter sequence and the start region of **ffh** (2.2).
3. Preparation of electro competent DY330 cells

The linear ffh-kan-ara fragment was transformed by electroporation into the DY330 strain. This strain allows temperature induced (42 °C) expression of the Rec proteins, which are necessary for homologous recombination. The electro competent DY330 cells were always prepared freshly before use.

100 ml LB were inoculated 1:100 with an overnight culture of DY330 and grown at 30 °C (no expression of Rec proteins) to an OD of 0.8. The culture was divided into halves, one was grown at 30 °C and the other one at 42 °C (expression of Rec proteins) for 15 min. The cells were cooled down on ice for 15 min and harvested in sterile JA 25.50 tubes for 8 min at 7 000 rpm, 4 °C. After two washing steps with 4 ml ddH2O and one step with 1,5 ml ddH2O in an Eppendorf tube, the cell pellets were resuspended in 400 µl ddH2O and immediately used for electroporation. It is possible to store the electro competent cells at -80 °C in a 20% glycerol solution, but then the transformation efficiency is strongly reduced.

4. Electroporation of DY330 with the ffh-kan-ara cassette

For the transformation of the linear DNA fragment 40 µl electro competent DY330 cells (42 °C, rec+) were incubated with 2 µl ffh-kan-ara cassette (Fig. 2.2) for 1 min on ice. As a negative control the same was performed with DY330 cells grown at 30 °C (rec -). The approaches were transferred into an Epo-cuvette and pulsed according to the manufacturer's protocol. Immediately after electroporation, 1ml SOC-Ara media (2,5 % tryptone, 0,5 % yeast, 10 mM NaCl, 2,5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 0,2% arabinose) was given to the cells, then transferred into a sterile plastic tube and incubated at 42 °C for 15 min in the shaker. Subsequently, the cells were grown for 1,5 more hours at 30 °C and then plated on LB-kanamycin (25 µg/ml) + arabinose (0,2 %) plates. The candidates were patched on LB-kan-arabinose and LB-kan-glucose (0,4 %) plates, respectively. Colonies which showed growth on arabinose plates, but not on glucose plates have the ffh-kan-ara cassette integrated in their chromosome by homologous recombination and were used to inoculate 2 ml LB-kan-ara media for the following P1-DY330 kan ara-lysate preparation.

5. Preparation of the P1DY330 kan ara lysate

The overnight cultures of each candidate were used to inoculate 2,5 ml LB-kan-ara media and were grown to an OD of 0.3 at 30 °C. At this point 10 mM CaCl₂ and 20 µl P1wt-lysate were added and incubated on the shaker for 3 h at 30 °C. To isolate the phages the lysed culture was mixed with 50 – 100 µl chloroform and stood on ice for 5 min until the
chloroform separated. 2 ml from the supernatant was transferred to an Eppendorf tube and centrifuged for 2 min at 16 000 g. The supernatant contained the P1<sub>DY330 kan ara</sub> lysate and was transferred to a new Eppendorf tube. The lysate was used to transduce the target strain MC1061.

6. Transduction of MC1061 with the P1<sub>DY330 kan ara</sub> lysate

An overnight culture of the recipient strain (MC1061) in 2 ml LB was used to inoculate (1:100) 4 ml LB media. The cells were grown at 37 °C to an OD<sub>600</sub> of 0.8 and 10 mM CaCl<sub>2</sub> was added. Three reactions were set up, each consisting of 1 ml recipient mixture and 100 µl P1<sub>DY330 kan ara</sub> lysate, mixed and incubated for 20 min, 30 min and 40 min at room temperature, respectively. To each reaction 100 µl 1 M sodium citrate was added to avoid additional phage adsorption, incubated for 1 h at 37 °C in the shaker and plated on LB plates with kanamycin, arabinose and sodium citrate. After an incubation over night at 37 °C the colonies were plated for another three times on LB-kan-ara-Na-citrate plates to remove the P1 phages completely. Finally, single colonies were patched on LB-ara and LB-glc plates and candidates who showed only growth on the arabinose plates were used for chromosomal control PCRs.

7. Control PCRs of the MC-Ffh candidates

To analyze whether the kan-ara cassette was correctly integrated into the chromosomal region downstream of <i>ffh</i>, four control PCRs were performed. Therefore chromosomal DNA of the respective MC-Ffh candidates was isolated and used as template for the PCR reaction with the following primers:

**PCR 1: ypjD-stop / ffh-rev**

- ypjD-stop: 5` GGA CAA TTC GGC TGC CGA AGT AGG 3`
- ffh-rev: 5` CGG TGC CAA AAT GGC AAA CAA GCC 3`

expected fragment length: 4622 bp

**PCR 2: kan / ffh rev**

- kan: 5` CCC AAT AGC AGC CAG TCC CTT CC 3`

expected fragment length: 3030 bp

**PCR 3: araC / ffh rev**

- araC: 5` GTA CCC GAT TAT CCA TCG GTG GAT GG 3`
expected fragment length: 2241 bp

PCR 4: ffh-fw / ffh-rev

ffh-fw: 5’ GGC CAA CCG TTT CCA CCC CAG G 3’

expected fragment length: no fragment

ffh-fw binds to the wild type promoter region of ffh which should be deleted in the correct candidate

As a negative control, the same PCR reactions were performed with chromosomal DNA from the MC1061 parent strain. The expected fragment sizes were 2154 bp (ypjD-stop / ffh-rev), 1362 bp (ffh-fw / ffh-rev) and no fragments for the kan and araC / ffh-rev PCRs.

Fig. 2.3 Overview of the chromosomal control PCRs. Four control PCRs were performed to check the correct localization of the kan-ara-cassette downstream of ffh. Each reaction contained the ffh-rev primer and one of the primers ypjD-stop, kan, araC and ffh-fw, respectively. The PCR reaction with ffh-fw should not give a PCR fragment, since the respective binding site is lost during the homologous recombination.

The depletion of Ffh in the MC-Ffh candidates that showed correct PCR fragments was tested via an expression test in LB-ara and LB-glc media and the Ffh level was detected with an antibody against the Ffh protein. The strain has to be grown for at least 3 h in glucose media to deplete Ffh completely. The expression of a plasmid encoded ffh rescued the lethal phenotype of MC-Ffh on LB-glc-IPTG plates in the complementation assay (2.7).
2.6.12 In vitro transcription

To synthesize the 4.5S RNA of the *E. coli* SRP the Thermo Scientific™ TranscriptAid™ T7 High Yield Transcription Kit was used.

1. DNA template preparation for in vitro transcription

Linearized plasmid DNA can be used as template for transcription if it contains a double-stranded RNA polymerase promoter region in the correct orientation. For the transcription of the 4.5S RNA the pUC18 plasmid (2.3 E1) was used with the integrated consensus promoter sequence for the T7 RNA polymerase (TAATACGACTCACTATAGGG). G will be the first base (+1) of the RNA transcript. The 4.5S RNA sequence was placed downstream of the promoter to synthesize sense RNA. To produce RNA transcripts of a defined length, the plasmid DNA has to be linearized by restriction digestion downstream of the insert prior to in vitro transcription. The pUC18-4.5S RNA plasmid was linearized by BamHI and gel purified with a DNA Gel Extraction Kit (GeneJet™, Thermo Fisher Scientific).

2. Phenol / chloroform extraction

The template DNA should have an A260/280 ratio of 1.8 - 2.0. Otherwise it is recommended to purify the DNA template by phenol/chloroform extraction using the following protocol:

- Add 1/10 volume of 3 M Sodium Acetate Solution to the DNA.
- Mix thoroughly.
- Extract with an equal volume of a 1:1 phenol/chloroform mixture, and then twice with an equal volume of chloroform. Collect the aqueous phase and transfer to a new tube.
- Precipitate the DNA by adding 2 volumes of ethanol. Incubate at – 20 °C for at least 30 min and collect the pellet by centrifugation.
- Remove the supernatant and rinse the pellet with 500 µl of cold 70 % ethanol.
- Resuspend the DNA in 20 µl of DEPC-treated Water (0,1 % (v/v) DEPC incubated overnight and autoclaved).
3. **In vitro transcription protocol**

The following reaction components were combined at room temperature in the given order:

<table>
<thead>
<tr>
<th>Component</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEPC-treated water</td>
<td>to 20 µl</td>
</tr>
<tr>
<td>5X TranscriptionAid Reaction Buffer</td>
<td>4 µl</td>
</tr>
<tr>
<td>ATP/CTP/GTP/UTP mix</td>
<td>8 µl</td>
</tr>
<tr>
<td>Template DNA</td>
<td>1 µg</td>
</tr>
<tr>
<td>TranscriptAid Enzyme Mix</td>
<td>2 µl</td>
</tr>
</tbody>
</table>

The reaction was mixed thoroughly, spun briefly to collect all drops and incubated at 37 °C for 3 h. Immediately after the *in vitro* transcription the RNA was purified by the RNA Clean & Concentrator™-25 Kit from Zymo Research and analyzed on a 2 % native agarose gel in 1X Tris-borate-EDTA buffer from Sigma Aldrich. For sample loading, 0,5 – 1 µg RNA were mixed with 2-fold RNA loading dye and incubated for 10 min at 70 °C and cooled down on ice for 3 min prior to loading. 0,5 mg/ml ethidium bromide was added to the gel and the running buffer, respectively. The gel was run for 2.5 h at a constant voltage of 40 V.

2.7 **in vivo complementation assay**

To analyze the functionality of the YidC homologs and the chimera (2.3.1), *in vivo* complementation assays in the YidC depletion strain MK6S (2.2.5) were performed. *E. coli* MK6S cells bearing the respective pGZ119EH plasmids expressing the protein of interest (2.3.1) were grown in LB medium with 0.2 % arabinose and 0.4 % glucose to an OD₆₀₀ of 1.0. Cells were washed once with LB, serially diluted in LB (10⁻², 10⁻⁴, 10⁻⁶) and spotted (5 µl) on LB plates containing 0.2 % arabinose to allow chromosomal YidC expression or 0.4 % glucose to repress chromosomal YidC expression. In the presence of 1 mM IPTG, the plasmid-encoded YidC proteins were expressed. The plates were incubated overnight at 37 °C.

2.8 **Fluorescence microscopy – GFP localization**

For the localization studies, an MscL-GFP fusion protein (2.3.4) was expressed from a pMS119 derivative. The strains IY26 (2.2.7) and MC-Ffh (2.2.6) were transformed with the MscL-GFP encoding plasmid and pGZ119EH derivatives encoding the YidC homologs or
YidC chimera (2.3.3), respectively. Strains were grown overnight at 37 °C in LB medium with 0.2 % arabinose and 0.4 % glucose, washed twice in LB and diluted 1/150 into fresh LB medium with arabinose (0.2 %) to allow normal membrane insertion or in glucose (0.4 %) to deplete FtsY or Ffh, respectively. Cells were grown at least 3 h to an OD$_{600}$ of 0.6 and protein expression was then induced with 1 mM IPTG for 2 h at 30 °C. Cells were collected by centrifugation, washed twice in LB medium, resuspended in 2 mM EDTA, 50 mM Tris-HCl pH 8.0 and incubated overnight at 4 °C. The cell suspension (corresponding to 50 µl cells) was applied to a polylysine-coated slide (Sigma-Aldrich) and examined by a Zeiss AXIO Imager M1 fluorescence microscope. Emission was detected with a filter set specific for GFP.

### 2.9 AMS derivatization – pulse chase & immuno-precipitation

In vivo AMS derivatization studies were used to analyze the proper insertion and topology of the MscL protein during the rescue experiments without a functional SRP system, while coexpressing the YidC proteins.

#### 2.9.1 Depletion of SRP-pathway components

For the AMS (4 – acetamido- 4’- maleimidylstilbene- 2,2’- disulfonylic acid, sodium salt) derivatization studies, *E. coli* MC-Ffh (Ffh depletion) and IY26 (FtsY depletion) were transformed with pMS-MscL168C and one of the pGZ-YidC plasmids (2.3.3), respectively. The MscL mutant and the YidC proteins were coexpressed by induction of IPTG. To deplete cells of Ffh or FtsY, the cells of overnight cultures were washed twice in LB to remove the arabinose and back diluted 1/150 into fresh LB medium containing 0.4 % glucose (w/v). Cells were grown for at least 3 h in glucose to an OD$_{600}$ of 0.5. 1.2 ml cells were washed 2 times (3 min, 6000 x g) in 1 ml 1x M9-salt medium (34 mM Na$_2$HPO$_4$, 22 mM KH$_2$PO$_4$, 18.7 mM NH$_4$Cl, 8.6 mM NaCl) and resuspended in 1 ml M9-minimal medium (1x M9-salt medium, 1 mM MgSO$_4$, 0.1 mM CaCl$_2$, 5 µg/ml Thiamin, 0.0005 % iron (II) citrate, 0.1 mM amino acid mix (19 aa -met)) supplemented with 0.2 % arabinose or 0.2 % glucose, respectively. For the AMS control (+/- AMS) the sample was divided in two 500 µl portions and incubated for another 60 min at 37 °C prior to pulse chase labelling.

#### 2.9.2 Pulse chase

Cells coexpressing the MscL mutant and the YidC proteins were induced for 10 min with 1 mM IPTG. Cells were radiolabeled with [³⁵S] Met (10 µCi ml$^{-1}$ culture) for 2 min. After labelling, cells were incubated with AMS (2.5 mM final concentration; Molecular Probes)
and non-radioactive L-methionine was added (final concentration, 500 µg/ml) for 10 min and then combined with 20 mM DTT for 10 min to quench the AMS reaction.

### 2.9.3 Immuno-precipitation

After quenching, samples were acid-precipitated and incubated overnight at 4 °C. Prior to immuno-precipitation the samples were washed with 1 ml acetone (10 min, 14 000 rpm, 4 °C) and the dried pellets (5 min, 95°C) were resuspended in 10 mM Tris pH 8, 2 % SDS (5 min, 95 °C shaker). The samples were mixed with 1 ml TEN-TX buffer (10 mM Tris pH 8, 1 mM EDTA, 150 mM NaCl, 2 % Triton X-100) and 15 µl Staph A (Staphylococcus aureus Protein A) and incubated for 1 h at 4 °C on a rotary wheel. After incubation the Staph A matrix was pelleted (10 min, 14 000 rpm, 4 °C) and the supernatant transferred to a new tube, mixed with 2 µl His antibody (monoclonal from Sigma, 1:10 dilution) and incubated overnight on a rotary wheel at 4 °C. Then 20 µl Staph A, binding the antigen-antibody complex, were added and incubated for another hour on the rotary wheel. The Staph A complex was washed two times (30 s, 8000 rpm, 4 °C) with 1 ml TEN-TX buffer and once (2 min, 11 000 rpm) with 1ml TEN buffer (10 mM Tris pH 8, 1 mM EDTA, 150 mM NaCl). The supernatant was discarded and the Staph A pellet resuspended in 50 µl 1x SDS sample buffer (5 parts solution 1 [200 mM Tris base, 20 mM EDTA] + 4 parts solution 2 [83,3 mM Tris base, 8,3 % SDS, 29 % glycerol, 0,15 % bromophenol blue] + 1 part 1 M DTT) by vortexing and boiling for 5 min at 95 °C.

### 2.9.4 Phosphorimaging and quantification

The samples were then analyzed in 14 % SDS-polyacrylamide gels. Therefor the samples were centrifuged immediately before loading for 2 min at 14 000 rpm and 30 µl of the supernatant was loaded on the SDS-polyacrylamide gel. For a good resolution of the AMS shift, the gel was run at 22 mA for 3,5 h. Then the gel was retained in a fixing solution (40 % MeOH, 7 % acetic acid) for 30 min, dried on a Whatman filter paper for 2 h at 80 °C and finally examined by phosphorimaging (Dürr Medical CR 35 bio). The imaging plate (Fujifilm BAS-IP) was incubated between 1 – 3 days and the samples were quantified with ImageJ (Schneider et al., 2012).
2.10 Protein purification

2.10.1 Purification of *E. coli* YidC_{6His10}

*Recombinant protein expression & membrane preparation*

*E. coli* C43 cells (2.2.4) were transformed (2.6.5) with pET16b-EcYidC (2.3.2, B1) and grown in 6 L LB containing 100 µg/ml ampicillin at 37 °C. Induction was initiated at an OD_{600} of 0.6 with 0.5 mM IPTG for 3 h. The cells were pelleted, resuspended in 2 ml/g cells buffer A_{EY} (20 mM Tris pH 7.5, 300 mM NaCl, 10 % glycerol) and lysed using a One Shot cell disruption instrument (Constant Systems LTD) at 1.3 kbar. Before cell disruption 1 mM DTT, 0.1 mM PMSF (phenylmethanesulfonyl fluoride) and 1 mM EDTA were added. The lysate was centrifuged for 12 min at 24 000 g to spin down the unbroken cells and centrifuged a second time under the same conditions to clear the supernatant from remaining cell debris. The membranes were pelletized by centrifugation for 1 h at 140 000 g, resuspended and homogenized in 20 ml buffer A_{EY} and centrifuged a second time. Membranes containing the YidC protein were homogenized in 25 ml buffer A_{EY} and extracted by the addition of 1 % n-dodecyl β-D maltoside (DDM; 10 % w/v stock solution) overnight at 4 °C on a rotary wheel.

*Immobilized metal ion affinity chromatography (IMAC)*

The YidC protein was purified by Ni chelating chromatography. To this end, the approach was filled up to 50 ml with buffer A_{EY} including the addition of 20 mM imidazole, 0.1 mM PMSF and 2 ml cV Ni-NTA agarose (Quiagen). After 2 h incubation at 4 °C on a rotary wheel, the batch was filled into an empty 10 ml column (90 µm filter, MoBiTec) by gravity flow and the matrix was washed with 20 ml buffer A_{EY} + 30 mM imidazole and 0.05 % DDM. The protein was eluted from the matrix with 20 ml buffer A_{EY} + 300 mM imidazole + 0.05 % DDM (10 x 2 ml fractions). The elution fractions (10 µl sample + 5 µl 5x sample buffer [4 parts solution 2 + 1 part 1M DTT]) were analyzed on 12 % SDS acrylamide gels and visualized by Coomassie staining (0.4 % Coomassie R250 [Serva], 45 % methanol puriss. [Sigma-Aldrich], 10 % acetic acid puriss. [Sigma-Aldrich]).

*Size exclusion chromatography (SEC)*

The main YidC-containing IMAC-fractions were further purified on a Superdex 200 16/60 column (GE Healthcare) with the Äkta-purifier System (GE Healthcare). The protein fractions from the IMAC were concentrated to 5 ml with Amicon® Ultra 15 ml Filters (Merck Millipore) and filtered prior to loading (Ultrafree® Durapore – PVDF 0.22 µm, Merck Millipore). The size exclusion chromatography (SEC) of *E. coli* YidC was performed
at room temperature in buffer GF_EY (20 mM Tris pH 7.5, 300 mM NaCl, 10 % glycerol, 0.05 % DDM, 1 mM DTT) at a flow rate of 0.5 ml / min. 2 ml fractions were collected and the peak fractions were analyzed on 12 % SDS acrylamide gels (10 µl sample) and visualized by Coomassie staining. The purified protein was stored at 4 °C. Typical YidC protein concentrations were ~ 0.5 mg/ml (Fig. 3.5 B).

2.10.2 Purification of YidC-Rb

Recombinant protein expression & inner membrane vesicle (IMV) preparation

E. coli C43 cells (2.2.4) were transformed (2.6.5) with pET16b-YidC-Rb (2.3.2, B3) and grown in 6 L LB containing 100 µg/ml ampicillin at 37 °C. Induction was initiated at an OD_600 of 0.6 by 0.5 mM IPTG for 2 h. The cells were pelleted, resuspended in 2 ml/g cells buffer A_ER (50 mM Tris pH 7.5, 10 % sucrose, 1 mM EDTA) and lysed using a One Shot cell disruption instrument (Constant Systems LTD) at 1,3 kbar. Before cell disruption, 0,1 mM PMSF (phenylmethanesulfonyl fluoride) was added. The lysate was centrifuged for 15 min at 10 000 g to spin down the unbroken cells and cell debris. The membranes were pelleted by centrifugation for 2 h at 140 000 g, resuspended and homogenized in 5 ml buffer A_ER. To isolate inner membrane vesicles containing the YidC-Rb protein, the resuspended membrane pellet was loaded onto a three step sucrose gradient (1 – 1.5 ml crude membrane suspension / gradient). The sucrose gradient consists of 3 sucrose steps (770 mM, 1440 mM and 2020 mM sucrose) in 20 mM Tris pH 7.5 and 1 mM EDTA. First, 12 ml of the 770 mM sucrose solution were filled into a SW28 centrifuge tube (Polyallomer 25 x 89 mm, Beckman) then 12 ml of the 1440 mM solution and finally 10 ml of the 2020 mM solution were underlayed. Prior to loading the gradient was incubated for 1 h at 4 °C. After centrifugation (SW28 rotor; 113 000 g, 16 h, 4 °C), the inner membrane vesicles accumulate at the boundary between 770 mM and 1140 mM sucrose. The IMV bands were pooled, diluted 1:5 with 50 mM Tris pH 7.5 and centrifuged for 2 h at 180 000 g. The IMV pellet was resuspended in 5 ml storage buffer (50 mM Tris pH 7.5, 10 % glycerol, 0,1 mM PMSF) and stored at - 80 °C or directly solubilized over night at 4 °C on a rotary wheel. For the solubilization 1,5 % LDAO (lauryldimethylamine-oxide), 500 mM NaCl, 1 mM DTT were added and filled-up to 10 ml with storage buffer. After the solubilization, the sample was centrifuged for 1 h at 100 000 g to remove aggregated or non-solubilized material.

Immobilized metal ion affinity chromatography (IMAC)

For the IMAC, the supernatant from the LDAO-solubilization (= load) was filled-up to 25 ml with buffer B_ER (50 mM Tris pH 9, 400 mM NaCl, 5 % glycerol) and 20 mM imidazole + 1
mM DTT were added. 2 ml (cV) NiNTA agarose (Quiagen) were equilibrated in buffer B<sub>ER</sub> with 20 mM imidazole, 1 mM DTT and 0.1 % LDAO and combined with the diluted load (= batch). The batch was incubated for 2 h at 4 °C on a rotary wheel and then filled into an empty 10 ml column (90 µm filter, MoBiTec) by gravity flow. The matrix was washed with 35 ml buffer W<sub>ER</sub> (50 mM Tris pH 9, 500 mM NaCl, 15 % glycerol, 50 mM imidazole, 1 mM DTT, 0.1 % LDAO) and the protein was eluted from the matrix with 20 ml buffer E<sub>ER</sub> (50 mM Tris pH 7.5, 300 mM NaCl, 5 % glycerol, 500 mM imidazole, 1 mM DTT, 0.2 % LDAO) in 10 x 2 ml fractions. The elution fractions (10 µl sample + 5 µl 5x sample buffer [4 parts solution 2 + 1 part 1M DTT]) were analyzed on 12 % SDS acrylamide gels and visualized by Coomassie staining (0.4 % Coomassie R250 [Serva], 45 % methanol puriss. [Sigma-Aldrich], 10 % acetic acid puriss. [Sigma-Aldrich]).

Size exclusion chromatography (SEC)

The main fractions from the IMAC were further purified on a Superdex 200 16/60 column (GE Healthcare) with the Äkta-purifier System (GE Healthcare). The protein fractions were concentrated to 5 ml with Amicon® Ultra 15 ml Filters (Merck Millipore) and filtered prior to loading (Ultrafree® Durapore – PVDF 0.22 µm, Merck Millipore). The size exclusion chromatography (SEC) of YidC-Rb was performed at room temperature in buffer G<sub>ER</sub> (50 mM Tris pH 7.5, 300 mM NaCl, 5 % glycerol, 0.2 % LDAO, 1 mM DTT) at a flow rate of 0.5 ml / min. 2 ml fractions were collected and the peak fractions were analyzed on 12 % SDS acrylamide gels (10 µl sample) and visualized by Coomassie staining. The purified protein was stored at 4 °C. Typical protein concentrations were 1.5 mg/ml (Fig. 3.7 B).

The YidC homolog from Oceanicaulis alexandrii and the respective hybrid protein YidC-Oa (2.3.2 B2 & B4) were purified as described for YidC-Rb with typical protein concentrations of ~ 0.5 mg/ml.

2.10.3 Purification of E. coli Ffh & Ffh C406S / M423C

Recombinant protein expression

Both, the wild type Ffh and the single cysteine mutant C406S / M423C were expressed in BL21 cells (2.2.2) and purified via a Strep-tag at the C-terminus of the protein (2.3.2, B12 & B14). The expression was done in 4 L LB medium by induction at an OD<sub>600</sub> of 0.6 with 0.5 mM IPTG for 3 h at 37°C. The cells were pelleted, resuspended in 60 ml buffer A<sub>Ffh</sub> (20 mM Hepes pH 8, 350 mM NaCl, 10 mM MgCl<sub>2</sub>, 10 mM KCl, 10 % glycerol) and lysed using a One Shot cell disruptor (Constant Systems LTD) at 1.3 kbar. Before cell disruption, 0.1 mM PMSF (phenylmethanesulfonyl fluoride) was added. The lysate was centrifuged
for 30 min at 20 000 g for two times to spin down the unbroken cells and cell debris. The supernatant, containing the soluble Ffh protein, was stored at - 80 °C.

**Strep-tag affinity chromatography**

The Strep-tag® II is a short peptide (8 amino acids, WSHPQFEK), which binds to Strep-Tactin® with a high selectivity. In this work, the Strep-Tactin matrix, an engineered streptavidin, from IBA was used.

The supernatant was loaded onto 4 ml Strep-Tactin matrix in a 10 ml column by gravity flow directly. The matrix was washed with 50 ml buffer W\textsubscript{Ffh} (20 mM Hapes pH 8, 500 mM NaCl, 10 mM MgCl\textsubscript{2}, 100 mM KCl) and the protein was eluted from the matrix with 20 ml buffer E\textsubscript{Ffh} (20 mM Hapes pH 8, 350 mM NaCl, 10 mM MgCl\textsubscript{2}, 10 mM KCl, 10 % glycerol, 2.5 mM desthiobiotin [Sigma]) in 10 x 2 ml fractions. The elution fractions (10 µl sample + 5 µl 5x sample buffer) were analyzed on 12 % SDS acrylamide gels and visualized by Coomassie staining. After the elution of the Ffh protein, the matrix was washed 3 times with 5 cV regeneration buffer (IBA, 100 mM Tris·HCl, 150 mM NaCl, 1 mM EDTA, 1 mM HABA (hydroxy-azophenyl-benzoic acid, pH 8.0) to displace desthiobiotin from the binding pocket. The regeneration buffer was removed from the matrix by extensive washing with buffer W (IBA, 100 mM Tris·HCl, 150 mM NaCl, 1 mM EDTA, pH 8.0). Once the red color of the regeneration buffer has disappeared, the column can be reused and was stored in buffer W at 4 °C.

