THE UNIVERSITY OF CHICAGO

# IDENTIFICATION OF A SUPERFAMILY OF MEMBRANE PROTEIN BIOGENESIS FACTORS

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE BIOLOGICAL SCIENCES
AND THE PRITZKER SCHOOL OF MEDICINE
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

GRADUATE PROGRAM IN CELL AND MOLECULAR BIOLOGY

BY

STEFAN ANDREI ANGHEL

CHICAGO, ILLINOIS

JUNE 2018

TO MY FAMILY, WHOSE SUPPORT HAS NEVER WAVERED.
I WON'T LET YOU DOWN.

## Table of Contents

## List of Figures

## List of Tables

## Acknowledgements

That this project ever got off the ground is a testament to the support I have received throughout the past six years. Performing a project that is part genetics, part evolutionary biology, part biochemistry in what has traditionally been a structural biology lab is a testament to how much my adviser Robert Keenan believed in me. He guided me at every step of the way, provided advice whenever I asked (and I asked often), and he was always much nicer to me than I was to him. Additionally, from the beginning this project was a day-to-day collaboration with Philip McGilvray, who was instrumental to most of the work done here and single-handedly responsible for saving the project from the dustbin on more than one occasion. It's always been Bob, Phil and I figuring out this stuff as we went along. We did OK.

My committee members Benjamin Glick (chair), Jonathan Staley and Michael Glotzer were more instrumental for this project than most, because on so many occasions I seemed lost and wandering aimlessly. Thank you for your advice and for providing direction. I apologize the food at committee meetings was not better.

Additionally, the main project described here would not have been possible without the work of Marta Borowska while in the Keenan lab and Pawel Dominik. Together they spearheaded the archaeal protein project and solved the structure on which this project rests. Marta, a great scientist as well as a labmate of immense kindness, taught me the ropes of the Mj0480 project. Virtually every other member of the Keenan lab made a mark on this project. Many thanks to Claire (who also taught me the ropes in the Keenan lab during and after my rotation), Brittney, Agnieszka, Ben, Matt, Szymon, Gosia and Frank (hope you join the lab!). Throughout this project I have benefited from scientific and technical advice from Wesley Clark,

Tyler Starr, Molly Evans, Edward Wallace, Katie Mika, Alicja Santos and Joanna Kalita (among others).

There are also many others at The University of Chicago who helped greatly keeping me on track. This includes Kristine Gaston in the CMB Office and Lisa Anderson in the BMB Office as well as the army of administrative and maintenance staff whose work is seldom mentioned but always crucial.

This project of course would not have been possible without those who believed in it and supported it. I am immensely grateful to the Boehringer Ingelheim Fonds who not only supported me financially for over two years but also helped me grow as a scientist and join a community of wonderful people. The foundation will remain a part of my identity forever. Many others have contributed to funding my work, including The BMB Department at The University of Chicago and the National Institute of Health through grants awarded to the Keenan lab on this work.

Successfully completing my PhD was of course only possible to those who helped me get there in the first place. This includes my mentor in Romania Florina Miricel (who helped me get started as a biologist and kept me inspired in the early days), the late Florina Maracine (without whom I would have never gotten into college in the US), my undergraduate mentor Anne Marie Stanley (who taught me most of what I know about science, and also taught me that scientists can be cool) and my undergraduate adviser Tom Rapoport (who helped me grow as a scientist in more ways than I can count).

Lastly, there are the friends who helped me remain sane(ish) throughout the process and sometimes even enjoy graduate school. Many thanks to Alex, Katie, Diedre, Phil and everyone else who inexplicably puts up with my snarky comments.

The gratitude I have for my parents and my sister for their support in life in general and in getting a PhD knows no bounds. Thank you, mom and dad, for letting me leave and follow my dreams. I am sorry the US isn't closer. I know it is difficult to have a child leave and build a new life half way around the world. The support you've always shown despite that reminds me just how much you love me. And thank you to my sister, who despite me being a grownup remains protective and supportive. The big duck forever, I guess.

Lastly, I am most grateful of all to my wife Lauren Lee. There is no part of my life in the past six years which she did not help with (she even helped me purify ribosomes!). Whenever I doubted myself she provided support, stability and the reassurance required to carry on. I only hope that I've done the same for her.

## Abstract

Cells possess specialized machinery to direct the insertion of membrane proteins into the lipid bilayer. In all organisms, the Sec61/SecYEG channel inserts most membrane proteins into the membrane in concert with various accessory subunits. Additionally, other membrane protein insertion systems have been described. In bacteria, the essential protein YidC inserts certain proteins into the plasma membrane. Eukaryotic orthologs of YidC are present in the mitochondrial inner membrane and the chloroplast thylakoid membrane (where they are called Oxa1 and Alb3 respectively) but no homologs have been described outside the endosymbiotic organelles. In eukaryotes, the Get1/2 complex inserts a subset of proteins with a single transmembrane domain at their C-terminus, but whether this pathway is conserved in prokaryotes is also unclear. Here, I present evidence that bacterial YidC and eukaryotic Get1 are both members of a superfamily of membrane protein biogenesis factors which I refer to as the Oxa1 superfamily. This superfamily includes a group of archaeal proteins annotated as DUF106, and two eukaryotic proteins named EMC3 and TMCO1. All superfamily members share a core membrane topology and structural fold in addition to key functional features. Archaeal superfamily members, like bacterial YidC, interact specifically with ribosomes translating a YidC substrate. Of the eukaryotic proteins, I show that the previously poorly understood TMCO1 protein functions in membrane protein biogenesis. TMCO1 associates with ribosomes translating multipass membrane proteins and functions in a complex with the Sec61 channel, some translocon accessory components and a group of poorly understood protein quality control factors. This work establishes a new paradigm for the evolution of membrane protein biogenesis factors and defines a biochemical function for TMCO1, paving the way for future mechanistic analysis of this novel pathway.

# 1 Introduction

## 1.1 An introduction to Transmembrane Proteins

A key feature that defines living cells is the presence of biological membranes that separate the intracellular and extracellular environments, and – in eukaryotes – define specific intracellular compartments. Biological membranes must be semipermeable – they must at the same time establish a physical barrier from the cellular environment and facilitate transmembrane flow of molecules and information. To accomplish these two seemingly contradictory roles, cellular membranes are composed of a mixture of lipids and proteins(Nicolson, 2014; Singer and Nicolson, 1972).

In the well-established fluid mosaic model of biological membrane architecture (Nicolson, 2014; Singer and Nicolson, 1972), membranes are described as amphipathic, with a core hydrophobic core flanked by hydrophilic regions on both sides. Both the lipid part – composed mostly of phospholipids, sphingolipids and sterols – and the protein part follow this basic amphipathic architecture. In particular, proteins can be divided into two categories. Peripheral, or extrinsic, proteins are soluble proteins which are only loosely attached to membranes through interactions with either lipid head groups or with other proteins. Intrinsic membrane proteins, on the other hand, are integral parts of the membrane that cross the lipid bilayer and are held in place by hydrophobic interactions with the surrounding lipids. Although some intrinsic membrane proteins are covalently anchored to a lipid molecule, the main class of such proteins are the transmembrane proteins – which have a part of the protein molecule physically cross the bilayer.

Transmembrane proteins (here referred to as simply “membrane proteins”) are the crucial functional component of biological membranes, as they perform most functions of transport of

small and large molecules, cell signaling and mediating interactions with other cells and the extracellular environment. These proteins share a common organizational principle – a hydrophobic core integrated into the bilayer and external hydrophilic regions which extend into the aqueous environment – but otherwise exhibit a wide diversity. The hydrophobic core itself can be composed of a β-sheeted barrel motif. This fold is present in bacterial and mitochondrial outer membrane "porin" proteins (Schulz, 1996) and in some bacterial toxins (Gouaux, 1997) but is otherwise not widely encountered and not discussed in this work. Instead, the membrane-spanning domain is usually composed of α-helices, with their number varying between one single helix in single-pass proteins to dozens of helices in the most complex multi-spanning membrane proteins.

Membrane proteins must be inserted into the correct lipid bilayer to perform their biological functions. In addition, their inherent hydrophobic nature makes membrane proteins prone to aggregation, which can be toxic (Chiti and Dobson, 2006; Kopito, 2000).
The existence of different types of membrane proteins and, in eukaryotes, different target membranes, serves to multiply the challenge of membrane protein insertion. To deal with this complexity efficiently, cells have evolved multiple dedicated targeting and insertion pathways. Despite different components and mechanisms, all pathways share the same conceptual steps. Membrane proteins are captured by a pretargeting factor during or soon after their synthesis in order to maintain their solubility and prevent aggregation. They are then targeted to their specific membranes, and then physically inserted into the bilayer (Figure 1).

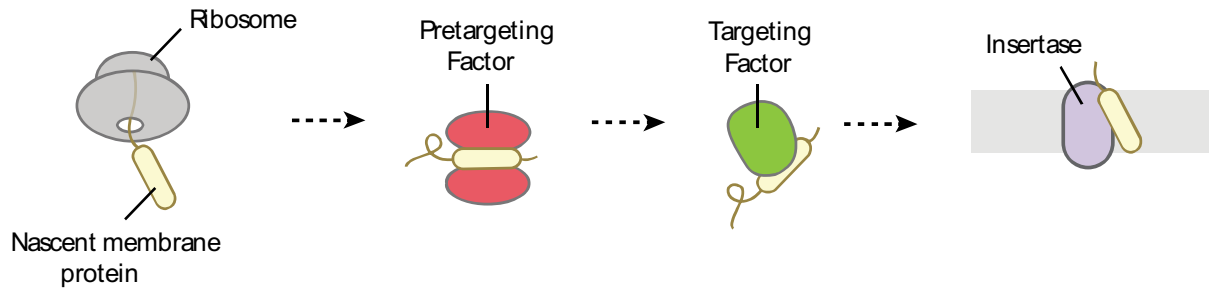


**Figure 1. General outline of a membrane protein targeting and insertion pathway**

All nascent membrane proteins must be captured by a pretargeting factor to prevent aggregation, after which they are passed on to a dedicated targeting machinery and actively inserted into the membrane. Sometimes two or more of these steps happen simultaneously.

## 1.2 The SRP-Sec61 Pathway Mediates the Insertion of Most Membrane Proteins

The best understood membrane protein insertion pathway is the co-translational pathway in which a protein with a transmembrane domain is detected while still on the ribosome, delivered to the membrane and inserted into the membrane before protein synthesis is completed (Figure 2).

The substrate capture, recognition and targeting steps are all accomplished by a soluble ribonucleoprotein complex called the Signal Recognition Particle (SRP) and the membrane-localized SRP Receptor. The process begins with the recognition of a stretch of hydrophobic residues on a nascent polypeptide by the SRP. The SRP is positioned adjacent to the ribosome exit tunnel, such that substrates bind co-translationally without escaping into the aqueous cytosol (Halic et al., 2004; Voorhees and Hegde, 2015). The SRP can accommodate a large variety of hydrophobic substrates, an ability which has been ascribed to the many methionine residues located inside the hydrophobic groove of the substrate binding domain (Bernstein et al., 1989; Janda et al., 2010; Keenan et al., 1998). In classical models, the SRP has been described to bind

the first stretch of 9-12 hydrophobic stretches it encounters, which could be either a signal sequence located at the N-terminus of soluble secretory proteins and some membrane proteins or the first transmembrane domain of a membrane protein. However, some recent studies have shown that in bacteria and yeast, SRP cannot efficiently bind moderately hydrophobic signal sequences and serves predominantly to bind transmembrane domains directly (Costa et al., 2018; Schibich et al., 2016; Valent et al., 1998). Binding of the SRP is traditionally thought to slow or even completely arrest translation until the ribosome is delivered to the membrane (Lakkaraju et al., 2008; Lipp et al., 1987; Mason et al., 2000; Walter and Blobel, 1981). This is a conceptually attractive model especially for multipass membrane proteins, as it would preclude exposure of more C-terminal TMDs to the aqueous environment. Recent studies, however, have questioned whether stalling happens *in vivo* (Chartron et al., 2016). SRP was also found associated with downstream TMDs in multipass proteins, suggesting the ribosome continues translating beyond the first TMD before being targeted to the membrane (Schibich et al., 2016). A lack of translational arrest implies that once the first transmembrane domain has been exposed, targeting to the membrane and aggregation are in direct competition.

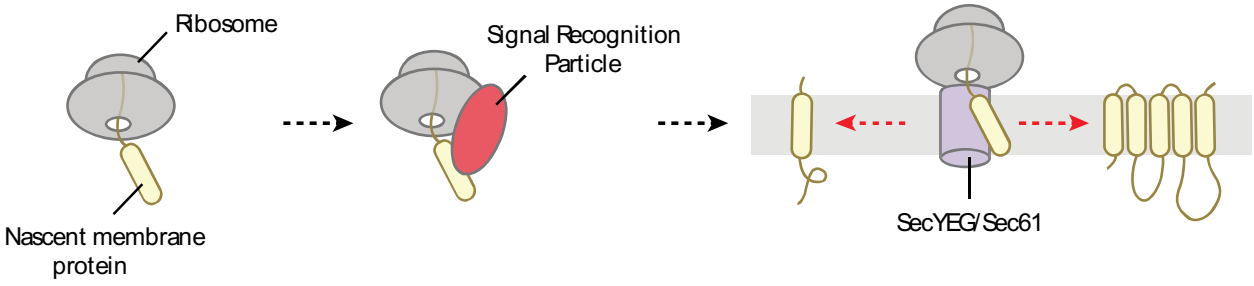


**Figure 2. Outline of the SRP/Sec Pathway**

The majority of membrane proteins are bound co-translationally by the Signal Recognition Particle, which performs both substrate capture and membrane targeting functions. The insertion step is the performed by the SecYEG/Sec61 channel, which can insert a wide range of singlepass and multipass membrane proteins.

The SRP-ribosome-nascent chain complex then docks on the membrane, where the SRP Receptor (SR) coordinates the handoff of the nascent chain to the Sec61 complex for insertion. The handoff step is still poorly understood, but structural studies show that SRP and Sec61 have overlapping binding sites on the ribosome, suggesting that extensive structural rearrangements must occur (Frauenfeld et al., 2011; Halic et al., 2006; Mitra et al., 2005; Schaffitzel et al., 2006; Voorhees and Hegde, 2015; Voorhees et al., 2014).

The channel which accomplishes the actual polypeptide translocation is always present as a trimeric protein, referred to as SecYEG in bacteria and Sec61αβγ in eukaryotes. The central subunit – SecY or Sec61α – has a central pore that accommodates the passage of nascent chains (Berg et al., 2004; Cannon et al., 2005; Kedrov et al., 2011; Park and Rapoport, 2012; Voorhees et al., 2014). The first step to insertion is the priming of the channel by docking of the translating ribosome. Structures of the channel bound to a nontranslating ribosome suggest this step involves little channel structural rearrangements beyond opening of the cytosolic part of the lateral gate, a region central to both nascent chain recognition and release (Gogala et al., 2014; Heinrich et al., 2000; Hessa et al., 2005; Voorhees et al., 2014). Priming is followed by actual channel opening. High resolution structures of a transmembrane domain bound to the channel are not available, but low-resolution structures of such complexes (Bischoff et al., 2014; Gogala et al., 2014) match well with high resolution structures of the channel engaging the signal sequence of a secreted protein (Li et al., 2016; Voorhees and Hegde, 2016). In both cases, the hydrophobic segment sits on the side of the lateral gate, in direct contact with the lipid phase of the membrane. This "lipid partitioning" mechanism may be utilized to expand the breadth of Sec channel substrates beyond what be accomplished through specific protein-protein interactions (Heinrich et al., 2000; Hessa et al., 2005). From this position, the transmembrane domain is

expected to then transition into the lipid phase, with the rest of the (hydrophilic) polypeptide passing through the channel pore. A similar mechanism is thought to occur for multi-pass membrane proteins, with each successive TMD contacting the lateral gate co-translationally and the being released laterally into the bilayer. This simple and elegant model cannot however account for all experimental observations. For example, some TMDs have been long known to exhibit a preferred topological orientation (Gafvelin and von Heijne, 1994; Locker et al., 1992). In one study, even mutating just the C-terminal residue of a protein can affect the topology of N-terminal TMDs (Seppälä et al., 2010). These experiments suggest that membrane protein topology can be locally and globally regulated, although factors involved are only starting to be characterized.

It is important to note that the Sec channel can also function post-translationally for the translocation of soluble proteins. Such a pathway is well described in prokaryotes, where the cytosolic chaperone SecA (sometimes together with a second chaperone named SecB) associates with substrates and delivers them to the SecYEG channel (Rapoport et al., 2017). However, it has long been known that membrane proteins prefer the co-translational SRP pathway (Valent et al., 1998), and at this time no SecA post-translational transmembrane substrates are known. Similarly, the Sec61 channel has been shown to function post-translationally in yeast (Deshaies and Schekman, 1987; Panzner et al., 1995; Rothblatt and Meyer, 1986) and more recently in metazoans (Lakkaraju et al., 2012b; Lang et al., 2012; Shao and Hegde, 2011a) but so far all known substrates are soluble.

## 1.3 YidC Family Proteins Co-translationally Insert Simple Membrane Proteins

A membrane protein insertion system that can operate completely independently of the Sec channel is found in bacteria and in the endosymbiotic organelles of eukaryotes. The central components of this system are the YidC family proteins (Figure 3). These proteins are represented by a single YidC gene in gram negative bacteria, while gram positive bacteria by two paralogs named YidC1 and YidC2. Multiple paralogs are also present in mitochondria (called Oxa1 and Oxa2) and plastids (called Alb3 and Alb4). Whether the presence of these multiple paralogous genes reflects the ancestral state of the last common ancestor of bacteria, mitochondria and chloroplasts or whether independent duplications have occurred is still unclear (Cavalier-Smith, 2013). The conservation of these proteins in archaea has also previously been suggested based on phylogenetic analyses (Luirink et al., 2001; Yen et al., 2001; Zhang et al., 2009) but the sequence similarity to the putative homologs is very low, and no experimental data exists to support this assignment.

YidC/Oxa1/Alb3 proteins all possess the same core protein fold, composed of 5 TMDs, a cytosolic coiled coil, and an N-out/C-in topology (Luirink et al., 2001). Some variations exist beyond this basic topology, with important functional implications. In certain bacteria such as *E. coli*, an extra TMD exists at the N-terminus (making the N-terminus cytosolic) and a large, folded periplasmic domain connects this TMD to the core protein (Sääf et al., 1998) . Additionally, mitochondrial Oxa1 and bacterial YidC2 (when present) have a long, positively charged C-terminal which can directly bind ribosomes through electrostatic interactions (Funes et al., 2009; Hell et al., 2001; Seitl et al., 2014).

YidC proteins can function to insert certain simple membrane proteins into the membrane in a Sec-independent co-translational mechanism. The first substrates found to use this pathway

were the Pf3 and M13 phage coat components (Chen et al., 2002; Samuelson et al., 2000, 2001). Two physiological substrates have also since been identified: the $F_0c$ subunit of the ATP Synthase and the MscL mechanosensitive channel (van Bloois et al., 2004; Facey et al., 2007; van der Laan et al., 2004; Yi et al., 2003). What these substrates have in common are very short translocated regions that lack significant charged or folded domains, suggesting the YidC-only pathway cannot handle the large free energy barrier of translocating such regions across the bilayer.

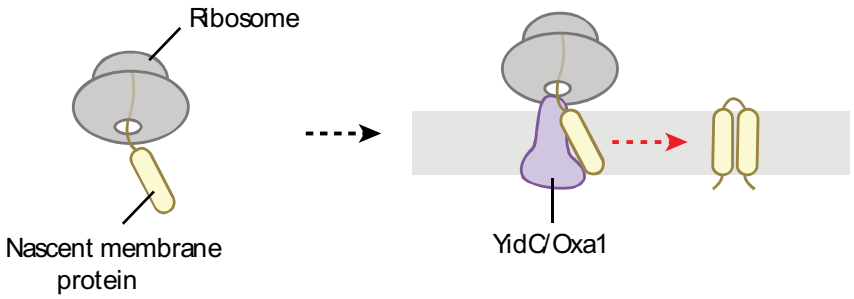


**Figure 3. YidC/Oxa1 directly inserts some simple membrane proteins.**
Some simple membrane proteins that have few transmembrane domains and short translocated regions can be inserted directly by the bacterial YidC insertase independent of the SecYEG channel. Substrate capture and targeting differ between species. In some bacterial species, this likely involves the SRP. In other bacterial species and mitochondria, the YidC/Oxa1 insertase binds ribosomes constitutively, making capture and targeting unnecessary.

Although conceptually the YidC pathway must have the same substrate capture, targeting and insertion phases as the Sec channel pathway, the initial stages are less understood. Bacterial YidC can bind ribosomes translating the $F_0c$ substrate, but not nontranslating ribsosomes (Kedrov et al., 2013), suggesting a need for active targeting. Targeting of substrates to YidC is thought to involve the SRP (van Bloois et al., 2004; Facey et al., 2007), although the mechanism of SRP handoff to YidC and the role of the SRP receptor are poorly defined. YidC homologs with extended C-terminal tails can bind ribosomes constitutively, eliminating the need for a nascent substrate or the SRP (Haque et al., 2010; Jia et al., 2003; Seitl et al., 2014; Szyrach et al.,

2003). This interaction in fact seems sufficient for recruitment of ribosomes to the membrane in general, such that YidC with a long tail can partially substitute for the SRP receptor *in vivo* (Seitl et al., 2014).

The insertion mechanism of YidC family proteins is still poorly understood mechanistically. Although initially suggested to act as a dimer with a hydrophilic pore in the middle (Kohler et al., 2009), recent YidC-ribosome structures and biochemical studies have shown convincingly that a YidC monomer can bind a translating ribosome (Kedrov et al., 2013; Seitl et al., 2014; Wickles et al., 2014). Functional studies employing an artificially fused YidC dimer also suggested that one single copy is sufficient *in vivo* (Spann et al., 2018). The monomer model implies no hydrophilic pore can be formed, further complicating mechanistic modelling. An answer to this puzzle was partially provided by high resolution YidC structures. Both the *Bacillus halodurans* and *E. coli* structures show a hydrophilic groove located deep inside the lipid bilayer (Kumazaki et al., 2014a, 2014b). In one functional model of *Bacillus* YidC, substrates were postulated to bind an essential arginine residue in the groove and then cross the membrane through a lipid-partitioning like mechanism (Kumazaki et al., 2014b). Although generalizing this model is difficult as the specific arginine residue is not essential in *E. coli* (Chen et al., 2014) the widely conserved hydrophilic groove may accomplish such a role through different residues in different species.

### 1.4 Post-translational Insertion of Transmembrane Proteins

Although far less common, post-translational insertion of membrane proteins has also been reported. Conceptually these pathways also involve dedicated substrate capture, targeting

and insertion steps, with the added complexity that the substrate must be quickly bound before it has the opportunity to aggregate in the cytosol.

By far the most widely established membrane proteins known to be inserted post-translationally are Tail Anchored (TA) proteins, which possess a single TMD at the very C-terminus. TA proteins exist in prokaryotes but have mostly been studied in eukaryotic cells where they much more common. It was recognized long ago that these proteins cannot engage the SRP pathway as by the time the TMD has exited the ribosome exit tunnel, translation has terminated (Kutay et al., 1993). Multiple hypotheses were initially proposed for how a TA protein might be inserted into the ER. These included the SRP/Sec61 system acting post-translationally (Abell et al., 2004), insertion catalyzed by Hsc70/HSp40 in the absence of membrane factors (Abell et al., 2007) and even completely spontaneous insertion (Brambillasca et al., 2005, 2006; Colombo et al., 2009). However, the existence of these pathways remains controversial, and the current consensus is that eukaryotic TA proteins are mainly inserted through the GET (Guided Entry of Tail Anchored) pathway.

In the GET pathway, the substrate protein is captured as soon as it is released from the ribosome, in what is known as the pre-targeting phase. In the well-understood yeast system, a three-protein complex composed of Get4, Get5 and Sgt2 first binds the substrate TA protein. Whether this step occurs on the ribosome *in vivo* is unclear, although these factors are known to associate with ribosomes (Fleischer et al., 2006; Zhang et al., 2016). Structural and biochemical analyses suggest a mechanism in which Sgt2 captures substrates relatively non-specifically and passes them on – through the a Get4/5 scaffold – to Get3, the specific targeting factor of the TA pathway (Bozkurt et al., 2010; Chang et al., 2010; Chartron et al., 2010; Wang et al., 2010). Get3 is a dimeric protein with a hydrophobic binding pocket that directly binds a transmembrane

domain, shielding it from the aqueous environment and maintaining its solubility (Bozkurt et al., 2009; Hu et al., 2009; Mateja et al., 2009, 2015; Suloway et al., 2009; Yamagata et al., 2010).

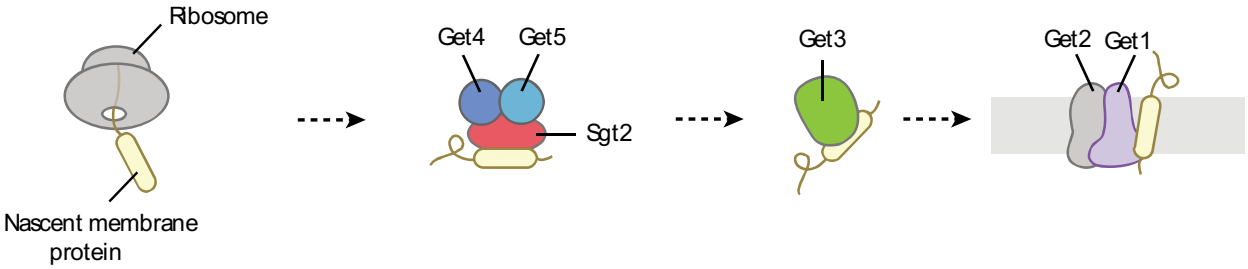


**Figure 4. The eukaryotic GET pathway inserts Tail Anchored proteins posttranslationally**
In the yeast GET pathway, TA proteins are captured by the Sgt2/Get4/Get5 targeting complex, handed off to the Get3 targeting factor, and then inserted into the ER membrane by the Get1/Get2 complex. The mammalian pathway has a more complex pretargeting step – which occurs on the ribosome – but highly conserved targeting and insertion steps.

After the Get3/TA complex migrates to the membrane of the ER, it binds to a specific receptor composed of the Get1 and Get2 proteins (Schuldiner et al., 2008). The Get1/2 complex functions both as Get3 receptor and insertase, as biochemical and genetic experiments have shown that Get1/2 is sufficient to accomplish insertion of a TA substrate comparable to that observed in wild type membranes (Mariappan et al., 2011; Wang et al., 2011a). The Get1/2 complex directly stimulates substrate release from Get3, and actively guides the substrate into the membrane, forming multiple interactions with the substrate as it is being inserted (Mariappan et al., 2011; Stefer et al., 2011; Wang et al., 2011a, 2014). *In vitro* reconstitution studies suggest that a single Get1/2 dimer is sufficient for maximal rates of insertion (Zalisko et al., 2017). The exact mechanism of Get1/2 mediated insertion, however, awaits high resolution structural data for the complex.

The mammalian TA protein insertion pathway is remarkably similar to the yeast one, although most steps are not as well understood. Pre-targeting is more complex, as a large, ribosome-attached protein named Bag6 forms a scaffold for the assembly of homologs of Get4,

Get5 and Sgt2 (referred to as TRC35, Ubl4a and SGTA respectively)(Leznicki et al., 2010; Mariappan et al., 2010). Bag6 seems to be a protein biogenesis hub, as in addition to the TA targeting pathway, it also serves a protein quality control factor, recruiting an E3 ligase which tags some substrate proteins for degradation (Hessa et al., 2011; Minami et al., 2010; Wang et al., 2011b). In the TA targeting pathway, Bag6 hands off the substrate to the soluble chaperone TRC40 (the Get3 homolog) (Mariappan et al., 2010; Stefanovic and Hegde, 2007). The substrate is then delivered to the ER membrane and inserted by a complex of WRB and CAML (the equivalents of Get1 and Get2), which just like in yeast are sufficient to support insertion in a purified system (Vilardi et al., 2011, 2014; Yamamoto and Sakisaka, 2012). Interestingly, although Get1 and WRB are clearly homologous, yeast Get2 and mammalian CAML share no similarity and appear to have evolved independently (Yamamoto and Sakisaka, 2012).

TA protein insertion in prokaryotes is much less understood than in eukaryotes. The Get3 targeting factor does have an archaeal homolog that can bind TA substrates and deliver them to yeast ER membranes *in vitro* (Borgese and Righi, 2010; Sherrill et al., 2011; Suloway et al., 2012). Nevertheless, membrane-associated components are not known, and the physiological relevance of this pathway in archaea remains unclear. In bacteria, insertion of some TA proteins seems to depend on the YidC insertase (Aschtgen et al., 2012; Peschke et al., 2018; Pross et al., 2016). YidC is an attractive candidate for TA protein insertion, as a hydrophilic groove-based insertion model could easily operate for TA substrates. Since the closest bacterial Get3 homolog is incapable of binding TA substrates (Borgese and Righi, 2010) some other mechanism must exist for targeting. So far studies have suggested that SRP and the Hsp70 protein DnaK might have this function, but this hypothesis remains speculative (Aschtgen et al., 2012; Peschke et al., 2018).

Much better understood is the post-translational insertion of certain membrane proteins by the YidC chloroplast homolog Alb3, the classical substrate being the Light Harvesting Complex Protein (LHCP) (Moore et al., 2000). LHCP is a nuclear-encoded protein synthesized as a precursor on cytosolic ribosomes that is first transported through the outer membrane of the chloroplast. In the chloroplast, LHCP is captured post-translationally by the chloroplast SRP (cpSRP) complex (Schuenemann et al., 1998) which then docks onto the SRP Receptor in the thylakoid membrane (Moore et al., 2003). Alb3 then captures the substrate by a direct interaction of the C-terminus of Alb3 with the cpSRP (Falk et al., 2010). Notably, the insertion step is fully independent of the chloroplast SecYE channel (Mori et al., 1999) even though LHCP is a multipass 3-TMD protein. If Alb3 is indeed able to catalyze insertion of this protein independent of other components, it would suggest that proteins of the YidC/Oxa1/Alb3 family are significantly more versatile than previously thought. However, such a conclusion awaits reconstitution of cpSRP/Alb3-mediated insertion with purified components.

## 1.5 Integral Membrane Accessory Factors of the Bacterial Translocon

Although the SecYEG channel is sufficient for some level of insertion of all substrates, *in vivo* the channel is often found as a protein supercomplex known as the holo-translocon. This larger complex contains, in addition to SecYEG, the YidC insertase and the trimeric SecDFYajC complex (Komar et al., 2016; Schulze et al., 2014; Scotti et al., 2000).

Functionally, the role of YidC in the holotranslocon is in facilitating the biogenesis of certain Sec-dependent membrane proteins. Specifically, YidC stimulates the translocation of specific TMDs in certain multipass proteins (Urbanus et al., 2001; Zhu et al., 2012). The physical location of YidC, directly contacting the SecYEG lateral gate (Petriman et al., 2018; Sachelaru et

al., 2013, 2017) suggest that YidC might prime the lateral gate for accepting a TMD. Such a model is supported by the observations that YidC alters the conductance of the SecY channel (Sachelaru et al., 2017) and that a nascent TMD displaces YidC from the lateral gate (Sachelaru et al., 2013).

Additionally, YidC has been shown to catalyze the folding of some multipass proteins. In the absence of YidC, the proteins LacY and MalF are inserted into the membrane but remain inactive (Beck et al., 2001; Nagamori et al., 2004; Wagner et al., 2008; Zhu et al., 2013). *In vitro* single molecule studies suggest YidC forms sequential, transient interactions with multipass proteins, stabilizing folding intermediates along the correct folding pathway at the expense of misfolded conformations (Serdiuk et al., 2016). Structural investigations of the holotranslocon suggest that its architecture may be adapted for this sequential interaction folding cycle, as the holotranslocon subunits define a central lipid-filled chamber which could accommodate protein folding (Botte et al., 2016).

The chloroplast YidC homolog Alb3 may also function similarly to bacterial YidC, as it associates with the chloroplast SecYE channel (Klostermann et al., 2002). The mechanism of Alb3 in this context has not been determined, however the most established substrate of this pathway – photosystem II D1 – is a multipass membrane protein with 5 TMDs (Walter et al., 2015). Other proposed substrates – based on their ability to bind Alb3 in a split ubiquitin system – are also multipass proteins: PSII D2 (5 TMDs), PSII CP43 (6 TMDs) and PS1-A (11 TMDs)(Pasch et al., 2005). The topology of these described substrates suggest that Alb3 may also support the insertion or folding of multipass proteins.

This SecY-dependent role of YidC family proteins does not survive in mitochondria in most organisms, as mitochondria usually lack a SecY gene (Glick and Heijne, 1996). Thus far,

no experiments have been done to investigate the existence of this pathway in organisms where mitochondria do contain a mitochondrial SecY gene (Lang et al., 1997). Nevertheless, the totality of the evidence suggests an ancestral role for YidC family proteins in Sec-dependent multipass membrane protein biogenesis.

Significantly less is known about the other components of the bacterial holotranslocon, SecDFYajC. The SecDF complex can conduct protons, and has been proposed to facilitate the harnessing of the trans-membrane proton motive force to catalyze the exit of translocated substrates from the SecYEG channel (Arkowitz and Wickner, 1994; Duong and Wickner, 1997; Economou et al., 1995; Matsuyama et al., 1993). One intriguing possibility is that these subunits harness the proton-motive force for a co-translational chaperone activity (Tsukazaki et al., 2011). No studies so far have linked these subunits to biogenesis of inner membrane proteins. The role of YajC in the holotranslocon likewise remains poorly understood.

