THE UNIVERSITY OF CHICAGO

STRUCTURAL AND FUNCTIONAL CHARACTERIZATION OF NOVEL EUKARYOTIC 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 BIOCHEMISTRY AND MOLECULAR BIOPHYSICS

BY

PHILIP THOMAS MCGILVRAY

CHICAGO, ILLINOIS

AUGUST 2019

To my grandfather Paul, who made sure I memorized my times tables and told me there's no such word as can't.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I think one of the most important and humbling lessons I've wrestled with in the last five years is that no one is an island. It is unfair to discuss any single project as though it were solely the efforts of one person, because in almost all cases projects can only be brought to completion by the efforts of a team. Perhaps this is such an important lesson because it is so broadly true in life as well. I am still trying very hard to learn this lesson.

Firstly, I would like to thank my mentor Dr. Robert J. Keenan for his patience, expertise, and bravery. This project is the definition of hypothesis-driven science: taking an interesting and complex evolutionary observation and testing it to describe a previously unappreciated fundamental biological process. It is the kind of challenging, impactful project that one always hopes to be a part of, and it has simply been an honor to be involved. What's more, it is to Bob's credit that he allowed me to use such a diverse set of techniques to answer our questions. No one at our university was doing single particle cryo-electron microscopy when we began, and yet he still trusted me to dive headlong into this revolutionary technique.

Next, I would like to thank the wonderful people I have gotten the chance to work with at UChicago. All the members of the Keenan lab, past and present, have played an important role in supporting me through hard times, and have lent their expertise to difficult projects. I would especially like to thank S. Andrei Anghel, my primary collaborator throughout the past five years. Andrei was a continuous source of valuable discourse and bright ideas; I doubt that these projects could have been completed without him.

A special thank you needs to be given to all the professional scientists who supported this research technically. Dr. Thomas Edwards and Dr. Ulrich Baxa at the NCI's National Cryo-EM Facility, Dr. Joe Austin and Dr. Tera Lavoie at the University of Chicago Electron Microscopy

Core, and Jonathan Remis from the Northwestern University Structural Biology facility were instrumental to the acquisition and analysis of all cryo-EM data. They taught me basically everything I know about cryo-EM. I would also like to thank all the support scientists and staff at the UChicago Resource Computing Center, especially Dr. Hossein Pourreza. This project has demanded more intensive computational resources than I had ever initially envisioned and having the technical expertise of the Resource Computing Center to call on was integral to the success of the project. Additionally, I would like to thank Dr. Mike Trnka at UCSF for his expansive expertise in mass spectrometry and for being an extremely reliable collaborator. It is a fact that without these brilliant people any scientific inferences developed from these works would remain unknown.

I would like to extend a special thank you to my friends and family. My wife Melissa was especially important to the success of my work, providing constant support and encouragement; never letting me give up, never letting me get down. I hope to be more like her when I grow up. Thank you to my cohort - Chris, Katherine, Alyssa, and Kevin - for being the best of friends and making all of the good parts of grad school happen.

Abstract

All cells possess hydrophobic plasma membranes which separate the cellular cytosol from the external environment. Anchored within the membrane are specialized proteins which perform a multitude of important tasks. As these membrane proteins have evolved to reside and function within the hydrophobic cellular membrane, the cell must complete the challenging task of trafficking them from the aqueous cytosol and inserting them into the membrane. If membrane proteins are mis-inserted, they may aggregate and become toxic to the cell. The difficulty of this problem is magnified in eukaryotes, which possess additional internal membranes as part of their endomembrane organellar system. The process by which membrane proteins are synthesized and inserted into the membrane is termed membrane protein biogenesis. Many proteins in eukaryotes are inserted co-translationally, or as they are synthesized, by the protein-conducting channel of the endoplasmic reticulum. However, it is becoming more and more evident that previously unappreciated factors play important roles in membrane protein trafficking, insertion, and quality control. Here, I present research elucidating the structure and function of a number of membrane protein biogenesis factors found in eukaryotes. Specifically, I present evidence that the poorly studied protein TMCO1 possesses functional and structural features similar to those of the better characterized membrane protein insertases it is evolutionarily related to. Additionally, we see that TMCO1 associates with NOMO, Nicalin, CCDC47, and TMEM147; pointing to a role in membrane protein biogenesis for these uncharacterized proteins. Finally, we show that Msp1, a mitochondrial protein, is a functional hexamer capable of directly extracting mislocalized tail anchored membrane proteins from the mitochondrial outer membrane; providing further evidence for its function in membrane protein quality control.

1 Introduction and Background

1.1 A General Introduction to Cellular Membranes and Membrane Proteins

Cells are the fundamental building block of all life as we know it. Importantly, all cells are bound by a membrane which divides the interior cytoplasm from the external chemical milieu. This membrane possesses two fundamental properties; it is able to keep unwanted chemicals from the environment out of the cell, while allowing those beneficial to the cell to be absorbed and processed.

The membrane is comprised of two major components; lipids and membrane proteins. Lipids, amphipathic molecules comprised of a hydrophilic head group and hydrophobic aliphatic tails, provide the gross structure of the membrane[1,2]. These lipids are organized in a bilayer such that their hydrophilic head groups face the cytosol and extracellular space, while their aliphatic tails form a hydrophobic core in the middle of the bilayer[1,2]. This hydrophobic core prevents hydrophilic molecules from crossing the membrane, either from the extra cellular space to cytosol or vice versa.

Membrane proteins comprise the second major component of the membrane. These proteins comprise approximately 30% of the human genome and perform critical functions for the cell including ion/nutrient transport, signal transduction, and enzymatic catalysis[3,4]. Membrane proteins come in two types, peripheral and integral membrane proteins. Peripheral membrane proteins are soluble proteins which bind to the membrane either through electrostatic interactions with lipid head groups or via post-translationally added chemical anchors which extend into the hydrophobic layer of the membrane[5,6]. Integral membrane proteins, or transmembrane proteins, are proteins which have specifically evolved to function directly within the hydrophobic core of the cell membrane[3,4]. These proteins possess specific hydrophobic

regions known as transmembrane domains (TMDs) which allow them to interact favorable with the hydrophobic regions of lipids and other integral membrane proteins in the bilayer. Transmembrane domains are usually comprised of hydrophobic alpha helices which span the membrane and anchor the membrane proteins. Membrane proteins may have a single alpha helical TMD, but frequently have more with some of the largest having more than 20[3,4]. The transmembrane domain may also made by a beta-barrel motif; however, this is less common being primarily found in the bacterial and the mitochondrial outer membrane[7,8].

1.2 Membrane Protein Biogenesis

Like all proteins, integral membrane proteins are synthesized by ribosomes in the cytosol. However, in order to fold and function properly they must be chaperoned to their target membrane and correctly inserted. If they are not, then the hydrophobic regions of integral membrane proteins may aggregate and become toxic to the cell[9,10]. As a whole we use the term membrane protein biogenesis to describe the cellular processes which manage membrane protein synthesis, chaperoning, insertion, and quality control.

While prokaryotes only possess the universal plasma membrane, eukaryotes have additional membranes that compartmentalize their various sub-cellular organelles, known as the endomembrane system. The membranes of each organelle are comprised of unique lipid compositions and specific membrane proteins which aid in the function of that organelle[11]. Thus, eukaryotic membrane proteins must be targeted to the correct membrane as well as properly synthesized. In eukaryotes, many integral membrane proteins are first inserted into the membrane of the endoplasmic reticulum (ER)[12]. From there, they are trafficked to their final target membrane. An exception to this rule is the mitochondria, where many membrane proteins

are synthesized in the cytosol and specifically chaperoned to the mitochondrial membrane in a pre-protein form[13,14].

1.3 Co-translational Membrane Protein Biogenesis and Protein Secretion

The best understood membrane protein insertion method is the co-translational pathway where integral membrane proteins are inserted into the membrane as they are being synthesized. Nascent membrane proteins are recognized by a soluble targeting factor while they are being translated by the ribosome. This targeting factor guides the translating ribosomes-nascent protein complex to its target membrane (Fig 1). At the target membrane the ribosomes is recognized and docks with a protein conducting channel. The central machinery for this process is conserved in all domains of life, indicating that co-translational biogenesis is a fundamental activity in all living cells.

The signal recognition particle (SRP) is a ribo-protein complex found in all cells which acts as the targeting factor for co-translational membrane protein biogenesis[15-17]. The SRP recognizes and binds to a stretch of nascent, hydrophobic peptides known as a signal sequence as they leave the exit channel, shielding them from the cytosol[15,16,18, 19]. The role of signal sequence is sometimes fulfilled by the first TMD of multi-TMD proteins[20,21]. SRP guides the translating ribosomes to their target membrane – the plasma membrane in bacteria and archaea, and the endoplasmic reticulum membrane in eukaryotes. At this membrane SRP binds a signal receptor (SR) which interacts with the protein conducting channel (PCC) on the cytosolic side of the membrane (Fig. 1)[22,23]. Binding of SR weakens the interaction between SRP and the ribosome-nascent chain complex (RNC), promoting the release of the RNC and hand off to the PCC.

To properly facilitate membrane protein insertion the PCC must complete two main tasks[24-27]. First, it must allow the passage of soluble protein domains through the membrane and into the ER lumen. Second, it must allow the entry of hydrophobic transmembrane domains into the membrane. These tasks are completed by the conserved heterotrimeric Sec complex – comprised of SecYEG in bacteria, SecYEβ in archaea, and Sec61αβγ in eukaryotes[27]. The central pore is made by SecY (in prokaryotes) or Sec61α (in eukaryotes), a 10 TMD integral membrane protein, which possesses an aqueous channel occluded by a hydrophobic ring of amino acids, a plug domain, and a "lateral gate" which opens into the membrane[27-31]. The function of SecE, SecG, Sec61β, and Sec61γ are less well understood, but their inclusion in the complex increases translation efficiency through an unknown mechanism[32,33]. Once a ribosome translating a new integral membrane protein has been delivered to its target membrane by SRP, it

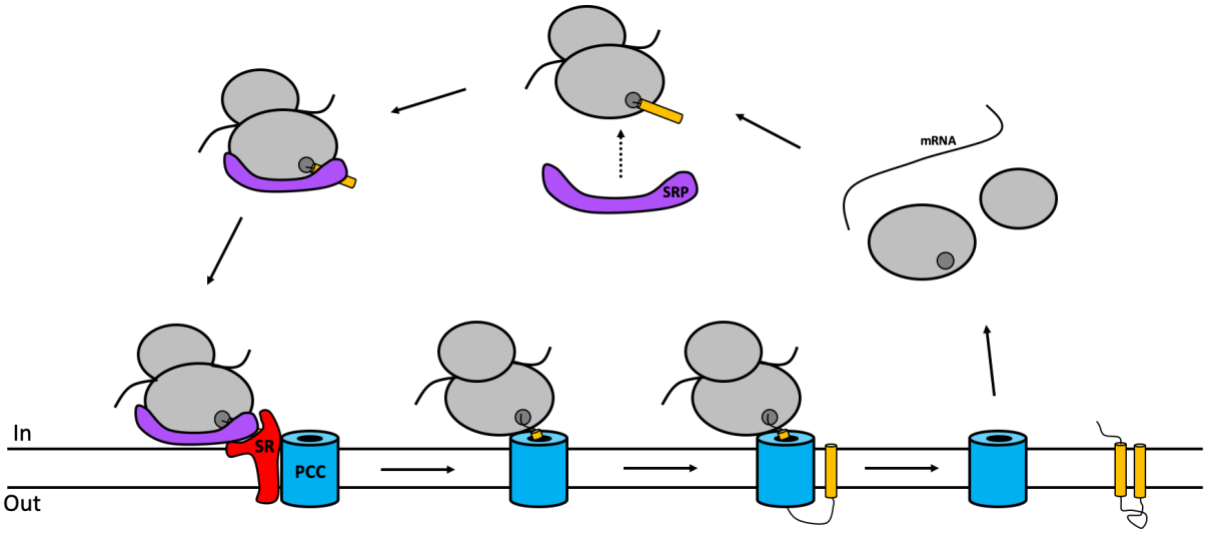


Figure 1. A Basic Model of Co-translational Membrane Protein Biogenesis. Ribosomes (gray) in the cytosol recognize begin translating mRNAs. SRP recognizes ribosomes translating secreted or integral membrane proteins via an exposed N-terminal hydrophobic sequence (signal sequence). SRP guides the ribosome to its target membrane where it docks with the SRP receptor (SR) and hands off the translating ribosome to the protein conducting channel (PCC). The PCC facilitates both entry of transmembrane domains into the membrane and passage of soluble domains across the membrane. Upon completion the ribosomal subunits split, releasing the mRNA, and are recycled for future rounds of translation.

docks with the Sec complex, primarily through interactions with SecY or Sec61α[25,28,34,35]. This docking causes the channel to enter a "primed" state, where interactions with the ribosome causes the lateral gate to open very slightly while the plug remains in the channel. The nascent hydrophobic peptide binds with the PCC and is inserted into the channel. This opens the channel by displacing a ring of hydrophobic residues and a plug helix which blocks entrance to the ER lumen[28,30,36]. The hydrophobic peptide from the nascent membrane protein will also interact with the helices around the channels lateral gate (helices 2,3,7, 8), displacing them and directly interacting with the hydrophobic region of the lipid bilayer[36]. As the membrane protein is translated by the bound ribosome, its soluble domains will pass through the channel while its transmembrane domains will partition into the membrane through the lateral gate. As mutipass membrane proteins, membrane proteins containing multiple TMDs, are translated, each TMD passes through the lateral gate and enter the membrane in this fashion[36-38].

1.4 The Translocon

While this model of co-translational insertion has proven valuable for understanding the basic process of co-translational membrane protein insertion, the reality is more complex. Membrane proteins are extremely diverse, possessing many disparate biophysical properties. In order to build these proteins into functioning entities the cell must cater to these different requirements during biogenesis. Thus, multiple other proteins termed translocon accessory factors interact with the central Sec complex to modulate its activity and facilitate the insertion requirements of different newly synthesized integral membrane proteins[39-42]. Together the Sec complex and its associated accessory factors make up the translocon.

1.4.1 Bacterial Translocon Accessory Factors

YidC is a 5 or 6 TMD integral membrane protein and translocon accessory factor found in the bacterial plasma membrane[43-45]. In the holo-translocon, YidC interacts directly with the lateral gate of SecY and chaperones the entry of TMDs from large, polytopic membrane proteins into the membrane[46-49]. Additionally, YidC can interact with ribosome independently of the Sec complex[50-52]. In this case, YidC interacts with the ribosomal exit channel and facilitates the entry of small, topologically simple membrane proteins into the plasma membrane. YidC's independent insertion activity is required for the biogenesis of multiple important topologically simple bacterial membrane proteins[43,53]. YidC is closely related to Oxa1 and Alb3, integral membrane proteins found in the mitochondrial inner membrane and the chloroplast thylakoid membrane respectively, and is theorized to share a common evolutionary origin[54,55].

SecD and SecF are each 6 TMD integral membrane proteins which bind to the Sec complex opposite the lateral gate[56-57]. SecD and F are important for efficient soluble protein secretion rather than membrane protein integration. SecDF stimulates general protein secretion via SecYEG, and mutants lacking SecDF have a cold-sensitive phenotype, and temperature sensitive secretion defects[57-59]. SecDF can conduct protons across the membrane and has been proposed to use the trans-membrane proton motive force to help facilitate protein passage across the membrane[56,59].

YajC is perhaps the most poorly characterized component of the holo-translocon. It is part of a translocon subcomplex with SecDF and interacts with SecYEG opposite the lateral gate, but its function remains largely unknown[40].

1.4.2 Eukaryotic Translocon Accessory Factors

The Translocon Associated Protein (TRAP) complex is a hetero-tetrameric membrane protein complex which associates stoichiometrically with the Sec61 complex in the presence and absence of translating ribosomes[25,60-62]. TRAP has been shown to interact directly with translating nascent chains and stimulate the insertion of membrane proteins in a signal sequence specific manner[63,64]. It is theorized that TRAP may stimulate the insertion of secreted proteins which engage poorly with Sec61 by properly orienting the nascent chain in Sec61 channel, or by modulating the structure or activity of the Sec61 complex to better accept the nascent chains[64].

The Oligosaccharyltransferase (OST) Complex is a crucial multi-protein complex highly conserved through-out eukaryotes and essential in most organisms. OST catalyzes the addition of N-linked glycans – one of the most abundant post-translational modifications - to secreted and membrane proteins containing in the consensus N-X-T/S sequence, where X can be any amino acid except proline[65-67]. Many secreted and membrane proteins are glycosylated, and defects in glycosylation have been associated with human diseases, and modulated protein stability[68-70].

The catalytic component of OST is referred to as STT3 in yeast. However, metazoans have two homologues of STT3, STT3A and STT3B[67,71]. These two homologues define two different types of OST complex[72-74]. The STT3A containing complex interacts directly with the Sec61 complex and ribosomes, allowing it to glycosylate proteins as they are translated. The STT3B containing complex does not interact with these biosynthetic machines and glycosylates substrates after they have been translated, or post-translationally.

Calnexin is a single pass membrane protein which associates directly with the Sec61 complex and acts as a chaperone for glycolsyated proteins in the ER lumen[75,76]. Calnexin recognizes glycosylated proteins and chaperones them during maturation in the ER lumen.

Proteins released from Calnexin have two fates. If the protein is correctly folded, its N-linked glycan is modified, preventing it from being re-bound by calnexin[77,78]. If it is not correctly folded, additional glucosyl residues will be added back to the misfolded proteins' glycan, causing re-association with Calnexin. Glycoprotein misfolding leads to multiple rounds of Calnexin reassociation and glycan addition, and eventual retro-translocation from the ER.

The Translocating-chain Associated Membrane protein (TRAM) is a single, multipass membrane protein which has been crosslinked to the signal sequences and TMDs of translating proteins[79-81]. TRAM has also been shown to stimulate the insertion activity of certain membrane proteins and soluble secreted proteins, however, its exact function and interactions with other biosynthetic machinery remain poorly understood.

While this list is not exhaustive, as having such a list would not impart additional pertinent information to the reader, it does present examples of how the major known translocon accessory factors facilitate membrane protein biogenesis and modulate the activity of the translocon.

1.5 Post-translational Membrane Protein Biogenesis

While the majority of membrane proteins are inserted co-translationally, a number of proteins possess biophysical qualities which preclude their insertion during translation. Membrane proteins inserted co-translationally by the SRP-Sec complex pathway are canonically thought to possess hydrophobic TMDs which are presented by the ribosome early on, and throughout synthesis. Some membrane proteins are too small to present a hydrophobic signal sequence or TMD during synthesis. Others have TMDs at their extreme C-termini, precluding their recognition until they leave the exit channel entirely. These proteins are inserted after they have

been fully translated, or post-translationally. While these proteins have quite different biophysical properties from co-translational inserted membrane proteins, the majority are still trafficked to the ER membrane for insertion.

1.5.1 Post-translational SecYEG Dependent Protein Translocation Across the Bacterial Plasma Membrane

In bacteria the Sec complex has been shown to function in both co- and post-translational protein secretion. Post-translationally secreted proteins are recognized by the chaperone SecB in the cytosol after translation[82,83]. SecB binds to hydrophobic stretches of amino acids exposed during protein translation. After translation SecB traffics the protein to SecA, a dimeric ATPase motor, which binds directly to the translocon on the plasma membrane[84-86]. SecA uses the energy of ATP hydrolysis to feed the delivered protein through the SecYEG channel and into the bacterial periplasm. Generally, proteins which require SecA for secretion are soluble proteins which reside in the bacterial periplasm, not integral membrane proteins.

1.5.2 Post-translational Membrane Protein Biogenesis at the Endoplasmic Reticulum

The best studied group of post translationally inserted proteins are the so-called tail-anchored (TA) proteins membrane proteins, a class of about 700 proteins[87]. Tail anchored proteins possess a C-terminal alpha helical TMD which anchors them in the membrane. This leaves their N-terminal domains to perform important biochemical tasks or make important interactions[88-90].

The pathway for TA protein insertion, The Guided Entry of Tail-anchored proteins (GET) pathway, has only recently been elucidated and has been best characterized in yeast[91-93].

Here, ribosomes translating TA proteins had off the substrate protein to a pre-targeting complex comprised of SGT2, Get4, and Get5[94-96]. This complex passes the TA protein to Get3, a soluble chaperoning factor which binds directly to the TA protein TMD and chaperones it to the ER. Membrane[97, 98]. At the ER membrane Get3 is captured by a membrane protein complex comprised of Get1 and Get2[99,100]. The long cytosolic arms of Get2 bind Get3 and tether it to membrane where it further interacts with the coiled coil of Get1[94,99,100]. This interaction seems to trigger a conformational rearrangement in Get3, which causes it to release the TA protein. Once released the insertion of the TA protein into the membrane is directly guided by Get1 and Get2[101].

All of the components found in yeast have homologues in metazoans; Sgt2 - SgtA, Get1 - WRB, Get2 - CAML, Get3 - TRC40, Get4 – TRC35, Get5 - UBL4A[94,102,103]. However, the metazoan pre-targeting complex contains an additional component BAG6. BAG6 interacts with TRC35, Sgt2, and UBL4A, but also with ribosomes, chaperones, and TA proteins directly. BAG6 appears to play the role of organizing and triage center. While TA proteins handed off to TRC40 from BAG6 continue on to be inserted in the ER membrane, TA proteins which continue to interact with BAG6 are marked for degradation, and TA protein handed to chaperones continue on an unidentified path, possibly to the mitochondria[104,105].

There is also evidence for certain proteins being inserted into the ER membrane by a translocon comprised of the Sec61 complex, Sec62, and Sec63. Sec62 and Sec63 are translocon accessory factors associated with ribosome free Sec61 complexes[106-108]. Sec62 and 63 have been well studied in yeast, where they interact with not only the Sec61 complex and the yeast specific proteins Sec71 and Sec72. Here, chaperone bound soluble, secreted proteins are recruited to the membrane by Sec72 and translocated across the membrane through the Sec61 complex[109]. New

studies are also shedding light on the ability of this complex translocate the soluble domains and insert the TMDs of post-translationally inserted membrane proteins[41,110].

1.6 Membrane Protein Biogenesis in Mitochondria

Mitochondria are endosymbiotic organelles found in almost all eukaryotes. It is theorized that these organelles were previously autonomous prokaryotic organisms (likely α-proteobacteria) which were engulphed by another cell, but somehow evaded destruction[111]. These organelles possess a special double membrane, leading to interesting problems in both membrane protein targeting and insertion. Additionally, mitochondria maintain a small genome which, in humans, encodes 13 proteins, 22 mitochondria specific tRNA, and the large and small subunit RNAs of the mitochondrial ribosome.

Note, this introduction on mitochondrial membrane protein biogenesis will only focus on the biogenesis of proteins containing alpha helical TMDs, as the biogenesis of beta-barrel containing integral membrane proteins is beyond the scope of this thesis.

1.6.1 Protein Trafficking Through the Mitochondrial Membrane System

Membrane proteins destined for the mitochondria outer membrane are translated in the cytosol but not inserted into the ER membrane like other membrane proteins of the endomembrane system. Broadly, mitochondrial membrane proteins are recognized as precursor proteins in the cytosol and trafficked to the mitochondria post-translationally[112,113]. A moderately hydrophobic presequence of 15-55 residues preceded by an N-terminal patch of basic amino acids is the most common targeting signal for mitochondrial proteins[113]. This presequence facilitates chaperoning by HSP70 to the mitochondria, as well as recognition by the outer mitochondrial insertion

machinery[114,115]. Mitochondrial signal anchored or multi-pass integral membrane proteins generally possess presequences closer to their N-terminus, while TA proteins contain this information within their C terminal TMDs[113,116].

The Translocase of the Outer Membrane (TOM) complex facilitates the passage of mitochondrial proteins from the cytosol through the outer mitochondrial membrane (OMM) and into the intermembrane space (IMS). TOM is comprised of three core components, TOM40,20, and 70, and multiple auxiliary components[113,117]. TOM40 is a beta barrel protein which acts as the central channel through which mitochondrial preproteins pass. TOM70 and TOM20 are integral membrane proteins which interact directly with TOM40 and recognize chaperones and preprotein presequences[117,118]. These interactions may act as a kind of quality control, only allowing those preproteins with the correct targeting sequence through the OMM. Mitochondrial preproteins are imported via TOM in an unfolded state, a fact likely facilitated by the chaperone-like activity of the insertion machinery.

Transportation of proteins from the OMM to the inner mitochondrial membrane (IMM) occurs concomitantly with passage through TOM. Here, unfolded preproteins are recognized by the Translocase of the Inner Membrane (TIM) complex, a second beta-barrel protein which acts as a channel through the IMM[113,119]. The core of the TIM complex is comprised of TIM23,17, and 50[113]. TIM23 is the central channel for protein transport[120]. TIM17 and TIM50 both have been proposed to regulate the opening of TIM23 to prevent the leakage of ions from the mitochondrial inner compartment, or mitochondrial matrix[121-123]. Additionally, TIM50 binds preproteins as they pass through TOM, and hands them off to TIM23 for translocation across the IMM. The membrane potential and interactions with different charged patches on TOM and TIM complex proteins provides the driving force for preprotein translocation across membranes[124,125].

The translocation of large hydrophilic soluble proteins, or domains of membrane proteins, destined for the mitochondrial matrix is assisted by the presequence translocase-associated import motor (PAM)[113]. PAM is a complex of integral membrane proteins and the mitochondiral HSP70 (mtHSP70) which associates directly with TIM[126,127]. ATP hydrolysis by mtHSP70 causes complex conformational changes which drive the preproteins translocation across the IMM[128,129]. The presequences of both soluble and integral membrane proteins which pass into the mitochondrial matrix are cleaved off by the mitochondrial processing peptidase (MPP)[130,131].

1.6.2 TMD Insertion in the Mitochondrial Outer and Inner Membrane

While the central tenants of protein translocation across the mitochondrial double membrane are well established, the mechanism of TMD insertion for the outer and inner mitochondrial membranes is poorly understood.

Signal anchored membrane proteins of the OMM rely on Mim1 for proper integration into the membrane[132,133]. Both TOM20 and TOM70 interact directly with Mim1 during biogenesis, however the mechanism by which Mim1 facilitates membrane integration is unclear[134]. The mechanism of TA protein TMD integration is even more unclear. It has been proposed that TA protein integration is influenced by the lipid composition of the membrane they are targeted too, and that some may be able to insert spontaneously[135,136].

Membrane proteins in the mitochondrial matrix are inserted into the IMM in a number of ways. The transmembrane domains of proteins translocated across both the IMM and the OMM can enter the membrane via a TIM dependent lateral sorting mechanism[137]. Alternatively, TMDs from both nuclear encoded inner membrane proteins, and mitochondrially encoded membrane

proteins translated by mitochondrial ribosomes can be inserted by Oxa1. Oxa1 is a five TMD IMM protein evolutionarily related to the bacterial plasma membrane and chloroplast thylakoid membrane protein biogenesis factors YidC and Alb3[138]. Oxa1 interacts directly with mitochondrial ribosomes to facilitate the co-translational insertion of mitochondrially encoded membrane proteins into the IMM and may insert some nuclear encoded proteins independently[139,140].

1.7 Identification of a YidC Homologues in the Archaea Plasma Membrane

As previously mentioned, the bacterial YidC, the mitochondrial IMM protein Oxa1, and the chloroplast protein Alb3 comprise an evolutionarily related family of proteins found in both bacteria and eukaryotes[138]. These proteins share common structural features, including 5 or 6 TMDs and a soluble coiled coil, and all three play a role in membrane protein biogenesis. Intriguingly, proteins with low sequence similarity to YidC/Oxa1/Alb3 had been identified in archaea[141]. These archaeal proteins were uncharacterized and shared low sequence similarity, but had been annotated as domain of unknown function 106 (DUF106) indicating shared predicted structural features.

Characterization of one archaeal DUF106 protein from Methanocaldococcus jannaschii, MJ0480, revealed that the archaeal DUF106 proteins are structurally and functionally related to YidC/Oxa1/Alb3[142]. Despite being extremely sequentially divergent (~10-15% similarity) and smaller than the YidC/Oxa1/Alb3 proteins, MJ0480 was predicted to share similar structural features. Indeed, the crystal structure of MJ0480 shows a conserved core three TMD bundle which overlaps well with that of the Bacillus halodurans YidC2 (RMSD = 3.9Å over 105 core residues)[142]. Additionally, MJ0480 was found to bind selectively to stalled ribosomes translating

a stalled membrane protein substrate in vitro and crosslinked directly to this stalled nascent chain via residues within a hydrophilic groove. These data indicate that the archaeal DUF106 proteins may be YidC-like biogenesis factors in the archaeal membrane and highlight the fact that many biologically important biogenesis factors remain to be found.