**Size exclusion chromatography (SEC)**

The main Strep-tag elution fractions were further purified on a Superdex 75 16/60 column (GE Healthcare) with the Äkta-purifier System (GE Healthcare). The protein fractions were concentrated up to 5 ml with Amicon® Ultra 15 ml Filters (Merck Millipore) and filtered prior to loading (Ultrafree® Durapore – PVDF 0.22 µm, Merck Millipore). The size exclusion chromatography (SEC) of Ffh was performed at room temperature in buffer GF\textsubscript{Ffh} (20 mM Hapes pH 8, 200 mM NaCl, 10 mM MgCl\textsubscript{2}, 10 mM KCl, 10 % glycerol) at a flow rate of 0,5 ml / min. 2 ml fractions were collected and the peak fractions were analyzed on 12 % SDS acrylamide gels (10 µl sample) and visualized by Coomassie staining. The purified protein was stored at - 80 °C with typical concentrations of ~ 1 mg/ml without further concentration (Fig. 3.31 A).
2.10.4 Purification of ribosomal proteins L24 and L29

**Recombinant protein expression**

The ribosomal proteins L24 and L29 were expressed from pET16b vectors (2.3.2, B10 & B11) in BL21 cells and purified via a Strep-tag at the N-terminus of the protein. The expression was performed in 2 L LB media and the cells were induced at an OD_{600} of 0.6 with 0.5 mM IPTG for 3 h at 30°C. The cells were pelleted, resuspended in 2 ml/g buffer ARib (100 mM Tris pH 8, 150 mM NaCl, 1mM EDTA, 10 % glycerol) and lysed using an One Shot cell disruptor (Constant Systems LTD) at 1,3 kbar. Before cell disruption, 0.1 mM PMSF was added. The lysate was centrifuged for 20 min at 10 000 g to spin down the unbroken cells and cell debris. Then the supernatant was centrifuged for 1 h at 150 000 g to pellet the membranes. The cleared supernatant, containing the soluble ribosomal proteins, was stored at - 80 °C.

**Strep-tag affinity chromatography**

The supernatant was loaded onto a 2 ml Strep-Tactin column matrix by gravity flow. The matrix was washed with 5 cV buffer WRib (100 mM Tris pH 8, 150 mM NaCl, 1mM EDTA) and the protein was eluted from the matrix with buffer ERib (100 mM Tris pH 8, 150 mM NaCl, 1mM EDTA, 2,5 mM desthiobiotin) in 6 x 1 ml fractions. The elution fractions (10 µl sample + 5 µl 5x sample buffer) were analyzed on 18 % SDS acrylamide gels and visualized by Coomassie staining. After the elution of the ribosomal proteins, the matrix was washed 3 times with 5 cV regeneration buffer (IBA, 100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, 1 mM HABA (hydroxy-azophenyl-benzoic acid, pH 8.0) to displace desthiobiotin from the binding pocket. The regeneration buffer was removed by extensive washing with buffer W (IBA, 100 mM Tris-HCl, 150 mM NaCl, 1 mM EDTA, pH 8.0). Once the red color of the regeneration buffer has disappeared the column can be reused and was stored in buffer W at 4 °C.

**Ion exchange chromatography (IEC)**

For further purification of the ribosomal proteins, a cation exchange chromatography (CEX) with a SP Sepharose FastFlow matrix (4 cV) followed.

First the main strep-tag elution fractions were pooled, concentrated to 1 ml with Amicon® Ultra 15 ml Filters (MWCO 3 kDa, Merck Millipore) and a buffer exchange was performed to reduce the Tris and NaCl concentration in the sample for better binding to the SP Sepharose matrix. The concentrated sample was loaded onto a 5 ml HiTrap desalting
column and eluted with 2 cV SP-bufferRb A (50 mM Tris-HCl pH 8, 100 mM NaCl, 1 mM EDTA, 1 mM DTT) at 2 ml/min in 0.5 ml fractions using again the Äkta-purifier system.

The main fractions of the desalting step were pooled and loaded onto the SP Sepharose FastFlow column (0.5 ml/min), followed by a washing step with 5 cV SP-bufferRb A to remove unbound proteins. Then the ribosomal proteins were eluted by a linear salt gradient from 100 mM NaCl to 1 M NaCl over 20 cV (0% - 100% SP-bufferRb B: 50 mM Tris-HCl pH 8, 1 M NaCl, 1 mM EDTA, 1 mM DTT). Flow through, washing step and the elutions were collected in 1 ml fractions and were analyzed by detecting the absorption at 280 nm and 260 nm. The SP-fractions of the ribosomal proteins were dialyzed against an appropriate buffer depending on the assay that will be performed (2.15 and 2.16).

### 2.10.5 Purification of the C-terminal insertase domains RbCT, OaCT, O1CT and A3CT

The C-terminal domains (CTDs) of the YidC homologs from *R. baltica* (RbCT) and *O. alexandrii* (OaCT), as well as the CTDs of the eukaryotic insertase homologs Oxa1 from *S. cerevisiae* (O1CT) and Alb3 from *A. thaliana* (A3CT) were cloned with an N- or C-terminal His10-Tag in pET16b or pET22b, respectively. In a pilot experiment the expression of all constructs was tested in C43, C41 and BL21 cells for the N- or C-terminal tagged versions of the proteins. The results are summarized in table 2.1.

<table>
<thead>
<tr>
<th>CTD</th>
<th>Plasmid</th>
<th>HisTag</th>
<th>Strain</th>
<th>Induction [h]</th>
</tr>
</thead>
<tbody>
<tr>
<td>RbCT</td>
<td>2.3.2 B5</td>
<td>N-terminal</td>
<td>C41</td>
<td>3h</td>
</tr>
<tr>
<td>OaCT</td>
<td>2.3.2 B6</td>
<td>N-terminal</td>
<td>C41</td>
<td>3h</td>
</tr>
<tr>
<td>A3CT</td>
<td>2.3.2 B7</td>
<td>C-terminal</td>
<td>BL21</td>
<td>3h</td>
</tr>
<tr>
<td>O2CT</td>
<td>2.3.2 B8</td>
<td>N-terminal</td>
<td>BL21</td>
<td>3h</td>
</tr>
</tbody>
</table>

*Table 2.1 Summary of the best expression conditions of the CTDs*

**Recombinant protein expression**

The CTDs were expressed according to the best conditions described in table 2.1. The expression was performed in 1l LB amp medium and the cells were induced at an OD600
of 0.6 with 0.5 mM IPTG for 3 h at 37°C. The cells were pelleted and resuspended in 2 ml/g buffer A_CT1 (50 mM Tris pH 9, 300 mM NaCl, 5 mg MgCl₂, 10 % (w/v) glycerol) for RbCT and OaCT. A3CT and O1CT were resuspended in buffer A_CT2 (100 mM Hepes pH 7.5, 300 mM NaCl, 5 mM MgCl₂, 10 % (w/v) glycerol, 1 mM DTT). All cell suspensions were lysed using an One Shot cell disruptor (Constant Systems LTD) at 1,3 kbar. Before cell disruption, 0.1 mM PMSF was added. The lysate was centrifuged for 30 min at 50 000 g to spin down the unbroken cells and cell debris. The cleared supernatant, containing the soluble CTDs, was stored at -80 °C.

**Immobilized metal ion affinity chromatography (IMAC)**

2 ml (cV) NiNTA agarose (Quiagen) were equilibrated in the appropriate buffer A_CT, supplemented with 10 mM imidazole and combined with the cleared supernatant. The batch was incubated for 2 h at 4 °C on a rotary wheel and then filled into an empty 10 ml column (90 µm filter, MoBiTec) by gravity flow. The matrix of the bound RbCT and OaCT purification step was washed with 20 ml buffer W₁ (50 mM Tris pH 9, 600 mM NaCl, 5 mM MgCl₂, 10 % (w/v) glycerol, 50 mM imidazole). For the A3CT and O1CT purification, three washing steps were performed each with 5 ml buffer W₂ (50 mM Hepes pH 7.5, 300 mM NaCl, 5 % (w/v) glycerol, 1 mM DTT) supplemented with 0 mM, 10 mM, 50 mM imidazole, respectively.

Then the CTDs were eluted from the matrix with 10 ml buffer E_Rb (50 mM Tris pH 8, 150 mM NaCl, 5 % (w/v) glycerol, 400 mM imidazole, 1 mM DTT) for RbCT, with 10 ml buffer E_Oa (50 mM Tris pH 7, 150 mM NaCl, 5 % (w/v) glycerol, 400 mM imidazole, 1 mM DTT) for OaCT and with 10 ml buffer E_A3/O1 (50 mM Hepes pH 7.5, 150 mM NaCl, 5 % (w/v) glycerol, 1 mM DTT, 300 mM imidazole) for the C-terminal domains of Alb3 (A3CT) and Oxa1 (O1CT). The proteins were collected in 5 x 2 ml fractions and analyzed on 15 % SDS acrylamide gels with subsequent Coomassie staining.

**Ion exchange chromatography (IEC)**

For further purification of the CTDs a cation exchange chromatography (CEX) with a SP Sepharose FastFlow matrix (4 cV) followed.

The main elution fractions were pooled and concentrated to 2 ml with Amicon® Ultra 15 ml Filters (MWCO 5 kDa, Merck Millipore). The concentrated sample was loaded onto the SP Sepharose FastFlow column (0.3 ml/min), followed by a washing step with 5 cV SP-buffer A_CT to remove unbound proteins. For the different CTDs, different SP-buffers were used:
RbCT: 50 mM Tris pH 8, 150 mM NaCl, 5 % (w/v) glycerol, 1 mM DTT

OaCT: 50 mM Tris pH 7, 150 mM NaCl, 5 % (w/v) glycerol, 1 mM DTT

A3CT & O1CT: 50 mM Hepes pH 7.5, 150 mM NaCl, 5 % (w/v) glycerol, 1 mM DTT

Then the CTD proteins were eluted with a linear salt gradient from 150 mM NaCl to 1 M NaCl over 20 cV (0 % - 100 %) SP-buffer $B_{CT}$ (1 M NaCl in the respective SP-buffer$_{CT}$). Flow through, washing step and the elution were collected in 1 ml fractions and analyzed by detecting the UV-absorption at 280 nm and 260 nm and Coomassie staining in SDS-gels.

The main fractions were pooled and concentrated to 1 ml for buffer exchange on a HiTrap desalting column (GE Healthcare) with the Äkta-purifier System (GE Healthcare) and filtered prior to loading (Ultrafree® Durapore – PVDF 0,22 µm, Merck Millipore). The buffer exchange of the CTDs was performed at room temperature in the respective CTD-buffer (20 mM Hepes pH 8 (RbCT) / pH 7 (OaCT) / pH 7.5 (A3CT + O1CT), 150 mM NaCl, 2 mM MgCl$_2$, 1 mM EDTA, 1 mM DTT) at a flow rate of 2 ml / min. 0,5 ml fractions were collected and the peak fractions were analyzed on 15 % SDS acrylamide gels (10 µl sample) and subsequent Coomassie staining. The purified protein was stored at - 80 °C.

For CD spectroscopy studies, the CTDs were dialyzed in Micro Float-A-Lyzer (MWCO 0,5 – 1 kDa; Spectra/Por®) to 20 mM K$_3$PO$_4$ pH 8 / 7.5 / 7, respectively.

2.11 Protein concentration determination according to Scopes

This method was used to determine the concentration of the small ribosomal proteins L24, L29 and the C-terminal YidC peptides RbCT, OaCT, O1CT and A3CT since they do not contain enough aromatic amino acid residues (Trp, Tyr) for the absorbance measurements at 280 nm. At a wavelength of 205 nm, the absorbance of the peptide bondage of the protein backbone is measured. First, the extinction coefficient ($\varepsilon_{205}$) was determined according to the empirical formula proposed by Scopes (Scopes, 1974):

$$\varepsilon_{205} [\text{mg/ml}] = 27 + 120 \times \frac{A_{280}}{A_{205}}$$

Then, the protein concentration was calculated according to the Lambert-Beer law:

$$c [\text{mg/ml}] = \frac{A_{205}}{(d \times \varepsilon_{205})}$$

2.12 in vitro ribosome binding assay

The ribosomes were isolated from a soluble cell extract of the E. coli strain MRE600 according to Maguire et al. (2008). The isolated ribosomes and the purified YidC proteins
were incubated on ice for 1 h (100 nM ribosomes; 500 nM protein) in ribosome-binding buffer (20 mM Tris-HCl pH 7.4, 25 mM MgOAc₂, 100 mM NaCl, 0.1 % DDM, 1 mM DTT), with a total reaction volume of 500 µl. Then the samples were centrifuged for 3 h at 200 000 g in a TLA 100.3 rotor. The pellet was directly resuspended in SDS-PAGE loading buffer, whereas the supernatant was precipitated with 10 % trichloroacetic acid (TCA) overnight and then subjected to SDS-PAGE analysis and subsequent Coomassie staining or immunoblotting and detection with antibodies against the His₁₀tag. For quantification, the total amount of applied protein in the assay (input) was loaded on the gel separately.

2.13 Reconstitution of an RNC-YidC complex

The purification of the ribosome nascent chain complexes (RNCs) was carried out as described (Seidelt et al., 2009) by our collaboration partner from the Roland Beckmann laboratory at the Ludwig-Maximilian’s-University in Munich, Germany. The translation system was programmed with an mRNA coding for an N-terminal His-tag followed by amino acid residues 1 – 115 of MscL and the TnaC stalling sequence at the C-terminus. This results in the first two transmembrane helices of MscL being fully emerged from the ribosomal exit tunnel. The sample complex was reconstituted by incubating 10 pMol MscL-RNCs with 50 pMol freshly purified YidC-Rb from our lab (2.10.2) in a final volume of 50 µL grid buffer (20 mM Hepes pH 7.2, 100 mM KOAc, 10 mM MgOAc₂, 0.1 % Cymal 6) for 30 min at 37 °C.

2.14 Electron microscopy and image processing

The here described electron microscopy and image processing was established and carried out by our collaboration partners from the Roland Beckmann laboratory at the Ludwig-Maximilian’s-University in Munich, Germany.

The reconstituted RNC-YidC-Rb (2.10.2) complex was vitrified on 2 mm pre-coated Quantifoil R3/3 holey carbon supported grids using a Vitrobot Mark IV (FEI Company). For automated data collection on the Titan Krios TEM (FEI Company) under low dose conditions (~ 20 e/Å²) the magnification was set to nominal 75.000x with a defocus range between -1 µm and -3.5 µm. The microscope was operated at 200 keV and a magnification of 148.721x at the plane of the CCD using a 4k x 4k TemCam-F416 CMOS camera (TVIPS GmbH) resulting in an image pixel size of 1.12 Å (object scale).

A total of 14 165 micrographs were collected of which 4 488 were selected manually for further processing based on the information content of the power spectra and particle density on the grid. The data processing was performed using the SPIDER software.
package (Frank et al., 1996). The defocus of each micrograph was determined using the TF ED command in SPIDER and particles were automatically selected using SIGNATURE (Chen and Grigorieff, 2007). The complete data sets of 140 266 particles was aligned to the structure of an empty 70S ribosome that was generated using the crystal structure of an E. coli ribosome (Schuwirth et al., 2005). Using semi-supervised classification (Penczek et al., 2006) it was possible to sort for subpopulations showing distinct ribosomal conformations and ligands (+/- E-site tRNA, +/- ligand at tunnel exit). A final dataset of 51 903 particles resulted in a density map refined to 8.6 Å resolution according to a Fourier Shell Correlation (FSC at 0.5 cut-off) showing P-site tRNA and high occupancy of YidC at the ribosomal exit site.

2.15 in vitro pull down assays

For the in vitro pull down assays 120 µg of purified E. coli YidC_{NH\text{is}} (2.10.1) or YidC-Rb_{NH\text{is}} protein was prebound to 50 µl cV NiNTA resin (Quiagen), respectively. Immobilization of the YidC proteins was performed in 1 ml binding buffer containing 50 mM Tris pH 7, 150 mM NaCl, 1 mM DTT and 0,05 % DDM (E.coli YidC) or 0,1 % LDAO (YidC-Rb) for 2 h at 4°C on a rotary wheel. Unbound protein was then collected by gravity flow. The ribosomal proteins L24 and L29 (2.10.4; ~ 500 µg) were loaded onto preloaded His-YidC columns. After 3 washing steps with 200 µl wash buffer (50 mM Tris pH 7, 100 mM NaCl, 1 mM DTT), all bound proteins were eluted by wash buffer supplemented with 500 mM imidazole by centrifugation.

2.16 Circular dichroism (CD) - spectroscopy

To analyze the secondary structure of the C-terminal domains RbCT and OaCT, a spectropolarimeter (J-715; Jasco, Tokyo, Japan) was used for CD measurements with 1 mm path-length quartz cuvettes (Hellma AG, Müllheim, Germany). All experiments were performed in 20 mM potassium phosphate buffer, pH 7 for OaCT or pH 8 for RbCT with a protein concentration of ~ 200 µg/ml. Background-corrected CD-spectra were recorded in the 190 - 260 nm range with a scan speed of 50 nm/min and a step size of 1 nm at 25 °C.

To facilitate the comparison of spectra from different samples, all spectra were normalized to mean residue weight ellipticity ($\theta_{MRW}$) [deg cm²/dmol], according to the formula:

$$\theta_{MRW} = \frac{MW \times \theta_{mdg}}{N \times c \times d \times 10000}$$
MATERIALS AND METHODS

$\theta_{\text{mdeg}}$ is the recorded ellipticity in millidegrees for a given wavelength, $MW$ is the molecular weight of the protein in Da, $d$ is the path length of the cuvette in cm, $N$ is the number of amino acid residues in the peptide and $c$ is the sample concentration in mg/ml.

Secondary structure predictions were performed on the DICHROWEB server (Whitmore and Wallace, 2004 & 2008) using the data set 4 (Sreerama and Woody, 2000) computed by the CDSSTR algorithm (Johnson, 1999).

Folding experiments of the C-terminal YidC domains upon titration of the ribosomal proteins L24 and L29 were performed by Emmanuelle Mboubi Kouaga and are described in detail in her Master thesis (Mboubi Kouaga, University of Hohenheim 2013).

2.17 Microscale thermophoresis (MST) measurements

MST measurements were performed to calculate the binding affinity of the isolated *E. coli* SRP to different purified internal signal peptides of KdpD (N22-48, W3A and 3Q).

2.17.1 Labeling of FfhM423C

Ffh$_{\text{M423C}}$ (2.10.3) was labeled using the NT-647-MALEIMIDE fluorescent dye and the RED-Maleimide Labeling kit (NanoTemper Technologies). The labeling reaction was performed according to the manufacturer’s instructions. Prior to labeling, a buffer exchange was performed to remove the glycerin from the protein sample (20 mM Hepes pH 8, 200 mM NaCl, 10 mM MgCl$_2$, 10 mM KCl) using the column provided in the kit. The protein concentration was adjusted to 10 µM and two labeling reactions were prepared with a 6 fold molar excess of the dye. After incubation for 30 min at room temperature both labeling reactions were combined, followed by the purification step to remove unreacted “free” dye. Therefore, the supplied dye removal columns were equilibrated with Ffh$_{\text{GF-buffer}}$ (20 mM Hepes pH 8, 200 mM NaCl, 10 mM MgCl$_2$, 10 mM KCl, 10 % glycerin). The degree of labeling (DOL, dye:protein ratio) was determined using an UV-Vis Spectrophotometer (NanoDrop 2000, ThermoScientific) at 650 and 280 nm. The DOL was calculated with:

$$\text{DOL} = \frac{A_{650} \times \varepsilon_{\text{Ffh}}}{(A_{280} - A_{650} \times \varepsilon_{\text{Ffh}}_{280}) \times \varepsilon_{\text{dye}}} \times \varepsilon_{\text{dye}}$$

$\varepsilon_{\text{Ffh}} = 8480$

$\varepsilon_{\text{dye}} = 250 000$

$\varepsilon_{\text{Ffh}}_{280 \text{mal}} = 0.05$

Thereby, a DOL of ~ 50 % was typically achieved. The labeled Ffh$_{\text{647NT}}$ protein was stored in aliquots at –80°C and protected from light.
2.17.2 Reconstitution of SRP

To get a functional SRP it is necessary to reconstitute the labeled Ffh$_{647NT}$ protein with the *in vitro* synthesized 4.5S RNA (2.6.12). The Ffh$_{647NT}$ protein was adjusted to 40 nM with MST$_{RNA}$ buffer (20 mM Hepes-KOH pH 7.5, 100 mM KCl, 10 mM MgCl$_2$, 0.5 mM EDTA, 0.05% Tween-20). Prior to use, the 4.5S RNA was 5-fold diluted, also in MST$_{RNA}$ buffer, to a final concentration of 9.48 µM. For refolding of the RNA, the diluted stock solution was heated to 75 °C for 2 min, and then cooled on ice for 1 min. Ffh$_{647NT}$ and 1.5 fold molar excess of 4.5S RNA (40 nM Ffh + 60 nM 4.5S RNA in 200 µl RNA buffer) were incubated together at 20 °C for 10 min and centrifuged for 10 min at 14 000 rpm, 4°C. The reconstituted SRP was kept on ice until start of the MST measurements.

2.17.3 MST measurement 1: Reconstitution of SRP

To analyze the correct assembly of SRP, the reconstitution and binding of Ffh to the 4.5S RNA was tested via MST measurement.

The labeled Ffh$_{647NT}$ protein was adjusted to 8 nM in MST$_{RNA}$ buffer (20 mM Hepes-KOH pH 7.5, 100 mM KCl, 10 mM MgCl$_2$, 0.5 mM EDTA, 0.05% Tween-20). The *in vitro* synthesized 4.5S RNA was diluted to a concentration of 400 nM in MST$_{RNA}$ buffer and a series of 16 1:1 dilutions was prepared in the same buffer, producing ligand concentrations ranging from 12.2 pM to 400 nM. For thermophoresis, each ligand dilution was mixed with one volume of labeled Ffh$_{647NT}$, which leads to a final concentration of fluorescently labeled Ffh$_{647NT}$ of 4 nM and a final ligand concentrations ranging from 6.1 pM to 200 nM. After 10 min incubation at room temperature, approximately 4 µl of each solution was filled into Monolith NT Premium Treated Capillaries (NanoTemper Technologies GmbH). Thermophoresis was measured using a Monolith NT.115 instrument at a temperature of 25 °C with 5 s/30 s/5 s laser off/on/off times, respectively. Instrument parameters were adjusted to 95 % LED power and 20 % MST power. Data of two independently pipetted measurements were analyzed (NT Affinity Analysis v2.0.1334, NanoTemper Technologies) using the signal from Thermophoresis + T-Jump.

2.17.4 MST measurement 2: Binding of SRP to signal peptides

To study the interaction of SRP to the putative SRP signal sequence at the very N-terminus of the KdpD protein different KdpD peptides (2.5), encompassing amino acid 22 - 48 of KdpD, were tested of their ability to bind to SRP via MST.
Therefore, the labeled SRP protein was reconstituted as described in 2.17.2 and adjusted to 40 nM in MSTRNA buffer. The \textit{in vitro} synthesized peptides were dissolved in MSTopt buffer (50 mM Tris pH 7.4, 150 mM NaCl, 10 mM MgCl2, 0.05 % Tween-20) and series of 16 1:1 dilutions were prepared in the same buffer, producing ligand concentrations ranging from 80 µM to 2.4 nM, respectively. For thermophoresis, each ligand dilution was mixed with one volume of labeled SRP, which leads to a final concentration of fluorescently labeled SRP of 20 nM and final ligand concentrations ranging from 1.2 nM to 40 µM. After 20 min incubation on ice, approximately 4 µl of each solution was filled into Monolith NT Premium Treated Capillaries (NanoTemper Technologies GmbH). Thermophoresis was measured using a Monolith NT.115 instrument at temperature of 23 °C with 5 s/30 s/5 s laser off/on/off times, respectively. Instrument parameters were adjusted to 50 % LED power and 20 % MST power. Data of 2 - 3 independently pipetted measurements were analyzed (NT Affinity Analysis v2.0.1334, NanoTemper Technologies) using the signal from Thermophoresis + T-Jump or manual evaluation settings (Cold region start / end: -1 s / 0 s; Hot region start / end: 4.51 s / 5.51 s).

2.17.5 $K_d$ Model: standard MST curve fitting mode

The $K_d$ Model is the standard fitting mode derived from law of mass action:

$$f(\text{Concentration}) = \text{Unbound} + \frac{(\text{Bound} - \text{Unbound}) \times (\text{concentration} + \text{TargetConc} + K_d - \sqrt{(\text{Concentration} + \text{TargetConc} + K_d)^2 - 4 \times \text{Concentration} \times \text{TargetConc}})}{2 \times \text{TargetConc}}$$

Unbound: response value of unbound state

Bound: response value of bound state

TargetConc: final concentration of fluorescent molecule

Fit quality was checked with the following parameters:

Response Amplitude = absolute difference between bound and unbound

$K_d$ Confidence = with a confidence of 68%, $K_d$ is within the given range.
The lower this number, the better the experimental data is represented by the fitted curve.

Standard Error of Regression = root mean square error (RMSE) of experimental data to fitted curve. The lower this number, the better the experimental data is represented by the fitted curve.
The measured data in this work were evaluated and presented as baseline-corrected normalized fluorescence ($\Delta F_{\text{norm}}$ in [%]). To obtain $\Delta F_{\text{norm}}$, the baseline $F_{\text{norm}}$ value (unbound molecule) is subtracted from all data points. Thus, by definition, $\Delta F_{\text{norm}}$ is 0 in the unbound state and can adopt positive or negative values, depending on an increase or decrease in $F_{\text{norm}}$ relative to the unbound state. This allows comparing both, amplitude of the binding curve as well as the $K_d$ of multiple experiments in one graph.
3 RESULTS

3.1 Chimera of the membrane insertase YidC with altered C-terminal regions

In comparison to the *E. coli* YidC (Ec-YidC), that has a charged C-terminal domain (CTD) of 16 aa only, the marine Gram-negative bacteria *Rhodopirellula baltica* and *Oceanicaulis alexandrii* feature membrane insertases with extended and highly positively charged C-terminal regions similar to the YidC homologs in mitochondria (Oxa1; CTD of 89 aa), chloroplasts (Alb3; CTD of 114 aa) and Gram-positive bacteria (YidC2; CTD of 64 aa).

The YidC homolog of the planctomycete *R. baltica* (Rb-YidC) has a size of 90 kDa and a hydrophilic C-terminal region of 80 amino acid residues. The α-proteobacterium *O. alexandrii* YidC homolog (Oa-YidC) is 83 kDa in size and has a positively charged C-terminal tail of 99 residues (Fig. 3.1).

![YidC homologues](image)

**Fig.3.1 Diversity of the C-terminal regions of YidC homologs.** The sequence after TM segment 6 is shown for *E. coli*, the marine bacteria *R. baltica*, *O. alexandrii* and, for comparison, the corresponding sequence of *S. cerevisiae* Oxa1. The length of the tail regions and the number of positively charged amino acid residues are indicated.