## 1.6 Integral Membrane Accessory Factors of the Eukaryotic Translocon

The eukaryotic translocon has a more complex subunit composition than its bacterial counterpart, and the different components in eukaryotes seem to assemble in a substrate specific manner (Conti et al., 2015). Most accessory subunits have poorly defined roles, although many seem to play some role in the biogenesis of membrane proteins.

### 1.6.1 The Oligosaccharyl Transferase Complex

The Oligosaccharyl Transferase (OST) complex is one of the most crucial translocon accessory factors. OST catalyzes N-linked glycosylation – one of the most common posttranslational modifications of proteins, predicted to be present in about 70% of secretory

pathway proteins (Zafar et al., 2011). Specifically, OST catalyzes the transfer of a glycan from a dolichol lipid carrier to an asparagine residue contained in the consensus Asn-X-Thr/Ser sequence, where X can be any amino acid except proline. (Tai and Imperiali, 2001).

The components of the OST are remarkably conserved throughout eukaryotic evolution. The central, catalytic component is referred to as STT3 in yeast, with two homologues in metazoans called STT3A and STT3B (Kelleher et al., 2003; Nilsson et al., 2003; Yan and Lennarz, 2002). Metazoan STT3A and STT3B have complementary and overlapping roles, and define different OST complexes. STT3A-containing complexes associate with the Sec61 channel and work co-translationally to glycosylate proteins as they exit the channel (Braunger et al., 2018; Chavan et al., 2005; Cherepanova and Gilmore, 2016; Ruiz-Canada et al., 2009; Shibatani et al., 2005; Yan and Lennarz, 2005). SST3B complexes are free-standing entities in the membrane that can function post-translationally on any sequences skipped by STT3A (and supposedly on TA proteins which never encounter the translocon) (Cherepanova and Gilmore, 2016; Ruiz-Canada et al., 2009). The physical location of the catalytic subunit is important for glycosylation of membrane proteins: at least 12-14 residues between the TMD and the glycosylation sequence are required for efficient glycosylation(Nilsson and Heijne, 1993; Wild et al., 2018).

The function of the remaining OST subunits is so far poorly understood. However, recent structural data suggest that most subunits – referred to as OST2, OST4, OST5, OST1, SWP1, WBP1 in yeast and Dad1, Ost4, TMEM258, Ribophorin I, Ribophorin II and Ost48 respectively in metazoans – play a structural role, stabilizing the large complex and positioning the catalytic subunit (Braunger et al., 2018; Pfeffer et al., 2014; Wild et al., 2018). In addition to these core structural subunits, the yeast OST contains one of two paralogous subunits called OST3 and

OST6. These subunits, which have oxidoreductase activity, can capture substrates by disulfide bond formation and present them to the catalytic core (Poljak et al., 2018; Schulz et al., 2009; Zacchi and Schulz, 2016). OST3 and OST6 seem to display somewhat different substrate specificities, and have been proposed to work to improve the efficiency of glycosylation by tethering a substrate region within a binding pocket, maintaining it in an unfolded state to allow glycosylation to occur (Schulz and Aebi, 2009; Schulz et al., 2009). The mammalian homologs of OST3 and OST6 – called Tusc3 and MagT1 respectively – are found exclusively in the Sec61 independent STT3B complex and, and are required for function by capturing and presenting a substrate to the STT3B enzyme (Cherepanova and Gilmore, 2016; Cherepanova et al., 2014). Instead of OST3/6, the metazoan Sec-associated STT3A complex contains two subunits called DC2 and KCP2, which mediate the interaction with the Sec61channel (Shrimal et al., 2017).

### 1.6.2 The Translocating Chain Associating Membrane Protein (TRAM)

Another important translocon accessory factor is the Translocating Chain Associated Membrane Protein, or TRAM. TRAM is a single multipass polypeptide (Tamborero et al., 2011) that was first identified as a crosslinking partner for soluble proteins with signal sequences (Conti et al., 2015; Görlich et al., 1992; Krieg et al., 1989; Saurí et al., 2007; Voigt et al., 1996) but then subsequently proven to also crosslink to TMDs (Do et al., 1996; Heinrich et al., 2000; McCormick et al., 2003). *In vitro* reconstitution studies have shown TRAM to be necessary for the successful translocation of some soluble secretory pathway proteins and also the insertion of some membrane proteins (Görlich and Rapoport, 1993). Functional interactions show that TRAM can directly influence the lipid partitioning mechanism of the Sec61 channel by retaining certain TMDs that are then released in a coordinated fashion into the lipid bilayer (Cross and

High, 2009; Do et al., 1996; Heinrich and Rapoport, 2003; Heinrich et al., 2000; Sadlish et al., 2005). Additionally, TRAM has been suggested to stimulate the insertion of some membrane proteins (Heinrich et al., 2000) and signal-sequence containing proteins (Voigt et al., 1996) of moderate hydrophobicity. In all, TRAM might work to adjust the inherent hydrophobicity of TMDs, to either facilitate lateral gate opening in the presence of a hydrophobic domain or keep the gate closed in its absence (Shao and Hegde, 2011b).

### 1.6.3 RAMP4

Another accessory factor that seems to have some specificity to membrane protein substrates is RAMP4. RAMP4 is a small Tail Anchored protein that was originally found in complex with the ribosome and the Sec61 channel but did not seem to have an appreciable impact on protein translocation *in vitro* (Görlich and Rapoport, 1993; Wang and Dobberstein, 1999). Functional studies then showed that induction of ER stress by tunicamycin treatment leads to upregulation of RAMP4 (Yamaguchi et al., 1999) and conversely that deletion of RAMP4 can induce the unfolded protein response (Hori et al., 2006). In cultured cells subjected to ER stress, overexpression of RAMP4 suppressed membrane protein aggregation and degradation, suggesting a role for RAMP4 in membrane protein biogenesis (Yamaguchi et al., 1999). Subsequent studies then showed that the presence of a TMD in the ribosome exit tunnel induces the recruitment of RAMP4 to the ribosome (Pool, 2009) and that RAMP4 can directly bind a substrate through a hydrophobic interaction (Schröder et al., 1999). Postulating the exact function of RAMP4 is difficult, partially due to its small size of ~7kDa, and more work is needed on this poorly-understood protein.

### 1.6.4 The Sec62/Sec63 Proteins

Two highly studied but still poorly understood translocon accessory factors are Sec62 and Sec63. Yeast Sec62 is a two TMD membrane protein with a positively charged N-terminus which directly interacts with the negatively charged Sec63 C-terminus (Deshaies and Schekman, 1989; Lang et al., 2012; Wittke et al., 2000). The human Sec62 homolog has gained the ability to bind ribosomes, an ability which seems to be mutually exclusive with Sec63 binding (Müller et al., 2010). Sec63 is a multipass membrane protein with a luminal Hsp40 J-domain, which interacts with the luminal chaperone Kar2/BiP (Brodsky et al., 1995; Feldheim et al., 1992; Skowronek et al., 1995).

The Sec62/Sec63 complex has been best described in the post-translational translocation of certain soluble proteins in yeast. In this pathway, Sec62/Sec63 work – together with the other membrane proteins Sec71 and Sec72 and the soluble Kar2 chaperone – to adapt the translocon for translocating certain proteins with moderately hydrophobic signal sequences post-translationally (Deshaies et al., 1991; Ng et al., 1996; Panzner et al., 1995). In this pathway, Sec71/72 recruit substrates bound to cytosolic Hsp70 chaperones and pass them onto the other translocon components(Plath and Rapoport, 2000; Tripathi et al., 2017). Sec62 then directly contacts the signal sequence of a soluble substrate at the same time as Sec61 (Plath et al., 2004). The described mechanisms suggest that Sec62/Sec63 post-translational translocation may operate on substrates that are not efficiently engaged by the SRP.

Recent work has shown that the post-translational pathway is conserved in mammalian cells (Johnson et al., 2013; Lakkaraju et al., 2012b; Lang et al., 2012) which do contain ribosome-free Sec61/Sec62/Sec63 complexes (Meyer et al., 2000). Like in yeast, there is some evidence this pathway is utilized by proteins that cannot efficiently recruit SRP, for example

because of their small size does not allow sufficient time for SRP binding to the ribosome (Johnson et al., 2013; Lakkaraju et al., 2012b; Shao and Hegde, 2011a). Unlike yeast, metazoans do not have a Sec71/72 complex – the targeting is instead accomplished by the cytosolic chaperone calmodulin (Shao and Hegde, 2011a).

Although much less understood, it is clear that Sec62/63 also play a role in co-translational translocation of soluble proteins in both yeast (Brodsky et al., 1995; Young et al., 2001) and metazoans (Conti et al., 2015; Lang et al., 2012; Müller et al., 2010; Tyedmers et al., 2000). The exact role of the Sec62/63 in this process is almost completely unknown, but they were recruited to the ribosome-Sec61 complex in a substrate dependent manner by a substrate that was stalled in the channel and unable to translocate (Conti et al., 2015). One simple model is that for some substrates Sec62/63 work to recruit BiP (the metazoan homolog of Kar2), which assists with substrate translocation by its characteristic ratcheting mechanism.

Perhaps related to this co-translational function, Sec62/63 have also recently been linked to the biogenesis of membrane proteins. The same substrate-dependent recruitment of Sec62/63 was also reported for Prion protein, which carries a weakly hydrophobic TMD (Conti et al., 2015). Functionally, Sec63 is required for the *in vivo* maturation of a 7-TMD G-protein coupled receptor (Fedeles et al., 2015) and its expression is negatively correlated with the steady state level of a number of multipass membrane proteins (Mades et al., 2012). These data suggest that Sec63 captures multipass proteins and slows their insertion, perhaps assisting with their folding. Evidence also exists that Sec62/63 insert specific TMDs of both single-pass and multipass membrane proteins (Jung et al., 2014; Reithinger et al., 2013) and may assist with establishing the proper topology of some single-pass proteins with a moderately hydrophobic TMD

(Reithinger et al., 2013). No functional studies have so far established which of these activities (if any) occur co-translationally as part of the translocon.

### 1.6.5 Calnexin

For glycosylated ER proteins, attachment of a glycan is only the initial step of the modification: successive trimming of the initially attached large, branched glycan is then performed (Aebi et al., 2010). The initial trimming is done by glucosidase I, followed by glucosidase II trimming one sugar monomer to result in an intermediate glycan ($Glc_1Man_9GlcNAc_2$). Glucosidase II then acts again on this substrate, resulting in the final branched glycan ($Man_9GlcNAc_2$) that marks proteins as competent for ER exit (Aebi et al., 2010).

Calnexin is a single-pass membrane protein that can associate directly with the Sec61 channel (Boisramé et al., 2002; Lakkaraju et al., 2012a). Together with its soluble homolog calreticulin, calnexin participates in the folding of nascent glycoproteins by forming direct interactions with proteins containing the intermediate $Glc_1Man_9GlcNAc_2$ glycan (Caramelo and Parodi, 2008). This substrate capture by calnexin is thought to give proteins time to fold appropriately and prevent aggregation and premature ER exit of unfolded proteins. Additionally, calnexin actively facilitates folding by recruiting the ER oxidoreductase ERp57, which catalyzes correct disulfide bond formation (Soldà et al., 2006). The membrane association is important for calnexin function: calnexin seems to mostly interact with glycans close to the membrane, while calreticulin associates with peripheral glycans (Molinari et al., 2004; Pieren et al., 2005). The association with the translocon why calnexin can act co-translationally as well as post-translationally (Chen et al., 1995). If after dissociating from calnexin a protein is still unfolded

and displays exposed hydrophobic patches, it is recognized by UGT1 which restores the $Glc_1Man_9GlcNAc_2$ glycan (Sousa and Parodi, 1995) and allows calnexin/calreticulin to rebind, in what is known as the calreticulin/calnexin cycle.

### 1.6.6 The Translocon-Associated Protein Complex (TRAP)

The Translocon-Associated Protein (TRAP) is a 4-subunit complex consisting of three single pass membrane protein subunits (α,β,δ) and one larger, 4-TMD subunit (γ) (Bañó-Polo et al., 2017; Hartmann et al., 1993). TRAP associates in a 1:1 ratio with the Sec61 channel both in the presence and absence of a substrate protein (Conti et al., 2015; Dejgaard et al., 2010; Ménétret et al., 2008; Shibatani et al., 2005; Snapp et al., 2004) but has been proven to interact directly with translocating substrates by chemical crosslinking (Görlich et al., 1992; Wiedmann et al., 1987, 1989). Additionally, TRAP forms a higher-order complex with the Sec61 channel, the ribosome, and the OST (Braunger et al., 2018; Pfeffer et al., 2017; Shibatani et al., 2005).

TRAP is still poorly understood functionally but studies have found effects associated with both soluble and membrane protein substrates. Loss of the TRAP complex causes a congenital disease of glycosylation where some glycosylation sites are skipped (Losfeld et al., 2014; Ng et al., 2015; Pfeffer et al., 2017). This suggests that the TRAP-OST interactions are important for optimal choice of glycosylation sites. TRAP also directly stimulates the translocation of some soluble proteins; TRAP dependence may depend on the efficiency of their signal sequence in transport initiation (Fons et al., 2003). For membrane proteins, TRAP seems important for the establishment of the topology of some single pass proteins where the topology is ambiguously encoded in the sequence (Sommer et al., 2013).

# 2 A YidC-like Protein in the Archaeal Plasma Membrane

## 2.1 Overview

The work presented in the chapter started as an attempt by the Keenan lab to solve a crystal structure of an archaeal Get1 homolog. The Keenan lab was successful in producing a structure of Mj0480 (a protein from *Methanocaldococcus jannaschii*) but numerous biochemical experiments had failed to show an interaction between Mj0480 and the *M. jannaschii* TRC40 (the archaeal homolog of Get3).

I started work on this project conducting a phylogenetic analysis that showed a possible distant homology between Mj0480 and bacterial YidC, largely reproducing earlier phylogenetic work. After unsuccessfully attempting functional complementation of bacterial YidC by Mj0480, I showed that Mj0480 binds ribosomes programmed with a YidC substrate. This showed the first functional similarity between an archaeal protein and YidC. Together with the obvious structural similarities between Mj0480 and the newly published structure of bacterial YidC, this established Mj0480 as a *bona fide* YidC homolog. Subsequent crosslinking work showed specific crosslinks between Mj0480 and the nascent substrate on the ribosome.

A version of this work was published as Borowska, M.T., Dominik, P.K., Anghel, S.A., Kossiakoff, A.A., and Keenan, R.J. (2015). A YidC-like Protein in the Archaeal Plasma Membrane. Structure 23, 1715-1724.

## 2.2 Contributions

In this work, I performed the phylogenetic analysis, complementation studies and ribosome binding analysis. M.T.B. carried out the production and characterization of *Mj*0480 and sABs, crystallization and data collection of the *Mj*0480-sAB complex, and the photocrosslinking

studies. P.K.D. performed the nanodisc reconstitution and sAB selections with guidance from A.A.K. R.J.K. conceived the project, guided experiments and solved and analyzed the crystal structure.

### 2.3 Phylogenetic Analysis Shows a Group of Archaeal Proteins Related to YidC

No archaeal homologs of YidC have been described, but proteins with low sequence similarity to the YidC/Oxa1/Alb3 family have been identified in archaea (Luirink et al., 2001; Makarova et al., 2015; Yen et al., 2001; Zhang et al., 2009). These proteins are annotated as 'Domain of Unknown Function 106' (DUF106), reflecting their highly diverged sequences and the absence of any experimental insight into their function. Thus, whether archaea possess *bona fide* members of the YidC/Oxa1/Alb3 family remains unknown.

To gain insight into the evolutionary distribution and mechanism of YidC/Oxa1/Alb3-type insertases, we carried out structural and functional studies of an archaeal DUF106 protein from *Methanocaldococcus jannaschii* (*Mj*0480). We determined the 3.5 Å resolution crystal structure of *Mj*0480, revealing an overall fold and unusual structural features that define it as a member of the YidC/Oxa1/Alb3 family. We also show that *Mj*0480, like YidC, binds selectively to stalled RNCs and can be crosslinked to the nascent chain via a conserved, lipid-exposed hydrophilic surface. These data provide direct experimental evidence that the archaeal DUF106 proteins function as YidC-like insertases within the archaeal plasma membrane.

We constructed a phylogenetic tree based on a multiple sequence alignment of a representative set of YidC/Oxa1/Alb3 and archaeal DUF106 sequences. Three separate clades are observed, corresponding to mitochondrial Oxa1, bacterial YidC (including chloroplast Alb3 family members) and the archaeal DUF106 proteins (Figure 5). Consistent with previous

analyses, the YidC/Oxa1/Alb3 homologs are more closely related to each other than to the highly diverged archaeal DUF106 sequences (Luirink et al., 2001; Zhang et al., 2009).

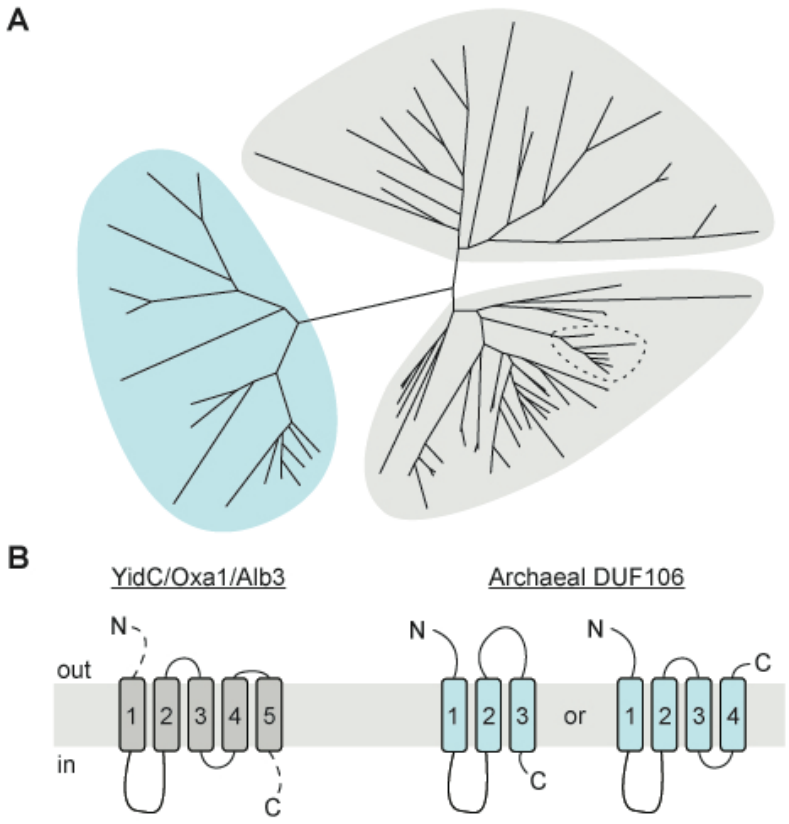


**Figure 5. Phylogenetic and topological analysis of the archaeal DUF106 family.**
(A) Phylogenetic tree of representative eukaryotic and bacterial YidC/Alb3/Oxa1 insertases (grey) and members of the archaeal DUF106 family of unknown function (light blue). Branch lengths for the three main clades are indicated; note the high divergence of the archaeal proteins. The archaeal DUF106 protein from *M. jannaschii* (*Mj*0480) is indicated with an asterisk.
(B) Members of YidC/Oxa1/Alb3 family are predicted to share a core topology comprising 5 TMs and a cytosolic-facing coiled-coil region between TM1 and TM2; these sequences are most divergent at their N- and C-termini (dashed lines). Depending on the algorithm, members of the archaeal DUF106 family are predicted to contain either three or four TMs, with a cytosolic-facing coiled coil between TM1 and TM2.

The archaeal DUF106 proteins share a low level of sequence identity (~10-15%) with members of the YidC/Oxa1/Alb3 family. Moreover, they are generally smaller than their eukaryotic and bacterial homologs, which show large variations in length, particularly at their N- and C-termini (Figure 5B). Nevertheless, the predicted membrane topology and secondary

structure reveal common features. The YidC/Oxa1/Alb3 proteins are thought to share a conserved core comprising five transmembrane (TM) helices (Kumazaki et al., 2014a, 2014b; Luirink et al., 2001; Sääf et al., 1998) and a cytosolic-facing coiled-coil region between the first two TMs. Likewise, COILS sequence analysis (Lupas et al., 1991) of the archaeal DUF106 proteins predicts a cytosolic-facing coiled-coil region between TM1 and TM2. However, in contrast to YidC/Oxa1/Alb3, the conserved transmembrane core of the shorter archaeal proteins is only predicted to contain three or four TMs (Figure 5B).

### 2.4 The structure of *Mj*0480

To gain insight into the function of the archaeal DUF106 family, we expressed an archaeal homolog from *M. jannaschii* (*Mj*0480) in *E. coli*. Although we could purify milligram quantities of n-dodecyl-β-D-maltopyranoside (DDM)-solubilized *Mj*0480, it formed oligomers in a concentration-dependent manner (Figure 6B), and failed to crystallize. To facilitate crystallization we generated high affinity synthetic antibody fragments (sABs) using a novel strategy to screen a phage library against biotinylated nanodiscs reconstituted with *Mj*0480 (Dominik and Kossiakoff, 2015). These sABs were screened for the ability to bind DDM-solubilized *Mj*0480 and promote crystallization. One such sAB ('M1'), which appears to form 1:1 and 2:2 complexes with *Mj*0480 in DDM (Figure 6C,D), gave rise to crystals that diffracted anisotropically to ~3.5 Å; the structure of this complex was solved by molecular replacement using an antibody fragment as a search model (Figure 7).

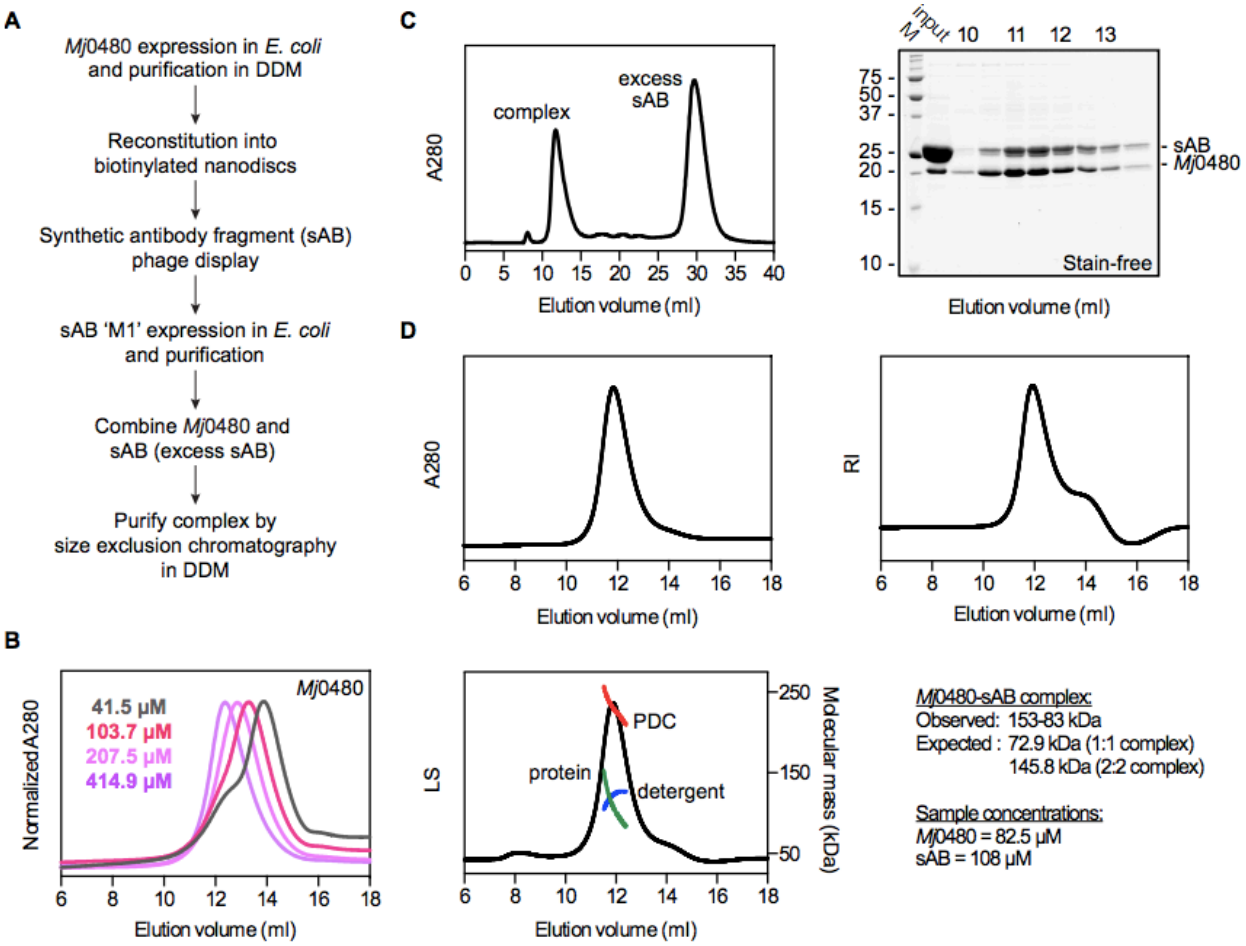


**Figure 6. Production and characterization of Mj0480 and Mj0480-sAB complexes.**

(A) Overview of the purification and sAB complex formation strategy.
(B) Size-exclusion chromatography of Mj0480 in DDM reveals concentration-dependent oligomerization. (C) Purification of the Mj0480-sAB 'M1' complex by gel filtration. Peak fractions were analyzed by SDS-PAGE and stain-free imaging (right panel). Note that the ~25 kDa band corresponds to overlapping heavy and light chains of the sAB.
(D) SEC-MALLS analysis of the Mj0480-sAB complex, showing ultraviolet (A280), refractive index (RI) and light scattering (LS) detector readings across the peak fractions. The red, green and blue lines in the LS trace indicate the observed molecular mass of the Mj0480-sAB-detergent complex (PDC), the Mj0480-sAB complex (protein) and detergent, respectively. Although the Mj0480-sAB complex elutes as a single, symmetric peak in DDM, it appears to be a mixture of 1:1 and 2:2 complexes.

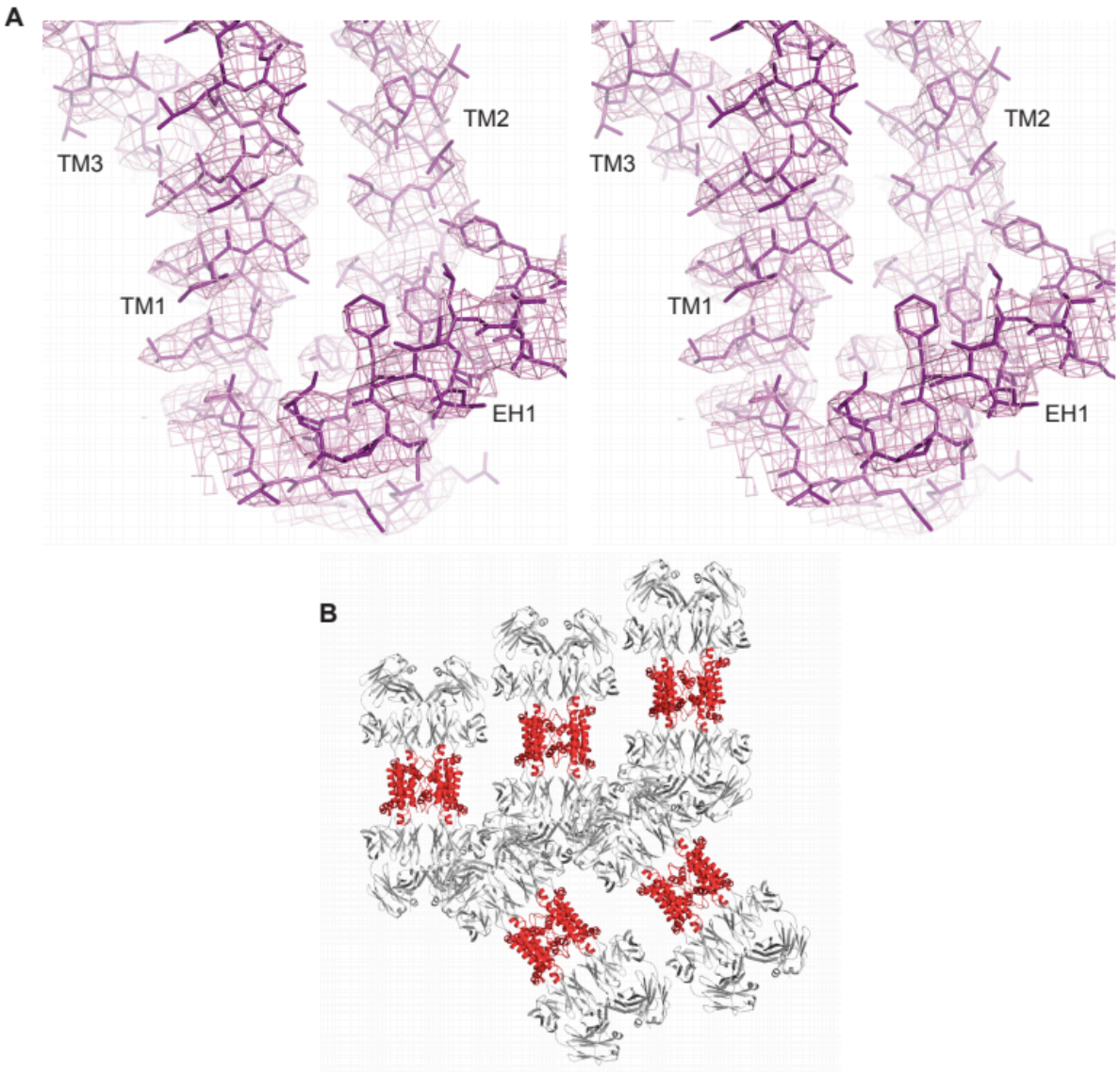


**Figure 7. Overview of the electron density map and crystal packing architecture**

(A) Cross-eyed stereo view of a σA-weighted 2Fo-Fc electron density map, calculated at 3.5 Å and contoured at 2.5 sigma after a single round of 4-fold NCS averaging. A portion of the final refined model, corresponding to part of the Mj0480 core, is superimposed (magenta sticks).
(B) Packing arrangement gives rise to alternating layers of Mj0480 (red) and sAB (grey) molecules in the crystal.

The asymmetric unit comprises four *Mj*0480-sAB complexes packed in an antiparallel arrangement (Figure 8A) that gives rise to alternating layers of sAB and *Mj*0480 molecules in the crystal (Figure 7B). Two of the *Mj*0480 molecules are in direct contact and form an antiparallel

dimer in which the C-terminal 50 residues from one monomer are swapped with those of the second monomer (Figure 8B).

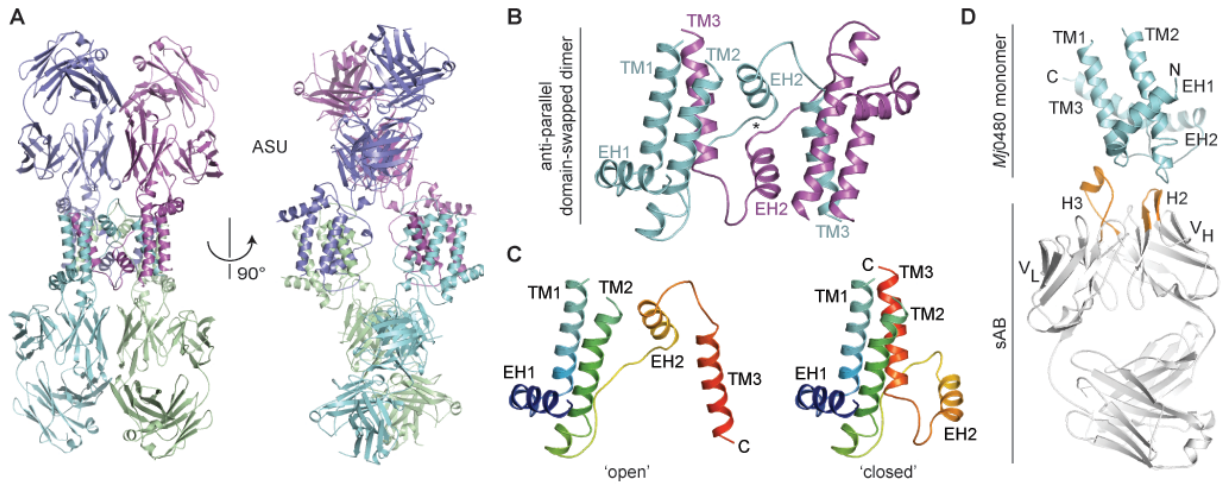


**Figure 8. A domain-swapped dimer in the Mj0480-sAB complex crystal.**

(A) The asymmetric unit comprises four sAB molecules sandwiched around four *Mj*0480 molecules; in this arrangement, the four *Mj*0480 molecules form two anti-parallel, domain-swapped dimers (cyan and magenta; blue and green).
(B) Close-up showing the anti-parallel arrangement of *Mj*0480 subunits (cyan, magenta) in the domain-swapped dimer. The swap is initiated at residue 149 (indicated by the asterisk), located in the loop following TM2, and extends through EH2 and TM3 to the C-terminus.
(C) Comparison of the crystallographic 'open' and modeled 'closed' monomers, color-ramped from N- (blue) to C- (red) terminus. The closed monomer model is generated by connecting residue 148 of one monomer to residue 149 of the other.
(D) Model of the closed *Mj*0480 monomer (cyan) bound to a single sAB molecule (grey). The sAB-*Mj*0480 contacts are mediated by residues within the H2 and H3 loops (orange).