2 Identification of Oxa1 homologs operating in the eukaryotic endoplasmic reticulum

2.1 Overview

After Borowska et al showed 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, were present in the ER membrane. We hypothesized that these proteins are all part of a superfamily of membrane protein biogenesis factors found in all domains of life, which have termed the Oxa1 superfamily.

The Oxa1 superfamily hypothesis predicts that these proteins share a common structure and function. Using evolutionary-based structural modeling and in vivo topology analysis, we showed 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 suggested that TMCO1 has an independent 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.

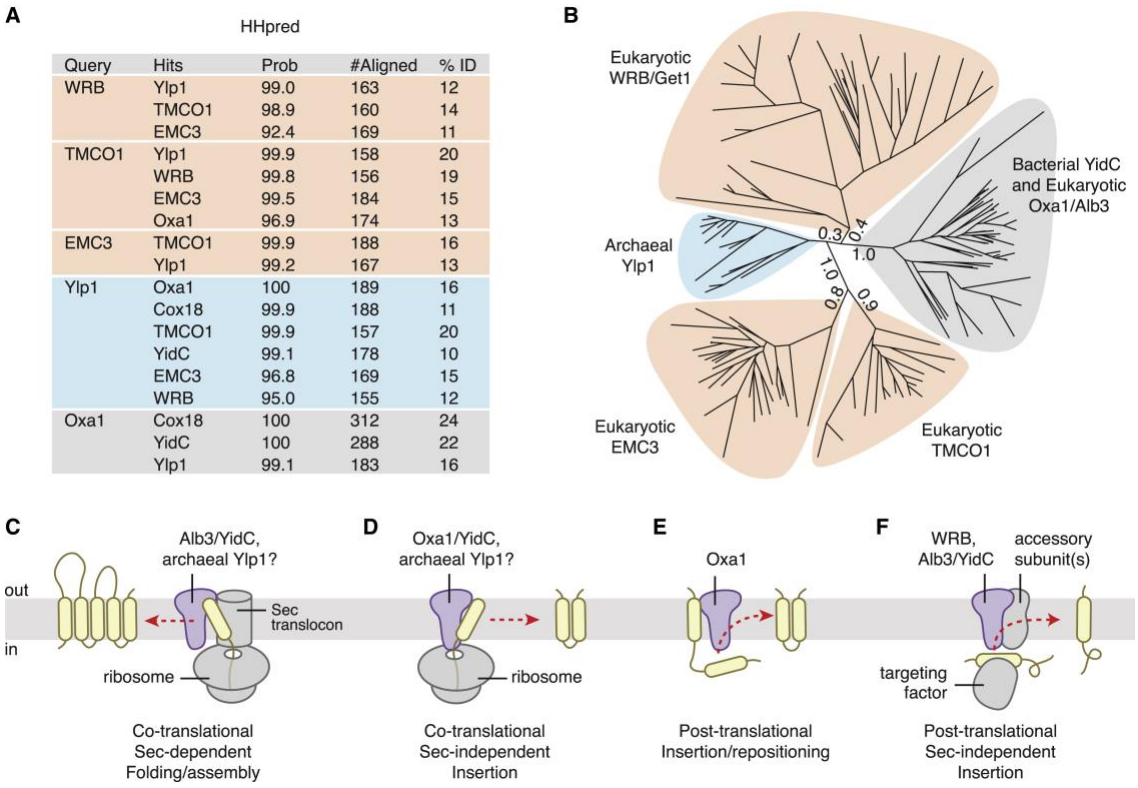
2.2 Contributions

In this work, I performed the recombinant TMCO1 purification and the in vitro ribosome binding experiments and assisted with the native purification of TMCO1-ribosome complexes. S.A.A performed the phylogenetic analysis, the glycosylation mapping and the biochemical analyses from HEK293 cells and canine microsomes. R.S.H. and R.J.K. conceived of the project. R.J.K.

performed the structural modeling. S.A.A wrote the manuscript together with R.J.K, with edits from me.

2.3 Phylogenetic and functional comparisons define the 'Oxa1 superfamily'

In bacteria, certain proteins are inserted into the plasma membrane by co- and post-translational, translocon-independent pathways mediated by YidC[143,144]. These substrates are generally small, topologically simple proteins that lack large or highly charged translocated regions[143]. YidC also functions in a translocon-dependent mode, where it facilitates the insertion, folding, and/or assembly of substrates containing multiple TMDs[145]. Homologs of YidC are present in the mitochondrial inner membrane (called Oxa1 and Cox18) and the chloroplast thylakoid membrane (Alb3 and Alb4)[146]. Like bacterial YidC, these proteins function in different contexts as insertases, chaperones, and assembly factors.



Query	Hits	Prob	#Aligned	% ID
WRB	Ylp1	99.0	163	12
	TMCO1	98.9	160	14
	EMC3	92.4	169	11
TMCO1	Ylp1	99.9	158	20
	WRB	99.8	156	19
	EMC3	99.5	184	15
	Oxa1	96.9	174	13
EMC3	TMCO1	99.9	188	16
	Ylp1	99.2	167	13
Ylp1	Oxa1	100	189	16
	Cox18	99.9	188	11
	TMCO1	99.9	157	20
	YidC	99.1	178	10
	EMC3	96.8	169	15
	WRB	95.0	155	12
Oxa1	Cox18	100	312	24
	YidC	100	288	22
	Ylp1	99.1	183	16

Figure 2. Phylogenetic and Functional Comparison Defines the Oxa1 Superfamily. (A) Identification of remote DUF106 homologs using HHpred. Eukaryotic, bacterial, and archaeal 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.

Although YidC homologs are widely conserved among bacteria and archaea, 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[147]. 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 (Fig. 2A,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[148,149].

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 (Fig 2C). 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[150,151]. Similarly, the archaeal DUF106 protein Mj0480 (henceforth called the "YidC-like protein 1" or Ylp1) binds RNCs, and can be crosslinked to a model substrate in vitro[142]. 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[103,146,152]. Finally, although its precise function remains to be defined, the EMC has been linked to ERAD and biosynthesis of multi-pass membrane proteins[153,154]. 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'.

2.4 Oxa1 superfamily members share membrane topology and 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, comparison of the crystal structures of bacterial YidC[150,151] and archaeal Ylp[142] 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 (Fig 3A).

Secondary structure and topology predictions for Get1, TMCO1 and EMC3 suggest they share this architecture (Fig 3B and 4A), 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[155]; this topology is incompatible with placement of TMCO1 into the Oxa1 superfamily.

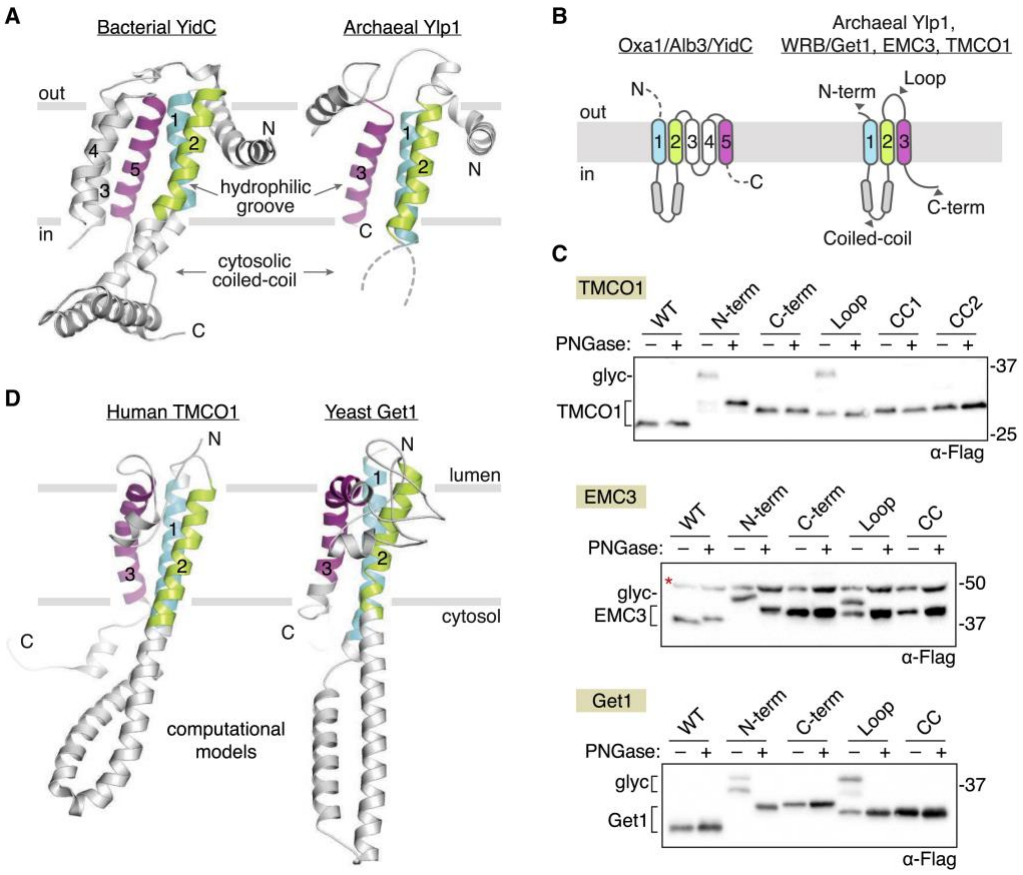


Figure 3. Oxa1 Superfamily Members Share a Conserved Membrane Topology and Core Structural Features. (A) Comparison of known structures from two clades: bacterial YidC (left; PDB: 3WO6) and archaeal Ylp1 (right; PDB: 5C8J). These proteins share a common N-out/C-in topology, a cytosolic-facing coiled 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 (gray arrowheads in B and Figure 4), 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). The resulting hybrid model is in good agreement with the covariation-based 3D model calculated for human WRB (Figure 4B).

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 (Fig 3B). 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 (Fig 3C). These data are consistent with the observation that the Get1 coiled-coil binds to the cytosolic targeting machinery[156-158], and with proteomic analyses showing that serine residues in the coiled-coil and C-terminal regions of TMCO1 are phosphorylated by cytosolic kinases[159,160]. We also performed an unbiased, 3D structure prediction of TMCO1, Get1 and EMC3 using distance restraints derived from evolutionarily coupled residue pairs[161]. 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 (Fig 3D and 4B,D). 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 4C); 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.

2.5 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[162]. Genetic variations around TMCO1 are linked to glaucoma and nonsense mutations cause a disorder associated with craniofacial dysmorphisms, skeletal anomalies, and intellectual disability[163-167].

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 complex and the ribosome (Fig 2C,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 (Fig 2C,D).

When digitonin-solubilized HEK293 membranes were fractionated on a sucrose gradient, TMCO1 and Sec61 were present in the 80S ribosome fraction (Fig 4A). 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 (Fig 6A) we observed a complex containing TMCO1, Sec61 and ribosomes (Figure 5B). Thus, TMCO1-Sec61-ribosome complexes can be isolated from cells under native conditions.

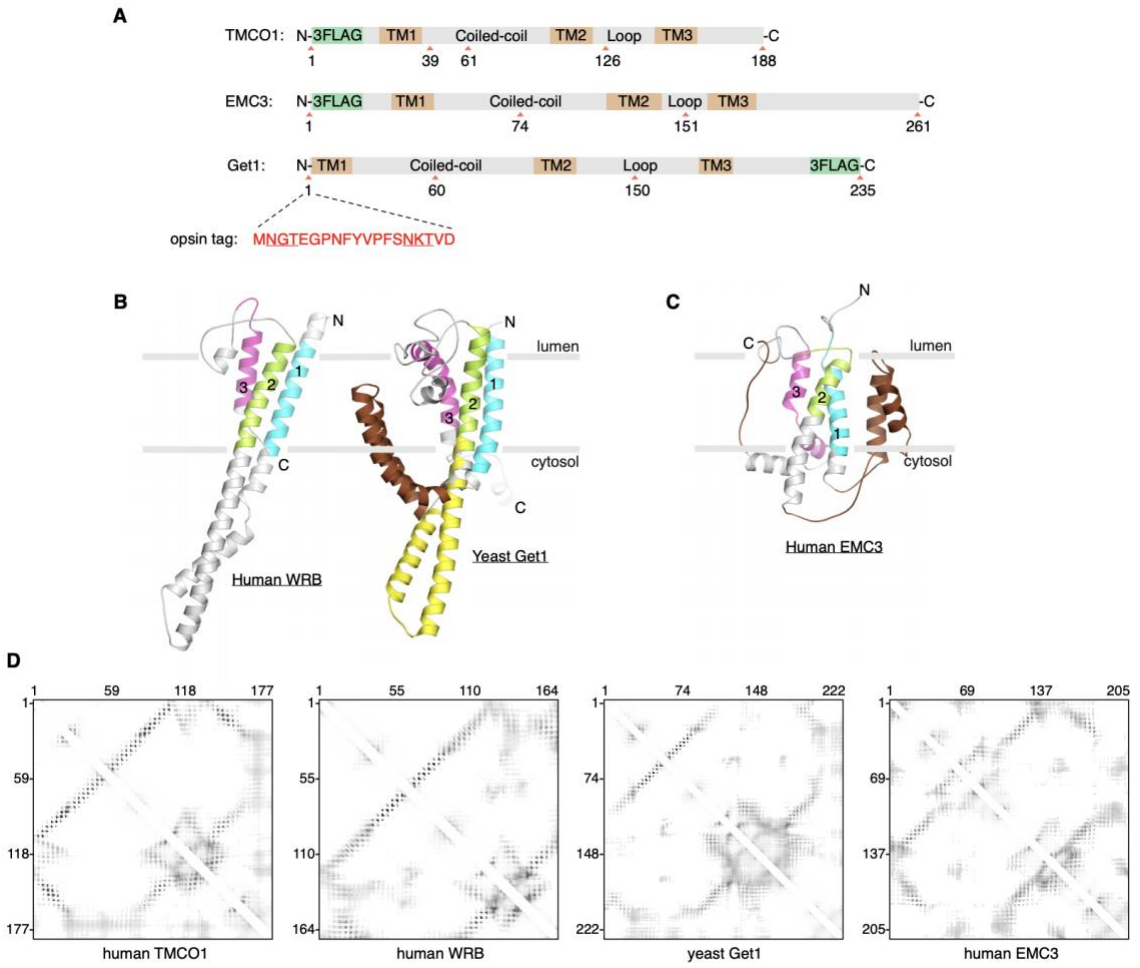


Figure 4. Additional details for the topology mapping experiments and 3D modeling. (A) Constructs used for glycosylation mapping. An opsin tag (red) containing two N-glycosylation sites (underlined) was inserted at the indicated positions of human TMCO1, human EMC3 and yeast Get1. Tag positions correspond 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 3D; 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 (see also Figure 3D). (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. (D) Heat maps of the RaptorX probabilities of two residues being in close proximity.

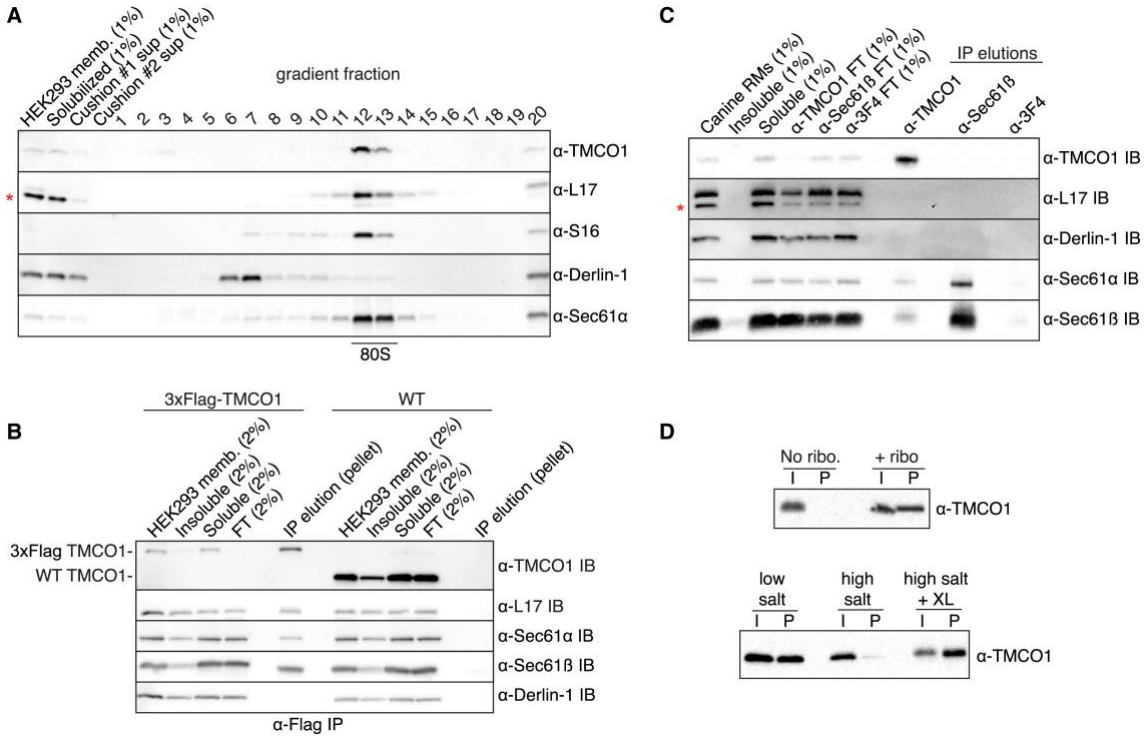


Figure 5. TMCO1 Forms a Complex with the Sec61 Translocon and RNCs. (A) HEK293 membranes were solubilized with digitonin, fractionated by sucrose cushion, separated on a high-resolution sucrose gradient, and analyzed by western blotting. TMCO1 co-fractionates with intact 80S particles and the Sec61 translocon, but not the unrelated ER membrane protein Derlin-1, 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 53 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[169,170]. 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 (Fig 5C). 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[170,171]. To test this, we incubated recombinant, purified TMCO1 (Fig 6B) with unprogrammed ribosomes from a rabbit reticulocyte lysate. After sedimentation through a sucrose cushion, we observed ribosome-dependent pelleting of TMCO1 (Fig 5D and 6C). 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 (Fig 5D and 6D). 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.

2.6 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 (Fig 2C-F). In particular, Oxa1/Alb3/YidC proteins facilitate the insertion, folding and/or assembly of a variety of membrane proteins, while the WRB/Get1 subunit of the GET pathway transmembrane complex mediates the insertion of TA membrane proteins into the ER

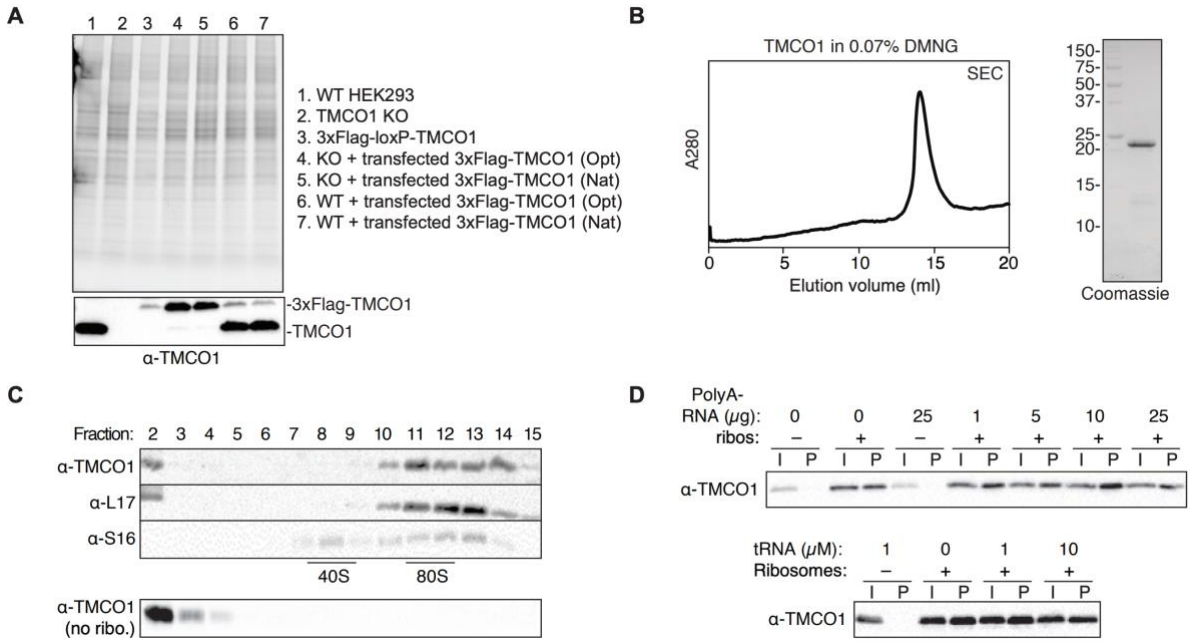


Figure 6. Additional characterization of the ribosome binding properties of TMCO1 in cells and in vitro. (A) Western blot analysis of TMCO1 expression levels in wild-type (WT) HEK293 cells, CRISPR/Cas9 generated knockout (KO) HEK293 cells, an integrated 3xFlag-tagged TMCO1 cell line and either KO or 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.

[146,103]. Similarly, the EMC3 subunit of the 'ER membrane complex' has been proposed to play a role in membrane protein quality control and biogenesis[153,154].

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 (Fig. 2C and D).

This assignment is not incompatible with the previous proposal that TMCO1 functions as a Ca^{2+} channel[155]. Indeed, other well-characterized membrane protein insertases, including the bacterial and eukaryotic Sec complexes[47, 24, 172, 173] and mitochondrial Oxa1[174], have 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[175,176] 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, TMCO1 or EMC3 is non-lethal[166,167,177-179]. Such functional redundancy would impart robustness to membrane protein biogenesis, particularly under conditions of stress[180,181]. Identifying the native substrates and molecular mechanisms underlying EMC3- and TMCO1-mediated biogenesis are important topics for future investigation.

3 A Novel Translocon Functioning During Co-translational Biogenesis of Multi-pass Membrane proteins

3.1 Overview

Previous work by Anghel et al indicated the TMCO1, a previously uncharacterized eukaryotic membrane protein, was part of a larger superfamily of suspected membrane protein biogenesis factors. What's more, TMCO1 was found to interact directly with ribosomes and the Sec61 complex. To learn more about its function in biogenesis, we sought to learn more about native TMCO1-Sec61-ribosome complexes.

Mass spectrometry of natively isolated TMCO1-contining RNC-translocon complexes revealed that TMCO1 associated with a number of unexpected membrane proteins; namely CCDC47, NOMO, TMEM147, and Nicalin. RNA sequencing of co-isolated mRNAs reveals a bias towards integral multi-pass membrane proteins. Single particle cryo-EM and crosslinking mass spectrometry show that these proteins form a loose complex at the ribosomal exit channel and indicate that biogenesis may proceed via TMD assembly within a central lipid filled cavity.

A version of this work is currently in submission as McGilvray, P.T., Anghel, S.A., Sundaram, A., Trnka, M.J., Zhong, F., Hu, H., Burlingame, A.L., and Keenan, R.J. (2019). A novel translocon functioning during co-translational biogenesis of multi-pass membrane proteins.

3.2 Contributions

In this work, I developed the protocol for all immunoprecipitation with S.A.A., established suspention HEK293 culture protocols generated all samples for cryo-EM and XLMS and carried out all EM analysis. S.A.A. analyzed all uncrosslinked mass spectrometry and, with H.H., the RNA sequencing data; R.J.K., built and analyzed the model; A.S. and F.Z. carried out the

EAAT1 expression analysis with help from P.T.M; M.J.T. carried out the cross-linking mass spectrometry, under the guidance of A.L.B; R.J.K. and P.T.M. wrote the manuscript with input from all authors.

3.3 Introduction

Human cells encode ~2,500 multi-pass membrane proteins, which play critical roles in nearly all aspects of cell physiology[3,4]. Most of these proteins are synthesized co-translationally at the endoplasmic reticulum (ER). The extreme diversity of multi-pass membrane proteins, which have different topological and biophysical requirements for insertion, folding, modification and assembly, presents a considerable challenge to the biosynthetic machinery of the cell. Biogenesis of multi-pass membrane proteins is coordinated by the 'translocon', a poorly defined and dynamic ensemble comprising the evolutionarily conserved Sec61 complex[12,22,23,27] in association with a variety of accessory proteins[39-42]. Despite its importance much of this machinery remains poorly understood, in part because translocon complexes are inherently difficult to isolate. Understanding how different types of translocon mediate multi-pass membrane protein biogenesis is a fundamental question in cell biology.

We recently identified TMCO1 as a eukaryotic member of the Oxa1 superfamily of membrane protein biogenesis factors[182]. TMCO1 is an ER-resident membrane protein, conserved in higher eukaryotes, which associates directly with ribosomes and the Sec61 complex[182]. Nonsense mutations in TMCO1 cause a developmental disorder associated with craniofacial dysmorphisms, skeletal anomalies and severe intellectual disability in mice and humans[167,183,184]. Unfortunately, the biochemical function remains unclear, as TMCO1 has been proposed to

function as a membrane protein biogenesis factor, and as a Ca2+ load activated Ca2+ channel[182,155].

Here, we use quantitative mass spectrometry, high-throughput RNA sequencing, genetic analysis and single-particle cryo-electron microscopy to identify and characterize a natively isolated human TMCO1-containing ribosome-translocon complex (RTC). We show that TMCO1 is part of a novel ~370 kDa translocon, unique to higher eukaryotes, that assembles on ribosomes to facilitate the co-translational biogenesis of ~20% of all multi-pass membrane proteins.

3.4 Identification of a TMCO1 containing translocon complex

To identify interaction partners of TMCO1 we affinity purified TMCO1-ribosome complexes from 3xFlag-TMCO1 cells for analysis by mass spectrometry (Fig. 7A). As expected, numerous ribosomal proteins, core subunits of the Sec61 complex, and TMCO1 itself were enriched relative to control cells. Most previously defined translocon accessory factors, including subunits of the OST and TRAP complexes, TRAM, Sec62/63 and the signal peptidase complex, were either absent or poorly enriched.

Three of the most enriched proteins were integral ER membrane proteins not previously reported to bind ribosomes. These include CCDC47 (calumin) and two subunits of the Nicalin-TMEM147-NOMO complex[185-187] (Fig. 7B). Immunoblotting of the pellet fraction following 3xFlag-TMCO1 affinity purification confirmed the presence of TMCO1, Nicalin, NOMO and CCDC47 in the ribosome-associated fraction (Fig 7C).

Like TMCO1, the molecular functions of CCDC47 and the Nicalin-TMEM147-NOMO complex are not well defined. CCDC47 has been proposed to function in various processes including ER-associated degradation (ERAD), and maintenance of Ca^{2+} homeostasis and

signaling in the ER[188,189]. Additionally, the Nicalin-TMEM147-NOMO complex plays a role in defining the localization and subunit stoichiometry of cell surface receptors including the nicotinic and muscarinic acetylcholine receptors, and the amiloride-sensitive sodium channel[185-187,190,191].