To analyze the role of the extended C-terminal tails, chimeric forms of *E. coli* YidC were constructed where the C-terminal region of YidC was replaced by the C-terminal tail of *R. baltica* (YidC-Rb), *O. alexandrii* (YidC-Oa) and *S. cerevisiae* (YidC-Oxa) (Fig. 3.2), respectively.
Fig. 3.2 Chimeric proteins of E. coli YidC with elongated C-terminal tails. (A) The respective sequence of the chimeric protein tails of Ec-YidC with the C-terminal tails of R. baltica (YidC-Rb), of O. alexandrii (YidC-Oa) and of S. cerevisiae (YidC-Oxa) is shown. (B) Schematic representation of the YidC proteins. The core domain of each YidC consists of 6 transmembrane segments and a large P1 domain, except Oxa1 from S. cerevisiae (TM 2-6).

Genetic construction of the YidC-chimera is described in Seidl, 2010 (University of Hohenheim). A brief summary of the cloning steps is listed in chapter 2.3.

3.1.1 Complementation of E. coli YidC

To analyze whether the YidC homologs and the chimera were functional, they were studied in the YidC depletion strain MK6S (2.2.5). In this strain, the yidc promoter is replaced by an araC-araBAD promoter on the chromosome. Growth in the presence of glucose for more than 2 h results in the depletion of YidC and finally to cell death, whereas the presence of arabinose allowed normal growth (Fig. 3.3 A and B). When the respective agar-plate contained 0.4 % glucose and 1 mM IPTG (Fig. 3.3 C) the plasmid-derived expression of Ec-YidC, Oa-YidC, YidC-Rb, YidC-Oa and YidC-Oxa promoted growth.

Fig. 3.3 Complementation of YidC in E. coli MK6. Serial dilutions of the E. coli MK6 YidC depletion strain cultures, bearing the plasmid-encoded YidC from the marine bacteria and the YidC chimera, were prepared. The cells were spotted onto agar-plates containing 0.2 % arabinose (A) to allow chromosomal YidC expression, 0.4 % glucose (B) to repress chromosomal YidC expression and 0.4 % glucose and 1 mM IPTG (C) to allow the sole plasmid-derived expression of YidC.
RESULTS

This clearly shows that the YidC homolog of *O. alexandrii* (Oa-YidC) and also the chimeric proteins YidC-Rb, YidC-Oa and YidC-Oxa are fully functional and can complement the wild-type YidC in *E. coli*, whereas the YidC homolog from *R. baltica* cannot. Therefore, only the three chimeras and the Oa-YidC homolog were used for the further studies.

### 3.1.2 Purification of the YidC proteins

For the *in vitro* assays and the cryo-EM studies it was necessary to purify N-terminally histidine-tagged versions of the YidC proteins in sufficient amounts. In the following chapter the purification of the Ec-YidC and the YidC-Rb chimera is shown.

**1. Purification of *E. coli* YidC_{NHis10} (2.10.1)**

The recombinant expression of the *E. coli* YidC from a pET16b vector (2.3.2 B1) was performed in a 6 L LB culture of *E. coli* C43 cells for 3 h at 37 °C. After the preparation of crude membranes, the Ec-YidC protein was solubilized by 1 % DDM and purified by Ni-chelating chromatography (IMAC). The protein was eluted in 10 x 2 ml fractions with buffer E_{EY} (20 mM Tris pH 7.5, 300 mM NaCl, 10 % glycerol, 0,05 % DDM, 1 mM DTT, 300 mM imidazole). 10 µl samples of elution fractions 1 – 7 were analyzed on a 12 % SDS acrylamide gel and Coomassie staining (Fig. 3.4).

![Fig. 3.4 IMAC elution fractions of Ec-YidC_{NHis}.
Samples, corresponding to the amount of protein of 1 ml cell culture, from the IMAC load (L), flow through (Ft) and the washing step (W), as well as 10 µl of each of the Ec-YidC elution fractions (1-7) were analyzed on a 12 % SDS gel by Coomassie staining.](Image)

Although the IMAC elutions were quite pure, the main fractions (Fig. 3.4 1-3) were further purified on a Superdex 200 16/60 column, especially to separate aggregates and higher oligomers of the Ec-YidC from the monomeric form to get a homogeneous protein sample. Additionally, the quick removal of imidazole from the sample is important for protein stability, since the Ec-YidC tends to aggregate in high imidazole concentrations. The size exclusion chromatography (SEC) of *E. coli* YidC was performed with a Superdex 200 16/60 column at room temperature in buffer GF_{EY} (20 mM Tris pH 7.5, 300 mM NaCl, 10 % glycerol, 0,05 % DDM, 1 mM DTT) at a flow rate of 0,5 ml / min. The elution profile of Ec-YidC was monitored using the UV absorbance at 280 nm (Fig. 3.5 A). 2 ml fractions were collected and samples of the three main peaks were analyzed (10 µl) on a
Coomassie stained 12 % SDS gel (Fig. 3.5 B). All samples showed a single Ec-YidC band. However, implicating the elution profile of the SEC, there were huge differences in homogeneity for the different fractions. The first peak, represented by fraction 2 with an elution volume at ~ 44 ml, corresponds to the void volume ($V_0$) with a molecular weight (MW) of > 2000 kDa, presumably containing aggregated Ec-YidC. Ec-YidC from fraction 5 of the second peak that is not completely separated from peak 1, has a high oligomeric form since the elution volume at ~ 51 ml corresponds to a MW of about 440 kDa. The third peak around an elution volume of 65 ml presumably contains the dimeric form of Ec-YidC with a MW of ~ 158 kDa. This calculated size matched perfectly for the 60 kDa Ec-YidC being a dimer in a DDM micelle with a theoretical MW of ~ 40 kDa. However, since the peak had a wide range from fractions 11 – 17, corresponding to a MW of 200 – 100 kDa, it is difficult to estimate if there are only dimers or a mixture of monomeric (later fractions around 70 ml) and dimeric (earlier fractions around 65 ml) forms of Ec-YidC.

Fig. 3.5 SEC elution profile of Ec-YidC on a Superdex 200 16/60 column. (A) Pooled Ec-YidC IMAC elution fractions (5 ml; Fig. 3.4) were loaded on the Superdex 200 16/60 column to separate aggregated (1. peak at 44.34 ml; > 2000 kDa) and higher oligomers (2. peak at 50.72 ml; ~ 440 kDa) from the dimeric or monomeric (3. peak at 65.34 ml; ~ 158 kDa) form of the protein. Flow rate: 0.5 ml/min System: Äkta purifier Detection: absorbance at 280 nm Fractionation: 2 ml (B) SEC fractions of Ec-YidC, corresponding to the three main peaks were analyzed on 12 % SDS gels (10 µl of each fraction) and Coomassie staining. Blue circles: fraction 2 from peak 1; fraction 5 from peak 2. Red box: fractions 11 – 17 of the third dimer / monomer Ec-YidC peak. All fractions showed pure Ec-YidC, but in putatively different oligomeric states.
RESULTS

Fractions 14 – 16 (Fig. 3.5 B) were concentrated by an Amicon® Ultra centrifugal unit 15 ml (MWCO 30 kDa) to a concentration of 0,641 mg/ml (= 10,6 µM). The concentration was determined by measurement of the absorbance at 280 nm with the NanoDrop 2000 UV-Vis spectrophotometer and calculated using the following parameters: \( \varepsilon_{YidC} = 93310 \); MW = 60 kDa. The concentrated Ec-YidC sample was stored in aliquots at 4 °C and used for further in vitro assays (3.2.2 & 3.3.2).

2. Purification of the YidC-Rb\textsubscript{NHis} chimera (2.10.2)

High effort was made to optimize the YidC-Rb purification, mainly for the cryo-EM structure (3.3.1). For the reconstitution of YidC-Rb with RNCs for the cryo-EM studies it was necessary to get a very pure and homogeneous protein sample of 1 ml with concentrations between 1 – 2 mg/ml.

The recombinant expression of the YidC-Rb from a pET16b vector (2.3.2 B3) was performed in a 6 L LB culture of E. coli C43 cells for 2 h at 37 °C. After the preparation of inner membrane vesicles (IMV), the YidC-Rb protein was solubilized with 1,5 % LDAO and purified by Ni chelating chromatography (IMAC). The protein was eluted in 10 x 2 ml fractions with buffer \( E_{ER} \) (50 mM Tris pH 7.5, 300 mM NaCl, 5 % glycerol, 0,2 % LDAO, 1 mM DTT, 500 mM imidazole). 10 µl sample of fractions 1 – 8 were analyzed on a 12 % SDS acrylamide gel and Coomassie staining (Fig. 3.6).

For further purification, the main elution fractions (Fig. 3.6 1-7) were pooled, concentrated (Amicon® Ultra 15 ml; MWCO 30 kDa) to 5 ml and loaded onto a Superdex 200 16/60 column. The SEC of YidC-Rb was performed in buffer \( GF_{Rb} \) (50 mM Tris pH 7.5, 300 mM NaCl, 5 % glycerol, 1 mM DTT, 0,2 % LDAO) at a flow rate of 0,75 ml/min at room temperature. The elution profile of YidC-Rb was monitored using UV absorbance at 280 nm (Fig. 3.7 A). 1 ml fractions were collected and samples of the main peaks were analyzed (10 µl each) on a Coomassie stained 12 % SDS gel (Fig. 3.7 B). All samples of the different fractions at the various retention times showed a single YidC-Rb band, indicating that the YidC-Rb load (Fig. 3.7 B load\textsubscript{SEC}) contained inhomogeneous oligomeric states of the protein. This became obvious by analyzing the elution profile of the SEC (Fig.
3.7 A) and calculating the respective molecular weights. YidC-Rb started eluting from the beginning of the void volume (= 43 ml) up to the end of the main elution peak (= 70 ml). However, there was a dominant peak with a maximum at 63,25 ml (fractions 22 – 25), corresponding to a MW of ~ 180 kDa. This leads to the assumption that the dimeric form of YidC-Rb, with a calculated monomeric mass of ~ 70 kDa and a LDAO micelle with the theoretical MW of ~13 kDa, was present in the second main elution peak. Again it cannot be excluded that in the later fractions (28 – 33) of the second peak monomeric YidC-Rb was present.

Fig. 3.7 SEC elution profile of YidC-Rb on a Superdex 200 16/60 column. (A) Pooled YidC-Rb IMAC elution fractions (5 ml concentrated) were loaded on the Superdex 200 16/60 column to separate aggregated (1. peak at 47,67 ml; > 2000 kDa) and higher oligomers (fraction 12 - 18) from the dimeric or monomeric (2. peak at 63,25 ml; ~ 158 kDa) active form of the protein. Flow rate: 0,75 ml/min System: Äkta purifier Detection: UV absorbance at 280 nm Fractionation: 1 ml (B) SEC fractions of YidC-Rb, corresponding to the main peaks were analyzed on 12 % SDS gels (10 µl of each fraction) and Coomassie staining. LoadSEC: 10 µl sample of 5 ml concentrated YidC-Rb IMAC elution. Dashed lines blue: fraction 7+8 from peak 1; fraction 12+13, 17+18 from the shoulder between peaks 1+2; fraction 22 – 29 of the main dimer / monomer YidC-Rb peak. Dashed lines red: dimeric YidC-Rb fractions 22-25. All fractions showed pure YidC-Rb, but in different oligomeric states.

The main peak fractions 22 – 25 (Fig. 3.7; red dashed lines) were concentrated to 1,5 mg/ml (= 23,07 µM). The concentration was determined by measurement of the UV
absorbance at 280 nm and calculated using the following parameters: $\varepsilon_{\text{YidC-Rb}} = 96260$; $\text{MW} = 70$ kDa. The concentrated YidC-Rb sample was stored in aliquots at 4 °C and used for the cryo-EM studies (3.3.1) and for further in vitro assays (3.2.2 & 3.3.2).

### 3.1.3 Ribosome binding of YidC is improved by the extended C-terminal tails

Since ribosome binding of the C-terminal region of the mitochondrial Oxa1 protein has been documented (Jia et al., 2003), it would be interesting to see whether the extended C-tails of YidC homologs from Gram-negative marine bacteria have a similar function.

The ribosome binding assays were mainly established within the scope of my diploma thesis (Seitl, 2010) and therefore will be discussed just briefly in this work. Ribosomes were isolated from *E. coli* MRE600 cells using a special chromatography system (Maguire et al., 2008) and incubated with the purified YidC proteins (Fig. 3.2) on ice. The ribosome-bound YidC proteins were separated by centrifugation and the pellet and supernatant fractions were analyzed by SDS-PAGE (Fig. 3.8). Whereas YidC-Rb and YidC-Oa fusion proteins readily bound to the 70S ribosomes and consequently were found in the pellet fraction (Fig. 3.8 A and C), YidC from *E. coli* was found in the supernatant (Fig. 3.8 A). Likewise, Oa-YidC, the insertase homolog from *O. alexandrii*, was found associated with the *E. coli* ribosomes in the pellet fraction (Fig. 3.8 D). When no ribosomes were present, virtually the bulk of all YidC versions was found in the supernatant.
Fig. 3.8 *Ribosome binding of YidC by the extended C-tails.* The N-terminally his-tagged, purified YidC proteins were incubated with isolated *E. coli* ribosomes and binding was analyzed by centrifugation. Pellet (P) and supernatant (S) were collected and analyzed by SDS-PAGE, stained with Coomassie blue (upper panels) or analyzed by Western blot using an anti-His antibody (lower panels). Ec-YidC (A), YidC-Rb (B), YidC-Oa (C) and Oa-YidC (D) were analyzed. As a control, the distribution of the YidC proteins was monitored in the absence of ribosomes. S1: ribosomal protein S1 of the 30S subunit.

The binding of the YidC proteins with the extended C-terminal tails to the ribosomes was salt sensitive, indicating that binding is mediated by ionic interactions (Fig. 3.9).

![Image](https://example.com/image1)

Fig. 3.9 *Binding of the YidC variants to *E. coli* 70S ribosomes is mediated by ionic interactions.* To test the salt-sensitivity of the interaction, the in vitro ribosome binding assay was performed with increasing concentrations of NaCl. In the presence of up to 300 mM NaCl constant amounts of YidC-Rb (A), YidC-Oa (B) and Oa-YidC (C) was found in the ribosome containing pellet. In the presence of 400 mM NaCl binding was significantly reduced for all three YidC proteins.

These data demonstrate that the charged C-terminal extensions are crucial for efficient binding of the ribosome to YidC. Therefore, the C-terminal tails of the marine YidC homologs most likely function as ribosome binding domains.
3.2 Structural studies of the YidC<sub>Rb</sub> - ribosome complex

Since the obtained results demonstrate increased ribosome affinity of the YidC derivatives with an extended hydrophilic C-terminal domain, we performed a structural analysis of YidC-ribosome complexes using cryo-EM and CD-spectroscopy.

3.2.1 Cryo-EM structure of the YidC<sub>Rb</sub> – RNC complex

In order to get high resolution cryo-EM structures of a YidC<sub>Rb</sub> – RNC complex, the preparation of a homogeneous sample was a prerequisite. The homogeneously purified YidC-Rb protein (3.1.2) was provided from our laboratory. The ribosome nascent chain (RNC) formation and purification, as well as the cryo-EM and 3D reconstruction was carried out by our collaboration partner Stephan Wickles from the Roland Beckmann laboratory (Gene Center and Department of Biochemistry) of the Ludwig-Maximillians-University in Munich, Germany.

In addition to the C-terminal ribosome binding domain of the YidC-Rb chimera, the use of RNCs instead of empty ribosomes stabilizes the insertase-ribosome complex (Kedrov et al., 2013). Therefore we used a TnaC stalled (Seidelt et al., 2009) nascent peptide chain of the mechanosensitive channel protein MscL, that is known to be a native substrate protein of the E. coli YidC. MscL is targeted to the membrane of E. coli by SRP and is then inserted by YidC into the inner membrane (Facey et al., 2007). The MscL protein consists of 136 amino acid residues and possesses two transmembrane helices (aa 17-45 and 75-94). In addition, there is a cytoplasmic amphipathic helix in the C-terminal segment (aa 105-119; Fig. 3.10 A). The in vitro translation system was programmed with an mRNA coding for an N-terminal His-tag followed by amino acid residues 1 – 115 of MscL and the TnaC stalling sequence at the C-terminus (Fig. 3.10 B). This results in a ribosome-coupled nascent protein chain bearing the first two transmembrane helices of MscL being fully emerged from the ribosomal exit tunnel. The RNCs were synthesized in a coupled in vitro transcription-translation reaction and purified as described (Seidelt et al., 2009), using the His<sub>6</sub>-tag for affinity purification.
For the cryo-EM analyses, isolated MscL-RNCs were reconstituted with the purified YidC-Rb (3.1.2) for 30 min at 37 °C. We determined the 3D structure of a complex consisting of a single YidC-Rb bound to a translating ribosome carrying the MscL protein with its first two TM segments as a nascent chain stalled by a C-terminally fused TnaC sequence (Fig. 3.11). The position of the extra density accounting for YidC-Rb agreed well with the density pattern observed in an earlier study of ribosome-bound E. coli YidC (Kohler et al., 2009). Moreover, we observed a similar interaction pattern between our C-terminally extended YidC-Rb and the ribosome. Ribosomal rRNA helix H59 and the ribosomal protein L24 appear to be the main contact sites with both types of YidC. Thus, the presence of the C-terminal extension increases the affinity of YidC to the ribosome without changing its overall interaction mode. An additional contact to the ribosomal protein L29 was observed for the bound YidC-Rb protein (Fig. 3.11 D).
Fig. 3.11 Cryo-EM reconstruction of an RNC-YidC-Rb complex. (A) The 30S and 50S ribosomal subunits are shown in yellow and grey, respectively; tRNA is shown in green and additional density accounting for YidC-Rb in red. (B) Cut through of the density to view the ribosomal tunnel and fragmented nascent chain density in green. (C and E) Close up view with the fitted molecular model of a 70S ribosome. L24 and L29 are shown in blue. YidC-Rb shows defined contacts to the ribosomal proteins L24, L29 and the helix H59. (D) Close up bottom view of the interaction area with a transparent density for YidC-Rb and the nascent MscL chain (green) surrounded by the YidC-Rb density.

For the 3D reconstruction a total of 14 165 micrographs were collected. 4 488 were selected manually for further processing based on the information content of the power spectra and particle density on the grid. The complete data sets of 140 266 particles was aligned to the structure of an empty ribosome that was generated using the crystal structure of an *E. coli* ribosome (Schuwirth et al., 2005). Using semi-supervised classification (Penczek et al., 2006) it was possible to sort for subpopulations showing distinct ribosomal conformations and ligands (+/- E-site tRNA, +/- ligand at tunnel exit). A final data set of 51 903 particles resulted in a density map refined to 8.6 Å resolution.
according to a Fourier Shell Correlation (FCS at 0.5 cut-off) showing P-site tRNA and high occupancy of YidC-Rb at the ribosomal exit site.

The TnaC-stalled ribosome has a strong density for a tRNA molecule located in the P-site indicating a high percentage of programmed RNCs in the final dataset. The path of the nascent chain can be traced from the peptidyl transferase center (PTC) through the ribosomal tunnel into an extra density at the tunnel exit representing YidC-Rb (Fig. 3.13).

Fig. 3.12 Local resolution of the RNC-YidC-Rb map. Surface and cross-section of the electron density map colored according to the local resolution. A reconstitution of the complex could be refined to 8.6 Å resolution after semi-supervised classification using competitive projection matching.
Fig. 3.13 3D structure of the isolated density for the nascent chain – YidC-Rb complex at 8.6 Å. (A) Bottom view of the isolated YidC-Rb density. Density blot of YidC-Rb shows entering of the nascent MscL chain (green) in the center of the YidC-Rb density (red). The asterisks mark the position of the ribosomal protein L29. (B) Side view of the translating ribosome carrying the MscL protein with its first 2 TM segments as a nascent chain stalled with an introduced TnaC sequence.

Notably, the presence and size of the detergent micelle (LDAO) surrounding the solubilized YidC-Rb was difficult to estimate at this resolution. Also, the large periplasmic loop of YidC-Rb between TM1 and TM2 is not resolved in the cryo-EM structures probably due to its dynamic motions relative to the membrane-embedded domains. The size of the density indicates the presence of only one copy of YidC-Rb under our conditions (Fig. 3.11). Even when assuming the presence of a minimal micelle, a YidC homodimer, as had been suggested in a previous study (Kohler et al., 2009) could not be accommodated. This suggests that a single copy of YidC is stably bound to the ribosome and may be sufficient to act as a ribosome-bound insertase complex for MscL as proposed by Kedrov et al., 2013.
3.2.2 The ribosomal protein L29 binds to the C-terminal tail of YidC-Rb

To investigate whether the observed contact between YidC-Rb and L29 is caused by the extended C-tail of the protein, the ribosomal proteins L24 and L29 were purified and their interaction with YidC-Rb was analyzed in pull-down experiments. The purified L24 protein was kindly provided by my master student Emmanuelle Mboubi. Since both ribosomal proteins were purified similarly (2.10.4), only the purification of L29 is shown here.

The ribosomal proteins were expressed in *E. coli* BL21 cell cultures and purified via a Strep-tag fused to the N-terminus of the protein. Because both proteins, L24 as well as L29, were soluble, the cleared supernatant after the cell disruption was directly applied to strep-tag affinity chromatography. L29\textsubscript{strep} was eluted from the matrix with buffer E\textsubscript{Rib} (100 mM Tris pH 8, 150 mM NaCl, 1mM EDTA, 2.5 mM desthiobiotin) and 10 µl of each fraction were analyzed on an 18 % SDS gel (Fig. 3.14). The L29 protein was isolated and ran on the gel at a height, corresponding to a molecular weight of ~ 10 kDa. Notably, the contaminating proteins in the elution fractions showed the typical pattern of 70S ribosomal proteins (Fig. 3.14). Therefore it seems that the recombinantly expressed L29\textsubscript{strep} protein was present in a complex with the endogenous *E. coli* ribosomes, so that the ribosomes co-eluted with L29\textsubscript{strep} from the matrix.

![Strep elution fractions of L29](image)

To separate the strep-tagged L29 from the endogenous ribosomal proteins and other contaminants, an ion exchange chromatography (IEC) followed. Since L29, as well as L24, has a rather basic isoelectric point (pI) of ~ 10, a cation exchange chromatography (CEX) using a SP Sepharose Fastflow column was performed. Prior CEX, the salt concentration of the L29 sample has to be reduced to ensure proper binding to the SP Sepharose matrix. The main fractions from the Strep-tag affinity purification (Fig. 3.14; fractions 2-4) were concentrated (Amicon® Ultra, MWCO 3 kDa) to 1 ml and a buffer
exchange with SP-bufferA (50 mM Tris pH 8, 100 mM NaCl, 1 mM EDTA, 1 mM DTT) on a HiTrap desalting column using the Äkta-purifier system was performed. The resulting HiTrap elution fractions were pooled and immediately loaded onto the SP Sepharose column for CEX (maximal load: 5 ml), followed by a washing step with SP-bufferA to remove unbound proteins. Subsequently, the L29 protein was eluted by a linear salt gradient from 100 mM (0% B) to 1 M NaCl (100% B) over 20 cV. L29 eluted at the beginning of the NaCl gradient between 5 – 10 % buffer B (Fig. 3.15 A), showing a defined peak including the fractions 3 - 8 that were analyzed on a 18 % SDS gel (Fig. 3.15 B).

The early elution point (5 – 10 % buffer B) suggests a weak binding of L29 to the cation exchange matrix, probably due to still too high salt concentration in the loaded sample and
in the SP-buffer (100 mM NaCl). However the peak fractions showed pure L29 protein in sufficient amounts and hence were used for the further pull down assays.

To study the interaction of the ribosomal proteins with the YidC insertases, the purified ribosomal proteins were loaded on Ni-NTA agarose matrix-immobilized YidC-Rb\textsubscript{NH\textsubscript{4}} and Ec-YidC\textsubscript{NH\textsubscript{4}}, respectively (3.1.2). The L24 as well as the L29 protein co-eluted from the resin together with the His-tagged YidC-Rb protein (Fig. 3.16 A; lane 4), demonstrating that both ribosomal proteins physically interact with YidC-Rb. In contrast, L29 was unable to bind and therefore did not co-elute from the resin together with the His-tagged wild type Ec-YidC (Fig. 3.16 B lane 4, lower panel), whereas L24 co-eluted in similar amounts as for YidC-Rb (Fig. 3.16 A and B lane 4, upper panels).

![Fig. 3.16 Interaction of the ribosomal proteins L24 and L29 with YidC-Rb and Ec-YidC. Pull down assays of L24 (upper panels) and L29 (lower panels) were performed with immobilized YidC-Rb (A) and Ec-YidC (B), respectively. About 120 µg of the respective YidC\textsubscript{His} proteins were attached to the NiNTA resin. Purified L24\textsubscript{Strep} or L29\textsubscript{Strep} (~ 500 µg) proteins were then loaded onto the immobilized YidC proteins. Lanes 1 show the total amount of added L24 and L29, respectively (input). Unbound protein was collected by gravity flow (lanes 2). After three washes (lanes 3 show the last washing steps), bound proteins were eluted with 500 mM imidazole by centrifugation (lanes 4). The proteins were acid-precipitated and analyzed by SDS-page and Coomassie staining.](image)

These results show that the enhanced affinity of YidC-Rb to the ribosome is at least partially caused by the interaction of the C-terminal domain of the \emph{R. baltica} YidC homolog.
with the ribosomal protein L29. But also features of the wild type Ec-YidC may contribute to the binding as suggested by the interaction of Ec-YidC with the ribosomal protein L24.

### 3.2.3 Structure of the YidC/Oxa1/Alb3 C-terminal domains (CTDs)

To further investigate the native architecture of the C-terminal domains of the diverse YidC homologs, the CTDs of both marine YidC homologs (RbCT, OaCT) as well as of Alb3 (A3CT) and Oxa1 (O1CT) were isolated as peptides without the insertase core domain and inspected by far-UV CD-spectroscopy and binding assays. The most important facts of the CTDs are summarized in Fig. 3.17 A. The amino acid sequence alignment of the CTDs showed no significant homology (Fig. 3.17 B). Therefore, it is assumed that the functional similarities are caused by shared structural features in the C-terminal extended domains, like the huge number of positively charged amino acid residues, the highly basic theoretical pI and probably a specific folding mechanism upon binding.