The observation of domain-swapping in protein crystals is not unusual (Liu and and Eisenberg, 2002) but it complicates assignment of the physiologically relevant species. Because the *Mj*0480 monomers are inverted (Figure 8B), it is unlikely that the domain swapped dimer exists in the archaeal membrane. More likely, the C-terminal swap arises from the high protein concentration used for crystallization and the destabilizing effect of detergent. We therefore constructed an unswapped ('closed') model of monomeric *Mj*0480 by connecting residue 148 of monomer A to residue 149 of monomer B (Figure 8B,C).

This *Mj*0480 monomer spans the membrane with a longest dimension of ~50 Å; viewed from the cytosol, its overall length and width are ~25 and 36 Å. The N-terminus of *Mj*0480 faces the extracellular space and is followed by an amphipathic helix (EH1) that lies nearly parallel to the plane of the membrane (Figure 9A). This helix ends in a sharp turn and is followed by the first of three *bona fide* TMs. The cytosolic-facing end of TM1 connects to TM2 via the predicted coiled-coil motif that is disordered in the crystal, presumably reflecting its flexibility. TM2 spans the membrane and packs against TM1 in an antiparallel orientation. This is followed by a short extracellular loop region and a second amphipathic helix (EH2) that also lies nearly parallel to the extracellular face of the membrane. Finally, an extracellular loop connects to TM3 which packs against TM1 and TM2 in the core of the protein and extends toward the disordered C-terminus, which ends in the cytosol.

The cytosolic-facing half of *Mj*0480, including TM1-3, is loosely packed and likely dynamic, as evidenced by its relatively high B-factors (Figure 9B). A striking feature of this region is the presence of a concave hydrophilic surface, presented by TM1-3, that is exposed to both the lipid bilayer and the cytosol (Figure 9C). Importantly, this surface is not accessible to the extracellular side of the membrane, by virtue of the tightly packed hydrophobic core comprising residues from each of the five helical elements and the extracellular loops. Moreover, the hydrophilic character (Figure 9D) of the surface is conserved in both archaea (Figure 10) and bacteria (Figure 11). Thus, a key structural feature of the archaeal DUF106 family is the presence of a conserved, lipid- and cytosol-exposed hydrophilic surface presented by three transmembrane helices.

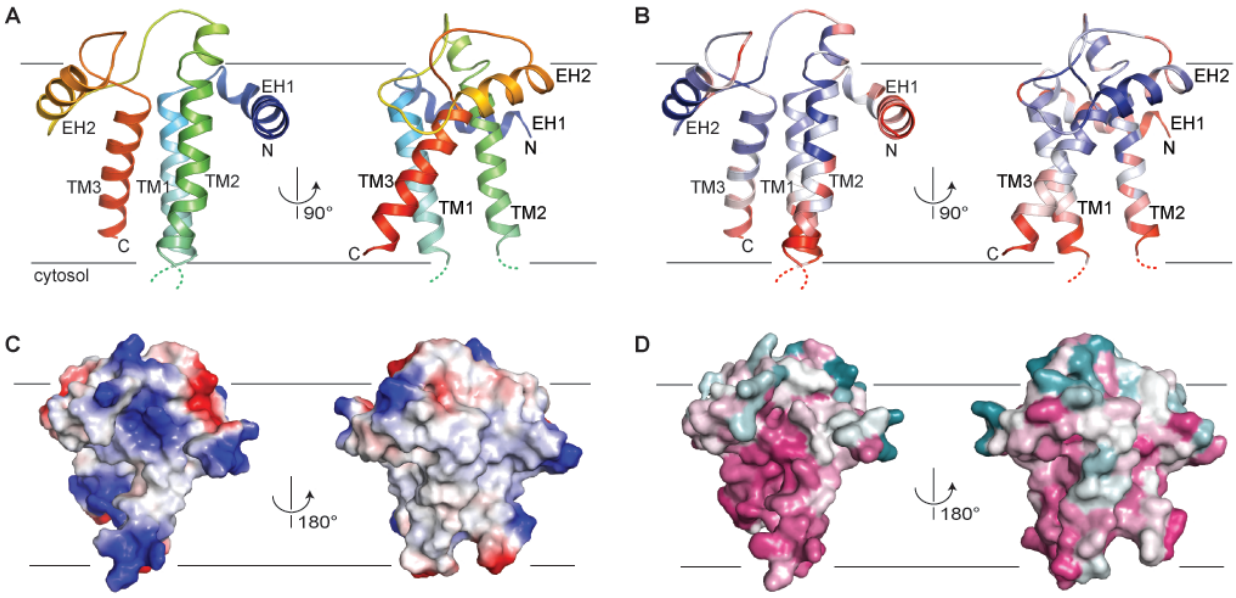


**Figure 9. Overall structure of the monomeric Mj0480 model.**

(A) Cartoon representation viewed from the plane of the membrane and color-ramped from the N- (blue) to the C- (red) terminus. The structure comprises three transmembrane helices (TM1, TM2, TM3) and two amphipathic helices (EH1, EH2) that lie nearly parallel to the plane of the membrane on the extracellular side of the membrane. The predicted coiled-coil region connecting TM2-TM3 is disordered in the crystal (dotted lines).
(B) Crystallographic B-factors are indicated using a three-color gradient from blue (60 $Å^2$) to red (160 $Å^2$).
(C) Electrostatic surface potential colored from negative (red) to positive (blue), reveals an exposed hydrophilic surface within the lipid bilayer.
(D) Sequence conservation of 33 archaeal DUF106-containing proteins mapped to the molecular surface of *Mj*0480 from most (pink) to least (cyan) conserved. The lipid exposed surface along one face of TM1-3 shows strong sequence conservation.

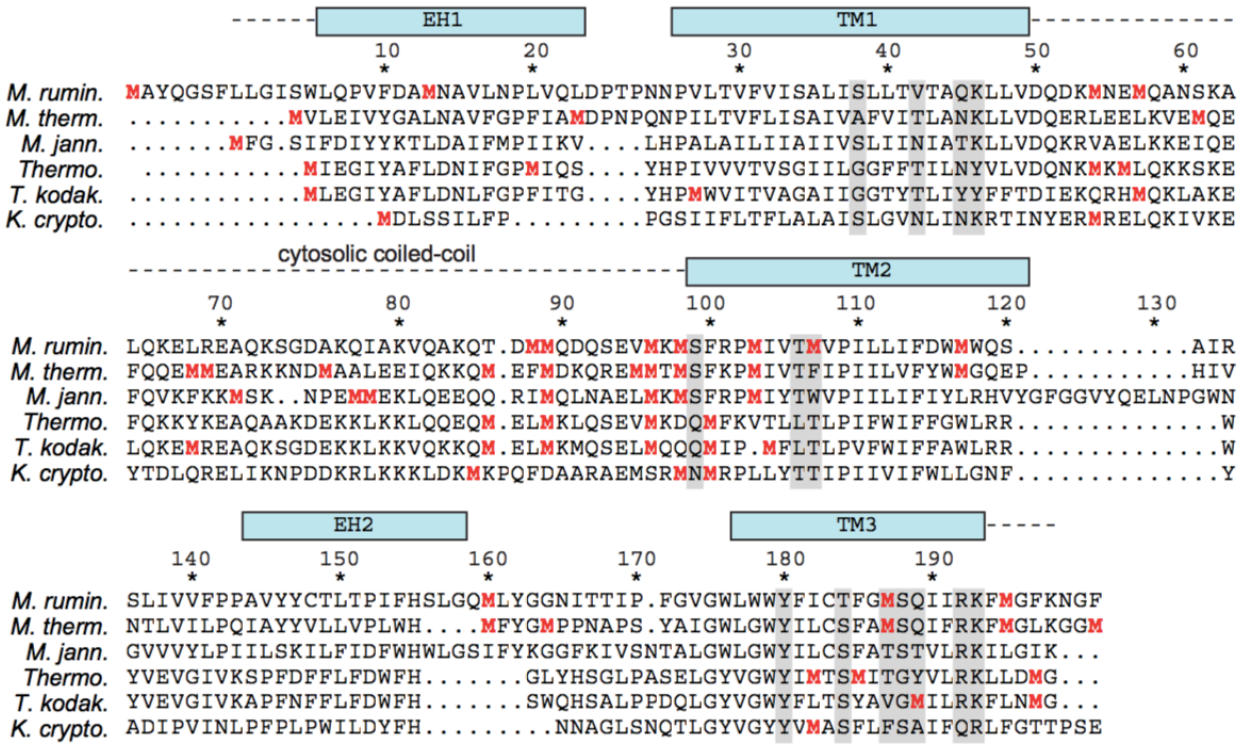


**Figure 10. Sequence alignment of representative archaeal DUF106 proteins**

Secondary structural elements and residue numbering are shown above each alignment for Mj0480. Hydrophilic residues lining the lipid-exposed groove (grey) and disordered regions (dashed lines) are indicated. All homologs are unusually rich in methionine residues (red), and these cluster near the cytosolic opening of the hydrophilic groove.

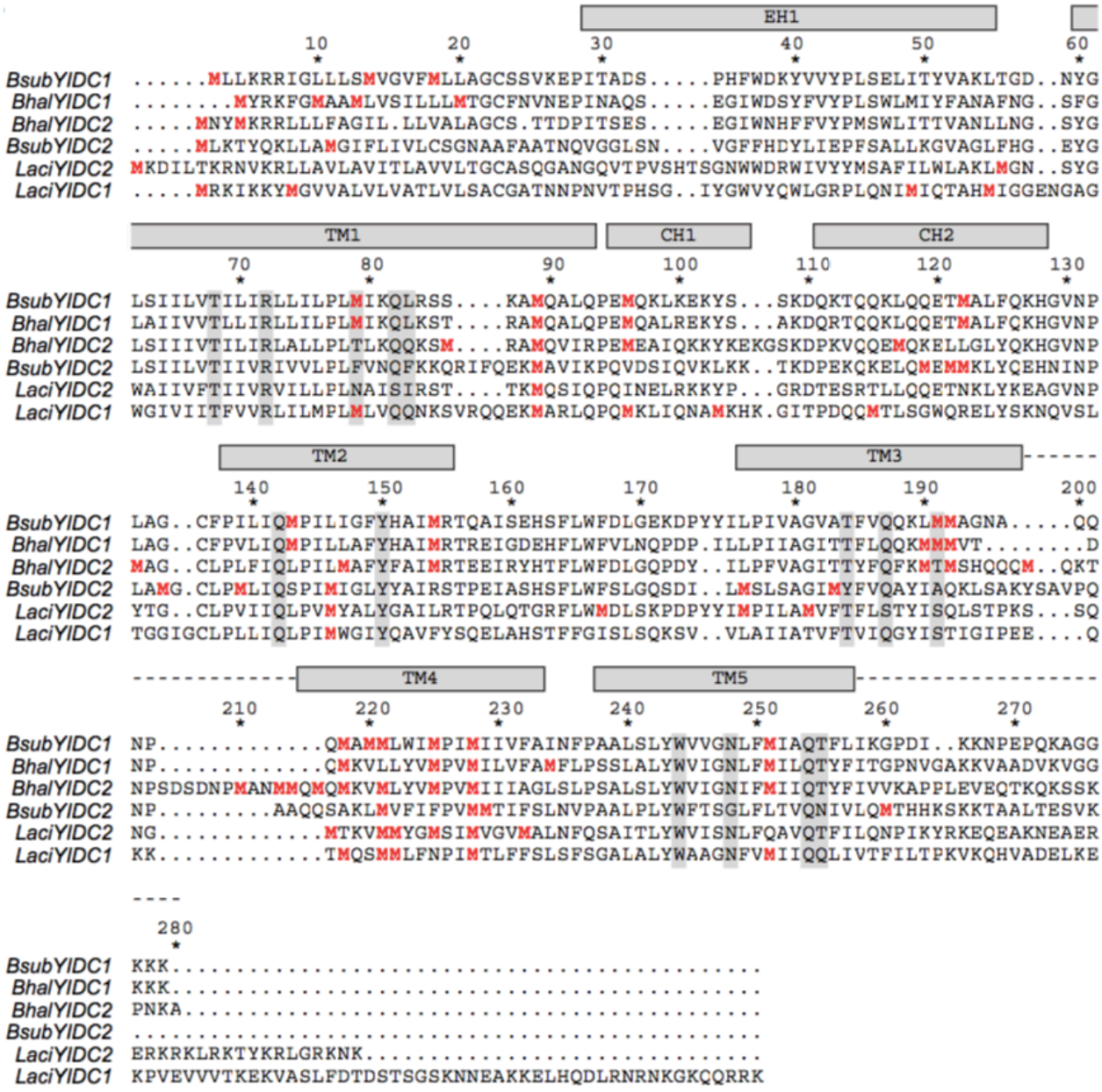


**Figure 11. Sequence alignment of representative bacterial YidC proteins**

Secondary structural elements and residue numbering are shown above each alignment for BhYidC. Hydrophilic residues lining the lipid-exposed groove (grey) and disordered regions (dashed lines) are indicated. Like the archaeal homologs, bacterial YidC proteins unusually rich in methionine residues (red), and these cluster near the cytosolic opening of the hydrophilic groove.

## 2.5 Comparison to bacterial YidC

Despite only ~14% overall sequence identity, the *Mj*0480 structure shares striking architectural similarities with the recently determined crystal structures of *B. halodurans* YidC2 (*Bh*YidC) (Kumazaki et al., 2014b) and *E. coli* YidC (Kumazaki et al., 2014a). The *Mj*0480 structure superimposes on *Bh*YidC with a root-mean-squared deviation of 3.9 Å (over 105 core residues) and a Dali similarity Z-score of 6.0 (Figure 12A) (Holm and and Park, 2000). This analysis defines a core structural motif that includes EH1, TM1, TM2 and TM3 in *Mj*0480, and EH1, TM1, TM2 and TM5 in *Bh*YidC. (Figure 12A).

In YidC, this core motif forms part of an unusual element—a hydrophilic, lipid- and cytosol-exposed groove lined with the sidechains of polar and positively charged residues (e.g., R72, Q187 and W244 in *Bh*YidC) (Figure 12C) that facilitate the insertion of membrane protein substrates (Kumazaki et al., 2014a, 2014b). Although it is constructed from three TMs (TM1-3) instead of five as in YidC (TM1-5), the *Mj*0480 groove is also lined with polar (e.g., S38, Y180, S184, S188) and positively charged (e.g., K46, R192) sidechains (Figures 12 and 13C). Thus, the presence of an unusual lipid- and cytosol-exposed hydrophilic groove is also conserved between bacterial YidC and *Mj*0480.

Members of the YidC family contain a dynamic coiled-coil motif near the cytosolic entrance of the hydrophilic groove (Figure 12A).This motif is essential for YidC function in vivo (Chen et al., 2014; Kumazaki et al., 2014b). The corresponding region in *Mj*0480 is predicted to form a coiled-coil, but is disordered in the crystal (Figure 12A). Sequence analysis reveals that the cytosolic entrance (including the coiled-coil motif and cytosolic ends of the TMs) contains an unusual abundance of methionine sidechains, a conserved feature of the archaeal DUF106 and bacterial YidC families (Figures 10 and 11). By analogy to other methionine-rich TM-binding

proteins including SRP54 (Bernstein et al., 1989; Hainzl et al., 2011; Janda et al., 2010) and Get3 (Mateja et al., 2015), we propose that this region facilitates interaction with different hydrophobic substrates. Thus, despite their low sequence identity, *Mj*0480 and the bacterial YidCs share an overall fold and key structural features.

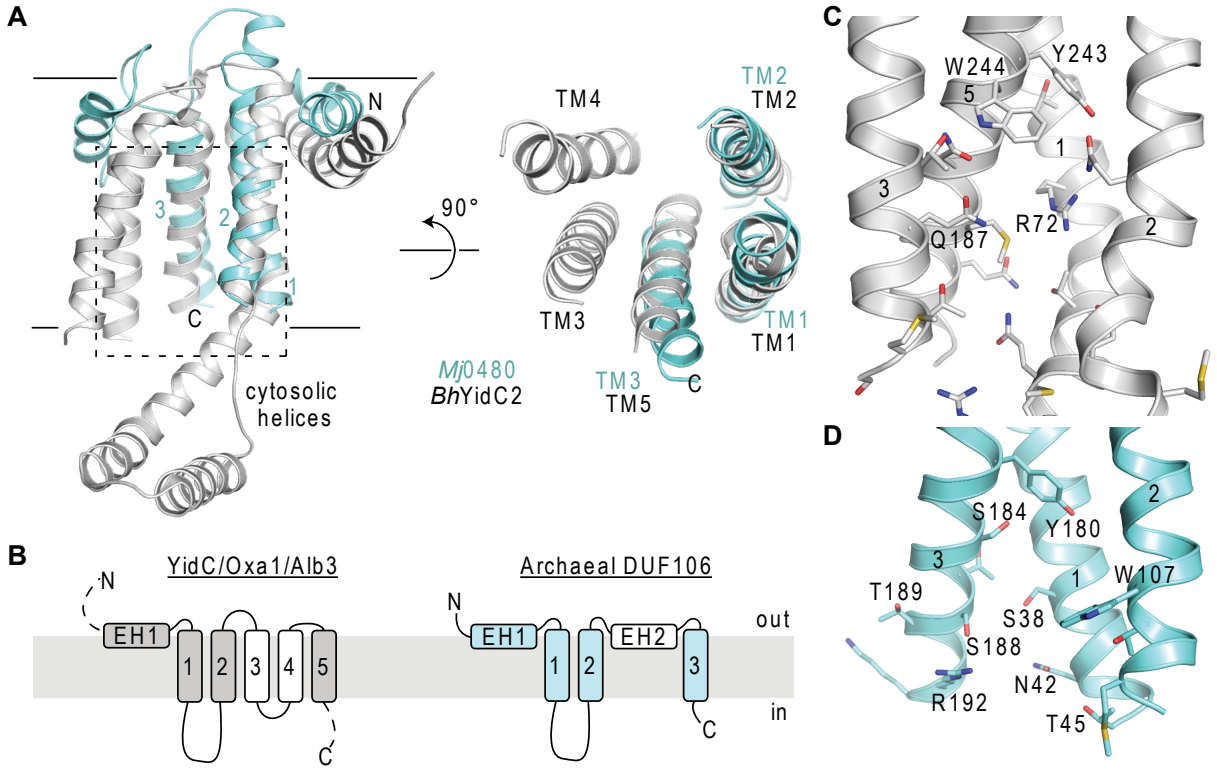


**Figure 12. Archaeal Mj0480 and bacterial YidC share key structural features.**

(A) Structure-based alignment of *Mj*0480 (light blue) and *Bh*YidC (grey; 3WO6), showing views from the plane of the membrane (left) and from the cytosol (right); the proteins superimpose with an RMSD of 3.9 Å over 105 equivalent residues (out of 141 visible).
(B) Structure-based topology cartoons for the YidC/Oxa1/Alb3 (left) and archaeal DUF106 (right) protein families.
(C) Close-up of the region highlighted in panel A (dashed box), showing details of the *Bh*YidC hydrophilic groove; TM4 has been removed for clarity. Despite its location within the lipid bilayer, the groove is lined with polar and charged residues; R72, Q187 and W244 are functionally important for membrane protein insertion (Kumazaki et al., 2014a). The cytosolic-facing vestibule and cytosolic coiled coil (only partially shown for clarity) are lined with methionine sidechains.
(D) Close-up of the corresponding region in *Mj*0480. The conserved, lipid-exposed surface in *Mj*0480 is lined with polar and charged residues (e.g., N42, T45, W107, Y180, S184, R192); the cytosolic-facing vestibule and the cytosolic coiled-coil (disordered) are also enriched in methionine residues.

## 2.6 Functional analysis of *Mj*0480

To directly test whether the archaeal DUF106 proteins function as YidC-like insertases, we attempted genetic complementation of YidC depletion in *E. coil* and of an Oxa1 knockout in *S. cerevisiae* mitochondria. Despite testing mesophilic and thermophilic archaeal DUF106 homologs, no complementation was observed. The inability of distantly related archaeal proteins to complement in these heterologous systems might reflect gross differences in function, or more subtle differences in lipid requirements, substrate specificity or interaction partners.

As an alternative to complementation we took advantage of the observation that during co-translational insertion of Sec-independent substrates, YidC and Oxa1 interact with the ribosome-nascent chain complex (RNC) (van Bloois et al., 2004; Jia et al., 2003; van der Laan et al., 2004). In *E. coli*, this is accomplished by selective binding of YidC to RNCs containing a substrate nascent chain (Kedrov et al., 2013), while mitochondrial Oxa1 recruits ribosomes even in the absence of a nascent chain (Jia et al., 2003).

To test this with archaeal *Mj*0480, we first isolated total 70S ribosomes and affinity-purified RNCs from *E. coli*. In the absence of a known archaeal substrate, we chose to monitor binding to RNCs containing the $F_0$c subunit of the $F_0F_1$ ATPase, a known substrate of both YidC and Oxa1 (van Bloois et al., 2005; Jia et al., 2007; Laan et al., 2004). To generate stable RNCs displaying only the first TM of $F_0$c, we used a previously characterized construct containing the stall sequence of SecM fused to the N-terminal region of $F_0$c (Figure 13) (Kedrov et al., 2013; Schaffitzel and and Ban, 2007).

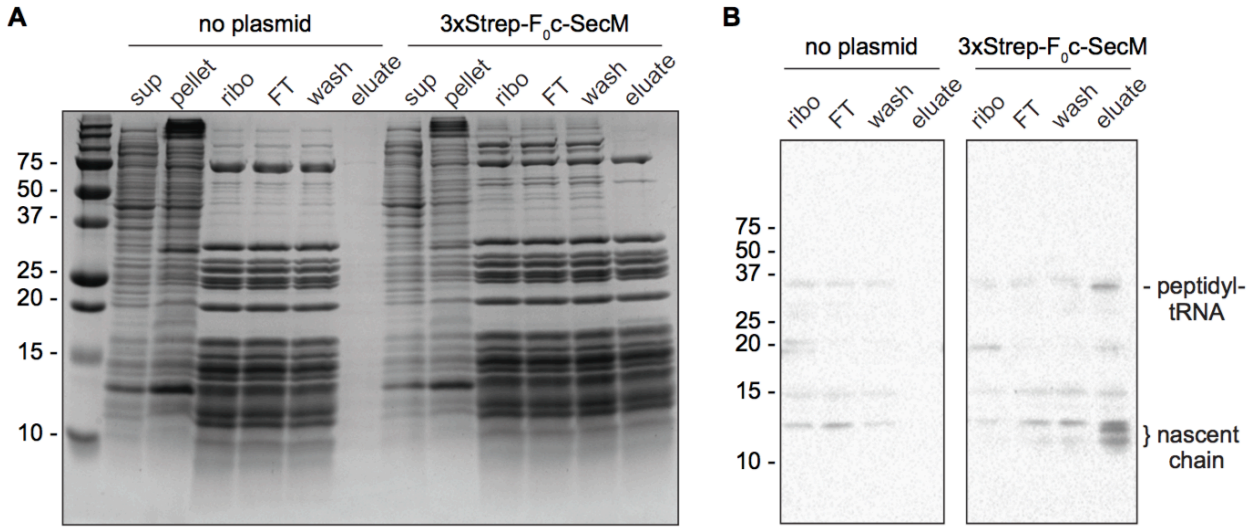


**Figure 13. Purification of stalled RNC-F$_0$c from E. coli.**

(A) Cells were transformed with a plasmid encoding an N-terminal triple Strep II tag fused to the first TMD of Foc and a SecM stall sequence. Affinity purified material was analyzed by SDS-PAGE and Coomassie blue staining. Note that multiple bands corresponding to ribosomal proteins are present in the sample containing the stalled substrate, but not in a control purification carried out in the absence of substrate.
(B) The presence of the stalled 3xStrep-F0c-SecM nascent chain (~13 kDa) and peptidyl-tRNA (~35 kDa) were confirmed on a Western blot using a StrepTactin-HRP conjugate

To determine whether Mj0480 binds ribosomes, total *E. coli* 70S ribosomes or affinity-purified, stalled RNCs were incubated with detergent-solubilized *Mj*0480, sedimented by ultracentrifugation, and analyzed for binding by immunostaining (Figure 14A). We observed binding only to stalled RNCs, suggesting that like *E. coli* YidC, *Mj*0480 binds selectively to ribosomes containing an exposed hydrophobic nascent chain.

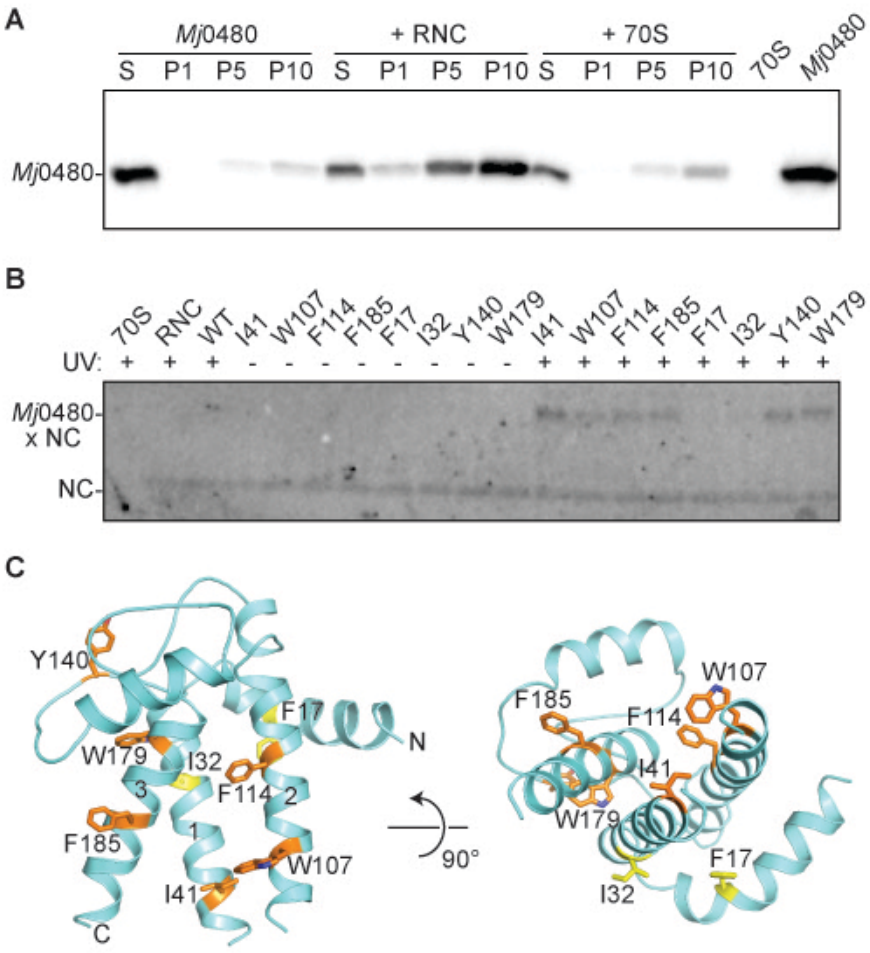


**Figure 14. Mj0480 binds selectively to RNCs displaying a YidC substrate and can be photocrosslinked to substrate via the hydrophilic groove.**

(A) Total *E. coli* 70S ribosomes (100 nM) or affinity-purified, stalled RNCs (100 nM) were incubated with recombinant *Mj*0480 (1 μM) and subjected to ultracentrifugation. Supernatant (S) and pellet (P) (1x, 5x and 10x volume equivalents of the supernatant) samples were analyzed by SDS-PAGE and immunoblotting (anti-His antibody).
(B) Stalled RNCs or total 70S ribosomes (100 nM) were incubated with proteoliposomes containing either wild-type or the indicated BpF-containing mutant (~100 nM), and subjected to UV crosslinking and analyzed by SDS-PAGE and immunoblotting (anti-Strep antibody). Uncrosslinked nascent chain (NC) and its adducts with *Mj*0480 are indicated.
(C) Photocrosslinking data mapped to the structure of *Mj*0480. Positions that crosslink to the stalled nascent chain are highlighted in orange; non-crosslinked sites are shown in yellow.

To test whether direct interactions occur between the stalled substrate and *Mj*0480, we carried out a site-specific photocrosslinking analysis. For this, we designed a series of *Mj*0480 variants containing p-benzoyl-L-phenylalanine (BpF) substitutions at eight different sites. Seven

of these are located within the bilayer, including four that line the hydrophilic groove, while the eighth site is located in a solvent accessible position on the extracellular side of *Mj*0480. After purification, BpF-containing *Mj*0480 variants were reconstituted into proteoliposomes, incubated with stalled RNCs and subjected to UV crosslinking (Figure 14B).

All four bilayer positions lining the hydrophilic groove (I41, W107, F114, F185) could be crosslinked to the stalled $F_0$c substrate, as expected for a membrane-embedded docking site (Figure 14C). In contrast, only one (W179) of the three bilayer positions distal to the hydrophilic groove (F17, I32, W179) could be crosslinked to substrate. The solvent-exposed extracellular site (Y140) could also be crosslinked, possibly via the soluble N-terminal region of the stalled $F_0$c substrate that would be accessible after translocation; however, because some *Mj*0480 might reconstitute into liposomes in the opposite orientation, the Y140 crosslink is not diagnostic for substrate insertion. Together, these data demonstrate that the stalled nascent chain makes direct contacts with the lipid-exposed hydrophilic groove of *Mj*0480, similar to what has been shown previously for *E. coli* and *B. halodurans* YidC (Klenner and and Kuhn, 2012; Kumazaki et al., 2014b; Yu et al., 2008).

## 2.7 Discussion

Our structural and biochemical analysis provides direct experimental evidence that the archaeal DUF106 proteins are members of a larger YidC/Oxa1/Alb3 superfamily. We determined the crystal structure of *Mj*0480 and found that it shares a fold with bacterial YidC (Kumazaki et al., 2014a, 2014b), including a core 'TM1-cytosolic coiled coil-TM2-X-TM3' motif (Figure 12). This scaffold presents a conserved hydrophilic surface that is exposed to the membrane bilayer and the cytosol. Importantly, we found that this surface directly contacts a stalled membrane

protein substrate (Figure 14B), similar to what has been observed previously with bacterial YidC (Klenner and and Kuhn, 2012; Kumazaki et al., 2014b; Yu et al., 2008).

A conserved function of many bacterial YidC and mitochondrial Oxa1 homologs is the co-translational, Sec-independent insertion of certain topologically 'simple' proteins into the membrane. This requires an interaction between YidC/Oxa1 and the ribosome. Oxa1 and some bacterial YidC homologs contain an extended, positively charged C-terminus that allows ribosome binding in the absence of an associated nascent chain (Jia et al., 2003; Seitl et al., 2014; Szyrach et al., 2003). In contrast, *E. coli* YidC, which lacks a C-terminal extension, binds selectively to ribosomes displaying hydrophobic nascent chains (Kedrov et al., 2013). We showed that *Mj*0480, which also lacks a C-terminal extension, binds to stalled RNC-$F_0$c but not 70S ribosomes. Taken together, these structural and functional similarities establish the universality of the YidC/Oxa1/Alb3 family in all three domains of life.

Our data are consistent with a role for *Mj*0480 in co-translational, Sec-independent insertion, but this remains to be demonstrated directly. Doing so will necessitate identification of archaeal substrates that require a DUF106 protein for insertion *in vivo,* and development of a reconstituted *in vitro* insertion assay that faithfully mimics the unique lipid composition of the archaeal plasma membrane.

It is intriguing that because *Mj*0480 lacks two of the five TMs present in bacterial YidC (*i.e.*, TM3 and TM4), its hydrophilic groove is smaller and more exposed to the lipid bilayer. This may reflect differences in substrate specificity or in mechanism of action. For example, the smaller hydrophilic groove might require *Mj*0480 to oligomerize, or to coordinate with the Sec machinery to facilitate insertion and/or folding of certain substrates, in a manner analogous to the Sec-dependent chaperone activity of bacterial YidC (Scotti et al., 2000; Urbanus et al., 2001).

Finally, we note that the core TM1-cytosolic coiled coil-TM2-X-TM3 topology observed here is predicted to be present in the eukaryotic integral membrane protein Get1/WRB, a subunit of a complex required for TA protein insertion into the ER membrane (Mariappan et al., 2011; Schuldiner et al., 2008; Vilardi et al., 2011; Wang et al., 2011a). TA proteins, like Sec-independent YidC/Oxa1/Alb3 substrates, lack large translocated regions. It will be of interest to determine if Get1/WRB belong to a superfamily of functionally diverse proteins including YidC/Oxa1/Alb3 and the archaeal DUF106 proteins, whose main similarity is mediating insertion of topologically simple membrane proteins.

# 3 Identification of Oxa1 Homologs Operating in the Eukaryotic Endoplasmic Reticulum

## 3.1 Overview

Having shown the existence of archaeal homologs of YidC, we extended our analysis to eukaryotes. Remarkably, phylogenetic analyses showed that distant homologs of YidC include the Get1 insertion factor as well as two less understood proteins, EMC3 and TMCO1. We hypothesized that these proteins are all part of a superfamily of membrane protein biogenesis factors. We refer to this here as the Oxa1 superfamily.

The two core predictions of the Oxa1 superfamily hypothesis are that the members share a common structure and function. Using evolutionary-based structural modeling and *in vivo* topology analysis, we show that Get1, EMC3 and TMCO1 share the same core fold as YidC and Ylp1. Functionally, TMCO1 was the only one of these proteins not previously linked to protein biogenesis. Here we showed that TMCO1 forms a native complex with both the Sec61 channel and the ribosome. Separate experiments show that TMCO1 has direct affinity for both the ribosome and the Sec61 channel.