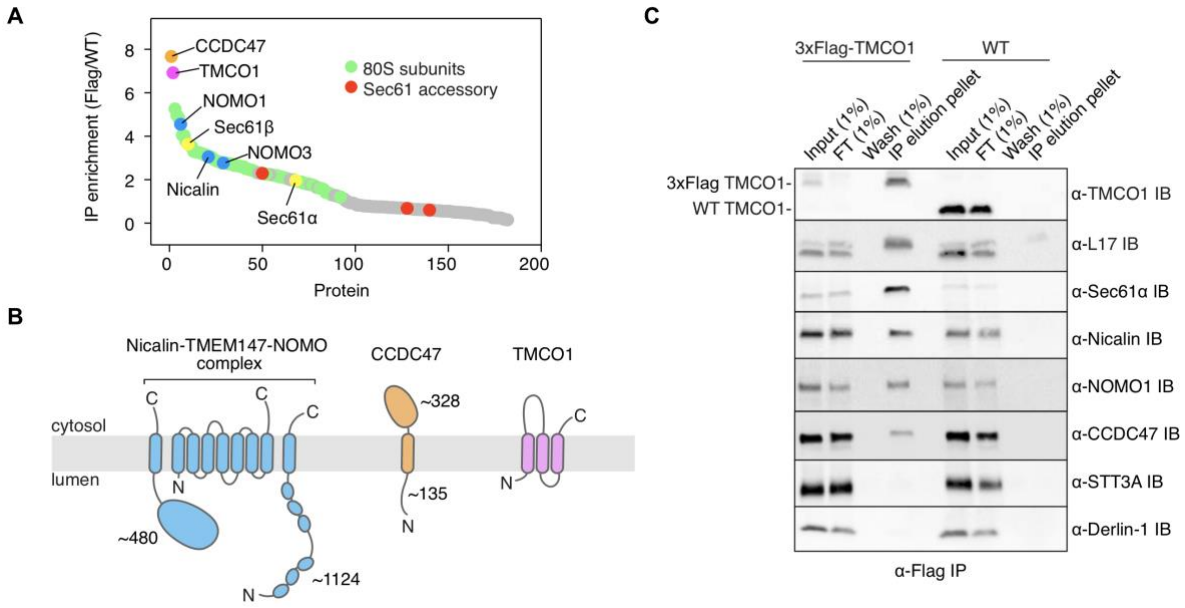


Figure 7. Ribosome-dependent assembly of a novel multi-component translocon. (A) The ribosome-associated fraction from emetine- and micrococcal nuclease-treated wild-type (WT) or 3xFlag-TMCO1 (Flag) HEK293 hRMs was digitonin-solubilized and analyzed by anti-Flag immunoprecipitation, sucrose cushion, and LC-MS/MS. Average fold-enrichment from three biological replicates is indicated. (B) Predicted topology and domain structure of the top hits (Sec61 complex not shown). Nicalin is predicted to contain a large globular lumenal domain, while the lumenal domains of NOMO and CCDC47 are predicted to be flexible. CCDC47 is predicted to contain a large cytosolic-facing domain harboring a conserved C-terminal coiled-coil. TMCO1 is predicted to contain a conserved, cytosolic-facing coiled-coil located between its first two TMDs. TMEM147, a known component of the Nicalin-TMEM147-NOMO complex, is a multi-pass protein that was not detected here, likely due to its short extra-membrane loops. (C) Top hits were confirmed by western blotting. Derlin-1, an abundant ER membrane protein that does not bind ribosomes, serves as a negative control. Note that STT3A, the catalytic subunit of the OST complex, is not detected. (D) Anti-flag immunoprecipitation of the non-ribosomally associated fraction shows that assembly of the five membrane components (but not Nicalin-TMEM147-NOMO) is largely ribosome-dependent.

3.5 TMCO1-ribosome complexes play a role in membrane protein biogenesis

We next sought to identify clients of TMCO1-bound ribosomes by sequencing their associated mRNAs. TMCO1 was purified from 3xFlagTMCO1 containing HEK293 cells via Flag immunoprecipitation and ribosome bound complexes were isolated by sedimentation. Control IPs to determine background mRNA levels were performed on wild-type HEK293 cells. In parallel, we prepared total RNA from the membrane fraction of wild-type cells. We then purified mRNA from all samples and performed high-throughput sequencing (RIP-seq).

We first asked whether this strategy specifically enriched for transcripts bound to TMCO1-associated ribosomes, or whether it simply enriched for the most abundant transcripts in the input sample. For this we calculated an "IP abundance" for each transcript, representing the difference in transcript abundance in the 3xFlag-TMCO1 and wild-type immunoprecipitations. When normalized by transcript abundance in the membrane fraction from wild-type cells ("input abundance"), we observed highly enriched transcripts (Fig. 8A; enrichment ≥ 2) across three orders of magnitude of input abundance, consistent with selective enrichment for TMCO1-linked transcripts.

Remarkably, we found that the set of most highly enriched transcripts shows a strong bias towards transmembrane proteins. Indeed, using a conservative enrichment threshold of ≥ 10, more than 98% (n=530) of the protein-encoding transcripts are predicted to encode membrane proteins of the secretory pathway (Fig. 8B). By contrast, soluble and mitochondrial proteins were not enriched.

We also analyzed the number of transmembrane domains present in proteins encoded by the most highly enriched transcripts (enrichment ≥ 10) from the TMCO1 immunoprecipitated

sample (Fig. 8C). Transcripts encoding single TMD proteins—by far the most abundant type of membrane protein in the human genome— were strongly depleted (~50-fold) in the TMCO1-associated dataset relative to the total membrane dataset. This was also true for tail-anchored membrane proteins, as expected for a co-translational process. By contrast, more than 96% of the most highly enriched transcripts (enrichment ≥ 10) encode multi-pass membrane proteins with three or more TMDs, including numerous transporters, receptors, transferases and hydrolases (Fig. 8D). This directly links TMCO1 and its associated factors to a co-translational process involving ~20% of all multi-pass membrane proteins encoded in the human genome.

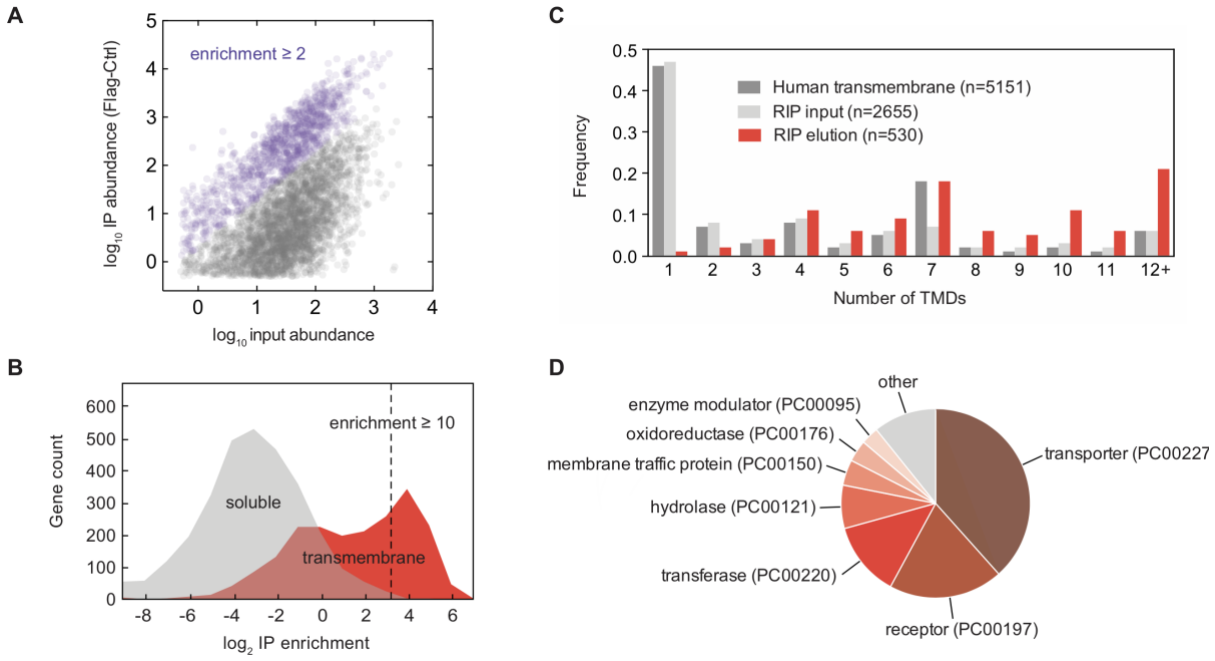


Figure 8. TMCO1-associated ribosomes are enriched for transcripts encoding multi-pass membrane proteins. (A) Scatterplot of mRNA transcript levels in the TMCO1 immunoprecipitated sample ("IP abundance") vs. mRNA transcript levels in the input sample ("input abundance"); average transcript abundance from three biological replicates is shown. Enrichment was calculated as the ratio of IP abundance to input abundance, for reliably expressed genes that were more abundant in the 3xFlag-TMCO1 pull-down than in the control pull-down. Transcripts with enrichment ≥2 (purple) are observed across three orders of magnitude of mRNA abundance in the input sample. (B) Histogram of log2 enrichment. Above a conservative enrichment threshold (e≥10) (dashed line), more than 98% of all transcripts encode integral membrane proteins (red). (C) Histogram showing the proportion of membrane proteins containing the indicated number of predicted (Uniprot) transmembrane domains (TMDs) in the human membrane proteome (dark gray), the RIP input sample (light gray), and in the subset of highly enriched (e≥10) membrane proteins from the RIP elution (red). Transcripts encoding single-pass membrane proteins are virtually absent from the TMCO1 immunoprecipitated sample, whereas those encoding multi-pass membrane proteins (≥4 TMDs) are enriched. (D) PANTHER protein classification for highly enriched (e≥10) membrane proteins.

3.6 Disruption of the TMCO1-translocon affects glutamate transporter expression

We next examined the steady-state expression levels of endogenous clients of the pathway in knockout cell lines. The largest class of membrane proteins identified in the RIP-seq experiment

are solute carrier (SLC) transporters. We focused on EAAT1 (SLC1A3), a homotrimeric glutamate transporter that helps limit neuronal excitation by clearing L-glutamate from the synaptic cleft.

We first constructed a set of knockout (KO) HEK293 cell lines by CRISPR/Cas9, as described previously[182], and characterized the levels of TMCO1, Nicalin-TMEM147-NOMO and CCDC47 by immunoblotting (Fig. 9A,B). We observed a slight but reproducible increase in the expression of Nicalin and NOMO in TMCO1 KO cells. Likewise, TMCO1 levels are modestly increased in Nicalin KO cells. Consistent with previous reports, we observed a severe reduction in NOMO protein levels in Nicalin KO cells[187]. Thus, the Nicalin KO cells are effectively deficient in all three components of the Nicalin-TMEM147-NOMO complex.

Next, we monitored the endogenous levels of EAAT1 monomers and multimers in wild-type and TMCO1 knockout cells by immunoblotting (Fig. 9A). Strikingly, the steady-state expression of EAAT1 monomers and multimers were reduced by ~2.4- and ~5.8-fold respectively (Fig. 9C). By contrast, the expression levels of a single-pass membrane protein (integrin α5) was unchanged in the KO cells. We also observed a strong reduction of EAAT1 monomer and multimer expression levels in Nicalin KO cells (2.0- and 5.0-fold, respectively), and a modest reduction in CCDC47 KO cells (Fig 9B). Similar results were observed TMCO1/Nicalin and TMCO1/CCDC47 double knockout cells. Notably, we found no significant change in mRNA levels of EAAT1 in TMCO1, Nicalin and CCDC47 single- and double-KO cells. Taken together with the RIP-seq data, these observations suggest that TMCO1-associated ribosomes play a role in multi-pass membrane protein biogenesis.

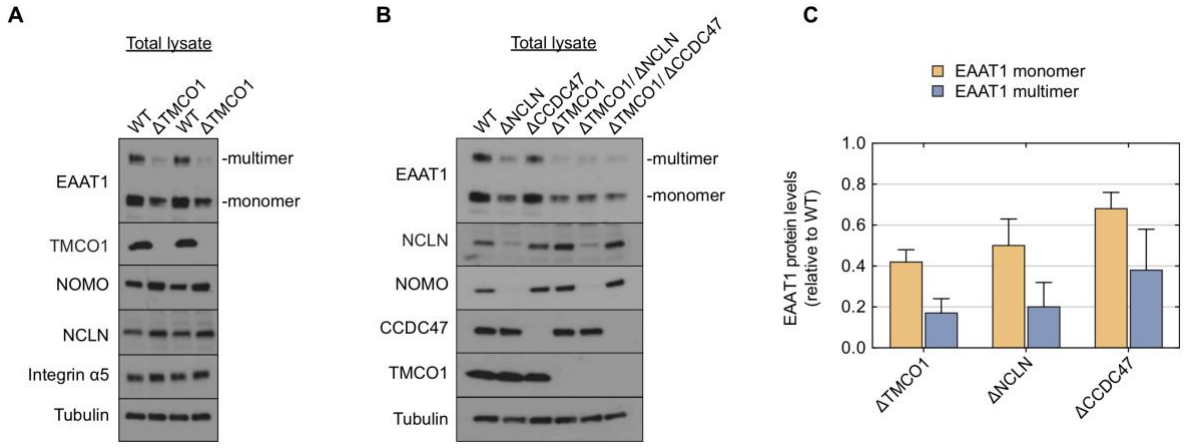


Figure 9. Steady-state levels of the multi-pass human glutamate transporter EAAT1 are decreased upon genetic disruption of TMCO1-associated ribosome complexes. (A) Western blot analysis of total cell lysate from wild-type (WT) and TMCO1 knockout HEK293 cells. Expression of the multi-pass membrane protein EAAT1, but not of the single-TMD protein integrin α5, is significantly decreased in the absence of TMCO1. Tubulin serves here as a loading control. Note the slight but reproducible increase in the steady state levels of NCLN and NOMO in TMCO1 knockout cells. (B) As in (A) but comparing EAAT1 expression levels in the indicated single- and double-knockout HEK293 cells. Deletion of NCLN phenocopies the EAAT1 expression defect observed in TMCO1 knockout cells. As shown previously[184,186], disruption of NCLN also reduces the expression levels of its binding partner, NOMO (and TMEM147, not shown). The faint band migrating just above NCLN in the ΔNCLN cells is a poorly resolved cross-reacting band. (C) Quantification of protein expression levels for monomeric and multimeric EAAT1 in ΔTMCO1 (n=9), ΔNCLN (n=4) and ΔCCDC47 (n=4) cells, measured by western blot; average (with standard deviation) levels are shown relative to wild-type cells.

3.7 Cross-linking mass spectrometry reveals a network of interactions surrounding the ribosomal exit channel

To gain insight into the organization of the TMCO1-associated ribosome complexes, we subjected affinity-purified complexes to chemical cross-linking and mass spectrometry (XL-MS). This yielded 1229 high-confidence intra- and inter-protein cross-links (SVM score ≥ 1.5) (Fig. 10A). Of the 715 cross-linked peptides that could be mapped to a high-resolution EM

structure of the human ribosome (PDB: 4UG0), more than 91% had Cα-Cα distances ≤35 Å (Fig. 10B), underscoring the high quality of the dataset.

We also detected 130 high-confidence cross-links involving the TMCO1 translocon components (Fig. 10C,D). As expected, the Sec61 complex is extensively cross-linked to ribosomal proteins around the exit channel. Likewise, specific cross-links to ribosomal subunits are detected for soluble regions of TMCO1, CCDC47, TMEM147 and NOMO that are predicted to face the cytosol. We also observed an extensive network of cross-links between Sec61, CCDC47, and TMCO1, and between these same components and ribosomal proteins RPL31, RPL17, RPL32 and RPL28 (Fig. 10C,D). Thus, in addition to confirming the predicted membrane topologies of TMCO1, CCDC47 and components of the Nicalin-TMEM147-NOMO complex, these data reveal a network of interactions surrounding the Sec61 complex at the ribosome exit channel.

3.8 Architecture of TMCO1-associated RTCs revealed by single-particle cryo-EM

We used single-particle cryo-electron microscopy (cryo-EM) to visualize the TMCO1-associated ribosome-translocon complex (RTC) (Table 1, Fig 11A). Reconstructions yielded an RTC with extensive membrane and lumenal density surrounding the exit channel (Fig. 11A). Well-defined density for the ribosome shows hybrid state A/P and P/E tRNAs, and a nascent polypeptide throughout the exit tunnel. The Sec61 complex is in a conformation similar to the previously observed "primed" state[29,192], with a closed lateral gate and the plug helix occluding the central pore (Fig. 11B). We modeled the remaining membrane components using a combination of cryo-EM, XL-MS and structure prediction.

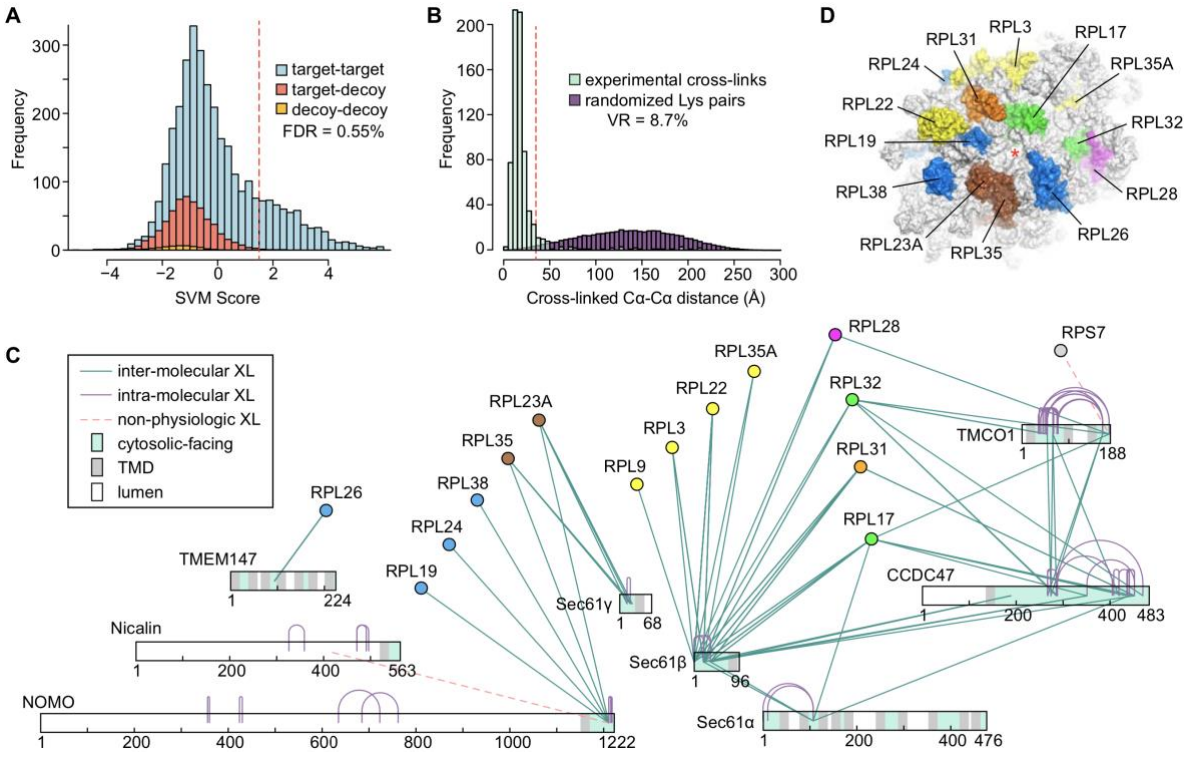


Figure 10. Cross-linking and mass spectrometry analysis of TMCO1-associated ribosomes. (A) Classification of cross-linked residue pairs. Crosslinks are reported at a false discovery rate of 0.55% corresponding to a classification score threshold of 1.5 (dashed red line). FDR was determined by modeling the distributions of crosslinks with zero (blue), one (red) or two (yellow) peptides randomly matching at given score using a decoy database (see methods). (B) Distance distributions for all ribosome-ribosome crosslinks mapping to the high-resolution cryo-EM structure of the human ribosome (PDB: 4ug0). After applying a conservative cutoff (SVM score ≥ 1.5), 715 crosslinks could be mapped to the EM structure with the majority (91.3%) of crosslink Cα-Cα distances lying within the 35 Å upper limit (red dashed line). A distribution sampled from all possible Lys-Lys pairs on the ribosome (purple) is shown for comparison. (C) Network diagram of the 130 strongest (SVM score ≥ 1.5) intra- (purple lines) and inter-molecular (cyan lines) cross-links involving the membrane components. Lumenal (white), transmembrane (grey) and cytosolic (cyan) regions of the membrane components are indicated. Crosslinks identified between the lumenal domain of Nicalin and the cytosolic domain of NOMO, as well as between the C-terminus of TMCO1 and RPS7, are likely non-physiologic (dashed red lines). (D) Surface representation of the human ribosomal large subunit highlighting the locations of proteins that can be crosslinked to one or more of the membrane components. Ribosomal subunits are colored as in (C), and the exit channel is marked with a red asterisk.

Density corresponding to a bundle of TMDs and a large lumenal domain was visible near the hinge region of Sec61α. To fit this density, we constructed a molecular model for Nicalin

and TMEM147 based on their evolutionary relationship to the APH-1-Nicastrin sub-complex of human γ-secretase[193]. The distinctive "tent" arrangement of TMEM147 TMDs fit well into the membrane density (Fig. 11C). In this configuration, the conserved TM3-TM4 loop of TMEM147 contacts rRNA H7 and RPL26, consistent with the XL-MS data (Fig. 10C,D and 11D).

We also observed two distinct regions of unexplained helical density in the cytosolic vestibule. The first of these was assigned to the conserved, cytosolic-facing coiled-coil of TMCO1, which is enriched in positively charged residues (Fig. 7B and 11E). This assignment places the coiled-coil in direct contact with ribosomal RNA helix H19, and in close proximity to RPL28 and RPL32, consistent with the XL-MS analysis (Fig. 10C,D). Using the docked coiled-coil as a constraint, we placed a model of full-length TMCO1 into the density. This satisfies extensive intra-molecular crosslinking between the TMCO1 C-terminus and coiled-coil, as well as crosslinks between the TMCO1 C-terminus and ribosomal proteins RPL17 and RPL28 (Fig. 10C,D and 11G).

The second region of unexplained density wedges between Sec61Sα and rRNA helix H24 (Fig 11F). A co-variation-based model of the cytosolic region of CCDC47 shows a helical "hook" at the extreme C-terminus that protrudes from a globular N-terminal domain. This unusual structural feature is strongly conserved and enriched in positively charged side chains, and could be modeled easily into the density with its C-terminus pointed at the mouth of the

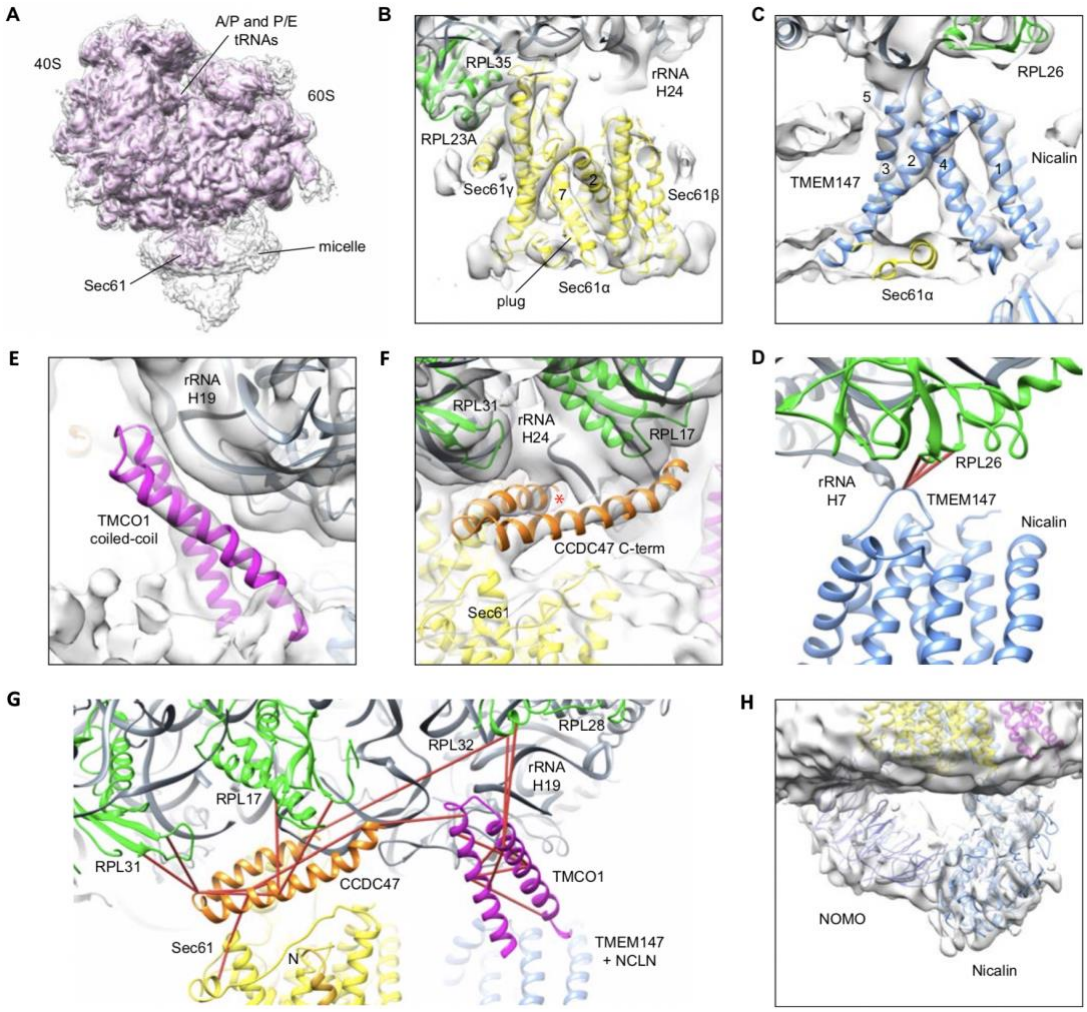


Figure 11. Localization of TMCO1, NCLN, TMEM147 and CCDC47 near the exit tunnel of actively translating ribosomes. (A) Globally refined, unsharpened map contoured at two different thresholds. (B) Model of the Sec61 complex (yellow). The model is superimposed on an unsharpened map generated through local refinement using a soft mask encompassing only the Sec61 complex; CCDC47 and TMCO1 are omitted for clarity. The plug and lateral gate helices are highlighted. (C) A homology model of the TMEM147-NCLN complex (blue) overlaid on an unsharpened map generated through local refinement using a soft mask around the whole translocon complex. Note the tent-like structure of TMEM147; the Sec61α hinge occupies part of this cavity. (D) Experimentally observed crosslinks (red) between residues in RPL26 (green) and the conserved cytosolic-facing TM3-TM4 loop of TMEM147. (E) Model of the TMCO1 coiled-coil and (F) the CCDC47 C-terminal Coiled Coil on an unsharpened map generated through local refinement using a soft mask around the whole translocon complex. Nascent chain density (marked with a red asterisk in panel F) extends to the exit channel, where the extreme C-terminal helix of CCDC47 inserts itself. (G) Extensive crosslinking between CCDC47, Sec61, TMCO1 and the ribosome. Additional crosslinking is observed to the N-terminal region of Sec61β (gold), to the globular cytosolic domain of CCDC47, and to the C-terminus of TMCO1, but these are not mappable due to disorder in the structure. (H) Portions of the NOMO and Nicalin lumenal domains modeled below Sec61α.

ribosomal exit tunnel. This configuration is consistent with multiple intra- and inter-molecular crosslinks observed to Sec61α, the flexible N-terminus of Sec61β, ribosomal subunits RPL17, RPL31 and RPL32, and to the TMCO1 coiled-coil and C-terminus (Fig. 10C,D and 11G).

At even lower contour levels we see additional density extending from the N-terminus of the CCDC47 hook motif towards density which we tentatively assign to the globular cytosolic domain of CCDC47. Although we did not explicitly model this domain, our assignment is consistent with the extensive crosslinking observed between the globular domain of CCDC47 and the N-terminus of Sec61β, TMCO1, and RPL17 (Fig. 10C,D).

Table 1 Collection and Reconstruction Data for the TMCO1-containing RTC

Data collection and processing	
Magnification	105,000x
Voltage (kV)	300
Electron exposure (e–/Å^2)	50
Defocus range (μm)	-1.0 to -3.0
Pixel size (Å)	0.66
Symmetry imposed	C1
Initial particle images (no.)	140,489
Final particle images (no.)	32,512
Map resolution (Å)	6.3
FSC threshold	0.143
Map resolution range (Å)	3.5 to 8.0

Additional density below Sec61 could be assigned to a portion of the lumenal domains of Nicalin and NOMO (Fig. 7B). We used our prior identification of TMEM147 to facilitate placement of the large N-terminal domain of Nicalin. Additionally, a co-variation-based model of the membrane-proximal region of NOMO, consisting of three IgG-like repeats connected by short peptide linkers, could be fit into the density with minor rigid-body adjustments. Notably, this places part of NOMO in direct contact with the Nicalin lumenal domain, consistent with previous reports that the Nicalin-NOMO interaction is mediated by their lumenal domains[186] (Fig. 11H). Although we did not explicitly model the rest of NOMO, this assignment is also consistent with the observed cross-links between the conserved cytosolic C-terminus of NOMO and ribosomal proteins L19, L24, L38, L35, and L23A (Fig. 10C,D).

The resulting structural model shows a loose assembly of accessory factors arranged around the core Sec61 complex and the ribosome exit channel (Fig. 12A). This organization is dictated by the direct interactions of TMCO1, CCDC47, and TMEM147 with the ribosome. By contrast, inter-subunit interactions between the cytosolic regions of TMCO1, CCDC47, Sec61, and between the lumenal domains of Nicalin and NOMO, appear to be more dynamic. A well-defined inter-subunit contact within the membrane occurs at the Sec61α hinge, which inserts into the lipid-exposed cavity of TMEM147. This interaction is structurally analogous to that observed between the presenilin-1 and APH-1 subunits of γ-secretase[193]. The functional significance of this interaction is unclear, but it may be important for recruiting the Nicalin-TMEM147-NOMO complex, or regulating Sec61 channel dynamics[194,195].