<table>
<thead>
<tr>
<th>CTD</th>
<th>organism</th>
<th>number of amino acids (aa)</th>
<th>positively charged aa</th>
<th>theoretical pI</th>
<th>MW [kD]</th>
</tr>
</thead>
<tbody>
<tr>
<td>RbCT</td>
<td>R. baltica</td>
<td>80</td>
<td>19</td>
<td>10.44</td>
<td>8.4</td>
</tr>
<tr>
<td>OaCT</td>
<td>O. alexandrii</td>
<td>99</td>
<td>17</td>
<td>9.52</td>
<td>10.4</td>
</tr>
<tr>
<td>O1CT</td>
<td>S. cerevisiae</td>
<td>91</td>
<td>22</td>
<td>10.73</td>
<td>10.7</td>
</tr>
<tr>
<td>A3CT</td>
<td>A. thaliana</td>
<td>114</td>
<td>26</td>
<td>9.39</td>
<td>12.7</td>
</tr>
</tbody>
</table>

Fig. 3.17 Comparison of the CTDs of diverse members of the YidC/Oxa1/Alb3 family. (A) All compared CTDs have in common that the elongated C-terminal domain (from 80 – 114 aa) contains a huge number of positively charged aa (from 19 – 26) and a basic theoretical isoelectric point (pI) of about 10. All parameters were calculated using the ExPASy ProtParam tool. (B) Sequence alignment of the C-terminal amino acid residues of RbCT, OaCT, O1CT and A3CT. There is no striking sequence homology within the CTDs. Only 1 lysine residue shows an identical
position in all 4 domains (asterisk). Two points indicate position of residues with very similar characteristics and one point indicates residues with minor similarities. Positively charged residues are shaded green. The alignment was performed using UniProt (UniProt Consortium, 2015).

The CTDs of RbCT, OaCT and O1CT were amplified by PCR using the YidC-chimera plasmids as template DNAs and cloned as fusion with either an N- or C-terminal His tag in pET16b or pET22b, respectively (2.3.2 B5, B6 + B8). The A3CT (2.3.2 B7) construct was kindly provided by the Irmgard Sinning group, Heidelberg (Falk et al., 2010). The CTDs were expressed according to the best conditions described in Table 2.1 and purified from 1 L cultures as described in the method section (2.10.5). The IMAC elution fractions of the CTDs showed a good yield of RbCT, OaCT and A3CT (Fig. 3.18 A, B and D), whereas the purification of O1CT could not be achieved with satisfying amounts (Fig. 3.18 C).

For further purification of the CTDs, a cation exchange chromatography (CEX) using a SP Sepharose FastFlow matrix was performed. The salt concentration of the CTD elution buffer from the IMAC (50 mM Tris, 150 mM NaCl) was low enough to bind the proteins to the CEX matrix and no buffer exchange has to be performed prior to the IEC. The main fractions of the IMAC elutions were concentrated to 2 ml and loaded onto the SP matrix, followed by a washing step with low salt (150 mM NaCl) buffer A (see method section 2.10.5 for buffer details). Finally, the CTDs were eluted by a linear gradient from 0 % - 100 % of high salt buffer B (1 M NaCl) over 20 cV (Fig. 3.19 A-D). RbCT started eluting from
RESULTS

the matrix at ~ 40 % buffer B (≈ 490 mM NaCl), indicating proper binding to the cation exchange matrix (Fig. 3.19 A). RbCT also eluted in a second peak at higher NaCl concentration (fractions 48 + 49), but only with a low protein yield compared to the main elution peak of fractions 38 – 43. Additionally, the first RbCT peak showed an increased UV absorption at 260 nm (Fig. 3.19 A, red curve), probably due to bound ribosomal RNA of co-eluting ribosomes (Fig. 3.20). The RNA possibly caused earlier elution of RbCT from the negatively charged matrix than the pure RbCT protein without bound RNA in the second peak. The elution of OaCT (Fig. 3.19 B) started at 20 % buffer B (≈ 320 mM NaCl) in a symmetric peak, including the fractions 18 – 24. Similar to RbCT, a second elution peak of OaCT was obtained (fraction 34 – 46), but without changes in the 260 nm absorption. O1CT (Fig. 3.19 C) and A3CT (Fig. 3.19 D) eluted at 22 % and 30 % buffer B, respectively, without a second elution peak at higher NaCl concentrations.

![Fig. 3.19 IEC elution profiles of the C-terminal YidC/Oxa1/Alb3 domains. The cation exchange profiles show the close up view of the main elution peaks and the respective fractions, analyzed on 15 % SDS-gels by Coomassie staining (insets).](image)

(A) RbCT showed the best binding on the SP Sepharose matrix and started eluting at ~ 40 % buffer B. Interestingly, the RbCT peak had an extended absorption at 260 nm (red curve). OaCT (B) and O1CT (C) eluted around 22 % of the salt gradient and the full length A3CT (D) at ~ 30 %. RbCT and OaCT showed two distinct elution peaks. The first (RbCT fractions 38-43; OaCT fractions 18-24) with higher protein amounts than the second peak (RbCT 48-49; OaCT 34-46).

In general, the CTD peptides showed only small peaks with heights between 10 – 20 mAU for the absorption at 280 nm. This is not surprising since the C-terminal domains are small
proteins with nearly no aromatic amino acid residues. RbCT for example contains no tryptophan (trp), tyrosine (tyr) or cysteine (cys) residues and is therefore expected to be barely visible. Only O1CT contains 1 tryptophan residue and A3CT, as well as OaCT, one tyrosine residue, respectively. Therefore, the protein concentration was not determined by UV absorption at 280 nm, but the peptide backbone was used to calculate the concentration of the CTDs by measuring the UV absorption at 205 nm and estimation of the value according to the empirical formula proposed by Scopes (Scopes, 1974) (2.11).

To highlight the strong binding affinity of the C-terminal tails to ribosomes, RbCT\textsubscript{NH\textsubscript{is}} was expressed in \textit{E. coli} and the IMAC of RbCT was performed under low salt conditions, namely 150 mM NaCl instead of 600 mM in the washing step. Within this purification procedure native ribosomes co-purified with his-tagged RbCT resulting in a complex which is stable during SEC (Fig. 3.20).

As known from the aforementioned studies of this work (3.2 Fig. 3.9) the binding of the C-terminal domains to ribosomes is mediated by ionic interactions and therefore is salt-sensitive. To avoid the massive co-elution of ribosomes during CTD purification, a high salt washing step of 600 mM NaCl was performed, resulting in separation of the CTD-ribosome complexes as shown in Fig. 3.18.

In previous studies it was shown that the C-terminal domain of the chloroplast Alb3 protein is intrinsically disordered and folds upon interaction with the chaperone cpSRP43, its native binding partner (Falk \textit{et al.}, 2010). To test if a ligand-triggered folding also holds for the C-terminal domains of the marine YidC homologs and for Oxa1, potential intrinsic

\begin{figure}
\centering
\includegraphics[width=\textwidth]{Fig_3.20}
\caption{\textbf{RbCT forms a stable complex with \textit{E. coli} 70S ribosomes.} Native 70S ribosomes from \textit{E. coli} were co-purified with RbCT\textsubscript{NH\textsubscript{is}} under low salt IMAC conditions (left panel). The RbCT-70S complex was stable during SEC with a SD75 16/60 column and eluted in a large void peak (middle panel). For comparison, the protein pattern of isolated 70S ribosomes from \textit{E. coli} is shown (right panel).}
\end{figure}
disorder of the CTDs was analyzed by theoretical structure predictors (PONDR) suited to identify unstructured regions in proteins (Fig. 3.21). RbCT and OaCT are similar to A3CT, their structure being predicted as intrinsically disordered regions for the last 70 amino acid residues (Fig. 3.21 A and B). In contrast to that, O1CT showed two distinct, shorter unfolded regions that were separated by a clearly folded region encompassing the amino acid residues 30 – 35. Also the very C-terminal region (aa 75 – 90) is predicted to be folded.

Fig. 3.21 PONDR Predictions of naturally disordered regions in the C-terminal domains of YidC/Oxa1/Alb3 family members. RbCT, OaCT and A3CT are clearly predicted as intrinsically disordered regions with an overall percentage of disordered portions of 95, 76 and 94, respectively. However, O1CT with only 53 % overall disorder has two regions (aa 30-45 and 75-90) that were predicted to be folded. Large segments of predicted disorder are indicated by horizontal black bars.

The potential intrinsic disorder of the isolated C-terminal domains was then experimentally analyzed by far UV CD-spectroscopy. For these purposes, the purified proteins (Fig. 3.19 A-D) were dialyzed against 20 mM K₃PO₄ CD-buffer and the protein concentration was then determined. Unfortunately, O1CT and A3CT were not stable in the CD-buffer and it was not possible to record a proper CD-spectrum. For RbCT and OaCT, far-UV CD-spectra could be measured in the range of 260 nm – 180 nm with 4 accumulations and a
speed of 50 nm/min (2.16), respectively. The far-UV spectra of RbCT (Fig. 3.22 A, blue curve) and OaCT (Fig. 3.22 B, blue curve) exhibited minima around 200 nm and only weak ellipticity above 210 nm, typical for unfolded proteins. However, the formation of an α-helical conformation of RbCT and OaCT was induced by the addition of trifluoroethanol (TFE), resulting in changed CD spectra of RbCT and OaCT with shifted minima to 207 nm, stronger ellipticity around 222 nm and positive maxima at 185 nm, respectively (Fig. 3.22 A and B, red curves). Secondary structure predictions confirmed the presence of regions in both CTDs that could possibly form α-helices (Fig. 3.22 C and D), whereas for OaCT, larger α-helical regions were predicted. This may be a reason for the extended α-helical signal of the OaCT spectrum after the addition of TFE compared to the RbCT-TFE curve.

Fig. 3.22 RbCT and OaCT are intrinsically disordered and fold upon addition of TFE. Analyses of RbCT and OaCT secondary structures using CD spectroscopy indicates that both C-terminal domains are unfolded in solution. Addition of TFE changed their conformation and induced the formation of α-helices in both proteins. Far-UV CD spectra of (A) RbCT and OaCT (B) in the absence of TFE (blue lines) or in the presence of 50 % TFE (red lines). Secondary structure prediction for RbCT (C) and OaCT (D) was generated using the PSIPRED server (McGuffin, 2000). Helices are symbolized by purple cylinders. Two shorter helices were predicted for RbCT and two longer helices for OaCT.
Although it was shown that RbCT and OaCT are intrinsically disordered in solution (Fig. 3.21 A and B; Fig. 3.22 A and B blue curves), the helices predicted in Fig. 3.22 C and D are likely to form upon interaction with their respective binding partner, namely the ribosome. In the context of the full length insertase proteins, membrane interaction might also influence the structure of the C-terminal domains.

As a direct interaction of the ribosomal protein L29 with the C-terminal domain of YidC-Rb was shown in the previous pull-down experiments (Fig. 3.16 A), we tested whether RbCT folds upon binding to L29. CD-spectra of the two individual binding partners were compared to a CD spectrum of the RbCT-L29 complex (Fig. 3.23 A). L29 shows a typical CD-spectrum of an α-helical conformation (red line), whereas RbCT is in a disordered formation as mentioned (3.22 A). The equimolar mixture of RbCT and L29 did not show a significant conformational change of the sum spectrum (green line) compared to the native RbCT spectrum (blue line). Accordingly, the difference spectrum obtained by subtracting the CD spectrum of L29 from the spectrum of the RbCT-L29 mixture revealed the unchanged minima around 200 nm of a disordered conformation (Fig. 3.23 B).

![Fig. 3.23 RbCT shows no conformational change upon L29 titration. (A) CD spectra of RbCT (blue, unfolded), L29 (red, α-helical) and RbCT+L29 mixture (green). (B) The difference spectrum obtained by subtracting the CD spectrum of L29 from the RbCT+L29 mixture shows the typical minima of an unfolded structure. This assay was performed in cooperation with E. Kouaga Mboubi during her master thesis (2013).](image-url)

Taken together, these data show that similar to Alb3 the isolated C-terminal domains of the marine YidC homologs are intrinsically disordered in solution, but no conformational change could be detected upon addition of the putative ribosomal binding partner L29.
3.3 *In vivo* studies: function of the C-terminal YidC regions

The discovery of YidC isoforms with a C-terminal extension in mitochondria (Oxa1) and Gram-positive bacteria (YidC2) suggested that they play a role in facilitating the co-translational membrane targeting activity independent of the SRP pathway (Funes *et al*., 2009). This might also hold true for the C-terminal extensions of the marine YidC homologs of Gram-negative bacteria. To analyze the cellular function of the C-terminally extended YidC proteins during the co-translational protein targeting and insertion, depletion strains of the SRP components were used for the *in vivo* studies. For depletion of Ffh, the protein component of the signal recognition particle, the depletion strain MC-Ffh was constructed (2.6.11). In this strain, the promoter of the chromosomal *ffh* gene has been exchanged with the *araC-araBAD* promoter cassette. Depletion of the chromosomally encoded Ffh was thus achieved by growth of MC-Ffh in the presence of 0.4% glucose. For depletion of the SRP receptor protein FtsY, the strain IY26 (2.2.7) was obtained from E. Bibi, depleting the chromosomal FtsY expression also under glucose conditions (0.4%).

3.3.1 Insertion of MscL via YidC-Rb in the absence of a functional SRP system

The C-terminally extended YidC proteins are not able to compensate completely for the loss of the SRP components in whole cells (data not shown). However, it is conceivable that the C-terminal domains may replace particularly the ribosome-assisted targeting function of components of the SRP system. Therefore, the ability of the C-tail modified YidC proteins to rescue either the targeting function or the insertion step mediated by SRP and its receptor FtsY was tested. The YidC substrate protein MscL that is known to be targeted via SRP to the membrane of *E. coli* and is then inserted by YidC into the inner membrane (Facey *et al*., 2007) was analyzed. To investigate whether an extended C-terminal region of YidC can compensate for a loss of distinct SRP functions, the FtsY depletion strain IY26 and the Ffh depletion strain MC-Ffh were transformed with pGZ119HE-N-his encoding the various YidC proteins (2.3.3 C1-6) and a second plasmid pSF147 encoding a single cysteine mutant (I68C) of the MscL protein (2.3.2 C7). The cells were grown in glucose media to deplete the SRP receptor protein FtsY or the Ffh protein component of SRP, respectively (2.9.1). Membrane insertion of MscL was monitored by chemical modification of the single cysteine residue at position 68 in the periplasmic loop of MscL with AMS (2.9), a membrane impermeable sulfhydryl-reagent that shifts the protein mobility on SDS-gels. Addition of AMS during the radioactive pulse chase experiments leads to a derivatization of the translocated cysteine residue at
position 68 of MscL and results in a shift in molecular size of ~ 0.5 kDa of the MscL cysteine mutant, indicating its proper membrane insertion (Facey et al., 2007).

It was tested whether the expression of Oa-YidC, YidC-Rb, YidC-Oa or YidC-Oxa can compensate for a loss of Ffh in the membrane targeting efficiency of MscL. Therefore, MC-Ffh cells coexpressing the MscL-68C single cysteine mutant and the respective YidC variant were pulse-labelled with [35S]-Met, chased with non-radioactive methionine and then incubated with AMS (2.9.2). After treatment with AMS, total proteins were acid-precipitated and immuno-precipitated to visualize MscL (2.9.3). Derivatized and underivatized proteins were separated by SDS-PAGE and examined by phosphorimaging (2.9.4). In the absence of Ffh, AMS derivatization of MscL is reduced by about 40% (Fig. 3.24 B; left column). As expected, this corroborates that the insertion process of MscL is affected by Ffh depletion (Facey et al., 2007). Coexpression of the YidC proteins did not significantly restore membrane insertion of MscL under Ffh depletion conditions (Fig. 3.24 A; lanes 3-12), in contrast to the control cells where Ffh was coexpressed from a plasmid (Fig. 3.24 A lane 14; Fig. 3.24 B right column).

Fig. 3.24 Ffh is indispensable for the proper membrane insertion of MscL and not replaceable by C-terminally tailed YidC proteins. The coexpression of the YidC variants could

![Graph](image-url)
not maintain the correct insertion of MscL in the absence of Ffh. The membrane insertion of MscL was monitored by AMS reactivity of a single cysteine residue (C68) in the periplasmic loop of MscL. (A) E. coli strain MC-Ffh expressing the MscL cysteine mutant was grown in M9 minimal medium containing arabinose (cFfh+, lane 1). For depletion of Ffh (cFfh-, lane 2), the cells were grown in the presence of glucose. IPTG (1 mM) was added for 10 min to induce expression. Cells were incubated in the presence of AMS, pulse-labeled with [35S] Met for 2 min and chased with non-radioactive Met for 10 min. After quenching with 20 mM DTT, the radiolabeled samples were acid-precipitated and immune-precipitated with an anti-His antibody and then subjected to SDS-PAGE and phosphorimaging. The arrowhead D denotes the AMS-derivatized MscL protein. Parallel experiments were performed with cells coexpressing Ec-YidC (lanes 3 and 4), Oa-YidC (lanes 5 and 6), YidC-Rb (lanes 7 and 8), YidC-Oa (lanes 9 and 10) or YidC-Oxa (lanes 11 and 12). As a control, the derivatization of MscL was monitored in MC-Ffh cells coexpressing plasmid-encoded Ffh (pFfh) in the absence of chromosomal Ffh (cFfh) (lane 14). The lower panels show immunoblots of the Ffh level under arabinose (+) and glucose (-) conditions, respectively. (B) The AMS-shift assay was quantified with ImageJ. In each experiment the amount of derivatized MscL in the presence of Ffh was set to 100% (control). The relative amount of the derivatized MscL in the absence of Ffh (red columns) was calculated after coexpression of the indicated YidC protein from plasmid. Cells without coexpression are shown in the very left bar (-). The data for the control cells with coexpression of plasmid-encoded Ffh protein are shown in the right column (pFfh). For several repeated experiments the standard deviations (error bars) were calculated.

Subsequently, it was analyzed whether the C-terminal extensions of the YidC proteins could replace the function of the SRP receptor protein FtsY. To assess the ability of YidC-Rb, YidC-Oa or YidC-Oxa to mediate membrane targeting and insertion of MscL in the absence of FtsY, the same experiment as described above for Ffh, was performed in the FtsY depletion strain IY26. When FtsY was present, about 50% of the MscL protein was shifted during the pulse time indicating that it was correctly inserted into the inner membrane (Fig. 3.25 A, lane 1). However, when cells were grown in the absence of arabinose to deplete FtsY, a much lower proportion of MscL was derivatized by AMS (lane 2), showing that the targeting of MscL to the membrane and the insertion process is inhibited, somewhat stronger than after Ffh depletion. To investigate whether coexpression of the YidC proteins can compensate for the loss of FtsY and promote membrane insertion of MscL, a parallel experiment was performed with IY26 cells that express MscL together with Ec-YidC, Oa-YidC, YidC-Rb, YidC-Oa or YidC-Oxa, respectively (Fig. 3.25, lanes 3-12). In the absence of FtsY, coexpression of the YidC-Rb and YidC-Oxa proteins with the C-terminal extensions allowed substantial AMS derivatization of MscL (lanes 8 and 12). In contrast, the coexpression of YidC-Oa or Oa-YidC showed no significant improvement of the membrane insertion of MscL in the absence of FtsY (lanes 6 and 10). To exclude that the overexpression of a functional YidC protein causes an enhanced MscL insertion under FtsY depletion conditions, wild type Ec-YidC was coexpressed (Fig. 3.25 A, lanes 3 and 4) and no improved MscL derivatization was observed (Fig. 3.25 B, second left column).
**RESULTS**

**Fig. 3.25** The C-terminal extensions of YidC-Rb and YidC-Oxa can mediate the membrane targeting of MscL in the absence of the SRP receptor protein FtsY. MscL is inserted into the membrane in the absence of FtsY when YidC-Rb or YidC-Oxa is coexpressed. Samples and AMS derivatization assays were prepared and processed as described for MC-Ffh (Fig. 3.24). (A) The FtsY depletion strain IY26 bearing the MscL I68C cysteine mutant was grown under either FtsY depletion (lane 2) or FtsY expression conditions (lane 1). Parallel experiments were performed with cells coexpressing Ec-YidC (lanes 3 and 4), Oa-YidC (lanes 5 and 6), YidC-Rb (lanes 7 and 8), YidC-Oa (lanes 9 and 10) and YidC-Oxa (lanes 11 and 12). The lower panels show immunoblots of the FtsY level under arabinose (+) and glucose (-) conditions, respectively. (B) The quantified AMS-shift assay (ImageJ) showed that the relative amount of derivatized MscL in the absence of FtsY is enhanced by 30% when YidC-Rb was coexpressed and by 20% for the YidC-Oxa coexpression, compared to cells without coexpression (left column -) or coexpressing Ec-YidC (second left column). The total amount of derivatized MscL in the presence of FtsY was set to 100%. For several repeated experiments the standard deviations (error bars) were calculated.

Taken together, these data indicate that the membrane targeting and the subsequent insertion of MscL in the absence of FtsY can be mediated by the C-terminal domain of the marine YidC homolog of *R. baltica* and partially also by the C-terminal domain of the mitochondrial YidC-homolog Oxa1.
3.3.2 Localization of MscL-GFP in the absence of Ffh or FtsY

To visualize the cellular localization of MscL, the protein was expressed as an MscL-GFP fusion protein (Maier et al., 2008). The localization of MscL-GFP was analyzed in the FtsY depletion strain Y26 and in MC-Ffh cells. The cells were grown in arabinose containing media to allow and analyze normal membrane targeting or in glucose to deplete FtsY or Ffh, respectively. The YidC proteins were coexpressed and the cells were then inspected for the localization of MscL-GFP by fluorescence microscopy (2.8). In the presence of Ffh and FtsY, the fluorescence was evenly distributed at the membrane surface (Figs. 3.26 A and 3.27 A, respectively). However, when the cells were grown in the presence of glucose to deplete Ffh, the MscL-GFP aggregated at the cell poles (Fig. 3.26 B). Under FtsY-depleted conditions, the MscL-GFP protein was found in patches, also mostly at the cell poles (Fig. 3.27 B). This is similar to the phenotype when YidC is depleted (Fig. 3.26 C) in MK6 cells. Remarkably, coexpression of YidC-Rb in the absence of Ffh (Fig. 3.26 D) restored the mislocalization of MscL-GFP and showed a distribution of the fluorescent signal similar to the coexpression of wild type Ffh encoded on a plasmid (Fig. 3.26 F).

Fig. 3.26 Localization of MscL-GFP in the absence of Ffh. Localization of MscL-GFP was studied in vivo by fluorescence microscopy. MC-Ffh cells bearing the MscL-GFP fusion plasmid were grown in LB medium either in the presence of arabinose (A) or in the presence of glucose to deplete Ffh (B, D, E and F). Cells were induced with 1 mM IPTG for 2 h at 30 °C and inspected by fluorescence microscopy. The coexpression of YidC-Rb under Ffh depletion conditions (D) showed a distribution of the fluorescent signal similar to the positive control of the plasmid-encoded expression of wild type Ffh (F). Coexpression of Ec-YidC (E) showed more MscL-GFP patches, but also a different cell localization compared to Ffh-depleted cells expressing MscL-GFP with no other coexpressing plasmid (B). (C) Localization of MscL-GFP in YidC depleted MK6 cells.
Targeting of the MscL-protein to the *E. coli* inner membrane is clearly dependent on the SRP-receptor FtsY (Fig. 3.27, compare A and B). Strikingly, the missing proper localization of MscL-GFP at the membrane in the absence of FtsY in IY26 cells (Fig. 3.27 B) was restored by the coexpression of YidC-Rb (Fig. 3.27 C). Expression of YidC-Rb together with MscL-GFP allowed a cellular distribution of MscL-GFP similar to wild type conditions (Fig. 3.27 A) despite depletion of FtsY.

These results, together with the data from the AMS derivatization assays (3.3.1) indicate that YidC-Rb supports membrane targeting of MscL also in the absence of Ffh or FtsY, while the correct insertion and membrane topology of MscL still requires Ffh. Thus the charged C-terminal ribosome-binding moiety of YidC-Rb can function as a membrane targeting factor.
3.4 Binding studies of SRP and the SRP signal sequence of KdpD

Similar to MscL, the KdpD protein of *E. coli* is targeted via SRP to the membrane but it does not require SecYEG nor YidC for its membrane insertion (Facey and Kuhn, 2003). This chapter focuses on the interaction of the signal recognition particle with SRP signal sequences to determine the specificity of SRP recognition in proteins.

The sensor protein KdpD of *E. coli* is composed of a large N-terminal hydrophilic region (aa 1-400), four transmembrane regions (aa 401-498) and a large hydrophilic region (aa 499-894) at the C-terminus (Fig. 3.28 A). Both termini are facing the cytoplasm. As in *E. coli* nearly all inner membrane proteins have N-terminal uncleaved signal sequences, the region that interacts with SRP is most likely the first transmembrane region. This provides that early after the translation start, SRP interacts with the nascent peptide chain at the ribosome and mediates its co-translational targeting to the membrane (Luirink and Sinning, 2004). Due to the long N-terminal hydrophilic region of KdpD, the first transmembrane region starts at amino acid residue 401, preventing an early membrane targeting of the protein. Therefore, an additional signal element might be present in the N-terminal region that is recognized by SRP and allows early co-translational membrane targeting of KdpD. *In vivo* studies of N-terminal KdpD-GFP fusion proteins revealed that residues 22-48 of KdpD (N22-48) contain a signal element that is capable to bind SRP and to target the GFP-fusion protein to the membrane (Maier *et al.*, 2008) (Fig. 3.28 B). This peptide contains five positively charged residues, three of which are closely spaced (aa 22-26), as well as a stretch of 10 hydrophobic residues (aa 27-36) followed by a lysine residue and then another six hydrophobic residues (aa 38-43). Within this peptide, also a Walker A motif was identified ranging from residues 30-38 (Jung and Altendorf, 1998) (Fig. 3.28 B, boxed region).
To explore the function of the N22-48 sequence of KdpD, two SRP signal mutants therein were generated (Fig. 3.29). Since basic amino acid residues are discussed to promote binding of SRP to a signal peptide (Peterson et al., 2003), the positively charged residues at positions 22, 24 and 26 were substituted by glutamine (Q) resulting in the 3Q mutant in which the net charge of the peptide is lowered from + 5 to + 2. Influence of the Walker A motif on SRP recognition of the KdpD signal element was investigated by mutating the conserved glycine residues at positions 30, 34 and 36 to alanine, respectively, resulting in the W3A mutant.

Fig. 3.29 SRP signal mutants of KdpD N22-48. To investigate the influence of the positively charged residues at the beginning of the N22-48 peptide, the 3Q mutant was generated, resulting in a reduced net charge of the peptide (+2). In the W3A mutant (+5), three conserved glycine residues in the Walker A motif were exchanged with alanine residues. The mutated amino acid residues are marked with boxes. The Walker A motif is underlined and the positively charged residues are displayed in red.
To investigate the involvement of the positively charged residues and the Walker A motif in binding of the signal sequence to SRP, the binding affinities of the wild type KdpD N22-48 fragment and the mutant peptides 3Q and W3A to SRP were determined via microscale thermophoresis (MST) measurements (2.17). The MST refers to the motion of molecules in microscopic temperature gradients and allows the quantification of biomolecule interactions by thermophoretic detection of even minute changes in conformation, charge and physical size of a molecule as they are induced by a binding event (Jerabek-Willemsen et al., 2011). The thermophoretical movement is monitored via a fluorescently labeled target molecule whose concentration is kept constant and low (1 – 100 nM), while the unlabeled binding partner (ligand) is titrated in a serial dilution (12 - 16 steps) with concentrations ranging from approximately 20 fold above the expected dissociation constant (K_d) down to sub-stoichiometric concentrations with respect to the labeled molecule.