A version of this work was published as Anghel, S.A., McGilvray, P.T., Hegde, R.S., and Keenan, R.J. (2017). Identification of Oxa1 Homologs Operating in the Eukaryotic Endoplasmic Reticulum. Cell Rep. 21, 3708–3716.

## 3.2 Contributions

In this work, I performed the phylogenetic analysis, the glycosylation mapping and the biochemical analyses from HEK293 cells and canine microsomes. P.T.M. performed the

recombinant TMCO1 purification and the *in vitro* ribosome binding experiments and assisted with the developing of the native purification of TMCO1-ribosome complexes. R.S.H. and R.J.K. conceived of the project. R.J.K. performed the structural modeling.

### 3.3 Phylogenetic and functional comparisons define the 'Oxa1 superfamily'

Although YidC homologs are widely conserved among bacteria and archaea (Borowska et al., 2015), none have yet been identified in the eukaryotic endomembrane system. The absence of any such homolog has been puzzling, since the eukaryotic endomembrane system is derived from invagination of the plasma membrane of a prokaryotic ancestor (Cavalier-Smith, 2002). Here we present evidence that the ER membrane possesses multiple proteins related to the Oxa1/Alb3/YidC family. These include the WRB/Get1 subunit of the TA protein insertase complex, and two less understood but highly conserved proteins, TMCO1 and EMC3. We propose that these proteins are members of a superfamily—which we designate the 'Oxa1 superfamily'—that all function broadly in membrane protein biogenesis.

In searching for archaeal homologs of the TA membrane protein insertion factor WRB/Get1, we identified a family of archaeal and eukaryotic membrane proteins annotated as 'Domain of Unknown Function 106' (DUF106) that are distantly related to the Oxa1/Alb3/YidC family (Figure 15A,B). The DUF106 group includes an archaeal family of uncharacterized membrane proteins, the eukaryotic 'ER Membrane Complex" (EMC) subunit 3 (EMC3) family, and the eukaryotic 'Transmembrane and Coiled-coil Domains 1' (TMCO1) family. DUF106 proteins appear to be phylogenetically ancient, as they are present in the Asgard Archaea, a group of organisms postulated to be the closest living relative of the last common ancestor of both archaeans and eukaryotes(Spang et al., 2015; Zaremba-Niedzwiedzka et al., 2017).

Consistent with these phylogenetic observations, there are clear functional similarities between members of the Oxa1/Alb3/YidC clade and members of the other clades for which some biochemical activity has been established (Figure 15C-F). For example, during co-translational, translocon-independent insertion of a substrate protein into the bacterial plasma membrane, YidC binds to ribosome-nascent chain complexes (RNC) and directly contacts the hydrophobic nascent chain(Kumazaki et al., 2014a, 2014b). Similarly, the archaeal DUF106 protein *Mj*0480 (henceforth called the "YidC-like protein 1" or Ylp1) binds RNCs, and can be crosslinked to a model substrate in vitro (Borowska et al., 2015). Moreover, the known translocon-independent substrates of YidC and Oxa1, and the post-translational substrates of Alb3 and WRB/Get1 are all simple membrane proteins with few transmembrane helices and small translocated regions (Aschtgen et al., 2012; Hegde and Keenan, 2011; Wang and Dalbey, 2011). Finally, although its precise function remains to be defined, the EMC has been linked to ERAD and biosynthesis of multi-pass membrane proteins (Richard et al., 2013; Satoh et al., 2015). Given these phylogenetic and functional similarities, we propose to assign these proteins as members of a superfamily, which we hereafter refer to as the 'Oxa1 superfamily'.

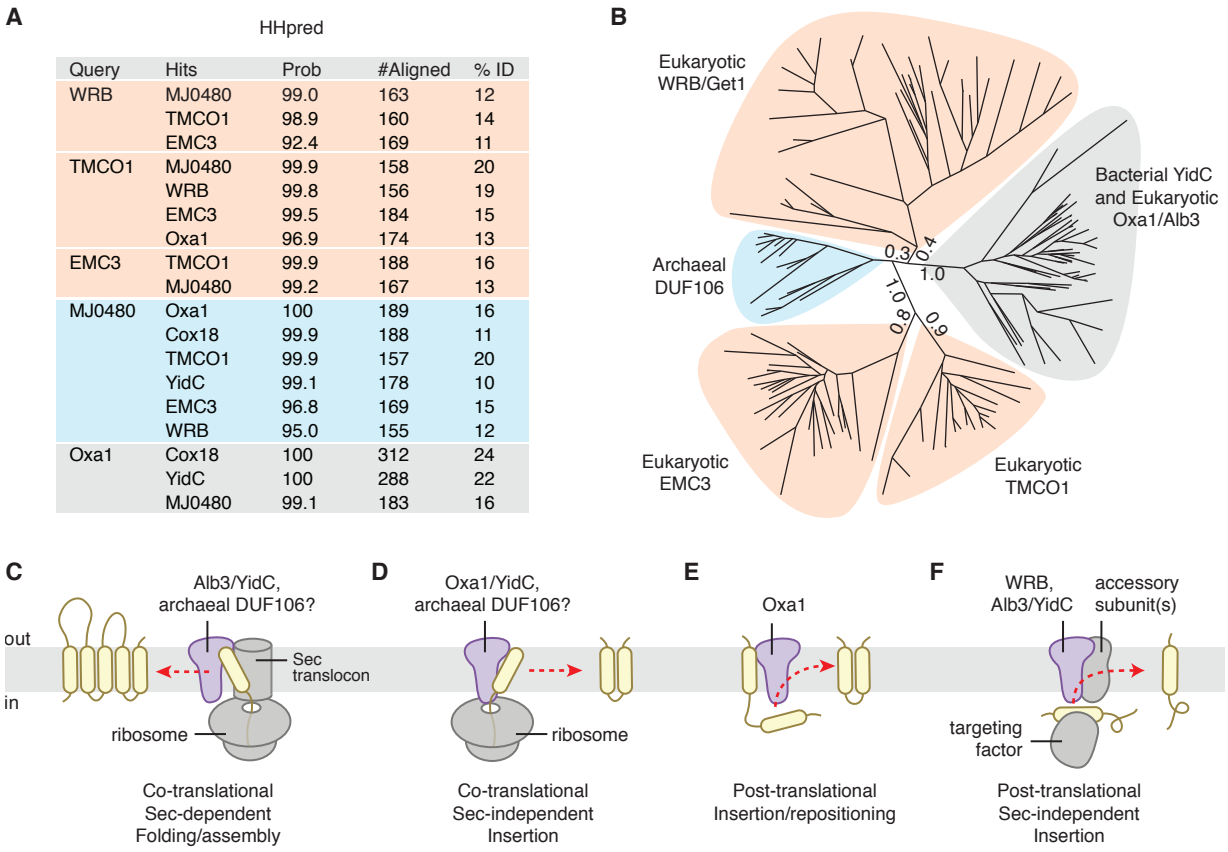


**Figure 15. Phylogenetic and functional comparison defines the Oxa1 superfamily.**

(A) Identification of remote DUF106 homologs using HHpred. Eukaryotic (human), bacterial (*E. coli*) and archaeal (*M. jannaschi*i) proteomes were searched for each query (UniProt ID: WRB=O00258; Oxa1=Q15070; TMCO1=Q9UM00; EMC3=Q9P0I2; Ylp1=Q57904) using default settings in HHpred in 'global' alignment mode. Top hits are listed, along with the HHpred probability score, the number of residues aligned, and the sequence identity.
(B) Maximum likelihood tree of representative sequences. Branch lengths for the five main clades are indicated.
(C) During Sec-dependent, co-translational assembly and folding, substrates are delivered to the membrane by the ribosome; insertion requires participation of the Sec translocon. Substrates of this pathway typically contain multiple TMDs and/or large translocated regions. Superfamily members exemplifying this activity include bacterial YidC and chloroplast Alb3.
(D) During Sec-independent, co-translational insertion, topologically 'simple' substrates that lack large or highly charged translocated regions are delivered to the membrane by the ribosome. Superfamily members exemplifying this activity include Oxa1 and YidC; archaeal Ylp1 proteins function similarly in vitro.
(E) Post-translational TMD repositioning, exemplified by Oxa1.
(F) During Sec-independent, post-translational insertion, topologically simple substrates are delivered to the membrane by soluble targeting factors. Superfamily members exemplifying this activity include WRB/Get1, which inserts tail-anchored proteins delivered by TRC40/Get3, chloroplast Alb3, which inserts specific proteins delivered to the thylakoid membrane by cpSRP43, and bacterial YidC.

### 3.4 Oxa1 superfamily members share membrane topology and core structural features

A key prediction is that, owing to their common ancestry and conserved function, all members of the Oxa1 superfamily share a common architecture. As noted previously (Borowska et al., 2015), comparison of the crystal structures of bacterial YidC (Kumazaki et al., 2014a, 2014b) and archaeal Ylp1 (Borowska et al., 2015) reveals considerable structural overlap, including a three-TMD core, an N-in/C-out orientation, a cytosolic coiled coil between the first two TMDs, and a lipid-exposed hydrophilic groove which has been shown to contact substrate proteins (Figure 17A).

Secondary structure and topology predictions for Get1, TMCO1 and EMC3 suggest they share this architecture (Figure 16B), but the topology of these proteins has not been conclusively established. Indeed, a recent study proposed that TMCO1 has an N-in/C-in topology, with only two TMDs and a lumenal-facing coiled-coil (Wang et al., 2016); this topology is incompatible with placement of TMCO1 into the Oxa1 superfamily.

To define the topology of Get1, TMCO1 and EMC3, we designed 3xFlag-tagged constructs containing a consensus glycosylation sequence at the N- or C-terminus, or within the predicted cytosolic coiled coil or lumenal loop regions (Figure 16B and 17). In all cases, we observed glycosylation of the N-terminus and the loop between the second and third TMDs, and no glycosylation of the C-terminus or the coiled-coil domain (Figure 16C). These data are consistent with the observation that the Get1 coiled-coil binds to the cytosolic targeting machinery (Mariappan et al., 2011; Stefer et al., 2011; Wang et al., 2011a), and with proteomic analyses showing that serine residues in the coiled-coil and C-terminal regions of TMCO1 are phosphorylated by cytosolic kinases (Dephoure et al., 2008; Olsen et al., 2010).

We also performed an unbiased, 3D structure prediction of TMCO1, Get1 and EMC3 using distance restraints derived from evolutionarily coupled residue pairs (Wang et al., 2017). Remarkably, the top-ranked models for human TMCO1 and yeast Get1 recapitulated the core structural features of bacterial YidC and archaeal Ylp1 proteins, including a lumenal N-terminus, cytosolic-facing coiled-coil and C-terminus, and a three TMD core (Figure 16D and 17). The top-ranked EMC3 models also possessed a three TMD core and a coiled-coil motif between the first two TMDs but showed physically implausible orientations for the coiled-coil and C-terminus (Figure 17); this may reflect the limited number of available sequence homologs, the relatively larger size of EMC3, and the absence of any membrane bilayer energy term. Nevertheless, these models suggest that members of the Oxa1 superfamily share a membrane topology and core architecture.

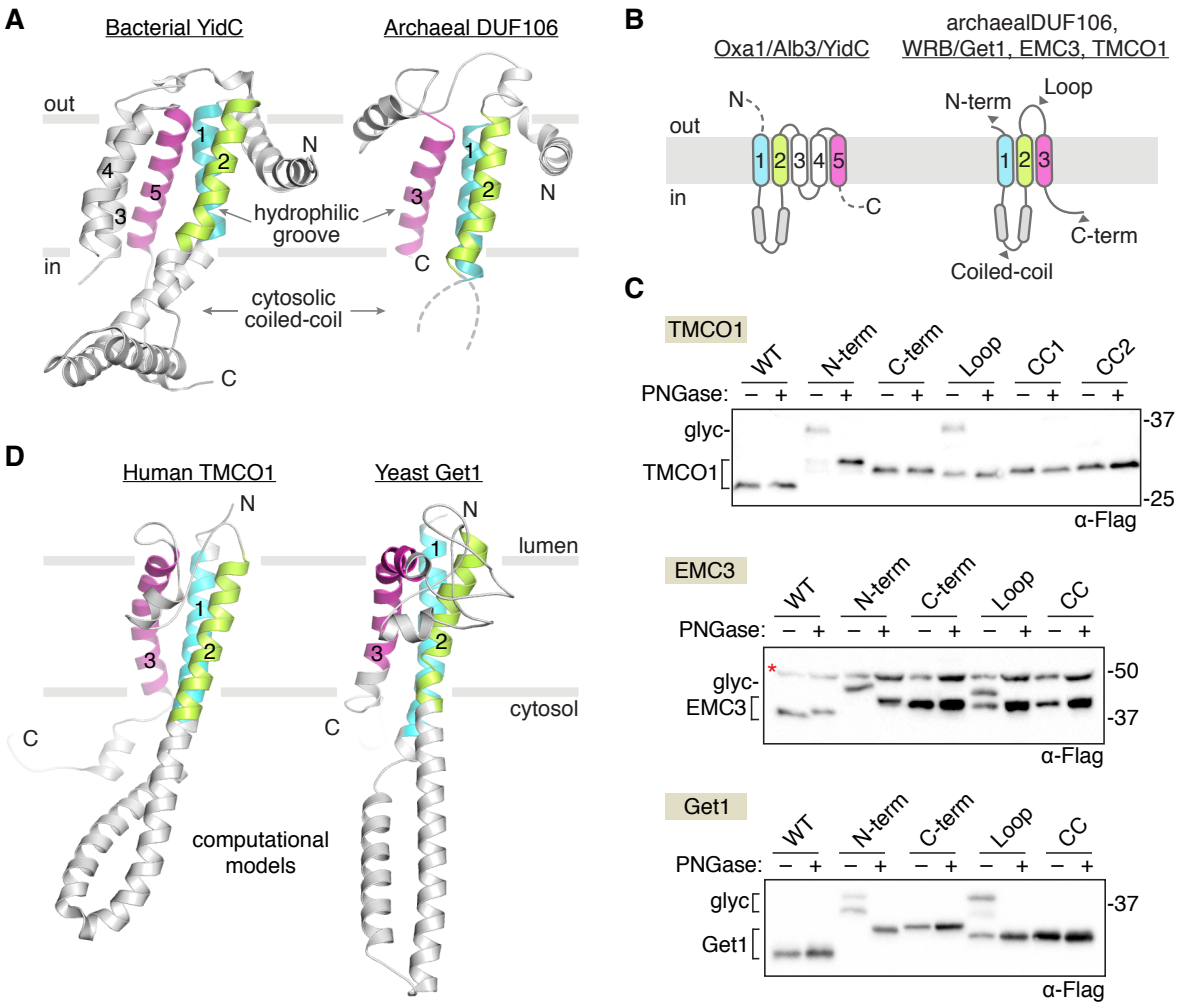


**Figure 16. Oxa1 superfamily members share a membrane topology and core structure.**

(A) Comparison of known structures from two clades: bacterial YidC (left; PDB 3WO6) and archaeal Ylp1 (right; PDB 5C8J). These proteins share a common N-out/C-in topology, a cytosolic-facing coiled-coil between TM1 and TM2 (disordered in the archaeal structure) and a three TMD core (colored) that harbors a lipid-exposed hydrophilic groove implicated in binding to nascent polypeptides during insertion.
(B) Predicted topology of the Oxa1 superfamily members.
(C) Experimentally defined topology of human TMCO1, EMC3 and yeast Get1. Glycosylation acceptor sequences were introduced at the indicated positions (grey arrowheads in (B)), and glycosylation was monitored by western blotting after treatment with or without PNGase F. All three proteins conform to the predicted Oxa1 superfamily topology. A non-specific, cross-reacting band visible in all EMC3 samples is marked (red asterisk).
(D) Evolutionary covariation-based computational 3D models of human TMCO1 and yeast Get1 recapitulate the core structural features of bacterial YidC and archaeal Ylp1: lumenal N-terminus, cytosolic-facing coiled-coil and C-terminus, and a three TMD core. Here the predicted coiled-coil region of Get1 has been replaced with the experimentally-determined structure of the Get1 coiled-coil (PDB 3ZS8).

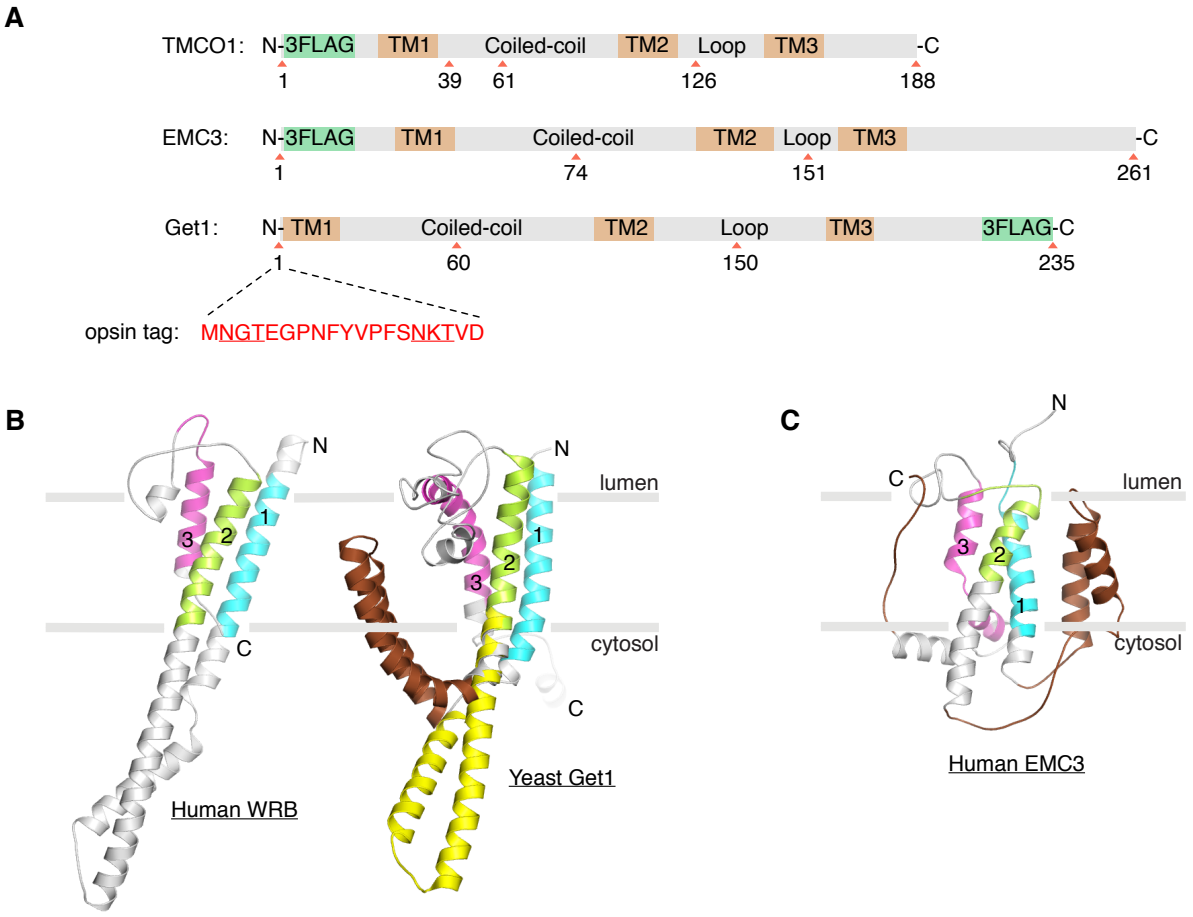


**Figure 17. Additional details for the topology mapping experiments and 3D modeling.**

(A) Constructs used for glycosylation mapping. An opsin tag (red) containing two N-glycosylation sites (underlined) was inserted at the indicated positions of human TMCO1, human EMC3 and yeast Get1. Tag positions correspond to the native (untagged) sequence. For the TMCO1 and EMC3 constructs, a GSS linker connects the 3xFlag tag and the protein sequence. For the N-terminally opsin-tagged Get1 sequence, a 3xGSS linker was inserted before the first TMD, as sufficient distance from the membrane is required for effective glycosylation.

(B) Co-variation-based 3D models of human WRB (left) and yeast Get1 (right), as in Figure 2D; note how the highly charged coiled-coil region of yeast Get1 (brown) bends back into the membrane bilayer (grey bars) in a non-physiologic conformation; this is likely due to the lack of a membrane bilayer energy term during 3D modeling (see Methods). In this case, a better, hybrid model is obtained by replacing the distorted coiled-coil (brown) with a crystallographically-defined Get1 coiled-coil (yellow; PDB 3ZS8) by manually docking it as a rigid body between TM1 and TM2

(C) Co-variation based 3D model of human EMC3 colored as in Figure 2D; a coiled-coil motif between TM1 and TM2, and the three TM core are both visible. However, similar to the yeast Get1 model, the coiled-coil and extended C-terminal region (both features colored brown) adopt physically implausible orientations in which they become embedded in the bilayer, despite being highly charged.

### 3.5 TMCO1 interacts with the ribosome and the Sec61 channel

A second prediction of the Oxa1 superfamily model is that all of the proteins function in some capacity in membrane protein biogenesis. To test this, we focused on human TMCO1, the only member of the superfamily not yet linked to membrane protein biogenesis. TMCO1 is an ER resident membrane protein that is conserved in most eukaryotes(Iwamuro et al., 1999). Genetic variations around *TMCO1* are linked to glaucoma (Burdon et al., 2011; Sharma et al., 2012) and nonsense mutations cause a disorder associated with craniofacial dysmorphisms, skeletal anomalies, and intellectual disability(Alanay et al., 2014; Caglayan et al., 2013; Xin et al., 2010).

We asked whether any of the interactions of TMCO1 are similar to those of the better characterized members of the Oxa1 superfamily. In the case of bacterial YidC, primary interaction partners include the Sec translocon and the ribosome (Figure 15C,D). We first explored whether TMCO1 is part of complex with translating ribosomes, as would be expected if it functions in co-translational insertion like some members of the Oxa1 superfamily (Figure 15C,D).

When digitonin-solubilized HEK293 membranes were fractionated on a sucrose gradient, TMCO1 and Sec61 were present in the 80S ribosome fraction (Figure 18A). In contrast, Derlin-1, an abundant ER membrane protein not known to bind the ribosome, did not co-migrate with ribosomes. Next, we tested whether TMCO1 and Sec61 are part of the same ribosome-bound complex. After immunoprecipitating digitonin-solubilized membranes prepared from a 3xFlag-tagged TMCO1 HEK293 cell line (Figure 19A), we observed a complex containing TMCO1, Sec61 and ribosomes (Figure 18B). Thus, TMCO1-Sec61-ribosome complexes can be isolated from cells under native conditions.

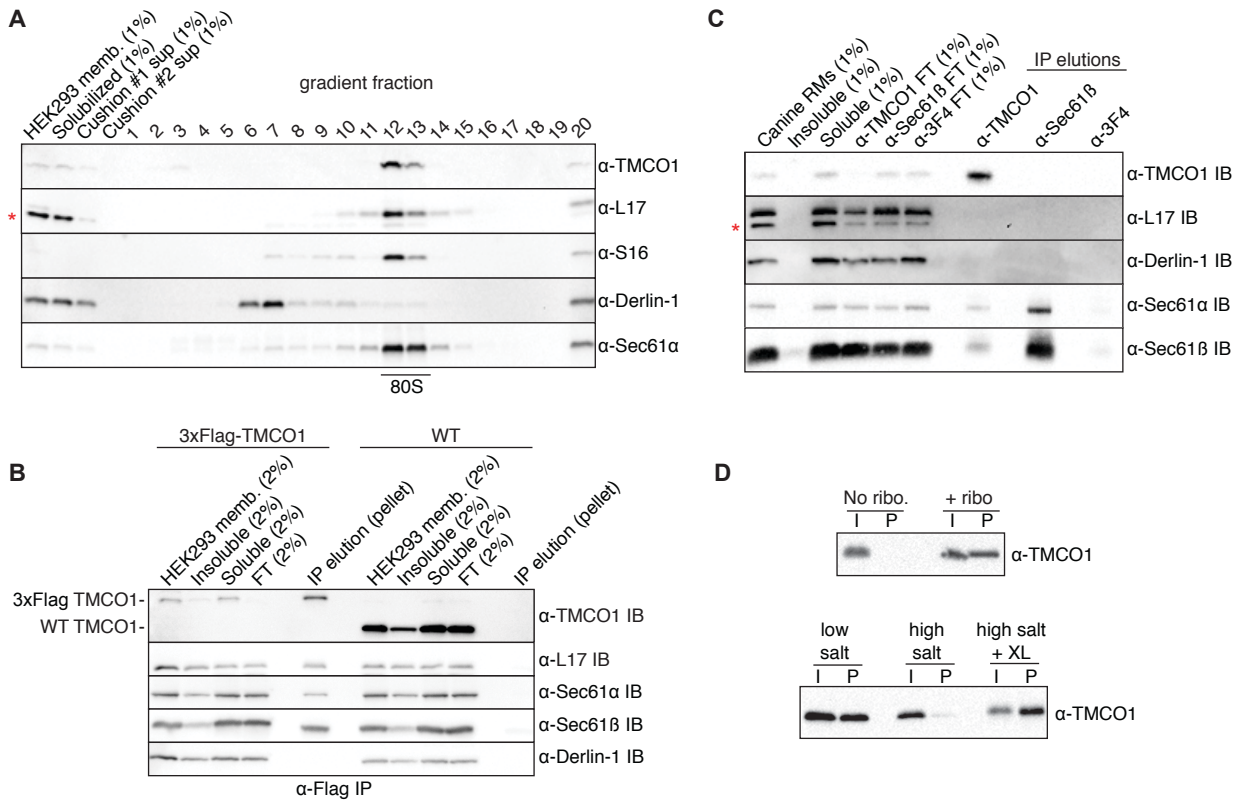


**Figure 18. TMCO1 forms a complex with the Sec61 channel and RNCs.**

(A) HEK293 membranes were solubilized with digitonin, fractionated by sucrose cushion, separated on a high resolution sucrose gradient and analyzed by western blotting. TMCO1 co-fractionates with intact 80S particles and the Sec61 channel, but not the unrelated ER membrane protein Derlin1, which does not bind to ribosomes. Blots for the large (L17) and small (S16) ribosomal subunits are also shown; a non-specific, cross-reacting band visible in the L17 blot is indicated with an asterisk.

(B) Digitonin-solubilized membranes from wild-type (WT) HEK293 cells or a HEK293 cell line containing an N-terminal 3xFlag-tagged TMCO1 allele were analyzed by anti-Flag immunoprecipitation, sucrose cushion and western blotting.

(C) Digitonin-solubilized canine pancreatic rough microsomes were tested for interaction between TMCO1 and Sec61 by co-immunoprecipitation and western blotting. An anti-TMCO1, but not a control anti-3F4 antibody, pulls down two components of Sec61. The absence of TMCO1 in the reciprocal pull-down is consistent with the higher levels of Sec61 in these membranes.

(D) Recombinant, purified TMCO1 co-sediments with unprogrammed ribosomes isolated from rabbit reticulocyte lysate (top panel). This interaction is salt-sensitive, and can be stabilized by chemical crosslinking (XL) (bottom panel). The pellet (P) fractions correspond to 5x volume equivalents of the input (I) fractions.

We next explored whether TMCO1 can exist in complex with Sec61 in the absence of ribosomes, as is true for YidC (Botte et al., 2016; Duong and and Wickner, 1997). To identify ribosome-independent complexes we used antibodies that recognize cytosolic epitopes on TMCO1 and Sec61β that are expected to be occluded by a bound ribosome. After immunoprecipitating digitonin-solubilized canine pancreatic microsomes (which contain high levels of Sec61), the anti-TMCO1 antibody pulled-down components of the Sec61 translocon (Figure 18C). As expected, none of the antibodies pulled-down ribosomes or the control protein, Derlin-1. This suggests that TMCO1 and Sec61 can exist in the same complex in the absence of ribosomes.

Finally, we asked whether TMCO1 has an intrinsic affinity for ribosomes, as is the case for Oxa1 and some YidC homologs with long, positively charged C-terminal regions(Jia et al., 2003; Seitl et al., 2014). To test this, we incubated recombinant, purified TMCO1 (Figure 19B) with unprogrammed ribosomes from a rabbit reticulocyte lysate. After sedimentation through a sucrose cushion, we observed ribosome-dependent pelleting of TMCO1 (Figure 18D). This interaction was salt-sensitive, could be stabilized by chemical crosslinking and is specific, since high concentrations of bulk RNA did not disrupt the interaction (Figure 18D and 19C,D). Thus, in addition to its conserved structural features, TMCO1 shares key functional properties with members of the Oxa1/Alb3/YidC family, consistent with the predictions of the Oxa1 superfamily hypothesis.

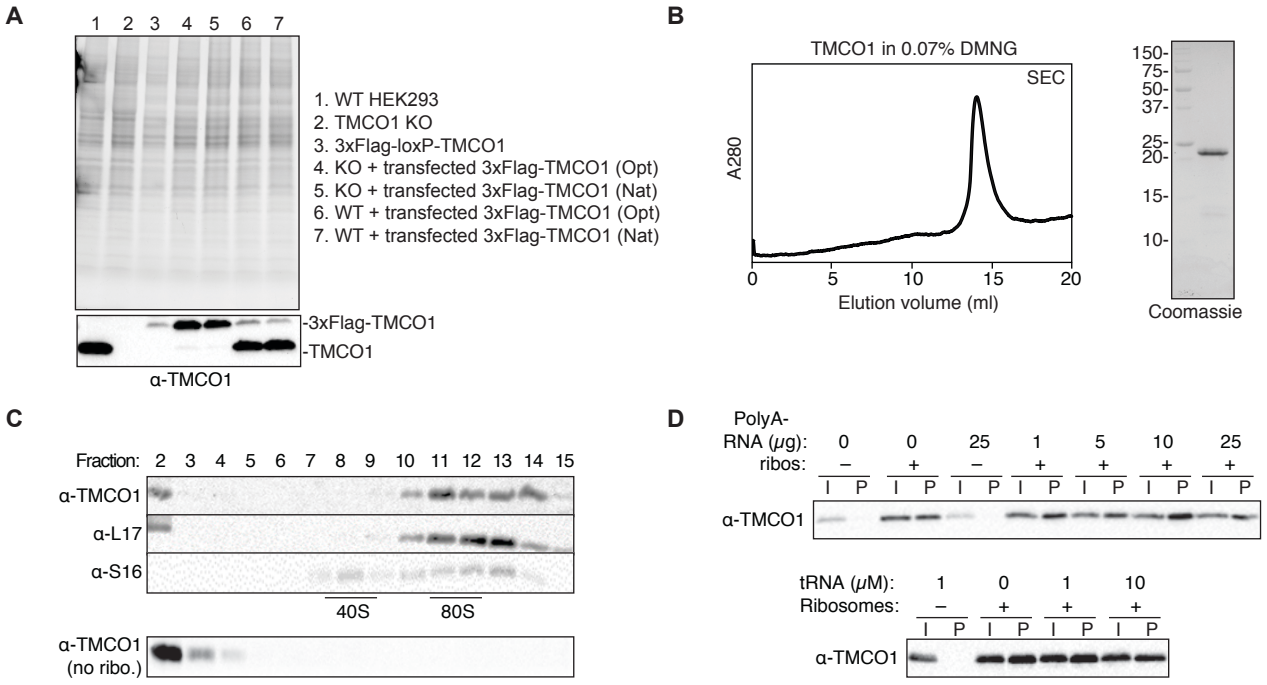


**Figure 19. Additional characterization of the ribosome binding properties of TMCO1.**

(A) Western blot analysis of TMCO1 expression levels in wild-type (WT) HEK293 cells, CRISPR/Cas9 generated knockout (KO) HEK293 cells, an integrated 3xFlag-tagged TMCO1 cell line and either KO or WT cells transfected with a 3xFlag-tagged TMCO1 construct either with ('Opt') or without ('Nat') codon optimization. A stain-free image of the gel prior to PVDF transfer shows that equal amounts of protein were loaded in each lane. Note that the transfected constructs express at lower levels than endogenous TMCO1 ('WT', lane 1).
(B) Size-exclusion chromatography (SEC) of Ni-NTA affinity purified, recombinant TMCO1 in DMNG; pooled fractions are shown at right.
(C) Sucrose gradient analysis of recombinant TMCO1 after chemical crosslinking to nuclease-treated rabbit reticulocyte lysate ribosomes. TMCO1 co-sediments with 80S ribosomes (but not the 40S ribosomal small subunit), while free TMCO1 remains at the top of the gradient.
(D) Sedimentation analysis of TMCO1-ribosome complexes in the presence of excess competitor RNA; assays contained 1 μM TMCO1, 0.1 μM ribosomes and the indicated concentrations of competitor RNA.

## 3.6 Discussion

Our phylogenetic, topological and functional data identify an unexpected evolutionary relationship among a diverse group of integral membrane proteins that together define the 'Oxa1 superfamily'. These proteins include bacterial YidC and its homologs in mitochondria and

chloroplasts, archaeal Ylp1 proteins, and three ER-resident proteins: WRB/Get1, EMC3 and TMCO1. The best characterized members of the superfamily function in membrane protein biogenesis (Figure 15C-F). In particular, Oxa1/Alb3/YidC proteins facilitate the insertion, folding and/or assembly of a variety of membrane proteins(Wang and Dalbey, 2011), while the WRB/Get1 subunit of the GET pathway transmembrane complex mediates the insertion of TA membrane proteins into the ER (Hegde and Keenan, 2011). Similarly, the EMC3 subunit of the 'ER membrane complex' has been proposed to play a role in membrane protein quality control (Richard et al., 2013) and biogenesis (Satoh et al., 2015).