3.9 Discussion

Together, these data describe a new translocon involved in multi-pass membrane protein biogenesis in the eukaryotic ER. The overall architecture of this novel TMCO1-associated translocon suggests that it interacts with different domains of nascent multi-pass membrane protein clients during their synthesis (Fig. 12A). On the cytosolic side of the membrane the emerging nascent chain would be shielded by the TMCO1 coiled-coil and CCDC47. Likewise, NOMO and Nicalin provide a flexible scaffold below Sec61 on the lumenal side of the membrane. Within the bilayer, the transmembrane regions of Sec61, TMCO1 and TMEM147 surround a large, presumably lipid-filled cavity (Fig. 12B). In this configuration the predicted hydrophilic groove of TMCO1, analogous to that found in other Oxa1 family members[182,142,150], could facilitate the passage of TMDs from the cytosolic vestibule into this lipid cavity. These

architectural features may facilitate the insertion, folding and/or assembly of multi-pass membrane protein clients in a protected environment.

Additionally, our data point to a model in which the translocon composition varies temporally during synthesis to meet the biophysical demands of its clients. For example, many of the multi-pass membrane protein clients of the TMCO1 translocon are glycosylated, and yet we see no evidence for OST components by cryo-EM or mass spectrometry. While it is possible that these are all STT3B clients, it is also likely that STT3A dynamically accesses these clients at the appropriate stage during synthesis. The structure suggests a possible mechanism: whereas steric overlap in the bilayer and lumenal regions would clearly preclude co-occupancy, the ribosome binding sites in the TMCO1-associated complex (i.e., conserved cytosolic regions of TMCO1, CCDC47 and TMEM147) are non-overlapping with those in the OST complex (RPN1) (Fig. 11G). This complementarity may facilitate exchange between the different components at the appropriate moment during synthesis possibly triggered by conformational changes in the ribosome or Sec61 complex (Fig. 12C).

Though our RIP-seq identified hundreds of potential substrates were not able to determine what features define a client of this new translocon. However, as is generally true of multi-pass membrane proteins, clients of the TMCO1 translocon are enriched in marginally hydrophobic TMDs, and TMDs connected by short intervening loops[196.197]. Such sequences are not rapidly accommodated by Sec61, instead accumulating in the cytosolic vestibule between the ribosome and the membrane (Fig. 12C)[64]. The TMCO1 translocon would be well positioned to act on these sub-optimal sequences (Fig. 12A,C). It is possible that the kinetics of translation and translocation/insertion through Sec61 may dictate whether the TMCO1 machinery is needed for a particular client. Intriguingly, we also see the coiled-coil of CCDC47 reaching directly into the

ribosomal exit channel. The very N-terminus of the CCDC47 coiled-coil may even be interacting with nascent proteins as they exit the ribosome.

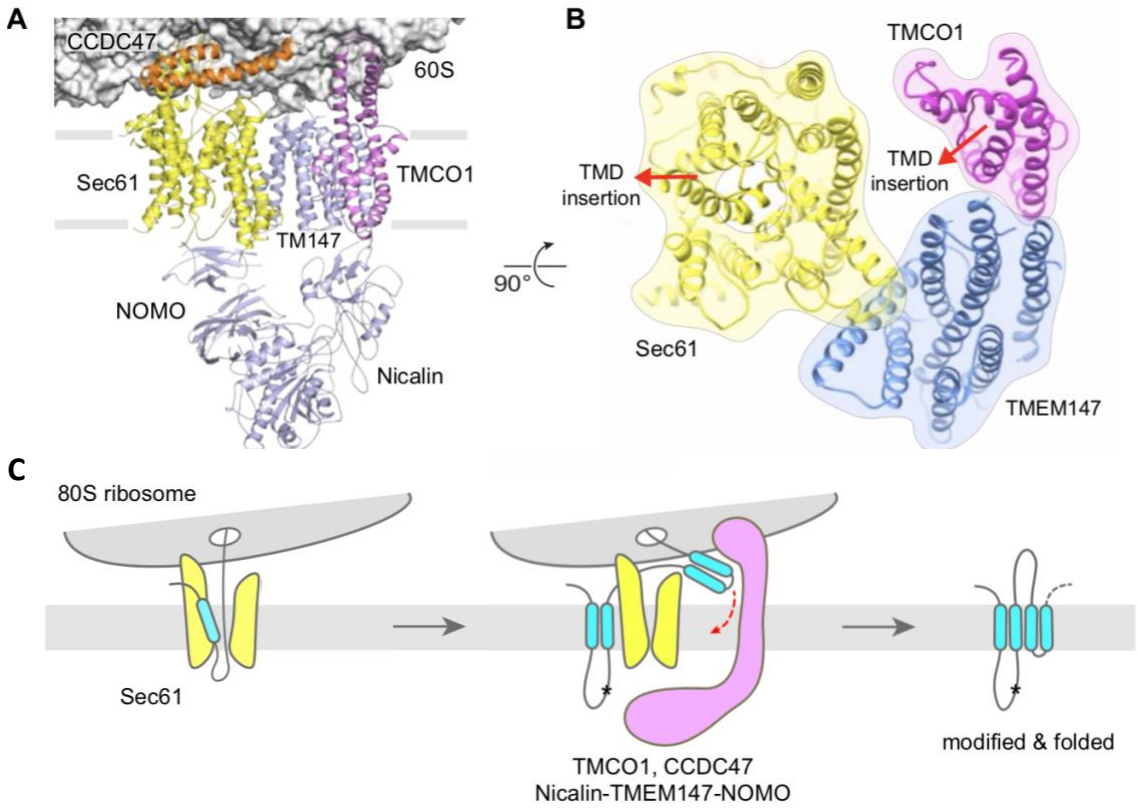


Figure 12. Low resolution model of the TMCO1 translocon. (A) Overview of the complex. (B) Slice through the membrane, viewed from the lumen. Locations of the Sec61 lateral gate and the TMCO1 hydrophilic groove are indicated. (C) Model for TMCO1 and associated factors inserting a TMD hairpin into the membrane via the central lipid cavity. In this model, the TMDs may forego insertion via Sec61alpha and move from the vestibule directly into the membrane.

From an evolutionary standpoint, there are a number of notable features of the TMCO1-containing translocon. First, our data demonstrate that all members of the Oxa1 superfamily play a role in membrane protein biogenesis, confirming the hypothesis that the Oxa1 superfamily is comprise do structurally and functionally related membrane proteins. Second, in contrast with

proposals for bacterial YidC, TMCO1 is not found at the Sec61 lateral gate[46,47]. This may reflect differences in mechanism and awaits higher resolution structural insight into the bacterial holo-translocon. Lastly, the structural mimicry of the Sec61α hinge – TMEM147 interaction (Fig. 12A), may indicate some form of convergent regulatory mechanism for complex activation and/or recruitment.

Lastly, it is worth speculating on the functional interaction between the TMCO1 translocon and the EMC complex, a subunit of which – EMC3 – is also a member of the Oxa1 superfamily[182]. The EMC also acts on multipass membrane proteins, although the comparatively small list of known EMC clients precludes a comparative analysis with TMCO1 clients[198]. Although EMC can function in the vicinity of the translating ribosome, no direct EMC-ribosome interaction has been demonstrated[199]. A reasonable model is thus one in which EMC and TMCO1 can work on the same substrates but at different times or on different regions. Similarly, these pathways may be able to compensate for one another under conditions of stress. These questions are important and worthy of future study.

4 Msp1 Is a Membrane Protein Dislocase for Tail-Anchored Proteins

4.1 Overview

The majority of TA proteins in eukaryotes are targeted either to the ER or mitochondrial membrane for insertion. TA proteins are brought to the ER by the conserved GET pathway, however disruption of this pathway can lead to mistargeting of ER resident TA proteins to the mitochondrial outer membrane. Two studies from 2014 identified Msp1, a signal anchored hexameric AAA-ATPase, as part of the quality control pathway which removes mis-directed TA proteins from the OMM. To better understand how this quality control process, we sought to structurally and biochemically characterize the activity of Msp1.

It remained unclear whether Msp1 could independently remove TA protein from the OMM or whether it required additional support factors to modulate its activity. Using in vitro assays with reconstituted Msp1 and TA containing proteoliposomes, we found that Msp1 was capable of independently removing TA proteins from membranes in an ATP dependent manner. Additionally, we used a combination of x-ray crystallography and single particle cryo-electron microscopy to characterize the structure of the Msp1 ring hexamer. This structural analysis allowed us to identify and characterized residues in Msp1 important to structure and function.

A version of this work was published as Wohlever, M. L., Mateja, A., McGilvray, P. T., Day, K. J., Keenan, R. J. 2017, Msp1 Is a Membrane Protein Dislocase for Tail-Anchored Proteins, Molecular Cell, 67:1-9

4.2 Contributions

In this work, I performed the sample generation, data collection, and 2D class analysis to determine the correct dimensions of the Msp1 hexamer. M.L.W purified the Msp1 proteins,

made all mutants, and performed all reconstituted in vitro activity assays. M.L.W and A.M. crystallized the Msp1 protein and collected the crystallographic data. K.J.D and M.L.W performed the in vivo yeast complementation assays and microscopy. M.L.W. and R.J.K conceived of the project and co-wrote the manuscript.

4.3 Introduction

Maintaining the correct intracellular distribution of membrane proteins is essential to the organization of eukaryotic cells. This is accomplished in part through biosynthetic pathways that recognize specific signals in nascent proteins and target them to their correct destination[12,200-202]. When these processes fail, quality control pathways help maintain order by identifying damaged or mislocalized membrane proteins and targeting them for degradation[203-205].

One example of the interplay between biosynthetic and quality control pathways is given by tail-anchored (TA) membrane proteins. TA proteins, which play critical roles in many different cellular processes[209-211], are defined by the presence of a single transmembrane domain (TMD) located near the C terminus[87]. Following synthesis, most TA proteins are initially targeted to the endoplasmic reticulum (ER) or the outer mitochondrial membrane (OMM). Targeting to the ER is mediated by the evolutionarily conserved GET pathway[91,93,99,212]. Disruption of this pathway results in accumulation of certain ER-destined TA proteins in the OMM[99,213-215].

Two recent studies have identified the yeast AAA ATPase Msp1 (ATAD1/Thorase in metazoans) as part of a new quality control pathway that eliminates these mislocalized TA proteins from the OMM[213,215]. Msp1 is a signal-anchored membrane protein comprising an N-terminal TMD that tethers its soluble AAA domain to the cytosolic face of the OMM and

peroxisomes[216]. In yeast, simultaneous disruption of Msp1 and the GET pathway leads to severe mitochondrial damage, including loss of mitochondrial proteins and DNA, and altered mitochondrial morphology. Similarly, depletion of mammalian ATAD1 compromises mitochondrial function[213,217]. Nevertheless, how the Msp1 quality control pathway prevents the accumulation of mislocalized TA proteins in the OMM is not known. To address this, we combine biochemical reconstitution with structural analysis to show that a ring hexamer of Msp1 is necessary and sufficient for ATP-dependent extraction of TA proteins from the membrane.

4.4 The Minimal Tail-Anchored Protein Extraction Machinery

To investigate whether Msp1 extracts integrated TA proteins from the membrane and to define the minimal number of components required for its activity, we sought to establish an in vitro assay using purified, recombinant proteins (Fig. 13A). For this, we devised methods to generate proteoliposomes containing TA proteins and Msp1 in the correct orientation and a soluble chaperone trap to capture dislocated TA proteins for facile isolation.

Full-length recombinant S. cerevisiae Msp1 and a model ER TA protein (SumoTMD)[94] were co-reconstituted from detergent solution into proteoliposomes with lipid composition mimicking that of the yeast OMM[218] (Fig. 14). Carbonate extraction of the reconstituted proteoliposomes showed that the majority of TA protein becomes integrated into the bilayer, with only a small amount present in the supernatant (Fig. 13B). This putatively non-integrated population of TA protein was removed from the proteoliposomes by binding to GST-tagged SGTA, a TMD-binding chaperone of the eukaryotic cytosol[94]. Removal of GST- SGTA via a glutathione-affinity resin resulted in ‘‘pre-cleared’’ proteoliposomes in which the TA protein is

quantitatively resistant to carbonate extraction (Fig. 13B) and thus integrated into the lipid bilayer.

To define the orientation of reconstituted Msp1 and TA protein, we examined the proteinase K (PK) sensitivity of N- and C-terminal epitope tags in pre-cleared proteoliposomes (Figure 13A). Consistent with its in vivo topology, the N-terminal 6xHis tag of the TA protein was sensitive to PK digestion in proteoliposomes, while its C-terminal (lumenal) opsin tag was protected (Figure 13C). Similarly, the C-terminal 6xHis tag of reconstituted Msp1 was sensitive to PK treatment, consistent with exposure of its AAA ATPase domain on the vesicle surface. Thus, both Msp1 and its TA protein client in the pre-cleared proteoliposomes are predominately oriented as they would be in the OMM.

With defined proteoliposomes in hand, we next monitored TA protein extraction under different conditions (Figure 13 and 14). In this assay, dislocated TA protein is captured via its exposed TMD by excess GST-SGTA and recovered by glutathione-affinity resin (Figure 13A). Proteoliposomes containing only TA protein and ATP showed background levels of dislocation (~1%–2% of input TA protein), but dislocation was stimulated by ~10-fold (to ~10% of input) in proteoliposomes that also contained Msp1 (Figure 13D).

Dislocation was strictly dependent on nucleotide hydrolysis, as no substrate capture was observed in the absence of ATP or when the hydrolysis-deficient Msp1 mutant (E193Q) was used instead of the wild-type protein (Figure 13D). This is consistent with the observation that yeast containing E193Q Msp1 fail to clear mislocalized TA proteins from the OMM[94,215].

Dislocation activity increased as a function of time and Msp1 concentration, and no ATP-dependent release of Msp1 from the membrane was observed (Fig. 14). Importantly, a mixture of separate Msp1- and TA-only proteoliposomes showed no TA protein dislocation (Figure 13D),

illustrating that Msp1 must be in the same membrane as its client. Taken together, these data indicate that Msp1 is the only membrane protein required for ATP-dependent dislocation of a TA protein, thereby defining a minimal reconstituted system for studying this process.

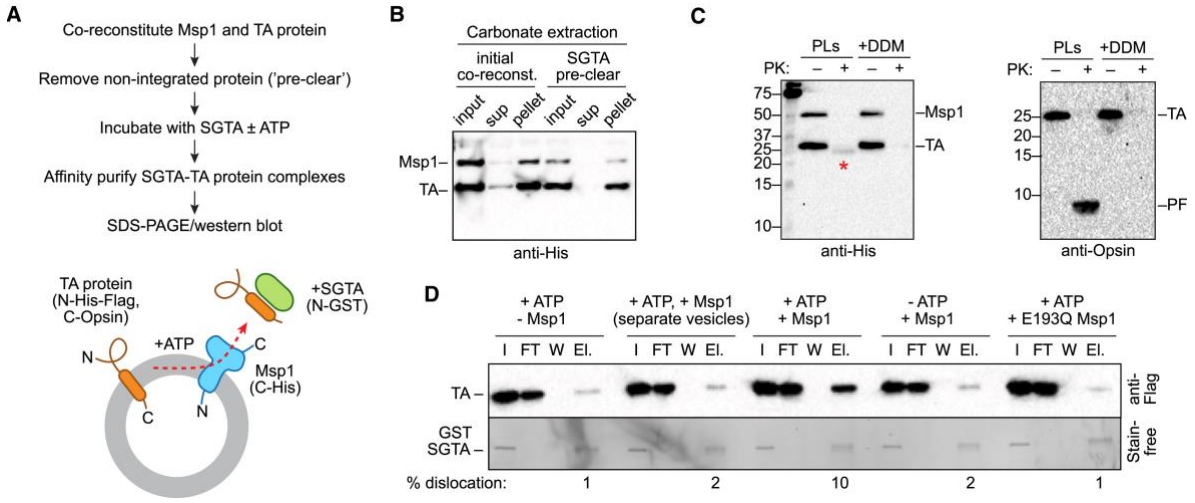


Figure 13. Msp1 Drives ATPase-Dependent TA Protein Dislocation in a Purified System. (A) Following co-reconstitution, non-integrated TA proteins are removed by incubation with GST-tagged SGTA and passage over a glutathione affinity resin. These ‘‘pre-cleared’’ proteoliposomes (PLs) are incubated with ATP and SGTA. SGTA-TA complexes are isolated by pull-down and analyzed by SDS-PAGE and WB. Epitope tags on Msp1 and TA protein are indicated (see also Figure 14). (B) The presence of non-integrated TA protein before and after pre-clearing was monitored by carbonate extraction, SDS-PAGE, and western blotting. A small amount of non-integrated TA protein is present in the carbonate extraction supernatant immediately following co-reconstitution but not after GST-SGTA pre-clearing. (C) The orientation of co-reconstituted TA protein and Msp1 in pre-cleared PLs was defined by proteinase K (PK) digestion in the presence or absence of DDM; products were analyzed by SDS-PAGE and western blotting. In vesicles, the N-terminal 6xHis tag of the TA protein and the C-terminal 6xHis tag of Msp1 are sensitive to digestion, while the C-terminal opsin tag remains protected (PF); this is diagnostic of proper orientation in the proteoliposomes. The red asterisk indicates a minor population of TA protein oriented with its N-terminus facing the vesicle lumen. (D) TA protein dislocation under different conditions. Samples were analyzed by SDS-PAGE and either WB (for TA protein) or stain-free gel (for GST-SGTA); the elution and wash fractions correspond to 53 volume equivalents of the input and flow-through fractions. Dislocation efficiencies were estimated using ImageJ.

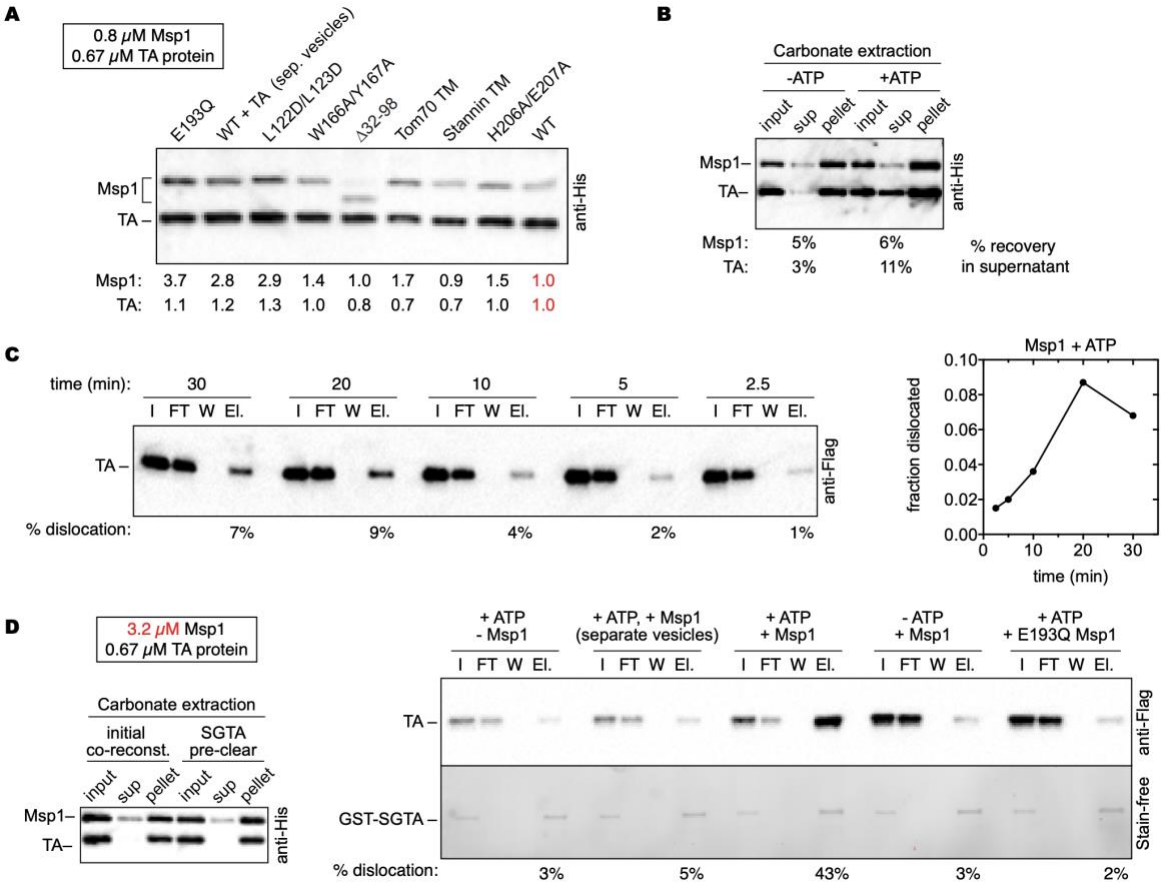


Figure 14. Additional characterization of the dislocation assay. (A) Following co-reconstitution with 0.67 µM TA protein and 0.8 µM of the indicated Msp1 construct, proteoliposomes were pre-cleared (see Figure 1B) and the relative amount of integrated Msp1 and SumoTMD was quantified by SDS-PAGE and anti-His western blotting. Band intensities were estimated using ImageJ. Relative levels vary by less than four- and two-fold for Msp1 and TA protein, respectively. (B) Anti-His western blot of a dislocation assay monitored by carbonate extraction instead of GST-SGTA pull-down. Here, a small amount of unincorporated Msp1 is observed in the carbonate supernatant (corresponding to residual Msp1 that sometimes remains after pre-clearing); importantly, it does not accumulate in an ATP dependent manner. In contrast, TA protein does accumulate in the carbonate supernatant in an ATPdependent manner. Thus, the proteoliposomes remain intact throughout the ATP-dependent dislocation reaction, and Msp1 itself is not dislocated from the proteoliposomes. (C) Time course of protein dislocation by wild-type Msp1, monitored by anti-Flag western blot. (D) A protein dislocation assay using proteoliposomes reconstituted with four-fold higher concentrations of Msp1 (3.2 µM instead of 0.8 µM). Under these conditions, very little non-integrated TA protein is detected in the carbonate extractable supernatant (left panel). In the presence of ATP, more than 40% of the TA protein is dislocated; this increased efficiency presumably reflects the larger number of proteoliposomes containing Msp1 hexamers.

4.5 Crystal Structure of Msp1

To gain insight into Msp1-mediated extraction of TA proteins, we sought to obtain structural information. Initial attempts to crystallize the soluble region of S. cerevisiae Msp1 in the presence or absence of nucleotide were unsuccessful. However, a cleaved construct containing a 3C protease site after residue 345 yielded diffraction quality crystals in the nucleotide-free state (‘‘D1-32+3C’’; Fig. 15A). Notably, an otherwise full-length construct of Msp1 containing the same C-terminal truncation retained the ability to complement an Msp1 deletion, indicating that the C-terminal region is not strictly required for function in vivo (Fig. 16). We solved these crystals at 2.6Å resolution by single-wavelength anomalous dispersion (Table 2) and found a single copy of Msp1 in the asymmetric unit. Analysis of the crystal packing revealed no hexamer-like interactions.

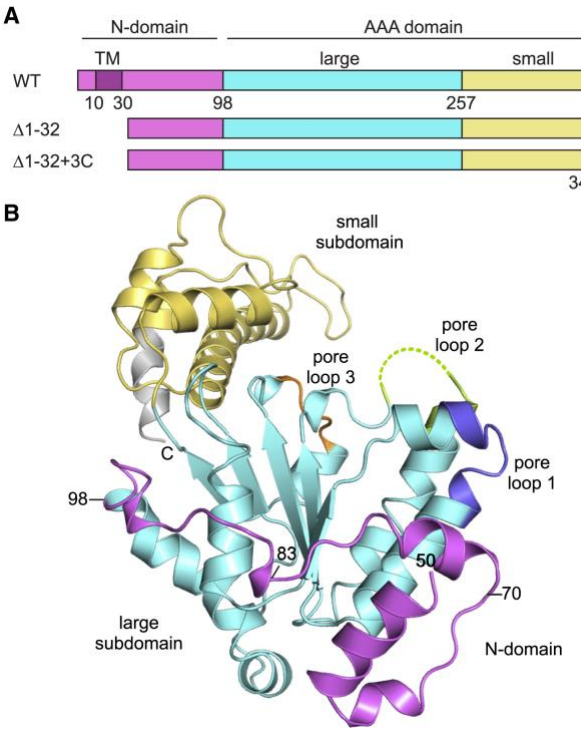


Figure 15. Crystal Structure of Soluble, Nucleotide-free S. cerevisiae Msp1. (A) Domain organization of full-length Msp1 (WT), a soluble construct lacking the transmembrane domain (D1-32), and the construct used for crystallization (D1-32+3C), which contains a 3C protease cleavage site after residue 345; following cleavage, the protein retains an eight-residue scar (gray). (B) Crystal structure of soluble Msp1 determined in its nucleotide-free, monomeric state. The protein is colored as in (A). The three ‘‘pore loops’’ are highlighted; disordered residues in pore loop 2 are indicated by dotted line.

The structure of the cytosolic Msp1 fragment comprises a helical N-domain and an AAA ATPase domain with characteristic ‘‘large’’ (nucleotide binding) and ‘‘small’’ (helical bundle) subdomains (Fig. 15B). Structure-based alignments of the large and small subdomains of Msp1

identified p97/Cdc48, FtsH, and members of the ‘‘meiotic’’ clade of AAA ATPases, including spastin, fidgetin, and Vps4[219], as the closest structural homologs of Msp1 (Fig. 16).

The Msp1 large subdomain contains the canonical Walker A (P loop) and B motifs and is connected to the small subdomain by a short flexible loop (Figure 15B). The small subdomain comprises a four-helix bundle; in some meiotic clade AAA ATPases, functionally important insertions are found between the third and fourth helices of the small subdomain[220,221]. In Msp1, the third helix protrudes from the bundle by virtue of an ~2º turn extension relative to the corresponding helix in closely related homologs (Fig. 16); however, the functional significance of this remains unclear, since there is little sequence conservation in this region among Msp1 homologs.

In the nucleotide-free, monomeric Msp1 crystal, the small subdomain is rotated 180º from its usual position in the ATP-bound state of other AAA ATPases such that the ATP-binding site is not

Table 2 Data Collection and Refinement Statistics

	Selenomet-Msp1 (Δ1-32+3C)
Data Collection	
Space group	*P*3,12
Cell dimensions	
a, *b*, *c* (Å)	56.4, 56.4, 206.7
α, β, γ (°)	90, 90, 120
Wavelength (Å)	0.9792
Resolution (Å)	48.8–2.63 (2.72–2.63)[a]
Redundancy	17.4 (9.4)
Completeness (%)	99 (100)
$I/\sigma I$	9.7 (1.2)
R_{pim} (%)	7.6 (77.1)
$CC_{1/2}$ (%)	99.5 (52.6)
Refinement	
Resolution range (Å)	48.8–2.63
No. Unique Reflections	11,555
R_{work}/R_{free} (%)	23.0/26.0
No. of non-H atoms	
Protein	2,362
Ligands	21
Solvent	11
Average B (Å^2)	
Protein	76.7
Ligands	63.1
Solvent	58.0
RMS Deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.56
Ramachandran analysis	
Favored (%)	97.0
Allowed (%)	2.7
Outliers (%)	0.0

Data were obtained from a single crystal.
[a]Values in parentheses refer to the high-resolution shell

fully formed (Fig. 17). The relative orientation of the large and small subdomains of AAA ATPases typically varies[222], but the ‘‘open’’ conformation observed here is unusual. This may be due to crystal packing forces or truncation of the final 17 C-terminal residues. In certain related AAA ATPases (including spastin and Vps4), these residues pack against the large subdomain; notably, however, truncation of the Msp1 C terminus has no effect on hexamerization or in vivo activity (Fig. 16B).

A unique structural feature of Msp1 is its N-domain (Figure 15B and 16). The first 16 residues of the crystallized construct, which immediately follow the TMD in the full-length protein, are disordered, likely due to intrinsic flexibility. Residues 50–80 form a small, three-helix bundle that packs loosely against the large subdomain, and the remainder of the N-domain wraps around the large subdomain in a mostly extended conformation. While the overall sequence conservation of the N-domain is modest, its amphipathic character is conserved and hydrophobic residues throughout the domain help anchor it against the surface of the large subdomain.