For this approach the SRP protein was labeled, while the unlabeled KdpD signal peptides N22-48, 3Q and W3A were titrated. Since the KdpD peptides are too small (2951 Da) to purify them efficiently from *E. coli* cells, the peptides were synthesized *in vitro* by the Custom peptide synthesis services from GENOSPHERE Biotechnologies (France) (2.5). The purification, labeling and reconstitution of SRP will be discussed in detail in the next chapter.

### 3.4.1 Purification and labeling of Ffh from *E. coli*

The *E. coli* wild type *ffh* gene was amplified from the *E. coli* MC1061 chromosome by PCR using primers coding for a C-terminal Strep-tag (WSHPQFEK) sequence and flanking Ncol / EcoR1 restriction sites. The *ffh* gene was cloned into the pMS119HE vector and the wild type cysteine residue at position 406 was mutated to serine. A single cysteine residue was introduced at position 423 resulting in the construct FfhCstrepC406S/M423C by site-directed mutagenesis (2.6.9). The recombinant protein expression was induced in a 4L LB culture of *E. coli* BL21 cells for 3 h at 37 °C and the FfhM423C mutant was purified as described in the method section 2.10.3 by strep-tag affinity chromatography. The FfhM423C protein was eluted in 2 ml fractions from the Strep-Tactin matrix in buffer E Ffh (20 mM Hepes pH 8, 350 mM NaCl, 10 mM MgCl₂, 10 mM KCl, 10 % glycerol, 2,5 mM desthiobiotin) in sufficient amounts with only few contaminants (Fig. 3.30).
RESULTS

Fig. 3.30 *Strep-tag elution fractions of FfhM423C.* 10 µl of each elution fraction (1-9) were analyzed on a 12 % SDS-gel and visualized by Coomassie staining. The FfhM423C protein was eluted by gravity flow in 2 ml fractions from the Strep-Tactin matrix (IBA) with 2.5 mM desthiobiotin in the elution buffer.

The main fractions 2 – 4 (Fig. 3.30) were further purified on a Superdex 75 16/60 gel filtration column with buffer GF_{Ffh} (20 mM Hepes pH 8, 200 mM NaCl, 20 mM MgCl₂, 10 mM KCl, 10 % glycerol). The SEC was monitored by tracking the UV absorptions at 280 nm and 260 nm and 2 ml fractions were collected (Fig. 3.31 A). The peak fractions were analyzed by a 12 % SDS-gel and Coomassie staining (Fig. 3.31 B). Two main peaks were detected during SEC, one representing the void volume at 45 ml and a second peak at an elution volume of 60 ml, corresponding to a molecular weight of ~ 50 kDa (Fig. 3.31 A). Presumably, the second peak contained the monomeric Ffh protein (48 kDa) with proper amount and purity (Fig. 3.31 B). The void volume peak showed a strong absorption at 260 nm (Fig. 3.31 A, red curve), indicating the presence of RNA or DNA. Indeed, the SDS-gel analysis of the void fractions 3 and 4 showed only minor amounts of Ffh protein (Fig. 3.31 B). Electrophoretic RNA analysis of these fractions revealed the presence of ribosomal 23S rRNA (2904 nt) and 16S rRNA (1540 nt) in the void volume, but not in the Ffh peak fraction 11 (Fig. 3.31 C). The assumption that the 260 nm peak contains 4.5S RNA (113 nt) was not confirmed since no RNA fragment at this size was visible on the RNA gel. Therefore, it is proposed that the Strep-Tactin matrix binds RNA unspecifically from the cell lysate during the purification process.
Fig. 3.31 SD75 16/60 SEC of Ffh\textsubscript{M423Cstrep}. (A) Chromatographic separation of the Ffh\textsubscript{M423Cstrep} elution fractions. The absorption was detected at 260 nm (red curve) and 280 nm (blue curve), monitoring 2 peaks at 45 ml (V\textsubscript{0}) and 60 ml (Ffh monomer; ~ 50 kDa). (B) SDS-gel analysis of the peak fractions 3+4 and 9-13. 10 µl of each fraction was loaded and analyzed by Coomassie staining. (C) RNA isolation of fractions 3, 4 and 11, analyzed on a 2 % agarose gel and visualized by ethidium bromide UV fluorescence, indicated the presence of ribosomal RNA at a size of ~ 3000 nt and ~ 1500 nt in the void volume (black circles, fraction 3 and 4), but not in the Ffh elution peak (black box, fraction 11). nt: nucleotides

The gel filtration fraction 11 (Fig. 3.31) was used for the subsequent labeling reaction. Ffh\textsubscript{M423Cstrep} was labeled using the NT-647-MALEIMIDE fluorescent dye and the RED-Maleimide labeling kit from NanoTemper Technologies (2.17.1). Prior to the labeling reaction, a buffer exchange had to be performed to remove the glycerol from the Ffh sample buffer since glycerol reduces the labeling efficiency drastically (Fig. 3.32, compare L1 with L2). L2 in figure 3.32 shows the Ffh protein from the labeling reaction with the same Ffh concentration in a buffer without glycerol demonstrating increased label efficiency. In the purification step removing the free dye from the reaction, the labeled Ffh protein eluted in fractions 3 and 4. The in gel fluorescence was measured by an ImageQuant\textsuperscript{TM} LAS400 system using a filter for Cy5 fluorescence (670 nm), detecting a distinct signal at the expected molecular weight of the Ffh protein in the SDS-gel without unspecifically labeled contaminants or free dye in the elution fractions (Fig. 3.32 E3 + E4).
RESULTS

Concentration determination and calculation of the DOL (degree of labeling) was performed by measuring the absorbance at 280 nm and 650 nm using an UV-Vis spectrophotometer (2.17.1). For fraction 4 (Fig. 3.32) a protein concentration of 2.5 µM and a labeling efficiency of ~ 47 % was calculated. The Ffh protein was stored protected from light in aliquots at – 80 °C and thawed on ice just prior to 4.5S RNA reconstitution.

3.4.2 Reconstitution of a functional SRP protein

At this point of the SRP preparation only the protein component Ffh was present (Fig. 3.31 C fraction 11) but for a functional signal recognition particle, the 4.5S RNA is indispensable. Therefore, the 114 nt long 4.5S RNA was synthesized in vitro (2.6.12) and reconstituted with the Ffh protein to a functional SRP (2.17.2). For the in vitro transcription of the 4.5S RNA, the template DNA had to be linearized by restriction digestion downstream of the insert. The pUC18-4.5S RNA plasmid (2.3.5) was therefore linearized by BamHI and gel purified using a DNA gel extraction kit. A phenol / chloroform extraction of the template DNA was performed to end up in an A260/280 ratio of 1.8 – 2. The in vitro transcription reaction was incubated for 3 h at 37 °C, the RNA was purified (RNA clean & concentrator-25 Kit) and analyzed on a 2 % native agarose gel stained with ethidium bromide (Fig. 3.33).

Fig. 3.32 In gel fluorescence measurement of the Ffh\textsubscript{647}mal labeling procedure. The first lane shows the fluorescent signal of the free dye (NT-647mal). L1 and L2 represent labeling reactions 1 (with 10 % glycerol) and 2 (without glycerol), respectively. E1-E4: elution fractions of Ffh\textsubscript{647}mal after removal of the free dye. ImageQuant settings: Fluorescence: Cy5; Light: Red (RGB); Filter: 670 nm; Iris: F 0.85; exposure time: 1 sec.

Fig. 3.33 In vitro transcription of 4.5S RNA. After the in vitro transcription the RNA was matrix-isolated and step wise eluted in 50 µl (A) and 25 µl (B) RNase free water. A total amount of 1,8 µg (A1 and B1) and 0.4 µg (A2 and B2) RNA was loaded onto a 2 % native agarose gel and detected by ethidium bromide fluorescence. UV exposure time: 0.2 sec. nt: nucleotides. Both elution fractions showed a single band <200 nt that corresponds to the 114 nt 4.5S RNA.
The *in vitro* transcription produced a defined RNA fragment with a nucleotide length of < 200 nt that was assigned to the 114 nt comprising 4.5S RNA. A molecular weight of 36.7 kDa was calculated and the UV absorbance measurement at 260 nm determined a RNA concentration of ~ 48 µM for both elution fractions.

Prior to the studies on the interaction of the SRP holo-particle with the N-terminal KdpD peptides by MST, it was necessary to establish conditions that ensure a stably reconstituted SRP protein and to exclude that the labeling affected the interaction between Ffh and its RNA. Therefore a molecular interaction experiment was performed between Ffh and the 4.5S RNA to determine the dissociation coefficient \(K_d\). A titration series of up to 16 dilutions was prepared, where the concentration of the fluorescent Ffh\(_{647\text{mal}}\) protein is kept constant and the concentration of the unlabeled 4.5S RNA is varied. For best and reliable performance, the concentration of the labeled component should be close to the expected \(K_d\) or less and should yield a range of 200 – 1500 fluorescence counts (typically equivalent to 1 – 200 nM). The unlabeled binding partner was prepared in a 16-step dilution series around the expected \(K_d\), starting from a concentration of > 20-fold of the expected \(K_d\) down to sub-stoichiometric concentrations. 10 µl of each sample of the dilution series was mixed with 10 µl of the stock solution of the labeled binding partner by pipetting up and down 10 times and the samples were incubated in the dark. Incubation times and temperatures differed between different binding assays. To design an appropriate MST experiment and find a proper concentration distribution, the concentration finder tool software from NanoTemper was used.

Published data for the binding affinity of the 4.5S RNA to Ffh ranges from \(K_d\) values of 7 pM (Siu et al., 2007) to 3 nM (Jagath et al., 2001), depending on the used methods and concentrations of monovalent and bivalent cations in the binding buffer (Batey and Doudna, 2002). I decided to use binding conditions as described for a filter binding assay (Batey et al., 2000) where a \(K_d\) of ~ 40 pM for the wild type 4.5S RNA binding to the *E. coli* Ffh protein was determined under physiological buffer conditions (20 mM Hepes-KOH pH 7.5, 100 mM KCl, 10 mM MgCl\(_2\), 0.5 mM EDTA). First the lowest possible concentration of Ffh\(_{647\text{mal}}\) that gives an adequate fluorescence signal (> 200 counts) was determined, since the labeled binding partner should be used in a concentration range around the expected \(K_d\). Solutions of Ffh with final concentrations of 2, 4 and 8 nM were tested, respectively (Fig. 3.34). The raw fluorescence counts of the different Ffh solutions were measured by the cap scan tool in premium coated capillaries (NanoTemper), without adjacent MST measurement.
RESULTS

Fig. 3.34 Cap scan of Ffh$_{647\text{mal}}$.
Raw fluorescence signals for Ffh were measured with concentrations of 2 nM (capillary positions 11+12), 4 nM (13+14) and 8 nM (15+16) in premium treated capillaries. All samples showed symmetric peaks and a concentration-dependent fluorescence increase. The minimum of 200 counts was reached in the 4 nM Ffh sample.

The Ffh$_{647\text{mal}}$ protein in MST$_{RNA}$ buffer (20 mM Hepes-KOH pH 7.5, 100 mM KCl, 10 mM MgCl$_2$, 0.5 mM EDTA, 0.05% Tween-20) showed a symmetric peak and a linear, concentration-dependent fluorescence signal increase (Fig. 3.34). This demonstrates that the protein was stable in this buffer and did not stick to the capillary surface. A 4 nM Ffh$_{647\text{mal}}$ sample had a raw fluorescence intensity of 200 counts and represents the lowest possible concentration to measure MST. Consequently, the labeled Ffh stock solution for the binding assay was adjusted to 8 nM with MST$_{RNA}$ buffer, resulting in a final concentration of 4 nM Ffh in the dilution series. Computer generated simulation (NanoTemper concentration-finder: Kd-fit) of the binding curve with 4 nM labeled molecule and a Kd of 40 pM suggested a ligand concentration of 200 nM for the dilution series. According to that, a stock solution of 400 nM 4.5S RNA in MST$_{RNA}$ buffer was prepared, producing final ligand concentrations ranging from 6.1 pM to 200 nM after mixing with one volume of the Ffh$_{647\text{mal}}$ stock solution. The mixed two binding partners were incubated for 20 min at room temperature, loaded into premium coated capillaries and analysis of their thermophoretic mobility was performed. Thermophoresis was measured using a Monolith NT.115 instrument (2.17.3) and data of two independent experiments were analyzed using the signals from Thermophoresis and T Jump. A significant change in MST response of the labeled Ffh in the presence of 4.5S RNA was observed that was attributed to the assembled SRP complex (Fig. 3.35). The dissociation constant was calculated to be 0.56 nM with a Kd confidence of 0.17 and a proper response amplitude of 14.3 (2.17.5). The standard error of regression was calculated to 0.55.
Fig. 3.35 Ffh-4.5S RNA assembly. In the MST experiment the concentration of the Ffh-labeled molecule was kept constant (4 nM), while the concentration of the non-labeled binding partner (4.5S RNA) was titrated between 6.1 pM and 200 nM. Thermophoresis was measured with 95% LED and 20% MST power in premium coated caps. (A) Original time traces display the affected thermophoretical mobility of labeled Ffh upon binding of 4.5S RNA. (B) A $K_d$ of 0.56 nM +/- 0.17 nM was determined for this interaction employing the Thermophoresis + T Jump signal for data analysis (2 independent measurements, error bars represent the standard deviation).

From a comparison of various interaction techniques, it is known that surface-based methods tend to overestimate the affinity of an interaction (Jecklin et al., 2009). With this in mind, the affinity of a $K_d = 0.56$ nM determined with MST for the 4.5S RNA binding to Ffh is in good agreement with previous studies. Additionally, it has to be mentioned that strong binding affinities in the pM range, as this is the case for SRP reconstitution, could not be exactly measured with the Monolith NT.115 instrument due to the required high fluorescence intensity. Thus, the observed $K_d$ of $\sim 500$ pM in this study, being about 10-fold higher than the expected dissociation constant (Batey et al., 2000), seems to represent a properly reconstituted SRP ribonucleoparticle under the assay conditions used in this study. Thus, the reconstituted SRP is suitable for further binding studies with the SRP signal sequence peptides.

3.4.3 Binding studies of SRP signal sequences to SRP using microscale thermophoresis

To explore the affinity of the wild type KdpD signal sequence N22-48 and the two mutant peptides W3A and 3Q (Fig. 3.29) to SRP, MST measurements were performed using the reconstituted SRP (Fig. 3.35) as the fluorescent binding partner at a constant concentration of 20 nM. The 27 amino acid comprising KdpD peptides with molecular weights of $\sim 3$ kDa were synthesized with a purity of $> 95 \%$ by GENOSPHERE Biotechnologies (2.5) and modified with an N-terminal amidation and a C-terminal acetylation. Prior to MST the lyophilized peptides were dissolved in ddH$_2$O by sonication to
a final concentration of about 1 mM. Based on fluorescence anisotropy measurements with SRP and the ΔEspP signal sequence (Bradshaw et al., 2009) a moderate, not too strong binding affinity of SRP to the signal peptides in the lower µM K_d range was expected. Therefore, the starting concentration of the KdpD peptides was set to 40 µM in the first titration step and was serially diluted to a lowest concentration of 1,2 nM.

First, the binding of N22-48 to SRP was tested in three different buffers whereas the SRP sample constantly was kept in RNA buffer (20 mM Hepes-KOH pH 7.5, 100 mM KCl, 10 mM MgCl_2, 0.5 mM EDTA, 0.05% Tween-20) that was also used for the SRP reconstitution measurements. The N22-48 peptide stock solutions were prepared in RNA-buffer, MST optimized buffer (50 mM Tris pH 7.4, 150 mM NaCl, 10 mM MgCl_2, 0.05% Tween-20) and Ffh- buffer (20 mM Hepes pH 7.4, 200 mM NaCl, 10 mM MgCl_2, 10 mM KCl, 0.05% Tween), respectively. The binding components were mixed, incubated for 20 min on ice and loaded into premium treated capillaries and analysis of their thermophoretic mobility was performed. A significant change in the MST response of the labeled SRP in the presence of the N22-48 peptide was observed in all three buffer types (Fig. 3. 36 A). However, the calculated dissociation constants differ between the single K_d fitted curves (Fig. 3.36 B). The binding affinities of N22-48 to SRP in RNA- and MST-buffer (green and orange data points) were in the same range with calculated K_d values of ~ 320 nM and ~ 480 nM, whereas the binding in Ffh- buffer was reduced significantly with a 3 fold higher K_d of ~ 1200 nM (purple data points). Comparison of the three fitting results showed that the highest binding affinity for N22-48 to SRP in RNA- buffer is represented also by the largest response amplitude, while with increasing K_d values, the response amplitudes decrease (Fig. 3.36 C). Nevertheless it was possible to perform a K_d fit of the averaged measurements resulting in a calculated dissociation constant around 390 nM with a K_d confidence of +/- 140 nM and a proper response amplitude of 9 (Fig. 3.36 D).
SRP binding to KdpD N22-48 in MST measurements. Concentration of the fluorescently labeled SRP molecule was kept constant at 20 nM and the concentrations of the titrated N22-48 peptides were varied between 1.2 nM to 40 µM. Thermophoresis was measured with an excitation power of 50 % and 20% MST power. (A) Original time traces of three independent MST measurements of SRP:N22-48 binding in different buffer systems, showing some bumpy regions in the curves at higher ligand concentrations. (B) Single Kd fitted curves of N22-48 titration in RNA- buffer (green), MSTopt- buffer (orange) and Ffh- buffer (purple). (C) Overview of the fitting results of the binding events in different buffer systems. Detailed explanation is given in chapter 0. (D) Kd of 388 nM +/- 142 nM was determined for the averaged Kd fit of the SRP:N22-48 interaction in different buffer compositions using the Thermophoresis + T Jump signal for data analysis (3 independent measurements, error bars represents the standard deviations).

The lower the standard error of regression or root mean square error (RMSE) is, the better the experimental data are represented by the fitted curve. For the averaged fitting result the RMSE constituted 1.1, indicating that the calculated Kd only roughly represents the actual binding affinity. Of course this could be due to the use of different buffer compositions of the averaged measurements, but the deviation between the signals in the low ligand concentration samples (1-100 nM N22-48) showed a good agreement of the independent measurements in the unbound state of the SRP protein. Therefore it is assumed that the bumpy signals in the bound state curves are caused by the high ligand concentrations, presumably inducing the aggregation of the SRP protein during the binding event to some extent. To define an exact Kd for this binding event, further assay optimization would be necessary. However, a distinct and specific binding of SRP to the N-terminal KdpD wild type peptide N22-48 was detected in all MST measurements under...
these conditions. Also, the binding affinity of ~ 400 nM calculated by the MST measurements are in good agreement with recent equilibrium measurements of SRP binding to RNCs stalled with nascent protein chains of 35 or more amino acid residues in length of the leader peptidase, indicating a very tight binding of ~ 1 nM, while K_d values around 50 nM were observed for vacant 70S ribosomes (Bornemann et al., 2008). Since I did not use ribosomes carrying the nascent peptide chain but used the free signal sequence peptide, it was expected that the binding affinity was lower due to the absence of the SRP binding sites on the ribosome. My data are therefore more comparable to the calculated binding affinity of ~ 1,5 µM of the ∆EspP signal peptide to SRP by fluorescence anisotropy (Bradshaw et al., 2009).

A crucial question is how the signal sequence is recognized and bound by SRP as no obvious consensus motifs are known for this process. SRP-dependent N-terminal signal sequences are highly diverse in amino-acid composition and length, but they all contain a core of at least eight consecutive hydrophobic amino acids that act as the major determinant for recognition by SRP (Valent et al., 1997 & 1998; Martoglio and Dobberstein 1998; Hegde and Bernstein 2006, Janda et al., 2010). In addition, the N-terminal region of a signal sequence typically contains positively charged residues with an as yet unknown function, suggesting a combination of hydrophobic interactions and electrostatic contacts during recognition and binding of SRP to signal sequences. To investigate the specificity of the SRP signal sequence of KdpD, the mutant W3A and 3Q peptides were tested of their ability to bind SRP under the same conditions as described for the wild type peptide, respectively. For the dilution series, 80 µM stock solutions of W3A and 3Q in MST_opt buffer (50 mM Tris pH 7.4, 150 mM NaCl, 10 mM MgCl_2, 0,05 % Tween-20) were prepared and mixed with 40 nM reconstituted SRP. The mixed binding components were incubated for 20 min on ice and MST measurements were performed in premium coated capillaries. Since the MST curves started to get bumpy after 10 s of thermophoretic movement, the affinity analysis was performed using the early MST signal of manually evaluated settings with start and end positions -1 s and 0 s for the cold region and 4.51 s / 5.51 s for start / end of the hot region (Fig. 3.37 A). This approach raised the K_d of the triplicate N22-48 measurement from ~ 390 nM (Fig. 3.36 D) to ~ 472 nM (Fig. 3.37 B) and also reduced the standard error of regression from 1.1 to 0.57 (Fig. 3.37 C), indicating that the manual K_d fitting represents more reliable affinity data for this assays than the Thermophoresis + T Jump evaluation. Both mutant peptides showed a change in MST response upon binding to SRP similar to the wild type peptide N22-48, however with reduced binding affinities. The calculation revealed a K_d of ~ 1,2 µM for 3Q binding and a
10-fold higher dissociation constant of ~ 5,4 µM, compared to the wild type peptide Kd, for the W3A mutant (Fig. 3.37 B).

**Fig. 3.37 Interaction of SRP with KdpD signal sequence mutants.** Concentration of the SRP-labeled molecule was kept constant at 20 nM and the concentration of the titrated KdpD peptides varied between 1.2 nM to 40 µM. Thermophoresis was measured with an excitation power of 50 % and 20 % MST power. (A) Original time traces monitor a total of 9 independent MST measurements of SRP:N22-48 (n=3); SRP:W3A (n=2); SRP:3Q (n=2); Ffh:N22-48 (n=2). All traces gave bumpy curves after 10 s of thermophoresis. Manual cursor settings for Kd evaluation are indicated. (B) Reduced binding affinities were measured for 3Q (Kd ~ 1,2 µM) and W3A (Kd ~ 5,4 µM) compared to the wild type N22-48 peptide (Kd ~ 0,5 µM). Additionally, MST was measured for N22-48 binding to labeled Ffh protein without 4.5S RNA reconstitution (brown curve), demonstrating the loss of binding ability to the SRP signal peptide. The error bars represent the s.d. of each data point. (C) Overview of the fitted binding events of the different KdpD peptides in MST opt buffer.

Strikingly, substitution of the three positive amino acid residues in the N-terminal part of the signal peptide with glutamine in the 3Q mutant increased the Kd value for SRP binding only two fold. Also, it did not affect the binding of SRP as much as for the W3A mutant exhibiting an altered Walker A motif in the core region. This observation is in good agreement with earlier in vivo studies with ΔEspP signal sequence mutants, indicating that single point mutations that slightly change the hydrophobicity of the core region profoundly affect SRP recognition, whereas mutations that alter the charge of the N- region have only marginal effects (Lee and Bernstein, 2001; Peterson et al., 2003).
To ensure that the interaction detected between SRP and the various signal peptides truly represents specific binding, the labeled Ffh protein was used instead of reconstituted SRP to further analyze the binding to N22-48. The absence of the 4.5S RNA leads to an inactive signal recognition particle and raised the calculated $K_d$ from $\sim 0.5 \, \mu M$ to more than $17 \, \mu M$, without reaching a saturated bound state within the peptide concentration range that was used for this assay (Fig. 3.37 B). This demonstrates the importance of a functional, fully assembled SRP and hence strongly suggests a specific SRP – substrate interaction in these measurements. Interestingly, the response amplitude changed from 5 to 10 for N22-48 and the 3Q mutant (R22Q, K24Q, K26Q), respectively. This is due to the fact that the thermophoretic amplitude not only depends on the size of the complex but contains contributions from both charge and structure. By replacing 3 charged arginine/lysine residues with structurally different and uncharged glutamine residues, a strong influence on the biophysical and thermodynamic features of the complex is very likely and can consequently be observed in the thermophoretic amplitude. Another explanation for the differing amplitudes could be that the conformational changes of SRP induced by binding of a proper signal sequence (Janda et al., 2010; Hainzl et al., 2011) are altered due to the mutant signal peptides and result in incorrect folded SRP:signal-peptide complexes.

In conclusion, it was shown that the SRP recognition sequence of KdpD encompasses the amino acid residues 22 – 48 at the very N-terminal region of the large cytoplasmic domain of the KdpD protein since a direct interaction of the isolated peptide N22-48 with purified SRP was demonstrated via microscale thermophoresis. A dissociation constant of $\sim 500 \, nM$ was calculated for this binding event (Fig. 3.36). In addition, MST measurements of the SRP signal sequence mutants 3Q and W3A revealed a two-fold and 10-fold reduced binding affinity to SRP (Fig. 3.37), respectively. This suggests a direct influence of the amino acid composition in the signal peptide on its SRP binding affinity, whereby the detailed analyses of additional signal sequence mutants are required to understand the importance of particular amino acid residues and their position within the sequence.
4 DISCUSSION

In this study, the molecular mechanisms of co-translational membrane insertion of inner membrane proteins in *E. coli* were investigated in detail. Membrane proteins comprise about one third of the bacterial proteome, many of them being involved in essential physiological processes such as the membrane-located electron transport in the respiratory chain, generation of ATP and the proton motive force, uptake of nutrients or sensory mechanisms. To functionally assemble the wide variety of membrane-located proteins and protein complexes, bacterial cells have evolved several different pathways using various membrane insertion and translocation machineries. For the insertion of inner membrane proteins in eubacteria, the membrane translocase SecYEG and the membrane insertase YidC, as well as the signal recognition particle (SRP) targeting system are of central importance. Insertion of integral membrane proteins into the lipid bilayer usually takes place co-translationally at membrane-associated ribosomes (Luirink and Sinning, 2004; Egea et al., 2005; Halic and Beckmann, 2005). Not fully clear, however is, how SRP recognizes the specific features of nascent chains of hydrophobic substrate proteins and how the ribosome nascent chain complexes (RNCs) are transferred to the adequate translocase or insertase, especially in the YidC-only insertion pathway. YidC is the prokaryotic member of the YidC/Oxa1/Alb3 protein family (Dalbey and Kuhn, 2004) and works either independently inserting a subset of specific inner membrane proteins or it functions cooperatively with the Sec translocon (Dalbey et al., 2011).