The function of TMCO1 has been less clear. Here we show that TMCO1 possesses an Oxa1-like architecture, and that TMCO1-Sec61-ribosome complexes can be isolated from HEK293 cells under native conditions. We also show that TMCO1 can be isolated in ribosome-free complexes with Sec61, and that TMCO1 has an intrinsic affinity for ribosomes. These properties suggest that TMCO1 functions most analogously to bacterial YidC, and may facilitate the co-translational insertion, folding and/or assembly of newly synthesized membrane proteins into the ER membrane (Figure 15C,D).

This assignment is not incompatible with the previous proposal that TMCO1 functions as a $Ca^{2+}$ channel (Wang et al., 2016). Indeed, other well-characterized membrane protein insertases, including the bacterial and eukaryotic Sec translocon (Sachelaru et al., 2017; Simon and Blobel, 1991; Simon et al., 1989; Wirth et al., 2003) and mitochondrial Oxa1(Krüger et al., 2012), have also been shown to conduct ions. This activity is likely related to their ability to translocate polypeptides across a membrane bilayer, and the same may be true for TMCO1. Alternatively, TMCO1 may modulate the $Ca^{2+}$ efflux properties of Sec61 (Erdmann et al., 2011; Lang et al., 2011) or facilitate the biogenesis of a protein that functions in $Ca^{2+}$-transport.

We speculate that Oxa1 superfamily proteins are all descendants of an ancestral machine that could insert topologically 'simple' membrane proteins into the bilayer. Over time, the need to handle more complex substrates with additional TMDs and/or larger translocated regions would have been satisfied by evolution of the translocon. Subsequently, Oxa1 superfamily members would have been freed to evolve more specialized functions in concert with other membrane-bound and soluble factors. This might manifest in the translocon-dependent chaperone activities of YidC and Alb3, and the evolution of eukaryotic WRB/Get1 and EMC3 to function in association with other integral membrane components. Likewise, adaptation of WRB/Get1 and Alb3 to post-translational insertion would have resulted from modification of their cytosolic-facing coiled-coil and C-terminus for binding to the TRC40/Get3 and cpSRP54 targeting factors, respectively, instead of the ribosome.

The Oxa1 superfamily illustrates how a single structural scaffold has been diversified to handle the insertion, folding and assembly of different proteins into different cellular membranes. The shared characteristics of Oxa1/Alb3/YidC and WRB/Get1 translocon-independent substrates raises the possibility that Oxa1 superfamily members might, under certain circumstances, act on overlapping sets of substrates in the ER. Consistent with this, it is notable that disruption of WRB (Sojka et al., 2014; Vogl et al., 2016), TMCO1 (Caglayan et al., 2013; Xin et al., 2010) or EMC3 (Ma et al., 2015) is non-lethal. Such functional redundancy would impart robustness to membrane protein biogenesis, particularly under conditions of stress (Aviram and and Schuldiner, 2014; Aviram et al., 2016). Identifying the native substrates and molecular mechanisms underlying EMC3- and TMCO1-mediated biogenesis are important topics for future investigation.

# 4 TMCO1 Works on Multispanning Membrane Proteins and May Connect Biogenesis and Quality Control

## 4.1 Overview

Our work up to this point had suggested TMCO1 functions as a protein biogenesis factor, as it directly associates with the ribosome and the Sec61 channel. However, the exact identity of TMCO1 clients was not known. Based on the Oxa1 superfamily framework, the expectation was that TMCO1 will function to assist with the insertion, folding or quality control of membrane proteins. Whether TMCO1 functions in direct insertion of small membrane proteins (like Get1 and the Sec-independent function of YidC) or in the folding/assembly of large, multispanning membrane proteins (like the Sec-dependent function of YidC) or has a yet undescribed function remained unknown. The observation that TMCO1 is in complex with the Sec61 channel favored a folding function for multispanning proteins. Here, we confirmed a role for TMCO1 in biogenesis of multispanning membrane proteins by showing that TMCO1-attached ribosomes specifically translate multispanning membrane proteins. Furthermore, mass spectrometry analysis of the TMCO1-ribosome complexes showed they contain low amounts of TRAP, Ribophorin I and Calnexin. Additionally, the complexes are rich in CCDC47, Nicalin and NOMO, proteins previously linked to protein quality control but not shown to act co-translationally. This suggests that TMCO1 may be part of a pathway that connects membrane protein biogenesis and quality control.

## 4.2 Contributions

I performed all experiments. RNA sequencing was performed at The University of Chicago Genomics Core Facility. Mass Spectrometry was performed at the Harvard University FAS Center for Systems Biology. R.J.K and I performed the analysis of the RNA sequencing data.

## 4.3 Profiling of TMCO1 Clients Reveals a Bias for Multipass Membrane Proteins

To accurately characterize TMCO1 function, the first step is a comprehensive characterization of TMCO1 client proteins. To accomplish this task, we took advantage of our Flag-pulldown approach which allows us to isolate native TMCO1/Ribosome/Sec61 complexes. Because TMCO1 only binds to a minority of ribosomes, we reasoned that these complexes contain ribosomes translating TMCO1 clients. We isolated 3xFlag-TMCO1/Ribosome complexes as described in Figure 19B, followed by a Trizol RNA extraction, cDNA library preparation and high throughput Illumina sequencing. A similar strategy was utilized recently to identify clients of the Signal Recognition Particle (Chartron et al., 2016; Schibich et al., 2016) and bacterial Trigger Factor (Oh et al., 2011). Because past experiments of this type did not employ membrane proteins and thus the effect of detergents was unclear, we performed this experiment with both digitonin and DDM. The results were qualitatively similar but usage of DDM resulted in higher background (as judged by the differences between the Flag pulldowns from 3xFlag-TMCO1 and control wild type cells).

We first asked whether this immunoprecipiation strategy selectively enriched certain transcripts. For this we first calculated a "corrected abundance" for each transcript, representing the difference in transcript abundance in the 3xFlag-TMCO1 and wild-type

immunoprecipitations. Regardless if digitonin or DDM was utilized, we observed a significant negative correlation between the abundance of a transcript in the membrane fraction overall and in the pulldown sample (Pearson's correlation coefficient of -0.47 and 0.61 for digitonin and DDM respectively) (Figure 20A). This suggested that the immunoprecipitation selectively enriched a subpopulation of membrane-associated mRNAs that are normally present at lower levels. Such a result is expected if TMCO1 clients are membrane proteins, which are expressed at a lower level than soluble proteins. Furthermore, when compared to the total membrane fraction, the pulldowns were highly enriched in secretory pathway proteins as judged by their Uniprot localization annotation (Figure 20B). Conversely, transcripts annotated as encoding for cytosolic, nuclear and mitochondrial proteins were highly depleted by the pulldown procedure. Transcripts annotated as "secreted" were also depleted even though they are part of the secretory pathway, again consistent with TMCO1 acting exclusively on membrane proteins.

Since our hypothesis is that TMCO1 functions in membrane protein biogenesis, we next asked whether genes encoding membrane proteins are overrepresented in TMCO1-associated transcripts. Of the protein-coding genes represented among the 1,000 transcripts most enriched by immunoprecipiation, 98% are predicted to be transmembrane proteins (Figure 21A). This compares favorably with the percentage of transmembrane proteins in the total membrane fraction RNA (35%) and the entire human genome (30%) (Linn et al., 2010). This strongly suggests that TMCO1, like all other proteins of the Oxa1 superfamily, functions in the biogenesis of transmembrane proteins.

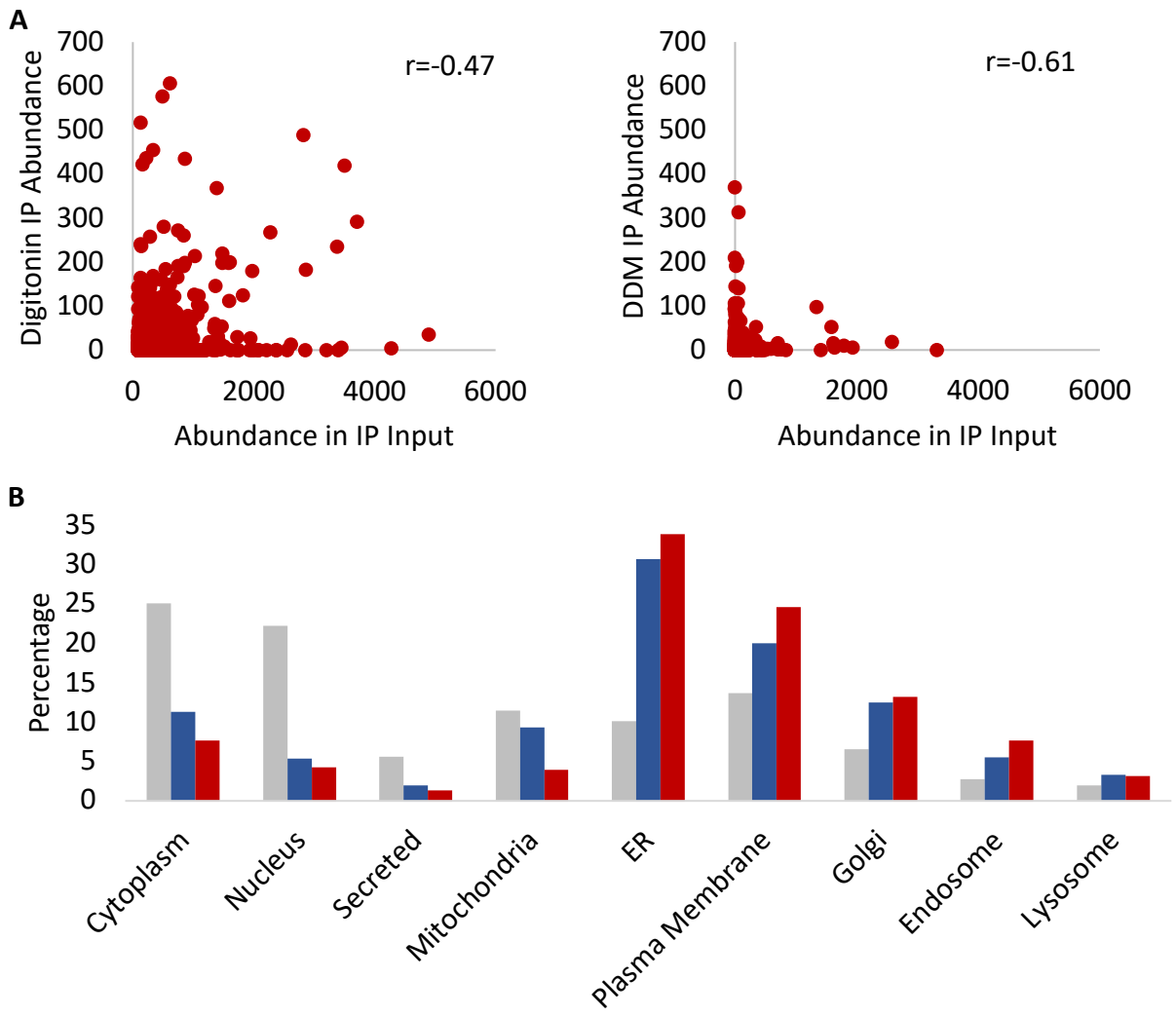


**Figure 20. Selective enrichment of a subset of TMCO1-associated mRNAs.**

(A) For the 1000 most abundant transcripts in the digitonin (left) and DDM (right) immunoprecipitation experiments, the IP corrected abundance was plotted against the overall abundance in the Hek293 membrane fraction. Values indicate transcripts per million (tpm). The Pearson correlation values (calculated for the entire datasets not just the 1000 displayed in the graph) are indicated.
(B) Uniprot annotated localization for the protein products of the transcripts in the total membrane fraction (grey) and the 1000 most enriched transcripts in the DDM (blue) and digitonin (red) IP experiments. Peroxisomal localization was omitted for clarity, as it accounts for less than 1% of the proteins in all datasets.

We next asked whether TMCO1 functions as a biogenesis factor for a particular type of membrane protein, as do Get1, YidC and likely EMC3. To this end, we compared the number of transmembrane domains – as predicted by TOPCONS (Tsirigos et al., 2015) – present in the

TMCO1 immunoprecipitated dataset and the total membrane transcript dataset (Figure 21B). Transcripts encoding single TMD proteins—by far the most abundant type of membrane protein in the human genome—were strongly depleted (>20-fold) in the TMCO1-associated dataset relative to the total membrane dataset. In contrast, multispanning membrane proteins were enriched among TMCO1-associated transcripts. Transmembrane proteins with 4 TMDs or more were enriched by at least a factor of 2 when compared to both the total membrane dataset and the previously reported distribution for human membrane proteins in general (Linn et al., 2010). This data suggests that TMCO1, like bacterial YidC, functions collaboratively with the Sec61 channel to promote the biogenesis of multispanning membrane proteins.

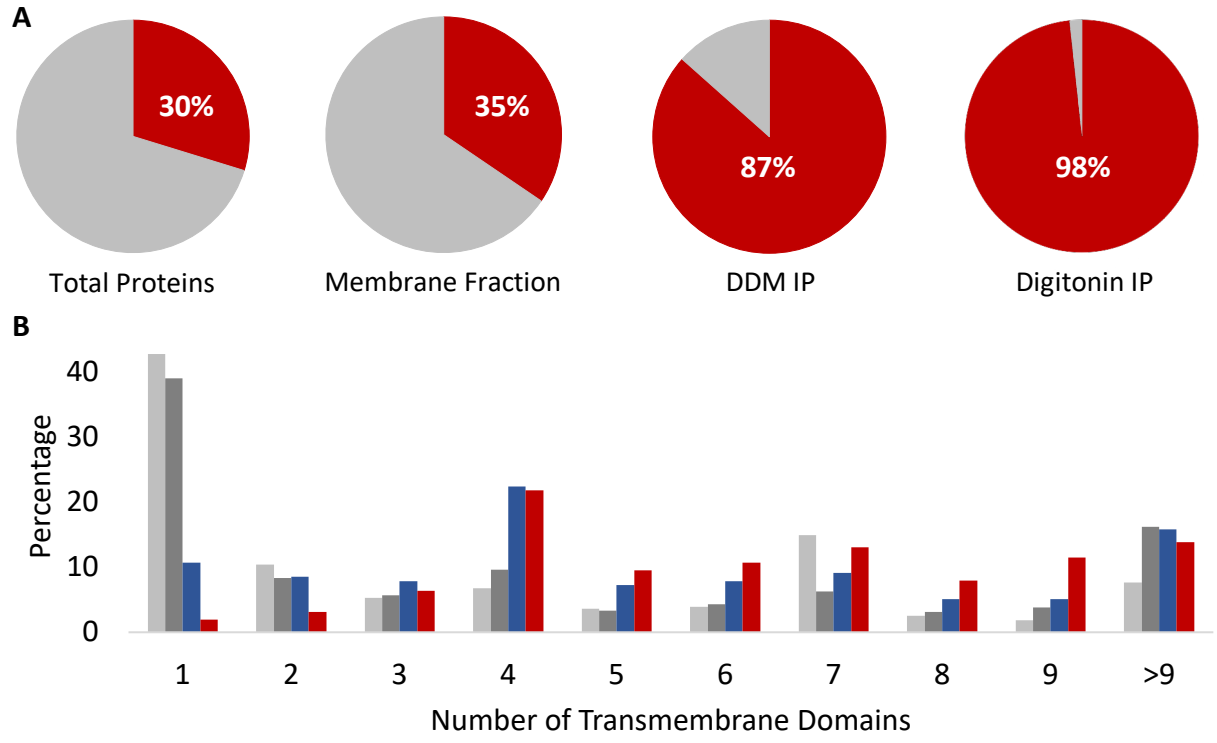


**Figure 21. TMCO1 pulldown selectively enriches for multipass membrane proteins.**

(A) Percentage of membrane proteins (red) in the total human proteome, the Hek293 membrane fraction and the DDM and Digitonin 3xFlag-TMCO1 pulldowns
(B) Division of membrane proteins by the number of transmembrane domains. Shown are distributions for all human membrane proteins (light grey), the Hek293 membrane fraction (dark grey), the DDM 3xFlag-TMCO1 pulldown (blue) and the digitonin 3xFlag-TMCO1 pulldown (red).

## 4.4 TMCO1-ribosome complexes contain CCDC47, NOMO and Nicalin

To begin probing the function of TMCO1, we set out to perform an exhaustive characterization of the protein components of TMCO1-ribosome complexes. To identify additional interaction partners of the TMCO1-ribosome complex, we prepared complexes from 3xFlag-TMCO1 and wild-type cells using either digitonin or DDM for solubilization. After immunoprecipitation, excess Flag peptide (used for elution) was removed by two successive sucrose cushions. These samples were then processed by trypsin digestion, TMT isobaric labeling and LC-MS/MS analysis, yielding a list of enrichment ratios (3xFlag/control) for each identified protein. As expected, we detected a large number of highly enriched ribosomal proteins, core subunits of the Sec61 complex and TMCO1 itself (Figure 22). Of the previously described translocon accessory complexes, the majority of OST complex subunits, as well as TRAM, Sec62, Sec63 and the signal peptidase complex were either absent or present at a higher level in the control pulldown (Figure 22 and Table 1). Some previously identified chaperones, however, looked moderately enriched by Flag pulldown (Table 1). These include the TRAP complex components, Calnexin, and the Ribophorin I subunit of the OST. When DDM was used instead of digitonin, the results were noisier with less clear enrichment of the translocon accessory subunits (Figure 22B), suggesting these subunits interact with the core complex in a detergent-sensitive manner.

Three poorly understood membrane proteins – CCDC47, NOMO and Nicalin – were very highly enriched in the TMCO1 pulldown (Figure 22 and Table 1). Each of these was separately validated by immunoblotting with the respective antibody (Figure 23). CCDC47, also known as Calumin, is a $Ca^{2+}$-binding protein which has been proposed to play a role in ER-associated degradation (ERAD) and $Ca^{2+}$ handling (Konno et al., 2012; Yamamoto et al., 2014; Zhang et

al., 2007). Nicalin and NOMO form a stable complex with a third protein called TMEM147. Together they have also been linked to membrane protein quality control, and to the Nodal signaling pathway, which plays an important role in vertebrate development(Almedom et al., 2009; Dettmer et al., 2010; Haffner et al., 2004; Kamat et al., 2014; Rosemond et al., 2011). We did not identify TMEM147 in our proteomic experiment, presumably because it lacks significant soluble regions (Figure 23) which are preferentially detected by LC-MS/MS. Attempts to identify TMEM147 by immunoblotting failed because of the poor quality of commercially available antibodies. Notably, none of these proteins has previously been reported to function co-translationally, or to interact with ribosomes, the Sec61 complex, or TMCO1. The presence of CCDC47, NOMO and Nicalin in this complex suggests TMCO1 may function to connect membrane protein biogenesis to quality control.

Another protein possibly enriched in the Flag pulldown is a poorly understood protein named C20Orf24. This protein has been associated with colorectal carcinoma progression (Carvalho et al., 2012) but its molecular function is completely unknown. Because the enrichment ratio is relatively low (Table 1), it is difficult to definitively assign it as a complex member. Attempts to verify enrichment by western blotting were unsuccessful, as commercial antibodies did not work.

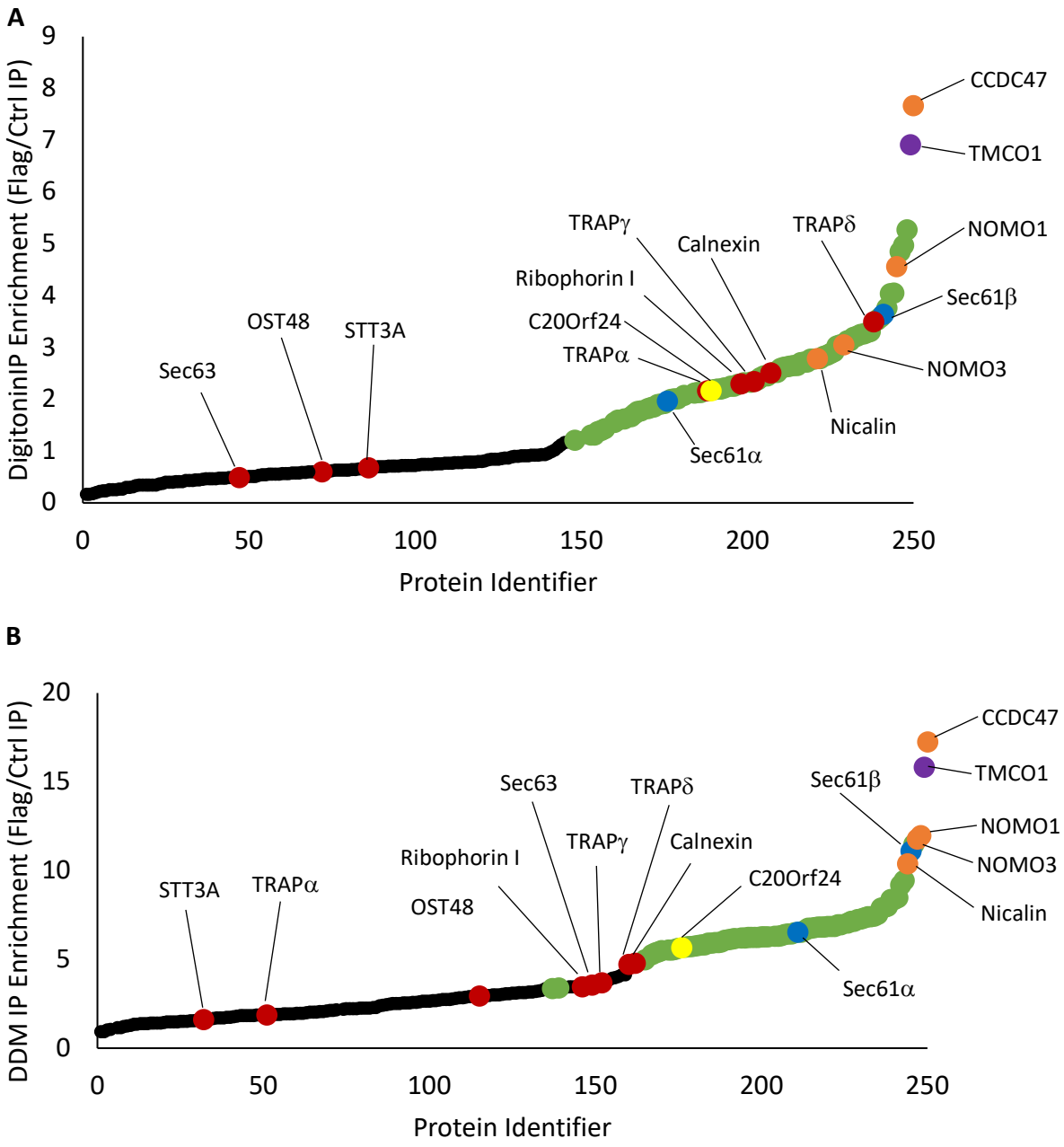


**Figure 22.TMCO1 forms defined ribosome-associated complexes.**

(A) After digitonin solubilization, 3xFlag-TMCO1-ribosome complexes were isolated and analyzed by mass spectrometry. Compared to control pulldowns, 3xFlag-TMCO1 pulldown enriched for ribosomal proteins (green), Sec61 subunits (blue), some translocon accessory subunits (red), some proteins loosely connected to protein quality control (orange) and a previously uncharacterized protein named C20Orf24 (yellow).
(B) Like (A) but using DDM instead of Digitonin during solubilization and pulldown.

| Subunit Role | Protein | Digitonin IP Enrichment (Flag/Ctrl IP) | DDM IP Enrichment (Flag/Ctrl IP) |
|---|---|---|---|
| TMCO1 | TMCO1 | 6.92 | 15.80 |
| Quality Control Factors | CCDC47 | 7.67 | 17.23 |
| | NOMO1 | 4.56 | 11.96 |
| | NOMO3 | 3.05 | 11.77 |
| | Nicalin | 2.78 | 10.38 |
| Sec61 Complex | Sec61β | 3.63 | 11.09 |
| | Sec61α | 1.96 | 6.52 |
| Chaperone | Calnexin | 2.51 | 4.77 |
| TRAP Complex | TRAPγ | 2.34 | 3.68 |
| | TRAPδ | 3.49 | 4.72 |
| | TRAPα | 2.16 | 1.86 |
| OST Complex | Ribophorin I | 2.3 | 3.46 |
| | STT3A | 0.68 | 1.59 |
| | Ost48 | 0.6 | 2.92 |
| Other | Sec63 | 0.49 | 3.56 |
| Unknown | C20Orf24 | 2.17 | 5.66 |

**Table 1. Enrichment of protein biogenesis and quality control factors by TMCO1 IP.**

3xFlag-TMCO1 and control pulldowns were analyzed by TMT isobaric labeling followed by mass spectrometry. As judged by ratios of ribosomal proteins, values to be considered significantly enriched are over 1 for the digitonin approach and over 5 for the DDM approach.

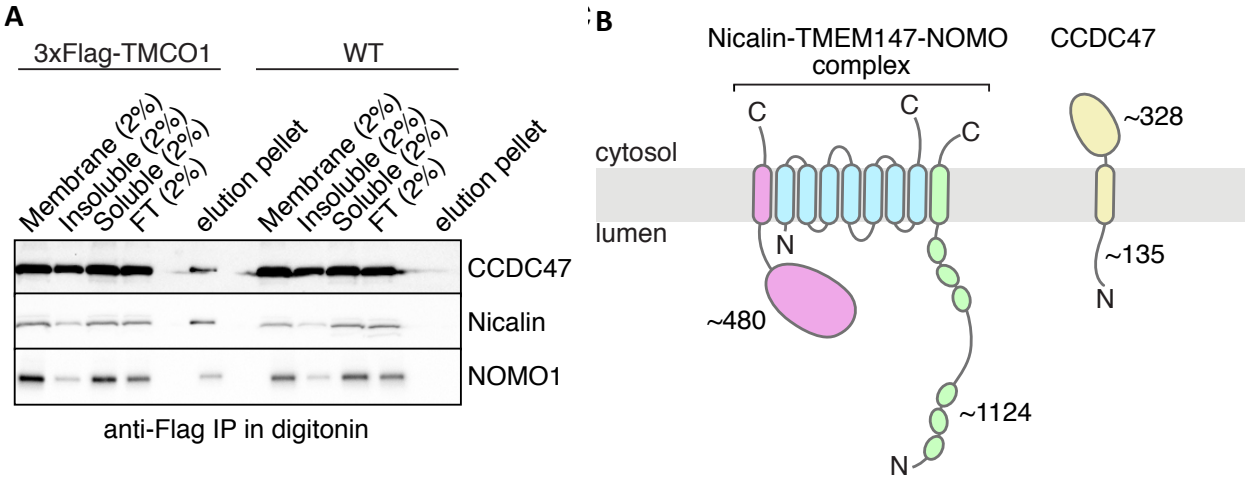


**Figure 23. Identification of a new ribosome-attached quality control complex.**

(A) Western blotting-based validation of quality control components identified by mass spectrometry. TMEM147 is also predicted to be present in this sample, based on its known association with Nicalin and NOMO, but none of the commercially available antibodies were of sufficient quality to detect it. This is likely due to the absence of useful epitopes in TMEM147, given its very short extra-membrane loops.
(B) Schematic diagram of quality control hits and their approximate domain structures.

## 4.5 Discussion

These observations support the hypothesis that TMCO1 functions as a membrane protein biogenesis factor, validating a key prediction of the Oxa1 superfamily hypothesis.

The identification of TMCO1 client proteins as multipass membrane proteins, together with the presence of TMCO1 in complex with both the Sec61 channel and ribosomes naturally prompts comparisons to the SecYEG-dependent function of bacterial YidC. Future studies are needed to establish the exact molecular mechanism of TMCO1. The YidC similiarity suggests TMCO1 may facilitate substrate folding, insertion of specific transmembrane domains, or both. From an evolutionary perspective, the retaining of the TMCO1-Sec61 interaction in eukaryotes serves to further support the idea that the Sec-dependent protein biogenesis is an ancestral function of Oxa1 superfamily proteins.

The interaction profile of TMCO1 in the ribosome-bound complex suggests that it may work to bridge the biogenesis of membrane proteins with quality control. The exact sequence of events is difficult to ascertain because the mass spectrometry enrichment results may represent an average of different subcomplexes that contain different components. Nevertheless, the presence of TRAP, Calnexin, and Ribophorin I on at least some particles is consistent with a role for the complex in the folding and/or insertion of nascent membrane proteins. Interestingly, we did not observe a wholesale enrichment of the OST complex, suggesting our procedure enriches for complexes containing TRAP but not OST. Such complexes have recently been observed by cryo-electron tomography of native ER membranes (Braunger et al., 2018). Structural analysis of the mammalian OST complex (Braunger et al., 2018) indicates that Ribophorin I and the catalytic STT3A subunit both contain large soluble domains that are expected to be identified by mass spectrometry. We did identify both, but only the former was enriched in the Flag pulldown (ratio>2); the latter was depleted (ratio 0.5). Ribophorin I has been suggested previously to function in membrane protein biogenesis independent of its role in glycosylation (Wilson et al., 2005). In this previous study, the association of Ribophorin I with a nascent transmembrane domain was independent of substrate glycosylation and specific to the identity of the TMD, as crosslinks could be observed to the first but not the fifth TMD of rhodopsin (Wilson et al., 2005). Whether other OST subunits are present in our TMCO1 affinity purified complexes remains to be determined. Within the OST complex, Ribophorin I and another subunit named TMEM258 form subcomplex I (Braunger et al., 2018; Kelleher and Gilmore, 2006; Wild et al., 2018). TMEM258 was not identified in our mass spectrometry approach (possibly because it has no soluble domains), so whether it is present in these complexes remains uncertain.

The identified quality control factors present an interesting possibility that TMCO1 might connect membrane protein biogenesis to quality control. CCDC47 is the single most enriched protein in our mass spectrometry dataset. This protein was originally identified by its ability to bind $Ca^{2+}$ through an acidic lumenal domain (Zhang et al., 2007) although the physiological relevance of this binding is unclear. A later study suggested a completely different role for CCDC47 in Endoplasmic Reticulum Associated Degradation (ERAD) based on its interaction with known ERAD components and modulation of degradation of known ERAD substrates (Yamamoto et al., 2014). The role of CCDC47 in $Ca^{2+}$ signaling may therefore be indirect. CCDC47 was shown to bind $Ca^{2+}$ release-activated $Ca^{2+}$ channels and affect their permeability (Konno et al., 2012), possibly in connection with its quality control function. Regardless, the involvement of CCDC47 in $Ca^{2+}$ flow may explain the similar phenotypes seen in TMCO1 knockout experiments (Wang et al., 2016).

The other identified proteins, NOMO and Nicalin, together with their binding partner TMEM147, have a better-defined role in folding and quality control, where they seem to act on transmembrane proteins. NOMO binds the heteropentameric nicotinic acetylcholine receptor in *C. elegans* (Gottschalk et al., 2005) and together with Nicalin controls its cell surface localization and subunit stoichiometry (Almedom et al., 2009). This may be a general function of this complex, as other studies showed a similar effect on the regulation of cell surface localization of other ion channels (Kamat et al., 2014; Rosemond et al., 2011). The exact role of this complex, however, remains poorly understood. The localization at the ribosome (described in this work) suggests it might capture membrane proteins as they are synthesized and coordinate proper folding and/or complex assembly in order to prevent immature or improperly assembled complexes from exiting the ER.

Finally, it is worth noting that a full mechanistic description of TMCO1 function depends on demonstrating an effect of TMCO1 on the insertion, folding or quality control of a particular substrate. The identification of hundreds of client proteins by mRNA sequencing makes this goal within reach, and the exhaustive component characterization of TMCO1-ribosome complexes opens the door towards step-by-step mechanistic analysis of the pathway.

# 5 Future Directions

This work provides a conceptual framework for understanding the evolution of certain membrane protein biogenesis factors and presents clear ways to frame future research questions. Conceptually, the future directions of research facilitated by this work fall into two broad categories: facilitating the discovery of new components in existing pathways and mechanistic dissection of known components.

In certain pathways likely to involve Oxa1 superfamily members, not all factors have yet been identified. This is perhaps most notable in the cases of Get3 homologs without known membrane receptors. Plants contain a Get3 homolog (known as Get3b) that localizes to the chloroplast and functions in a different pathway from cytosolic Get3 (Xing et al., 2017). This raises the possibility that a separate pathway targets Tail Anchored or similar proteins to the thylakoid membrane. Based on the distant homology of YidC and Get1 family members, the prime candidates would be one of the two superfamily members Alb3 or Alb4. Remarkably, some plants contain a third homolog Get3 homolog, known as Get3c, that localizes to the mitochondrial matrix (Xing et al., 2017), raising the possibility of yet another targeting pathway. The role of Get3c remains unstudied, but targeting a subset of mitochondrial inner membrane proteins and delivering them to either of two Oxa1 mitochondrial paralogs is an attractive possibility (Benz et al., 2013).

Similarly, most archaea contain a Get3 homolog (usually referred to as TRC40). Archaeal Get3 homologs have been shown to be bind TA substrates and can substitute yeast Get3 in an artificial insertion assay with yeast membranes (Sherrill et al., 2011; Suloway et al., 2012). Despite numerous attempts by the Keenan lab to demonstrate binding of Ylp1 to the archaeal Get3 homolog, no interaction was ever observed. One hypothesis based on the difference

between bacterial YidC and archaeal Ylp1 that shows two extra transmembrane helices in YidC, is that Ylp1 requires an as-yet unidentified binding partner as an adaptor for a post-translational function. Such a mechanism would be analogous to the requirement for Get2 in the yeast GET pathway. However, without further experimental data such an explanation remains speculative. The archaeal Get3 homolog might function solely as a soluble chaperone without a need for a membrane-bound insertase.