4.6 Model of the Msp1 Hexameric Ring

Most AAA ATPases function as oligomers. To determine the oligomeric state of the Msp1 cytosolic region, we used size-exclusion chromatography and multi-angle laser light scattering. A soluble Msp1 construct lacking the first 32 residues (Fig. 19) exists as a mixture of monomers and dimers in the absence of nucleotide but forms hexamers in the presence of saturating concentrations of the non-hydrolyzable ATP analog, ATPγS (Fig. 18A and 19). Likewise, a soluble construct containing the hydrolysis-inactivating mutation E193Q hexamerizes in the presence of ATP (Fig. 18A and 19).

We obtained low-resolution structural information on the E193Q hexamer by single-particle cryo-electron microscopy (cryo-EM). In the presence of ATP, Msp1 hexamers are clearly visible (Fig 19G); however, because the particles are preferentially oriented, we could not generate high-resolution 3D reconstructions. Nevertheless, 2D class averages generated from nearly 10,000 particles reveal a hexameric ring structure with a diameter of ~13.1 nm and distinctive protrusions emanating from the central ring (Figure 18B).

To generate a pseudo-atomic model of the Msp1 hexamer, we individually superimposed the large and small subdomains of the monomeric crystal structure onto the corresponding subdomains of the D2 ring of hexameric p97 (PDB: 5C18) (Fig. 16E). The resulting model contains few inter-subunit steric clashes, and these would presumably be relieved by subtle conformational changes accompanying ATP binding and hexamer formation. Within the hexameric assembly, the large and small subdomains are arranged in approximately the same plane, while the N-domain lies below this plane, along what is likely the mem- brane proximal surface of Msp1 (Figure 18C). Consistent with the 2D class averages, the elongated small subdomain protrudes outward from the center of the ring, giving the overall assembly a longest dimension of ~13.4 nm. The Msp1 hexamer model has a well-formed ATP-binding site with the characteristic ‘‘second region of homology’’ (SRH) of ‘‘meiotic clade’’ AAA ATPases[219] close to the nucleotide-binding pocket of the adjacent subunit.

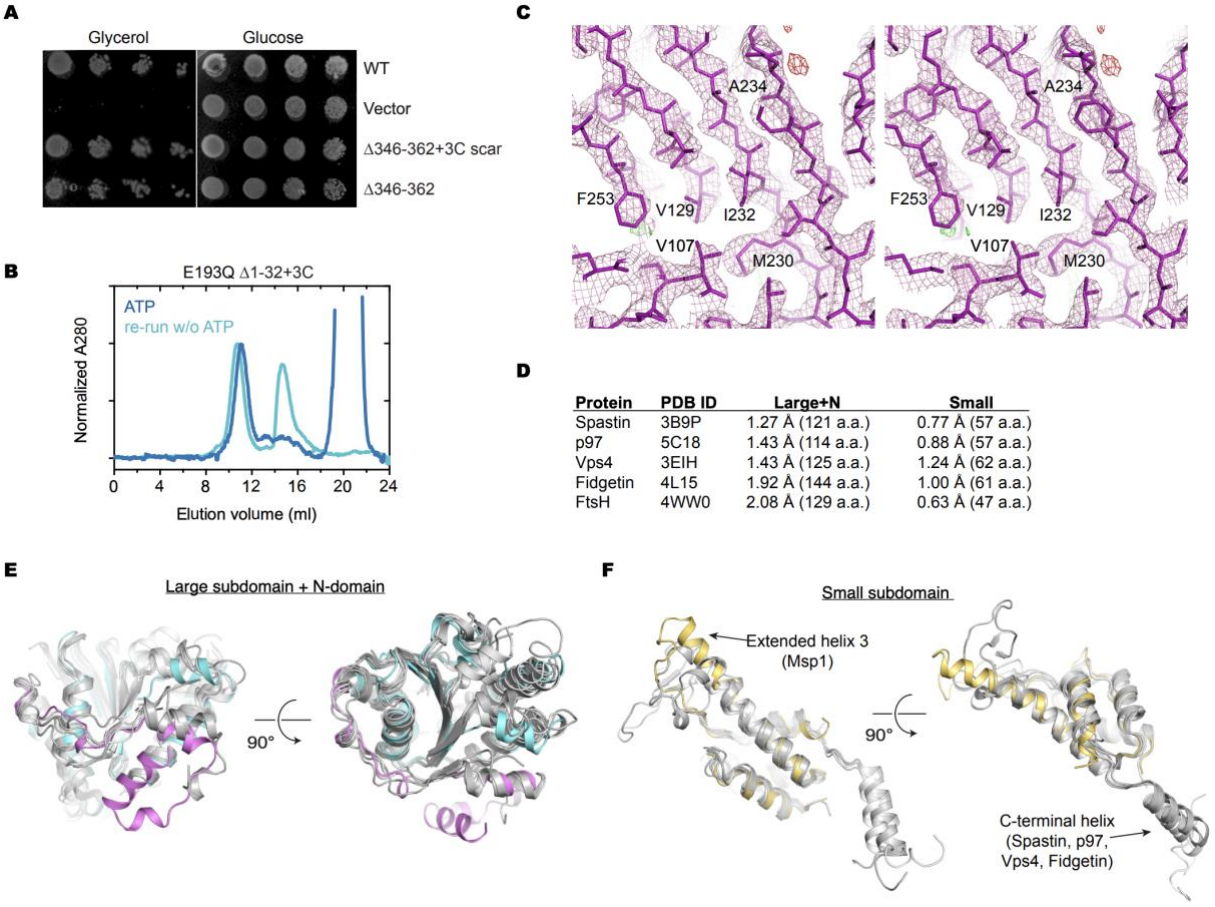


Figure 16. Characterization of the crystallization construct (D1-32+3C) and structural comparison. (A) Full-length Msp1 constructs truncated at position 345 and either containing (D346-362 + 3C scar) or lacking (D346-362) the 8-residue 3C protease cleavage site scar complements an Msp1 knockout in yeast. (B) Gel filtration analysis of an ATPase-deficient form of the crystallization construct following cleavage with 3C protease. This construct forms ATP-dependent hexamers (dark blue trace). Re-running the main peak in buffer lacking ATP demonstrates the reversibility of hexamer formation (light blue trace). (C) Cross-eyed stereo view of σA-weighted 2Fo-Fc (pink) and Fo-Fc (green, red) electron density maps, calculated at 2.6 Å and contoured at 1.5 and 3.5 sigma, respectively. A portion of the final refined model, including part of the central beta sheet in the Msp1 large subdomain, is superimposed (magenta sticks). (D) RMSD (on Cα) for five of the closest structural homologs of S. cerevisiae Msp1. The number of residues included in the alignment is given in parentheses. (E) Structural overlap of the Msp1 large subdomain (cyan) and N-domain (magenta), with the homologs in (D). (F) As in (E), but with the Msp1 small subdomain (yellow). The extended helix 3 of Msp1 is highlighted here; in some meiotic-clade AAA ATPases, this region contains functionally important insertions. The C-terminal helix of other meiotic-clade members is also highlighted; this region is not present in the Msp1 soluble construct crystallized in this work.

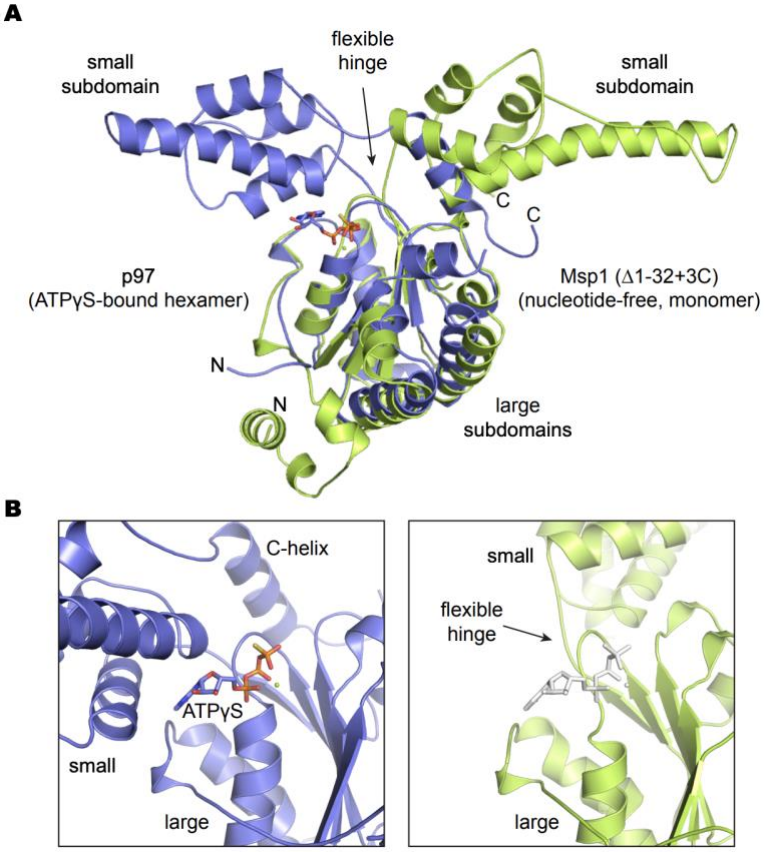


Figure 17. Related to Figure 15. Conformation of the Msp1 small subdomain observed in the nucleotide-free, monomeric crystal. (A) Structural overlap highlights the unusual ~180° rotation of the Msp1 (green) small subdomain observed in the monomeric, nucleotide-free crystal relative to its position in hexameric, ATPgS-bound p97 (blue). (B) In the crystallized conformation, the nucleotide binding site of Msp1 is only partially formed. Closeup of the composite ATP binding sites of p97 (left panel), which was crystallized as an ATPgS-bound hexamer, and Msp1 (right panel) which was crystallized as a nucleotidefree monomer; for reference, an ATPgS molecule was modeled into the monomeric Msp1 active site (grey). Whether the open conformation represent a physiologically relevant on-pathway conformation that could be exploited to regulate Msp1 activity, or simply reflects the interdomain flexibility seen in many AAA+ ATPases (here captured in an unusual conformation by crystal packing forces), remains to be clarified.

To further validate the model, we sought to disrupt the predicted hexamer interface by site-directed mutagenesis. We identified a hydrophobic interface between the large subdomain of one subunit and the small subdomain of an adjacent subunit (Figure 18D) and introduced a double leucine-to-aspartate mutation at positions 122 and 123. Using a soluble, ATPase-deficient construct containing the double mutation, we observed reduced hexamer formation in the presence of ATP (Figure 18E). Likewise, a full-length Msp1 construct containing the double mutation was unable to complement an msp1 deletion in vivo (Figure 18F) and failed to extract TA substrate in vitro (Figure 18G). Taken together, these data validate the structural model and indicate that Msp1 functions as a ring hexamer.

To locate functionally important surfaces, we mapped electro- static potential and sequence conservation from 94 fungal Msp1 homologs to the surface of the hexamer model. The positive charge character of the ‘‘bottom’’ surface of the hexamer (including the N-domain) is compatible with the negatively charged surface of the OMM, supporting assignment of this face as the membrane proximal surface of the Msp1 hexamer (Fig. 19). Most of the hexamer periphery, including the N-domain and small subdomain, shows limited sequence conservation (Figure 18C). In contrast, the ‘‘top’’ and bottom surfaces, particularly toward the center of the ring, are highly conserved, suggesting a role in the capture and extraction of TA substrates.

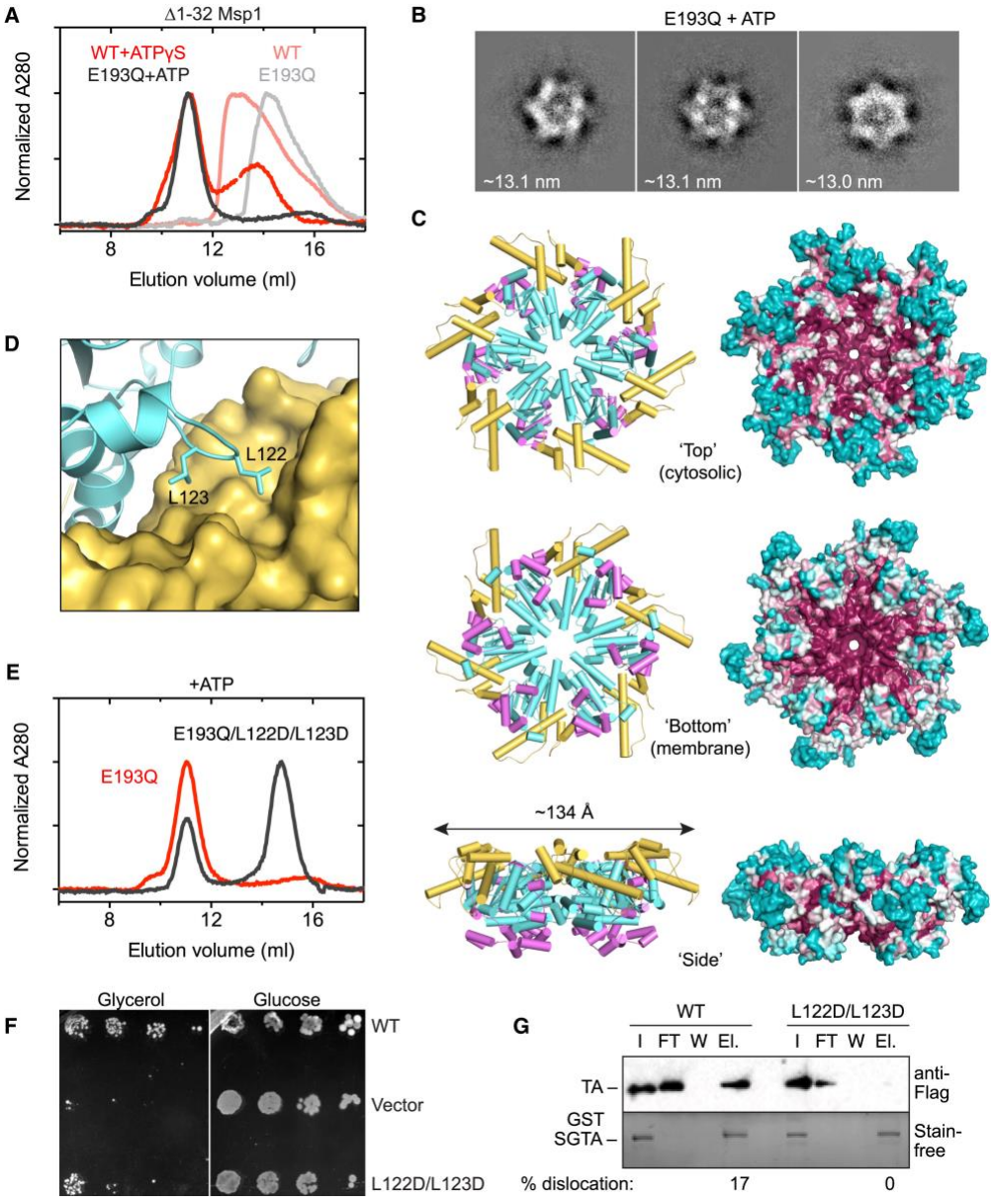


Figure 18. Msp1 Functions as a Ring Hexamer. (A) Size-exclusion chromatography (SEC) of wild-type (WT) D1-32 Msp1 and its ATPase-deficient mutant (E193Q) demonstrates ATPgS- and ATP-dependent hexamer formation (see also Figure 19). (B) Cryo-EM 2D class averages of D1-32 Msp1E193Q incubated with Mg^{2+}ATP reveal a hexameric structure with protrusions emanating from the central ring. (C) A hexameric model generated by separately superimposing the large and small subdomains of Msp1 with the corresponding subdomains in the D2 ring of hexameric, ATP-bound p97 (PDB: 5C18). Right, the sequence conservation of 94 fungal Msp1 proteins is mapped to the molecular surface of the Msp1 hexamer model from most (magenta) to least (cyan) conserved. Note, the sequence conservation is strongest on the ‘‘bottom’’ (membrane proximal) face of the hexamer and in the central pore. (D) Closeup of a portion of the predicted hexamer interface between hydrophobic residues in the large subdomain of one subunit (cyan) and the small subdomain of a second subunit (yellow). (E) SEC analysis shows that a double mutation within this interface (L122D/L123D) disrupts hexamerization of the soluble E193Q construct. The corresponding mutations in full-length, wild-type Msp1 also disrupt function *in vivo* and *in vitro*.

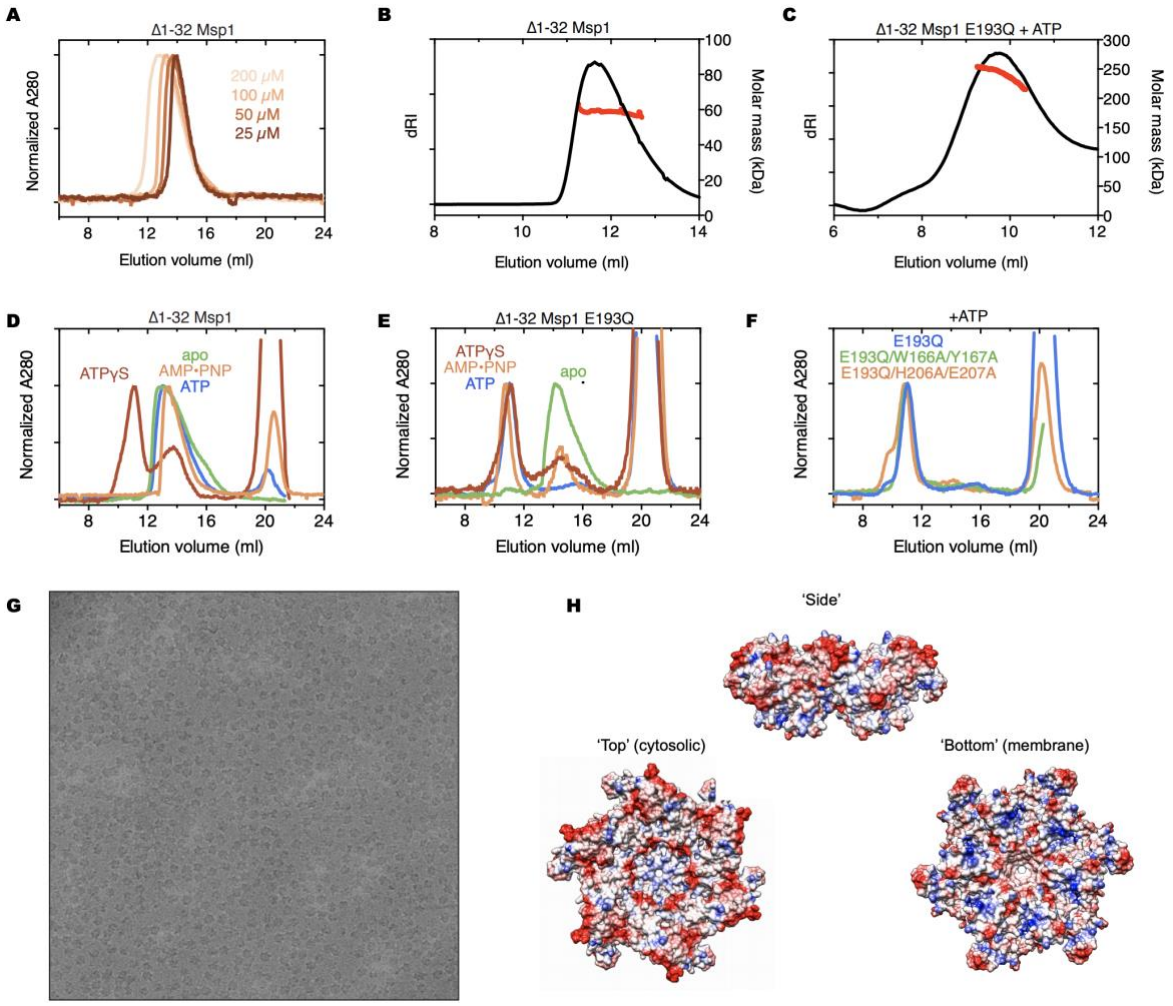


Figure 19. Additional analysis of soluble Msp1 hexamers. (A) SEC traces for the soluble region of Msp1 (D1-32) reveal a concentration-dependent shift towards a dimeric species in the absence of nucleotide. (B) SEC-MALS trace of soluble Msp1 shows a mixture of monomeric and dimeric species at 250 μM; the calculated mass of monomeric, soluble Msp1 is 39.1 kDa and the observed mass is 58 kDa. (C) SEC-MALS trace of the ATPase-deficient E193Q mutant of soluble Msp1 obtained in the presence of Mg^{2+}ATP. The observed molecular mass of 241 kDa agrees with the expected mass for a hexamer (235 kDa). Note that SEC was performed on an S200 Increase column at 4° C whereas SEC-MALS was performed on an S200 column at room temperature, so the elution volumes are not directly comparable. (D) SEC traces of soluble wild-type Msp1 in the presence of different nucleotides. Hexamers are observed in the presence of ATPγS but not ATP or AMP•PNP. (E) As in D, but with the E193Q mutant of Msp1; this construct forms hexamers in the presence of ATP as well as the non-hydrolzyable analogs ATPγS and AMP•PNP. (F) Pore loop mutations introduced into the soluble E193Q construct form hexamers in the presence of ATP. Note that ATP was included in the running buffer for the pore loop 1 mutant. (G) Single-particle cryo-EM of Msp1 hexamers. Representative raw micrograph of E193Q Msp1 (Δ1-32) particles in the presence of Mg^{2+}ATP. Hexameric particles with a diameter of ~13 nm are clearly visible in the raw image. Note that the majority of particles adopt a preferred orientation in the vitreous ice. (H) Electrostatic surface representation of the Msp1 hexamer model colored from negative (red) to positive (blue).

4.7 Functional Role of the Central Pore

Many AAA ATPases translocate polypeptide substrates through their central pore, a process that involves a highly conserved aromatic-hydrophobic-glycine pore loop motif[223]. In Msp1, the central pore is lined with conserved residues from three loops (Figure 20A). Loop 1 contains the canonical ‘‘WYG’’ sequence and is well defined in the crystal structure of the Msp1 monomer (Figure 20B). In the hexamer, loop 1 lies on the membrane proximal side, adjacent to the N-domain at the pore entrance. Loop 2, comprising mostly polar and charged side chains, lies at the center of the pore. Although residues 200–205 are disordered in the structure, loop 2 residues contribute to the main constriction point of the pore, separating the membrane-proximal and cytosolic-facing vestibules. Loop 3, enriched in negatively charged sidechains, lies on the cytosolic-facing side of the ring.

To explore the role of the Msp1 pore loops, we generated two double mutants—W166A/Y167A in loop 1 and H206A/E207A in loop 2—and tested their activity. In each case, the double mutants failed to complement an msp1 deletion in yeast, although the loop 1 mutant was more severe than the loop 2 mutant (Figure 20C). Likewise, both mutants failed to extract TA protein in vitro (Figure 20D). This loss of activity is not due to impaired hexamer formation, as soluble, ATPase-deficient constructs containing the double pore loop mutations formed hexamers in the presence of ATP (Fig. 19). Taken together, these data point to an essential functional role for the central pore, most likely involving TA substrate translocation during extraction from the membrane.

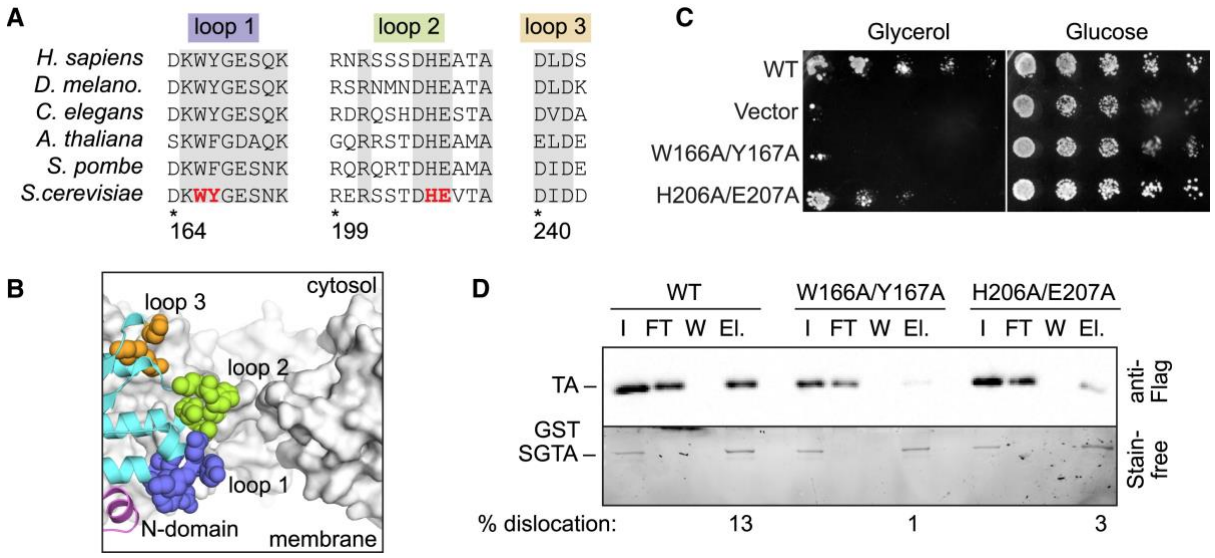


Figure 20. Central Pore Loop Mutations Disrupt TA Protein Dislocation. (A) Alignment of pore loop sequences in homologs from yeast Msp1 to human ATAD1. The most conserved residues are highlighted in gray; residues mutated for functional analysis in yeast Msp1 are in red. (B) Cutaway of the central pore in the hexamer model. Three subunits are shown, two in surface representation (gray) and the third in cartoon with pore loop residues shown as spheres. Note the constriction in center of the pore by residues in pore loop 2 (some of which are disordered in the crystal and thus not shown); this constriction separates the membrane proximal and cytosolic faces of Msp1. (C and D) Mutation of conserved residues in pore loops 1 and 2 disrupts activity of full-length Msp1 *in vivo* (C) and i*n vitro* (D) but does not disrupt ATP-dependent hexamer formation by the soluble domain.

4.8 Functional Role of the N-domain

The Msp1 N-domain comprises a lumenal sequence of ~10 residues, an ~20 residue TMD, and a cytosolic ~68 residue linker that tethers the AAA domain to the OMM (Figures 16A and 21A). Residues within these regions are not strongly conserved in Msp1, but the linker region has been implicated in substrate binding in related AAA ATPases[224,225]. Given this and its location at the entrance to the central pore, we sought to clarify the role of the Msp1 linker. We constructed a series of deletion mutants (Figure 21A) and found that the linker region is largely dispensable for function in vivo; indeed, removing the entire linker (D32-98) resulted in only a slight defect in

vivo (Figure 21B) and in vitro (Figure 21C). Thus, consistent with its lack of sequence conservation, the N-domain linker region of Msp1 is not strictly required for function.

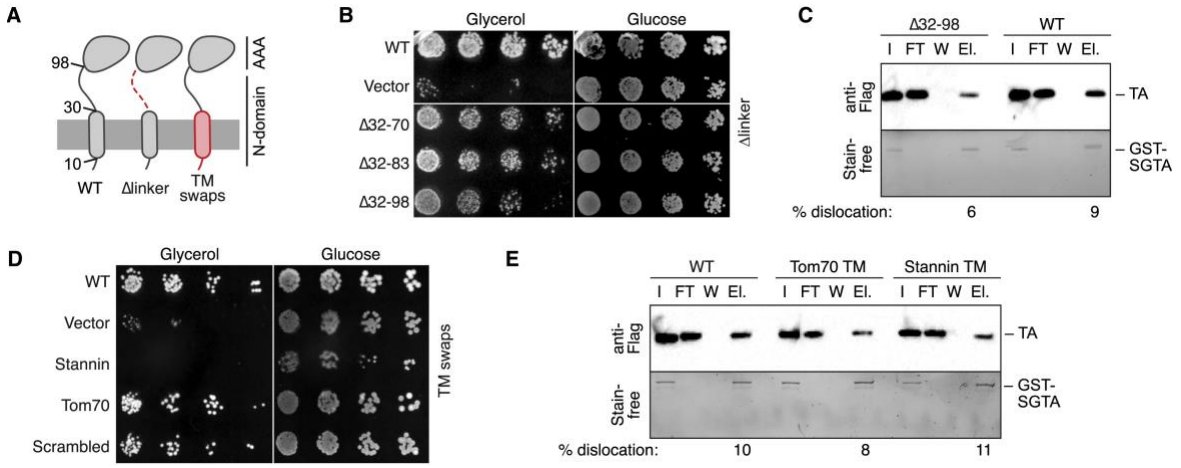


Figure 21. Role of the N-domain Linker and TM Regions. (A) Schematic of the N-domain constructs used for functional analysis. (B and C) Progressive deletion of the N-domain linker (‘‘Dlinker’’) connecting the Msp1 AAA domain to its TM has little effect on *in vivo* (B) or *in vitro* (C) activity; a slight reduction in activity is observed for the longest deletion (D32-98). Note that all samples were grown on the same plate. (D) Chimeric constructs (‘‘TM swaps’’), in which the Msp1 TM sequence is scrambled or swapped with the TM from yeast Tom70 or human Stannin, show differing effects on activity *in vivo*; whereas the scrambled and Tom70 TM chimeras are functional, the Stannin chimera mislocalizes (Figure S5) and thus fails to complement. (E) *In vitro*, however, the Tom70 and Stannin chimeras are both functional for extraction.