In the first, main part of this thesis, the function of the prolonged C-terminal putative ribosome-binding domain in certain YidC proteins and its actual role in ribosome-interaction and in SRP-mediated membrane insertion in Gram-negative bacteria was analyzed. Evidence is provided for a possible (ancestral) function of this extended tail by biochemical as well as structural approaches. The second part focuses on the interaction of the signal recognition particle with the signal sequences. Isolated mutant signal sequence peptides were used to determine the specificity of SRP recognition in proteins. The interaction studies were established in an *in vitro* system with purified SRP and binding affinities were determined via microscale thermophoresis (MST), a new technique that enables immobilization-free, in-solution kinetic measurements of biomolecular interactions. Taken together, this study contributes to the understanding of the molecular mechanisms of co-translational membrane protein biogenesis in bacteria and also highlights evolutionary aspects of this complex vital cellular process.
4.1 C-terminally extended YidC variants complement *E. coli* YidC *in vivo*

The fundamental findings of this thesis are provided by analyzing the engineered YidC chimera, comprised of the insertase core domain of *E. coli* YidC and the C-terminal putative ribosome binding domains (CTDs) of marine, Gram-negative YidC homologs (Fig. 3.2), resulting in the C-terminally extended YidC variants YidC-Rb (CTD of *R. baltica* YidC) and YidC-Oa (CTD of *O. alexandrii* YidC). The C-terminal regions of YidC from *Rhodopirellula baltica* (80 aa; +20), a bacterium belonging to the planctomycetes, and from the α-proteobacterium *Oceanicaulis alexandrii* (99 aa; +18), are extended compared to the *E. coli* YidC insertase. The marine CTDs have numerous positively charged amino acid residues and show similarities to the C-terminal region of the mitochondrial Oxa1 protein (Fig. 3.1), as well as to the homologous Alb proteins in chloroplasts and YidC2 found in Gram-positive bacteria. In contrast to that, the C-terminus of *E. coli* YidC (Ec-YidC) is substantially shorter and only moderately positively charged (16 residues; +8).

At first, the ability of the chimeric YidC-Rb and YidC-Oa proteins to complement the *E. coli* YidC was tested *in vivo* (Fig. 3.3 C). Both YidC chimera promoted growth of *E. coli* MK6 cells when the wild type chromosomal YidC was depleted, with the same efficiency as the positive control (Fig. 3.3 A). Accordingly, this leads to the assumption that the elongated C-terminal tail on YidC does not interfere with its insertase function both in the YidC-only and in the Sec-YidC pathways of *E. coli*. This is consistent with the finding that Oxa1 of *Saccharomyces cerevisiae* (van Bloois *et al.*, 2005) as well as YidC2 of *Streptococcus mutants* (Dong *et al.*, 2008) and *Arabidopsis thaliana* Alb3 (Jiang *et al.*, 2002), all containing alike C-terminally extended tails, are also able to rescue a YidC deficient phenotype in *E. coli*. However, it was shown that Oxa1 is unable to take over the Sec-associated function of YidC, while a direct association of Alb3 with SecY has been demonstrated in *A. thaliana* and a functional complementation for both pathways was observed (van Bloois *et al.*, 2005). Importantly, the substitution of *E. coli* YidC was only successful when parts of its N-terminal region (aa 1-57 fused to Alb3; aa 1-247 fused to Oxa1 and YidC2), that contains an uncleaved signal sequence, had been fused to the respective homologous proteins. Hence, while the C-terminally extended YidC chimera used in this study could almost fully replace the *E. coli* protein, the mitochondrial and chloroplast, as well as the Gram-positive insertase homologs show differences in their activity and efficiency of complementation depending on growth and species background. This suggests that the complete functional spectrum of each insertase is optimized for the specific organism and environment in which they reside. A striking and explicit example
for this differentiation is the finding that Oxa1 cannot facilitate Sec-associated insertion, probably due to the fact that mitochondria lack a Sec translocon (Glick and van Heijne, 1996) but instead Oxa1 itself most likely forms an intricate homo-oligomeric insertion complex in the inner mitochondrial membrane (Nargang et al., 2002). The property of Oxa1 to interact directly with the Sec translocon might have been lost, concomitantly with the Sec-translocase, during evolution of mitochondria from endosymbiotic bacteria. In reciprocal experiments published by Preuss et al. (2005) it was shown that YidC expressed in mitochondria can functionally replace Oxa1, only when the C-terminal ribosome-binding domain of Oxa1 had been appended onto YidC. It would be interesting to test whether the C-terminally added ribosome-binding domains of the marine YidC homologs of the chimeric proteins YidC-Rb and YidC-Oa could substitute for Oxa1 in mitochondria.

4.2 The C-terminal tails of marine YidC homologs facilitate ribosome-binding independently of a nascent polypeptide chain

The YidC homolog Oxa1 of *S. cerevisiae* binds via its C-terminal domain to the large subunit of the mitochondrial ribosome in proximity to the protein exit tunnel (Szyrach et al., 2003; Jia et al., 2003; Gruschke et al., 2010). Besides, the C-terminal tail of *S. mutants* YidC2 has been shown to be crucial for direct ribosome interaction (Funes et al., 2009; Wu et al., 2013). Therefore, the two C-terminal YidC extensions from marine Gram-negative bacteria (*R. baltica* & *O. alexandrii*) were tested for their ability to function also as ribosome binding domains. Ribosome binding was analyzed by in vitro ribosomal pull down assays (2.12) with purified YidC proteins and isolated *E. coli* ribosomes (Seitl, 2010; Maguire et al., 2008). A purification protocol for the YidC-chimera YidC-Rb and YidC-Oa, as well as for the *O. alexandrii* YidC homolog Oa-YidC, was successfully established (2.10.2), resulting in sufficient protein amounts and purity for the ribosome-binding assays (Fig. 3.8, input). Both chimeras with the hydrophilic C-tails of the marine YidC homologs fused to the *E. coli* YidC moiety (YidC-Rb and YidC-Oa) as well as the *O. alexandrii* YidC homolog Oa-YidC efficiently bound to *E. coli* ribosomes (Fig. 3.8 B-D). The results also verify that *E. coli* YidC (Ec-YidC), without an extended C-terminal tail, binds very poorly to empty ribosomes (Fig. 3.8 A).

However, contradictory data are published for the ribosome binding activity of Ec-YidC and the role of its short C-terminal tail in this regard. A cross-linking study published by the group of Koch in 2012 suggested that Ec-YidC interacts with both, empty and
translating ribosomes, while the interaction sites appear to be distributed over the large and the small subunits (Welte et al., 2012). However, the affinity of YidC to purified ribosomes was very low and the specificity could not be clearly shown in the study by Welte et al. (2012). On the other hand, in binding studies of a detergent-solubilized Ec-YidC to translating ribosomes, visualized by cryo-EM, the C-terminal region of YidC was assigned to be crucial for this interaction (Kohler et al., 2009). Corroborating the results presented in this thesis, the detailed analysis of the interaction between Ec-YidC and ribosomes by fluorescence correlation spectroscopy (FCS) delineated by Kedrov et al., demonstrated that under physiological buffer conditions (pH 7.4) Ec-YidC does not bind to non-programmed, so called empty, ribosomes (Kedrov et al., 2013). Furthermore, they showed that the addition of just six histidine residues to the C-terminus of YidC is sufficient to bind YidC to ribosomes particularly upon low pH (pH 6.2) when the histidine residues are positively charged. Since C-terminally histidine-tagged YidC protein was used by Kohler et al. (2009) for the cryo-EM reconstitution of the Ec-YidC:ribosome complex under acidic conditions, this would be an explanation for their observation of a strong and C-terminal tail-dependent binding of Ec-YidC to empty and translating ribosomes. Thus, these experiments may not reflect naturally occurring interactions. Hence, the binding assays in this present study were performed under physiological buffer conditions (pH 7.4) with N-terminally histidine-tagged YidC proteins and empty ribosomes, confirming the results of Kedrov et al. (2013) with YidC not binding empty ribosomes. Interestingly, membrane-reconstituted Ec-YidC without any artificial affinity tag was able to bind to ribosome nascent chain complexes that expose the N-terminal transmembrane segment of F_oC, a natural substrate of Ec-YidC, at physiological pH (Kedrov et al., 2013). While the mitochondrial Oxa1 protein binds permanently to ribosomes even in the absence of nascent chains and therefore tethers mitochondrial ribosomes perpetually to the membrane (Szyrach et al., 2003; Jia et al., 2003), the lack of an interaction between Ec-YidC and non-programmed ribosomes probably provides a more flexible targeting of RNCs to the membrane insertase, depending on the emerging nascent substrate chain. Further studies are necessary to elucidate the specificity of YidC to nascent chains and the involvement of additional factors, particularly the signal recognition particle and its receptor. The SRP components might further contribute to the YidC-nascent chain interaction and support the discrimination between YidC-dependent and -independent substrates in vivo. The C-terminal region of Ec-YidC is not essential for its activity as a truncated variant fully complements the insertase function in vivo (Jiang et al., 2003). Concordantly, Kedrov et al. (2013) showed that the Ec-YidC:RNC binding is not completely abolished upon removal of the YidC C-terminal tail, suggesting that additional docking sites on Ec-YidC also contribute to ribosome binding. Recently Gen et al. (2015)
stated that the cytosolic loop C2 of Ec-YidC ensures stable docking of a translating ribosome to YidC while modifications in this region lead to enhanced ribosome dissociation. Remarkably, simultaneous dual deletions within the C2 loop and at the C-terminus of Ec-YidC abolished the insertion of the YidC-only substrates F_oC and MscL, underlining the important role of a direct YidC:RNC interaction for co-translational protein insertion.

Notably, the in vitro binding studies with empty 70S ribosomes in this study here clearly showed a strong affinity of the YidC hybrid protein with C-terminal extension sequences from the marine bacteria to E. coli ribosomes without the need of an emerging polypeptide chain, in contrast to the wild type Ec-YidC. Consequently, this leads to the assumption that the elongated, charged C-terminal tails of the marine YidC homologs from R. baltica and O. alexandrii functions as ribosome binding domains. This specific binding is therefore crucial for permanent association of the YidC homologs to ribosomes independent of their translational state and similar to the essential binding function of the C-terminal Oxa1 domain to mitochondrial ribosomes.

4.3 Visualization of the ribosome-bound insertase complex

In order to understand and accurately depict how YidC inserts proteins in a co-translational manner into the inner membrane of E. coli, it is essential to solve the structure of YidC bound to substrate specific RNCs at high resolution. The structural information for Ec-YidC was then limited to an X-ray structure of the large periplasmic domain (Oliver and Paetzel, 2008; Ravaud et al., 2008), a 10 Å 2D projection map of the membrane-integrated Ec-YidC dimer (Lotz et al., 2008) and a low-resolution (14.4 Å) cryo-EM reconstruction of an Ec-YidC:RNC complex (Kohler et al., 2009). None of those structures, though, had enough performance to answer for instance the question about the functional oligomeric state or to give further insights how YidC acts at a molecular level. To address these questions, we performed an advanced structural analysis of YidC-ribosome complexes using high-resolution cryo-EM with the C-terminally elongated YidC-chimera YidC-Rb. Due to the extension of Ec-YidC with the C-terminal ribosome binding domain, that drastically increases the ribosome affinity in vitro (Fig. 3.8), we were able to isolate stable YidC-Rb:RNC complexes (2.13). Those complexes were visualized using high-performance cryo-EM and the reconstruction of the complex could be refined to 8.6 Å (2.14). This improved resolution now allows a more detailed interpretation of the structural and functional features of the YidC-insertase complex. The TnaC-stalled ribosomes carried nascent polypeptide chains encompassing the first two TM segments of MscL, a
native YidC substrate in *E. coli* (Fig. 3.10). The path and position of the nascent chain within and at the ribosome can be traced from the peptidyl transferase center (PTC) through the ribosomal exit tunnel into an extra density at the tunnel exit representing YidC-Rb (Fig. 3.13). Although our overall resolution was in the subnanometer range, the alpha-helical secondary structures were barely resolved in the YidC density, most probably due to some flexibility in this region (Fig. 3.12). Interestingly, the presence of the C-terminal extension increases the affinity of YidC to the ribosome without changing its overall ribosome interaction mode. The spatial position and topology of the YidC-Rb chimera in this study agreed well with the position of Ec-YidC observed in the cryo-EM structure from an earlier study of ribosome-bound *E. coli* YidC (Kohler *et al.*, 2009). Moreover, we observed a similar interaction pattern between our C-terminally extended YidC-Rb and the ribosome. We could resolve in molecular detail the contact sites of YidC-Rb to the ribosome: The helix H59 of the 23S rRNA showed the strongest contact to YidC-Rb followed by close interaction with the two ribosomal proteins L24 and L29. Surprisingly, we observed a relatively weak connecting density with L23, the main contact site for SRP and TF, and also for a proposed contact derived from the lower resolution structure of the Ec-YidC-RNC with a F0C nascent chain (Kohler *et al.*, 2009). However, the ribosomal proteins L24 and L29, together with L23, also surround the ribosomal tunnel exit and provide binding sites for diverse factors involved in co-translational processing, folding, targeting and membrane insertion of nascent chains (Fig. 1.3). Since an enhanced interaction interface of YidC-Rb to the ribosomal protein L29 was observed (Fig. 3.11 C and D), compared to the Kohler *et al.* (2009) structure, the YidC-Rb protein as well as the Ec-YidC were tested for their ability to interact directly with isolated ribosomal proteins in *in vitro* pull down assays (2.15). Indeed, L29 only co-eluted with YidC-Rb (Fig. 3.16, lower panels), whereas L24 co-eluted both with Ec-YidC and in similar amounts as with YidC-Rb (Fig. 3.16, upper panels). Thus, L24 seems to be a major contact site for the YidC-insertase core domain, while the enhanced affinity of YidC-Rb to the ribosome is at least partially caused by the interaction between the C-terminal *R. baltica* YidC tail domain with the ribosomal protein L29, suggesting that the C-tail of YidC-Rb faces towards the L29 moiety of the ribosome.

Recently, Wickles *et al.* (2014) reconstituted YidC-Rb with RNCs exposing the first TM helix of F0C and subjected the purified complex to cryo-EM and single particle analysis to a resolution of ~ 8 Å. In contrast to our first cryo-EM structure of ribosome-bound YidC-Rb (Seitl *et al.*, 2014), it was now possible to separate the weaker electron density of the detergent micelle from that of the YidC-Rb protein moiety (Fig. 4.1 B). Furthermore, Wickles *et al.* (2014) calculated a structural model of *E. coli* YidC via the intramolecular
co-variation analysis that could be docked in a distinct orientation into the cryo-EM structure of the YidC-Rb:RNC complex (Fig. 4.1 B and C). Shortly after, the X-ray structure of Ec-YidC was published and confirmed that the built Ec-YidC model (Wickles et al., 2014) appeared in good agreement with the experimentally solved molecular structure (Kumazaki et al., 2014-b). Interestingly, in both YidC-Rb:RNC densities an additional density which is aligned with the ribosomal exit tunnel, neighboring TM3, was found that could be attributed to the TM helix of the nascent MscL- or F0c- chain, respectively (Fig. 4.1 A and C). This hypothesis is supported by data from several independent studies which show that YidC substrates, and also nascent F0c chains (Wickles et al., 2014), can be crosslinked to TM3 (Klenner et al., 2008; Yu et al., 2008; Neugebauer et al., 2012; Klenner and Kuhn, 2012). Strikingly, at the same relative position nascent chains have been observed inside the SecY channel (Frauenfeld et al., 2011; Wickles et al., 2014). Independently of the nascent chain substrate used in our experiments, YidC-Rb revealed an almost identical interaction pattern on the ribosomal exit site (Fig. 4.1) for both cryo-EM reconstructions (Seitl et al., 2014; Wickles et al., 2014).

![Fig. 4.1](image_url)

**Fig. 4.1** Comparison of two YidC-Rb structures bound to RNCs carrying MscL (8.6 Å) – and F0c (8 Å) nascent chains, respectively. YidC-Rb revealed an almost identical interaction pattern on the ribosomal exit site for both cryo-EM reconstructions and both nascent chains enter the insertase core within particular proximity of TM3. (A) Density blot (bottom view) of YidC-Rb shows the entering of the nascent MscL chain (green) in the center of the YidC-Rb density (red) and the proposed TM domain arrangement (1-6; red), relative to the indicated ribosomal proteins L23, L24, L29 and the rRNA helix H59 (blue). The assumed position of the C-terminal ribosome-binding
domain (CTD) was assigned to an additional density near L29. Cytoplasmic loops C1 (47aa) and C2 (13aa) are schematically (dotted lines) included in the figure for orientation. (B) Close up view from within the membrane region of YidC-Rb bound to F_{OC}-RNC, highlighting the predicted contacts between the C1 and C2 loops of YidC and the ribosome, indicated by magenta spheres. The detergent micelle is shown in blue and the F_{OC} TM helix in pink. (C) Periplasmic view of the active ribosome-bound YidC model, with the YidC contour outlined in red. The polypeptide exit tunnel is indicated with an asterisk. Structure (A) was adapted from Seidl et al., 2014 and structure (B/C) from Wickles et al., 2014.

Although the C-terminally extended domain of YidC-Rb was excluded in the built Ec-YidC model (Fig. 4.1 B and C), we suggest that the additional density, located between the ribosomal proteins L23 and L29, accounts for the C-terminal ribosome-binding domain (CTD) of the marine *Rhodopirellula* YidC homolog (Fig. 4.1 A). As discussed above, YidC-Rb shows a much higher affinity to L29 than Ec-YidC in pull down assays (Fig. 3.16). Therefore, we assume that the C-tail of YidC-Rb faces the ribosomal rRNA helix H59, which is located next to TM6, up to the opposite site of the exit tunnel. There it contacts the ribosomal protein L29 next to TM1, stabilizing the whole insertase-ribosome complex. This model also explains why we observed in both YidC-Rb structures only weak contacts to the ribosomal protein L23, compared to the Ec-YidC:RNC structure (Kohler et al., 2009). It is possible that the large CTD, positioned by the high affinity to L29, shields the L23 majority and therefore prevents to a large extent the interaction with the YidC-insertase core. As additional contacts to the ribosome, Wickles et al. (2014) suggested residues in the positively charged C1 (Y370; Y377, contacts to rRNA helix H59) and C2 (D488, contact to L23) loops of YidC (Fig. 4.1 B). Indeed, mutations in these residues compromised the growth of YidC-depleted *E. coli* cells, emphasizing their functional significance. However, as discussed in chapter 4.2, Geng et al. (2015) probed the influence of the C1 and C2 loop regions in direct ribosome-binding efficiency and showed that only C2 contributes to YidC:RNC assembly, while the C1 loop is involved in other vital functions. Due to the local proximity of the C1 region to the nascent chain portal site in the YidC-Rb density (Fig. 4.1 A), it is feasible that the C1 loop directly contacts the emerging polypeptide chain and assists the correct folding of inserting membrane proteins into the lipid bilayer.

Another striking finding of the cryo-EM YidC-Rb:RNC structures was that only a single monomer of YidC was bound to the translating ribosome (Fig. 4.1; Seidl et al., 2014; Wickles et al., 2014). Until then, results from the early studies on YidC extracted from bacterial membranes suggested that the protein is present in both monomeric and dimeric forms (van der Laan et al., 2001; Heuberger et al., 2002). Also crystallographic analysis showed that YidC forms symmetric dimers in the membrane (Lotz et al., 2008). Additionally, the low resolution cryo-EM structure on a YidC:RNC-F_{OC} complex also
suggested that two copies of YidC bind to the ribosomal tunnel exit (Kohler et al., 2009). However, this view changed during the last years due to the YidC-Rb:RNC structures and also by the recently published X-ray structures of both Gram-positive and Gram-negative YidC proteins, clearly showing YidC to be in a monomeric state (Kumazaki et al., 2014-a & 2014-b). Kedrov et al. (2013) probed the oligomeric state of Ec-YidC in its free form and bound at the ribosome via fluorescence cross-correlation spectroscopy (FCCS). In this study it was shown that a single membrane-embedded YidC copy is sufficient to bind a substrate-translating ribosome and YidC only oligomerizes on the ribosome when YidC was applied in excess, non-physiological concentrations, compared to the amount of applied RNCs. Since a 10-fold excess of YidC was used to form YidC:RNC complexes in the previously cryo-EM reconstitution (Kohler et al., 2009), the corresponding structure likely represented a concentration-dependent oligomer of YidC. Taken together our observation of a single YidC-Rb protein bound to a 70S E. coli ribosome is in agreement with the recent literature, showing clearly that a single YidC copy is fully active as monomer, thus being the minimal and probably true-to-life functional unit for YidC-dependent, co-translational insertion of membrane proteins.

4.4 The C-terminal ribosome-binding domains of marine YidC homologs are intrinsically disordered

Since the C-terminal domain (CTD) of YidC-Rb was not resolved in the cryo-EM structure (Fig. 3.13), the native architecture of the CTDs from R. baltica YidC (RbCT) and O. alexandrii YidC (OaCT) without their insertase core domains was further investigated by CD-spectroscopy (2.16). Therefore, RbCT and OaCT encompassing 80 and 99 amino acid residues, respectively, were separately cloned and purified via an N-terminal His$_{10}$-tag to near homogeneity by Ni-NTA affinity and strong cation exchange chromatography (2.10.5). Similar to the isolated CTDs of Alb3 (A3CT; Falk et al., 2009), yeast Oxa1p (O1CT; Szyrach et al., 2003) and human Oxa1L (Haque et al., 2010), both RbCT and OaCT revealed far-UV CD spectra typical for unfolded proteins (Fig. 3.22) in aqueous solution under low salt conditions. The results obtained by CD-spectroscopy are consistent with the prediction of the intrinsically disordered nature of RbCT and OaCT. The computer models identified unstructured regions in both protein domains (Fig. 3.21 A and B), similar to the Alb3 C-terminal segment (Fig. 3.21 D). However, formation of an α-helical conformation was induced by the addition of trifluoroethanol (TFE) for all C-terminal insertase domains (Fig. 3.22 A and B; Szyrach et al., 2003, Falk et al., 2009; Haque et al., 2010). TFE is a co-solvent that is known to disrupt coiled-coil interactions in proteins leading to non-interacting single helices, but also stabilizes the helices in single
α-helical segments (Lau et al., 1984; Frère et al., 1995). In accordance, secondary structure prediction of the CTDs from various organisms using the PSIRED server, confirmed the presence of regions in these domains that could potentially form α-helices (Fig. 4.2 A), but they show differences in lengths and numbers of the predicted helical regions as well as in the structural motifs these regions are predicted to form. However, all CTDs compromise helices within the first 20 amino acid residues of the extended regions, followed by another second, larger α-helical region, starting approximately at the position of amino acid residue 40, except in the case of the S. mutans YidC2 CTD (Fig. 4.2 A; Y2CT). For the most extended A3CT, the C-tail of the chloroplast insertase homolog, two more regions were predicted to adopt a helical conformation. Alternative prediction analyses revealed coiled-coil interaction domains (COILS; Lupas et al., 1991) almost certainly in O1CT and in A3CT (Fig. 4.2 A; dark green boxes) and with less probability for the second helical motif in OaCT (Fig. 4.2 A; light green boxes). RbCT and Y2CT analyses rendered no bias to form coiled-coil motifs. However, helical wheel projections (RZ lab) of the predicted α-helical regions suggested the formation of single amphipathic helices in RbCT (aa 42-52), Y2CT (aa 10-20) and also for the first putative helical domain (5-19) in OaCT (Fig. 4.1 B).

Fig. 4.2 Secondary structure predictions of C-terminally extended insertase domains (CTDs). (A) Secondary structure prediction was generated computationally using the PSIRED server.
Several α-helices of diverse lengths are predicted for all CTDs, but the number and length of the α-helical regions vary among the different organisms (A3CT: CTD of Alb3 from A. thaliana; OaCT: CTD of YidC from O. alexandrii; O1CT: CTD of Oxa1 from S. cerevisiae; RbCT: CTD of YidC from R. baltica; Y2CT: CTD of YidC2 from S. mutans). Pink bars indicate putative helical regions and yellow arrows predict strand motifs. Amino acid residues predicted to form coiled-coil motifs (Lupas et al., 1991) are boxed in green (dark green: most likely; light green: less likely). Regions predicted to form amphipathic helices are boxed in red. (B) Helical wheel projections (RZ-lab) of the bacterial helical CTD regions suggest the formation of amphipathic helices for the second helix in RbCT, the first helix in OaCT (aa 5-19) and the sole helical region in Y2CT. The region encompassing amino acid residues 40-56 in OaCT is more likely to form a coiled-coil motif. Hydrophilic residues are presented as circles, hydrophobic residues as diamonds, potentially negatively charged as triangles, and potentially positively charged residues as pentagons. Hydrophobicity is color coded: the most hydrophobic residue is green, and the portion of green is decreasing proportionally to the declining hydrophobicity, with zero hydrophobicity coded as yellow. Hydrophilic residues are coded red with pure red being the most hydrophilic (polar) residue, and the amount of red decreasing proportionally to the hydrophilicity. The charged residues are marked in light blue.

Despite the low primary sequence homology observed for all CTDs (Fig. 3.17 B), it seems that the organellar insertase homologs favor a coiled-coil formation in their C-terminal extended domains, whereas the prokaryotic ribosome-binding domains generally encompass only shorter helical regions, presumably forming single amphipathic helices for their interaction with the ribosomes or the membrane environment. It is likely that the differences between the organellar and the prokaryotic ribosomes lead to significant alterations in the interaction of ribosomes from different organisms with the respective insertase homologs. These differences may reflect the individual paths trodden in evolving the distinct eukaryotic and the prokaryotic membrane translocale systems, respectively.

Although mitochondrial ribosomes developed from those of their bacterial ancestors, over the long period of eukaryotic evolution, the translational system in the organelle changed considerably (Smits et al., 2007). The catalytic core of the mitochondrial ribosome and the region around the polypeptide exit tunnel of the large subunit still resembles those of bacterial ribosomes, but many mitochondria-specific proteins have been added through the eras of evolution (Gruschke et al., 2010). Particularly in animals, the rRNA content has considerably decreased (Sharma et al., 2003; O'Brien, 2003). Interestingly, differences in the C-terminal domains of Oxa1 homologs from higher and lower eukaryotes, studied by Haque et al. (2010), revealed structural differences in the CTD from mammalian Oxa1L compared to the yeast Oxa1p. Also an alternative interaction pattern of the Oxa-homolog on mammalian mitochondrial ribosomes was found. However, even though these structural features may be organism specific, they obviously have sufficient relatedness to mediate the binding for example of yeast Oxa1 to E. coli ribosomes (Kohler et al., 2009), indicating an accommodative and conserved binding mode to some extent.