Based on the common ancestry of EMC3 with YidC, TMCO1 and Get1, this work also allows for specific mechanistic predictions of how the ER Membrane Complex (EMC) might function. The EMC as a whole has very recently shown to work as an insertase for TA proteins with a TMD that is not hydrophobic enough to engage the GET targeting machinery (Guna et al., 2017). However, the EMC is a ten-subunit complex in which subunits depend on one another for expression (Guna et al., 2017), which would likely hamper future efforts at mechanistic characterization of the individual subunits. We predict that EMC3 is one of the catalytic subunits, and that it possesses the ability to stabilize a transmembrane domain during the insertion process. Designing specific mutations in the membrane-enclosed hydrophilic binding pocket of EMC3 may allow for both activity-modulating variants and substrate crosslinks to be identified.

Another promising research avenue opened by this work is the discovery of the TMCO1-mediated membrane protein biogenesis pathway. The immediate next goal of this project will be to demonstrate a specific phenotype for TMCO1 null cells in the biogenesis of at least one substrate. This can be shown through a combination *in vitro* and *in vivo* analyses using previously published methods. Cell surface localization of the identified TMCO1 clients can be assayed directly using antibodies or previously published protocols utilizing chemical

biotinylation followed by mass spectrometry (Belleannee Clémence et al., 2011; Shin et al., 2003). It is possible, however, that very few proteins depend on TMCO1 for insertion and TMCO1 instead works to accelerate insertion and/or folding. Such a role is supported by the observation that TMCO1 is recruited relatively non-specifically to ribosomes translating any one of a very large number of multipass membrane proteins. This type of activity would not be easily detected by cell surface membrane protein analysis, but would be well suited for *in vitro* analysis. In this long-established assay, translation is performed *in vitro* in the presence of purified cell membranes. By comparing the insertion of a substrate protein into wild type and TMCO1 KO membranes, activity can be investigated along a precise time course and under various external conditions which cannot be easily varied in a live cell.

The identification of hundreds of multi-spanning membrane proteins as cotranslational clients of TMCO1-containing ribosome-Sec61 complexes also raises a series of mechanistic questions about TMCO1 functions in the cell. For example, what client features are recognized by the TMCO1 machinery? How do these clients engage the Sec61 complex, TMCO1 and other components of the complex at different stages of their biogenesis? The protocols we establish in this work will make tackling these questions possible. For example, programming an *in vitro* translation reaction with a series of truncated, stop codon-less client mRNAs of various lenghts will generate ribosome-nascent chain complexes stalled at defined stages of translation. When translation is carried out in the presence of membranes derived from 3xFlag-TMCO1 cells, our co-IP/western blot strategy can be utilized to define when TMCO1 (and the other components we have identified) is recruited to the ribosome-Sec61 complex. Such an approach has been used successfully for generic Sec61 substrates (Conti et al., 2015). Incorporating site-specific crosslinkers into the nascent substrate will then allow for mapping of the substrate interactions

with the different biogenesis factors. These types of experiments will define the specific features that define TMCO1 clients, which can then be tested by mutational analysis in the same assays. As a complementary goal, the affinity purification of TMCO1 complexes will allow for the reconstitution of TMCO1-mediated biogenesis with purified components. This will allow for definitive determination of the role of TMCO1 complexes in multi-spanning membrane protein biogenesis. Immunoaffinity purification followed by co-reconstitution of all different components into proteoliposomes was recently demonstrated for the EMC (Guna et al., 2017). By purifying TMCO1 complexes from cells that are selectively depleted for various components (e.g. Nicalin, TMEM147, NOMO, CCDC47 etc.) these experiments will allow for the identification of the minimal machinery required for biogenesis of TMCO1 substrates.

Finally, little is known more generally about how different accessory factors coordinate with the core Sec61 complex to facilitate membrane protein insertion. Our native purification of TMCO1-containing ribosome-Sec61 complexes from cells provides an excellent starting point for structural studies by single particle cryo-EM. Programming ribosomes with stalled substrates of various lengths would then also allow for the structural analysis of the Sec61-ribosome complexes that recruit the various components we have identified, an approach that has been used recently for the determination of the Sec61/ribosome/OST structure (Braunger et al., 2018). High resolution structural analysis of these complexes will provide valuable mechanistic information on the highly conserved TMCO1 membrane protein biogenesis pathway and reveal how factors not known to act co-translationally interact with the translating ribosome. Additionally, a high resolution structure of the TMCO1-ribosome complex will provide a case study of the substrate-specific assembly of a specific holotranslocon complex in human cells.

# 6 Materials and Methods

## 6.1 Protocols related to Borowska et al. *Structure* (2015)

### 6.1.1 Phylogenetic Analysis

1. Archaeal DUF106 sequences were obtained using *Mj*0480 as a query in NCBI Blast with an E-value cutoff of $10^{-30}$
2. Sequences were aligned with representative bacterial YidC, eukaryotic Oxa1 and eukaryotic Alb3 sequences using T-Coffee (Notredame et al., 2000) with the default parameters.
    a. Note: Multiple MSA programs can be used. In our experience, T-Coffee performs best (ie: results in shortest branch lengths and highest level of statistical support) for YidC and related proteins. For larger alignments that exceed the capacity of T-Coffee, we have used MUSCLE with good results.
3. Gaps were trimmed automatically using the Software TrimAl (Capella-Gutierrez et al., 2009) with a cutoff of 0.55.
    a. Note: Gap trimming cutoff is determined empirically. We have mostly kept this value relatively low (0.3 - 0.6) as the alignments we present here include highly divergent sequences
4. The resulting alignment was tested in Prottest (Darriba et al., 2011) to find the best evolutionary model, which for this data was: LG substitution matrix, empirical amino acid frequencies, fixed gamma shape parameter (1.491) with 4 substitution rate categories.
5. The phylogenetic tree was then assembled in PhyML (Guindon et al., 2010) using determined parameters

a. Note: It is now possible to perform steps 4 and 5 at once using the new PhyML Smart Model Selection feature.

### 6.1.2 Bacterial YidC Complementation

1. The *Mj*0480 gene and two closely related mesophilic archaeal homologs (from *Methanococcus aeolicus* and *Methanococcus maripaludis*) were cloned into pQE81-L vector encoding a C-terminal 6xHis tag using Gibson assembly (Gibson et al., 2009); we also cloned chimeras in which the gene of interest was fused to the first 330 residues of *E. coli* YidC. Full-length YidC was used as a positive control and protein expression was verified by western blot using an anti-6xHis antibody.
    a. Chimeras were included because *E. coli* YidC has an extra TMD and a large periplasmic domain. The 1-330 YidC fusion gene conserves these features
2. Transform *E. coli* JS7131 YidC depletion strain (gift from R. Dalbey) (Samuelson et al., 2000) with the appropriate experimental construct, incubate overnight at 37 degrees.
3. Pick single colonies and grow for 4-6 hours in LB media supplemented with 0.2% arabinose, 25 μg/mL spectinomycin and 100 μg/mL ampicillin, taking care that cells never entered stationary phase.
4. Pellet 0.2 $OD_{600}$ of cells for 5 min at 5000 x g at 4 °C
5. Wash with 1 mL LB, resuspend to a final $OD_{600}$ of 0.1
6. Spot 2 μL of 1:10 serial dilutions on LB plates supplemented with spectinomycin, ampicillin, and either 0.2% arabinose (positive control) or 0.2% glucose (negative control).
    a. The latter condition was necessary as the JS7131 strain sometimes spontaneously reverts to a wild-type phenotype.
    b. For the experimental condition, we tested IPTG concentrations between 20 and 500 μM.

### 6.1.3 Yeast Oxa1 Complementation

1. To test Oxa1 complementation, we generated a yeast W303-1A *Δoxa1* knockout strain by homologous recombination. This strain was verified by PCR and exhibited an inability to grow on nonfermentable carbon source as previously reported (Preuss et al., 2005).
2. Complementation was tested by cloning either the *Mj*0480 gene or its *M. aeolicus* homolog into a Yeplac195 vector (gift from B. Glick) with a C-terminal 6xHis tag. Each gene was fused to the Oxa1 mitochondrial targeting sequence and optionally also the long Oxa1 C-terminal tail, which is important for its function (Preuss et al., 2005). A vector carrying a C-terminally 6xHis-tagged Oxa1 was used as a positive control and empty vector was used as a negative control.
3. Transform the knockout strain with the appropriate construct using the Lithium Acetate method (Gietz and and Woods, 2002), incubate at 30 degrees until colonies are observed (2-3 days)
4. Restreak a single colony on fresh plates, incubate 1 day at 30 degrees
5. Inoculate CSM-Ura cultures, grow at 30 degrees until they became cloudy ($OD_{600} < 2$)
6. Centrifuge 0.5 $OD_{600}$ of cells for 4 min at 2,600 x g at 4 °C and resuspend in 1 mL cold water.
7. Spot 2 μL of 5 x 1:10 serial dilutions on YP agar plates supplemented with either 2% glucose or 3% glycerol
    a. We incubated plates at room temperature, 30 °C and 37 °C, with similar results.
    b. Only strains transformed with the Oxa1 gene grew on 3% glycerol plates.

### 6.1.4 Ribosome binding assays

1. We purified 70S *E. coli* ribosomes according to a published protocol (Wu et al., 2013) with the exception that we used *E. coli* BL21(DE3) cells.
2. Stalled ribosome-nascent chain complexes (RNCs) were obtained using a construct containing a 3x Strep-tag II sequence fused to the first transmembrane segment of the YidC substrate $F_0c$ and to the SecM stalling sequence, similar to a previously utilized construct (Kedrov et al., 2013). An artificially synthesized gene was Gibson assembled into pET21a, and transformed into BL21(DE3) cells. Stalled RNC-$F_0c$ were expressed and purified according to the published protocol (Wu et al., 2013) (Figure S4).
    a. Note: Both the total 70S ribosomes and the stalled RNCs were quantified by $A_{260}$, flash frozen and stored at -80 °C. They seem to be stable for at least a month.
3. Incubate 100 nM ribosomes or RNCs with 1 μM *Mj*0480 in 1.5 mL of assay buffer containing 50 mM Tris pH 7.5, 150 mM KCl, 5 mM $MgCl_2$ and 0.04% DDM for 1 hour on ice.
4. Pellet for 3 h at 100,000 x g at 4°C in a TLA100.3 rotor. Save an aliquot of the supernatant for analysis.
    a. Note: We did see some pelleting of Mj0480 in the absence of ribosomes, likely due to the fact that we did not include a sucrose cushion in this step. A modified pelleting protocol was used in later studies with mammalian ribosomes where a sucrose cushion was included, and that did seem to make a difference.
5. Wash pellets with 300 μL assay buffer, resuspend in 50 μL SDS sample buffer and then adjusted to 1x, 5x and 10x volume equivalents of the supernatant sample.

6. Analyze supernatant and pellet samples by SDS-PAGE and western blotting with a Penta-His HRP conjugated antibody (Qiagen catalog #34460).

### 6.1.5 Plasmids

| Purpose | Identifier | Plasmid schematic |
|---|---|---|
| Ribosome-RNC Purification | bSA013 | pET21c-2StrepTag-AtpE-SecM |
| YidC complementation | bSA004 | pQE81L-MjYlp1-6xHis |
| | bSA005 | pQE81L-YidC(1-330)-Mj0480-6xHis |
| | bSA009 | pQE81L-MaYlp1-6xHis |
| | bSA010 | pQE81L-MmYlp1-6xHis |
| | bSA016 | pQE81L-YidC(1-330)-MtYlp1-6xHis |
| | bSA017 | pQE81L-YidC(1-330)-MmYlp1-6xHis |
| | bSA018 | pQE81L-YidC(1-330)-MaYlp1-6xHis |
| | bSA039 | pQE81L-YidC-6xHis |
| Oxa1 Complementation | bSA056 | Yeplac195-Oxa1-6xHis |
| | bSA057 | Yeplac195-Oxa1ts-MjYlp1-6xHis |
| | bSA058 | Yeplac195-Oxa1ts-MjYlp1-Oxa1tail-6xHis |
| | bSA059 | Yeplac195-Oxa1ts-MaYlp1-Oxa1tail-6xHis |
| | bSA060 | Yeplac195-Oxa1ts-MaYlp1-6xHis |
| | bSA061 | Yeplac195-Oxa1ts-YidC(1-330)-Oxa1tail-6xHis |

**Table 2. Plasmids used in the subset of this work described in Borowska et al. 2015**

## 6.2 Protocols related to Anghel et al. *Cell Reports* (2017)

### 6.2.1 Phylogenetic analysis

1. Distant homologs of Ylp1 were identified using the *M. jannaschii* Ylp1 sequence (Mj0480) using HHpred (Soding et al., 2005)
2. Sequences were obtained using *Mj*0480 as a query in NCBI Blast with an E-value cutoff of $10^{-30}$
    a. Note: We made an effort to include organisms as diverse as possible phylogenetically, as this resulted in shorter branch lengths.
3. Proteins in the master list were then aligned using MUSCLE (Edgar, 2004) with standard parameters
4. Gaps in the alignment were trimmed using TrimAl (Capella-Gutierrez et al., 2009) with a cutoff of 0.4
5. The phylogenetic tree was then assembled in PhyML-SMS (Guindon et al., 2010) using nearest-neighbor interchange (NNI) and the Akaike Information Criterion

### 6.2.2 Generation of the TMCO1 antibody

Note: These services were performed by Lampire Biologicals

1. A peptide representing a region of the TMCO1 cytosolic coiled coil (EKKKETITESAGRQQKK) was chemically synthesized and conjugated to KLH.
2. Two rabbits were immunized and antibody was generation through a 50 day protocol, with final exsanguination bleed at day 57.
3. Exsanguination bleed was supplemented with 0.02% sodium azide, flash-frozen in liquid nitrogen and stored at -80°C
4. A minority of experiments used antibody that was purified against the EKKKETITESAGRQQKK peptide using the SulfoLink Kit (Fisher Scientific).
    a. The affinity purified antibody can be used at a 1:10,000 dilution in a standard western blot (as opposed to a 1:1000 dilution for the raw antiserum) and provides much better signal and specificity.

### 6.2.3 Hek293 Transfection

1. Grow cells until ~90% confluent.
2. For one 10 cm dish, mix 10 µg of DNA with 20 µL of Trans-It 293 reagent (Mirus Bio LLC) in serum-free DMEM medium.
3. Incubate complexes 20 minutes at room temperature. While complexes are incubated, add fresh DMEM supplemented with 10% FBS (no antibiotics) to cells.
4. Add complexes drip-wise to cells, incubate 24 hours in a cell culture incubator
5. Subcultivate cells 1:2 and incubate a further 24 hours in a cell culture incubator
6. Harvest cells (48 hours after original transfection)

### 6.2.4 Generation of a TMCO1 KO cell line by CRISPR/Cas9 Genome Editing

Note: This cell line was generated with the technical services of the Genome Engineering Core Facility at the University of Chicago.

1. Hek293-Cas9 cells were generated by the Hegde lab by integrating a 3xFlag-Cas9 gene into the Flp-In site of Hek293 Flp-In T-REx cells (Invitrogen).
    a. For genome editing, we supplement the media with 15 µg/mL Blasticidin and 100 µg/mL Hygromycin B although we do not use these antibiotics otherwise.
2. Cas9 expression was induced by addition doxycycline at 10 ng/mL, followed by transfection of a plasmid expressing the guide RNA 5'-GAAACAATAACAGAGTCAGCTGG-3'
3. Single cells were seeded into 96 well plates and allowed to grow clonally
4. The final TMCO1 knockout line was verified by both genomic DNA sequencing and immunoblotting with an α-TMCO1 antibody
    a. Three clones were isolated, labeled D3, D5 and G11.
        i. D5 is a true knockout as it carries a frameshift mutation on both choromosomes. This clone is our designated TMCO1 KO clone.
        ii. G11 is heterozygotic (it carries a frameshift mutation on only one chromosome).
        iii. D3 resembles a true knockout by immunoblotting. but in fact one chromosome carries a 12 residue deletion in the epitope for our α-TMCO1 but the protein is predicted to be made.

### 6.2.5 Generation of a 3xFlag-TMCO1 cell line by CRISPR/Cas9 Genome Editing

Note: This cell line was generated with the technical services of the Genome Engineering Core Facility at the University of Chicago.

1. N-terminally Flag tagged TMCO1 was also generated at the same facility using a previously described two step strategy with transfections and selection identical to the TMCO1 KO generation procedure.
2. The 3xFlag-TMCO1 lines were verified by both genomic DNA sequencing and immunoblotting with α-TMCO1 and α-Flag antibodies.
    a. Multiple clones were obtained, four of which (P1A6, P1F1, P1F3, P2C4) have one nonfunctional TMCO1 allele and one allele containing a 3xFlag-tagged TMCO1 with a 13 residue linker (ITSYNVCYTKLSG, from the Cre-lox recombination) before the TMCO1 ORF. We use P2C4 as our primary tagged cell line.

### 6.2.6 Preparation of the total membrane fraction from Hek293 cells

1. Harvest cells at a density of 70-100% while growing
    a. Do not supplement media of Hek293 TRex cells with Blasticidin and Hygromycin B when growing cells if membranes will be used for translation. Supplementing media with these antibiotics makes resulting membranes incompetent for translation.
    b. The yield of this procedure is about 10 µL membrane pellet from one 15 cm cell culture dish at ~80% confluence.
    c. Keep cells at room temperature until ice cold DPBS is added. Afterwards, perform procedure on ice.
    d. After preparation, membranes can be frozen and stored at -80°C for at least several months without negative consequences. Do not flash freeze membranes repeatedly.
2. Aspirate media, and add 10 mL DPBS to one 15 cm dish. Scrape cells and collect into 50 mL conical tubes.
    a. If cells will be used for purification of ribosome complexes, I add 50 µg/mL emetine to DPBS and throughout the procedure. Do not add emetine if membranes will be used for translation
3. Pellet cells 5 minutes at 500 x g centrifugation at 4°C
4. Resuspend in HM Buffer (10 mM Hepes pH 7.5, 10 mM potassium chloride, 1 mM magnesium chloride) equal to 3.5x the volume of the cell pellet.
5. Incubate on ice for 15 minutes

    a. If desired, analyze swelling by dilution 5 µL of cell suspension with 5 µL and observing under an inverted microscope.

6. Dounce for 15 strokes (up-down is considered one stroke) in a douncer with a tight-fitting "B" pestle (Kontess)
    a. If desired, analyze lysis by dilution 5 µL of cell suspension with 5 µL and observing under an inverted microscope. Less than ~20% cells should remain unlysed, and no significant clumping of nuclei should be present.
7. Add sucrose to 250 mM to balance osmolarity.
8. Pellet 3 minutes at 700 x g to remove nuclei and unlysed cells. Save supernatant.
9. Pellet 10 minutes at 10,000 x g to pellet membrane fraction
    a. Contrary to previous studies, in our hands this was sufficient to pellet most biological membranes of interest, including the endoplasmic reticulum, Golgi, plasma membrane and mitochondria.
10. Wash membranes with assay buffer (150 mM potassium acetate, 50 mM Hepes pH 7.4, 5 mM magnesium acetate) and centrifuge again 10 minutes at 10,000 x g to remove any residual cytosolic proteins.
11. If using membranes for sucrose cushions, gradients, pull-downs or translation, digest polysomes by adding Calcium Acetate to 1 mM and 100 Units of Microccocal Nuclease (NEB)
12. Incubate 10 minutes at 25°C, and then quenche by addition of EGTA to 2 mM.
    a. If membranes will be used for translation, nuclease treatment is performed for 30 minutes at 37°C as in this case partial digestion of ribosomes is not problematic.
13. Wash membranes one more time with assay buffer to remove nuclease.

14. If freezing, supplement membrane suspension with sucrose to 250 mM and flash freeze. Store membrane suspensions at -80°C

### 6.2.7 Topology analysis by glycosylation mapping in Hek293 cells

1. Design N-terminally 3xFlag-tagged construct into the pGFP plasmid (Clontech). Add an opsin N-glycosylation tag (MNGTEGPNFYVPFSNKTVD) at the positions to be tested.
    a. For TMCO1, we have used a bacterially codon-optimized DNA sequence. EMC3 plasmids were identical, but contained a cDNA-derived EMC3 sequence.
    b. The resulting constructs encode a 3xFlag-tagged protein under the control of a CMV promoter and an SV40 polyA signal.
2. Transfect 10 µg of DNA with 20 µL of Trans-It 293 reagent (Mirus Bio LLC) as usual
3. Subcultivate cells 1:2 the next day and harvest 48 hrs after transfection as usual
4. Prepare membrane fraction as usual.
5. Resuspend membranes in 100 µL of 1% SDS with 100 mM DTT. Incubate for 5 minutes at 95 °C, and then cool at room temperature.
6. Adjust buffer to 50 mM Hepes pH 7.4, 150 mM NaCl, 1% NP40, 0.1% SDS and supplement with 50 Units of Benzonase (Sigma, E1014). Benzonase is required for reducing viscosity.
    a. Final volume does not seem relevant. 800 µL has worked well in the past.
7. Split reactions in half. To half, add 20 Units of PNGase F (Promega).
8. Incubated for 4 hours at 37 °C with 200 rpm shaking to mix well.
9. TCA precipitate.
10. Resuspend in Lamelli Sample Buffer and analyze by SDS-PAGE.

### 6.2.8 Topology analysis by glycosylation mapping in yeast cells

11. Design a 3xFlag-tagged construct with the endogenous promoter in the Yeplac195 vector. Add an opsin N-glycosylation tag (MNGTEGPNFYVPFSNKTVD) at the positions to be tested.
    a. For Get1, we used a C-terminal tag as that is better tolerated for this protein
2. Transform plasmid into BY4741 yeast using lithium acetate/PEG (Gietz and and Woods, 2002). Incubate at 30°C for 2-3 days.
3. Pick a single colony from transformation plates and restreak onto a fresh plate. Incubate at 30°C for an extra day
4. Inoculate a 4 mL SD –URA +2% glucose culture with about half of the amount of yeast on the restreaked plate. Incubate for 1 hour at room temperature with 225 rpm shaking
5. Measure. $A_{600}$. Collect 4 $A_{600}$ units and mix with sodium azide to a final concentration of 0.01%. All manipulations should be done on ice from here on.
6. Pellet cells 3 minutes at 16,000 x g and resuspend in 350 mM freshly diluted NaOH supplemented with 1 mM PMSF and 1x cOmplete, Mini, EDTA-free Protease Inhibitor Cocktail tablets (Roche)
7. Incubate 5 minutes on ice
8. Resuspend pellet in 120 µL of 1% SDS, 100 mM DTT, 50 mM Tris pH 6.8. Incubate 5 minutes at 95 °C and then cool to room temperature
9. Pellet 3 minutes at 16,000 x g to remove insoluble material. Save supernatant and discard pellet.

10. Adjust buffer composition of supernatant to 5 mM Tris pH 6.8, 50 mM Hepes pH 7.4, 150 mM NaCl, 1% NP40, 0.1% SDS and supplement with 25 Units of Benzonase (Sigma, E1014).
11. Split reactions in half. To half, add 20 Units of PNGase F (Promega).
12. Incubated for 4 hours at 37 °C with 200 rpm shaking to mix well.
13. TCA precipitate.
14. Resuspend in Lamelli Sample Buffer and analyze by SDS-PAGE.

### 6.2.9 Co-fractionation of native TMCO1-ribosome complexes from Hek293 cells

1. Prepare membrane fraction from Hek293 cells.
    a. This entire procedure should be done on ice.
    b. Usually I add emetine to 50 μg/mL to all buffers to inhibit ribosome dissociation, although this does not seem to make a difference in yield of ribosome/TMCO1 complexes.
2. Resuspend membranes in assay buffer (150 mM potassium acetate, 50 mM Hepes pH 7.4, 5 mM magnesium acetate). Add recrystallized digitonin from a 5% stock to a final concentration of 2%.
    a. Digitonin displays large lot-to-lot variability. For the work in *Anghel et al.* (Cell Reports 2017) we used Calbiochem Lot #2913883 that had been recrystallized in-house
    b. Even after recrystallization, digitonin is only moderately soluble in water. Do not attempt to make stock solutions at higher concentrations than 5%. All stocks should be used within ~4 hours of preparation, or digitonin impurities will start to precipitate.
    c. Resolubilization volume appears relatively insensitive to dilution. I usually use a resolubilization volume of 3 times the membrane pellet volume to conserve digitonin but much higher volumes work well and may increase yield slightly.
3. Incubate for 30 minutes at 4°C with gentle end over end mixing (wheel in the Keenan Lab cold room set at 20-30 rpm works well)
4. Pellet 10 minutes at 10,000 x g to remove insoluble material.

5. Layer supernatant on a 1 mL sucrose cushion 150 mM KCl, 50 mM Tris pH 7.4, 5 mM $MgCl_2$, 1 M Sucrose, 0.1% digitonin)
6. Pellet for 2 hours at 250,000 x g in a TLA100.3 rotor
7. Resuspend pellet in 400 μL assay buffer supplemented with 0.1% digitonin and spin again over a cushion as before
    a. Ribosome pellets from cushions are extremely difficult to resuspend. Gently pipetting up and down for ~15 minutes (until no more particles are visible) is the best way that does not damage ribosome complexes.
8. Resuspend pellet in 400 μL assay buffer supplemented with 0.1% digitonin and layer over a sucrose gradient (10-50% sucrose, 150 mM potassium acetate, 50 mM Tris pH 7.5, 5 mM $MgCl_2$, 0.1% digitonin) at 130,000 x g (SW28.1, Beckman-Coulter) for 12 hrs at 4°C
    a. Some digitonin may precipitate overnight at the bottom of the gradient. This does not appear to negatively affect outcomes
    b. A 12 hour centrifugation time results in 80S ribosomes migrating about 2/3 of the way through the gradient. A centrifugation time as short as 4 hours may sufficient depending on the experiment.
9. Collect fractions from the top of the gradient.
    a. 900 uL fraction volumes work well as they result in 20 fractions total.
10. TCA precipitate.
11. Resuspend in Lamelli Sample Buffer and analyze by SDS-PAGE.

### 6.2.10 Co-immunoprecipitation analysis from Canine Pancreatic Membranes

1. Resuspend 50 μL membranes in a total of 300 μL of 250 mM potassium acetate, 50 mM Hepes pH 7.4, 5 mM magnesium acetate, 15% glycerol and 3% recrystallized digitonin
    a. Perform the entire procedure on ice, except final elution
    b. Digitonin displays large lot-to-lot variability. For the work in *Anghel et al.* (Cell Reports 2017) we used Calbiochem Lot #2913883 that had been recrystallized in-house
    c. Even after recrystallization, digitonin is only moderately soluble in water. Do not attempt to make stock solutions at higher concentrations than 5%. All stocks should be used within ~4 hours of preparation, or digitonin impurities will start to precipitate.
2. Incubate for 30 minutes on ice
3. Pellet 10 minutes at 10,000 x g to remove insoluble material.
4. Divided soluble material in three and layer on top of Protein A resin that had been crosslinked to antibodies against TMCO1, Sec61β or 3F4 (as control)
5. IP reactions were incubated for 2 hours at 4°C with end-over-end mixing (wheel in the Keenan Lab cold room set at 20 rpm works well)
6. Wash six times with 500-1000 μL 250 mM potassium acetate, 50 mM Hepes pH 7.4, 5 mM magnesium acetate, 15% glycerol and 0.1% digitonin
7. Elute in three successive 10 minute incubations with 200 μL of 1 M Glycine pH 3 supplemented with 0.1% Fos-choline-12. Add to 100 μL of 1 M Tris pH 9 or above.
    a. Resin can in fact be saved and used in a subsequent experiment, although capacity seems to decrease somewhat. Do not save resin for longer than one week

8. TCA precipitate.

9. Resuspend in Lamelli Sample Buffer and analyze by SDS-PAGE.

### 6.2.11 Co-immunoprecipitation analysis from 3xFlag Hek293 cells

1. Prepare membrane fraction from Hek293 cells.
2. Wash membrane pellet twice in assay buffer, recovering membranes by spinning
3. Resuspend membrane pellet in buffer containing 250 mM sucrose, 300 mM potassium acetate, 50 mM Hepes pH 7.4, 10 mM magnesium acetate, 2% digitonin.
    a. A higher salt concentration helps reduce background signal in these experiments.
4. Incubate for 30 minutes at 4°C with gentle end over end mixing (wheel in the Keenan Lab cold room set at 20-30 rpm works well)
5. Pellet 10 minutes at 10,000 x g to remove insoluble material.
6. Add soluble fraction to Flag M2 resin (Sigma) and incubate for 1 hour at 4°C with gentle end-over-end mixing (wheel in the Keenan Lab cold room set at 20 rpm works well)
7. Wash four times with 350 mM potassium acetate, 50 mM Hepes pH 7.4, 5 mM magnesium acetate, 250 mM sucrose and 0.1% digitonin.
8. Elute by 2 successive 30 minute incubations with same buffer as the wash but supplemented with 0.5 mg/mL 3xFlag peptide (ApexBio).
9. Clear residual resin from elutions by passing through a Paper Spincup (Pierce).
10. Layer elution on a 1 mL sucrose cushion 150 mM KCl, 50 mM Tris pH 7.4, 5 mM $MgCl_2$, 1 M Sucrose, 0.1% digitonin)
11. Pellet for 2 hours at 250,000 x g in a TLA100.3 rotor
10. Discard supernatant, resuspend pellet in Lamelli Sample Buffer and analyze by SDS-PAGE.

### 6.2.12 Plasmids

| Purpose | Identified | Plasmids Schematic |
|---|---|---|
| TMCO1 topology mapping | bSA143 | pGFP-3xFlag-Tmco1 |
| | bSA149 | pGFP-Opsin-3xFlag-Tmco1 |
| | bSA150 | pGFP-3xFlag-Tmco1-Opsin (Cterm) |
| | bSA152 | pGFP-3xFlag-Tmco1-Opsin (Loop) |
| | bSA153 | pGFP-3xFlag-Tmco1-Opsin (CC) |
| | bSA154 | pGFP-3xFlag-Tmco1-Opsin (CC2) |
| EMC3 topology mapping | bSA156 | pGFP-3xFlag-EMC3 |
| | bSS165 | pGFP-Opsin-3xFlag-EMC3 |
| | bSA168 | pGFP-3xFlag-EMC3-Opsin (Cterm) |
| | bSA167 | pGFP-3xFlag-EMC3-Opsin (Loop) |
| | bSA166 | pGFP-3xFlag-EMC3-Opsin (CC) |
| Get1 topology mapping | bSA169 | Yeplac195-Get1p-Get1-3xFlag |
| | bSA171 | Yeplac195-Get1p-3xGSS-Get1-3xFlag |
| | bSA172 | Yeplac195-Get1p-Get1-Opsin(Loop)-3xFlag |
| | bSA173 | Yeplac195-Get1p-Get1-Opsin(CC)-3xFlag |
| | bSA174 | Yeplac195-Get1p-Get1-3xFlag-Opsin(Cterm) |

**Table 3. Plasmids used in the subset of this work described in Anghel et. al. 2017**

## 6.3 Other protocols

### 6.3.1 TCA Precipitation of Protein Samples

This protocol is useful if starting with a large sample volume

1. Add TCA from a 100% stock to a final concentration of 25% and mix
2. Incubate 10 min at 4°C. Perform the entire procedure on ice until drying step.
3. Spin for 5 minutes at 20,000 x g.
4. Remove supernatant, leaving protein pellet intact.
5. If precipitating sucrose-containing fractions, wash pellet with 500 µL 30% TCA and spin again
    a. This removes residual sucrose. As sucrose is not soluble in acetone, it will otherwise precipitate in acetone and contaminate the gel sample
6. Wash pellet with 200µL ice-cold acetone.
7. Spin 10 minutes at 20,000 x g.
    a. Remove acetone quickly. Pellet is much more fragile at this step.
8. Repeat acetone wash.
    a. Failure to remove residual TCA will negatively affect the pH of the gel sample, making the pellet much harder to resuspend and the resulting SDS-PAGE gel run poorly.
9. Air dry pellet
10. Resuspend pellet in Lamelli Sample Buffer by incubating 20 minutes at 50°C
    a. Incubating gel samples at significantly higher temperature or for significantly longer will cause membrane proteins to aggregate.