Next, we investigated the role of the N-domain TMD by generating chimeric constructs in which the Msp1 TMD and lumenal region were swapped with the corresponding regions from yeast Tom70 or human Stannin, unrelated single-pass membrane proteins of the OMM with the same topology as Msp1. The Tom70 chimera localized to the OMM (Fig. 22) and complemented the msp1 deletion in vivo (Figure 21D). In contrast, the Stannin chimera failed to target properly to the OMM and thus failed to complement in vivo (Figure 21D). However, using the in vitro

extraction assay, where localization is artificially imposed, we found that both chimeras were functional for extraction (Figure 21E). Thus, while TA protein extraction does not appear to depend on the precise sequence of the Msp1 TMD, the TMD is required for proper targeting and anchoring of Msp1 to the OMM.

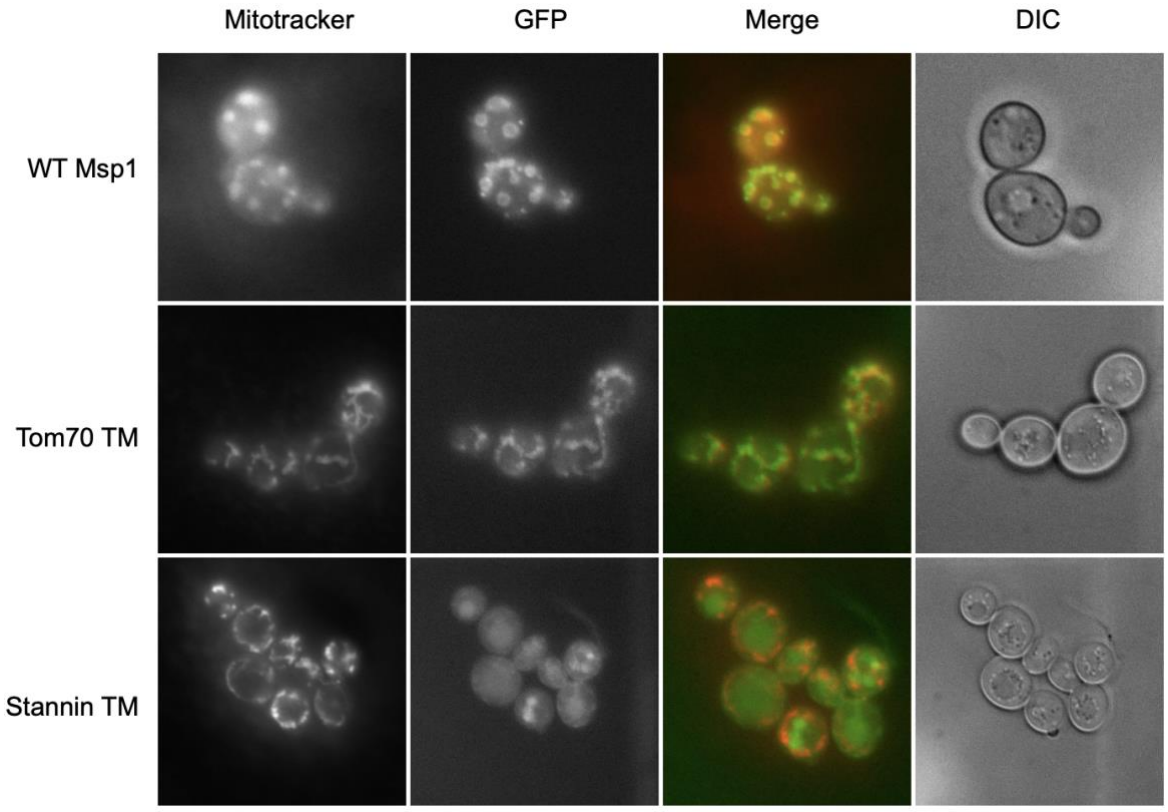


Figure 22. Localization analysis. The sub-cellular distribution of GFP-tagged Msp1 and its chimeric mutants (Tom70 TM and Stannin TM) were monitored by fluorescence microscopy in live cells. Wild-type Msp1 and the Tom70 TM chimera co-localize with mitotracker, but the Stannin TM chimera does not.

4.9 Discussion

Here we provide direct evidence that Msp1 extracts TA proteins from the membrane. Using reconstituted proteoliposomes containing only recombinant Msp1 and an integrated model ER

TA protein, we show that TA proteins are completely extracted from the bilayer in an ATPase-dependent process (Figure 13). This extraction activity is fully encoded in Msp1 and does not require additional factors or post-translational modification, including ubiquitination or proteolysis.

Our functional and structural data provide insight into the mechanism of extraction. First, Msp1 appears to function as ATP-dependent ring hexamer. Thus, mutations that disrupt the Msp1 hexamer in vitro also disrupt its ability to extract TA proteins from the membrane (Figure 18). Whether Msp1 forms a constitutive hexamer in vivo or assembles in a regulated process is not clear. An intriguing possibility is that the unusual open conformation observed in the monomeric, nucleotide-free crystal structure of Msp1 (Figure 15 and 17) represents an inactive ground-state conformation, allowing for regulated activation of Msp1 through conformational switching.

Second, the central pore appears to play a critical role in TA protein binding and/or translocation during extraction. Thus, mutations within the conserved pore loops, which do not impair hexamer formation, disrupt the ability of Msp1 to extract TA proteins from the bilayer (Figure 20). Third, despite its location near the entrance to the central pore, the N-domain linker region is largely dispensable for function (Figure 20). Finally, the N-terminal TMD is minimally required for localization and anchoring of Msp1 to the OMM. The observation that membrane protein extraction can occur in the presence of a heterologous TMD sequence (Figure 21) is similar to that reported for the bacterial AAA+ protease FtsH[226]. Notably, while dispensable for proteolysis of soluble clients, the TMDs of FtsH and the mitochondrial protease m-AAA[227] are required for extraction of their integral membrane protein clients. An intriguing possibility is that

the TMDs of these membrane-anchored dislocases form channel-like structures or perturb the bilayer around a client to facilitate extraction from the bilayer.

How Msp1 selects TA membrane protein clients for extraction remains unclear. This is a challenging task because sequence differences between the TMDs of mislocalized and endogenous TA proteins are likely to be subtle, and genuine clients are likely to be rare relative to the number of endogenous TA proteins. Although heterologous sequences can functionally substitute for the Msp1 TMD, the TMD may play an active role in substrate selection by sensing hydrophobic mismatch or surfaces that become exposed in the absence of appropriate binding partners. While our data show that Msp1 functions autonomously to extract TA proteins in vitro, additional factors may play a role in substrate se- lection in vivo. Systematic analyses are needed to define how the cell ensures that only genuine substrates are extracted; these studies will be facilitated by our in vitro assay.

The immediate fate of TA proteins following extraction also remains unclear. During ER-associated degradation (ERAD), the cytosolic chaperone Bag6 (BCL2-associated athanogene 6) is recruited to the ER to capture aggregation-prone clients extracted by the Cdc48/p97 AAA+ ATPase[228-231]. Intriguingly, the cytosolic-facing surface of the Msp1 hexamer is conserved, suggesting a potential binding site for a cytosolic chaperone to rapidly engage TA protein clients during extraction from the membrane.

In addition to extracting mislocalized TA proteins from the OMM, the Msp1/ATAD1 AAA ATPases may perform additional functions in the cell. This is suggested by the broad subcellular distribution of Msp1/ATAD1, including the OMM, peroxisomes, plasma membrane, and the cytosol. By analogy with the mem- brane-anchored AAA+ proteases FtsH, i-AAA, and m-AAA, whose clients include peripheral and integral membrane protein[227,232,233] ,

Msp1/ADAD1 AAA ATPases likely act on different classes of substrates. Consistent with this, putative non-TA protein clients have been identified for yeast Msp[213], while mammalian ATAD1 has been proposed to mediate internalization of AMPA receptor complexes by dissociating a peripherally associated cytosolic factor from the receptor[217]. Future studies are needed to define how Msp1/ATAD1 family members function in different cellular contexts.

5 Future Directions

5.1 Relating to Anghel et al., 2017 and McGilvray et al. 2019

The data presented in Anghel et al., 2017, and McGilvray et al., 2019 describe a set of unexpected human membrane protein biogenesis factors. We find that TMCO1, a previously uncharacterized member of the Oxa1 superfamily, is involved in the co-translational biogenesis of integral multi-pass membrane proteins. What's more, we find that it possesses the expected structural features of other Oxa1 superfamily members, including a soluble coiled-coil, and a core three TMD core. These data provide further evidence that the Oxa1 superfamily is comprised of membrane protein biogenesis factors. What's more, we assign ribosome binding and potential membrane protein biogenesis functions to four previously uncharacterized membrane proteins; CCDC47, NOMO, TMEM147, and Nicalin.

Improved understanding of the TMCO1-containing translocon would come from a higher resolution structure of the TMCO1-containing RNC-translocon complex. To achieve this, generation of RNC-translocon complexes containing stalled client proteins should be generated by in vitro translation, and isolated via a combination of IP and centrifugation. Similar methods have allowed for high resolution structures of different RNC-translocon structures to be generated[28,191]. Stabilization of this TMCO1-cointaining RNC-translocon complexes by inclusion of a stalled client may also allow for the identification of additional complex components which may have been lost from the natively isolated complex during purification.

Though this data describes new biogenesis machinery, many questions remains. For instance, what precise chemical features define a client of a TMCO1-containing translocon and the exact mechanism by which it facilitates multi-pass membrane protein biogenesis is still to be determined. Determining when TMCO1 engages different clients would allow us to compare the

biophysical properties of the regions TMCO1 recognizes. Ribosome profiling would allow a medium to high resolution analysis of the regions TMCO1 recognizes. This analysis could be confirmed biochemically by modifying TMCO1 client and non-client membrane proteins to include the biophysical features identified in a ribosome profiling screen and testing their dependence on a TMCO1-containing translocon for proper biogenesis in cells and in vitro.

Additionally, by examining the defects imparted by the loss of these new biogenesis factors we can learn more about the actual mechanism by which TMCO1-containing translocons facilitate biogenesis. It is possible to track biogenesis using in vitro translation into prepared microsomes. By studying changes in TMD insertion in WT and TMCO1 KO microsomes we may learn what particular ways TMCO1-conatining translocons insert TMDs. TMD insertion can be tracked by glycosylation, looking at the parts of a protein translocated across the membrane, or by pegylation of engineered cysteine residues, via PEG-maleimide, tracking which parts of the protein does not translocate across the membrane. To gain insight into which components perform what function, we can also crosslink newly translated membrane proteins to biogenesis machinery in vitro via either chemical methods, engineered cysteine pairs, or the inclusion of UV activatable amino acids at different locations in the nascent chains.

Our data also suggests that the translocon is a dynamic complex, with different biogenesis factors engaging the Sec complex in response to the needs of different client proteins. How then is the TMCO1-containing translocon recruited? One possibility is that it responds to changes in the confirmation of the ribosome or Sec61 complex. TMCO1, TMEM147, and CCDC47 all seem to have interactions with the ribosome. One could determine which component plays the primary role in recruitment by mutating or deleting the residues in their ribosome binding domains and examining changes in ribosome binding biochemically via IP, or

changes in client biogenesis in cells. Similarly, one could examine recruitment by mutating or deleting residues in components which contact the Sec61 complex and performing similar experiments. It may also be possible, through very careful single amino acid resolution crosslinking experiments, to determine the exact times in biogenesis hand off between different translocon accessory factors occurs. By extending the nascent chain one amino acid at a time and tracking crosslinking by any of the previously mentioned methods, it may be possible to see when the nascent chain is engaged by different translocon accessory factors.

Separately, our data provides further evidence that EMC3, another Oxa1 family member, is likely involved in membrane protein biogenesis. EMC3 is part of the EMC complex, a unbiquitous ER membrane complex in all eukaryotes. This complex has recently been shown to facilitate the biogenesis of both TA and multi-pass membrane proteins. It is possible that as an Oxa1 superfamily member, EMC3 is the catalytic core component of the EMC complex. Direct crosslinking of EMC3 to client proteins during insertion would provide added evidence that it directly functions during insertion. Additionally, it may be possible to use our understanding of Oxa1 family member structure to generate EMC3 mutants which do not disrupt complex assembly but do disable EMC insertase activity.

5.2 Relating to Wohlever et al. 2017

These data tell us that Msp1 is a functional hexamer which is capable of independently removing TA proteins form lipid membranes. It also tells us the TMDs of Msp1 are not inherently required for its dislocase activity. However, it does leave some questions on the mechanism of activity open.

While Msp1 can recognize TA proteins and dislocate them from the membrane in an ATP dependent manner, we do not have a clear picture as to how it recognizes mis-localized TA proteins. Recent work has indicated that Msp1 may recognize mislocalized TA proteins through a combination of hydrophobic and charged patches on its cytosolic membrane facing side and short luminal coil respectively. This explanation would preclude the need for its TMDs in substrate recognition.

However, what happens to the TA proteins Msp1 removes from the membrane. Surprisingly, it appears that their degradation is controlled by the ER ubiquitin ligase Doa10. But TA proteins cannot float through the cytosol unchaperoned lest their TMDs aggregate and become toxic to the cells. Therefore, I would theorize that the TA proteins Msp1 removes are bound by chaperones in the cytosol and brought to Doa10 for degradation. To identify these chaperones one could over express tagged TA proteins in a GET pathway deficient cell line, and isolate the mitochondrial, ER, and cytosolic fractions. IP on the tag and mass spec of the TA protein and co-isolated proteins should how what proteins interact with dislocated TA proteins.

6 Materials and Methods

6.1 Protocols Relating to Anghel et al. Cell Reports (2017)

6.1.1 Assay for In Vivo Association of TMCO1 with Ribosomes

The total HEK293 cell membrane fraction (in assay buffer: 150 mM potassium acetate; 50 mM HEPES [pH 7.4]; and 5 mM magnesium acetate) was solubilized by addition of recrystallized digitonin (Calbiochem; lot no. 2913883) from a 5% stock to a final concentration of 2%. Solubilization was allowed to proceed for 30 min at 4C with end over end mixing and then insoluble material was removed by centrifugation for 10 min at 10,000 3 g. The soluble material was then layered over a 1-mL sucrose cushion (150 mM KCl, 50 mM Tris [pH 7.4], 5 mM MgCl2, 1 M sucrose, and 0.1% digitonin). Sucrose cushions were pelleted for 2 hr at 250,000 3 g in a TLA100.3 rotor (Belin et al., 2010). The pellet was resuspended in the same buffer, re-run over a cushion again, and then finally the resuspended pellet was pelleted through a gradient (10%–50% sucrose, 150 mM potassium acetate, 50 mM Tris [pH 7.5], 5 mM MgCl2, and 0.1% digitonin) at 130,000 3 g (SW28.1; Beckman Coulter) for 12 hr at 4C. 900 mL fractions were collected manually from the top of the gradient, TCA precipitated, and analyzed by SDS-PAGE.

6.1.2 Co-immunoprecipitation Analyses

For co-immunoprecipitations from canine pancreatic membranes (Promega), the membranes were resuspended in a buffer containing 250 mM potassium acetate, 50 mM HEPES (pH 7.4), 5 mM magnesium acetate, 15% glycerol, and 3% recrystallized digitonin (Calbiochem; Kun et al., 1979). Solubilization was allowed to proceed for 30 min on ice, and then insoluble material was removed by centrifugation for 10 min at 10,000 3 g. The soluble material was then

divided equally and layered over protein A resin that had been crosslinked to antibodies against TMCO1, Sec61b, or 3F4 (as control). Immunoprecipitation (IP) reactions were incubated for 2 hr at 4C with end-over-end mixing and then washed six times with 250 mM potassium acetate, 50 mM HEPES (pH 7.4), 5 mM magnesium acetate, 15% glycerol, and 0.1% digitonin. Bound proteins were eluted by three successive 10-min incubations with 1 M glycine (pH 3) supplemented with 0.1% Fos-choline-12. Elutions were then TCA precipitated, resuspended in Laemmli Sample Buffer, and analyzed by SDS-PAGE.

For co-immunoprecipitations from 33Flag-TMCO1 HEK293 TRex cells, the membrane fraction was isolated and washed twice with 250 mM potassium acetate, 50 mM HEPES (pH 7.4), 10 mM magnesium acetate, and 250 mM sucrose. Membranes were then resuspended in buffer containing 250 mM sucrose, 300 mM potassium acetate, 50 mM HEPES (pH 7.4), and 10 mM magnesium acetate. Solubilization was allowed to proceed for 30 min on ice, and then insoluble material was removed by centrifugation for 10 min at 10,000 3 g. The soluble fraction was then added to anti-Flag M2 resin (Sigma) and incubated for 1 hr at 4C with end-over-end mixing and then washed four times with 350 mM potassium acetate, 50 mM HEPES (pH 7.4), 5 mM magnesium acetate, 250 mM sucrose, and 0.1% digitonin. Bound proteins were eluted by 2 successive 30-min incubations with same buffer as the wash but supplemented with 0.5 mg/mL 33Flag peptide (ApexBio). The elutions were then layered over a 1-mL sucrose cushion (150 mM KCl, 50 mM Tris [pH 7.4], 5 mM MgCl2, 1 M sucrose, and 0.1% digitonin) and then pelleted for 2 hr at 250,000 3 g at 4C in a TLA100.3 rotor (Belin et al., 2010). The supernatant was discarded, and the ribosome pellet was resuspended in 23 Laemmli Sample Buffer for analysis.

6.1.3 Antibodies

Antiserum against human TMCO1 was generated by Lampire Biologicals. Rabbits were immunized with a KLH conjugated EKKKETITESAGRQQKK peptide, located in the cytosolic coiled-coil of TMCO1. Exsanguination bleed was supplemented with 0.02% sodium azide, flashfrozen in liquid nitrogen and stored at -80°C. For immunoprecipitation experiments, antibody was thawed and used immediately without further purification. For western blotting, initial experiments used unpurified serum; other experiments used peptide affinity purified antibody.

Antibodies against L17 (Abgent), S16 (Santa Cruz) and Derlin-1 (Abcam) were purchased, and antibodies against Sec61a and Sec61b were characterized previously (Gorlich et al., 1992).

6.1.4 Cell culture

HEK293-Cas9 cells containing a 3xFlag-Cas9 construct integrated into the genome were generated from HEK293 Flp-In T-REx cells (Invitrogen). A TMCO1 knockout line derived from these cells was generated at the Genome Engineering Core Facility at the University of Chicago, using a guide RNA with the sequence 5'-GAAACAATAACAGAGTCAGCTGG-3'. Cas9 expression was induced by addition of doxycycline at 10 ng/mL, followed by transfection of a gRNA-expressing plasmid. Single cells were then seeded into 96 well plates allowed to grow clonally. The final TMCO1 knockout line was verified by both genomic DNA sequencing and immunoblotting with an α-TMCO1 antibody.

A separate cell line containing an N-terminally Flag tagged TMCO1 was also generated at the same facility using a previously described two step strategy (Xi et al., 2015). The resulting

cell line has one nonfunctional TMCO1 allele and one allele containing a 3xFlag-tagged TMCO1 with a 13 amino acid linker (ITSYNVCYTKLSG, from the Cre-lox recombination) before the TMCO1 ORF. The 3xFlag-TMCO1 lines were verified by both genomic DNA sequencing and immunoblotting with α-TMCO1 and α-Flag antibodies.

Cells were grown in DMEM supplemented with 10% Fetal Bovine Serum (Gemini Benchmark; Lot #A99D00E) and penicillin/streptomycin mixture (Invitrogen). The culture medium was also supplemented with 15 µg/mL Blasticidin and 100 µg/mL Hygromycin B for the TMCO1 knockout and 3xFlag-TMCO1 cell line generation procedure, but not when growing cells for other applications.

6.1.5 Isolation of total membrane fraction from HEK293 cells

Cells were harvested at a density of 70-100% while growing. Media was removed and cells were scraped into DPBS. Cells were collected by 5 min at 500 x g centrifugation at 4°C, and then lysed osmotically (Sabatini, 2014) by resuspending in a volume of 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. Cells were allowed to swell on ice for 15 minutes, followed by 15 strokes of a douncer with a tight-fitting pestle (Kontess). Sucrose was then added to 250 mM to balance osmolarity. Nuclei were then removed by pelleting 3 minutes at 700 x g, and the supernatant was centrifuged 10 minutes at 10,000 x g to collect the membrane fraction. 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. The membranes were then washed with assay buffer (150 mM potassium acetate, 50 mM Hepes pH

7.4, 5 mM magnesium acetate) and centrifuged again 10 minutes at 10,000 x g to remove any residual cytosolic proteins.

Membranes used for sucrose cushions, gradients, and pull-downs were further treated with micrococcal nuclease to digest polysomes as follows: reaction was supplemented with calcium acetate to 1 mM and 100 Units of micrococcal nuclease (NEB), incubated 10 minutes at 25°C, and then quenched by addition of EGTA to 2 mM. Membranes were then washed again with assay buffer to remove nuclease.

6.1.6 Recombinant TMCO1 production

The gene encoding human TMCO1 was amplified by PCR from total human testicular cDNA (Biosettia), subcloned into a pET28b vector (Novagen) encoding an N-terminal 6xHis tag followed by a TEV protease site, and verified by DNA sequencing. TMCO1 encoding vectors were transformed into E. coli BL21(DE3) and colonies from these transformations were used to inoculated terrific broth (TB, Fisher) starter cultures in baffled flasks containing 50 µg/mL kanamycin. 50 mL starter cultures were grown overnight at 37°C and 250 rpm. 1 L TB cultures containing 50 µg/mL kanamycin were inoculated with 3 mL of starter culture, grown at 37°C, and shaken at 250 rpm until they reached an A260 of 0.6. Expression was induced by addition of 0.1 mM isopropyl-β-d-thiogalactoside (IPTG, Sigma) and growth was continued for 4 hrs at room temperature and 250 rpm. Cells were harvested by centrifugation and pellets frozen at -80°C.

Frozen cell pellets were resuspended in 35 mL ice cold lysis buffer (500 mM NaCl, 50 mM Hepes pH 7.5, 10 mM imidazole pH 7.5, 20 µM EDTA pH 8, 1 mM PMSF, 2 mM DTT, 5% glycerol (v/v)) supplemented with 10 µg/mL DNaseI and 0.5 mg/mL of lysozyme. Resuspended

pellets were dounced five times on ice and lysed by passages twice through a high pressure microfluidizer. Lysate was clarified by centrifugation at 18,500 x g for 45 min at 4°C. To pellet bacterial membranes, the crude lysate supernatant was subjected to centrifugation at 120,000 x g for 1 hr at 4°C. Pelleted membranes were resuspended gently with a paintbrush in 40 mL ice cold lysis buffer, supplemented with 1% Decyl Maltose Neopentyl Glycol (DMNG, Anatrace), and incubated overnight (~14 hrs) at 4°C with gentle end-over-end mixing. Detergent soluble material was isolated by centrifugation at 120,000 x g for 1 hr at 4°C and batch purified by TALON affinity chromatography (Clonetech). The column was washed with 10 column volumes of lysis buffer supplemented with 15 mM Imidazole pH 7.5 (25 mM Imidazole total) and 0.07% DMNG. Protein was eluted in elution buffer (500 mM NaCl, 50 mM Hepes pH 7.5, 2 mM DTT, 300 mM imidazole pH 7.5, 0.07% DMNG) and further purified by size exclusion chromatography (Superdex 200, 10/300 GL, GE Healthcare) in 500 mM NaCl, 50 mM Hepes pH 7.4, 2 mM DTT, 0.07% DMNG at room temperature. Desired fractions were pooled and concentrated in a 50 kDa MWCO Amicon ultra centrifugal filter (Millipore). 10% glycerol was added before flash freezing and storage in aliquots at -80°C. Protein concentration was determined by Bradford assay.

6.1.7 Assays for in vitro association of TMCO1 with ribosomes

High-salt stripped ribosomes were prepared from rabbit reticulocyte lysate (Green Hectares Farm). After supplementing with 350 mM KCl, the lysate was layered on top of a high density, high salt sucrose cushion (1 M sucrose, 500 mM KCl, 50 mM Tris pH 7.4, 5 mM MgCl2), and subjected to centrifugation at 250,000 x g for 2 hrs at 4°C (TLA100.3, Beckman-Coulter). After incubating the pellet with ribosome buffer (250 mM sucrose, 150 mM KCl, 50

mM Tris pH 7.4, 5 mM MgCl2) for 1 hr on ice, KCl was added to 500 mM and ribosomes were again pelleted through a high density, high salt sucrose cushion. Ribosome pellets were gently resuspended in ribosome buffer, aliquoted, and flash frozen for storage at -80°C.

Ribosome binding assays were carried out in binding buffer (150 mM KCl, 50 mM Tris pH 7.4, 5 mM MgCl2, 0.07% DMNG), with 100 nM purified rabbit reticulocyte ribosomes and a 10-fold molar excess (1 µM) of purified, recombinant TMCO1 in a total volume of 100 µL. After incubating for 1 hr at 4 °C, 80 µL of the binding reaction was pelleted through a sucrose cushion (1 M sucrose, 150 mM KCl, 50 mM Tris pH 7.4, 5 mM MgCl2, 0.07% DMNG) for 2 hr 250,000 x g at 4°C (TLA100.3, Beckman-Coulter). Pellets were washed with 1 mL of ice cold water and resuspended in 40 µL of 1x lithium dodecyl sulphate sample buffer supplemented with 100 mM β-mercaptoethanol. Competition assays were performed as described above, but with the addition of either tRNA or polyA RNA at the indicated concentrations before incubation.

In vitro crosslinking was performed by adding fresh DSP (in DMSO) to a final concentration of 250 µM, followed by incubation for 10 minutes at room temperature. Reactions were quenched by the addition of Tris pH 7.4 to a final concentration of 100 mM, followed by an additional 10 min incubation on ice. NaCl was added to 500 mM to dissociate uncrosslinked TMCO1 from the ribosome. To separate ribosomal subunits after crosslinking, samples were incubated with 2 mM puromycin and 1 mM PMSF for 30 min on ice, then 20 minutes at 37°C. Crosslinked, puromycin-treated samples were separated by centrifugation through a high salt sucrose gradient (10-50% sucrose, 500 mM NaCl, 50 mM Hepes pH 7.5, 0.07% DMNG, 5 mM MgCl2) at 130,000 x g (SW28.1, Beckman-Coulter) for 14 hrs at 4°C. 1 mL fractions were collected manually from the top of the gradient, TCA precipitated, and analyzed by SDS-PAGE.

6.2 Protocols Relating to McGilvray et al. in submission (2019)

6.2.1 Cell Culture and RNAi

HEK293 cells were cultured with DMEM Complete (DMEM supplemented with 4.5g/L glucose, L-glutamine, sodium pyruvate, 10% FBS, 100 U/ml Penicillin, and 0.1mg/ml streptomycin) in 10 or 15cm cell culture treated dishes at 37^{o}C, 5% CO_2.

RNAi knockdown was performed in 6-well plates, with ~500,000 cells/well. Plates we first treated with 0.02mg/ml poly-L-lysine in DiH_2O for 15min at RT, the solution being swirled around the whole well. Wells were then washed 3x with 2ml DiH_2O and allowed to dry for at least an hour open in the hood. Cells were added dropwise to treated wells, swirled 20x clockwise, counter clockwise, forwards/backwards, and side to side, then allowed to adhere overnight in incubator. Don't look at the cells before the next morning.

30pmol of siRNA in 150ul of OptiMEM media was mixed with 2.5ul of Lipofectamine RNAiMAX in 150ul OptiMEM, and incubated for 10min at RT. Media in the wells were replaced with 2ml DMEM Complete, then the siRNA-Lipofectamine-Optimem solution was added drop wise to all areas of the well. The solution in the well was swirled 20x clockwise, counter clockwise, forwards/backwards, and side to side. Cells were allowed to incubate for 24hrs in the incubator. Then the media was replaced and the siRNA transfection was performed again. After another 24 – 48hrs, cells were harvested.