Interestingly, as demonstrated by Falk et al. (2009) via CD-spectroscopy, the intrinsically disordered C-terminal region of Alb3 folds upon binding to cpSRP43, its native chaperone-
like interaction partner in chloroplasts. In a similar experiment, it was investigated whether
the isolated C-terminal ribosome binding domain of YidC-Rb (RbCT) adopts an α-helical
conformation upon binding to the ribosomal protein L29, since a direct interaction of L29
with the C-terminal domain of YidC-Rb has been shown (Fig. 3.16 A). However, we did
not observe any conformational change and RbCT remained in its disordered formation
upon addition of the putative ribosomal binding partner L29 (Fig. 3.23 B). Alternative
explanations are possible for this observation: (I) the ribosomal protein L29 is not the sole
determinant for folding of the YidC-Rb C-tail, and other ribosomal components like the
rRNA helix 59 may induce instead or in cooperation with L29 the conformational change
of the C-terminal insertase domain. (II) The context of the full-length insertase protein as
well as (III) the membrane, particularly lipid interaction might also influence the structure
of the C-terminal domain. Especially for the formation of an amphipathic helix the
presence and binding of lipids may be important. In general, the intrinsic disorder of these
domains might be advantageous to provide orientational freedom and flexibility in
scanning for and binding to RNCs at the cramped cytoplasmic surface of the inner
bacterial membrane. Intrinsically disordered regions are associated with a broad
repertoire of biological functions including cellular control mechanisms and signaling. It is
proposed for the C-terminal domain of a potassium channel, that its disordered structure
might modulate the kinetics of channel activation (Magidovich et al., 2007). Similar
modulating functions are also proposed for the A3CT-cpSRP interaction (Falk et al.,
2009). Triggered folding of intrinsically disordered regions upon binding to their interaction
partners has been described for a number of physiologically relevant protein-protein
interactions and it is suggested to be a general regulatory mechanism (Wright and Dyson,
2009). At present it is unclear whether such a mechanism can also be featured for the
YidC:RNC assembly and further studies will be necessary to evaluate the binding-induced
conformational changes of the intrinsically disordered C-terminal ribosome binding
domains of the Gram-negative YidC homologs. Particularly a possible functional role of
this trigger yet stays to be elucidated.

4.5 Stable ribosome:YidC association can partially substitute for
an SRP-mediated targeting process in *E. coli*

Since the *in vitro* studies clearly showed an enhanced interaction of C-terminally
elongated YidC proteins with empty ribosomes (Fig. 3.8) and likewise a stable complex
formation with programmed RNCs (Fig. 3.11), I wanted to test a possible functional role of
the C-terminal YidC extensions in co-translational protein targeting *in vivo*. For that
purpose the membrane targeting of the SRP-dependent YidC-substrate protein MscL
(Facey et al., 2007) was analyzed as a GFP fusion protein by fluorescence microscopy (2.8). In addition, the proper membrane insertion of MscL was probed in radioactive pulse chase experiments via AMS gel shift assays (2.9) in the absence of a functional SRP pathway. To investigate whether an extended C-terminal region of YidC can compensate for a loss of SRP function, either the protein component of the signal recognition ribonucleoparticle Ffh or the SRP receptor protein FtsY was depleted (2.9.1) in MscL-plasmid transformed E. coli strains.

The AMS gel shift assays demonstrated that insertion of MscL in E. coli depends on the R. baltica tail of YidC when FtsY is depleted (Fig. 3.25, YidC-Rb). Similarly, coexpression of a YidC chimera extended by the C-terminal tail of Oxa1 resulted in an enhanced membrane insertion efficiency under FtsY-depleted conditions (Fig. 3.25, YidC-Oxa). Cellular localization of MscL as a GFP fusion protein showed that it accumulated as fluorescent patches mostly close to the cell poles when the receptor protein FtsY was depleted (Fig. 3.27 B). In accordance with results from the AMS assay, coexpression of YidC-Rb prevented the formation of these MscL-GFP aggregates and the fluorescence was found at the membrane (Fig. 3.27 C). In contrast, coexpression of Ec-YidC (Fig. 3.27 D) had no effect on the localization of MscL-GFP, indicating that only the C-terminal tail of YidC-Rb (or of Oxa1) can restore membrane targeting of MscL in the absence of FtsY. It should be kept in mind that under FtsY depletion conditions, Ffh is still present and possibly required under these conditions. When Ffh was depleted, the MscL-GFP fusion protein clearly aggregated at the cell poles (Fig. 3.26 B) and coexpression of YidC-Rb only partially restored membrane targeting (Fig. 3.26 D). Similar results were observed for the membrane insertion of MscL assayed by AMS derivatization in which coexpression of YidC-Rb or YidC-Oxa only marginally stimulated MscL insertion in Ffh-depleted cells (Fig. 3.24). Taken together, the extended C-tails of the YidC derivatives can replace mainly the FtsY receptor function but not the Ffh function. It is assumed that the presence of Ffh might still be required to keep the newly synthesized MscL protein in an insertion-competent form, a function that obviously cannot be fulfilled by YidC. Without a functional SRP-system the hydrophobic non-inserted MscL protein molecules most likely aggregate in the cytoplasm and are then prone to proteolytic degradation, which mostly occurs at the cell poles. In summary, I could show that the C-terminal ribosome-binding domain of YidC found in the marine Gram-negative bacterium R. baltica can partially substitute for the SRP receptor protein FtsY, in contrast to the native E. coli YidC but comparable to YidC-Oxa. Therefore, it is suggested that the C-terminally extended YidC-tail putatively acts as an adaptor for a specialized membrane-targeting function during the co-translational insertion of the YidC-only pathway.
This is in accordance with recent data that show a functional overlap of the SRP-machinery and the elongated C-terminal tails of YidC/Oxa1/Alb3 family members in co-translational protein insertion (Funes et al., 2009). However, also in post-translational insertion processes, a functional and spatial proximity of FtsY and YidC seems to be important for an efficient SRP-dependent membrane insertion of hydrophobic substrates (Robinson and Woolhead, 2013). The SRP pathway thus allows a dynamic, substrate-triggered membrane association of cytosolic ribosomes with the translocation machinery at the endoplasmic reticulum of eukaryotic cells as well as with translocases and insertases at the cytoplasmic inner membrane surface of prokaryotes. In addition to its role in ribosome recruitment to the membrane, the SRP ribonucleoprotein might shield hydrophobic sectors on nascent chains from adverse interactions, and thereby improve and regulate the ongoing synthesis of membrane proteins. Genome analyses of fully sequenced eukaryotes revealed that mitochondria generally lack an SRP system. Comparison of the different translational systems suggests that, during evolution of mitochondria, the universal SRP-mediated co-translational insertion system was replaced by an SRP-independent machinery that relies on direct binding of ribosomes to the membrane-embedded insertase Oxa1 (Preuss et al., 2005; Funes et al., 2009). This development was favored by the small number (13 polypeptides in humans / animals; 8 in yeast) of almost exclusively very hydrophobic membrane proteins that are synthesized by the mitochondrial translation apparatus. Presumably as a consequence of its specialization for the synthesis of hydrophobic membrane proteins, the translation machinery of mitochondria is closely coupled to the inner membrane (Liu and Spremulli, 2000; Ott and Herrmann, 2010). Permanent association of mitochondrial ribosomes to the inner membrane surface is mainly mediated by the C-terminal domain of Oxa1 (Jia et al., 2003; Gruschke et al., 2010) and by the ribosome receptor Mba1. Presumably, Mba1 does not mechanistically contribute to membrane insertion of the hydrophobic substrates, but ensures the positioning of the ribosomal polypeptide exit tunnel to the membrane-embedded insertase (Preuss et al., 2001; Ott et al., 2006). Mutants lacking both Mba1 and the CTD of Oxa1 show severe defects in the assembly of the respiratory chain (Ott et al., 2006). Similar to the observation that C-terminal ribosome binding domains on YidC can partially but significantly substitute an SRP-mediated targeting pathway in E. coli, Funes et al. (2013) showed that E. coli Ffh binds to the large subunit of mitochondrial ribosomes and could relieve, to some degree, the defect during double-depletion of the Mba1 ribosome receptor and the C-terminal ribosome-binding domain of Oxa1. Since yeast Mrp20, the homolog of L23 in bacterial ribosomes that is critical for the contact to Ffh, was identified as an interaction partner of Oxa1 (Jia et al., 2009) it is conceivable that Oxa1 and Ffh employ similar ribosomal binding sites that are in direct proximity to the
polypeptide exit tunnel. As Funes and co-workers were unable to detect direct contacts between Ffh and FtsY during their rescue experiments expressing the bacterial SRP components in mitochondria, it is assumed that prokaryotic Ffh and FtsY are not able to form a complex in mitochondria, presumably due to the absence of the functionally important 4.5S RNA (1.2.1). Therefore, it is suggested that expression of Ffh in mutant mitochondria lacking Mba1 and the Oxa1 C-tail, might reduce the risk of newly synthesized proteins aggregating irreversibly and increase the time window during which nascent chains can be productively integrated into the inner membrane (Funes et al., 2013). This conclusion is in good agreement with the observation that in Ffh-depleted E. coli cells MscL aggregated predominantly at the membrane, for the most part unaffected by the introduction of an additional ribosome-binding domain in YidC (Fig. 3.26 B and D).

Yet, in presence of Ffh, but under FtsY depletion, the targeting defect was rescued by the YidC C-terminal ribosome-binding domain for targeting (Fig. 3.27 C) and insertion (Fig. 3.25) of MscL. This underlines the importance of a chaperone-like function of the free, cytosolic SRP in addition to its targeting function in cooperation with its receptor FtsY.

Another important evidence for the specialized co-translational functions of particular members of the YidC/Oxa1/Alb3 family is that the development of C-terminal ribosome binding domains does not only occur in mitochondria, but also during evolution in different prokaryotic systems. In Gram-positive bacteria usually two YidC homologs coexist. The oral pathogen Streptococcus mutans for example, possesses one homolog with a ribosome-binding domain (YidC2) while the other one lacks it (Funes et al., 2009). Both paralogs have certain overlapping functions (Hasona et al., 2005; Funes et al., 2009), but also reveal strong differences conferred by the presence or absence of the C-terminal domain (Palmer et al., 2012). Interestingly, the phenotype of the deletion of YidC2 in this organism is strikingly similar to that of SRP pathway mutants, including stress-sensitivity and diminished genetic competence. Deletion of SRP is tolerated when YidC2 is present, while double mutants lacking both YidC2 and SRP components are not viable, even in the absence of environmental stressors. This suggests that YidC2 overlaps functionally with the SRP pathway and supports a co-translational membrane insertion pathway that does not rely on the SRP only (Hasona et al., 2005). Furthermore, it was shown that YidC2 expressed in yeast binds to mitochondrial ribosomes and promotes co-translational integration of proteins into the inner mitochondrial membrane in the absence of Oxa1. Reciprocally YidC2 can be functionally replaced by Oxa1 in S. mutans (Funes et al., 2009). Interestingly, E. coli YidC was not able to complement Oxa1 until the C-terminal tail of YidC2 had been appended to it.
The studies employed in mitochondria provide detailed, yet specialized knowledge for a system that lacks an SRP pathway, whereas my *in vivo* studies in *E. coli* gave new insights also on the molecular level on how C-terminal ribosome binding domains in general can assist co-translational protein insertion in an organism in which the SRP components are essential for viability. Depletion of the essential SRP components in *E. coli* bears drastic physiological consequences for a wide range of synthesized proteins, especially for the SRP-mediated insertion/translocation of proteins via the SecYEG translocase. Since targeting of by far the most inner membrane proteins to the Sec-translocase is mediated by the SRP pathway, it was not surprising that the C-terminally elongated YidC proteins were only able to compensate partially for the loss of the SRP components in *E. coli* cells. To evade this complexity the assays in this study were focused on the targeting and insertion of the MscL protein, a native SRP-dependent, but SecYEG-independent YidC substrate in *E. coli*.

Based on the current structural (4.3 and 4.4) and biochemical (4.1, 4.2 and 4.5) data a combined targeting and insertion model of YidC-substrates (e.g. MscL) is proposed involving the SRP, the nascent ribosome and the C-terminally extended YidC making the SRP-receptor FtsY dispensable (Fig. 4.3). For the initial step it is assumed that the nascent chain is recognized by SRP, either on free cytosolic RNCs (1a) or on membrane-associated ribosomes (1b). Possibly, a stable association of a certain pool of ribosomes at the membrane, irrespective of their translational state, is mediated by the C-terminally extended YidC due to additional contacts of its C-terminal ribosome binding domain and the flexible binding platform at the ribosomal exit tunnel (H59, L24 and L29). For both initial stages (1ab), the ribosomal protein L23 offers a transient binding site for the N-domain of SRP while the M-domain binds the emerging nascent chain, preventing the aggregation or misfolding of the growing polypeptide. The binding of SRP to an RNC with a proper SRP signal sequence repositions the NG domain and thereby clears the ribosomal protein L29 (Estrozi et al., 2011). This then allows an enhanced YidC:RNC interaction probably due to the expanded contact-interface of the C-terminally extended YidC domain to L29. As a result the proper positioning of the RNC onto the YidC insertase is promoted and SRP can be displaced from the RNC:insertase complex (2). Finally, the nascent chain is directly and co-translationally inserted by YidC-Rb into the inner membrane of *E. coli* (3), alike it is proposed for the SRP-FtsY pathway for Sec-independent YidC-substrates such as MscL.
Further studies are required to explicitly analyze a possible tripartite complex of RNC:SRP:YidC-Rb and also to dissect at which stage SRP is displaced from the ribosome. It also would be interesting to analyze if YidC insertases with an extended C-terminal tail in prokaryotes are predominantly occupied by ribosomes, similar to their mitochondrial homologs, or if the transfer of the RNC to the YidC insertase occurs mainly co-translationally from free cytosolic ribosomes in cooperation with other components, such as SRP.

4.6 Evolutionary aspects of the accessory YidC domains

The combination of a conserved membrane-spanning core domain with subtype-specific flanking regions makes the YidC/Oxa1/Alb3 protein family an interesting example to study the evolution of membrane proteins (Fig. 1.12; Yen et al., 2001; Funes et al., 2009; Zhang...
et al., 2009). Phylogenetic tree analysis reveal three separate clades corresponding to mitochondrial Oxa proteins, bacterial YidC proteins including chloroplast Alb3 family members, and the more distantly related archaeal DUF106 proteins (Borowska et al., 2015), respectively. These clades are further divided into several subbranches: YidC/YidC1/SpoIIIJ and YidC2/YqjG in bacteria, Oxa1 and Cox18/Oxa2 in mitochondria, Alb3 and Alb4 in plastids (Funes et al., 2011) and the highly divergent archaeal DUF106 proteins (Borowska et al., 2015). These subbranches presumably represent monophyletic subgroups of common ancestry and it is suggested that the different subgroups have been evolutionarily developed through three independent gene duplication events that occurred in Gram-positive bacteria, mitochondria and chloroplasts coincidently. Specialization then shaped the different subgroups and allowed the diverse insertase variants to acquire specific flanking regions, individual functional properties and their particular substrate specificities (Funes et al., 2009).

Among eubacteria, there is a clear structural and topological distinction between Gram-negative and Gram-positive YidC homologs. The Gram-positives have only 5 TM domains, similar to the organellar homologs and all the bacteria with two YidC homologs genes are generally Gram-positive bacteria. However not every Gram-positive bacterial species possesses two YidC genes (Zhang et al., 2009). There are some Streptomyces and Frankia genera where only one YidC gene is found on the genomes (Zhang et al., 2009). The YidC homologs within the group of the Gram-negative bacteria are easily recognized by the additional N-terminal TM domain and a large periplasmic loop P1 connecting TM1 and TM2. Thus, only these YidC proteins feature six transmembrane segments. The P1 domain varies in size from 232 residues in Thermotoga maritima to 500 in Rhodopirellula baltica. Intriguingly, the C-terminal regions of many marine Gram-negative bacteria are extended and show unique similarities to the C-terminal region of Oxa1 (Kiefer and Kuhn, 2007). One important question that arose during this work is why solely the marine prokaryotes, among them the model organisms of this thesis, R. baltica and O. alexandrii, exhibit C-terminal extensions on their YidC homologs in contrast to all the other Gram-negative bacteria. R. baltica and O. alexandrii both have homologous Ffh and FtsY genes on their chromosome, which most likely operate in a similar mode as the E. coli SRP components. However, particularly in the A-domain of the marine SRP receptor protein, there are some striking differences, compared to the E. coli homolog. Sequence alignments reveal that in contrast to the strongly conserved N and G domains of FtsY (Fig. 4.4 B), the A domain is poorly conserved (Pohlschröder et al., 2004). Moreover, several bacterial and archaeal FtsY homologs lack an A domain (Egea et al., 2004; Eitan and Bibi 2004; Focia et al., 2004). For the FtsY protein of Haloferax volcanii,
the N-terminal A-domain, a typically prokaryotic feature of the SRP receptor homolog, is namely present in this archaeon, but is not essential for growth (Haddad et al., 2005). In *E. coli*, two conserved helical stretches were identified which are proposed to constitute the lipid-binding sites of FtsY. One is located in the very N-terminal segment of the A-domain (aa 1-14) (Weiche et al., 2008), and a second one near the N-terminal end of the conserved N-domain of FtsY (aa 195-207), which has been shown to form an amphipathic helix (Fig. 4.4 A) (Parlitz et al., 2007).

**Fig. 4.4 Comparison of Gram-negative FtsY homologs.** (A) Domain structure of *E. coli* FtsY. The two conserved helices are indicated by their amino acid position (1-14; 195-207). Helical wheel representation shows a typically amphipathic conformation for helix 2. Residues are color coded like in Fig. 4.2. (B) Sequence alignment of the NG domains of the FtsY homologs from *E. coli* and *R. baltica* shows a moderate conservation, while the N-terminal A-domain of *R. baltica* is strikingly truncated (aa 1-45 in Rb-FtsY; aa 1-195 in Ec-FtsY).

Both helices together allow FtsY to execute an exceptionally stable interaction with the membrane (Braig et al., 2009). Additionally, the first-14-amino acid residues comprising region of the A-domain was shown to stabilizes the FtsY:SecYEG interaction. It has a highly positive net charge, compared to the remainder of the acidic A-domain (Weiche et al., 2008). Interestingly, sequence comparison of *E. coli* FtsY (Ec-FtsY) with the *R. baltica* homolog revealed an identity score of 43 % for their NG-domains, but the sequence alignment also revealed that the NG-domain of *R. baltica* FtsY (Rb-FtsY) already starts at amino acid position 45, compared to the nearly 200 amino acid residues long A domains of Ec-FtsY (Fig. 4.4 B) or the *H. volcanii* archaeal homolog. Secondary structure prediction of the first 45 amino acid residues of Rb-FtsY revealed a coil-conformation without exhibition of helical features. Conclusively, the A-domain of Rb-FtsY is either very small, operating without helical motifs or even is basically nonexistent.
Due to the fact that the C-terminal ribosome binding domain of the *R. baltica* YidC homolog (Rb-YidC) fused to Ec-YidC can mediate the co-translational insertion of an inner membrane YidC substrate protein even in the absence of FtsY (Fig. 4.3) and the Rb-FtsY seems to be an N-terminally truncated version of the SRP receptor, I hypothesize that the accessory C-terminal YidC domain of Rb-YidC possibly fulfills similar functions in the marine prokaryote as does the FtsY A domain during co-translation protein biogenesis in other organisms. Particularly, the local targeting and arrangement of RNCs at the membrane and the subsequent selective delivery of the nascent chain to the respective translocation or insertion machineries. This may be promoted and quality-controlled by involving either the A-domain or alternatively the C-terminal YidC domains in marine bacteria. Interestingly, *in vitro* experiments suggested that the A-domain of FtsY may also be involved in recruiting the SRP to the haloarchaeal cytoplasmic membrane (Lichi et al., 2004). This is also in accordance with the observation that SRP apparently plays an important role during the co-translational membrane insertion of MscL via the C-terminally extended YidC in the absence of FtsY (Fig. 4.3), underlining the functional similarities of the FtsY A domain and the C-terminally extended YidC domain. The marine planctomycete *Rhodopirellula* shows some unique features concerning the overall nature of its proteins and particularly its membrane proteins (Kiefer and Kuhn, 2007). The extraordinary features of exhibiting intracellular organelle-like compartments with sterol ladderane lipids like e.g. the anammoxosomes and the strikingly complex and huge membrane proteins puts the planctomycetes more close to eukaryotes as any other group of bacteria. Especially the anammoxosomes are suggested to be functional and structural homologs to mitochondria (Fuerst and Sagulenko, 2011). If the C-terminally elongated YidC of marine bacteria has a higher affinity to ribosomes, more ribosomes should be firmly bound to the membrane surface. This could be an advantage for cells that synthesize a large number of hydrophobic membrane proteins, similar to the specialized translation and insertion machineries in mitochondria. Marine prokaryotic organisms are, in many cases, unique in their genomic organization (Serres et al., 2009). Particularly in the planctomycete group, several gene duplications and protein motifs are found that only have counterparts in the archaeal or eukaryotic phyla (Studholme et al., 2004). Thus, the C-terminally extended tail of YidC in marine bacteria may be an ancestral remnant of a primordial translocation system operating without a SRP receptor or a truncated, minimal version of it. Later in evolution this system might have shifted to the mitochondria. Possibly, the phylogenetically ancient function of the SRP was mainly chaperone-like rather than having a targeting function. The targeting function may first have been acquired by an ancestral YidC protein with an extended C-terminal tail mediating membrane binding of ribosome-associated substrates. Later on in evolution the SRP
receptor became more efficient in membrane targeting and thus also allowed interaction with the Sec translocase. At this evolutionary stage, the C-terminal extensions of the YidC proteins were dispensable. Above all, marine bacterial groups may be a key to understanding how complex cellular processes like membrane targeting and protein translocation may have evolved and spread over the whole organismic world.

4.7 Deciphering the SRP signal sequence of KdpD

The aim of the second part of this study was to determine the specificity of SRP recognition in substrate proteins. In E. coli, the ribonucleoprotein SRP is essential for cell viability (Phillips and Silhavy, 1992) and comprises a single protein, Ffh, bound to the 4.5S RNA (Poritz et al., 1990; Ribes et al., 1990). During protein synthesis, the SRP binds to a hydrophobic signal sequence of the emerging polypeptide at the ribosomal exit tunnel of RNCs and targets the whole ribosome-associated complex in a GTP-dependent process to the inner membrane (de Gier et al., 1996; Ulbrandt et al., 1997). Two functional domains in SRP contribute to this activity: the NG domain of the Ffh protein, containing a Ras-like GTPase motif (Freymann et al., 1997), and the 4.5S RNA-Ffh M domain complex, which is responsible for signal sequence recognition (Zopf et al., 1990).

In earlier studies of our group the SRP-dependent membrane targeting of the sensor protein KdpD of E. coli was analyzed. KdpD had been shown to insert into the inner membrane independently both of Sec and YidC (Facey and Kuhn, 2003; Rothenbücher et al., 2006; Maier et al., 2008). The four-spanning membrane topology of KdpD provides an interesting aspect of signal sequence recognition by SRP. As in E. coli nearly all inner membrane proteins contain an N-terminal uncleaved signal sequence, the region that interacts with SRP is most likely the first transmembrane region, ensuring that early after the translation start, SRP mediates the co-translational targeting to the membrane. However, this is not possible for KdpD, since the protein encompasses a huge N-terminal hydrophilic region of 400 amino acid residues (Fig. 3.28 A) and recognition of a signal sequence in the first TM domain following that long hydrophilic region would prevent an early membrane targeting of the protein. Maier et al. (2008) identified via GFP-localization studies of N-terminal KdpD-GFP fusion proteins, a minimal peptide of residues 22-48 (N22-48) that is required to target KdpD to the membrane. This peptide contains three positively charged residues at its very N-terminus (aa 22-26), a stretch of 10 hydrophobic residues (aa 27-36), followed by a lysine residue and then another 6 hydrophobic residues (aa 38-43) and accessorially a predicted Walker A motif (aa 30-38) (Fig. 3.28 B). A crucial question is how the KdpD signal sequence is recognized by SRP as no obvious
consensus motifs are known for this interaction. SRP-dependent N-terminal signal sequences are highly diverse in their amino-acid composition and length, but they all contain a core of at least eight consecutive hydrophobic amino acid residues that act as the major determinant for recognition by SRP (Valent et al., 1997 & 1998; Martoglio and Dobberstein 1998; Hegde and Bernstein 2006, Janda et al., 2010). In addition, the N-terminal tail of a signal sequence typically contains positively charged residues with yet unknown function, suggesting a combination of hydrophobic interactions and electrostatic contacts during recognition and binding of SRP to signal sequences. Optionally, the positively charged residues may also interact with the negatively charged head groups of the membrane phospholipids.

To investigate the involvement of the positively charged residues and the Walker A motif in binding of SRP to the signal sequence, the binding affinities of the wild type KdpD N22-48 fragment and the mutant peptides 3Q (reduced net charge from +5 to +2) as well as the W3A mutant signal sequence (destroyed Walker A motif) to E. coli SRP were determined via microscale thermophoresis (MST) (2.17). I established purification (2.10.3) and labeling (2.17.1) protocols for the Ffh protein and stably reconstituted SRP (2.17.2) by the addition of in vitro synthesized 4.5S RNA (2.6.12). To ensure a properly reconstituted SRP under the assay conditions, MST measurements were performed to analyze the correct binding of labeled Ffh to the synthesized 4.5S RNA. A dissociation constant (K_d) of ~ 0.56 nM was determined for the Ffh:4.5S RNA holo SRP assembly (Fig. 3.35), which is in good agreement with the values cited in the relevant literature. Published data for the in vitro SRP assembly ranges from K_d values of 7 pM (Siu et al., 2007) to 3 nM (Jagath et al., 2001), depending on the method and on buffer components (Batey and Doudna, 2002; Jecklin et al., 2009). Thus, the reconstituted SRP here is highly suitable for the projected binding studies with the variant KdpD SRP signal sequence peptides. As expected, the wild type KdpD peptide N22-48 revealed the highest affinity to SRP with calculated K_d of ~ 500 nM, while the signal sequence mutants 3Q and W3A showed two-fold and 10-fold reduced binding affinities to SRP (Fig. 3.37), respectively. To ensure that the measured binding affinities truly represent specific binding of the signal sequence peptides to the 4.5S RNA-Ffh M domain complex, the interaction of the wild type N22-48 peptide with labeled Ffh protein, without the 4.5S RNA, was tested. When using this RNA-free SRP the calculated K_d raised from ~ 0.5 µM to more than 17 µM. This demonstrates the importance of using a functional, fully assembled SRP and hence strongly suggests the specificity of the SRP-substrate recognition in these in vitro MST-measurements.