### 6.3.2 Preparation of 10-50% Continuous Sucrose Gradients

1. Prepare 10% and 50% sucrose solutions in the buffer to be used for the final experiment
    a. For ribosomes, we use 150 mM potassium acetate, 50 mM Tris pH 7.5, 5 mM $MgCl_2$, supplemented with any detergents depending on the experiment
2. Fill a SW28 or SW28.1 tube with 10% sucrose solution until the half way mark by pouring
3. Fill a syringe with the required amount of 50% sucrose solution for the remaining half of the centrifuge tube and attach a metal cannula
4. Place syringe upside down such that the tip of the metal cannula is in the bottom of the centrifuge tube
5. Slowly extrude 50% sucrose solution taking care to minimize air bubbles that will flow to the top and disturb the interface between the two layers
6. Place tube(s) in dedicated holder on the Gradient Maker Station and run program
    a. Standard preset program for a 10-50% linear gradient takes around 7 minutes
7. Gradients should be stored at 4°C and used as soon as possible after they are poured, as diffusion will degrade the gradient over time

### 6.3.3 *In vitro* Translation and Insertion Assays

This protocol uses the Rabbit Reticulocyte Lysate system perfected in the Hegde Lab (Sharma et al., 2010)

1. For transcription, start by designing a DNA construct coding for a substrate
    a. Transcription will proceed either by a T7 or SP6 RNA polymerase. They should both work, although in the Keenan lab we have always used SP6. Cloning the substrate into a pS64 vector works well, as it includes the promoter as well as the Kozak sequence in the vector
    b. Include tags if desired. If including tags for detection, 3xFLAG and 3xHA work well as the triple nature makes them possible to detect by Western Blotting even at the low levels they are synthesized in this system
    c. If generating a truncated product to enrich for translating ribosomes, replace the stop codon with a valine.
    d. The transcription will be done from a PCR reaction (use at 50 ng/uL) or from the plasmid itself (use at >200 ng/uL). We generally use a column purified PCR product.
2. Set up Transcription (T1) reaction on ice by mixing 7.6 μL T1 master mix with 0.2 SuperaseIn, 2 μL PCR and 0.2 μL Polymerase
3. Incubate at for 1 hour at 40°C (for SP6 polymerase) or 37°C (for T7 polymerase)
4. Set up Translation (T2) reaction on ice by mixing 5.6 μL T2 master mix with 0.2 μL SuperaseIn, 2 μL T1 reaction to a total volume of 10 μL
    a. For some templates the $Mg^{2+}$ concentration may need to adjusted upwards by 1-2 mM

b. If doing co-translational insertion, supplement with 1 μL of 50% microccocal nuclease treated Hek293 membrane suspension. Note that some preparations of membranes can inhibit translation to a significant degree. If desired, include a positive control of Canine Pancreatic Membranes (Promega)

c. If the substrate is to be monitored by phosphor imaging, add 0.2 μL of $S^{35}$-labeled methionine at this stage

5. Incubate T2 reaction at 32°C

   a. Incubation time will dependent on the template length and translation difficulty, but 10 minutes for 10 kDa of substrate is a good starting point. Substrates larger than ~50kDa are increasingly difficult to translate in this system

   b. Translation will work at temperatures between 25 and 34°C although synthesis time must be adjusted empirically

6. If more than 0.2 μL is to be loaded on a gel (which is sufficient if detecting substrate by phosphor imaging), remove hemoglobin before proceeding as too much hemoglobin will interfere with SDS-PAGE

   a. If only the membrane fraction of a co-translational experiment is of interest, pellet membranes for 5 minutes at 13,000 x g, discard the supernatant and keep the pellet. Pellet should be chlorophorm-methanol extracted to remove lipids.

   b. An alternative method is to add an equal volume of the equilibrated Ni-NTA resin to the translation reaction, incubate 10 minutes at room temperature with occasional mixing, spin in tabletop picofuge and keep supernatant (resin should be red). This method cannot be employed if the substrate or any other proteins

that are desired are 6xHis tagged. If supernatant volume is too large, TCA precipitate it.

7. Mix with Lamelli Sample Buffer and denature by incubating 20 minutes at 50°C
8. Analyze by SDS-PAGE
    a. If sample is to be analyzed by immunoblotting, take note that the T2 mixture contains high levels of rabbit antibodies and therefore the use of an anti-rabbit-HRP secondary antibody is suboptimal

### 6.3.4 Chloroform-methanol extraction of proteins

This protocol was kindly provided by D. Allan Drummond. It is useful for extracting proteins from mixtures containing a large amount of lipids and also for purifying samples for mass spectrometry analysis.

1. To up to 100 µL of sample, add 400 µL methanol and vortex thoroughly.
2. Add 100 µL chloroform and vortex.
3. Add 300 µL water and vortex.
    a. The mixture should become cloudy after addition of water
4. Centrifuge 1 minute at 14,000 x g.
    a. Three layers should form: a large aqueous layer on top, a circular flake of protein in the interphase, and a smaller chloroform layer at the bottom.
5. Remove top aqueous layer carefully, trying not to disturb the protein flake.
6. Add 400 µL methanol and vortex.
7. Centrifuge 10 minutes at 20,000 x g
    a. Pellet will be very delicate after this step
8. Remove as much methanol as possible by pipetting without disturbing the pellet
9. Air dry the pellet to remove residual methanol
    a. If submitting samples to mass spectrometry, cover loosely with foil to prevent too much keratin from contaminating the sample
10. If analyzing by SDS-PAGE, resuspend pellet in Lamelli Sample Buffer by incubating 20 minutes at 50°

## References


Abell, B.M., Pool, M.R., Schlenker, O., Sinning, I., and High, S. (2004). Signal recognition particle mediates post-translational targeting in eukaryotes. EMBO J. *23*, 2755–2764.

Abell, B.M., Rabu, C., Leznicki, P., Young, J.C., and High, S. (2007). Post-translational integration of tail-anchored proteins is facilitated by defined molecular chaperones. J. Cell Sci. *120*, 1743–1751.

Aebi, M., Bernasconi, R., Clerc, S., and Molinari, M. (2010). N-glycan structures: recognition and processing in the ER. Trends Biochem. Sci. *35*, 74–82.

Alanay, Y., Ergüner, B., Utine, E., Haçarız, O., Kiper, P.O.S., Taşkıran, E.Z., Perçin, F., Uz, E., Sağıroğlu, M.Ş., Yuksel, B., et al. (2014). TMCO1 deficiency causes autosomal recessive cerebrofaciothoracic dysplasia. Am. J. Med. Genet. A. *164*, 291–304.

Almedom, R.B., Liewald, J.F., Hernando, G., Schultheis, C., Rayes, D., Pan, J., Schedletzky, T., Hutter, H., Bouzat, C., and Gottschalk, A. (2009). An ER-resident membrane protein complex regulates nicotinic acetylcholine receptor subunit composition at the synapse. EMBO J. *28*, 2636–2649.

Arkowitz, R. a., and Wickner, W. (1994). SecD and SecF are required for the proton electrochemical gradient stimulation of preprotein translocation. EMBO J. *13*, 954–963.

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

Aviram, N., and and Schuldiner, M. (2014). Embracing the void–how much do we really know about targeting and translocation to the endoplasmic reticulum? Curr Opin Cell Biol *29*, 8–17.

Aviram, N., Ast, T., Costa, E.A., Arakel, E.C., Chuartzman, S.G., Jan, C.H., Hassdenteufel, S., Dudek, J., Jung, M., and Schorr, S. (2016). The SND proteins constitute an alternative targeting route to the endoplasmic reticulum. Nature *540*, 134–138.

Bañó-Polo, M., Martínez-Garay, C.A., Grau, B., Martínez-Gil, L., and Mingarro, I. (2017). Membrane insertion and topology of the translocon-associated protein (TRAP) gamma subunit. Biochim. Biophys. Acta BBA - Biomembr. *1859*, 903–909.

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

Belleannee Clémence, Belghazi Maya, Labas Valérie, Teixeira-Gomes Ana-Paula, Gatti Jean Luc, Dacheux Jean-Louis, and Dacheux Françoise (2011). Purification and identification of sperm surface proteins and changes during epididymal maturation. PROTEOMICS *11*, 1952–1964.

Benz, M., Soll, J., and Ankele, E. (2013). <Emphasis Type="Italic">Arabidopsis thaliana</Emphasis> Oxa proteins locate to mitochondria and fulfill essential roles during embryo development. Planta *237*, 573–588.

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

Bernstein, H.D., Poritz, M.A., Strub, K., Hoben, P.J., Brenner, S., and and Walter, P. (1989). Model for signal sequence recognition from amino-acid sequence of 54K subunit of signal recognition particle. Nature *340*, 482–486.

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

van Bloois, E., Jan Haan, G., de Gier, J.-W., Oudega, B., and Luirink, J. (2004). F1F0 ATP synthase subunit c is targeted by the SRP to YidC in the E. coli inner membrane. FEBS Lett. *576*, 97–100.

van Bloois, E., Nagamori, S., Koningstein, G., Ullers, R.S., Preuss, M., Oudega, B., Harms, N., Kaback, H.R., Herrmann, J.M., and and Luirink, J. (2005). The Sec-independent function of Escherichia coli YidC is evolutionary-conserved and essential. J Biol Chem *280*, 12996–13003.

Boisramé, A., Chasles, M., Babour, A., Beckerich, J.-M., and Gaillardin, C. (2002). Sbh1p, a subunit of the Sec61 translocon, interacts with the chaperone calnexin in the yeast Yarrowia lipolytica. J. Cell Sci. *115*, 4947–4956.

Borgese, N., and Righi, M. (2010). Remote Origins of Tail-Anchored Proteins. Traffic *11*, 877–885.

Borowska, M.T., Dominik, P.K., Anghel, S.A., Kossiakoff, A.A., and Keenan, R.J. (2015). A YidC-like Protein in the Archaeal Plasma Membrane. Structure *23*, 1715–1724.

Botte, M., Zaccai, N.R., Nijeholt, J.L. à, Martin, R., Knoops, K., Papai, G., Zou, J., Deniaud, A., Karuppasamy, M., Jiang, Q., et al. (2016). A central cavity within the holo-translocon suggests a mechanism for membrane protein insertion. Sci. Rep. *6*, 38399.

Bozkurt, G., Stjepanovic, G., Vilardi, F., Amlacher, S., Wild, K., Bange, G., Favaloro, V., Rippe, K., Hurt, E., Dobberstein, B., et al. (2009). Structural insights into tail-anchored protein binding and membrane insertion by Get3. Proc. Natl. Acad. Sci. *106*, 21131–21136.

Bozkurt, G., Wild, K., Amlacher, S., Hurt, E., Dobberstein, B., and Sinning, I. (2010). The structure of Get4 reveals an α-solenoid fold adapted for multiple interactions in tail-anchored protein biogenesis. FEBS Lett. *584*, 1509–1514.

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

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

Braunger, K., Pfeffer, S., Shrimal, S., Gilmore, R., Berninghausen, O., Mandon, E.C., Becker, T., Förster, F., and Beckmann, R. (2018). Structural basis for coupling protein transport and N-glycosylation at the mammalian endoplasmic reticulum. Science eaar7899.

Brodsky, J.L., Goeckeler, J., and Schekman, R. (1995). BiP and Sec63p are required for both co- and posttranslational protein translocation into the yeast endoplasmic reticulum. Proc. Natl. Acad. Sci. *92*, 9643–9646.

Burdon, K.P., Macgregor, S., Hewitt, A.W., Sharma, S., Chidlow, G., Mills, R.A., Danoy, P., Casson, R., Viswanathan, A.C., Liu, J.Z., et al. (2011). Genome-wide association study identifies susceptibility loci for open angle glaucoma at TMCO1 and CDKN2B-AS1. Nat Genet *43*, 574–578.

Caglayan, A.O., Per, H., Akgumus, G., Gumus, H., Baranoski, J., Canpolat, M., Calik, M., Yikilmaz, A., Bilguvar, K., and Kumandas, S. (2013). Whole-exome sequencing identified a patient with TMCO1 defect syndrome and expands the phenotic spectrum. Clin Genet *84*, 394–395.

Cannon, K.S., Or, E., Clemons, W.M., Shibata, Y., and Rapoport, T.A. (2005). Disulfide bridge formation between SecY and a translocating polypeptide localizes the translocation pore to the center of SecY. J. Cell Biol. *169*, 219–225.

Capella-Gutierrez, S., Silla-Martinez, J.M., and and Gabaldon, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics *25*, 1972–1973.

Caramelo, J.J., and Parodi, A.J. (2008). Getting In and Out from Calnexin/Calreticulin Cycles. J. Biol. Chem. *283*, 10221–10225.

Carvalho, B., Sillars-Hardebol, A.H., Postma, C., Mongera, S., Droste, J.T.S., Obulkasim, A., Wiel, M. van de, Criekinge, W. van, Ylstra, B., Fijneman, R.J.A., et al. (2012). Colorectal adenoma to carcinoma progression is accompanied by changes in gene expression associated with ageing, chromosomal instability, and fatty acid metabolism. Cell. Oncol. *35*, 53–63.

Cavalier-Smith, T. (2002). The phagotrophic origin of eukaryotes and phylogenetic classification of Protozoa. Int J Syst Evol Microbiol *52*, 297–354.

Cavalier-Smith, T. (2013). Symbiogenesis: Mechanisms, Evolutionary Consequences, and Systematic Implications. Annu. Rev. Ecol. Evol. Syst. *44*, 145–172.

Chang, Y.-W., Chuang, Y.-C., Ho, Y.-C., Cheng, M.-Y., Sun, Y.-J., Hsiao, C.-D., and Wang, C. (2010). Crystal Structure of Get4-Get5 Complex and Its Interactions with Sgt2, Get3, and Ydj1. J. Biol. Chem. *285*, 9962–9970.

Chartron, J.W., Suloway, C.J.M., Zaslaver, M., and Clemons, W.M. (2010). Structural characterization of the Get4/Get5 complex and its interaction with Get3. Proc. Natl. Acad. Sci. *107*, 12127–12132.

Chartron, J.W., Hunt, K.C., and Frydman, J. (2016). Cotranslational signal-independent SRP preloading during membrane targeting. Nature *536*, 224.

Chavan, M., Yan, A., and Lennarz, W.J. (2005). Subunits of the translocon interact with components of the oligosaccharyl transferase complex. J. Biol. Chem. *280*, 22917–22924.

Chen, M., Samuelson, J.C., Jiang, F., Muller, M., Kuhn, A., and Dalbey, R.E. (2002). Direct Interaction of YidC with the Sec-independent Pf3 Coat Protein during Its Membrane Protein Insertion. J. Biol. Chem. *277*, 7670–7675.

Chen, W., Helenius, J., Braakman, I., and Helenius, A. (1995). Cotranslational folding and calnexin binding during glycoprotein synthesis. Proc. Natl. Acad. Sci. *92*, 6229–6233.

Chen, Y., Soman, R., Shanmugam, S.K., Kuhn, A., and Dalbey, R.E. (2014). The Role of the Strictly Conserved Positively Charged Residue Differs among the Gram-positive, Gram-negative, and Chloroplast YidC Homologs. J. Biol. Chem. *289*, 35656–35667.

Cherepanova, N.A., and Gilmore, R. (2016). Mammalian cells lacking either the cotranslational or posttranslocational oligosaccharyltransferase complex display substrate-dependent defects in asparagine linked glycosylation. Sci. Rep. *6*, 20946.

Cherepanova, N.A., Shrimal, S., and Gilmore, R. (2014). Oxidoreductase activity is necessary for N-glycosylation of cysteine-proximal acceptor sites in glycoproteins. J Cell Biol *206*, 525–539.

Chiti, F., and Dobson, C.M. (2006). Protein Misfolding, Functional Amyloid, and Human Disease. Annu. Rev. Biochem. *75*, 333–366.

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

Conti, B.J., Devaraneni, P.K., Yang, Z., David, L.L., and Skach, W.R. (2015). Cotranslational Stabilization of Sec62/63 within the ER Sec61 Translocon Is Controlled by Distinct Substrate-Driven Translocation Events. Mol. Cell *58*, 269–283.

Costa, E.A., Subramanian, K., Nunnari, J., and Weissman, J.S. (2018). Defining the physiological role of SRP in protein-targeting efficiency and specificity. Science eaar3607.

Cross, B.C.S., and High, S. (2009). Dissecting the physiological role of selective transmembrane-segment retention at the ER translocon. J. Cell Sci. *122*, 1768–1777.

Darriba, D., Taboada, G.L., Doallo, R., and and Posada, D. (2011). ProtTest 3: fast selection of best-fit models of protein evolution. Bioinformatics *27*, 1164–1165.

Dejgaard, K., Theberge, J.-F., Heath-Engel, H., Chevet, E., Tremblay, M.L., and Thomas, D.Y. (2010). Organization of the Sec61 Translocon, Studied by High Resolution Native Electrophoresis. J. Proteome Res. *9*, 1763–1771.

Dephoure, N., Zhou, C., Villen, J., Beausoleil, S.A., Bakalarski, C.E., Elledge, S.J., and Gygi, S.P. (2008). A quantitative atlas of mitotic phosphorylation. Proc. Natl. Acad. Sci. *105*, 10762–10767.

Deshaies, R.J., and Schekman, R. (1987). A yeast mutant defective at an early stage in import of secretory protein precursors into the endoplasmic reticulum. J. Cell Biol. *105*, 633–645.

Deshaies, R.J., and Schekman, R. (1989). SEC62 encodes a putative membrane protein required for protein translocation into the yeast endoplasmic reticulum. J. Cell Biol. *109*, 2653–2664.

Deshaies, R.J., Sanders, S.L., Feldheim, D.A., and Schekman, R. (1991). Assembly of yeast Sec proteins involved in translocation into the endoplasmic reticulum into a membrane-bound multisubunit complex. Nature *349*, 806–808.

Dettmer, U., Kuhn, P.-H., Abou-Ajram, C., Lichtenthaler, S.F., Krüger, M., Kremmer, E., Haass, C., and Haffner, C. (2010). Transmembrane Protein 147 (TMEM147) Is a Novel Component of the Nicalin-NOMO Protein Complex. J. Biol. Chem. *285*, 26174–26181.

Do, H., Falcone, D., Lin, J., Andrews, D.W., and Johnson, A.E. (1996). The Cotranslational Integration of Membrane Proteins into the Phospholipid Bilayer Is a Multistep Process. Cell *85*, 369–378.

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

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

Economou, A., Pogliano, J.A., Beckwith, J., Oliver, D.B., and Wickner, W. (1995). SecA membrane cycling at SecYEG is driven by distinct ATP binding and hydrolysis events and is regulated by SecD and SecF. Cell *83*, 1171–1181.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res *32*, 1792–1797.

Erdmann, F., Schauble, N., Lang, S., Jung, M., Honigmann, A., Ahmad, M., Dudek, J., Benedix, J., Harsman, A., and Kopp, A. (2011). Interaction of calmodulin with Sec61alpha limits Ca2+ leakage from the endoplasmic reticulum. EMBO J *30*, 17–31.

Facey, S.J., Neugebauer, S.A., Krauss, S., and Kuhn, A. (2007). The Mechanosensitive Channel Protein MscL Is Targeted by the SRP to The Novel YidC Membrane Insertion Pathway of Escherichia coli. J. Mol. Biol. *365*, 995–1004.

Falk, S., Ravaud, S., Koch, J., and Sinning, I. (2010). The C Terminus of the Alb3 Membrane Insertase Recruits cpSRP43 to the Thylakoid Membrane. J. Biol. Chem. *285*, 5954–5962.

Fedeles, S.V., So, J.-S., Shrikhande, A., Lee, S.H., Gallagher, A.-R., Barkauskas, C.E., Somlo, S., and Lee, A.-H. (2015). *Sec63* and *Xbp1* regulate IRE1α activity and polycystic disease severity. J. Clin. Invest. *125*, 1955–1967.

Feldheim, D., Rothblatt, J., and Schekman, R. (1992). Topology and functional domains of Sec63p, an endoplasmic reticulum membrane protein required for secretory protein translocation. Mol. Cell. Biol. *12*, 3288–3296.

Fleischer, T.C., Weaver, C.M., McAfee, K.J., Jennings, J.L., and Link, A.J. (2006). Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. Genes Dev. *20*, 1294–1307.

Fons, R.D., Bogert, B.A., and Hegde, R.S. (2003). Substrate-specific function of the translocon-associated protein complex during translocation across the ER membrane. J Cell Biol *160*, 529–539.

Frauenfeld, J., Gumbart, J., Van Der Sluis, E.O., Funes, S., Gartmann, M., Beatrix, B., Mielke, T., Berninghausen, O., Becker, T., and Schulten, K. (2011). Cryo-EM structure of the ribosome–SecYE complex in the membrane environment. Nat. Struct. Mol. Biol. *18*, 614.

Funes, S., Hasona, A., Bauerschmitt, H., Grubbauer, C., Kauff, F., Collins, R., Crowley, P.J., Palmer, S.R., Brady, L.J., and Herrmann, J.M. (2009). Independent gene duplications of the YidC/Oxa/Alb3 family enabled a specialized cotranslational function. Proc. Natl. Acad. Sci. *106*, 6656–6661.

Gafvelin, G., and von Heijne, G. (1994). Topological "frustration" in multispanning E. coli inner membrane proteins. Cell *77*, 401–412.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison Iii, C.A., and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat. Methods *6*, 343–345.

Gietz, R.D., and and Woods, R.A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. Methods Enzym. *350*, 87–96.

Glick, B.S., and Heijne, G.V. (1996). *Saccharomyces cerevisiae* mitochondria lack a bacterial-type Sec machinery. Protein Sci. *5*, 2651–2652.

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

Görlich, D., and Rapoport, T.A. (1993). Protein translocation into proteoliposomes reconstituted from purified components of the endoplasmic reticulum membrane. Cell *75*, 615–630.

Görlich, D., Hartmann, E., Prehn, S., and Rapoport, T.A. (1992). A protein of the endoplasmic reticulum involved early in polypeptide translocation. Nature *357*, 47–52.

Gottschalk, A., Almedom, R.B., Schedletzky, T., Anderson, S.D., Yates, J.R., and Schafer, W.R. (2005). Identification and characterization of novel nicotinic receptor-associated proteins in Caenorhabditis elegans. EMBO J. *24*, 2566–2578.

Gouaux, E. (1997). Channel-forming toxins: tales of transformation. Curr. Opin. Struct. Biol. *7*, 566–573.

Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., and and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst. Biol. *59*, 307–321.

Guna, A., Volkmar, N., Christianson, J.C., and Hegde, R.S. (2017). The ER membrane protein complex is a transmembrane domain insertase. Science.

Haffner, C., Frauli, M., Topp, S., Irmler, M., Hofmann, K., Regula, J.T., Bally-Cuif, L., and Haass, C. (2004). Nicalin and its binding partner Nomo are novel Nodal signaling antagonists. EMBO J. *23*, 3041–3050.

Hainzl, T., Huang, S., Merilainen, G., Brannstrom, K., and and Sauer-Eriksson, A.E. (2011). Structural basis of signal-sequence recognition by the signal recognition particle. Nat Struct Mol Biol *18*, 389–391.

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

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

Haque, M.E., Elmore, K.B., Tripathy, A., Koc, H., Koc, E.C., and Spremulli, L.L. (2010). Properties of the C-terminal Tail of Human Mitochondrial Inner Membrane Protein Oxa1L and Its Interactions with Mammalian Mitochondrial Ribosomes. J. Biol. Chem. *285*, 28353–28362.

Hartmann, E., Görlich, D., Kostka, S., Otto, A., Kraft, R., Knespel, S., Bürger, E., Rapoport, T.A., and Prehn, S. (1993). A tetrameric complex of membrane proteins in the endoplasmic reticulum. FEBS J. *214*, 375–381.

Hegde, R.S., and Keenan, R.J. (2011). Tail-anchored membrane protein insertion into the endoplasmic reticulum. Nat Rev Mol Cell Biol *12*, 787–798.

Heinrich, S.U., and Rapoport, T.A. (2003). Cooperation of transmembrane segments during the integration of a double-spanning protein into the ER membrane. EMBO J. *22*, 3654–3663.

Heinrich, S.U., Mothes, W., Brunner, J., and Rapoport, T.A. (2000). The Sec61p Complex Mediates the Integration of a Membrane Protein by Allowing Lipid Partitioning of the Transmembrane Domain. Cell *102*, 233–244.

Hell, K., Neupert, W., and Stuart, R.A. (2001). Oxa1p acts as a general membrane insertion machinery for proteins encoded by mitochondrial DNA. EMBO J. *20*, 1281–1288.

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

Hessa, T., Sharma, A., Mariappan, M., Eshleman, H.D., Gutierrez, E., and Hegde, R.S. (2011). Protein targeting and degradation are coupled for elimination of mislocalized proteins. Nature *475*, 394–397.

Holm, L., and and Park, J. (2000). DaliLite workbench for protein structure comparison. Bioinformatics *16*, 566–567.

Hori, O., Miyazaki, M., Tamatani, T., Ozawa, K., Takano, K., Okabe, M., Ikawa, M., Hartmann, E., Mai, P., Stern, D.M., et al. (2006). Deletion of SERP1/RAMP4, a Component of the Endoplasmic Reticulum (ER) Translocation Sites, Leads to ER Stress. Mol. Cell. Biol. *26*, 4257–4267.

Hu, J., Li, J., Qian, X., Denic, V., and Sha, B. (2009). The Crystal Structures of Yeast Get3 Suggest a Mechanism for Tail-Anchored Protein Membrane Insertion. PLOS ONE *4*, e8061.

Iwamuro, S., Saeki, M., and Kato, S. (1999). Multi-Ubiquitination of a Nascent Membrane Protein Produced in a Rabbit Reticulocyte Lysate1. J. Biochem. (Tokyo) *126*, 48–53.

Janda, C.Y., Li, J., Oubridge, C., Hernández, H., Robinson, C.V., and Nagai, K. (2010). Recognition of a signal peptide by the signal recognition particle. Nature *465*, 507.

Jia, L., Dienhart, M., Schramp, M., McCauley, M., Hell, K., and and Stuart, R.A. (2003). Yeast Oxa1 interacts with mitochondrial ribosomes: the importance of the C-terminal region of Oxa1. EMBO J *22*, 6438–6447.

Jia, L., Dienhart, M.K., and and Stuart, R.A. (2007). Oxa1 directly interacts with Atp9 and mediates its assembly into the mitochondrial F1Fo-ATP synthase complex. Mol Biol Cell *18*, 1897–1908.

Johnson, N., Haßdenteufel, S., Theis, M., Paton, A.W., Paton, J.C., Zimmermann, R., and High, S. (2013). The Signal Sequence Influences Post-Translational ER Translocation at Distinct Stages. PLOS ONE *8*, e75394.

Jung, S., Kim, J.E.H., Reithinger, J.H., and Kim, H. (2014). The Sec62–Sec63 translocon facilitates translocation of the C-terminus of membrane proteins. J Cell Sci *127*, 4270–4278.

Kamat, S., Yeola, S., Zhang, W., Bianchi, L., and Driscoll, M. (2014). NRA-2, a Nicalin Homolog, Regulates Neuronal Death by Controlling Surface Localization of Toxic Caenorhabditis elegans DEG/ENaC Channels. J. Biol. Chem. *289*, 11916–11926.

Kedrov, A., Kusters, I., Krasnikov, V.V., and Driessen, A.J. (2011). A single copy of SecYEG is sufficient for preprotein translocation. EMBO J. *30*, 4387–4397.

Kedrov, A., Sustarsic, M., de Keyzer, J., Caumanns, J.J., Wu, Z.C., and Driessen, A.J.M. (2013). Elucidating the Native Architecture of the YidC: Ribosome Complex. J. Mol. Biol. *425*, 4112–4124.

Keenan, R.J., Freymann, D.M., Walter, P., and Stroud, R.M. (1998). Crystal Structure of the Signal Sequence Binding Subunit of the Signal Recognition Particle. Cell *94*, 181–191.

Kelleher, D.J., and Gilmore, R. (2006). An evolving view of the eukaryotic oligosaccharyltransferase. Glycobiology *16*, 47R-62R.

Kelleher, D.J., Karaoglu, D., Mandon, E.C., and Gilmore, R. (2003). Oligosaccharyltransferase Isoforms that Contain Different Catalytic STT3 Subunits Have Distinct Enzymatic Properties. Mol. Cell *12*, 101–111.

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

Klostermann, E., Helling, I.D. gen, Carde, J.-P., and Schünemann, D. (2002). The thylakoid membrane protein ALB3 associates with the cpSecY-translocase in Arabidopsis thaliana. Biochem. J. *368*, 777–781.

Kohler, R., Boehringer, D., Greber, B., Bingel-Erlenmeyer, R., Collinson, I., Schaffitzel, C., and Ban, N. (2009). YidC and Oxa1 Form Dimeric Insertion Pores on the Translating Ribosome. Mol. Cell *34*, 344–353.

Komar, J., Alvira, S., Schulze, R.J., Martin, R., Nijeholt, J.A.L. a, Lee, S.C., Dafforn, T.R., Deckers-Hebestreit, G., Berger, I., Schaffitzel, C., et al. (2016). Membrane protein insertion and assembly by the bacterial holo-translocon SecYEG–SecDF–YajC–YidC. Biochem. J. *473*, 3341–3354.

Konno, M., Shirakawa, H., Miyake, T., Sakimoto, S., Nakagawa, T., and Kaneko, S. (2012). Calumin, a Ca2+-binding protein on the endoplasmic reticulum, alters the ion permeability of Ca2+ release-activated Ca2+ (CRAC) channels. Biochem. Biophys. Res. Commun. *417*, 784–789.

Kopito, R.R. (2000). Aggresomes, inclusion bodies and protein aggregation. Trends Cell Biol. *10*, 524–530.

Krieg, U.C., Johnson, A.E., and Walter, P. (1989). Protein translocation across the endoplasmic reticulum membrane: identification by photocross-linking of a 39-kD integral membrane glycoprotein as part of a putative translocation tunnel. J. Cell Biol. *109*, 2033–2043.

Krüger, V., Deckers, M., Hildenbeutel, M., van der Laan, M., Hellmers, M., Dreker, C., Preuss, M., Herrmann, J.M., Rehling, P., Wagner, R., et al. (2012). The Mitochondrial Oxidase Assembly Protein1 (Oxa1) Insertase Forms a Membrane Pore in Lipid Bilayers. J. Biol. Chem. *287*, 33314–33326.

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

Kumazaki, K., Chiba, S., Takemoto, M., Furukawa, A., Nishiyama, K., Sugano, Y., Mori, T., Dohmae, N., Hirata, K., Nakada-Nakura, Y., et al. (2014b). Structural basis of Sec-independent membrane protein insertion by YidC. Nature *509*, 516–520.

Kutay, U., Hartmann, E., and Rapoport, T. (1993). A class of membrane proteins with a C-terminal anchor. Trends Cell Biol. *3*, 72–75.

van der Laan, M., Bechtluft, P., Kol, S., Nouwen, N., and Driessen, A.J.M. (2004). F1F0 ATP synthase subunit c is a substrate of the novel YidC pathway for membrane protein biogenesis. J. Cell Biol. *165*, 213–222.

Lakkaraju, A.K., Abrami, L., Lemmin, T., Blaskovic, S., Kunz, B., Kihara, A., Peraro, M.D., and Goot, F.G. van der (2012a). Palmitoylated calnexin is a key component of the ribosome–translocon complex. EMBO J. *31*, 1823–1835.

Lakkaraju, A.K.K., Mary, C., Scherrer, A., Johnson, A.E., and Strub, K. (2008). SRP Keeps Polypeptides Translocation-Competent by Slowing Translation to Match Limiting ER-Targeting Sites. Cell *133*, 440–451.

Lakkaraju, A.K.K., Thankappan, R., Mary, C., Garrison, J.L., Taunton, J., and Strub, K. (2012b). Efficient secretion of small proteins in mammalian cells relies on Sec62-dependent posttranslational translocation. Mol. Biol. Cell *23*, 2712–2722.

Lang, B.F., Burger, G., O'kelly, C.J., Cedergren, R., Golding, G.B., Lemieux, C., Sankoff, D., Turmel, M., and Gray, M.W. (1997). An ancestral mitochondrial DNA resembling a eubacterial genome in miniature. Nature *387*, 493–497.

Lang, S., Erdmann, F., Jung, M., Wagner, R., Cavalie, A., and and Zimmermann, R. (2011). Sec61 complexes form ubiquitous ER Ca2+ leak channels. Channels Austin *5*, 228–235.

Lang, S., Benedix, J., Fedeles, S.V., Schorr, S., Schirra, C., Schäuble, N., Jalal, C., Greiner, M., Haßdenteufel, S., Tatzelt, J., et al. (2012). Different effects of Sec61α, Sec62 and Sec63 depletion on transport of polypeptides into the endoplasmic reticulum of mammalian cells. J Cell Sci *125*, 1958–1969.

Leznicki, P., Clancy, A., Schwappach, B., and High, S. (2010). Bat3 promotes the membrane integration of tail-anchored proteins. J Cell Sci *123*, 2170–2178.

Li, L., Park, E., Ling, J., Ingram, J., Ploegh, H., and Rapoport, T.A. (2016). Crystal structure of a substrate-engaged SecY protein-translocation channel. Nature *531*, 395–399.

Linn, F., von Heijne, Gunnar, Uhlén, Mathias, and Berglund, Lisa (2010). Prediction of the human membrane proteome. PROTEOMICS *10*, 1141–1149.

Lipp, J., Dobberstein, B., and Haeuptle, M.-T. (1987). Signal recognition particle arrests elongation of nascent secretory and membrane proteins at multiple sites in a transient manner. J. Biol. Chem. *262*, 1680–1684.

Liu, Y., and and Eisenberg, D. (2002). 3D domain swapping: as domains continue to swap. Protein Sci *11*, 1285–1299.

Locker, J.K., Rose, J.K., Horzinek, M.C., and Rottier, P. (1992). Membrane assembly of the triple-spanning coronavirus M protein. Individual transmembrane domains show preferred orientation. J. Biol. Chem. *267*, 21911–21918.

Losfeld, M.E., Ng, B.G., Kircher, M., Buckingham, K.J., Turner, E.H., Eroshkin, A., Smith, J.D., Shendure, J., Nickerson, D.A., Bamshad, M.J., et al. (2014). A new congenital disorder of glycosylation caused by a mutation in SSR4, the signal sequence receptor 4 protein of the TRAP complex. Hum. Mol. Genet. *23*, 1602–1605.

Luirink, J., Samuelsson, T., and de Gier, J.-W. (2001). YidC/Oxa1p/Alb3: evolutionarily conserved mediators of membrane protein assembly. FEBS Lett. *501*, 1–5.