6.2.2 Isolation of native TMCO1 containing translocon-RNC complexes

T-Rex-293 cells expressing 3xFlag-TMCO1 from the endogenous TMCO1 promoter were culture on 10cm plates for at least 2 splits post-thawing and to at least 80% confluency. Cell were removed from plates by pipetting and put into 25ml of suspension culture media (Freestyle

293, supplemented with 10mM Hepes pH 7.5, 10mM L-Glutamine, 0.3µg/ml Penicillin, 0.5µg/ml Streptomycin, and 1% FBS) in a 125ml PETG square media bottle. Cells were grown at 37°C, 5% CO_2, and shaking at 135rpm to a density of $1x10^6$cells/ml. Cells were then expanded into a 1L PETG square media bottle in 250ml suspension culture media to a density of $1x10^6$cells/ml. Cells were harvested by centrifugation and washed in ice cold 1x PBS. Cells were resuspended in ice cold hypotonic lysis buffer (10mM Hepes pH 7.4, 10mM KoAc, 1mM $MgCl_2$) and incubated on ice for 15min. Cells were lysed with 20 strokes of a pre-chilled dounce tissue grinder on ice. 250mM sucrose was added to lysate and nuclei were removed by centrifugation at 700xg for 3 min. The supernatant was kept on ice as the microsome fraction and 1mM PMSF was added. The pelleted nuclei were washed with 1ml ice cold assay buffer (250mM Sucrose, 250mM KoAc, 50mM Hepes pH 7.4, 10mM $MgCl_2$). Nuclei were removed by centrifugation as before and the wash supernatant was combined with the microsome fraction. Microsomes were sedimented at 10,000xg 10min 4°C and resuspended in assay buffer to an A260 of ~30. In order to convert polysomes into monosomes, and degrade aberrant DNA, microsomes were supplemented with 1mM $CaCl_2$, and treated with 10,000U/ml Micrococcal Nuclease and 5U/ml DNase I for 15min at room temperature. Microsomes were then treated with 2mM EGTA, and separated from the nucleases by centrifugation as before.

Nuclease treated microsomes were solubilized in ice cold assay buffer containing 2.5% digitonin for 15min on ice. Unsolubilized material was removed by centrifugation at 10,000xg 10min at 4°C. Affinity purification of TMCO1 containing translocon-RNC complexes was carried out by incubating the solubilized material with anti-Flag resin for 1hr at 4'C with gentle end-over-end mixing. Flag resin was separated from the flow through by centrifugation in a picofuge. The resin was washed twice with 5 bed volumes of wash buffer (250mM Sucrose,

350mM KoAc, 50mM Hepes pH 7.4, 10mM $MgCl_2$, 0.25% Digitonin), and twice with 5 bed volumes of assay buffer supplemented with 0.25% Digitonin. Samples were eluted in two successive 30min incubations with two bed volumes of ice cold assay buffer supplemented with 0.25% Digitonin and 0.5mg/ml 3xFlag peptide, at 4°C with gentle end-over-end mixing.

TMCO1 containing translocon-RNC containing complexes were isolated by centrifugation through a 300ul sucrose cushion (0.5M Sucrose, 150mM KoAc, 50mM Hepes pH 7.4, 5mM $MgCl_2$, 0.25% Digitonin) in a TLA120.1 rotor at 355,000xg for 45min, 4°C. The pellet was resuspended in ribosome buffer (150mM KoAc, 50mM Hepes pH 7.4, 5mM $MgCl_2$, 0.25% Digitonin), and the concentration brought to ~100nM as judged by A260.

6.2.3 Cryo-EM Sample Preparation and Data Acquisition

Clean ~8cm by ~4cm mica sheets were held in a vacuum of at least $5x10^{-5}$ mbar and coated with a 3-4nm thick layer of amorphous carbon, as judged by a quartz crystal thickness monitor. After at least 1 day of maturation, this carbon was floated off of the sheets in distilled waters and used to batch coat Quantifoil 1.2/1.3 200 mesh EM grids. Grids were allowed to dry for at least one day before use.

Carbon coated grids were glow discharged for 30seconds immediately before use. Using an FEI Vitrobot, 2.5ul of ~100nM TMCO1 containing translocon-RNC sample was applied to each grid, incubated for 10 seconds at 22°C and 100% humidity, before blotting for 11seconds and flash freezing in liquid ethane.

Data were collected on a FEI Titan Krios at 300 KV using Latitude S (Gatan) at defocus values of 1-3 µm. Images were recorded using a K2 Summit direct electron detector in super

resolution mode with a calibrated magnification of 105,000x (super resolution pixel size of 0.66Å^2) and a total exposure of 50e$^-$/Å^2 fractionated over 40 frames (5e$^-$/Å^2/s).

6.2.4 Image processing

1,502 super resolution movies were motion corrected using Motioncor2 with 2x binning, generating corrected micrographs with a pixel size of 1.32 Å^2. Contrast transfer function parameters were estimated using GCTF. Particles were picked using the semi-autonomous particle picking algorithm in Relion3.0. All 2D classification, 3D classification, and refinements were performed in Relion3.0. Reference-free 2D classification was used to discard non-ribosome particles, resulting in 97,022 ribosome containing particles. 3D refinement against an 80S human ribosome (EMD-5592) low pass filtered to 60 Å was used to generate an initial model for further classification. 3D classification of all refined particles was performed using a mask around the 60S and translocon components. Particles from a 3D class showing density for the TMCO1 coiled-coil were subjected to focused refinement and further rounds of 3D classification without alignment using a mask containing only the translocon. Final particles were subjected to an unmasked local refinement to generate the final map.

6.2.5 Crosslinking-Mass Spec Sample preparation

Pelleted, purified TMCO1 containing translocon-RNC complexes were resuspended in ribosome buffer to a final concentration of 0.5mg/ml. Sample was crosslinked with 500μM disuccimidyl suberate for 30min at 35^{o}C with intermittent agitation. Crosslinking was quenched by the addition of 100mM Tris pH 7.5 and incubated on ice for 10min. Samples for

uncrosslinked and crosslinked material was kept for analysis by western blot and negative stain EM. Crosslinked samples were TCA precipitated and pelleted at 21,000g for 10min.

The precipitate was brought up in 8M Urea, 10 mM TCEP, heated at 56° C for 20 min, and then alkylated with 15 mM iodoacetamide (30 min at room temperature) which was quenched with 15 mM dithiothreiotol (15 min at room temperature). The sample was diluted to 2M Urea and digested overnight with 1 µg trypsin (Promega Gold) for 4 hours at 37° C. A second aliquot of 1 µg trypsin was then added and digestion was allowed to proceed overnight. The digestion mixture was acidified to 0.5% TFA and diluted 6-fold prior to desalting on a Peptide C18 MacroTrap column (Michrom Bioresources) controlled by Akta Purifier (GE Healthcare Life Sciences) and evaporated to dryness. Crosslinked products were brought up in 10 µl of SEC buffer (70:30 H_2O:ACN with 0.1% TFA) and enriched by size-exclusion chromatography (Superdex Peptide, GE Healthcare Life Sciences) as in (Leitner et al., 2012). 100 µl fractions eluting between 0.9 and 1.4 ml were dried, resuspended in 0.1% formic acid for MS analysis. The fractions starting at 0.9 ml and 1.3 ml were combined prior to evaporation to make four MS fractions.

6.2.6 Mass Spectrometry

Samples were reconstituted in 5 µl of 0.1% formic acid for mass spectrometry. LC-MS analysis was performed with an Orbitrap Fusion Lumos mass spectrometer (Thermo Scientific) coupled with a nanoelectrospray ion source (Easy-Spray, Thermo) and M-Class NanoAcquity UPLC system (Waters). Crosslink enriched fractions were separated on a 50 cm x 75 µm ID PepMap C18 column (Thermo). 2.5 µl of sample was loaded onto the column and eluted running a gradient from 3.5% solvent B (A: 0.1% formic acid in water, B: 0.1% in ACN) to 25% B in

175 minutes followed by a second gradient to 30% B over 10 minutes. Precursor scans were acquired in the Orbitrap from 375-1500 m/z (resolution: 120000, AGC Target: 4.0e5, max injection time: 50ms). Precursor ions were selected for dissociation using the following criteria: peptide monoisotopic precursor determination, charge state between 3-9, intensity greater than 5e4, and a 30 sec dynamic exclusion window. Each precursor that passed the selection criteria was subjected to subsequent HCD and ETD MS2 scans (resolution: 30000, quadrupole isolation window: 1.6 m/z units, HCD NCE: 28%, HCD AGC Target: 1.0e5, HCD max injection time: 150 ms, ETD collision time: calibrated charge dependent ETD parameters, ETD supplemental activation: 10% EThcD, ETD AGC Target: 2.0e5, ETD max injection time: 200 ms). Nine product ion scans of each type were performed for each precursor scan.

Crosslinked Peptdie ID and analysis

Separate peaklists were generated for ETD and HCD scans using Proteome Discoverer 2.2 (Thermo) and searched using Protein Prospector 5.23.0 (Trnka et al., 2014). The search database consisted of the sequences of 82 human ribosomal protein components in addition to 10 sequences corresponding to the membrane associated components: TMCO1, NCLN, NOMO1, NOMO2, NOMO3, S61A1, S61A2, SC61B, SC61G, TM147 and CCD47. The sequence of TMCO1 contained the N-terminal Flag tag (reported crosslinked residue numbers reference the endogenous sequence). The three NOMO isoforms are highly homologous and in most cases cross-links to NOMO could not be assigned a specific isoform. These proteins were confirmed to be the dominant components of the sample by MS analysis of late eluting SEC fractions (corresponding to linear peptides). The 92 target proteins were concatenated with a decoy database consisting of 10 randomized amino acid sequences of for each target sequence (1012 total protein sequences searched). ETD peaklists were searched using Prospector instrument type

ESI-ETD-high-res and HCD peaklists were searched using ESI-Q-high-res. Other search parameters were: mass tolerance of 7 ppm (precursor) and 15 ppm (product); fixed modifications of carbamidomethylation on cysteine; variable modifications of peptide N-terminal glutamine conversion to pyroglutamate, oxidation of methionine, and "dead-end" modification of lysine and the protein N-terminus by semi-hydrolyzed BS3, protein N-terminal acetylation, protein N-terminal methionine loss, and incorrect monoisotopic precursor selection (neutral loss of 1Da); crosslinking reagent was DSS/BS3; trypsin specificity was used with 3 missed cleavages and three variable modifications per peptide were allowed. The top 85 product ion signals were used for the search. Searches were performed using 64 cores on an HPC cluster and took about 4 hours to complete.

Cross-link spectral matches (CSM) were initially kept with peptide scores above 20, score difference above 0, and length of each peptide between 4-25 residues. A linear support vector machine (SVM) model was constructed to classify CSMs between decoy and target classes (Trnka et al., 2014). Features selected for the SVM classifier were: score difference, percent of ions matched, precursor charge state, rank of each peptide, and length of each peptide. Models were trained on half of the dataset and parameters were chosen to give a specificity of 90% tested on the other half of the data. Separate classifiers were built for ETD and HCD results. The best scoring CSM per unique cross-linked residue pair was selected and the ETD and HCD results were merged. The distribution of cross-linked residue pairs with one and two incorrectly identified peptides was modeled using essentially the same logic as (Fischer et al., 2017) but extending their analysis to account for the 10x increased size of decoy database. The number of target-target hits with one wrong peptide is given by:

$$\mathrm{tf(TT)} = \frac{1}{\mathrm{k}}\,\mathrm{TD} - \frac{2}{\mathrm{k}^2}\mathrm{DD}$$

and the number with both wrong is given by:

$$\mathrm{ff(TT)} = \frac{1}{k^2}\,\mathrm{DD}$$

where TT, TD, and DD are the number of target-target, target-decoy, and decoy-decoy hits, and k is the scaling factor describing the ratio in size of the decoy database to size of the target database. In this case k=10. The final list of cross-links was reported at an SVM score of 1.5 which corresponded to a 0.55% FDR. Distance analysis was performed by measuring the Cα-Cα distances between all ribosome cross-links against an EM reconstruction of the human 80S Ribosome (pdb code 4ug0). At the reporting threshold of 1.5, the violation rate (fraction of mappable cross-links > 35Å) was 10.2%.

Annotated spectra corresponding to the reported crosslinks are available at the MS-Viewer website (http://msviewer.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer) with the following accession keys:

HCD data: 7s2yb4zfjw

ETD data: vdibnsypj7

Raw data files are available through the MassiVE repository:

Accession: [not yet deposited]

6.3 Protocols Relating to Wohlever et al. Molecular Cell (2017)

6.3.1 Production of Soluble Protein Constructs - Msp1

The gene encoding the soluble region of S. cerevisiae Msp1 (D1-32) was PCR amplified from genomic DNA and subcloned into a pET28a derivative (Novagen) encoding an N-terminal 6xHis tag followed by a TEV or 3C protease cleavage site. All insertions and deletions were

performed by standard PCR techniques. Site-specific mutagenesis was carried out by QuickChange PCR. All constructs were verified by DNA sequencing.

Plasmids encoding the soluble region of S. cerevisiae Msp1 (D1-32) or its mutants were transformed into E. coli BL21(DE3) containing a pRIL plasmid and expressed in terrific broth at 37C until an OD600 of 0.6-1.0, cultures were induced with 1 mM IPTG and grown at room temperature for an additional 3-4 hr. Cells were harvested by centrifugation, and resuspended in Msp1 Lysis Buffer (20 mM Tris pH 7.5, 100 mM NaCl, 20 mM Imidazole, 0.01 mM EDTA, 1 mM DTT) supplemented with 0.05 mg/mL lysozyme (Sigma), 1 mM phenylmethanesulfonyl fluoride (PMSF) and 500 U of benzonase (Sigma), and lysed by sonication. The supernatant was isolated by centrifugation for 40 min at 4C at 18,500 x g and purified by Ni-NTA affinity chromatography (Novagen) on a gravity column. Ni-NTA resin was washed with 10 column volumes (CV) of Msp1 Lysis Buffer and then 10 CV of Wash Buffer (Msp1 Lysis buffer with 30 mM Imidazole) before elution with Lysis Buffer supplemented with 250 mM imidazole.

The protein was further purified by size exclusion chromatography (SEC) (Superdex 200 Increase 10/300 GL, GE Healthcare) in 20 mM Tris pH 7.5, 100 mM NaCl, 0.1 mM TCEP. Peak fractions were pooled, concentrated to 5-15 mg/ml in a 30 kDa MWCO Amicon Ultra centrifugal filter (Millipore) and aliquots were flash-frozen in liquid nitrogen and stored at –80C. Protein concentrations were determined by A280 using a calculated extinction coefficient (Expasy) for monomeric samples or a Bradford assay (samples containing nucleotide). The soluble E193Q construct was purified similarly, but with an additional overnight dialysis step (against 20 mM Tris pH 7.5, 200 mM NaCl, 0.1 mM EDTA, 1 mM DTT) to remove any bound nucleotide before SEC.

The crystallization construct (D1-32+3C) was generated by subcloning the soluble region of S. cerevisiae Msp1 (D1-32) into a pET21b derivative containing a C-terminal 6xHis tag, and then inserting a 3C protease site after residue 345. This construct was expressed and purified as described above with the following changes. After Ni-NTA purification, 3C protease was added at a 1:100 molar ratio to Msp1 and dialyzed for 16 hr at 4C against 20 mM Tris pH 7.5, 150 mM NaCl, 5 mM bME. Uncleaved protein and 3C protease were then removed by subtractive Ni-NTA chromatography. Cleaved protein was further purified by SEC on an S200 Increase column equilibrated in Buffer B (20 mM HEPES pH 7.5, 100 mM NaCl, 1 mM DTT). Selenomethionine (Se-Met) labeled protein was purified similarly, except for an additional 16 hr dialysis step against Buffer B after SEC.

6.3.2 Crystallization and Data Collection

Native and selenomethioine-containing Msp1 crystals were grown at room temperature by hanging drop vapor diffusion. A protein solution containing 9 mg/mL protein was mixed in a 1:1 (2:1 for Se-Met) ratio with a reservoir solution containing 16% PEG 3350 and 0.6 M Sodium Thiocyanate. Crystals were cryoprotected in 16% PEG 3350, 0.6 M Sodium Thiocyanate, 20% ethylene glycol, and flash frozen in liquid nitrogen.

Native and selenium SAD data were collected at 100K at APS beamline 24-IDC (l = 0. 97918 A˚) on a Pilatus 6M pixel-array detector. Data were processed using the xia2 DIALS pipeline (Evans, 2006; Winter, 2010).

6.3.3 Structure Determination, Refinement, and Analysis

The 2.6 A˚ structure of the Msp1 soluble region (D1-32+3C) was determined by SAD using PHENIX (Adams et al., 2010); after phasing and density modification the resulting electron density maps were of good quality, allowing 242 residues to be placed automatically using Autobuild. The initial model was improved by manual building and refinement with COOT (Emsley et al., 2010) and PHENIX. The final model contains one Msp1 and 11 water molecules; no electron density was observed for residues 33-49 and 200-205. Refinement and validation statistics are listed in Table 1.

Sequence conservation was analyzed from an alignment of 94 fungal Msp1 homologs identified by a BLASTp search (NCBI) using S. cerevisiae Msp1 as the search sequence. After aligning in Clustal Omega (Sievers et al., 2011) (http://www.ebi.ac.uk/), sequence conservation was mapped to the surface of Msp1 using the ConSurf server (Glaser et al., 2003) (http://consurf.tau.ac.il). Structure figures were generated with PyMOL (http://www.pymol.org).

6.3.4 Electron Microscopy and Image Analysis

For cryo-EM, 2 ml of soluble (D1-32) S. cerevisiae Msp1 E193Q at 1 mg/ml was incubated for 30 s on glow discharged C-Flat holey carbon grids (CF-1.2/1.3-2C, EMS), blotted for 10 s at 100% humidity and plunge frozen in liquid ethane using an FEI Vitrobot. Cryo-EM samples were imaged using a JEOL 3200FS operating at 300 KeV, equipped with a K2 direct electron detector camera (Gatan). Images were collected manually at a nominal magnification of 30,000x with a pixel size of 1.19A˚ and defocus range of 2-4 mm. Total exposure time was 4 s with an accumulated dose of 48 e- /A˚ 2 . Particle selection, CTF correction, 2D class averaging, and measurements were performed in EMAN 2.1 (Bell et al., 2016) on the University of Chicago

Midway computing cluster. Cryo-EM 2D class averages were calculated from 9,897 particles picked from 77 micrographs. Diameter measurements were made automatically with ImageJ (Schneider et al., 2012) after highlighting by applying a grayscale threshold to the class averages.

6.3.5 Size Analysis by Size-Exclusion Chromatography and Multi-angle Laser Light Scattering

Analytical SEC was carried out on a Superdex 200 Increase 10/300 GL column (GE Healthcare) with 500 mL of 2 mg/mL protein supplemented with 2 mM of the appropriate nucleotide and 2 mM MgCl2, as needed. Unless indicated otherwise, the column was equilibrated in 20 mM Tris pH 7.5, 100 mM NaCl, 0.1 mM TCEP; nucleotide and MgCl2 were not present in the SEC running buffer.

For size analysis by SEC-MALS, 100 mL of 10 mg/mL protein sample were injected onto a Superdex 200 10/300 GL column (GE Healthcare) equilibrated in 20 mM Tris pH 7.5, 100 mM NaCl, 0.1 mM TCEP. The purification system was coupled to an online, static, light scattering detector (Dawn HELEOS II, Wyatt Technology), a refractive-index detector (Optilab rEX, Wyatt Technology), and an ultraviolet-light detector (UPC-900, GE Healthcare). Absolute weight-averaged molar masses were calculated using the ASTRA software (Wyatt Technology).

7 References

1. 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.

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

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

4. Fagerberg, L., Jonasson, K., von Heijne, G., Uhlen, M. Berglund, L. (2010). Prediction of the human membrane proteome. Proteomics 10, 1141–1149.

5. Ferguson, Michael A.J. Lipid Anchors on Membrane Proteins. (1991). Current Opinion in Structural Biology 1, 4: 522–29.

6. Paulick, Margot G., Carolyn R. Bertozzi. (2008). The Glycosylphosphatidylinositol Anchor: A Complex Membrane-Anchoring Structure for Proteins. Biochemistry 47, 27: 6991–7000.

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

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

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

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

11. van Meer G., de Kroon A. I. P. M. (2011). Lipid map of the mammalian cell. Journal of Cell Science. 124, 5-8.

12. Shao, S., Hegde, R. S. (2011). Membrane protein insertion at the endoplasmic reticulum. Annu. Rev. Cell Dev. Biol. 27, 25–56.

13. Hacinska, A., Koehler, C. M., Milenkovic, D., Lithgow, T. & Pfanner, N. (2009). Importing mitochondrial proteins: machineries and mechanisms. Cell. 138, 628–644.

14. Pfanner, N., and Geissler, A. (2001). Versatility of the Mitochondrial Protein Import Machinery. Nature Reviews Molecular Cell Biology. 2, 5: 339–49.

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

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

17. Shan S. O., Walter, P. (2005). Co-translational protein targeting by the signal recognition particle. FEBS Letters. 579: 921–926.

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

19. von Heijne, G. (1985). Signal sequences: The limits of variation. J. Mol. Biol. 184, 99–105.

20. Costa, E.A., Subramanian, K., Nunnari, J., Weissman, J.S. (2018). Defining the physiological role of SRP in protein-targeting efficiency and specificity. Science. 3607.

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

22. Luirink, J. (2004). SRP-Mediated Protein Targeting: Structure and Function Revisited. Biochimica et Biophysica Acta. 1694(1-3):17-35.

23. Fu, H. Y. H., Huang, W. Y. C., Shen K., Groves J. T., Miller T., Shan, S. (2017). Two-Step Membrane Binding by the Bacterial SRP Receptor Enable Efficient and Accurate Co-Translational Protein Targeting. ELife. 6: e25885.

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

25. Pfeffer, S., Burbaum, L., Unverdorben, P., Pech, M., Chen, Y., Zimmermann, R., Beckmann, R., Förster, F. (2015). Structure of the native Sec61 protein-conducting channel. Nat. Commun. 6, 8403.

26. Martoglio, B., Hofmann M. W., Brunner J., Dobberstein, B. (1995). The Protein-Conducting Channel in the Membrane of the Endoplasmic Reticulum Is Open Laterally toward the Lipid Bilayer. Cell. 81, 2: 207–14.

27. Osborne, A. R., Rapoport T. A., van den Berg, B. (2005). Protein Translocation by the Sec61/SecY. Annual Review of Cell and Developmental Biology. 21, 1:529–50.

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

29. 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.

30. Li, Long, Eunyong Park, JingJing Ling, Jessica Ingram, Hidde Ploegh, and Tom A. Rapoport. (2016). Crystal Structure of a Substrate-Engaged SecY Protein-Translocation Channel. Nature. 531:395.

31. Tanaka, Y., Sugano, Y., Takemoto, M., Mori, T., Furukawa, A., Kusakizako, T., Kumazaki, K. Kashima, A., Ishitani, R., Sugita, Y., Nureki, O., Tsukazaki, T. (2015). Crystal Structures of SecYEG in Lipidic Cubic Phase Elucidate a Precise Resting and a Peptide-Bound State. Cell Reports. 13, 8:1561–68.

32. Nishiyama K., Mizushima S., Tokuda H. (1993). A novel membrane protein involved in protein translocation across the cytoplasmic membrane of Escherichia coli. EMBO J. 12:3409–3415.

33. Pohlschröder, M., Prinz W. A., Hartmann, E., Beckwith, J. Protein Translocation in the Three Domains of Life: Variations on a Theme. Cell. 91, 5: 563–66.

34. Park. E., Ménétret, J., Gumbart, J. C., Ludtke, S. J., Li, W., Whynot, A., Rapoport, T. A., Akey, C. W. (2014). Structure of the SecY Channel during Initiation of Protein Translocation. Nature. 506:102.

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

36. Elisabet C. M., Trueman, S. F., Gilmore, R. (2009). Translocation of Proteins through the Sec61 and SecYEG Channels. Current Opinion in Cell Biology. 21, no. 4:501–7.

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

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

39. 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.

40. 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.

41. Conti, B.J., Devaraneni, P.K., Yang, Z., David, L.L., 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.

42. Aviram, N., Schuldiner, M. (2017). Targeting and Translocation of Proteins to the Endoplasmic Reticulum at a Glance. Journal of Cell Science. 130, 24:4079.

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

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

45. 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.

46. Sachelaru, I., Petriman, N.A., Kudva, R., Kuhn, P., Welte, T., Knapp, B., Drepper, F., Warscheid, B., 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.

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

48. 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., Luirink, J. (2001). Sec-dependent membrane protein insertion: sequential interaction of nascent FtsQ with SecY and YidC. EMBO Rep. 2, 524–529.

49. Zhu, L., Klenner, C., Kuhn, A., 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.

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

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

52. Wickles, S, Singharoy, A., Andreani, J., Seemayer, S., Bischoff, L., Berninghausen, O., Soeding, J., Schulten, K., van der Sluis, E. O., Beckmann, R. (2014). A Structural Model of the Active Ribosome-Bound Membrane Protein Insertase YidC. eLife. 3:e03035.

53. Yuan, J., Phillips, G. J., Dalbey, R. E. (2007). Isolation of Cold-Sensitive YidC Mutants Provides Insights into the Substrate Profile of the YidC Insertase and the Importance of Transmembrane 3 in YidC Function. Journal of Bacteriology. 189, 24: 8961–72.

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

55. Preuss, M., Ott, M., Funes, S., Luirink, J., 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.

56. Tsukazaki, Tomoya, Hiroyuki Mori, Yuka Echizen, Ryuichiro Ishitani, Shuya Fukai, Takeshi Tanaka, Anna Perederina, et al. (2011). Structure and Function of a Membrane Component SecDF That Enhances Protein Export. Nature. 474: 235.

57. Pogliano, J. A. Beckwith, J. (1994). SecD and SecF facilitate protein export in Escherichia coli. EMBO J. 13, 554–561.

58. Nouwen, N., Piwowarek, M., Berrelkamp, G. Driessen, A. J. (2005). The large first periplasmic loop of SecD and SecF plays an important role in SecDF functioning. J. Bacteriol. 187, 5857–5860.

59. Arkowitz, R. A, Wickner, W. (1994). SecD and SecF are required for the proton electrochemical gradient stimulation of preprotein translocation. EMBO J. 13(4):954–963.

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

61. Ménétret, J., Hegde R. S., Aguiar, M., Gygi, S. P., Park, E., Rapoport, T. A., Akey, C. W., (2008). Single Copies of Sec61 and TRAP Associate with a Nontranslating Mammalian Ribosome. Structure. 16, 7:1126–37.

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

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

64. Fons, Ryen D., Brigitte A. Bogert, Ramanujan S. Hegde. (2003). Substrate-Specific Function of the Translocon-Associated Protein Complex during Translocation across the ER Membrane. The Journal of Cell Biology. 160, 4: 529–39.

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

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

67. Kelleher, Daniel J., Gilmore R. (2006) An Evolving View of the Eukaryotic Oligosaccharyltransferase. Glycobiology. 16, 4: 47R-62R.

68. Shrimal S, Ng BG, Losfeld ME, Gilmore R, Freeze HH. (2013). Mutations in STT3A and STT3B cause two congenital disorders of glycosylation. Hum Mol Genet. 22(22):4638-45.

69. Jayaprakash, N. G., Surolia, A. (2017). Role of Glycosylation in Nucleating Protein Folding and Stability. Biochemical Journal. 474, 14: 2333–47.

70. Lee, H. S., Qi Y., Im, W. (2015). Effects of N-Glycosylation on Protein Conformation and Dynamics: Protein Data Bank Analysis and Molecular Dynamics Simulation Study. Scientific Reports. 5, 1: 8926.

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

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

73. Cherepanova, N.A., 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.

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

75. Boisramé, A., Chasles, M., Babour, A., Beckerich, J.-M., 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.

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

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

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

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

80. Krieg, U.C., Johnson, A.E., 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.

81. Voigt, S., Jungnickel, B., Hartmann, E., 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.

82. Valent, Q. A. (1998) The Escherichia Coli SRP and SecB Targeting Pathways Converge at the Translocon. EMBO J. 17, 9: 2504–12.

83. Sala, A., Bordes, P., Genevaux P. (2014). Multitasking SecB Chaperones in Bacteria. Frontiers in Microbiology 5, 5:666.

84. Lycklama a Nijeholt, J. A., Driessen, A. J. M. (2012). The Bacterial Sec-Translocase: Structure and Mechanism. Philosophical Transactions of the Royal Society B: Biological Sciences. 367, 1592:1016–28.