In general, it is obviously very difficult to estimate a universally valid affinity of SRP to signal sequences since several previous studies already revealed contradictory results of
this highly dynamic process. Nearly all of the studies investigating the SRP:substrate interaction, were performed with ribosome nascent chain complexes. This is due to the physiological significance of the ribosome during substrate recognition of SRP. On the other hand, this bears difficulties concerning the translationally stalled state of the RNCs and especially the influence of the nascent chain length is discussed controversially (Flanagan et al., 2003; Bornemann et al., 2008; Holtkamp et al., 2012; Noriega et al., 2014-1) compared to SRP binding to actively translating ribosomes (Noriega et al., 2014-2). Considerably influenced by the applied method, binding affinities of SRP to RNCs displaying hydrophobic signal sequences range from 0.1 nM – 10 nM (Flanagan et al., 2003; Bornemann et al., 2008; Zhang et al., 2010; Saraogi et al., 2011; Holtkamp et al., 2012; Noriega et al., 2014-1), in all cases indicating a very tight binding. Interestingly, SRP still binds to vacant, non-translating ribosomes with a $K_d$ around 50 nM, while over long nascent chains, especially from SRP-independent proteins, weaken the binding affinity of SRP to the RNCs down to a $K_d$ value of > 200 nM (Bornemann et al., 2008). Since I did not use RNCs in the MST measurements but instead analyzed the direct SRP:signal sequence interaction with isolated components, it was expected that the binding affinities were lower due to the absence of the stabilizing SRP binding sites on the ribosome. Indeed, the MST-data are in good agreement to the calculated binding affinity of $\sim$ 1.5 µM of an isolated $\Delta$EspP signal peptide to SRP, measured by fluorescence anisotropy (Bradshaw et al., 2009). These observed binding affinities are likely due to multiple hydrophobic interactions between the SRP M domain and the signal sequence (Janda et al., 2010), but they cannot account for the high affinity of SRP to RNCs by only adding the SRP:signal sequence interactions to the SRP:ribosome interaction. This suggests that the tight binding of SRP to RNCs displaying a proper signal sequence is mainly caused by the conformational changes of SRP upon binding to a nascent chain (Hainzl et al., 2011), changing the interaction mode of SRP on the ribosome (Estrozi et al., 2011). These structural changes of SRP seems to be an important feature of SRP:substrate discrimination, since it is supposed that SRP also binds to RNCs displaying an incorrect cargo with moderate $K_d$ values of around 13 nM, but resulting in a falsely assembled SRP:FtsY targeting complex (von Loeffelholz et al., 2013). However, whether these observations represents the native mechanisms in vivo remains still unclear since recent interaction studies with actively translating ribosomes and dynamic single-molecule measurements at physiologically relevant substrate concentrations showed that SRP only engages translating RNCs that expose a functional signal sequence (Noriega et al., 2014-2). Further studies will be necessary to answer the question how a limited, sub-stoichiometric pool of cellular SRP ($\sim$ 400 nM in E. coli,
compared to 40 – 50 µM ribosomes) effectively distinguishes RNCs that display a proper signal sequence from those that do not.

Surprisingly, the measurements of the SRP interaction with isolated signal sequence peptides in this work revealed striking differences in the binding affinities of the wild type sequence compared to both, only slightly changed, mutant peptides. Substitution of the three positively charged amino acid residues in the N-terminal part of the signal peptide with glutamine residues in the 3Q mutant decreased the binding affinity for SRP two-fold. Initially, it had been proposed that the SRP 4.5S RNA interacts with the positively charged residues in the N-region of a signal peptide via electrostatic interactions (Batey et al., 2000; Wild et al., 2004), but crystal structures of SRP and cryo-EM structures of SRP:RNC complexes elucidated that the N-region is presumably too far away from the SRP RNA to make direct contacts (Janda et al., 2010). The positively charged residues in the N-region of the signal sequences may be conserved for other reasons, for example for its interaction with the phospholipid head groups of the membrane lipids during protein translocation (van den Berg et al., 2004). This is also consistent with preliminary results of localization studies with GFP fused to the KdpD-peptides, analyzed by fluorescence microscopy (data not shown). It was observed that the 3Q:GFP construct, although the isolated peptide showed direct interaction with SRP in vitro in the MST measurements, was distributed mainly in patches over the whole cell in vivo, probably due to a reduced interaction ability with the membrane. In contrast, the N22-48 (wild type) and W3A peptides fused to GFP were localized mainly at the inner membrane of E. coli. This also demonstrates that, although a 10-fold higher Kd for SRP:W3A interaction was observed in vitro, SRP efficiently targets the W3A:GFP fusion to the membrane in vivo, emphasizing the difficulty of direct conclusions from SRP:signal sequence affinities to proper targeting via SRP. For instance, there is no conclusive explanation for the strongly reduced binding affinity of the W3A mutant to SRP in the MST measurements, although in vivo studies showed that single point mutants that slightly change the hydrophobicity of the core region profoundly affect SRP recognition, while an altered charge in the N-region only has weak influence (Lee and Bernstein, 2001; Peterson et al., 2003). Since the grand average of hydropathicity (GRAVY) actually showed a slightly increased hydrophobicity of the W3A mutant (0.3 N22-48; 0.36 3Q; 0.54 W3A), in which three glycine residues were substituted for alanine residues, hydrophobicity alone could not account for the different binding affinities. One possible explanation could be that the destroyed Walker A motif in this mutant causes an altered conformation of the peptide, which is then disadvantaged in the direct interaction with the hydrophobic groove of the signal sequence binding cavity of SRP. It is proposed that the exclusion of the N- and C-terminal regions of the signal
sequence from the hydrophobic groove likewise restricts the longitudinal position of the signal peptide with respect to the groove (Janda et al., 2010). Secondary structure prediction of the three KdpD-signal sequence peptides confirm a presumable structural change of the signal sequence conformation due to the substitution of the three glycine residues in the W3A mutant, compared to the wild type N22-48 and 3Q peptides (Fig. 4.5).

![Secondary structure prediction of the KdpD-signal sequence peptides.](image)

**Fig. 4.5 Secondary structure prediction of the KdpD-signal sequence peptides.** The wild type peptide N22-48 and the 3Q mutant show the same predicted conformation, while the structure prediction of W3A shows an enhanced tendency for helix formation (purple boxes). Yellow boxes indicate sheet formation. Structure prediction was performed by the PSIRED server.

Therefore, it can be assumed that the alterations in the helical stretch of the hydrophobic core region of the signal sequence in W3A mainly cause the low binding affinity to SRP in the in vitro MST assays, which is in agreement with the observation that hydrophobicity may not be the sole factor for SRP recognition (Huber et al., 2005). Taken together, my results confirm that the SRP recognition sequence of KdpD encompasses the amino acid residues 22-48 at the very N-terminal tail of the large cytoplasmic domain of KdpD since a direct interaction of the isolated peptide with purified SRP was demonstrated via MST. A direct influence of the amino acid composition in the signal peptide on its SRP binding affinity in vitro was also shown, thereby confirming a minor influence of an altered charge in the N-terminal region while mutations in the hydrophobic core region cause strongly reduced binding affinities. These, however, probably are not due to the changed hydrophobicity but conformational changes in the helical region. However, more detailed analyses of additional signal sequence mutants and their SRP affinity, in combination with in vivo localization studies are required to understand the importance of particular amino acid residues and their position within the sequence.
SUMMARY

Members of the YidC/Oxa1/Alb3 protein family catalyze the insertion of integral membrane proteins into the lipid bilayer of the bacterial plasma membrane (YidC), the inner mitochondrial membrane (Oxa1), and the chloroplast thylakoid membrane (Alb3) (Saller et al., 2012; Dalbey et al., 2014). The insertase homologs are comprised of a conserved core region of 5 transmembrane domains, but are provided with additionally flanking N- and C-terminal regions of variable lengths and functions. The Gram-negative YidC is characterized by an additional N-terminal domain, while Gram-positive bacteria, mitochondria and plastids developed C-terminally extended insertase-domains. These domains are involved e.g. in direct interaction with ribosomes and facilitate a functional overlap with the co-translational SRP-targeting pathway. An extended C-terminal highly positively charged tail region was also found in the YidC homologs of the Gram-negative marine bacteria Rhodopirellula baltica and Oceanicaulis alexandrii, but not in Escherichia coli.

The primary subject of this work was to characterize and analyze in detail the C-terminally extended YidC chimera, composed of the E. coli YidC and the C-terminally extended domains of the marine YidC homologs. Biochemical binding assays with the purified YidC proteins and isolated, vacant E. coli 70S ribosomes showed that the C-tails mediate specific binding to ribosomes independently of the translational state of the ribosome. Furthermore, a ribosome-bound insertase complex was visualized by cryo-electron microscopy. The enhanced affinity of the C-terminally extended YidC was used to isolate stable complexes with stalled ribosomes, carrying a nascent polypeptide chain of a YidC substrate protein (MscL). The cryo-EM structure of a YidC-ribosome nascent chain complex (RNC) was solved to a 8,6 Å resolution and allowed the visualization of the nascent chain from the peptidyl transferase center through the ribosomal exit tunnel into the YidC density. The structure revealed the helix H59 of the 23S rRNA and the two ribosomal proteins L24 and L29 as the major contacts sites of YidC at the ribosomal tunnel exit. Pull down assays confirmed a significantly interaction of the C-terminal ribosome binding domain and the ribosomal protein L29, while L24 seems to be a universal contact site for the YidC-insertase core domain. Strikingly, the cryo-EM structure clearly showed a single monomer of YidC bound to the translating ribosome. This suggests that monomeric YidC might be the minimal functional unit for YidC-dependent, co-translational insertion of inner membrane proteins.
In addition to the in vitro tests, a possible role of the C-terminal YidC extensions in co-translational protein targeting was tested in vivo in E. coli. For that purpose the targeting and localization of the SRP-dependent YidC-substrate protein MscL (Facey et al., 2007) was investigated as a GFP fusion protein via fluorescence microscopy. In addition, the proper membrane insertion of MscL was analyzed in radioactive pulse chase experiments via AMS gel shift assays, either in the absence of a functional SRP or SRP receptor (FtsY). Both in vivo assays clearly showed that the C-terminal ribosome binding domain of the R. baltica YidC homolog can partially substitute for the SRP receptor function in E. coli, while the cytosolic signal recognition particle is still required for correct insertion of the MscL protein. Therefore, a new co-translational targeting and insertion model of YidC-only substrates was proposed. This model includes the binding of SRP to the ribosome nascent chain complex (RNC), preventing the aggregation and misfolding of the polypeptide chain. Together with the C-terminal ribosome binding domain that stably positions the RNC:SRP complex on the YidC insertase due to the interaction with the ribosomal protein L29, the SRP-receptor became dispensable. This works also highlights evolutionary aspects of the accessory YidC domains and indicates that the C-terminal extended tail of YidC in the planctomycete group may be an ancestral remnant of a primordial translocation system operating without a typical SRP receptor.

The second part focuses on the interaction of the signal recognition particle with SRP signal sequences. Isolated mutant signal sequence peptides were used to determine the specificity of SRP recognition in proteins. The interaction studies were established in an in vitro system and binding affinities of purified SRP to the isolated signal sequence peptides were determined via microscale thermophoresis (MST). A short sequence of 27 amino acid residues at the very N-terminal tail of the large cytoplasmic domain of KdpD was identified as a SRP signal sequence. Furthermore, a direct influence of the amino acid composition in the signal peptide on its SRP binding affinity in vitro was demonstrated. This confirms a low influence of an altered charge in the N-terminal region while mutations in the hydrophobic core region causes significantly reduced binding affinities to SRP.

Taken together, this study contributes to the understanding of the molecular mechanisms of co-translational membrane protein biogenesis in bacteria.
Proteine aus der YidC/Oxa1/Alb3-Familie vermitteln die direkte Insertion von integralen Membranproteinen in den Lipid-Bilayer der bakteriellen Plasmamembran (YidC), in die innere Mitochondrien-Matrix (Oxa1) und in die Thylakoidmembran von Chloroplasten (Alb3) (Saller et al., 2012; Dalbey et al., 2014). Die Homologe dieser Insertase-Familie besitzen eine konservierte Kernregion aus 5 Transmembrandomänen und zusätzliche N- und C-terminale Regionen, die sich in ihrer Länge und Funktion stark voneinander unterscheiden. Charakteristisch für das YidC aus Gram-negativen Bakterien ist eine zusätzliche N-terminale Transmembrandomäne, während Mitochondrien und Plastide C-terminal verlängerte, hydrophile Domänen an ihren Insertasen besitzen. Diese C-terminalen Domänen sind zum Beispiel an der direkten Bindung von Ribosomen beteiligt und verleihen der Insertase co-translationale Funktionen die sich mit dem SRP-Targetingmechanismus überschneiden. Im Gegensatz zur YidC-Insertase aus E. coli, besitzen die Gram-negativen, marinen Bakterien Rhodopirellula baltica und Oceanicaulis alexandrii YidC Homologe mit C-terminal erweiterten, positiv geladenen, hydrophilen Domänen.

Ziel dieser Arbeit war die detaillierte Charakterisierung und Analyse von C-terminal erweiterten YidC-Chimären. Die YidC-Chimären bestehen aus Fusionen des E. coli YidC Proteins und den C-terminalen Domänen der marinen YidC Homologe. Biochemische Bindungsstudien mit den gereinigten YidC-Proteinen und isolierten E. coli 70S Ribosomen zeigten, dass die C-terminalen Domänen eine spezifische Bindung an die Ribosomen vermitteln. Ähnlich zu der essentiellen Funktion der C-terminalen Domäne von Oxa1 war die Ribosomenbindung der YidC-Chimären unabhängig von einer translatierten Polypeptidkette am Ribosom. Des Weiteren, wurde ein Komplex aus naszierendem Ribosom (RNC) und gebundener YidC-Chimäre mittels Kryo-Elektronenmikroskopie auf 8,6 Å gelöst. In der Struktur konnte man den Weg der wachsenden Polypeptidkette vom Peptidyltransferase-Zentrum, durch den Ausgangstunnel des Ribosoms bis hinein in das YidC Protein verfolgen. Mit Hilfe der Kryo-EM Struktur wurde die Helix H59 der 23S rRNA und die beiden ribosomalen Proteine L24 und L29 als Hauptkontakte des YidCs am Ribosom identifiziert. Durch Pull-Down Versuche konnte eine signifikante Interaktion der C-terminalen Domäne mit L29 gezeigt werden, wohingegen L24 vermutlich eine Interaktion mit der YidC-Kerndomäne eingeht. Außerdem lieferte die Struktur des YidC:RNC Komplexes den Beweis, dass ein einzelnes YidC-Molekül an translatierende Ribosomen bindet und somit vermutlich die minimalste, funktionelle Einheit für die YidC-abhängige, co-translationale Insertion von Membranproteinen darstellt.
Um die in vivo-Funktion der C-terminalen YidC Domänen während des co-translationalen Targetings in *E. coli* zu untersuchen, wurde das Membrantargeting des SRP-abhängigen YidC-Substratproteins MscL (Facey *et al.*, 2007) als GFP-Fusionsprotein mittels Fluoreszenzmikroskopie untersucht. Parallel dazu wurde die korrekte Membraninsertion von MscL über eine AMS-Modifizierung in radioaktiven pulse-chase Experimenten analysiert. Beide in vivo Untersuchungen zeigten, dass die C-terminale Ribosomenbinededomäne des *R. baltica* YidC Homologs zu einem gewissen Teil die Funktion des SRP-Rezeptors in *E. coli* übernehmen kann. Dabei wird jedoch SRP weiterhin für die korrekte Insertion von MscL benötigt. Aufgrund der Erkenntnisse dieser Arbeit kann ein neues Modell der co-translationalen Insertion von YidC-abhängigen Proteinen, wie z.B. MscL, postuliert werden. Daran beteiligt ist das SRP, welches durch die Bindung an RNC-Komplexe die Aggregation und Missfaltung der Polypeptidkette verhindert. Zusammen mit der C-terminalen Ribosomenbinededomäne, kann in Abwesenheit des SRP-Rezeptors der RNC:SRP-Komplex an der YidC-Insertase durch die Interaktion mit dem ribosomalen Protein L29 positioniert und stabilisiert werden. In dieser Arbeit werden ebenfalls evolutionäre Aspekte der akzessorischen YidC-Domänen behandelt, die darauf schließen lassen, dass die C-terminale YidC-Domäne in der Gruppe der Planktomyceten eine Art Überbleibsel eines Vorfahren mit primitivem Translokationssystem ohne klassischem SRP-Rezeptor ist.


Beide Teilspekte dieser Arbeit tragen zum Verständnis des molekularen Mechanismus der co-translationalen Biogenese von integralen Membran-proteinen in Bakterien bei.
5 REFERENCES


Becker T., Bhushan S., Jarasch A., Armache J.P., Funes S., Jossinet F., Gumbart J.,
Mielke T., Berninghausen O., Schulten K., Westhof E., Gilmore R., Mandon E.C.
and Beckmann R. (2009) Structure of monomeric yeast and mammalian Sec61
complexes interacting with the translating ribosome. *Science* 326(5958):1369-1373

D.N. and Beckmann R. (2015) Translational arrest by a prokaryotic signal
recognition particle is mediated by RNA interactions. *Nat Struct Mol Biol* doi:
10.1038/nsmb.3086

Beckmann R., Spahn C.M., Eswar N., Helmers J., Penczek P.A., Sali A., Frank J. and
the translating 80S ribosome. *Cell* 107(3):361-372

to the specific depletion of the light-harvesting system. *Plant Cell* 14(9):2303-2314

Benz M., Bals T., Gügel I.L., Piotrowski M., Kuhn A., Schünemann D., Soll J. and Ankele
E. (2009) Alb4 of Arabidopsis promotes assembly and stabilization of a non-
chlorophyll-binding photosynthetic complex, the CF1CF0-ATP synthase. *Mol Plant*
2(6):1410-1424

Benz M., Soll J. and Ankele E. (2013) Arabidopsis thaliana Oxa proteins locate to
mitochondria and fulfill essential roles during embryo development. *Planta*
237(2):573-588

for signal sequence recognition from amino-acid sequence of 54K subunit of signal

the signal recognition particle 54-kDa subunit by its Escherichia coli homolog. *Proc
Natl Acad Sci USA* 90(11):5229-5233

Bibi E. (2011) Early targeting events during membrane protein biogenesis in Escherichia
coli. *Biochim Biophys Acta* 1808(3):841-850

terminal status of Escherichia coli proteins at the steady-state and upon

Bingel-Erlenmeyer R., Kohler R., Kramer G., Sandikci A., Antolic S., Maier T., Schaffitzel
complex reveals mechanism of nascent chain processing. *Nature* 452(7183):108-
111

Bischoff L., Berninghausen O. and Beckmann R. (2014-2) Molecular basis for the

Bischoff L., Wickles S., Berninghausen O., van der Sluis E.O. and Beckmann R.
(2014-1) Visualization of a polytopic membrane protein during SecY-mediated
membrane insertion. *Nat Commun* 5:4103

1500

*Biomembranes* pp 193-195

Bohnert M., Rehling P., Guiard B., Herrmann J.M., Pfanner N. and van der Laan M.
(2010) Cooperation of stop-transfer and conservative sorting mechanisms in
REFERENCES


REFERENCES


REFERENCES


162
REFERENCES


164


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


Yi L., Jiang F., Chen M., Cain B., Bolhuis A. and Dalbey R.E. (2003) YidC is strictly required for membrane insertion of subunits a and c of the F(1)F(0)ATP synthase and SecE of the SecYEG translocase. *Biochemistry* 42(35):10537-10544


6 ABBREVIATIONS

µM   micromolar
A    Ångstroem
A. thaliana Arabidopsis thaliana
A3CT  C-terminal domain of A. thaliana Alb3
aa   amino acid
aa-tRNA aminoacyl-tRNA
AC   affinity chromatography
Alb3  Albino-3, chloroplast insertase homolog
amp  ampicillin
AMS  4-acetamido-4´-maleimidylstilbene-2,2´-disulfonic acid
ara  arabinose
A-site acceptor-site of the ribosome
ATP  adenosine triphosphate
bp   base pairs
CD   circular dichroism
cpSRP chloroplast signal recognition particle
cryo-EM cryo-electron microscopy
CTD  C-terminal domain
C-terminal carboxyl-terminal
cV   column volume
Da   dalton
DDM  n-dodecyl-β-d-maltoside
DMSO dimethylsulfoxide
DNA  deoxyribonucleic acid
DTT  dithiothreitol
DUF  archaeal insertase homolog
E. coli Escherichia coli
EcYidC E. coli YidC
EDTA ethylenediaminetetraacetic acid
EF   elongation factor
ER   endoplasmic reticulum
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-site</td>
<td>exit-site of the ribosome</td>
</tr>
<tr>
<td>et al.,</td>
<td><em>et alii; and others</em></td>
</tr>
<tr>
<td>EtOH</td>
<td>ethanol</td>
</tr>
<tr>
<td>Ffh</td>
<td>fifty-four-homolog, SRP54-homolog</td>
</tr>
<tr>
<td>FSC</td>
<td>fourier shell correlation</td>
</tr>
<tr>
<td>FtsY</td>
<td>filamentous temperature sensitive Y, prokaryotic SRP-receptor</td>
</tr>
<tr>
<td>g</td>
<td>acceleration of gravity (9.81 m/s²)</td>
</tr>
<tr>
<td>GFP</td>
<td>green fluorescent protein</td>
</tr>
<tr>
<td>glc</td>
<td>glucose</td>
</tr>
<tr>
<td>GTP</td>
<td>guanosine triphosphate</td>
</tr>
<tr>
<td>h</td>
<td>hour</td>
</tr>
<tr>
<td>H₂Oᵈᵈ</td>
<td>double-deionized water</td>
</tr>
<tr>
<td>HCl</td>
<td>hydrochloric acid</td>
</tr>
<tr>
<td>HEPES</td>
<td>4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid</td>
</tr>
<tr>
<td>His-tag</td>
<td>histidinertag</td>
</tr>
<tr>
<td>IEC</td>
<td>ion exchange chromatography</td>
</tr>
<tr>
<td>IF</td>
<td>initiation factor</td>
</tr>
<tr>
<td>IMAC</td>
<td>immobilized metal ion affinity chromatography</td>
</tr>
<tr>
<td>IMV</td>
<td>inner membrane vesicle</td>
</tr>
<tr>
<td>Inc.</td>
<td>incorporated</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-β-D-thiogalactoside</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani</td>
</tr>
<tr>
<td>LDAO</td>
<td>lauryldimethylamine oxide</td>
</tr>
<tr>
<td>min</td>
<td>minute</td>
</tr>
<tr>
<td>mM</td>
<td>millimolar</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger ribonucleic acid</td>
</tr>
<tr>
<td>MscL</td>
<td>mechanosensitive channel of large conductance</td>
</tr>
<tr>
<td>MST</td>
<td>microscale thermophoresis</td>
</tr>
<tr>
<td>MW</td>
<td>molecular weight</td>
</tr>
<tr>
<td>MWCO</td>
<td>molecular weight cut-off</td>
</tr>
<tr>
<td>NaCl</td>
<td>natriumchloride</td>
</tr>
<tr>
<td>Ni²⁺</td>
<td>nickel</td>
</tr>
</tbody>
</table>
nm  nanometer
nM  nanomolar
N-terminal amino-terminal
O. alexandrii Oceanicaulis alexandrii
O1CT C-terminal domain of S. cerevisiae Oxa1
OaCT C-terminal domain of O. alexandrii YidC
Oa-YidC YidC homolog of O. alexandrii
OD optical density
Oxa1 oxidase assembly protein 1, mitochondrial insertase homolog
PAGE polyacrylamide gel electrophoresis
PBS phosphate buffered saline
PCR polymerase chain reaction
pH potential hydrogen
PMF proton motive force
PMSF phenylmethylsulfonyl fluoride
P-site peptidyl-site of the ribosome
PTC peptidyl transferase center
R. baltica Rhodopirellula baltica
RbCl rubidium chloride
RbCT C-terminal domain of R. baltica YidC
Rb-YidC YidC homolog of R. baltica
RNA ribonucleic acid
RNC ribosome nascent chain
rpm revolutions per minute
RRF ribosome recycling factor
rRNA ribosomal ribonucleic acid
RT room temperature
S sedimentation coefficient (Svedberg)
S. cerevisiae Saccharomyces cerevisiae
S. mutans Streptococcus mutans
SDS sodium dodecyl sulfate
Sec secretion
SEC size exclusion chromatography
SRP  signal recognition particle
Staph A  \textit{Staphylococcus aureus} protein A
Strep-tag  streptavidin-tag
TAE  tris-acetate-EDTA buffer
TAT  twin arginine translocation
TCA  trichloroacetic acid
TF  trigger factor
TM  transmembrane (segment)
Tris  trishydroxymethylaminomethane
tRNA  transfer ribonucleic acid
UV  ultraviolet
wt  wild type
YidC  bacterial membrane insertase
YidC-Oa  \textit{E. coli} YidC with the C-terminal domain of \textit{O. alexandrii} YidC
YidC-Oxa  \textit{E. coli} YidC with the C-terminal domain of \textit{S. cerevisiae} Oxa1
YidC-Rb  \textit{E. coli} YidC with the C-terminal domain of \textit{R. baltica} YidC
\alpha  anti, alpha
ACKNOWLEDGMENTS

First of all, I would like to thank Prof. Andreas Kuhn for the possibility to perform my PhD thesis in his laboratory and to work on fascinating and also challenging projects in an excellent research environment. Especially, I want to thank Prof. Kuhn for giving me the opportunity and the support to present my work on many national and international conferences.

Exceptionally, I want to express my deep gratitude to Dr. Dorothee Kiefer for supervising this work, thereby giving me all the freedom and support I needed to conduct my research. Thank you for all the help, proof-reading of this work, enriching discussions and for sharing the passion of C-terminal tails!

Special thanks also to the whole “Doro-Lab” for the great time and excellent teamwork. Lutz and Renate, it was a pleasure of working with you. Furthermore, I want to thank my Master students Emanuelle, Markus and Paul for their outstanding work, contributing to this thesis.

I would like to thank all past and present members of the entire Institute of Microbiology for providing such a great working atmosphere during my time in the lab. Special thanks go to Chris, Dirk and Eva for the funny times in the lab and also for critical discussions, great support and motivation throughout the past 5 years. Domenico and Sebastian, thank you for all the answered biochemical and biophysical questions, sharing protocols and critical proof-reading of my application. I also want to thank Andrea for the help in all important things beyond the scientific stuff.

Additionally, I want to thank the Beckmann group for the great collaboration with special thanks to Stephan for the excellent work in pushing the YidC:RNC complexes to a successful cryo-EM structure.

Schließlich möchte ich mich noch ganz herzlich bei meinen Eltern für die bedingungslose Unterstützung in all den Jahren bedanken, ebenso bei meiner Schwester, die einfach immer an mich glaubt! Vielen Dank Andy, dass du immer für mich da bist und die ganzen Jahre über mein Ausgleich warst.
Eidesstattliche Versicherung gemäß §7 Absatz 7 der Promotionsordnung der Universität Hohenheim zum Dr. rer. nat.


Ich versichere an Eides Statt, dass ich nach bestem Wissen die reine Wahrheit erklärt und nichts verschwiegen habe.

____________________  ______________________
Ort, Datum                      Unterschrift
Personal data

Name: Ines Seitl
Date of birth: 17.06.1985
Place of birth: Esslingen a.N., Germany
Nationality: German

Education

2015 / 10 – to date: Post Doc at the University of Hohenheim, Institute of Food Science and Biotechnology, Prof. Dr. rer. nat. Lutz Fischer
2015 / 09 – 2010 / 03: PhD student at the University of Hohenheim, Institute of Microbiology, Prof. Dr. rer. nat. Andreas Kuhn
2010 / 01 – 2009 / 03: Diploma thesis “Charakterisierung der Ribosomenbindedomäne von Membraninsertasen mariner Bakterien”, University of Hohenheim, Institute of Microbiology, Prof. Dr. rer. nat. Andreas Kuhn
2010 / 03 – 2004 / 10: Studies of Biology, Dipl. Biol., University of Hohenheim

Publication