Lupas, A., Van Dyke, M., and and Stock, J. (1991). Predicting coiled coils from protein sequences. Science *252*, 1162–1164.

Ma, H., Dang, Y., Wu, Y., Jia, G., Anaya, E., Zhang, J., Abraham, S., Choi, J.G., Shi, G., and Qi, L. (2015). A CRISPR-Based Screen Identifies Genes Essential for West-Nile-Virus-Induced Cell Death. Cell Rep *12*, 673–683.

Mades, A., Gotthardt, K., Awe, K., Stieler, J., Döring, T., Füser, S., and Prange, R. (2012). Role of Human Sec63 in Modulating the Steady-State Levels of Multi-Spanning Membrane Proteins. PLOS ONE *7*, e49243.

Makarova, K.S., Galperin, M.Y., and and Koonin, E.V. (2015). Comparative genomic analysis of evolutionarily conserved but functionally uncharacterized membrane proteins in archaea: Prediction of novel components of secretion, membrane remodeling and glycosylation systems. Biochimie.

Mariappan, M., Li, X., Stefanovic, S., Sharma, A., Mateja, A., Keenan, R.J., and Hegde, R.S. (2010). A ribosome-associating factor chaperones tail-anchored membrane proteins. Nature *466*, 1120–1124.

Mariappan, M., Mateja, A., Dobosz, M., Bove, E., Hegde, R.S., and Keenan, R.J. (2011). The mechanism of membrane-associated steps in tail-anchored protein insertion. Nature *477*, 61–66.

Mason, N., Ciufo, L.F., and Brown, J.D. (2000). Elongation arrest is a physiologically important function of signal recognition particle. EMBO J. *19*, 4164.

Mateja, A., Szlachcic, A., Downing, M.E., Dobosz, M., Mariappan, M., Hegde, R.S., and Keenan, R.J. (2009). The structural basis of tail-anchored membrane protein recognition by Get3. Nature *461*, 361–366.

Mateja, A., Paduch, M., Chang, H.-Y., Szydlowska, A., Kossiakoff, A.A., Hegde, R.S., and Keenan, R.J. (2015). Structure of the Get3 targeting factor in complex with its membrane protein cargo. Science *347*, 1152–1155.

Matsuyama, S., Fujita, Y., and Mizushima, S. (1993). SecD is involved in the release of translocated secretory proteins from the cytoplasmic membrane of Escherichia coli. EMBO J. *12*, 265–270.

McCormick, P.J., Miao, Y., Shao, Y., Lin, J., and Johnson, A.E. (2003). Cotranslational Protein Integration into the ER Membrane Is Mediated by the Binding of Nascent Chains to Translocon Proteins. Mol. Cell *12*, 329–341.

Ménétret, J.-F., Hegde, R.S., Aguiar, M., Gygi, S.P., Park, E., Rapoport, T.A., and Akey, C.W. (2008). Single Copies of Sec61 and TRAP Associate with a Nontranslating Mammalian Ribosome. Structure *16*, 1126–1137.

Meyer, H.-A., Grau, H., Kraft, R., Kostka, S., Prehn, S., Kalies, K.-U., and Hartmann, E. (2000). Mammalian Sec61 Is Associated with Sec62 and Sec63. J. Biol. Chem. *275*, 14550–14557.

Minami, R., Hayakawa, A., Kagawa, H., Yanagi, Y., Yokosawa, H., and Kawahara, H. (2010). BAG-6 is essential for selective elimination of defective proteasomal substrates. J. Cell Biol. jcb.200908092.

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

Molinari, M., Eriksson, K.K., Calanca, V., Galli, C., Cresswell, P., Michalak, M., and Helenius, A. (2004). Contrasting Functions of Calreticulin and Calnexin in Glycoprotein Folding and ER Quality Control. Mol. Cell *13*, 125–135.

Moore, M., Harrison, M.S., Peterson, E.C., and Henry, R. (2000). Chloroplast Oxa1p Homolog Albino3 Is Required for Post-translational Integration of the Light Harvesting Chlorophyll-binding Protein into Thylakoid Membranes. J. Biol. Chem. *275*, 1529–1532.

Moore, M., Goforth, R.L., Mori, H., and Henry, R. (2003). Functional interaction of chloroplast SRP/FtsY with the ALB3 translocase in thylakoids: substrate not required. J. Cell Biol. *162*, 1245–1254.

Mori, H., Summer, E.J., Ma, X., and Cline, K. (1999). Component Specificity for the Thylakoidal Sec and Delta Ph–Dependent Protein Transport Pathways. J. Cell Biol. *146*, 45–56.

Müller, L., Escauriaza, M.D. de, Lajoie, P., Theis, M., Jung, M., Müller, A., Burgard, C., Greiner, M., Snapp, E.L., Dudek, J., et al. (2010). Evolutionary Gain of Function for the ER Membrane Protein Sec62 from Yeast to Humans. Mol. Biol. Cell *21*, 691–703.

Nagamori, S., Smirnova, I.N., and Kaback, H.R. (2004). Role of YidC in folding of polytopic membrane proteins. J Cell Biol *165*, 53–62.

Ng, B.G., Raymond, K., Kircher, M., Buckingham, K.J., Wood, T., Shendure, J., Nickerson, D.A., Bamshad, M.J., University of Washington Center for Mendelian Genomics, Wong, J.T.S., et al. (2015). Expanding the Molecular and Clinical Phenotype of SSR4-CDG. Hum. Mutat. *36*, 1048–1051.

Ng, D.T., Brown, J.D., and Walter, P. (1996). Signal sequences specify the targeting route to the endoplasmic reticulum membrane. J. Cell Biol. *134*, 269–278.

Nicolson, G.L. (2014). The Fluid—Mosaic Model of Membrane Structure: Still relevant to understanding the structure, function and dynamics of biological membranes after more than 40years. Biochim. Biophys. Acta BBA - Biomembr. *1838*, 1451–1466.

Nilsson, I.M., and Heijne, G. von (1993). Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. J. Biol. Chem. *268*, 5798–5801.

Nilsson, I., Kelleher, D.J., Miao, Y., Shao, Y., Kreibich, G., Gilmore, R., Heijne, G. von, and Johnson, A.E. (2003). Photocross-linking of nascent chains to the STT3 subunit of the oligosaccharyltransferase complex. J. Cell Biol. *161*, 715–725.

Oh, E., Becker, A.H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., Nichols, R.J., Typas, A., Gross, C.A., Kramer, G., et al. (2011). Selective Ribosome Profiling Reveals the Cotranslational Chaperone Action of Trigger Factor In Vivo. Cell *147*, 1295–1308.

Olsen, J.V., Vermeulen, M., Santamaria, A., Kumar, C., Miller, M.L., Jensen, L.J., Gnad, F., Cox, J., Jensen, T.S., Nigg, E.A., et al. (2010). Quantitative Phosphoproteomics Reveals Widespread Full Phosphorylation Site Occupancy During Mitosis. Sci. Signal. *3*, ra3–ra3.

Panzner, S., Dreier, L., Hartmann, E., Kostka, S., and Rapoport, T.A. (1995). Posttranslational protein transport in yeast reconstituted with a purified complex of Sec proteins and Kar2p. Cell *81*, 561–570.

Park, E., and Rapoport, T.A. (2012). Bacterial protein translocation requires only one copy of the SecY complex in vivo. J Cell Biol jcb-201205140.

Pasch, J.C., Nickelsen, J., and Schünemann, D. (2005). The yeast split-ubiquitin system to study chloroplast membrane protein interactions. Appl. Microbiol. Biotechnol. *69*, 440–447.

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

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

Pfeffer, S., Dudek, J., Gogala, M., Schorr, S., Linxweiler, J., Lang, S., Becker, T., Beckmann, R., Zimmermann, R., and Förster, F. (2014). Structure of the mammalian oligosaccharyl-transferase complex in the native ER protein translocon. Nat. Commun. *5*, 3072.

Pfeffer, S., Dudek, J., Schaffer, M., Ng, B.G., Albert, S., Plitzko, J.M., Baumeister, W., Zimmermann, R., Freeze, H.H., Engel, B.D., et al. (2017). Dissecting the molecular organization of the translocon-associated protein complex. Nat. Commun. *8*, 14516.

Pieren, M., Galli, C., Denzel, A., and Molinari, M. (2005). The use of calnexin and calreticulin by cellular and viral glycoproteins. J. Biol. Chem. *280*, 28265–28271.

Plath, K., and Rapoport, T.A. (2000). Spontaneous Release of Cytosolic Proteins from Posttranslational Substrates before Their Transport into the Endoplasmic Reticulum. J Cell Biol *151*, 167–178.

Plath, K., Wilkinson, B.M., Stirling, C.J., and Rapoport, T.A. (2004). Interactions between Sec Complex and Prepro-α-Factor during Posttranslational Protein Transport into the Endoplasmic Reticulum. Mol. Biol. Cell *15*, 1–10.

Poljak, K., Selevsek, N., Ngwa, E., Grossmann, J., Losfeld, M.E., and Aebi, M. (2018). Quantitative Profiling of N-linked Glycosylation Machinery in Yeast Saccharomyces cerevisiae. Mol. Cell. Proteomics *17*, 18–30.

Pool, M.R. (2009). A trans-membrane segment inside the ribosome exit tunnel triggers RAMP4 recruitment to the Sec61p translocase. J. Cell Biol. *185*, 889–902.

Preuss, M., Ott, M., Funes, S., Luirink, J., and and Herrmann, J.M. (2005). Evolution of mitochondrial oxa proteins from bacterial YidC. Inherited and acquired functions of a conserved protein insertion machinery. J Biol Chem *280*, 13004–13011.

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

Rapoport, T.A., Li, L., and Park, E. (2017). Structural and mechanistic insights into protein translocation. Annu. Rev. Cell Dev. Biol. *33*, 369–390.

Reithinger, J.H., Kim, J.E.H., and Kim, H. (2013). Sec62 Protein Mediates Membrane Insertion and Orientation of Moderately Hydrophobic Signal Anchor Proteins in the Endoplasmic Reticulum (ER). J. Biol. Chem. *288*, 18058–18067.

Richard, M., Boulin, T., Robert, V.J.P., Richmond, J.E., and Bessereau, J.-L. (2013). Biosynthesis of ionotropic acetylcholine receptors requires the evolutionarily conserved ER membrane complex. Proc. Natl. Acad. Sci. *110*, E1055–E1063.

Rosemond, E., Rossi, M., McMillin, S.M., Scarselli, M., Donaldson, J.G., and Wess, J. (2011). Regulation of M3 Muscarinic Receptor Expression and Function by Transmembrane Protein 147. Mol. Pharmacol. *79*, 251–261.

Rothblatt, J.A., and Meyer, D.I. (1986). Secretion in yeast: Reconstitution of the translocation and glycosylation of α-factor and invertase in a homologous cell-free system. Cell *44*, 619–628.

Ruiz-Canada, C., Kelleher, D.J., and Gilmore, R. (2009). Cotranslational and Posttranslational N-Glycosylation of Polypeptides by Distinct Mammalian OST Isoforms. Cell *136*, 272–283.

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

Sachelaru, I., Petriman, N.A., Kudva, R., Kuhn, P., Welte, T., Knapp, B., Drepper, F., Warscheid, B., and Koch, H.-G. (2013). YidC Occupies the Lateral Gate of the SecYEG Translocon and Is Sequentially Displaced by a Nascent Membrane Protein. J. Biol. Chem. *288*, 16295–16307.

Sachelaru, I., Winter, L., Knyazev, D.G., Zimmermann, M., Vogt, A., Kuttner, R., Ollinger, N., Siligan, C., Pohl, P., and Koch, H.-G. (2017). YidC and SecYEG form a heterotetrameric protein translocation channel. Sci. Rep. *7*.

Sadlish, H., Pitonzo, D., Johnson, A.E., and Skach, W.R. (2005). Sequential triage of transmembrane segments by Sec61α during biogenesis of a native multispanning membrane protein. Nat. Struct. Mol. Biol. *12*, 870–878.

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

Samuelson, J.C., Jiang, F., Yi, L., Chen, M., Gier, J.-W. de, Kuhn, A., and Dalbey, R.E. (2001). Function of YidC for the Insertion of M13 Procoat Protein in *Escherichia coli* . J. Biol. Chem. *276*, 34847–34852.

Satoh, T., Ohba, A., Liu, Z., Inagaki, T., and Satoh, A.K. (2015). dPob/EMC is essential for biosynthesis of rhodopsin and other multi-pass membrane proteins in Drosophila photoreceptors. ELife *4*, e06306.

Saurí, A., McCormick, P.J., Johnson, A.E., and Mingarro, I. (2007). Sec61α and TRAM are Sequentially Adjacent to a Nascent Viral Membrane Protein during its ER Integration. J. Mol. Biol. *366*, 366–374.

Schaffitzel, C., and and Ban, N. (2007). Generation of ribosome nascent chain complexes for structural and functional studies. J. Struct. Biol. *158*, 463–471.

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

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

Schröder, K., Martoglio, B., Hofmann, M., Hölscher, C., Hartmann, E., Prehn, S., Rapoport, T.A., and Dobberstein, B. (1999). Control of glycosylation of MHC class II-associated invariant chain by translocon-associated RAMP4. EMBO J. *18*, 4804–4815.

Schuenemann, D., Gupta, S., Persello-Cartieaux, F., Klimyuk, V.I., Jones, J.D.G., Nussaume, L., and Hoffman, N.E. (1998). A novel signal recognition particle targets light-harvesting proteins to the thylakoid membranes. Proc. Natl. Acad. Sci. *95*, 10312–10316.

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

Schulz, G.E. (1996). Porins: general to specific, native to engineered passive pores. Curr. Opin. Struct. Biol. *6*, 485–490.

Schulz, B.L., and Aebi, M. (2009). Analysis of Glycosylation Site Occupancy Reveals a Role for Ost3p and Ost6p in Site-specific N-Glycosylation Efficiency. Mol. Cell. Proteomics *8*, 357–364.

Schulz, B.L., Stirnimann, C.U., Grimshaw, J.P.A., Brozzo, M.S., Fritsch, F., Mohorko, E., Capitani, G., Glockshuber, R., Grütter, M.G., and Aebi, M. (2009). Oxidoreductase activity of oligosaccharyltransferase subunits Ost3p and Ost6p defines site-specific glycosylation efficiency. Proc. Natl. Acad. Sci. *106*, 11061–11066.

Schulze, R.J., Komar, J., Botte, M., Allen, W.J., Whitehouse, S., Gold, V.A.M., Nijeholt, J.A.L. a, Huard, K., Berger, I., Schaffitzel, C., et al. (2014). Membrane protein insertion and proton-motive-force-dependent secretion through the bacterial holo-translocon SecYEG–SecDF–YajC–YidC. Proc. Natl. Acad. Sci. *111*, 4844–4849.

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

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

Seppälä, S., Slusky, J.S., Lloris-Garcerá, P., Rapp, M., and Heijne, G. von (2010). Control of Membrane Protein Topology by a Single C-Terminal Residue. Science *328*, 1698–1700.

Serdiuk, T., Balasubramaniam, D., Sugihara, J., Mari, S.A., Kaback, H.R., and Müller, D.J. (2016). YidC assists the stepwise and stochastic folding of membrane proteins. Nat. Chem. Biol. *12*, 911–917.

Shao, S., and Hegde, R.S. (2011a). A Calmodulin-Dependent Translocation Pathway for Small Secretory Proteins. Cell *147*, 1576–1588.

Shao, S., and Hegde, R.S. (2011b). Membrane Protein Insertion at the Endoplasmic Reticulum. Annu. Rev. Cell Dev. Biol. *27*, 25–56.

Sharma, A., Mariappan, M., Appathurai, S., and Hegde, R.S. (2010). In Vitro Dissection of Protein Translocation into the Mammalian Endoplasmic Reticulum. Methods Mol. Biol. Clifton NJ *619*, 339–363.

Sharma, S., Burdon, K.P., Chidlow, G., Klebe, S., Crawford, A., Dimasi, D.P., Dave, A., Martin, S., Javadiyan, S., Wood, J.P.M., et al. (2012). Association of Genetic Variants in the *TMCO1* Gene with Clinical Parameters Related to Glaucoma and Characterization of the Protein in the Eye. Investig. Opthalmology Vis. Sci. *53*, 4917.

Sherrill, J., Mariappan, M., Dominik, P., Hegde, R.S., and Keenan, R.J. (2011). A Conserved Archaeal Pathway for Tail-Anchored Membrane Protein Insertion. Traffic *12*, 1119–1123.

Shibatani, T., David, L.L., McCormack, A.L., Frueh, K., and Skach, W.R. (2005). Proteomic Analysis of Mammalian Oligosaccharyltransferase Reveals Multiple Subcomplexes that Contain Sec61, TRAP, and Two Potential New Subunits. Biochemistry (Mosc.) *44*, 5982–5992.

Shin, B.K., Wang, H., Yim, A.M., Naour, F.L., Brichory, F., Jang, J.H., Zhao, R., Puravs, E., Tra, J., Michael, C.W., et al. (2003). Global Profiling of the Cell Surface Proteome of Cancer Cells Uncovers an Abundance of Proteins with Chaperone Function. J. Biol. Chem. *278*, 7607–7616.

Shrimal, S., Cherepanova, N.A., and Gilmore, R. (2017). DC2 and KCP2 mediate the interaction between the oligosaccharyltransferase and the ER translocon. J Cell Biol *216*, 3625–3638.

Simon, S.M., and Blobel, G. (1991). A protein-conducting channel in the endoplasmic reticulum. Cell *65*, 371–380.

Simon, S.M., Blobel, G., and Zimmerberg, J. (1989). Large aqueous channels in membrane vesicles derived from the rough endoplasmic reticulum of canine pancreas or the plasma membrane of Escherichia coli. Proc. Natl. Acad. Sci. *86*, 6176–6180.

Singer, S.J., and Nicolson, G.L. (1972). The Fluid Mosaic Model of the Structure of Cell Membranes. Science *175*, 720.

Skowronek, M.H., Rotter, M., and Haas, I.G. (1995). Molecular Characterization of a Novel Mammalian DnaJ-Like Sec63p Homolog. Biol. Chem. *380*, 1133–1138.

Snapp, E.L., Reinhart, G.A., Bogert, B.A., Lippincott-Schwartz, J., and Hegde, R.S. (2004). The organization of engaged and quiescent translocons in the endoplasmic reticulum of mammalian cells. J. Cell Biol. *164*, 997–1007.

Soding, J., Biegert, A., and and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res *33*, 244–248.

Sojka, S., Amin, N.M., Gibbs, D., Christine, K.S., Charpentier, M.S., and and Conlon, F.L. (2014). Congenital heart disease protein 5 associates with CASZ1 to maintain myocardial tissue integrity. Development *141*, 3040–3049.

Soldà, T., Garbi, N., Hämmerling, G.J., and Molinari, M. (2006). Consequences of ERp57 Deletion on Oxidative Folding of Obligate and Facultative Clients of the Calnexin Cycle. J. Biol. Chem. *281*, 6219–6226.

Sommer, N., Junne, T., Kalies, K.-U., Spiess, M., and Hartmann, E. (2013). TRAP assists membrane protein topogenesis at the mammalian ER membrane. Biochim. Biophys. Acta BBA - Mol. Cell Res. *1833*, 3104–3111.

Sousa, M., and Parodi, A.J. (1995). The molecular basis for the recognition of misfolded glycoproteins by the UDP-Glc:glycoprotein glucosyltransferase. EMBO J. *14*, 4196–4203.

Spang, A., Saw, J.H., Jorgensen, S.L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A.E., van Eijk, R., Schleper, C., Guy, L., and Ettema, T.J.G. (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. Nature *521*, 173–179.

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

Stefanovic, S., and Hegde, R.S. (2007). Identification of a Targeting Factor for Posttranslational Membrane Protein Insertion into the ER. Cell *128*, 1147–1159.

Stefer, S., Reitz, S., Wang, F., Wild, K., Pang, Y.-Y., Schwarz, D., Bomke, J., Hein, C., Löhr, F., Bernhard, F., et al. (2011). Structural Basis for Tail-Anchored Membrane Protein Biogenesis by the Get3-Receptor Complex. Science *333*, 758–762.

Suloway, C.J., Rome, M.E., and and Clemons, W.M., Jr (2012). Tail-anchor targeting by a Get3 tetramer: the structure of an archaeal homologue. EMBO J *31*, 707–719.

Suloway, C.J.M., Chartron, J.W., Zaslaver, M., and Clemons, W.M. (2009). Model for eukaryotic tail-anchored protein binding based on the structure of Get3. Proc. Natl. Acad. Sci. *106*, 14849–14854.

Szyrach, G., Ott, M., Bonnefoy, N., Neupert, W., and Herrmann, J.M. (2003). Ribosome binding to the Oxa1 complex facilitates co-translational protein insertion in mitochondria. EMBO J. *22*, 6448–6457.

Tai, V.W.-F., and Imperiali, B. (2001). Substrate specificity of the glycosyl donor for oligosaccharyl transferase. J. Org. Chem. *66*, 6217–6228.

Tamborero, S., Vilar, M., Martínez-Gil, L., Johnson, A.E., and Mingarro, I. (2011). Membrane Insertion and Topology of the Translocating Chain-Associating Membrane Protein (TRAM). J. Mol. Biol. *406*, 571–582.

Tripathi, A., Mandon, E.C., Gilmore, R., and Rapoport, T.A. (2017). Two alternative binding mechanisms connect the protein translocation Sec71-Sec72 complex with heat shock proteins. J. Biol. Chem. *292*, 8007–8018.

Tsirigos, K.D., Peters, C., Shu, N., Kall, L., and and Elofsson, A. (2015). The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Res *43*, 401–407.

Tsukazaki, T., Mori, H., Echizen, Y., Ishitani, R., Fukai, S., Tanaka, T., Perederina, A., Vassylyev, D.G., Kohno, T., Maturana, A.D., et al. (2011). Structure and function of a membrane component SecDF that enhances protein export. Nature *474*, 235–238.

Tyedmers, J., Lerner, M., Bies, C., Dudek, J., Skowronek, M.H., Haas, I.G., Heim, N., Nastainczyk, W., Volkmer, J., and Zimmermann, R. (2000). Homologs of the yeast Sec complex subunits Sec62p and Sec63p are abundant proteins in dog pancreas microsomes. Proc. Natl. Acad. Sci. *97*, 7214–7219.

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

Valent, Q.A., Scotti, P.A., High, S., de Gier, J.L., von Heijne, G., Lentzen, G., Wintermeyer, W., Oudega, B., and Luirink, J. (1998). The *Escherichia coli* SRP and SecB targeting pathways converge at the translocon. EMBO J. *17*, 2504.

Vilardi, F., Lorenz, H., and Dobberstein, B. (2011). WRB is the receptor for TRC40/Asna1-mediated insertion of tail-anchored proteins into the ER membrane. J. Cell Sci. *124*, 1301.

Vilardi, F., Stephan, M., Clancy, A., Janshoff, A., and Schwappach, B. (2014). WRB and CAML are necessary and sufficient to mediate tail-anchored protein targeting to the ER membrane. PloS One *9*, e85033.

Vogl, C., Panou, I., Yamanbaeva, G., Wichmann, C., Mangosing, S.J., Vilardi, F., Indzhykulian, A.A., Pangrsic, T., Santarelli, R., and Rodriguez-Ballesteros, M. (2016). Tryptophan-rich basic protein (WRB) mediates insertion of the tail-anchored protein otoferlin and is required for hair cell exocytosis and hearing. EMBO J *35*, 2536–2552.

Voigt, S., Jungnickel, B., Hartmann, E., and Rapoport, T.A. (1996). Signal sequence-dependent function of the TRAM protein during early phases of protein transport across the endoplasmic reticulum membrane. J. Cell Biol. *134*, 25–35.

Voorhees, R.M., and Hegde, R.S. (2015). Structures of the scanning and engaged states of the mammalian SRP-ribosome complex. Elife *4*.

Voorhees, R.M., and Hegde, R.S. (2016). Structure of the Sec61 channel opened by a signal sequence. Science *351*, 88–91.

Voorhees, R.M., Fernández, I.S., Scheres, S.H., and Hegde, R.S. (2014). Structure of the mammalian ribosome-Sec61 complex to 3.4 Å resolution. Cell *157*, 1632–1643.

Wagner, S., Pop, O., Haan, G.-J., Baars, L., Koningstein, G., Klepsch, M.M., Genevaux, P., Luirink, J., and Gier, J.-W. de (2008). Biogenesis of MalF and the MalFGK2 Maltose Transport Complex in Escherichia coli Requires YidC. J. Biol. Chem. *283*, 17881–17890.

Walter, P., and Blobel, G. (1981). Translocation of proteins across the endoplasmic reticulum III. Signal recognition protein (SRP) causes signal sequence-dependent and site-specific arrest of chain elongation that is released by microsomal membranes. J. Cell Biol. *91*, 557.

Walter, B., Hristou, A., Nowaczyk, M.M., and Schünemann, D. (2015). In vitro reconstitution of co-translational D1 insertion reveals a role of the cpSec–Alb3 translocase and Vipp1 in photosystem II biogenesis. Biochem. J. *468*, 315–324.

Wang, L., and Dobberstein, B. (1999). Oligomeric complexes involved in translocation of proteins across the membrane of the endoplasmic reticulum. FEBS Lett. *457*, 316–322.

Wang, P., and Dalbey, R.E. (2011). Inserting membrane proteins: The YidC/Oxa1/Alb3 machinery in bacteria, mitochondria, and chloroplasts. Biochim. Biophys. Acta BBA - Biomembr. *1808*, 866–875.

Wang, F., Brown, E.C., Mak, G., Zhuang, J., and Denic, V. (2010). A Chaperone Cascade Sorts Proteins for Posttranslational Membrane Insertion into the Endoplasmic Reticulum. Mol. Cell *40*, 159–171.

Wang, F., Whynot, A., Tung, M., and Denic, V. (2011a). The Mechanism of Tail-Anchored Protein Insertion into the ER Membrane. Mol. Cell *43*, 738–750.

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

Wang, Q., Liu, Y., Soetandyo, N., Baek, K., Hegde, R., and Ye, Y. (2011b). A Ubiquitin Ligase-Associated Chaperone Holdase Maintains Polypeptides in Soluble States for Proteasome Degradation. Mol. Cell *42*, 758–770.

Wang, Q.-C., Zheng, Q., Tan, H., Zhang, B., Li, X., Yang, Y., Yu, J., Liu, Y., Chai, H., Wang, X., et al. (2016). TMCO1 Is an ER Ca2+ Load-Activated Ca2+ Channel. Cell *165*, 1454–1466.

Wang, S., Sun, S., Li, Z., Zhang, R., and Xu, J. (2017). Accurate De Novo Prediction of Protein Contact Map by Ultra-Deep Learning Model. PLoS Comput Biol *13*, 1005324.

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

Wiedmann, M., Kurzchalia, T.V., Hartmann, E., and Rapoport, T.A. (1987). A signal sequence receptor in the endoplasmic reticulum membrane. Nature *328*, 830–833.

Wiedmann, M., Goerlich, D., Hartmann, E., Kurzchalia, T.V., and Rapoport, T.A. (1989). Photocrosslinking demonstrates proximity of a 34 kDa membrane protein to different portions of preprolactin during translocation through the endoplasmic reticulum. FEBS Lett. *257*, 263–268.

Wild, R., Kowal, J., Eyring, J., Ngwa, E.M., Aebi, M., and Locher, K.P. (2018). Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic *N*-glycosylation. Science.

Wilson, C.M., Kraft, C., Duggan, C., Ismail, N., Crawshaw, S.G., and High, S. (2005). Ribophorin I Associates with a Subset of Membrane Proteins after Their Integration at the Sec61 Translocon. J. Biol. Chem. *280*, 4195–4206.

Wirth, A., Jung, M., Bies, C., Frien, M., Tyedmers, J., Zimmermann, R., and Wagner, R. (2003). The Sec61p Complex Is a Dynamic Precursor Activated Channel. Mol. Cell *12*, 261–268.

Wittke, S., Dünnwald, M., and Johnsson, N. (2000). Sec62p, A Component of the Endoplasmic Reticulum Protein Translocation Machinery, Contains Multiple Binding Sites for the Sec-Complex. Mol. Biol. Cell *11*, 3859–3871.

Wu, Z.C., de Keyzer, J., Berrelkamp-Lahpor, G.A., and and Driessen, A.J. (2013). Interaction of Streptococcus mutans YidC1 and YidC2 with translating and nontranslating ribosomes. J Bacteriol *195*, 4545–4551.

Xin, B., Puffenberger, E.G., Turben, S., Tan, H., Zhou, A., and Wang, H. (2010). Homozygous frameshift mutation in TMCO1 causes a syndrome with craniofacial dysmorphism, skeletal anomalies, and mental retardation. Proc. Natl. Acad. Sci. *107*, 258–263.

Xing, S., Mehlhorn, D.G., Wallmeroth, N., Asseck, L.Y., Kar, R., Voss, A., Denninger, P., Schmidt, V.A.F., Schwarzländer, M., Stierhof, Y.-D., et al. (2017). Loss of GET pathway orthologs in Arabidopsis thaliana causes root hair growth defects and affects SNARE abundance. Proc. Natl. Acad. Sci. *114*, E1544–E1553.

Yamagata, A., Mimura, H., Sato, Y., Yamashita, M., Yoshikawa, A., and Fukai, S. (2010). Structural insight into the membrane insertion of tail-anchored proteins by Get3. Genes Cells *15*, 29–41.

Yamaguchi, A., Hori, O., Stern, D.M., Hartmann, E., Ogawa, S., and Tohyama, M. (1999). Stress-Associated Endoplasmic Reticulum Protein 1 (Serp1)/Ribosome-Associated Membrane Protein 4 (Ramp4) Stabilizes Membrane Proteins during Stress and Facilitates Subsequent Glycosylation. J. Cell Biol. *147*, 1195–1204.

Yamamoto, Y., and Sakisaka, T. (2012). Molecular Machinery for Insertion of Tail-Anchored Membrane Proteins into the Endoplasmic Reticulum Membrane in Mammalian Cells. Mol. Cell *48*, 387–397.

Yamamoto, S., Yamazaki, T., Komazaki, S., Yamashita, T., Osaki, M., Matsubayashi, M., Kidoya, H., Takakura, N., Yamazaki, D., and Kakizawa, S. (2014). Contribution of calumin to embryogenesis through participation in the endoplasmic reticulum-associated degradation activity. Dev. Biol. *393*, 33–43.

Yan, A., and Lennarz, W.J. (2005). Two oligosaccharyl transferase complexes exist in yeast and associate with two different translocons. Glycobiology *15*, 1407–1415.

Yan, Q., and Lennarz, W.J. (2002). Studies on the Function of Oligosaccharyl Transferase Subunits Stt3p Is Directly Involved In The Glycosylation Process. J. Biol. Chem. *277*, 47692–47700.

Yen, M.-R., Harley, K.T., Tseng, Y.-H., and Saier, M.H. (2001). Phylogenetic and structural analyses of the oxa1 family of protein translocases. FEMS Microbiol. Lett. *204*, 223–231.

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 F1F0ATP Synthase and SecE of the SecYEG Translocase. Biochemistry (Mosc.) *42*, 10537–10544.

Young, B.P., Craven, R.A., Reid, P.J., Willer, M., and Stirling, C.J. (2001). Sec63p and Kar2p are required for the translocation of SRP-dependent precursors into the yeast endoplasmic reticulum in vivo. EMBO J. *20*, 262–271.

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

Zacchi, L.F., and Schulz, B.L. (2016). SWATH-MS Glycoproteomics Reveals Consequences of Defects in the Glycosylation Machinery. Mol. Cell. Proteomics *15*, 2435–2447.

Zafar, S., Nasir, A., and Bokhari, H. (2011). Computational analysis reveals abundance of potential glycoproteins in Archaea, Bacteria and Eukarya. Bioinformation *6*, 352–355.

Zalisko, B.E., Chan, C., Denic, V., Rock, R.S., and Keenan, R.J. (2017). Tail-Anchored Protein Insertion by a Single Get1/2 Heterodimer. Cell Rep. *20*, 2287–2293.

Zaremba-Niedzwiedzka, K., Caceres, E.F., Saw, J.H., Bäckström, D., Juzokaite, L., Vancaester, E., Seitz, K.W., Anantharaman, K., Starnawski, P., Kjeldsen, K.U., et al. (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. Nature *541*, 353–358.

Zhang, M., Yamazaki, T., Yazawa, M., Treves, S., Nishi, M., Murai, M., Shibata, E., Zorzato, F., and Takeshima, H. (2007). Calumin, a novel Ca2+-binding transmembrane protein on the endoplasmic reticulum. Cell Calcium *42*, 83–90.

Zhang, Y., Schäffer, T., Wölfle, T., Fitzke, E., Thiel, G., and Rospert, S. (2016). Cotranslational Intersection between the SRP and GET Targeting Pathways to the Endoplasmic Reticulum of Saccharomyces cerevisiae. Mol. Cell. Biol. *36*, 2374–2383.

Zhang, Y.-J., Tian, H.-F., and Wen, J.-F. (2009). The evolution of YidC/Oxa/Alb3 family in the three domains of life: a phylogenomic analysis. BMC Evol. Biol. *9*, 137.

Zhu, L., Klenner, C., Kuhn, A., and Dalbey, R.E. (2012). Both YidC and SecYEG Are Required for Translocation of the Periplasmic Loops 1 and 2 of the Multispanning Membrane Protein TatC. J. Mol. Biol. *424*, 354–367.

Zhu, L., Kaback, H.R., and Dalbey, R.E. (2013). YidC Protein, a Molecular Chaperone for LacY Protein Folding via the SecYEG Protein Machinery. J. Biol. Chem. *288*, 28180–28194.