85. Suo, Y., Hardy S. J. S., Randall, L. L. (2015). The Basis of Asymmetry in the SecA:SecB Complex. Journal of Molecular Biology. 427, 4: 887–900.

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

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

88. Beilharz T, Egan B, Silver PA, Hofmann K, Lithgow T. (2003). Bipartite signals mediate subcellular targeting of tail-anchored membrane proteins in Saccharomyces cerevisiae. J Biol Chem. 278:8219–8223.

89. Kalbfleisch T, Cambon A, Wattenberg BW. (2007) A bioinformatics approach to identifying tail-anchored proteins in the human genome. Traffic. 8:1687–1694.

90. Kriechbaumer V, et al. (2009). Subcellular distribution of tail-anchored proteins in Arabidopsis. Traffic. 10:1753–1764.

91. Stefanovic S, Hegde RS. (2007) Identification of a targeting factor for posttranslational membrane protein insertion into the ER. Cell. 128:1147–1159.

92. Favaloro V, Spasic M, Schwappach B, Dobberstein B. (2008) Distinct targeting pathways for the membrane insertion of tail-anchored (TA) proteins. J Cell Sci. 121:1832–1840.

93. Hegde, Ramanujan S., and Robert J. Keenan. (2011) Tail-Anchored Membrane Protein Insertion into the Endoplasmic Reticulum. Nature Reviews Molecular Cell Biology. 12, 12:787–98.

94. Wang, F., Brown, E. C., Mak, G., Zhuang, J. Denic, V. (2010). A chaperone cascade sorts protein for posttranslational membrane insertion into the endoplasmic reticulum. Mol. Cell 40, 159–171.

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

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

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

98. Mateja, A., M. Paduch, H.-Y. Chang, A. Szydlowska, A. A. Kossiakoff, R. S. Hegde, R. J. Keenan. (2015) Structure of the Get3 Targeting Factor in Complex with Its Membrane Protein Cargo." Science 347, 6226;1152–55.

99. Schuldiner, Maya, Jutta Metz, Volker Schmid, Vladimir Denic, Magdalena Rakwalska, Hans Dieter Schmitt, Blanche Schwappach, and Jonathan S. Weissman. (2008). The GET Complex Mediates Insertion of Tail-Anchored Proteins into the ER Membrane. Cell 134, 4: 634–45.

100. B.E. Zalisko, C. Chan, V. Denic, R.S. Rock, R.J. (2017). Keenan Tail-anchored protein insertion by a single Get1/2 heterodimer. Cell Rep. 20, 2287-2293

101. Wang F., Chan C., Weir N. R., Denic V. (2014). The Get1/2 transmembrane complex is an endoplasmic-reticulum membrane protein insertase. Nature; 512(7515):441-4.

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

103. Hegde, Ramanujan S., Robert J. Keenan. (2011). Tail-Anchored Membrane Protein Insertion into the Endoplasmic Reticulum. Nature Reviews Molecular Cell Biology 12, 12:787–98.

104. Hessa, T. et al. (2011). Protein targeting and degradation are coupled for elimination of mislocalized proteins. Nature 475, 394–397.

105. Minami, R., Hayakawa, A., Kagawa, H., Yanagi, Y., Yokosawa, H., Kawahara. H. (2010). BAG-6 Is Essential for Selective Elimination of Defective Proteasomal Substrates. The Journal of Cell Biology 190, 4: 637–50.

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

107. 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.

108. Deshaies, R.J., Sanders, S.L., Feldheim, D.A., 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.

109. Plath, K., 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.

110. Brodsky, J.L., Goeckeler, J., 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.

111. Gray, M. W. (2012). Mitochondrial Evolution. Cold Spring Harbor Perspectives in Biology 4, 9: a011403–a011403.

112. Walther, D. M., Rapaport, D. (2009). Biogenesis of Mitochondrial Outer Membrane Proteins. Biochimica et Biophysica Acta. 1793, 1: 42–51.

113. Dudek, J., Rehling, P., van der Laan, M. (2013) Mitochondrial Protein Import: Common Principles and Physiological Networks. Biochimica et Biophysica Acta. 1833, 2: 274–85.

114. Young, J. C., Hoogenraad, N. J., Hartl, F. U. (2003). Molecular Chaperones Hsp90 and Hsp70 Deliver Preproteins to the Mitochondrial Import Receptor Tom70. Cell 112, 1: 41–50.

115. Bruggisser, J., Käser, S., Mani, J., Schneider, A. (2017). Biogenesis of a Mitochondrial Outer Membrane Protein in Trypanosoma Brucei: TARGETING SIGNAL AND

DEPENDENCE ON A UNIQUE BIOGENESIS FACTOR. Journal of Biological Chemistry. 292, 8:3400–3410.

116. Pfanner, N. (2000). Protein Sorting: Recognizing Mitochondrial Presequences. Current Biology 10, 11: R412–15.

117. K. Hill, K. Model, M.T. Ryan, K. Dietmeier, F. Martin, R. Wagner, N. Pfanner. (1998). Tom40 forms the hydrophilic channel of the mitochondrial import pore for preproteins. Nature 395, 516–521.

118. Brix, J., Dietmeier, K., Pfanner. N. (1997). Differential recognition of preproteins by the purified cytosolic domains of the mitochondrial import receptors Tom20, Tom22, and Tom70, J. Biol. Chem. 272, 20730–20735.

119. Moczko, M., Bömer, U., Kübrich, M., Zufall, N., Hönlinger, A., Pfanner. N. (1997). The intermembrane space domain of mitochondrial Tom22 functions as a trans binding site for preproteins with N-terminal targeting sequences. Mol. Cell. Biol., 17, 6574-6584.

120. Truscott, K. N., Kovermann, P., Geissler, A., Merlin, A., Meijer, M., Driessen, A. J., Rassow, J., Pfanner, N., Wagner. R. (2001). A presequence- and voltage-sensitive channel of the mitochondrial preprotein translocase formed by Tim23. Nat. Struct. Biol., 8, 1074-1082.

121. Meinecke, M., Wagner, R., Kovermann, P., Guiard, B., Mick, D. U., Hutu, D. P., Voos, W., Truscott, K. N., Chacinska, A., Pfanner, N., Rehling. P. (2006) Tim50 maintains the permeability barrier of the mitochondrial inner membrane. Science. 312, 1523-1526.

122. Chacinska, A., Lind, M., Frazier, A. E., Dudek, J., Meisinger, C., Geissler, A., Sickmann, A., Meyer, H. E., Truscott, K. N., Guiard, B., Pfanner, N., Rehling. P. (2005) Mitochondrial presequence translocase: switching between TOM tethering and motor recruitment involves Tim21 and Tim17. Cell, 120, 817-829.

123. Meier, S., Neupert, W., Herrmann. J. M. (2005). Conserved N-terminal negative charges in the Tim17 subunit of the TIM23 translocase play a critical role in the import of preproteins into mitochondria. J. Biol. Chem., 280, 7777-7785.

124. Schleyer, M., Neupert. W. (1985). Transport of proteins into mitochondria: translocational intermediates spanning contact sites between outer and inner membranes. Cell, 43 (1985), 339-350.

125. Schwaiger, M., Herzog, V., Neupert. W. (1987). Characterization of translocation contact sites involved in the import of mitochondrial proteins. J. Cell Biol. 105, 235-246.

126. Li, Y., Dudek, J., Guiard, B., Pfanner, N., Rehling, P., Voos. W., (2004). The presequence translocase-associated protein import motor of mitochondria: Pam16 functions in an antagonistic manner to Pam18. J. Biol. Chem. 279, 38047-38054.

127. Truscott, K. N., Voos, W., Frazier, A. E., Lind, M., Li, Y., Geissler, A., Dudek, J., Müller, H., Sickmann, A., Meyer, H. E., Meisinger, C., Guiard, B., Rehling, P., Pfanner. N. (2003). A J-protein is an essential subunit of the presequence translocase-associated protein import motor of mitochondria. J. Cell Biol. 163, 707-713.

128. Ungermann, C., Neupert, W., Cyr. D. W. (1994). The role of Hsp70 in conferring unidirectionality on protein translocation into mitochondria. Science. 266, 1250-1253.

129. Liu, Q., D'Silva, P., Walter, W., Marszalek, J., Craig, E. A. (2003) Regulated cycling of mitochondrial Hsp70 at the protein import channel. Science. 300, 139-141.

130. Akio. I. (1999). Mitochondrial Processing Peptidase: Multiple-Site Recognition of Precursor Proteins. Biochemical and Biophysical Research Communications. 265: 611–16.

131. Taylor, A., Smith, B., Kitada, S., Kojima, K., Miyaura, H., Otwinowski, Z., Ito, A., Deisenhofer, J. (2001). Crystal structures of mitochondrial processing peptidase revealthe mode for specific cleavage of import signal sequences, Structure, 9 615–625.

132. Becker, T., Pfannschmidt, S., Guiard, B., Stojanovski, D., Milenkovic, D., Kutik, S., Pfanner, N., Meisinger, C., Wiedemann. N. (2008). Biogenesis of the mitochondrial TOM complex: Mim1 promotes insertion and assembly of signal-anchored receptors. J. Biol. Chem. 283, 120-127.

133. Popov-Celeketić, J., Waizenegger, T., Rapaport. D. (2008). Mim1 functions in an oligomeric form to facilitate the integration of Tom20 into the mitochondrial outer membrane. J. Mol. Biol. 376, 671-680.

134. Hulett, J. M., Lueder, F., Chan, N. C., Perry, A. J., Wolynec, P., Likić, V. A., Gooley, P. R., Lithgow. T. (2008). The transmembrane segment of Tom20 is recognized by Mim1 for docking to the mitochondrial TOM complex. J. Mol. Biol. 376, 694-704.

135. Setoguchi, K., Otera, H., Mihara, K. (2006). Cytosolic factor- and TOM-independent import of C-tail-anchored mitochondrial outer membrane proteins. EMBO J. 25, 5635-5647

136. Kemper, C., Habib, S. J., Engl, G., Heckmeyer, P., Dimmer, K. S., Rapaport. D. (2008). Integration of tail-anchored proteins into the mitochondrial outer membrane does not require any known import components. J. Cell Sci. 121, 1990-1998.

137. Bohnert, M., Rehling, P., Guiard, B., Herrmann, J. M., Pfanner, N., van der Laan, M. (2010). Cooperation of Stop-Transfer and Conservative Sorting Mechanisms in Mitochondrial Protein Transport. Current Biology 20, 13:1227–32.

138. Luirink, J., Samuelsson, T., de Gier, J. (2001). YidC/Oxa1p/Alb3: Evolutionarily Conserved Mediators of Membrane Protein Assembly. FEBS Letters 501, 1: 1–5.

139. Keil, M., Bareth, B., Woellhaf, M. W., Peleh, V., Prestele, M., Rehling, P., Herrmann, J. M. (2012). Oxa1-Ribosome Complexes Coordinate the Assembly of Cytochrome c Oxidase in Mitochondria. Journal of Biological Chemistry. 287, 41: 34484–93.

140. Bonnefoy, N., Chalvet, F., Hamel, P., Slonimski, P. P., Dujardin, G. (1994). OXA1, a Saccharomyces Cerevisiae Nuclear Gene Whose Sequence Is Conserved Form Prokaryotes to Eukaryotes Controls Cytochrome Oxidase Biogenesis. Journal of Molecular Biology. 239, 2: 201–12.

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

142. Borowska, M. T., Dominik, P. K., Anghel, S. A., Kossiakoff, A. A., Keenan, R. J., (2015). A YidC-like Protein in the Archaeal Plasma Membrane. Structure 23, 9: 1715–24.

143. 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.

144. Kuhn, A., Koch, H. G., Dalbey, R. E. (2017). Targeting and insertion of mem- brane proteins. Ecosal Plus 7.

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

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

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

148. Zaremba-Niedzwiedzka, K., Caceres, E. F., Saw, J. H., Backstrom, 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.

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

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

151. Aschtgen, M. S., Zoued, A., Lloubes, R., Journet, L., Cascales, E. (2012). The C-tail anchored TssL subunit, an essential protein of the enteroaggrega- tive Escherichia coli Sci-1 Type VI secretion system, is inserted by YidC. Microbiology Open. 1, 71–82.

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

153. 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.

154. 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 Ca(2+) load-activated Ca(2+) channel. Cell 165, 1454–1466.

155. 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.

156. Stefer, S., Reitz, S., Wang, F., Wild, K., Pang, Y.Y., Schwarz, D., Bomke, J., Hein, C., Lo¨hr, F., Bernhard, F., et al. (2011). Structural basis for tail-anchored membrane protein biogenesis by the Get3-receptor complex. Science. 333, 758–762.

157. Wang, F., Whynot, A., Tung, M., and Denic, V. (2011). The mechanism of tail- anchored protein insertion into the ER membrane. Mol. Cell. 43, 738–750.

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

159. 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.

160. Wang, S., Sun, S., Li, Z., Zhang, R., Xu, J. (2017). Accurate de novo pre- diction of protein contact map by ultra-deep learning model. PLoS Comput. Biol. 13, e1005324.

161. Iwamuro, S., Saeki, M., Kato, S. (1999). Multi-ubiquitination of a nascent membrane protein produced in a rabbit reticulocyte lysate. J. Biochem. 126, 48–53.

162. 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.

163. Sharma, S., Burdon, K. P., Chidlow, G., Klebe, S., Crawford, A., Dimasi, D. P., Dave, A., Martin, S., Javadiyan, S., Wood, J.P., 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. Invest. Ophthalmol. Vis. Sci. 53, 4917–4925.

164. Alanay, Y., Erguener, B., Utine, E., Hacarız, O., Kiper, P. O. S., Tasxkıran, E. Z., Percin, F., Uz, E., Sagıroglu, M. S., Yuksel, B., et al. (2014). TMCO1 deficiency causes autosomal recessive cerebrofaciothoracic dysplasia. Am. J. Med. Genet. A. 164, 291–304.

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

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

167. Botte, M., Zaccai, N. R., Nijeholt, J. L., 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.

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

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

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

171. Simon, S. M., Blobel, G., 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.

172. 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.

173. Kruger, V., Deckers, M., Hildenbeutel, M., van der Laan, M., Hellmers, M., Dreker, C., Preuss, M., Herrmann, J. M., Rehling, P., Wagner, R., and Meinecke, M. (2012). The mitochondrial oxidase assembly protein1 (Oxa1) insertase forms a membrane pore in lipid bilayers. J. Biol. Chem. 287, 33314–33326.

174. Erdmann, F., Schauble, N., Lang, S., Jung, M., Honigmann, A., Ahmad, M., Dudek, J., Benedix, J., Harsman, A., Kopp, A., et al. (2011). Interaction of calmodulin with Sec61a limits Ca2+ leakage from the endoplasmic reticulum. EMBO J. 30, 17–31.

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

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

177. Vogl, C., Panou, I., Yamanbaeva, G., Wichmann, C., Mangosing, S. J., Vilardi, F., Indzhykulian, A. A., Pangrsic, T., Santarelli, R., Rodriguez-Ballesteros, M., et al. (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.

178. Ma, H., Dang, Y., Wu, Y., Jia, G., Anaya, E., Zhang, J., Abraham, S., Choi, J.G., Shi, G., Qi, L., et al. (2015). A CRISPR-based screen identifies genes essential for West-Nile-virus-induced cell death. Cell Rep. 12, 673–683.

179. Aviram, N., 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.

180. Aviram, N., Ast, T., Costa, E.A., Arakel, E.C., Chuartzman, S.G., Jan, C.H., Haßdenteufel, S., Dudek, J., Jung, M., Schorr, S., et al. (2016). The SND pro- teins constitute an alternative targeting route to the endoplasmic reticulum. Nature 540, 134–138.

181. Anghel, S. A., McGilvray, P. T., Hegde, R. S., Keenan. R. J. (2017). Identification of Oxa1 Homologs Operating in the Eukaryotic Endoplasmic Reticulum. Cell Reports. 21, 13:3708–16.

182. Centers for Mendelian Genomics, Davut Pehlivan, Ender Karaca, Hatip Aydin, Christine R Beck, Tomasz Gambin, Donna M Muzny, et al. (2014). Whole-Exome Sequencing Links TMCO1 Defect Syndrome with Cerebro-Facio-Thoracic Dysplasia. European Journal of Human Genetics. 22, 9: 1145–48.

183. Yates, M., Ng, T. O., Offiah, A. C., Willoughby, J., Berg, J. N., DDD Study, Johnson. D. S. (2019). Cerebrofaciothoracic Dysplasia: Four New Patients with a Recurrent TMCO1 Pathogenic Variant. American Journal of Medical Genetics Part A 179, 1: 43–49.

184. Haffner, C., Dettmer, U., Weiler, T., Haass. C. (2007). The Nicastrin-like Protein Nicalin Regulates Assembly and Stability of the Nicalin-Nodal Modulator (NOMO) Membrane Protein Complex. Journal of Biological Chemistry. 282, 14: 10632–38.

185. Haffner, C., Frauli, M., Topp, S., Irmler, M., Hofmann, K., Regula, J. T., Bally-Cuif, L., Haass. C. (2004). Nicalin and Its Binding Partner Nomo Are Novel Nodal Signaling Antagonists. The EMBO Journal. 23, 15:3041–50.

186. Dettmer, U., Kuhn, P., Abou-Ajram, C., Lichtenthaler, S. F., Krüger, M., Kremmer, E., Haass, C., Haffner. C. (2010). Transmembrane Protein 147 (TMEM147) Is a Novel Component of the Nicalin-NOMO Protein Complex. Journal of Biological Chemistry. 285, 34: 26174–81.

187. Yamamoto, S., Yamazaki, T., Komazaki, S., Yamashita, T., Osaki, M., Matsubayashi, M., Kidoya, H., Takakura, N., Yamazaki, D., Kakizawa. S. (2014). Contribution of Calumin to Embryogenesis through Participation in the Endoplasmic Reticulum-Associated Degradation Activity. Developmental Biology 393, 1: 33–43.

188. Morimoto, M., Waller-Evans, H., Ammous, Z., Song, X., Strauss, K. A., Pehlivan, D., Gonzaga-Jauregui, C., et al. (2018). Bi-Allelic CCDC47 Variants Cause a Disorder Characterized by Woolly Hair, Liver Dysfunction, Dysmorphic Features, and Global Developmental Delay. The American Journal of Human Genetics 103, 5: 794–807.

189. Gottschalk, A., Almedom, R. B., Schedletzky, T., Anderson, S. D., Yates, J. R., Schafer. W. R. (2005). Identification and Characterization of Novel Nicotinic Receptor-Associated Proteins in Caenorhabditis Elegans. The EMBO Journal. 24, 14: 2566–78.

190. Almedom, R. B., Liewald, J. F., Hernando, G., Schultheis, C., Rayes, D., Pan, J., Schedletzky, T., Hutter, H., Bouzat, C., Gottschalk. A. (2009). An ER-Resident Membrane Protein Complex Regulates Nicotinic Acetylcholine Receptor Subunit Composition at the Synapse. The EMBO Journal. 28, 17: 2636–49.

191. Braunger, K., Pfeffer, S., Shrimal, S., Gilmore, R., Berninghausen, O., Mandon, E. C., Becker, T., Förster, F., Beckmann. R. (2018). Structural Basis for Coupling Protein Transport and N-Glycosylation at the Mammalian Endoplasmic Reticulum. Science. 360, 6385: 215–19.

192. Bai, X., Yan, C., Yang, G., Lu, P., Ma, D., Sun, L., Zhou, R., Scheres, S. H. W., Shi. Y. (2015). An Atomic Structure of Human γ-Secretase. Nature. 525, 7568: 212–17.

193. Lang, S., Pfeffer, S., Lee, P., Cavalié, A., Helms, V., Förster, F., Zimmermann. R. (2017). An Update on Sec61 Channel Functions, Mechanisms, and Related Diseases." Frontiers in Physiology. 8:887.

194. Pereira, F., Rettel, M., Stein, F., Savitski, M. M., Collinson, I., Römisch. K. (2019). Effect of Sec61 Interaction with Mpd1 on Endoplasmic Reticulum-Associated Degradation. PLOS ONE. 14, 1: e0211180.

195. Goder, V., Bieri, C., Spiess, M. (1999). Glycosylation can influence topogenesis of membrane proteins and reveals dynamic reorientation of nascent polypeptides within the translocon. The Journal of Cell Biology. 147(2):257-66.

196. De Marothy, M. T., Elofsson. A. (2015). Marginally Hydrophobic Transmembrane α - Helices Shaping Membrane Protein Folding: Marginally Hydrophobic Transmembrane α-Helices. Protein Science. 24, 7: 1057–74.

197. Chitwood, P. J., Juszkiewicz, S., Guna, A., Shao, S., Hegde, R. S. (2018). EMC Is Required to Initiate Accurate Membrane Protein Topogenesis. Cell. 175, 6:1507-1519.e16.

198. Shurtleff, M. J., Itzhak, D. N., Hussmann, J. A., Oakdale, N. T. S., Costa, E. A., Jonikas, M., Weibezahn, J. et al. (2018). The ER Membrane Protein Complex Interacts Cotranslationally to Enable Biogenesis of Multipass Membrane Proteins. eLife. 7:e37018.

199. Chacinska, A., Koehler, C. M., Milenkovic, D., Lithgow, T., Pfanner, N. (2009). Importing mitochondrial proteins: machineries and mechanisms. Cell. 138, 628–644.

200. Heiland, I., Erdmann, R. (2005). Biogenesis of peroxisomes. Topogenesis of the peroxisomal membrane and matrix proteins. FEBS J. 272, 2362–2372.

201. Inaba, T., Schnell, D.J. (2008). Protein trafficking to plastids: one theme, many variations. Biochem. J. 413, 15–28.

202. Brodsky, J.L., Skach, W.R. (2011). Protein folding and quality control in the endoplasmic reticulum: Recent lessons from yeast and mammalian cell systems. Curr. Opin. Cell Biol. 23, 464–475.

203. Foresti, O., Rodriguez-Vaello, V., Funaya, C., and Carvalho, P. (2014). Quality control of inner nuclear membrane proteins by the Asi complex. Science. 346, 751–755.

204. Hamon, M.P., Bulteau, A.L., Friguet, B. (2015). Mitochondrial proteases and protein quality control in ageing and longevity. Ageing Res. Rev. 23, 56–66.

205. Khmelinskii, A., Blaszczak, E., Pantazopoulou, M., Fischer, B., Omnus, D.J., Le Dez, G., Brossard, A., Gunnarsson, A., Barry, J.D., Meurer, M., et al. (2014). Protein quality control at the inner nuclear membrane. Nature. 516, 410–413.

206. Okiyoneda, T., Barriere, H., Bagdany, M., Rabeh, W. M., Du, K., Ho hfeld, J., Young, J.C., Lukacs, G.L. (2010). Peripheral protein quality control re- moves unfolded CFTR from the plasma membrane. Science. 329, 805–810.

207. Tong, Z., Kim, M.S., Pandey, A., Espenshade, P.J. (2014). Identification of candidate substrates for the Golgi Tul1 E3 ligase using quantitative diGly proteomics in yeast. Mol. Cell. Proteomics. 13, 2871–2882.

208. Beilharz, T., Egan, B., Silver, P. A., Hofmann, K., Lithgow, T. (2003). Bipartite signals mediate subcellular targeting of tail-anchored membrane proteins in Saccharomyces cerevisiae. J. Biol. Chem. 278, 8219–8223.

209. Borgese, N., Brambillasca, S., Colombo, S. (2007). How tails guide tail- anchored proteins to their destinations. Curr. Opin. Cell Biol. 19, 368–375.

210. Wattenberg, B., Lithgow, T. (2001). Targeting of C-terminal (tail)-anchored proteins: understanding how cytoplasmic activities are anchored to intracellular membranes. Traffic. 2, 66–71.

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

212. Chen, Y. C., Umanah, G. K., Dephoure, N., Andrabi, S. A., Gygi, S. P., Dawson, T. M., Dawson, V. L., Rutter, J. (2014). Msp1/ATAD1 maintains mitochondrial function by facilitating the degradation of mislocalized tail-anchored proteins. EMBO J. 33, 1548–1564.

213. Jonikas, M. C., Collins, S. R., Denic, V., Oh, E., Quan, E. M., Schmid, V., Weibezahn, J., Schwappach, B., Walter, P., Weissman, J. S., Schuldiner, M. (2009). Comprehensive characterization of genes required for protein folding in the endoplasmic reticulum. Science. 323, 1693–1697.

214. Okreglak, V., Walter, P. (2014). The conserved AAA-ATPase Msp1 confers organelle specificity to tail-anchored proteins. Proc. Natl. Acad. Sci. 111, 8019–8024.

215. Nakai, M., Endo, T., Hase, T., Matsubara, H. (1993). Intramitochondrial protein sorting. Isolation and characterization of the yeast MSP1 gene which belongs to a novel family of putative ATPases. J. Biol. Chem. 268, 24262–24269.

216. Zhang, J., Wang, Y., Chi, Z., Keuss, M.J., Pai, Y. M., Kang, H. C., Shin, J. H., Bugayenko, A., Wang, H., Xiong, Y., et al. (2011). The AAA+ ATPase Thorase regulates AMPA receptor-dependent synaptic plasticity and behavior. Cell. 145, 284–299.

217. Kale, J., Chi, X., Leber, B., Andrews, D. (2014). Examining the molecular mechanism of bcl-2 family proteins at membranes by fluorescence spectroscopy. Methods Enzymol. 544, 1–23.
218. Frickey, T., Lupas, A.N. (2004). Phylogenetic analysis of AAA proteins. J. Struct. Biol. 146, 2–10.

219. Peng, W., Lin, Z., Li, W., Lu, J., Shen, Y., Wang, C. (2013). Structural in- sights into the unusually strong ATPase activity of the AAA domain of the Caenorhabditis elegans fidgetin-like 1 (FIGL-1) protein. J. Biol. Chem. 288, 29305–29312.

220. Scott, A., Chung, H.Y., Gonciarz-Swiatek, M., Hill, G. C., Whitby, F. G., Gaspar, J., Holton, J.M., Viswanathan, R., Ghaffarian, S., Hill, C.P., Sundquist, W. I. (2005). Structural and mechanistic studies of VPS4 proteins. EMBO J. 24, 3658–3669.

221. Stinson, B. M., Nager, A. R., Glynn, S. E., Schmitz, K. R., Baker, T. A., Sauer, R. T. (2013). Nucleotide binding and conformational switching in the hexameric ring of a AAA+ machine. Cell. 153, 628–639.

222. Martin, A., Baker, T. A., Sauer, R. T. (2008). Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. Nat. Struct. Mol. Biol. 15, 1147–1151.

223. Monroe, N., Hill, C. P. (2016). Meiotic clade AAA ATPases: protein polymer disassembly machines. J. Mol. Biol. 428 (9 Pt B), 1897–1911.

224. Wohlever, M. L., Baker, T. A., Sauer, R. T. (2014). Roles of the N domain of the AAA+ Lon protease in substrate recognition, allosteric regulation and chaperone activity. Mol. Microbiol. 91, 66–78.

225. Akiyama, Y., Ito, K. (2000). Roles of multimerization and membrane association in the proteolytic functions of FtsH (HflB). EMBO J. 19, 3888–3895.

226. Korbel, D., Wurth, S., Kaeser, M., Langer, T. (2004). Membrane protein turnover by the m-AAA protease in mitochondria depends on the transmembrane domains of its subunits. EMBO Rep. 5, 698–703.

227. Claessen, J. H., Ploegh, H. L. (2011). BAT3 guides misfolded glycoproteins out of the endoplasmic reticulum. PLoS ONE. 6, e28542.

228. Kawahara, H., Minami, R., Yokota, N. (2013). BAG6/BAT3: emerging roles in quality control for nascent polypeptides. J. Biochem. 153, 147–160.

229. Lee, J.G., and Ye, Y. (2013). Bag6/Bat3/Scythe: a novel chaperone activity with diverse regulatory functions in protein biogenesis and degradation. BioEssays. 35, 377–385.

230. Wang, Q., Liu, Y., Soetandyo, N., Baek, K., Hegde, R., Ye, Y. (2011). A ubiquitin ligase-associated chaperone holdase maintains polypeptides in soluble states for proteasome degradation. Mol. Cell. 42, 758–770.

231. Graef, M., Seewald, G., Langer, T. (2007). Substrate recognition by AAA+ ATPases: distinct substrate binding modes in ATP-dependent protease Yme1 of the mitochondrial intermembrane space. Mol. Cell. Biol. 27, 2476–2485.

232. Ito, K., Akiyama, Y. (2005). Cellular functions, mechanism of action, and regulation of FtsH protease. Annu. Rev. Microbiol. 59, 211–231.